

UNIVERSIDAD NACIONAL DE SAN ANTONIO

ABAD DEL CUSCO

ESCUELA DE POSGRADO

MAESTRÍA EN ESTADÍSTICA



**MODELOS AUTORREGRESIVOS INTEGRADOS DE MEDIA MÓVIL Y
RED NEURONAL RECURRENTE PARA PREDECIR LA DEMANDA DE
MEDICAMENTOS, HOSPITAL REGIONAL CUSCO, 2013 - 2018**

TESIS PRESENTADA POR:

Br. AMILCAR PUMA CÁRDENAS

**PARA OPTAR AL GRADO ACADÉMICO DE
MAESTRO EN ESTADÍSTICA**

ASESOR: Dr. ALFREDO VALENCIA TOLEDO

CUSCO - PERÚ

2020

DEDICATORIA

A mis padres Samuel y Marina por sus enseñanzas que coadyuvan a dirigir mi vida

*A mi hermana Mildred y mi abuela Rosenda(†)
por sus buenos consejos*

A mi estimado amigo Marcelo (†)

A todos los programadores que propugnan el uso del software libre como sinónimo de libertad y solidaridad.

AGRADECIMIENTO

A la Escuela de Posgrado de la Universidad Nacional de San Antonio Abad del Cusco por la instrucción recibida a través de la Maestría en Estadística y sus docentes.

Al Dr. Alfredo Valencia Toledo, docente del departamento académico de Matemática y Estadística de la Universidad Nacional de San Antonio Abad del Cusco, por la asesoría y el tiempo dedicado para la realización del presente trabajo.

Al Q.F. Néstor Arzubialde Zamalloa, jefe del departamento de farmacia de hospital Regional del Cusco por la entrega del consolidado de la demanda de los productos farmacéuticos y afines de los años 2013 al 2018.

A los dictaminantes, el Dr. Cleto de la Torre Dueñas y al Mg. Guillermo Paucar Carlos, por sus observaciones pertinentes para la mejora del presente trabajo.

Al Ingeniero en Estadística e Informática Eulios Mamani Barrios, por su apoyo en el manejo del software libre R

Por último, a mi familia por su apoyo para la conclusión de la maestría.

PRESENTACIÓN

Señor Director de la Escuela de Posgrado de la Universidad Nacional de San Antonio Abad del Cusco, Señores miembros del Jurado examinador:

En cumplimiento de lo establecido en el Reglamento de Grados de la Escuela de Posgrado de la Universidad Nacional de San Antonio Abad del Cusco, presento la tesis intitulada “**MODELOS AUTORREGRESIVOS INTEGRADOS DE MEDIA MÓVIL Y RED NEURONAL RECURRENTE PARA PREDECIR LA DEMANDA DE MEDICAMENTOS, HOSPITAL REGIONAL CUSCO, 2013 - 2018**”, para optar al grado académico de Maestro en Estadística.

La investigación que se presenta, pretende ofrecer un instrumento alternativo que permita pronosticar la demanda de los medicamentos con el fin de disminuir el desabastecimiento de los mismos.

RESUMEN

El objetivo principal de la presente investigación es analizar los modelos autorregresivo integrado de media móvil (ARIMA) o la red neuronal recurrente tipo de memoria a corto y largo plazo (RNNLSTM) para predecir adecuadamente la demanda de medicamentos en el Hospital Regional del Cusco, 2018. El estudio es de tipo descriptivo y predictiva-longitudinal, donde se utilizó los modelos ARIMA y RNNLSTM. La población y la muestra de estudio está constituida por cinco medicamentos con mayor demanda entre enero 2013 a diciembre 2017. Para el procesamiento de los datos se utiliza el software libre R Project (R). Los resultados obtenidos son los siguientes: en cuanto a los modelos ARIMA se tiene Cloruro de sodio 900 mg/1000 ml inyectable con un MAPE de 4.50; Ibuprofeno 400 mg tableta con un MAPE de 27.57; Metamizol sódico 1g/2 ml inyectable con un MAPE de 4.66; Paracetamol 500 mg tableta con un MAPE de 23.47; y Sevoflurano 250 ml solución con un MAPE de 10.54. Respecto a los modelos RNNLSTM se tiene Cloruro de sodio 900 mg/1000 ml inyectable con un MAPE de 4.78; Ibuprofeno 400 mg tableta con un MAPE de 25.00; Metamizol sódico 1g/2 ml inyectable con un MAPE de 1.85; Paracetamol 500 mg tableta con un MAPE de 22.19; y Sevoflurano 250 ml solución con un MAPE de 9.53. Se concluye que el modelo de la RNNLSTM permite predecir adecuadamente la demanda de medicamentos frente al modelo ARIMA con una reducción del MAPE en promedio de 1.92 y 0.28 respectivamente.

Palabras clave: autorregresivo, media móvil, red neuronal, error porcentual, medicamentos.

ABSTRACT

The main aim of the present research is to analyze the models integrated autoregressive of moving average (ARIMA) or the recurrent neural network type of short and long-term memory (RNNLSTM) to adequately predict the demand for medicines in the Regional Hospital of Cusco, 2018. This study is descriptive and predictive-longitudinal, where ARIMA and RNNLSTM models were used. The population and the study sample consisted of five drugs with the highest demand between January 2013 and December 2017 R Project (R) is used to process the data. The results achieved are as follows: regarding the ARIMA models, sodium chloride 900 mg / 1000 ml for injection with a MAPE of 4.50; Ibuprofen 400 mg tablet with a MAPE of 27.57; Metamizole sodium 1g / 2 ml for injection with a MAPE of 4.66; Paracetamol 500 mg tablet with a MAPE of 23.47; and Sevoflurane 250 ml solution with a MAPE of 10.54. Concerning the RNNLSTM models, sodium chloride 900 mg / 1000 ml for injection with a MAPE of 4.78; Ibuprofen 400 mg tablet with a MAPE of 25.00; Metamizole sodium 1g / 2 ml injectable with a MAPE of 1.85; Paracetamol 500 mg tablet with a MAPE of 22.19; and Sevoflurane 250 ml solution with a MAPE of 9.53. It is concluded that the RNNLSTM model allows to adequately predict the demand for drugs compared to the ARIMA model with a reduction in MAPE on average of 1.92 and 0.28 respectively.

Key words: autoregressive, moving average, neural network, percentage error, drugs.

ÍNDICE

DEDICATORIA.....	ii
AGRADECIMIENTO.....	iii
PRESENTACIÓN.....	iv
RESUMEN.....	v
ABSTRACT.....	vi
ÍNDICE.....	vii
INTRODUCCIÓN.....	1

CAPÍTULO I

PLANTEAMIENTO DEL PROBLEMA

1.1 SITUACIÓN PROBLEMÁTICA.....	3
1.2 FORMULACIÓN DEL PROBLEMA.....	4
1.2.1 Problema general.....	4
1.2.2 Problemas específicos.....	4
1.3 ANTECEDENTES.....	4
1.4 JUSTIFICACIÓN.....	7
1.5 OBJETIVOS.....	8
1.5.1 Objetivo general.....	8
1.5.2 Objetivos específicos.....	8
1.6 HIPÓTESIS.....	8
1.6.1 Hipótesis general.....	8
1.6.2 Hipótesis específicas.....	9

CAPITULO II

MARCO TEÓRICO

2.1 BASES TEÓRICAS.....	10
2.1.1 Series de tiempo.....	10
2.1.1.1 Definición.....	10
2.1.1.2. Componentes.....	11
2.1.1.3 Esquema de integración.....	11

2.1.2 Serie Temporal Univariada.....	11
2.1.3 Procesos estocásticos.....	13
2.1.3.1 Definición.....	13
2.1.3.2 Propiedades.....	13
2.1.3.3 Caracterización.....	14
2.1.3.4 Momentos.....	14
2.1.4 Procesos estocásticos estacionarios.....	15
2.1.4.1 Estacionariedad estricta.....	15
2.1.4.2 Estacionariedad débil.....	15
2.1.5 Proceso ruido blanco.....	17
2.1.6 Modelos autorregresivos integrados de media móvil (ARIMA).....	17
2.1.6.1 Modelo autorregresivo (AR(p)).....	17
2.1.6.1.1 Modelo AR(1).....	18
2.1.6.1.2 Modelo AR(2).....	20
2.1.6.2 Modelo de media móvil (MA(q)).....	22
2.1.6.2.1 Modelo MA(1).....	23
2.1.6.2.2 Modelo MA(2).....	25
2.1.6.3 Modelo autorregresivo de media móvil (ARMA(p,q)).....	26
2.1.6.3.1 Modelo ARMA(1,1).....	27
2.1.6.4 Modelo autorregresivos integrado de media móvil (ARIMA(p,d,q)).....	29
2.1.6.4.1 Transformación Box-Cox.....	29
2.1.6.4.2 Diferenciación.....	30
2.1.6.4.3 Modelamiento ARIMA.....	31
2.1.6.5 Modelo autorregresivos integrado de media móvil estacional (ARIMA(p,d,q) (P,D,Q) _s).....	32
2.1.6.5.1 Series estacionales.....	32
2.1.6.5.2 Modelamiento ARIMA(p,d,q) (P,D,Q) _s	33
2.1.7 Metodología Box-Jenkins.....	34
2.1.7.1 Identificación.....	35
2.1.7.2 Estimación.....	36
2.1.7.2.1 Estimación por Máxima Verosimilitud exacta o no condicional (MVE).....	36
2.1.7.3 Validación.....	39

2.1.7.3.1	Análisis de los residuales.....	39
2.1.7.3.2	Bondad de ajuste.....	40
2.1.7.3.3	Criterio de información Akaike (AIC).....	40
2.1.7.4	Predicción.....	40
2.1.7.4.1	Error porcentual absoluto medio (MAPE).....	41
2.1.7.4.2	Raíz del error cuadrático medio (RMSE).....	41
2.1.8	Modelo de la red neuronal recurrente de corta y larga memoria (RNNLSTM).....	42
2.1.8.1	Cuestiones previas sobre las redes neuronales artificiales.....	42
2.1.8.1.1	¿Qué son las redes neuronales artificiales?.....	42
2.1.8.1.2	Perceptrón simple.....	42
2.1.8.1.3	Componentes de redes neuronales.....	43
2.1.8.1.3.1	Neurona.....	43
2.1.8.1.3.2	Arquitectura de la red neuronal.....	44
2.1.8.2	Red Neuronal Recurrente (RNN).....	47
2.1.8.2.1	Estructura.....	47
2.1.8.2.2	Entrenamiento de la red.....	49
2.1.8.2.2.1	Propagación hacia atrás dinámica.....	49
2.1.8.2.3	Gradientes explosivos y gradientes desvanecidos.....	52
2.1.8.3	Redes Neuronales Recurrentes de Memoria a corto y largo plazo (RNNLSTM).....	52
2.1.8.3.1	Arquitectura de la RNNLSTM.....	52
2.1.8.3.2	Arquitectura del bloque memoria.....	53
2.1.8.3.2.1	Componentes del bloque memoria.....	56
2.1.8.3.3	Funcionamiento de la red.....	56
2.1.8.3.4	Entrenamiento de la red.....	57
2.1.8.3.4.1	Propagación hacia adelante.....	57
2.1.8.3.4.2	Función del error cuadrático medio.....	58
2.1.8.3.4.3	Propagación hacia atrás.....	58
2.1.8.3.4.4	Descenso del gradiente.....	59
2.1.8.3.4.5	Actualización de los pesos.....	59
2.1.8.3.5	Propagación hacia atrás de la RNNLSTM.....	60
2.2	MARCO CONCEPTUAL.....	62
2.2.1	Series de tiempo.....	62

2.2.2 Red neuronal.....	63
2.2.3 Demanda de medicamentos.....	63
2.2.4 Tiempo.....	63

CAPÍTULO III METODOLOGÍA

3.1 AMBITO DE ESTUDIO.....	64
3.1.1 Localización política.....	64
3.1.2 Localización geográfica.....	64
3.2 TIPO Y NIVEL DE INVESTIGACIÓN.....	64
3.2.1 Tipo de investigación.....	64
3.2.2 Nivel de investigación.....	65
3.3 IDENTIFICACIÓN Y OPERACIONALIZACIÓN DE VARIABLES E INDICADORES.....	65
3.3.1 Identificación de variables.....	65
3.3.2 Operacionalización de variables.....	66
3.4 UNIDAD DE ANÁLISIS.....	66
3.5 POBLACIÓN DE ESTUDIO.....	66
3.6 TAMAÑO DE LA MUESTRA.....	66
3.7 TÉCNICA DE RECOLECCIÓN DE LA INFORMACIÓN.....	67
3.8 VALIDACIÓN Y PREDICCIÓN.....	67
3.9 ANÁLISIS COMPARATIVO PARA DEMOSTRAR LA VERDAD O FALSEDAD DE LAS HIPÓTESIS PLANTEADAS.....	68
3.10 PROCEDIMIENTO.....	68
3.11 PROCESAMIENTO DE DATOS.....	69
3.11.1 Modelo de serie de tiempo autorregresivo integrado de media móvil.....	69
3.11.2 Modelo de red neuronal recurrente de memoria a corto y largo plazo.....	77

CAPÍTULO IV RESULTADOS

IDENTIFICACIÓN DE LOS MEDICAMENTOS CON MAYOR DEMANDA.....	84
DEMANDA DE CLORURO DE SODIO (DCS) 900 MG/1000 ML INYECTABLE.....	86
DEMANDA DE IBUPROFENO (DI) 400 MG TABLETA.....	100
DEMANDA DE METAMIZOL SODICO (DMS) 1 G - 2 ML INYECTABLE.....	114
DEMANDA DE PARACETAMOL (DP) 500 MG TABLETA.....	128
DEMANDA DE SEVOFLURANO (DS) 250 ML SOLUCIÓN.....	142
DISCUSIÓN.....	158
CONCLUSIONES.....	162
RECOMENDACIONES.....	163
REFERENCIAS.....	164
ANEXOS.....	176

ÍNDICE DE TABLAS

IDENTIFICACIÓN DE LOS MEDICAMENTOS CON MAYOR DEMANDA

Tabla 1	
<i>Demanda de medicamentos del año 2013 al 2017.....</i>	<i>84</i>

DEMANDA DE CLORURO DE SODIO (DCS) 900 MG/1000 ML INYECTABLE

Tabla 2	
<i>Estadísticos descriptivos de la DCS por año.....</i>	<i>86</i>
Tabla 3	
<i>Estadísticos descriptivos de toda la serie de la DCS</i>	<i>87</i>
Tabla 4	
<i>Pruebas de estacionariedad, tendencia, estacionalidad y normalidad de la DCS.....</i>	<i>90</i>
Tabla 5	
<i>Identificación del modelo para la DCS.....</i>	<i>93</i>
Tabla 6	
<i>Modelo ARIMA para la DCS.....</i>	<i>93</i>
Tabla 7	
<i>Pruebas de normalidad de los residuos y comportamiento de ruido blanco del modelo ARIMA (0,1,1) de la DCS.....</i>	<i>94</i>
Tabla 8	
<i>Ppronóstico del modelo ARIMA para la DCS.....</i>	<i>95</i>
Tabla 9	
<i>Valores reales y pronosticados de la DCS.....</i>	<i>96</i>
Tabla 10	
<i>Estimación del modelo de la RNNLSTM de la DCS.....</i>	<i>97</i>
Tabla 11	
<i>Valores reales y pronosticados de la DCS.....</i>	<i>98</i>
Tabla 12	
<i>Pronóstico de los modelos para la DCS.....</i>	<i>99</i>

DEMANDA DE IBUPROFENO (DI) 400 MG TABLETA

Tabla 13	
<i>Estadísticos descriptivos de la DI por año.....</i>	<i>100</i>
Tabla 14	
<i>Estadísticos descriptivos de toda la serie de la DI.....</i>	<i>101</i>
Tabla 15	
<i>Pruebas de estacionariedad, tendencia, estacionalidad y normalidad de la DI.....</i>	<i>104</i>
Tabla 16	
<i>Identificación del modelo para la DI.....</i>	<i>107</i>
Tabla 17	
<i>Modelo ARIMA para la DI.....</i>	<i>107</i>
Tabla 18	
<i>Pruebas de normalidad de los residuos y comportamiento de ruido blanco del modelo ARIMA (1,0,1) de la DI.....</i>	<i>108</i>
Tabla 19	
<i>Pronóstico del modelo ARIMA para la DI.....</i>	<i>109</i>
Tabla 20	
<i>Valores reales y pronosticados de la DI.....</i>	<i>110</i>
Tabla 21	
<i>Estimación del modelo de la RNNLSTM de la DI.....</i>	<i>111</i>
Tabla 22	
<i>Valores reales y pronosticados de la DI.....</i>	<i>112</i>
Tabla 23	
<i>Pronóstico de los modelos para la DI.....</i>	<i>113</i>

DEMANDA DE METAMIZOL SODICO (DMS) 1 G - 2 ML INYECTABLE

Tabla 24	
<i>Estadísticos descriptivos de la DMS por año.....</i>	<i>114</i>
Tabla 25	
<i>Estadísticos descriptivos de toda la serie de la DMS.....</i>	<i>115</i>
TABLA 26	
<i>Pruebas de estacionariedad, tendencia, estacionalidad y normalidad de la DMS.....</i>	<i>118</i>

Tabla 27	
<i>Identificación del modelo para la DMS.....</i>	121
Tabla 28	
<i>Modelo ARIMA para la DMS.....</i>	121
Tabla 29	
<i>Pruebas de normalidad de los residuos y comportamiento de ruido blanco del modelo ARIMA (3,1,1)(0,0,1)₁₂ de la DMS.....</i>	122
Tabla 30	
<i>Pronóstico del modelo ARIMA para la DMS.....</i>	123
Tabla 31	
<i>Valores reales y pronosticados de la DMS.....</i>	124
Tabla 32	
<i>Estimación del modelo de la RRNNLSTM de la DMS.....</i>	125
Tabla 33	
<i>Valores reales y pronosticados de la DMS.....</i>	126
Tabla 34	
<i>Pronóstico de los modelos para la DMS.....</i>	127
DEMANDA DE PARACETAMOL (DP) 500 MG TABLETA	
Tabla 35	
<i>Estadísticos descriptivos de la DP por año.....</i>	128
Tabla 36	
<i>Estadísticos descriptivos de toda la serie de la DP.....</i>	129
Tabla 37	
<i>Pruebas de estacionariedad, tendencia, estacionalidad y normalidad de la DP.....</i>	132
Tabla 38	
<i>Identificación del modelo para la DP.....</i>	135
Tabla 39	
<i>Modelo ARIMA para la DP.....</i>	135
Tabla 40	
<i>Pruebas de normalidad de los residuos y comportamiento de ruido blanco del modelo ARIMA (1,1,0)(0,0,1)₁₂ de la DP.....</i>	136

Tabla 41	
<i>Pronóstico del modelo ARIMA para la DP.....</i>	<i>137</i>
Tabla 42	
<i>Valores reales y pronosticados de la DP.....</i>	<i>138</i>
Tabla 43	
<i>Estimación del modelo de la RNNLSTM de la DP.....</i>	<i>139</i>
Tabla 44	
<i>Valores reales y pronosticados de la DP.....</i>	<i>140</i>
Tabla 45	
<i>Pronóstico de los modelos para la DP.....</i>	<i>141</i>
DEMANDA DE SEVOFLURANO (DS) 250 ML SOLUCIÓN	
Tabla 46	
<i>Estadísticos descriptivos de la DS por año.....</i>	<i>142</i>
Tabla 47	
<i>Estadísticos descriptivos de toda la serie de la DS.....</i>	<i>143</i>
Tabla 48	
<i>Pruebas de estacionariedad, tendencia, estacionalidad y normalidad de la DS.....</i>	<i>146</i>
Tabla 49	
<i>Identificación del modelo para la DS.....</i>	<i>149</i>
Tabla 50	
<i>Modelo ARIMA para la DS.....</i>	<i>149</i>
Tabla 51	
<i>Pruebas de normalidad de los residuos y comportamiento de ruido blanco del modelo ARIMA (3,1,2) de la DS.....</i>	<i>150</i>
Tabla 52	
<i>Pronóstico del modelo ARIMA para la DS.....</i>	<i>151</i>
Tabla 53	
<i>Valores reales y pronosticados de la DS.....</i>	<i>152</i>
Tabla 54	
<i>Estimación del modelo de la RNNLSTM de la DS.....</i>	<i>153</i>
Tabla 55	
<i>Valores reales y pronosticados de la DS.....</i>	<i>154</i>

Tabla 56	
<i>Pronóstico de los modelos para la DS.....</i>	155
Tabla 57	
<i>Pronóstico de los modelos ARIMA y RNNLSTM</i>	
<i>en función del MAPE.....</i>	156
Tabla 58	
<i>Pronóstico de los modelos ARIMA y RNNLSTM</i>	
<i>en función del RMSE.....</i>	157

ÍNDICE DE FIGURAS

<i>Figura 1. Esquema de la Metodología Box-Jenkins</i>	35
<i>Figura 2. Perceptrón simple</i>	43
<i>Figura 3. Red simétrica totalmente conectada</i>	45
<i>Figura 4. Redes en capas ocultas</i>	46
<i>Figura 5. Red acíclicas</i>	46
<i>Figura 6. Red de alimentación hacia adelante</i>	47
<i>Figura 7. Red neuronal recurrente</i>	48
<i>Figura 8. Arquitectura de una RNNLSTM simple</i>	53
<i>Figura 9. Propagación hacia adelante del bloque memoria de una RNNLSTM</i>	54
<i>Figura 10. Propagación hacia atrás del bloque memoria de una RNNLSTM</i>	60
<i>Figura 11. Porcentaje de la demanda de medicamentos del año 2013 al 2017</i>	84
DEMANDA DE CLORURO DE SODIO (DCS) 900 MG/1000 ML INYECTABLE	
<i>Figura 12. DCS mensual de enero 2013 a diciembre de 2017</i>	86
<i>Figura 13. Función de autocorrelación de la DCS</i>	87
<i>Figura 14. Diagrama de caja de la DCS</i>	88
<i>Figura 15. Diagrama de caja por mes de DCS</i>	88
<i>Figura 16. Descomposición de la DCS</i>	89
<i>Figura 17. Evaluación de la normalidad de la DCS</i>	91
<i>Figura 18. Función de autocorrelación simple y parcial de la serie con transformación BoxCox y diferenciación de la DCS</i>	92
<i>Figura 19. Evaluación de la normalidad y comportamiento de ruido blanco de los residuos del modelo ARIMA (0,1,1) de la DCS</i>	94
<i>Figura 20. Valores del entrenamiento y pronóstico del modelo ARIMA (0,1,1) para la DCS</i>	96
<i>Figura 21. Valores del entrenamiento y pronóstico del modelo de la RNNLSTM para la DCS</i>	98
<i>Figura 22. Serie con datos reales, entrenamiento y pronóstico de los modelos ARIMA y RNNLSTM</i>	99

DEMANDA DE IBUPROFENO (DI) 400 MG TABLETA

<i>Figura 23.</i> DI mensual de enero 2013 a diciembre de 2017.....	100
<i>Figura 24.</i> Función de autocorrelación de la DI.....	101
<i>Figura 25.</i> Diagrama de caja de la DI.....	102
<i>Figura 26.</i> Diagrama de caja por mes de DI.....	102
<i>Figura 27.</i> Descomposición de la DI.....	103
<i>Figura 28.</i> Evaluación de la normalidad de la DI.....	105
<i>Figura 29.</i> Función de autocorrelación simple y parcial de la serie de la DI.....	106
<i>Figura 30.</i> Evaluación de la normalidad y comportamiento de ruido blanco de los residuos del modelo ARIMA (1,0,1) de la DI.....	108
<i>Figura 31.</i> Valores del entrenamiento y pronóstico del modelo ARIMA (1,0,1) para la DI.....	110
<i>Figura 32.</i> Valores del entrenamiento y pronóstico del modelo de la RNNLSTM para la DI.....	112
<i>Figura 33.</i> Serie con datos reales, entrenamiento y pronóstico de los modelos ARIMA y RNNLSTM.....	113

DEMANDA DE METAMIZOL SODICO (DMS) 1 G - 2 ML INYECTABLE

<i>Figura 34.</i> DMS mensual de enero 2013 a diciembre de 2017.....	114
<i>Figura 35.</i> Función de autocorrelación de la DMS.....	115
<i>Figura 36.</i> Diagrama de caja de la DMS.....	116
<i>Figura 37.</i> Diagrama de caja por mes de la DMS.....	116
<i>Figura 38.</i> Descomposición de la DMS.....	117
<i>Figura 39.</i> Evaluación de la normalidad de la DMS.....	119
<i>Figura 40.</i> Función de autocorrelación simple y parcial de la serie con diferenciación de la DMS.....	120
<i>Figura 41.</i> Evaluación de la normalidad y comportamiento de ruido blanco de los residuos del modelo ARIMA (3,1,1)(0,0,1) ₁₂ de la DMS.....	122
<i>Figura 42.</i> Valores del entrenamiento y pronóstico del modelo ARIMA (3,1,1)(0,0,1) ₁₂ para la DMS.....	124
<i>Figura 43.</i> Valores del entrenamiento y pronóstico del modelo de la RNNLSTM para la DMS.....	126

Figura 44. Serie con datos reales, entrenamiento y pronóstico de los modelos ARIMA y RNNLSTM.....	127
DEMANDA DE PARACETAMOL (DP) 500 MG TABLETA	
Figura 45. DP mensual de enero 2013 a diciembre de 2017.....	128
Figura 46. Función de autocorrelación de la DP.....	129
Figura 47. Diagrama de caja de la DP.....	130
Figura 48. Diagrama de caja por mes de DP.....	130
Figura 49 Descomposición de la DP.....	131
Figura 50. Evaluación de la normalidad de la DP.....	133
Figura 51. Función de autocorrelación simple y parcial de la serie con transformación BoxCox y diferenciación de la DP.....	134
Figura 52. Evaluación de la normalidad y comportamiento de ruido blanco de los residuos del modelo ARIMA (1,1,0)(0,0,1)₁₂ de la DP.....	136
Figura 53. Valores del entrenamiento y pronóstico del modelo ARIMA (1,1,0)(0,0,1)₁₂ para la DP.....	138
Figura 54. Valores del entrenamiento y pronóstico del modelo de la RNNLSTM para la DP.....	140
Figura 55. Serie con datos reales, entrenamiento y pronóstico de los modelos ARIMA y RNNLSTM.....	141
DEMANDA DE SEVOFLURANO (DS) 250 ML SOLUCIÓN	
Figura 56. DS mensual de enero 2013 a diciembre de 2017.....	142
Figura 57. Función de autocorrelación de la DS.....	143
Figura 58. Diagrama de caja de la DS.....	144
Figura 59. Diagrama de caja por mes de la DS.....	144
Figura 60 Descomposición de la DS.....	145
Figura 61. Evaluación de la normalidad de la DS.....	147
Figura 62. Función de autocorrelación simple y parcial de la serie con diferenciación de la DS.....	148
Figura 63. Evaluación de la normalidad y comportamiento de ruido blanco de los residuos del modelo ARIMA (3,1,2) de la DS.....	151
Figura 64. Valores del entrenamiento y pronóstico del modelo	

ARIMA (3,1,2) para la DS.....	152
<i>Figura 65. Valores del entrenamiento y pronóstico del modelo de la RNNLSTM para la DS.....</i>	<i>154</i>
<i>Figura 66. Serie con datos reales, entrenamiento y pronóstico de los modelos ARIMA y RNNLSTM.....</i>	<i>155</i>

ABREVIATURAS

- RNNLSTM Red neuronal recurrente tipo de memoria a corto y largo plazo
- AR Autoregresivo
- AR(P) Autoregresivo de orden p
- ARMA Autorregresivo de media móvil
- ARMA(p,q) Autorregresivo de orden p y de media móvil de orden q
- ARIMA Autorregresivo integrado de media móvil
- ARIMA(p,d,q) Autorregresivo de orden p, integración de orden d y de media móvil de orden q
- SARIMA Autorregresivo integrado de media móvil estacional
- ARIMA(P,D,Q)_s Autorregresivo estacional P; integración estacional D; media y media móvil estacional Q
- RMSE Raíz del error cuadrático medio
- MAPE Error porcentual absoluto medio
- DCS Demanda de cloruro de sodio
- DMS Demanda de metamizol sódico
- DI Demanda de ibuprofeno
- DP Demanda de paracetamol
- DS Demanda de sevoflurano
- AIC Criterio de información Akaike
- BIC Criterio de información Bayesiano

INTRODUCCIÓN

La carencia de medicamentos esenciales es un problema de salud que se presenta a nivel mundial, el cual se ha venido incrementando en los últimos años. Se afirma que es un problema de salud porque tiene repercusiones peligrosas en la salud de los pacientes debido a la falta de medicamentos para tratar la afección que presenta y los posibles errores que se originan al intentar sustituir los medicamentos faltantes. Uno de los factores que contribuye a la escasez de medicamentos es el inadecuado aprovisionamiento frente a la demanda de los pacientes (WHO, 2016).

Sin embargo, el disponer de herramientas metodológicas adecuadas de estadística, como los modelos Autorregresivo Integrado de Media Móvil (ARIMA) y la Red Neuronal Recurrente tipo de Memoria a Corto y Largo Plazo (RNNLSTM) para pronosticar la demanda de los medicamentos a futuro, contribuiría a disminuir la escasez de los medicamentos en los establecimientos de salud. Por tanto, en la presente investigación se plantea utilizar ambos modelos para estudiar la demanda de medicamentos en el hospital regional del Cusco en el 2018.

El proyecto de investigación está estructurado en cuatro capítulos, que a continuación se detalla:

En el primer capítulo el planteamiento del problema, contiene la situación problemática, formulación del problema, antecedentes, justificación, objetivos y las hipótesis.

En el segundo capítulo el marco teórico, se expone las bases teóricas y el marco conceptual.

En el tercer capítulo la metodología, contiene el ámbito de estudio, tipo y nivel de investigación, identificación y operacionalización de las variables e indicadores, unidad de análisis, población de estudio, tamaño de la muestra, técnica de recolección de la información, validación y predicción, análisis comparativo para demostrar la verdad o falsedad de las hipótesis planteadas, procedimiento y procesamiento de datos.

En el cuarto capítulo los resultados, los cuales se presentan a través de tablas y gráficos con su respectiva interpretación.

Finalmente se presenta la discusión, las conclusiones, las recomendaciones, las referencias y los anexos.

CAPÍTULO I

PLANTEAMIENTO DEL PROBLEMA

1.1 SITUACIÓN PROBLEMÁTICA

El desabastecimiento de medicamentos en los establecimientos de salud es un problema que se presenta a nivel mundial el cual afecta seriamente la salud y la economía de los pacientes (Agirrezabala et al., 2015).

Los grupos terapéuticos de los medicamentos con mayor desabastecimiento son los anestésicos, antibióticos, analgésicos, productos nutricionales y electrolíticos, y agentes de quimioterapia; a nivel clínico dicho problema causo en el país de Estados Unidos 15 muertes documentadas (Fox, Sweet y Jensen, 2014).

Desabastecimiento de medicamentos a nivel nacional en general, se tiene para el año 2015 con 13.38 %, 2016 con 10.86 y enero a mayo del 2017 con 4.8 % (DIGEMID, 2018a).

El porcentaje de desabastecimiento de medicamentos en la Dirección Regional de Salud del Cusco a nivel de los Establecimiento de Salud de julio a diciembre 2017 con 1.18 % y enero a junio de 2018 con 3.08 % (DIGEMID, 2018b).

Con respecto al desabastecimiento de medicamentos en el Hospital Regional del Cusco se tiene para los meses de agosto a diciembre de 2016 con 3.99 %, para el año 2017 con 16.96 % (DIGEMID, 2018c).

Asimismo, indicar que, en el Hospital Regional del Cusco el Área de Almacenamiento y Gestión de la Programación de productos farmacéuticos, dispositivos médicos quirúrgicos y productos sanitarios, es la encargada de realizar y aplicar los métodos de estimación de necesidades y programación de

productos farmacéuticos y afines de acuerdo al Método de Consumo Histórico Ajustado¹.

Con la información antes mencionada se verifica que el problema del desabastecimiento de medicamentos en el Hospital Regional del Cusco es un problema actual y vigente.

1.2 FORMULACIÓN DEL PROBLEMA

1.2.1 Problema general

¿Cuál de los modelos de pronóstico, entre el modelo autorregresivo integrado de media móvil y la red neuronal recurrente permite pronosticar adecuadamente la demanda de los medicamentos en el Hospital Regional del Cusco?

1.2.2 Problemas específicos

- ¿Qué modelos autorregresivo integrado de media móvil proporcionan un mejor ajuste para el pronóstico de la demanda de los medicamentos?
- ¿Qué modelos red neuronal recurrente proporcionan un mejor ajuste para el pronóstico de la demanda de los medicamentos?

1.3 ANTECEDENTES

Kaushik et al. (2017) utilizaron la metodología de los métodos de pronóstico como el modelo autorregresivo integrado de promedio móvil (ARIMA), la red neuronal de larga memoria a corto plazo estándar (LSTM) y la red neuronal de LSTM apilado para predecir el gasto mensual en medicamentos para más de 50,000 pacientes en los EE.UU. entre los años 2011 y 2015. Los resultados en relación al modelo ARIMA con los mejores parámetros encontrados fueron $p=2$, $d=1$ y $q=1$, con una raíz del error cuadrático medio (RMSE) de 16.855; en cuanto

1 Manual de Organización y Funciones del Servicio de Farmacia del Hospital Regional del Cusco 2016.

al modelo LSTM estándar se tuvo un RMSE de 14.617 con 7 celdas de memoria oculta, y para el modelo LSTM apilado con un RMSE de 13.693 con 4 celdas de memoria oculta. Obteniendo la conclusión que con el modelo LSTM apilado se obtiene mejores predicciones frente al modelo LSTM estándar y el modelo ARIMA.

Li y Cao (2018) emplearon la metodología de los métodos de pronóstico como el modelo auto regresión integrada de promedio móvil (ARIMA), la red neuronal de propagación de regreso (BPNN), la red neuronal de larga memoria a corto plazo simple (LSTM) y la red neuronal de larga memoria a corto plazo apilada (LSTM) para predecir el flujo de turismo de enero 2013 a diciembre 2015 del museo de Xi'an de la ciudad de Shaanxi en China. Los resultados en cuanto al modelo ARIMA con una raíz del error cuadrático medio (RMSE) de 1488.0 y un error absoluto medio (MAE) de 962.37; con respecto a la red neuronal BPN con un RMSE de 1365.35 y un MAE de 912.70; en relación al modelo de red neuronal LSTM simple un RMSE de 988.4 y un MAE de 702.11; en cuanto a la red neuronal LSTM apilada con un RMSE de 1195.76 y un MAE de 837.05. Arribando a la conclusión que el modelo de la red neuronal LSTM simple es más efectiva en la predicción frente al modelo ARIMA, al modelo de red neuronal de BPN y al modelo de red neuronal LSTM apilado.

Li et al. (2017) utilizaron la metodología de los métodos de pronóstico como el modelo auto regresión de promedio móvil (ARMA), la red neuronal de memoria larga a corto plazo extendido (LSTME), la red neuronal de larga memoria a corto plazo simple (LSTM), el modelo de aprendizaje profundo espacio temporal (STDL), el modelo de red neuronal de retraso de tiempo (TDNN) y el modelo de regresión vectorial de apoyo (SVR), para la predicción de la concentración de contaminantes en el aire de la ciudad de Beijing de China del 01/01/2014 al 28/05/2016. Los resultados con respecto al modelo ARMA con una raíz del error cuadrático medio (RMSE) de 24.40 y un error porcentual absoluto medio (MAPE) de 27.54; en cuanto al modelo de LSTME con una RMSE de 12.60 y un MAPE de 11.93, en relación al modelo de LSTM con una RMSE de 17.94 y un MAPE de 15.84; en cuanto al modelo de STDL con una RMSE de 14.96 y un MAPE de 21.75; en cuanto al modelo de TDNN con una RMSE de 16.19 y un

MAPE de 26.87 y en cuanto al modelo de SVR con una RMSE de 22.04 y un MAPE de 28.45. Llegando a la conclusión que el modelo de la red neuronal LSTM es capaz de modelar series de tiempo con dependencias de larga duración y puede determinar automáticamente los retardos de tiempo óptimos frente a los modelos LSTM, ARMA, STDL, TDNN Y SVR.

Fu, Zhang y Li (2016) utilizaron la metodología de los métodos de pronóstico como el modelo de auto regresión integrado de promedio móvil (ARIMA), el modelo de red neuronal de memoria a corto plazo a largo plazo (LSTM) y los métodos de redes neuronales de unidades recurrentes seguras (GRU), para predecir el flujo de tráfico de transporte utilizando la base de datos del sistema de medición de rendimiento del área de la Bahía, Alameda, Oakland, EE. UU.. Los resultados con respecto al modelo ARIMA con un error cuadrático medio (MSE) de 841.0065 y un error absoluto medio (MAE) de 19.1753; en cuanto al modelo de red neuronal LSTM con un MSE de 710.0502 y un MAE de 18.127758 y en relación al modelo de red neuronal GRU con un MSE de 668.9304 y un MAE de 17.2116. Obteniendo la conclusión que el modelo de la red neuronal GRU tiene un mejor rendimiento de predicción frente al modelo de red neuronal de LSTM y el modelo ARIMA.

Adamowski et al. (2012) emplearon la metodología de los métodos de pronóstico como el modelo de regresión lineal múltiple, el modelo de regresión no lineal, múltiple, el modelo de media móvil integrada autorregresiva, el modelo de la red neuronal artificial y el modelo red neuronal artificial wavelet, para la previsión de demanda de agua urbana en Montreal, Canadá. Los resultados con respecto al modelo de regresión lineal múltiple con una raíz del error cuadrático medio (RMSE) de 0.059 y un error cuadrático medio relativo (RRMSE) de 3.491; en cuanto al modelo de regresión no lineal múltiple con un RMSE de 0.058 y un RRMSE de 3.323, en relación al modelo de media móvil integrada autorregresiva con un RMSE de 0.045 y un RRMSE de 2.663, en cuanto al modelo de redes neuronales artificiales con un RMSE de 0.035 y un RRMSE de 2.056 y en relación al modelo de redes neuronales artificiales wavelet con un RMSE de 0.027 y un RRMSE de 1.591. Obteniendo la conclusión que el modelo de la red neuronal wavelet tiene un mejor rendimiento de predicción frente a los modelos

de redes neuronales artificiales, de media móvil integrada autorregresiva, de regresión no lineal múltiple y de regresión lineal múltiple.

1.4 JUSTIFICACIÓN

El derecho a la salud esta subsumido dentro de los derechos de segunda generación que son los económicos, sociales y culturales, asimismo está reconocido en la Declaración Universal de los Derechos Humanos de la Organización de las naciones Unidas, de este hecho tan trascendental a la fecha ya han transcurrido cien años (Quijano, 2016).

En el artículo 25 numeral 1 de la Declaración Universal de los Derechos Humanos (1948) señala “Toda persona tiene derecho a un nivel de vida adecuado que le asegure, así como a su familia, la salud y el bienestar, y en especial la alimentación, el vestido, la vivienda, la asistencia médica y los servicios sociales necesarios...”, por lo tanto para el ejercicio total del derecho a la salud de los ciudadanos se debe considerar a la infraestructura, el personal asistencial y primordialmente el acceso a medicamentos eficaces, seguros y a bajo costo (Cordero, 2006).

La Constitución Política del Perú (1993) en los artículos 7 Derecho a la salud, donde señala “Todos tienen derecho a la protección de su salud...” y el artículo 9 Política Nacional de salud, donde indica “El Estado determina la política nacional de salud. El poder Ejecutivo norma y supervisa su aplicación. Es responsable de diseñarla y conducirla en forma plural y descentralizada para facilitar a todos los accesos equitativos a los servicios de salud.”

Asimismo, la Ley General de Salud 26842 (1997), Título preliminar, numeral VI señala “Es de interés público la provisión de servicios de salud, cualquiera sea la persona o institución que los provea. Es responsabilidad del Estado promover las condiciones que garanticen una adecuada cobertura de prestaciones de salud a la población, en términos socialmente aceptables de seguridad, oportunidad y calidad.”

Con todo lo señalado el presente estudio de investigación es relevante que se realice ya que los funcionarios públicos del Hospital Regional del Cusco, contarán con un instrumento con el cual se podrá realizar el pronóstico de la demanda de los medicamentos con un error bajo lo cual contribuirá a disminuir el desabastecimiento de medicamentos y ayudar al restablecimiento de la salud de los ciudadanos ya que en la actualidad el pronóstico lo realizan utilizando el método del consumo histórico el cual tiene la desventaja de realizar pronósticos en crecimiento perpetuo y constante.

1.5 OBJETIVOS

1.5.1 Objetivo general

Analizar los modelos autorregresivo integrado de media móvil o la red neuronal recurrente para predecir adecuadamente la demanda de medicamentos en el Hospital Regional del Cusco.

1.5.2 Objetivos específicos

- Estimar los modelos autorregresivo integrado de media móvil que proporcionen el mejor ajuste para el pronóstico de la demanda de medicamentos.
- Estimar los modelos red neuronal recurrente que proporcionen el mejor ajuste para el pronóstico de la demanda de medicamentos.

1.6 HIPÓTESIS

1.6.1 Hipótesis general

El modelo de la red neuronal recurrente permite pronosticar adecuadamente la demanda de los medicamentos en el Hospital Regional del Cusco.

1.6.2 Hipótesis específicas

- Los modelos autorregresivo integrado de media móvil proporcionan un mejor ajuste para el pronóstico de la demanda de los medicamentos.
- Los modelos red neuronal recurrente proporcionan un mejor ajuste para el pronóstico de la demanda de los medicamentos.

CAPITULO II MARCO TEÓRICO

2.1 BASES TEÓRICAS

2.1.1 Series de tiempo

2.1.1.1 Definición

Conjunto de observaciones realizadas secuencialmente durante un tiempo determinado (Chatfield, 2003, pág. 1).

Sucesión de observaciones efectuadas en forma ordenada en el tiempo (Box et al., 2016, pág. 1),

Cumulo de eventos cuantitativos los cuales son registrados en forma ordenada en un tiempo determinado (Kirchgässner, Wolters y Hassler, 2013, pág.1).

Agrupación de observaciones X_t , hechas en un tiempo específico t . (Brockwell y Davis, 2016, pág.1),

Observaciones efectuadas en forma ordenada las cuales están anexadas a un momento de tiempo (González, 2009, pág.1).

En base a las definiciones antes indicadas se infiere que la trayectoria que sigue una determinada serie de tiempo es la realización de un proceso estocástico, porque una serie de tiempo delinea el desarrollo aleatorio de una variable en el tiempo, una representación matemática al respecto sería mediante la función siguiente $t = Y_t(\omega)$; donde $t = \{0,1,2, \dots \dots n\}$ representa a un momento del tiempo y $Y_t(\omega)$ para cada Y_t le corresponde un valor $\omega \in \Omega$.

2.1.1.2 Componentes

Tomando como referencia a Hanke y Wichen (2006, pág. 158) y Ramón y López (2016, pág. 14), se tiene los siguientes:

Tendencia (T_t), representa el crecimiento o descenso de un evento.

Cíclico (C_t), representa una serie de fluctuaciones o ciclos de más de un año de duración.

Estacional (E_t), representa las fluctuaciones estacionales se encuentran en datos trimestrales, mensuales o semanales.

Irregular (I_t), representa las fluctuaciones impredecibles aleatorias.

2.1.1.3 Esquema de integración

La variable (Y_t) en un tiempo específico t , se puede poner en función de las componentes antes mencionadas de la siguiente manera $Y_t = f(T_t, C_t, E_t, I_t)$, resultando tres esquemas de integración que se describen a continuación (Ramón y López, 2016, pág. 15).

a) Modelo Multiplicativo $Y_t = T_t \cdot C_t \cdot E_t \cdot I_t$

b) Modelo Aditivo $Y_t = T_t + C_t + E_t + I_t$

c) Modelo Mixto $Y_t = T_t \cdot C_t \cdot E_t + I_t$

2.1.2 Serie Temporal Univariada

Antes de entrar de lleno al tema, conviene mencionar que los modelos de series temporales pueden ser univariado, cuando el análisis se realiza a una sola serie de tiempo en relación a las observaciones pasadas y multivariado cuando el análisis se efectúa a varias series de tiempo (González, 2009, pág. 4, Gujarati y Porter, 2010, pág. 775).

Una serie temporal univariada viene a ser una colección de observaciones de la variable Y , realizadas en forma cronológica en un tiempo t , la cual se puede representar de la siguiente manera

$$Y_1, Y_2, Y_3, \dots Y_t \quad (2.1)$$

donde Y es la observación y el sub índice t de cada valor de Y representa el momento de la observación realizada; cuando dichas observaciones se efectúan a intervalos iguales de tiempo (mensual, trimestral, etc.), se denomina como serie temporal discreta (Arnau, 2001, pág. 24).

Considerando a González (2009, pág. 19) , se entiende que el comportamiento de las variables Y_t en (2.1) queda representada con la ecuación siguiente

$$Y_t = PS_t + a_t \quad t = 1, 2, \dots \quad (2.2)$$

donde PS_t es la parte sistemática (comportamiento regular de la variable), la cual se obtiene en función de la información del pasado de la serie

$$PS_t = f(Y_t, Y_{t-1}, Y_{t-2}, Y_{t-3}, \dots) \quad (2.3)$$

y a_t es la parte aleatoria (innovación).

Por otro lado la herramienta principal para realizar el análisis de una serie temporal es el coeficiente de autocorrelación, se tiene T observaciones $Y_1, Y_2, Y_3 \dots Y_T$, de la cual se puede originar $(T - 1)$ pares de observaciones $(Y_1, Y_2), \dots, (Y_{T-1}, Y_T)$, entonces ahora se tiene dos variables $Y_1, Y_2, Y_3 \dots Y_{T-1}$ y $Y_2, Y_3 \dots Y_T$; luego se puede establecer la correlación entre las variables Y_t y Y_{t+1} de la siguiente manera

$$r_{Y_t Y_{t+1}} = \frac{\sum_{t=1}^{T-1} (Y_t - \bar{Y}_{(1)}) (Y_{t+1} - \bar{Y}_{(2)})}{\sqrt{\sum_{t=1}^{T-1} (Y_t - \bar{Y}_{(1)})^2 \sum_{t=1}^{T-1} (Y_{t+1} - \bar{Y}_{(2)})^2}} \quad (2.4)$$

donde $\bar{Y}_{(1)}$ es la media muestral de las $T - 1$ primeras observaciones y $\bar{Y}_{(2)}$ es la media muestral de las $T - 1$ ultimas observaciones, la ecuación (2.4) se aproxima en forma general con la siguiente ecuación

$$r_k = \frac{\sum_{t=1}^{T-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^T (Y_t - \bar{Y})^2} \quad (2.5)$$

donde r_k es el coeficiente de autocorrelación de orden k el cual mide la correlación entre las observaciones separadas por k intervalos de tiempo (González, 2009, pág. 6-7).

2.1.3 Procesos estocásticos

2.1.3.1 Definición

El proceso estocástico (aleatorio), implica alguna variable respuesta, que toma valores variando aleatoriamente en el tiempo, dicho proceso estocástico puede ser un escalar o un vector (Lindsey, 2004, pág. 3).

Un proceso estocástico se describe como un fenómeno estadístico que se desarrolla en el tiempo de acuerdo con leyes probabilísticas (Chatfield, 2003, pág. 33).

Un proceso estocástico se define como una familia de variables aleatorias $\{X(t): -\infty \leq t < \infty\}$, proceso en tiempo continuo y $\{X_t: t = 0, \pm 1, \pm 2, \dots\}$, proceso en tiempo discreto; además todas las X 's están definidos en un espacio de probabilidad (Ω, F, P) (Madsen, 2008, pág. 97, Hang, 2010, pág. 16, Chatfield, 2003, pág. 33).

2.1.3.2 Propiedades

Se tiene las siguientes (Lindsey, 2004, pág. 3):

- a) La frecuencia o periodicidad de las observaciones se efectúan a lo largo del tiempo.
- b) El conjunto de todos los valores observables es denominado como espacio de estado.
- c) Las fuentes y formas de aleatoriedad presentes, incluida la naturaleza de la dependencia entre los valores de la serie se denominan como realizaciones de la variable aleatoria Y_t .
- d) La cantidad de copias del proceso disponible (solo una o varias), determinará cómo se puede obtener información adecuada para modelar.

2.1.3.3 Caracterización

Se logra mediante su función de distribución o por sus momentos, en este caso se recurre a los dos primeros momentos por resultar más práctico el uso de estos (Kirchgässner, Wolters y Hassler, 2013, pág.12).

2.1.3.4 Momentos

Al respecto González (2009, pág. 12) indica:

a) Primer momento

Corresponde a la media o valor esperado

$$E(Y_t) = \mu_t < \infty, \quad t = 0, \pm 1, \pm 2, \dots, \quad (2.6)$$

b) Segundo momento

Conciérne a la varianza y covarianza

$$V(Y_t) = E [Y_t - \mu_t]^2 = \sigma_t^2 < \infty, \quad t = 0, \pm 1, \pm 2, \dots, \quad (2.7)$$

$$\text{cov}(Y_t Y_s) = E[Y_t - \mu_t][Y_s - \mu_s] = \gamma_{t,s}, \quad \forall t, s (t \neq s) \quad (2.8)$$

2.1.4 Procesos estocásticos estacionarios

Con el fin de obtener una predicción coherente, no es posible hacer uso de cualquier proceso estocástico, por tal motivo se hace uso de los procesos estocásticos estacionarios los cuales se basan en el supuesto de que el proceso se encuentra en un estado particular de equilibrio estadístico (Box et al., 2016, pág. 24).

2.1.4.1 Estacionariedad estricta

Para que un proceso estocástico Y_t sea estrictamente estacionario debe cumplir que la distribución conjunta no debe ser afectada por el cambio de los k periodos en el tiempo (Madsen, 2008, pág. 99).

$$F[Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}] = F[Y_{t_1+k}, Y_{t_2+k}, \dots, Y_{t_n+k}] \quad (2.9)$$

2.1.4.2 Estacionariedad débil

Al respecto Box et. al. (2016, pág. 24), indica lo siguiente:

a) Estacionario en media

Cuando $n=1$ (n son observaciones), implica que la distribución de probabilidad de $p(Y_t)$ es la misma para todos los tiempos de t y puede escribirse como $p(Y)$, en consecuencia, el proceso estocástico presenta una media constante.

$$\mu = E[Y_t] = \int_{-\infty}^{\infty} Yp(Y)dY \quad (2.10)$$

b) Estacionario en varianza

Es cuando la dispersión es constante alrededor de la media.

$$\sigma_Y^2 = E[(Y_t - \mu)^2] = \int_{-\infty}^{\infty} (Y - \mu)^2 p(Y) dY \quad (2.11)$$

Asimismo, la media y la varianza de un proceso estocástico se puede estimar con la media muestral y la varianza muestral.

$$\bar{Y} = \frac{1}{N} \sum_{t=1}^N Y_t \quad (2.12)$$

$$\sigma_{\bar{Y}}^2 = \frac{1}{N} \sum_{t=1}^N (Y_t - \bar{Y})^2 \quad (2.13)$$

c) Autocovarianza y coeficiente de autocorrelación

El asumir la estacionariedad implica que la probabilidad de distribución $p(Y_{t_1}, Y_{t_2})$ es la misma para los tiempos t_1, t_2 los cuales están en un intervalo constante aparte. En consecuencia, los valores Y_t y Y_{t+k} separados por k intervalos de tiempo o por lag k , debe ser el mismo para todos los t dicha covarianza se denomina autocovarianza en el lag k y se define

$$\gamma_k = cov[Y_t, Y_{t+k}] = E[(Y_t - \mu)(Y_{t+k} - \mu)] \quad (2.14)$$

En forma similar la autocorrelación en el lag k es

$$\rho_k = \frac{E[(Y_t - \mu)(Y_{t+k} - \mu)]}{\sqrt{E[(Y_t - \mu)^2]E[(Y_{t+k} - \mu)^2]}} = \frac{E[(Y_t - \mu)(Y_{t+k} - \mu)]}{\sigma_Y^2} \quad (2.15)$$

Para un proceso estacionario la varianza $\sigma_Y^2 = \gamma_0$ es el mismo en el tiempo $t + k$ que en el tiempo t . Entonces la autocorrelación en el lag k es la correlación entre Y_t y Y_{t+k} , es decir

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (2.16)$$

2.1.5 Proceso ruido blanco

Un proceso estocástico Y_t se denomina proceso de ruido blanco si los componentes son independientes y para cualquier t las variables aleatorias Y_t tienen la misma distribución $F(Y)$, con el valor esperado $E[Y_t] = 0$. Lo cual implica que la media es cero, la varianza es constante y las covarianzas son cero (Florescu, 2015, pág. 298).

2.1.6 Modelos Autoregresivos integrados de de media móvil (ARIMA)

2.1.6.1 Modelo autorregresivo (AR(p))

Según lo planteado por Hanke y Wichern (2006, pág 386), un modelo autorregresivo de orden p (AR(p)), de manera general es representado de la siguiente forma

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + \varepsilon_t \quad (2.17)$$

donde

Y_t = variable respuesta en el tiempo t

$Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ = variables respuesta en los retrasos $t - 1, t - 2, \dots, t - p$

$\varphi_0, \varphi_1, \varphi_2, \dots, \varphi_p$ = coeficientes a estimar

ε_t = error en el tiempo t , además $\varepsilon_t \sim N(0, \sigma^2)$

Haciendo uso de un operador de desplazamiento hacia atrás B , un modelo Ar(p) es representado de la forma siguiente (Madsen, 2008, pág. 119).

$$(1 + \varphi_1 B + \dots + \varphi_p B^p) Y_t = \varepsilon_t \quad (2.18)$$

En cuanto a las propiedades Pérez (2007, pág. 50), Tsay (2013, pág. 59), Montgomery, Jennings y Kulachi (2008, pág. 247) y Guidolin y Pedio (2018, pág. 54) manifiestan las siguientes:

a) La media de una serie estacionaria es

$$E(Y_t) = \frac{\varphi_0}{1 - \varphi_1 - \dots - \varphi_p} \quad (2.19)$$

siempre que el denominador no sea cero, es decir $|\varphi_1 + \varphi_2 + \dots + \varphi_p| < 1$.

b) Autocovarianzas

$$\gamma_k \begin{cases} \varphi_1 Y_1 + \varphi_2 Y_2 + \dots + \varphi_p Y_p + \sigma^2, & k = 0 \\ \varphi_1 Y_{k-1} + \varphi_2 Y_{k-2} + \dots + \varphi_p Y_{k-p}, & k \geq 1 \end{cases} \quad (2.20)$$

c) Función de autocorrelación simple

$$\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} + \dots + \varphi_p \rho_{k-p}, \quad k \geq 1 \quad (2.21)$$

d) Función de autocorrelación parcial

$$\varphi_{kk} = \begin{cases} \rho_1, & k = 1 \\ \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}, & k = 2 \\ \vdots & \\ \varphi_p, & k = p \\ 0, & k \geq p + 1 \end{cases} \quad (2.22)$$

A continuación, determinaremos media, varianza, función de autocovarianza, función de autocorrelación simple y función de autocorrelación parcial; de los modelos AR(1) y AR(2) en base a las condiciones de estacionariedad.

2.1.6.1.1 Modelo AR(1)

Un modelo AR(1) queda expresado de la siguiente manera

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varepsilon_t$$

a) Media

$$E(Y_t) = E(\varphi_0 + \varphi_1 Y_{t-1} + \varepsilon_t)$$

$$E(Y_t) = \varphi_0 + \varphi_1 E(Y_{t-1}) + 0$$

$$\mu = \varphi_0 + \varphi_1 \mu$$

$$\mu - \varphi_1 \mu = \varphi_0$$

$$\mu(1 - \varphi_1) = \varphi_0$$

$$\mu = \frac{\varphi_0}{1 - \varphi_1}$$

de la ecuación anterior se deduce que $-1 < \varphi_1 < 1$ para que el modelo AR(1) sea estacionario.

b) Varianza

$$Var(Y_t) = (\varphi_1 Y_{t-1} + \varepsilon_t)(\varphi_1 Y_{t-1} + \varepsilon_t)$$

$$Var(Y_t) = \varphi_1^2 Y_{t-1}^2 + \varphi_1 Y_{t-1} \varepsilon_t + \varphi_1 Y_{t-1} \varepsilon_t + \varepsilon_t^2$$

$$Var(Y_t) = \varphi_1^2 Y_{t-1}^2 + 2\varphi_1 Y_{t-1} \varepsilon_t + \varepsilon_t^2$$

$$Var(Y_t) = \varphi_1^2 Y_{t-1}^2 + \varepsilon_t^2$$

$$Var(Y_t) = \varphi_1^2 Var(Y_t) + \sigma_\varepsilon^2$$

$$Var(Y_t) - \varphi_1^2 Var(Y_t) = \sigma_\varepsilon^2$$

$$Var(Y_t)(1 - \varphi_1^2) = \sigma_\varepsilon^2$$

$$Var(Y_t) = \frac{\sigma_\varepsilon^2}{(1 - \varphi_1^2)}$$

c) Función de autocovarianza

$$\gamma_0 = E(Y_t Y_t) = E(Y_t^2) = Var(Y_t)$$

$$\gamma_1 = E(Y_t Y_{t-1}) = E[Y_{t-1}(\varphi_0 + \varphi_1 Y_{t-1} + \varepsilon_t)] = E[\varphi_0 Y_{t-1} + \varphi_1 Y_{t-1} Y_{t-1} + Y_{t-1} \varepsilon_t] = \varphi_1 \gamma_0$$

$$\gamma_2 = E(Y_t Y_{t-2}) = E[Y_{t-2}(\varphi_0 + \varphi_1 Y_{t-1} + \varepsilon_t)] = E[\varphi_0 Y_{t-2} + \varphi_1 Y_{t-1} Y_{t-2} + Y_{t-1} \varepsilon_t] = \varphi_1 \gamma_1$$

$$\gamma_3 = E(Y_t Y_{t-3}) = E[Y_{t-3}(\varphi_0 + \varphi_1 Y_{t-1} + \varepsilon_t)] = E[\varphi_0 Y_{t-3} + \varphi_1 Y_{t-3} Y_{t-1} + Y_{t-3} \varepsilon_t] = \varphi_1 \gamma_2$$

d) Función de autocorrelación

$$\rho_0 = \frac{Cov(Y_t, Y_t)}{\sqrt{Var(Y_t) Var(Y_t)}} = \frac{Cov(Y_t, Y_t)}{\sqrt{Var(Y_t)^2}} = \frac{Var(Y_t)}{Var(Y_t)} = \frac{\gamma_0}{\gamma_0} = 1$$

$$\rho_1 = \frac{Cov(Y_t, Y_{t-1})}{\sqrt{Var(Y_t) Var(Y_{t-1})}} = \frac{Cov(Y_t, Y_{t-1})}{\sqrt{Var(Y_t)^2}} = \frac{Cov(Y_t, Y_{t-1})}{Var(Y_t)} = \frac{\gamma_1}{\gamma_0} = \frac{\varphi_1 \gamma_0}{\gamma_0} = \varphi_1$$

$$\rho_2 = \frac{Cov(Y_t, Y_{t-2})}{\sqrt{Var(Y_t) Var(Y_{t-2})}} = \frac{Cov(Y_t, Y_{t-2})}{\sqrt{Var(Y_t)^2}} = \frac{Cov(Y_t, Y_{t-2})}{Var(Y_t)} = \frac{\gamma_2}{\gamma_0} = \frac{\varphi_1 \gamma_1}{\gamma_0} = \frac{\varphi_1 \varphi_1 \gamma_0}{\gamma_0} = \varphi_1^2$$

$$\rho_3 = \frac{Cov(Y_t, Y_{t-3})}{\sqrt{Var(Y_t) Var(Y_{t-3})}} = \frac{Cov(Y_t, Y_{t-3})}{\sqrt{Var(Y_t)^2}} = \frac{Cov(Y_t, Y_{t-3})}{Var(Y_t)} = \frac{\gamma_3}{\gamma_0} = \frac{\varphi_1 \gamma_2}{\gamma_0} = \frac{\varphi_1 \varphi_1 \varphi_1 \gamma_0}{\gamma_0} = \varphi_1^3$$

d) Funcion de autocorrelación parcial

Haciendo uso de la regla de Cramer se tiene

$$Y_t = \varphi_{11} Y_{t-1} + \varepsilon_t$$

$$\varphi_{11} = \text{corr}(Y_t, Y_{t-1}) = \frac{|\rho_1|}{|1|} = \rho_1$$

2.1.6.1.2 Modelo AR(2)

Un modelo AR(2) queda expresado de la siguiente manera

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varepsilon_t$$

a) Media

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varepsilon_t$$

$$E(Y_t) = E(\varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varepsilon_t)$$

$$E(Y_t) = \varphi_0 + \varphi_1 E(Y_{t-1}) + \varphi_2 E(Y_{t-2}) + 0$$

$$\mu = \varphi_0 + \varphi_1 \mu + \varphi_2 \mu$$

$$\mu - \varphi_1 \mu - \varphi_2 \mu = \varphi_0$$

$$\mu(1 - \varphi_1 - \varphi_2) = \varphi_0$$

$$\mu = \frac{\varphi_0}{1 - \varphi_1 - \varphi_2}$$

de la anterior ecuación se deduce que

$$\varphi_2 + \varphi_1 < 1$$

$$\varphi_2 - \varphi_1 < 1$$

$$-1 < \varphi_2 < 1$$

b) Varianza

$$\begin{aligned}
 \text{Var}(Y_t) &= (\varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varepsilon_t)(\varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varepsilon_t) \\
 &= (\varphi_1^2 Y_{t-1}^2 + \varphi_1 \varphi_2 Y_{t-1} Y_{t-2} + \varphi_1 Y_{t-1} \varepsilon_t + \varphi_1 \varphi_2 Y_{t-1} Y_{t-2} + \varphi_2^2 Y_{t-2}^2 + \varphi_2 Y_{t-2} \varepsilon_t + \varphi_1 Y_{t-1} \varepsilon_t \\
 &\quad + \varphi_2 Y_{t-2} \varepsilon_t + \varepsilon_t^2) \\
 &= (\varphi_1^2 Y_{t-1}^2 + \varphi_1 \varphi_2 Y_{t-1} Y_{t-2} + \varphi_1 \varphi_2 Y_{t-1} Y_{t-2} + \varphi_2^2 Y_{t-2}^2 + \varepsilon_t^2) \\
 &= (\varphi_1^2 Y_{t-1}^2 + 2\varphi_1 \varphi_2 Y_{t-1} Y_{t-2} + \varphi_2^2 Y_{t-2}^2 + \varepsilon_t^2) \\
 \text{Var}(Y_t) &= \varphi_1^2 \text{Var}(Y_t) + 2\varphi_1 \varphi_2 \text{Cov}(Y_{t-1} Y_{t-2}) + \varphi_2^2 \text{Var}(Y_t) + \text{Var}(\varepsilon_t) \\
 \text{Var}(Y_t) - \varphi_1^2 \text{Var}(Y_t) - \varphi_2^2 \text{Var}(Y_t) &= 2\varphi_1 \varphi_2 \text{Cov}(Y_{t-1} Y_{t-2}) + \text{Var}(\varepsilon_t) \\
 \text{Var}(Y_t)(1 - \varphi_1^2 - \varphi_2^2) &= 2\varphi_1 \varphi_2 \text{Cov}(Y_{t-1} Y_{t-2}) + \text{Var}(\varepsilon_t) \\
 \text{Var}(Y_t) &= \frac{2\varphi_1 \varphi_2 \text{Cov}(Y_{t-1} Y_{t-2}) + \text{Var}(\varepsilon_t)}{(1 - \varphi_1^2 - \varphi_2^2)} \\
 \text{Var}(Y_t) &= \frac{2\varphi_1 \varphi_2 \gamma_1 + \text{Var}(\varepsilon_t)}{(1 - \varphi_1^2 - \varphi_2^2)}
 \end{aligned}$$

c) Función de autocovarianza

$$\begin{aligned}
 \gamma_0 &= E(Y_t Y_t) = E(Y_t^2) = \text{Var}(Y_t) \\
 \gamma_1 &= E(Y_t Y_{t-1}) = E[Y_{t-1}(\varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varepsilon_t)] \\
 &= E[\varphi_0 Y_{t-1} + \varphi_1 Y_{t-1} Y_{t-1} + \varphi_2 Y_{t-2} Y_{t-1} + Y_{t-1} \varepsilon_t] = \varphi_1 \gamma_0 + \varphi_2 \gamma_1 \\
 \gamma_2 &= E(Y_t Y_{t-2}) = E[Y_{t-2}(\varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varepsilon_t)] \\
 &= E[\varphi_0 Y_{t-2} + \varphi_1 Y_{t-1} Y_{t-2} + \varphi_2 Y_{t-2} Y_{t-2} + Y_{t-2} \varepsilon_t] = \varphi_1 \gamma_1 + \varphi_2 \gamma_0 \\
 \gamma_3 &= E(Y_t Y_{t-3}) = E[Y_{t-3}(\varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varepsilon_t)] \\
 &= E[\varphi_0 Y_{t-3} + \varphi_1 Y_{t-1} Y_{t-3} + \varphi_2 Y_{t-2} Y_{t-3} + Y_{t-3} \varepsilon_t] = \varphi_1 \gamma_2 + \varphi_2 \gamma_1
 \end{aligned}$$

d) Función de autocorrelación

$$\begin{aligned}
 \rho_0 &= \frac{\text{Cov}(Y_t, Y_t)}{\sqrt{\text{Var}(Y_t) \text{Var}(Y_t)}} = \frac{\text{Cov}(Y_t, Y_t)}{\sqrt{\text{Var}(Y_t)^2}} = \frac{\text{Var}(Y_t)}{\text{Var}(Y_t)} = \frac{\gamma_0}{\gamma_0} = 1 \\
 \rho_1 &= \frac{\text{Cov}(Y_t, Y_{t-1})}{\sqrt{\text{Var}(Y_t) \text{Var}(Y_{t-1})}} = \frac{\text{Cov}(Y_t, Y_{t-1})}{\sqrt{\text{Var}(Y_t)^2}} = \frac{\text{Cov}(Y_t, Y_{t-1})}{\text{Var}(Y_t)} = \frac{\varphi_1 \gamma_0 + \varphi_2 \gamma_1}{\gamma_0} = \varphi_1 + \varphi_2 \rho_1 \\
 \rho_2 &= \frac{\text{Cov}(Y_t, Y_{t-2})}{\sqrt{\text{Var}(Y_t) \text{Var}(Y_{t-2})}} = \frac{\text{Cov}(Y_t, Y_{t-2})}{\sqrt{\text{Var}(Y_t)^2}} = \frac{\text{Cov}(Y_t, Y_{t-2})}{\text{Var}(Y_t)} = \frac{\gamma_2}{\gamma_0} = \frac{\varphi_1 \gamma_1 + \varphi_2 \gamma_0}{\gamma_0} \\
 &= \varphi_1 \rho_1 + \varphi_2
 \end{aligned}$$

$$\begin{aligned}\rho_3 &= \frac{Cov(Y_t, Y_{t-3})}{\sqrt{Var(Y_t) Var(Y_{t-3})}} = \frac{Cov(Y_t, Y_{t-3})}{\sqrt{Var(Y_t)^2}} = \frac{Cov(Y_t, Y_{t-3})}{Var(Y_t)} = \frac{\gamma_3}{\gamma_0} = \frac{\varphi_1\gamma_2 + \varphi_2\gamma_1}{\gamma_0} \\ &= \varphi_1\rho_2 + \varphi_2\rho_1\end{aligned}$$

d) Función de autocorrelación parcial

Utilizando la regla de Cramer se tiene

$$\begin{aligned}Y_t &= \varphi_{21}Y_{t-1} + \varphi_{22}Y_{t-2} + \varepsilon_t \\ \varphi_{22} &= \text{corr}(Y_t, Y_{t-2}) = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2\rho_1^2}{1 - \rho_1^2}\end{aligned}$$

2.1.6.2 Modelo de media móvil (MA(q))

Tomando como referencia a Hanke y Wichern (2006, pág 388), de forma general el modelo de media móvil de orden q (MA(q)) se puede representar de la siguiente manera

$$Y_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q} \quad (2.23)$$

donde

Y_t = variable respuesta en el tiempo t

μ = valor promedio constante

$\theta_1, \theta_2, \dots, \theta_q$ = coeficientes a estimar

ε_t = error en el tiempo t , además $\varepsilon_t \sim N(0, \sigma^2)$

$\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$ = errores en periodos anteriores al tiempo t

Otra manera de representar un modelo MA(q), es haciendo uso de un operador de desplazamiento hacia atrás B de la manera siguiente (Madsen, 2008, pág. 117).

$$Y_t = (1 + \theta_1B + \dots + \theta_qB^q)\varepsilon_t = \theta(B)\varepsilon_t \quad (2.24)$$

Sobre las propiedades Brooks (2008, pág. 212) y Mills (2019, pág. 42), señalan las siguientes:

$$a) E(Y_t) = \mu \quad (2.25)$$

$$b) var(Y_t) = \gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \theta_3^2) \sigma_a^2 \quad (2.26)$$

$$c) covarianza \gamma_s \begin{cases} \theta_s + \theta_{s+1} \theta_1 + \theta_{s+2} \theta_2 + \dots + \theta_q \theta_{q-s} \sigma_a^2, & s = 1, 2, \dots, q \\ 0, & s > q \end{cases} \quad (2.27)$$

$$d) autocorrelación \rho_s \begin{cases} \frac{-\theta_k + \theta_1 + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2}, & s = 1, 2, \dots, q \\ 0, & s > q \end{cases} \quad (2.28)$$

Seguidamente, se determina media, varianza, función de autocovarianza, función de autocorrelación simple y función de autocorrelación parcial; de los modelos MA(1) y MA(2) aplicando las condiciones de estacionariedad.

2.1.6.2.1 Modelo MA(1)

Un modelo MA(1) queda expresado de la siguiente manera

$$Y_t = \varepsilon_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

a) Media

$$\begin{aligned} Y_t &= \varepsilon_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} \\ E(Y_t) &= E(\varepsilon_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1}) \\ E(Y_t) &= E(\varepsilon_0) + E(\varepsilon_t) - E(\theta_1 \varepsilon_{t-1}) \\ E(Y_t) &= \varepsilon_0 \end{aligned}$$

b) Varianza

$$\begin{aligned} Var(Y_t) &= (\varepsilon_t - \theta_1 \varepsilon_{t-1})(\varepsilon_t - \theta_1 \varepsilon_{t-1}) \\ Var(Y_t) &= \varepsilon_t^2 - \theta_1 \varepsilon_{t-1} \varepsilon_t - \theta_1 \varepsilon_{t-1} \varepsilon_t + \theta_1^2 \varepsilon_{t-1}^2 \\ Var(Y_t) &= \varepsilon_t^2 - 2\theta_1 \varepsilon_{t-1} \varepsilon_t + \theta_1^2 \varepsilon_{t-1}^2 \\ Var(Y_t) &= Var(\varepsilon_t) + \theta_1^2 Var(\varepsilon_{t-1}) \\ Var(Y_t) &= Var(\varepsilon_t)(1 + \theta_1^2) \end{aligned}$$

c) Función de autocovarianza

$$\gamma_0 = E(Y_t Y_t) = E(Y_t^2) = \text{Var}(Y_t)$$

$$\begin{aligned} \gamma_1 &= E(Y_t Y_{t-1}) = E[(\varepsilon_t - \theta_1 \varepsilon_{t-1})(\varepsilon_{t-1} - \theta_1 \varepsilon_{t-2})] \\ &= E[\varepsilon_t \varepsilon_{t-1} - \theta_1 \varepsilon_{t-2} \varepsilon_t - \theta_1 \varepsilon_{t-1}^2 + \theta_1^2 \varepsilon_{t-1} \varepsilon_{t-2}] = -\theta_1 \text{Var}(\varepsilon_t) \end{aligned}$$

$$\begin{aligned} \gamma_2 &= E(Y_t Y_{t-2}) = E[(\varepsilon_t - \theta_1 \varepsilon_{t-1})(\varepsilon_{t-2} - \theta_1 \varepsilon_{t-3})] \\ &= E[\varepsilon_t \varepsilon_{t-2} - \theta_1 \varepsilon_{t-3} \varepsilon_t - \theta_1 \varepsilon_{t-1} \varepsilon_{t-2} + \theta_1^2 \varepsilon_{t-1} \varepsilon_{t-3}] = 0 \end{aligned}$$

$$\begin{aligned} \gamma_3 &= E(Y_t Y_{t-3}) = E[(\varepsilon_t - \theta_1 \varepsilon_{t-1})(\varepsilon_{t-3} - \theta_1 \varepsilon_{t-4})] \\ &= E[\varepsilon_t \varepsilon_{t-3} - \theta_1 \varepsilon_{t-4} \varepsilon_t - \theta_1 \varepsilon_{t-1} \varepsilon_{t-3} + \theta_1^2 \varepsilon_{t-1} \varepsilon_{t-4}] = 0 \end{aligned}$$

d) Función de autocorrelación simple

$$\rho_0 = \frac{\text{Cov}(Y_t, Y_t)}{\sqrt{\text{Var}(Y_t) \text{Var}(Y_t)}} = \frac{\text{Cov}(Y_t, Y_t)}{\sqrt{\text{Var}(Y_t)^2}} = \frac{\text{Var}(Y_t)}{\text{Var}(Y_t)} = \frac{\gamma_0}{\gamma_0} = 1$$

$$\begin{aligned} \rho_1 &= \frac{\text{Cov}(Y_t, Y_{t-1})}{\sqrt{\text{Var}(Y_t) \text{Var}(Y_{t-1})}} = \frac{\text{Cov}(Y_t, Y_{t-1})}{\sqrt{\text{Var}(Y_t)^2}} = \frac{\text{Cov}(Y_t, Y_{t-1})}{\text{Var}(Y_t)} = \frac{\gamma_1}{\gamma_0} = \frac{-\theta_1 \text{Var}(\varepsilon_t)}{\text{Var}(\varepsilon_t)(1 - \theta_1^2)} \\ &= \frac{-\theta_1}{1 - \theta_1^2} \end{aligned}$$

$$\rho_2 = \frac{\text{Cov}(Y_t, Y_{t-2})}{\sqrt{\text{Var}(Y_t) \text{Var}(Y_{t-2})}} = \frac{\text{Cov}(Y_t, Y_{t-2})}{\sqrt{\text{Var}(Y_t)^2}} = \frac{\text{Cov}(Y_t, Y_{t-2})}{\text{Var}(Y_t)} = \frac{\gamma_2}{\gamma_0} = \frac{0}{\text{Var}(\varepsilon_t)(1 - \theta_1^2)} = 0$$

$$\rho_3 = \frac{\text{Cov}(Y_t, Y_{t-3})}{\sqrt{\text{Var}(Y_t) \text{Var}(Y_{t-3})}} = \frac{\text{Cov}(Y_t, Y_{t-3})}{\sqrt{\text{Var}(Y_t)^2}} = \frac{\text{Cov}(Y_t, Y_{t-3})}{\text{Var}(Y_t)} = \frac{\gamma_3}{\gamma_0} = \frac{0}{\text{Var}(\varepsilon_t)(1 - \theta_1^2)} = 0$$

e) Función de autocorrelación parcial

Debido a que un modelo MA(1) es invertible se puede obtener el modelo equivalente AR(1) y luego utilizando la regla de Cramer.

$$\begin{aligned} Y_t &= \varphi_{11} Y_{t-1} + \varepsilon_t \\ \varphi_{11} &= \text{corr}(Y_t, Y_{t-1}) = \frac{|\rho_1|}{|1|} = \rho_1 = \frac{-\theta_1}{1 - \theta_1^2} \end{aligned}$$

2.1.6.2.2 Modelo MA(2)

Un modelo MA(2) queda expresado de la siguiente manera

$$Y_t = \varepsilon_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

a) Media

$$\begin{aligned} Y_t &= \varepsilon_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \\ E(Y_t) &= E(\varepsilon_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}) \\ E(Y_t) &= E(\varepsilon_0) + E(\varepsilon_t) - E(\theta_1 \varepsilon_{t-1}) - E(\theta_2 \varepsilon_{t-2}) \\ E(Y_t) &= \varepsilon_0 \end{aligned}$$

b) Varianza

$$\begin{aligned} \text{Var}(Y_t) &= (\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2})(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}) \\ \text{Var}(Y_t) &= \varepsilon_t^2 - \theta_1 \varepsilon_{t-1} \varepsilon_t - \theta_2 \varepsilon_{t-2} \varepsilon_t - \theta_1 \varepsilon_{t-1} \varepsilon_t + \theta_1^2 \varepsilon_{t-1}^2 + \theta_1 \varepsilon_{t-1} \theta_2 \varepsilon_{t-2} - \theta_2 \varepsilon_{t-2} \varepsilon_t \\ &\quad + \theta_2 \varepsilon_{t-2} \theta_1 \varepsilon_{t-1} + \theta_2^2 \varepsilon_{t-2}^2 \\ \text{Var}(Y_t) &= \varepsilon_t^2 - 2 \theta_1 \varepsilon_{t-1} \varepsilon_t - 2 \theta_2 \varepsilon_{t-2} \varepsilon_t + \theta_1^2 \varepsilon_{t-1}^2 + 2 \theta_1 \theta_2 \varepsilon_{t-1} \varepsilon_{t-2} + \theta_2^2 \varepsilon_{t-2}^2 \\ \text{Var}(Y_t) &= \varepsilon_t^2 + \theta_1^2 \varepsilon_{t-1}^2 + \theta_2^2 \varepsilon_{t-2}^2 \\ \text{Var}(Y_t) &= \text{Var}(\varepsilon_t) - \theta_1^2 \text{Var}(\varepsilon_{t-1}) - \theta_2^2 \text{Var}(\varepsilon_{t-2}) \\ \text{Var}(Y_t) &= \text{Var}(\varepsilon_t)(1 + \theta_1^2 + \theta_2^2) \end{aligned}$$

c) Función de autocovarianza

$$\begin{aligned} \gamma_0 &= E(Y_t Y_t) = E(Y_t^2) = \text{Var}(Y_t) \\ \gamma_1 &= E(Y_t Y_{t-1}) = E[(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2})(\varepsilon_{t-1} - \theta_1 \varepsilon_{t-2} - \theta_2 \varepsilon_{t-3})] \\ &= (-\theta_1 + \theta_1 \theta_2) \text{Var}(\varepsilon_t) \\ \gamma_2 &= E(Y_t Y_{t-2}) = E[(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2})(\varepsilon_{t-2} - \theta_1 \varepsilon_{t-3} - \theta_2 \varepsilon_{t-4})] = -\theta_2 \text{Var}(\varepsilon_t) \\ \gamma_3 &= E(Y_t Y_{t-3}) = E[(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2})(\varepsilon_{t-3} - \theta_1 \varepsilon_{t-4} - \theta_2 \varepsilon_{t-5})] = 0 \end{aligned}$$

d) Función de autocorrelación simple

$$\rho_0 = \frac{\text{Cov}(Y_t, Y_t)}{\sqrt{\text{Var}(Y_t) \text{Var}(Y_t)}} = \frac{\text{Cov}(Y_t, Y_t)}{\sqrt{\text{Var}(Y_t)^2}} = \frac{\text{Var}(Y_t)}{\text{Var}(Y_t)} = \frac{\gamma_0}{\gamma_0} = 1$$

$$\begin{aligned}
\rho_1 &= \frac{Cov(Y_t, Y_{t-1})}{\sqrt{Var(Y_t) Var(Y_{t-1})}} = \frac{Cov(Y_t, Y_{t-1})}{\sqrt{Var(Y_t)^2}} = \frac{Cov(Y_t, Y_{t-1})}{Var(Y_t)} = \frac{\gamma_1}{\gamma_0} = \frac{(-\theta_1 + \theta_1\theta_2)Var(\varepsilon_t)}{Var(\varepsilon_t)(1 - \theta_1^2 - \theta_2^2)} \\
&= \frac{(-\theta_1 + \theta_1\theta_2)}{(1 - \theta_1^2 - \theta_2^2)} \\
\rho_2 &= \frac{Cov(Y_t, Y_{t-2})}{\sqrt{Var(Y_t) Var(Y_{t-2})}} = \frac{Cov(Y_t, Y_{t-2})}{\sqrt{Var(Y_t)^2}} = \frac{Cov(Y_t, Y_{t-2})}{Var(Y_t)} = \frac{\gamma_2}{\gamma_0} = \frac{-\theta_2 Var(\varepsilon_t)}{Var(\varepsilon_t)(1 - \theta_1^2 - \theta_2^2)} \\
&= \frac{-\theta_2}{(1 - \theta_1^2 - \theta_2^2)} \\
\rho_3 &= \frac{Cov(Y_t, Y_{t-3})}{\sqrt{Var(Y_t) Var(Y_{t-3})}} = \frac{Cov(Y_t, Y_{t-3})}{\sqrt{Var(Y_t)^2}} = \frac{Cov(Y_t, Y_{t-3})}{Var(Y_t)} = \frac{\gamma_3}{\gamma_0} = \frac{0}{Var(\varepsilon_t)(1 - \theta_1^2 - \theta_2^2)} \\
&= 0
\end{aligned}$$

e) Función de autocorrelación parcial

Debido a que un modelo MA(2) es invertible se puede obtener el modelo equivalente AR(2) y después utilizando la regla de Cramer.

$$Y_t = \varphi_{21}Y_{t-1} + \varphi_{22}Y_{t-2} + \varepsilon_t$$

$$\varphi_{22} = corr(Y_t, Y_{t-2}) = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2\rho_1^2}{1 - \rho_1^2}$$

2.1.6.3 Modelo autorregresivo de media móvil (ARMA(p,q))

Según Hanke y Wichern (2006, pág. 389), un modelo autorregresivo de media móvil de orden p y q (ARMA (p,q)), se puede representar de la siguiente forma

$$Y_t = \varphi_0 + \varphi_1Y_{t-1} + \varphi_2Y_{t-2} + \dots + \varphi_pY_{t-p} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q} \quad (2.29)$$

Sin embargo, una representación más compacta de un modelo ARMA(p,q) y en base al uso de un operador de retraso sería (Guidolin y Pedio, 2018, pág.58)

$$Y_t = \varphi_0 + \sum_{i=1}^p \varphi_i Y_{t-i} + \sum_{i=0}^q \theta_i \varepsilon_{t-i}$$

$$\left(1 - \sum_{i=1}^p \varphi_i B^i\right) Y_t = \varphi_0 + \left(1 + \sum_{i=0}^q \theta_i B^i\right) \varepsilon_t$$

$$Y_t = \frac{\varphi_0 + (1 + \sum_{i=0}^q \theta_i B^i) \varepsilon_t}{(1 - \sum_{i=1}^p \varphi_i B^i)} \quad (2.30)$$

Sobre las propiedades de dicho modelo Ramón y López (2016, pág. 79) indican los siguientes:

- Los modelos ARMA (p,q) serán estacionarios e invertibles cuando las soluciones de los polinomios $\varphi(L) = 0$ y $\theta(L) = 0$ sean superiores a uno.
- La función de autocorrelación no se anula nunca.
- La función de autocorrelación parcial no se anula, debido a su parte MA(q).

En seguida, se determina media, varianza, función de autocovarianza, función de autocorrelación simple y función de autocorrelación parcial; del modelo ARMA(1,1) haciendo uso de las condiciones de estacionariedad.

2.1.6.3.1 Modelo ARMA(1,1)

Un modelo ARMA(1,1) queda representado

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

a) Media

$$\begin{aligned} Y_t &= \varphi_0 + \varphi_1 Y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t \\ E(Y_t) &= E(\varphi_0 + \varphi_1 Y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t) \\ E(Y_t) &= \varphi_0 + \varphi_1 E(Y_{t-1}) + \theta_1 (\varepsilon_{t-1}) + E(\varepsilon_t) \\ \mu &= \varphi_0 + \varphi_1 \mu + 0 + 0 \\ \mu(1 - \varphi_1) &= \varphi_0 \\ \mu &= \frac{\varphi_0}{(1 - \varphi_1)} \end{aligned}$$

b) Varianza

$$\begin{aligned}
 Var(Y_t) &= (\varphi_1 Y_{t-1} - \theta_1 \varepsilon_{t-1} + \varepsilon_t)(\varphi_1 Y_{t-1} - \theta_1 \varepsilon_{t-1} + \varepsilon_t) \\
 Var(Y_t) &= (\varphi_1^2 Y_{t-1}^2 - \varphi_1 Y_{t-1} \theta_1 \varepsilon_{t-1} + \varphi_1 Y_{t-1} \varepsilon_t - \varphi_1 Y_{t-1} \theta_1 \varepsilon_{t-1} + \theta_1^2 \varepsilon_{t-1}^2 - \theta_1 \varepsilon_{t-1} \varepsilon_t \\
 &\quad + \varphi_1 Y_{t-1} \varepsilon_t - \theta_1 \varepsilon_{t-1} \varepsilon_t + \varepsilon_t^2) \\
 Var(Y_t) &= (\varphi_1^2 Y_{t-1}^2 - 2 \varphi_1 Y_{t-1} \theta_1 \varepsilon_{t-1} + 2 \varphi_1 Y_{t-1} \varepsilon_t + \theta_1^2 \varepsilon_{t-1}^2 - 2 \theta_1 \varepsilon_{t-1} \varepsilon_t + \varepsilon_t^2) \\
 Var(Y_t) &= (\varphi_1^2 Y_{t-1}^2 - 2 \varphi_1 \theta_1 Y_{t-1} \varepsilon_{t-1} + \theta_1^2 \varepsilon_{t-1}^2 + \varepsilon_t^2) \\
 Var(Y_t) &= \varphi_1^2 Var(Y_{t-1}) - 2 \varphi_1 \theta_1 Var(\varepsilon_t) + \theta_1^2 Var(\varepsilon_{t-1}) + Var(\varepsilon_t) \\
 Var(Y_t) - \varphi_1^2 Var(Y_{t-1}) &= -2 \varphi_1 \theta_1 Var(\varepsilon_t) + \theta_1^2 Var(\varepsilon_{t-1}) + Var(\varepsilon_t) \\
 Var(Y_t)(1 - \varphi_1^2) &= Var(\varepsilon_t) (1 + \theta_1^2 - 2 \varphi_1 \theta_1) \\
 Var(Y_t) &= \frac{Var(\varepsilon_t) (1 + \theta_1^2 - 2 \varphi_1 \theta_1)}{(1 - \varphi_1^2)}
 \end{aligned}$$

c) Función de autocovarianza

$$\begin{aligned}
 Y_t Y_t &= Y_t(\varphi_1 Y_{t-1} - \theta_1 \varepsilon_{t-1} + \varepsilon_t) \\
 Y_t Y_t &= \varphi_1 Y_{t-1} Y_t - \theta_1 \varepsilon_{t-1} Y_t + \varepsilon_t Y_t \\
 E(Y_t Y_t) &= \varphi_1 E(Y_{t-1} Y_t) - \theta_1 E(\varepsilon_{t-1} Y_t) + E(\varepsilon_t Y_t) \\
 E(Y_t Y_{t-1}) &= \varphi_1 E(Y_{t-1} Y_{t-1}) - \theta_1 E(\varepsilon_{t-1} Y_{t-1}) + E(\varepsilon_t Y_{t-1}) \\
 E(Y_t Y_{t-2}) &= \varphi_1 E(Y_{t-1} Y_{t-2}) - \theta_1 E(\varepsilon_{t-1} Y_{t-2}) + E(\varepsilon_t Y_{t-2}) \\
 E(Y_t Y_{t-s}) &= \varphi_1 E(Y_{t-1} Y_{t-s}) - \theta_1 E(\varepsilon_{t-1} Y_{t-s}) + E(\varepsilon_t Y_{t-s})
 \end{aligned}$$

resolviendo las ecuaciones anteriores se tiene

$$\begin{aligned}
 \gamma_0 &= \varphi_1 \gamma_1 + \theta_1 (\varphi_1 + \theta_1) Var(\varepsilon_t) + Var(\varepsilon_t) \\
 \gamma_1 &= \varphi_1 \gamma_0 + \theta_1 Var(\varepsilon_t) \\
 \gamma_2 &= \varphi_1 \gamma_1 \\
 &\quad \vdots \\
 \gamma_s &= \varphi_1 \gamma_{s-1} \\
 \gamma_0 &= \varphi_1 (\varphi_1 \gamma_0 + \theta_1 Var(\varepsilon_t)) + \theta_1 (\varphi_1 + \theta_1) Var(\varepsilon_t) + Var(\varepsilon_t) \\
 \gamma_0 &= \frac{(1 + \theta_1^2 + 2\theta_1 \varphi_1) Var(\varepsilon_t)}{1 - \varphi_1^2} \\
 \gamma_1 &= \frac{(1 + \gamma_1 \varphi_1) (\gamma_1 + \varphi_1) Var(\varepsilon_t)}{1 - \theta_1^2}
 \end{aligned}$$

$$\gamma_k = \varphi_1 \gamma_{k-1}, \quad k \geq 2$$

d) Función de autocorrelación simple

$$\rho_0 = \frac{\gamma_0}{\gamma_0} = 1$$

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(1 + \theta_1 \varphi_1)(\theta_1 + \varphi_1)}{(1 + \theta_1^2 + 2\theta_1 \varphi_1)}$$

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \varphi_1 \rho_{k-1}, \quad k \geq 2$$

e) Función de autocorrelación parcial

Haciendo uso de la regla de cramer se tiene

$$\varphi_{11} = \text{corr}(Y_t, Y_{t-1}) = \frac{|\rho_1|}{|1|} = \rho_1$$

2.1.6.4 Modelo autorregresivos integrado de media móvil (ARIMA(p,d,q))

En el mundo real las series de tiempo no son estacionarias es decir pueden presentar tendencia en media y varianza (Garcia, Rio y Coto, 2000, pág. 70).

2.1.6.4.1 Transformación Box-Cox

Box y Cox en 1964 plantearon transformaciones de potencia dentro de estas se tiene la logarítmica la cual se utiliza para estabilizar la varianza (Cajiao, Melo y Parra, 2014, pág. 40).

Se puede considerar a la transformación Box-Cox como un conjunto de posibles cambios en los valores de una serie de tiempo, claro está sin alterar lo característico de la serie lo cual se logra realizando un cambio de escala y compensación todo esto con el fin de mitigar la inestabilidad de la varianza (Aguirre, 1994, pág. 193).

Tomando en cuenta a Hipel y McLeod (1994, pág. 122), Pérez (2006, pág. 172) y Aguirre (1994, pág. 193) se tiene la serie de tiempo X_t y su correspondiente transformación W_t que se relacionan de la manera siguiente

$$W_t = \frac{(X_t + \lambda_2)^{\lambda_1} - 1}{\lambda_1 g^{\lambda_1 - 1}} \quad \text{si } \lambda_1 \neq 0 \quad (2.31)$$

$$W_t = g \ln(X_t + \lambda_2) \quad \text{si } \lambda_1 = 0 \quad (2.32)$$

donde

W_t es la serie transformada

Z_t es la serie de tiempo original

g es la media geométrica $X_t + \lambda_2$, es decir el factor de escala utilizado en la transformación

λ_1 es el parámetro de la transformación, si $\lambda_1 = 0$ le corresponde la transformación (2.32) y cuando $\lambda_1 > 0$ le corresponde la transformación (2.31)

λ_2 es una constante, si los datos de la serie son positivos $\lambda_2 = 0$ y en el caso que en la serie se presente valores negativos o el cero λ_2 es igual al valor absoluto del valor más negativo

Considerando a Perez (2006, pág. 172) se puede inferir que la transformación Box-Cox es un conjunto de transformaciones que dependen del parámetro λ_1 y que al asignarle un valor determinado se tendrá una transformación específica es decir

$$\lambda_1 \begin{cases} -1 & \text{transformación recíproca} \\ -1/2 & \text{recíproca de la raíz cuadrada} \\ 0 & \text{logarítmica} \\ 1/2 & \text{raíz cuadrada} \\ 1 & \text{identidad} \end{cases}$$

2.1.6.4.2 Diferenciación

Cuando una serie temporal es no estacionaria en media, se le debe de separar la tendencia valiéndose de las diferenciaciones en la serie con el fin de lograr un nivel constante (Arnau, 2001, pág. 49).

En muchos casos, una sola diferenciación es suficiente para conseguir la estacionariedad en media y en varianza (Ramón y López, 2016, pág. 88).

De acuerdo a Pérez (2006, pág. 171) si se tiene una serie temporal no estacionaria en media Y_t , la cual presenta una tendencia lineal, se debe aplicar una diferenciación $Z_t = Y_t - Y_{t-1}$ y de esta manera la serie se convierte en estacionaria es decir

$$Z_t = Y_t - Y_{t-1} \quad \text{primera diferencia o integrada de primer orden } I(1).$$

$$W_t = Z_t - Z_{t-1} \quad \text{segunda diferencia o integrada de segundo orden } I(2).$$

2.1.6.4.3 Modelamiento ARIMA

Se define la serie Y_t denominada modelo autorregresivo integrado de media móvil ARIMA (p,d,q), donde p y q indican el orden y d representa la diferenciación, dicho modelo se representa haciendo uso de los operadores de desplazamiento hacia atrás de la siguiente manera

$$\varphi(B)\nabla^d Y_t = \theta(B)\varepsilon_t, (d \in \mathbb{N}) \quad (2.33)$$

donde

ε_t es el ruido blanco

$\varphi(Z^{-1})$ es el polinomio de orden p

$\theta(Z^{-1})$ es el polinomio de orden q

de ambos polinomios sus raíces se encuentran dentro del círculo unitario (Madsen, 2008, pág. 130).

No obstante, se puede plantear el modelo ARIMA de otra manera tomando como guía a Hyndman y Athanasopoulos (2018, pág. 234) y Arnau (2001, pág. 27), quedaría expresado de la forma siguiente

$$\underbrace{\left(1 - \sum_{i=1}^p \varphi_i B^i\right)}_{\text{AR}(p)} \underbrace{\left(\prod_{i=1}^d (1 - B)^d\right)}_d Y_t = \underbrace{\left(1 + \sum_{i=1}^q \theta_i B^i\right)}_{\text{MA}(q)} \varepsilon_t \quad (2.34)$$

donde:

B : Operador retroactivo

Y_t : Serie de tiempo a encontrar

p : Orden autorregresivo (no estacional)

d : Indica el número de veces a diferenciar (no estacional)

q : Orden de promedio móvil (no estacional)

φ : Parámetro autorregresivo (no estacional)

θ : Parámetro de promedio móvil (no estacional)

ε_t : Ruido blanco

2.1.6.5 Modelo autorregresivos integrado de media móvil estacional (ARIMA(p,d,q) (P,D,Q)_s)

2.1.6.5.1 Series estacionales

Las series estacionales son aquellas donde se encuentran un comportamiento cíclico el cual se produce con un período inferior o igual a un año, lo cual puede ser detectado en los diferentes años, en consecuencia, el objetivo es desestacionalizar la serie, con el fin de obtener la verdadera distribución de la serie; para realizar una diferenciación estacionalmente se debe analizar la función de auto correlación simple (FAC) considerando sólo los rezagos donde se presenta los períodos estacionales m y sus múltiplos, es decir si los coeficientes de la FAC no decaen rápidamente en los rezagos múltiplos del período estacional m se debe diferenciar estacionalmente (Pérez , 2006, pág. 172)

Para un buen entendimiento y considerando a Pérez (2006, pág. 173) se plantea las series de tiempo Y_t , y Y_{ik}

$$Y_t \text{ donde } t (t = 1,2, \dots, T) \text{ es el tiempo}$$
$$Y_{ik} \text{ donde } \begin{cases} i (i = 1,2, \dots, N) \text{ es el año} \\ k (k = 1,2, \dots, m) \text{ la estación del año} \end{cases}$$
$$T = Nm$$

es decir, si la estación es el año $m = 12$ y si la estación es el trimestre $m = 4$, en consecuencia $T = Nm$; por ejemplo, se pretende desestacionalizar la serie Y_t

donde la estacionalidad se presenta al año, se debe aplicar una diferenciación de orden $m = 12$, representado por z_t

$$z_t = Y_t - Y_{t-12} \quad (2.35)$$

Tomando como guía a Tsay (2013, pág. 100), una diferenciación estacional de una serie estacionaria, utilizando los operadores de desplazamiento hacia atrás se expresa

$$\Delta_s Y_t = Y_t - Y_{t-s} = (1 - B^s)Y_t \quad (2.36)$$

donde

Δ_s es el operador diferencia estacional

s es el periodo estacional

Y_t es la variable a determinar

2.1.6.5.2 Modelamiento ARIMA(p,d,q) (P,D,Q)_s

Con este tipo de modelo se busca tener una herramienta con la cual se pueda abarcar los nexos entre periodos (años) y los nexos dentro de los periodos (intra anuales) y la relación entre ambas (González, 2009, pág. 109).

Madsen (2007, pág. 132) define a los modelos ARIMA estacionales multiplicativos como el proceso $\{Y_t\}$ denominado modelo multiplicativo estacional (p,d,q) X (P,D,Q)_s si

$$\phi(B)\Phi(B^s)\nabla^d\nabla_s^D Y_t = \theta(B)\theta(B^s)\varepsilon_t \quad (2.37)$$

donde $\{\varepsilon_t\}$ es ruido blanco, ϕ y θ son polinomios de orden p y q respectivamente, Φ y Θ son polinomios de orden P y Q, las raíces de los polinomios se encuentran dentro del círculo unitario.

Shumway y Stoffer (2011, pág. 154,157) describen ∇_s^D , $\Phi(B^s)$ y $\Theta(B^s)$ como los operadores estacionales diferencia, autorregresivo y promedio móvil respectivamente como se muestra a continuación

$$\nabla_s^D = (1 - B^S)^D$$

$$\Phi(B^S) = 1 - \Phi_1 B^S - \dots - \Phi_p B^{pS}$$

$$\theta(B^S) = 1 + \theta_1 B^S + \dots + \theta_Q B^{QS}$$

Tomando como guía a Hyndman y Athanasopoulos (2018, pág. 252), Shumway y Stoffer (2011, pág. 154,157) y Arnau (2001, pág. 27); se plantea el modelo ARIMA(p,d,q) (P,D,Q)_s de la siguiente forma

$$\underbrace{\left(1 - \sum_{i=1}^p \varphi_i B^i\right)}_{\text{AR}(p)} \underbrace{\left(1 - \sum_{i=1}^P \Phi_i B^{iS}\right)}_{\text{AR}(P)} \underbrace{\left(\prod_{i=1}^d (1 - B)^d\right)}_d \underbrace{\left(\prod_{i=1}^D (1 - B^S)^D\right)}_D Y_t = \underbrace{\left(1 + \sum_{i=1}^q \theta_i B^i\right)}_{\text{MA}(q)} \underbrace{\left(1 + \sum_{i=1}^Q \theta_i B^{iS}\right)}_{\text{MA}(Q)} \varepsilon_t \quad (2.38)$$

donde:

B : Operador retroactivo

Y_t : Serie de tiempo a encontrar

p : Orden autorregresivo (no estacional)

d : Indica el número de veces a diferenciar (no estacional)

q : Orden de promedio móvil (no estacional)

P : Orden autorregresivo (estacional)

D : Señala las veces a diferenciar (estacional)

Q : Orden de promedio móvil (estacional)

φ : Parámetro autorregresivo (no estacional)

θ : Parámetro de promedio móvil (no estacional)

Φ : Parámetro autorregresivo (estacional)

Θ : Parámetro de promedio móvil (estacional)

S : Indica el periodo estacional

ε_t : Ruido blanco

2.1.7 Metodología Box-Jenkins

La metodología Box-Jenkins implica cuatro fases identificación, estimación, examen de diagnóstico y pronóstico (Gujarati y Porter, 2010, pág. 777).

Por otro lado, Uriel (1985, pág. 72) concuerda con lo antes señalado, solamente difiriendo en la tercera etapa, que sería la fase de validación.

Sin embargo, se debe de tomar en cuenta que para la construcción de un modelo se debe precisar de un tratamiento iterativo de tres pasos identificación, estimación y un examen de diagnóstico (Box Jenkins, 1976, pág. 171).

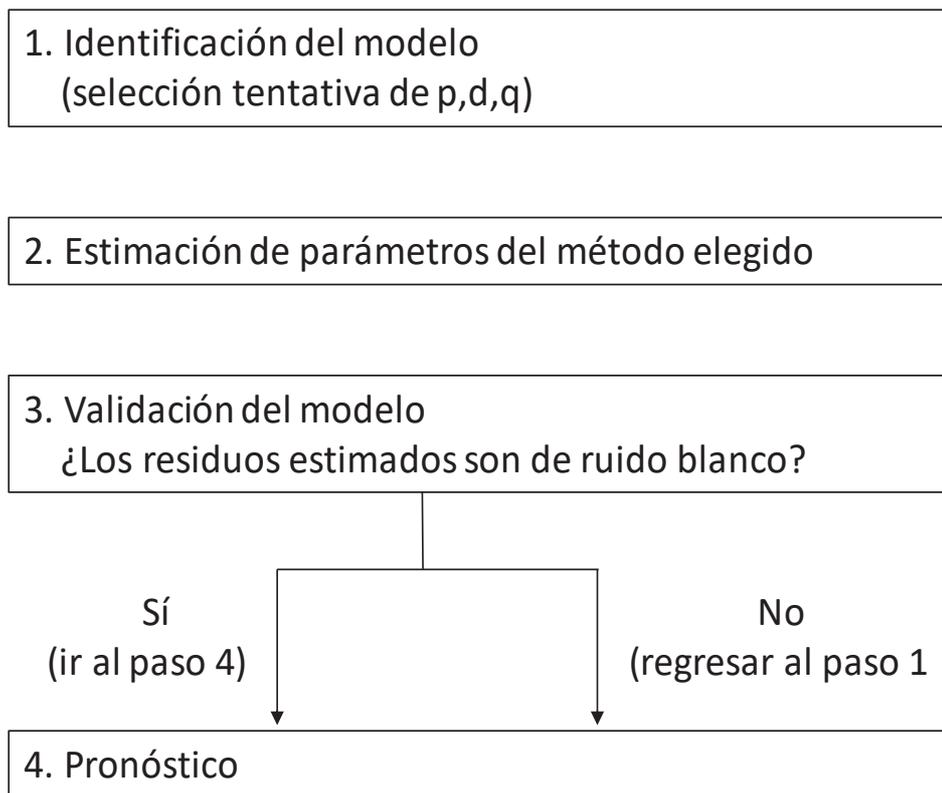


Figura 1. Esquema de la Metodología Box-Jenkins (Gujarati y Porter, 2010, pág. 778).

2.1.7.1 Identificación

Se debe verificar si la serie es estacionaria usando la prueba de KPSS (Kwaitowski, Phillips, Schmidt, Shin), en el caso que no sea estacionaria se presenta dos casos la no estacionariedad en media y la no estacionariedad en varianza para lo cual se debe aplicar la diferenciación de orden d y la transformación Box-Cox mediante el parámetro λ , seguidamente se debe analizar las funciones de autocorrelación simple y parcial de la serie estacionaria, básicamente en esta etapa se tendrá propuestas respecto a que valores puede tomar p, d y q (Uriel, 1985, pág. 73, 93, 95).

Para entender la prueba KPSS, se parte del modelo siguiente

$$Y_t = \beta' D_t + \mu_t + \mu_t$$

$$\mu_t = \mu_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{ruido blanco}(0, \sigma_\varepsilon^2)$$

donde D_t contiene componentes deterministas (constante o constante con tendencia en el tiempo), μ_t puede ser heterocedástico y β' vienen a ser los coeficientes del modelo; el estadístico de la prueba KPSS es el multiplicador de Lagrange donde las hipótesis respecto de Y_t vienen formulados como $H_0 : \sigma_\varepsilon^2 = 0$ (μ_t es constante) y $H_a : \sigma_\varepsilon^2 > 0$ (μ_t no es constante) y está dada por

$$KPSS = \left(T^{-2} \sum_{t=1}^T \hat{S}_t^2 \right) / \hat{\lambda}^2 \quad (2.39)$$

donde $\hat{S}_t = \sum_{j=1}^t \hat{\mu}_j$, $\hat{\mu}_t$ son los residuales de la regresión de Y_t en D_t y $\hat{\lambda}^2$ es una estimación consistente de la varianza a largo plazo de μ_t , asimismo cabe precisar que la hipótesis $H_0 : \sigma_\varepsilon^2 = 0$ converge a una función de movimiento browniano estándar que depende de la forma de los términos deterministas de D_t y no de sus valores de los coeficientes de β' (Zivot y Wang, 2005, pág.129).

2.1.7.2 Estimación

En esta etapa se pretende determinar los estimadores de los parámetros $\varphi_1, \varphi_2, \dots, \varphi_p, \omega_1, \omega_2, \dots, \omega_p, \mu_t, \sigma_\varepsilon^2$ es decir buscar un valor para los parámetros los cuales correspondan lo más cercano posible a los observados (Hang, 2010, pág. 44 y Uriel, 1985, pág. 109).

2.1.7.2.1 Estimación por Máxima Verosimilitud exacta o no condicional (MVE)

Para abordar el tema de la estimación por MVE se toma como referencia Woodward, Gray y Elliott (2016, pág. 286-288) de la forma siguiente:

Para la estimación por MVE suponemos que a_t tiene una distribución normal y por lo tanto $X = X_1, \dots, X_n$ tiene una distribución normal multivariante con μ y con una matriz de covarianza Σ_n donde

$$\mu = (\mu, \mu, \dots, \mu)' \quad y \quad \Sigma_n = \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \cdots & \gamma_0 \end{pmatrix}$$

luego se analiza $X_t - \bar{X}$ y asumimos que $\mu = 0$

si $\sigma_a^2 M_n^{-1} = \Sigma_n$ la función de densidad de probabilidad conjunta de $x = (x_1, \dots, x_n)'$ viene dada por

$$f(x_1, \dots, x_n) = \frac{1}{(2\pi\sigma_a^2)^{n/2}} |M_n|^{1/2} \exp\left\{-\frac{x' M_n x}{2\sigma_a^2}\right\} \quad (2.40)$$

la función de densidad conjunta de la anterior ecuación se puede escribir como

$$f(x_1, \dots, x_n) = f(x_{p+1}, \dots, x_n | x_1, \dots, x_p) f(x_1, \dots, x_p)$$

donde $f(\cdot | \cdot)$ es la función de densidad de probabilidad condicional, las a_t siguen una distribución de ruido blanco $N(0, \sigma_a^2)$ se deduce que

$$f(a_{p+1}, \dots, a_n) = \frac{1}{(2\pi\sigma_a^2)^{(n-p)/2}} \exp\left\{-\frac{1}{2\sigma_a^2} \sum_{t=p+1}^n a_t^2\right\}$$

después se tiene que

$$a_t = X_t - \varphi_1 X_{t-1} - \cdots - \varphi_p X_{t-p} + \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q}$$

Si $X_1 = x_1, \dots, X_p = x_p$ y $a_1 = a_1, \dots, a_p = a_p$, entonces la anterior ecuación (por conveniencia $q \leq p$) que para una elección de φ_i 's y θ_i 's, las variables aleatorias a_{p+1}, \dots, a_n y X_{p+1}, \dots, X_n están relacionadas por el sistema de ecuaciones siguientes

$$\begin{aligned}
a_{p+1} &= X_{p+1} - \varphi_1 X_p - \dots - \varphi_p X_1 + \theta_1 a_p + \dots + \theta_q a_{p+1-q} \\
a_{p+2} &= X_{p+2} - \varphi_1 X_{p+1} - \dots - \varphi_p X_2 + \theta_1 a_{p+1} + \dots + \theta_q a_{p+2-q} \\
&\vdots \\
a_n &= X_n - \varphi_1 X_{n-1} - \dots - \varphi_p X_{n-p} + \theta_1 a_{n-1} + \dots + \theta_q a_{n-q}
\end{aligned}$$

Los valores dados para a_1, \dots, a_p anteriormente se considera que son cero, por otro lado el valor absoluto del Jacobiano de esta transformación es uno entonces se deduce que

$$\begin{aligned}
&f(x_{p+1}, \dots, x_n | x_1, \dots, x_p) \\
&= \frac{1}{(2\pi\sigma_a^2)^{(n-p)/2}} \exp\left\{-\frac{1}{2\sigma_a^2} \sum_{t=p+1}^n (x_t - \varphi_1 x_{t-1} - \dots - \varphi_p x_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q})^2\right\} \\
&= \frac{1}{(2\pi\sigma_a^2)^{(n-p)/2}} \exp\left\{-\frac{S_c}{2\sigma_a^2}\right\}
\end{aligned}$$

donde S_c es

$$S_c = (x_t - \varphi_1 x_{t-1} - \dots - \varphi_p x_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q})^2$$

seguidamente

$$f(x_1, \dots, x_p) = \frac{1}{(2\pi\sigma_a^2)^{p/2}} |M_p|^{1/2} \exp\left\{\frac{x_p' M_p x_p}{2\sigma_a^2}\right\} \quad (2.41)$$

donde $x_p = (x_1, \dots, x_p)'$, después se multiplica a la anterior ecuación por $f(x_{p+1}, \dots, x_n | x_1, \dots, x_p)$ se tiene

$$\begin{aligned}
f(x_1, \dots, x_n) &= f(x_{p+1}, \dots, x_n | x_1, \dots, x_p) f(x_1, \dots, x_p) \\
&= \frac{1}{(2\pi\sigma_a^2)^{(n-p)/2}} \exp\left\{-\frac{S_c}{2\sigma_a^2}\right\} \frac{1}{(2\pi\sigma_a^2)^{p/2}} |M_p|^{1/2} \exp\left\{\frac{x_p' M_p x_p}{2\sigma_a^2}\right\} \\
&= \frac{1}{(2\pi\sigma_a^2)^{(n)/2}} |M_p|^{1/2} \exp\left\{-\frac{S_u}{2\sigma_a^2}\right\}
\end{aligned}$$

donde

$$S_u = S_c + x_p' M_p x_p$$

S_u es la suma incondicional de cuadrados y S_c es la suma condicional de cuadrados, finalmente, la estimación exacta por máxima verosimilitud queda expresada de la forma siguiente

$$l(\varphi_1, \dots, \varphi_p | \theta_1, \dots, \theta_p, \sigma_a^2) = -\frac{n}{2} \ln(2\pi\sigma_a^2) + \frac{1}{2} \ln |M_p| - \frac{S_u}{2\sigma_a^2} \quad (2.42)$$

2.1.7.3 Validación

2.1.7.3.1 Análisis de los residuales

Se verifica el comportamiento de ruido blanco de los residuales para lo cual se hace uso de las pruebas de normalidad de Shapiro-Wilk con el fin de evaluar la homocedasticidad y la prueba de Ljung-Box para evaluar la incorrelación de los residuales.

Tomando en cuenta a Basu (2016, pág. 356) la Prueba de Ljung-Box queda expresada de la siguiente forma

$$L = n(n+2) \sum_{k=1}^h \frac{\rho_k^2}{(n-k)'} \quad (2.43)$$

donde

n es el tamaño de la serie

ρ_k es la autocorrelación en el k -ésimo retraso y también prueba si todas las autocorrelaciones hasta el retraso h no son significativas, las hipótesis se plantean de la forma siguiente

$H_0: \rho_1 = \rho_2 = \dots = \rho_k = 0$ (los datos se distribuyen de forma independiente)

$H_a: \text{al menos } \rho_k \neq 0$ (los datos no se distribuyen independientemente)

Con respecto a la prueba de Shapiro-Wilk (Sahai, 2000, pág. 93 y Verzani, 2014, pág. 357), queda expresada

$$W = \frac{(\sum_{i=1}^n a_i y_{(i)})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.44)$$

donde

$y_1 \leq y_2 \leq \dots \leq y_n$ representa el orden del estadístico

a_i son los pesos óptimos para el estimador de mínimos cuadrados ponderados de la desviación estándar para una población normal.

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

en cuanto a las hipótesis se formulan de la manera siguiente

H_0 : Los datos se distribuyen en forma normal

H_a : Los datos no se distribuyen en forma normal

2.1.7.3.2 Bondad de ajuste

Se evalúa bajo el criterio del AIC

2.1.7.3.3 Criterio de información Akaike (AIC)

El AIC se expresa mediante la ecuación matemática siguiente

$$AIC = -2 \ln ML + 2k \quad (2.45)$$

donde

ML es la máxima verosimilitud

k es el número de parámetros ajustados

el AIC viene a ser una medida de calidad de un buen ajuste estadístico y a su vez la parsimonia del modelo (Hipel y McLeod, 1994, pág. 210).

2.1.7.4 Predicción

En esta etapa el pronóstico obtenido se evalúa con el error porcentual absoluto medio y la raíz del error cuadrático medio.

2.1.7.4.1 Error porcentual absoluto medio (MAPE)

El MAPE queda representado mediante la siguiente expresión matemática

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{A_t - P_t}{A_t} \right| \quad (2.46)$$

donde A_t es el valor actual, P_t es el valor pronosticado y n es el número de observaciones, por otro lado, se recomienda usar el MAPE cuando los datos en la serie no presenten cambios bruscos como aumento o disminución repentinos, debido a que se obtendrá valores distorsionados del error (Khurram, 2010, pág. 13).

2.1.7.4.2 Raíz del error cuadrático medio (RMSE)

El RMSE se expresa con la ecuación matemática siguiente

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (A_t - P_t)^2}{n}} \quad (2.47)$$

donde A_t es el valor actual, P_t es el valor pronosticado y n es el número de observaciones; los errores se elevan al cuadrado antes de ser promediados motivo por el cual se obtiene valores altos cuando se tiene datos atípicos, es decir hay una sobreestimación de los errores grandes pero dicho efecto es mitigado en algo al aplicar la raíz cuadrada al final (Khurram, 2010, pág. 15).

2.1.8 Modelo de la red neuronal recurrente de corta y larga memoria (RNNLSTM)

2.1.8.1 Cuestiones previas sobre las redes neuronales artificiales

2.1.8.1.1 ¿Qué son las redes neuronales artificiales?

Las redes neuronales artificiales son una inteligencia artificial que emula a las neuronas del cerebro humano el cual procesa y guarda la información, siendo la unidad básica la neurona (Graupe, 2007, pág. 1).

2.1.8.1.2 Perceptrón simple

El perceptrón simple fue creado por Rosenblatt en 1958, el cual consta de un solo nodo o neurona que aplica una función de paso a la suma ponderada neta de sus entradas, viene hacer el modelo más simple de una red neuronal (Menhrotra, Mohan y Ranka, 1996, pág. 43).

El perceptron simple se puede modelar de la siguiente forma (Lewis, 2017, pág. 26):

$$f(u) = \sum_{i=1}^n w_{ij}x_j + b_j \quad (2.48)$$

donde

x_j es el dato de entrada,

w_{ij} son los pesos asociados con cada conexión a la neurona

b_j es el sesgo que es similar al intercepto de una un modelo de regresión lineal

$f(.)$ es la función de transferencia que permite el avance de la información

$u = \sum_{i=1}^n w_{ij}x_j + b_j$ es la neurona o nodo

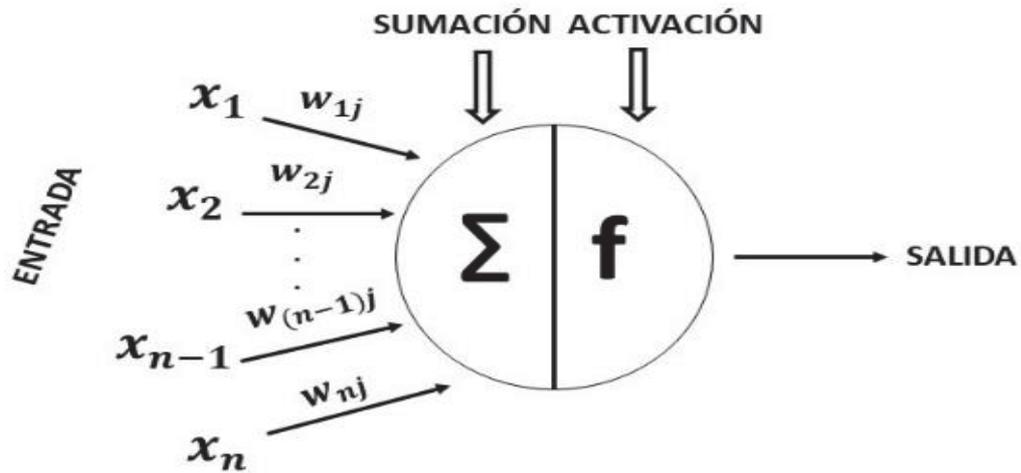


Figura 2. Perceptrón simple

Fuente: Lewis (2017 pág. 26)

2.1.8.1.3 Componentes de redes neuronales

Castillo (1999, pág. 7-11), señala que son la neurona y la forma como están organizadas las neuronas.

2.1.8.1.3.1 Neurona

Es la unidad de procesamiento donde arriba la información procedente de otras neuronas, que luego es procesada mediante la actividad lineal que viene a ser la sumatoria del producto del dato de entrada (x_n) asociado a un peso (w_{nj}) y finalmente obtener una información de salida mediante el uso de la función de activación (Ciaburro y Venkateswaran, 2017, pág. 10).

En cuanto a las funciones de activación se tiene los siguientes:

a) Función lineal

Dicha función pasa la señal sin cambios, y matemáticamente es representada como la función identidad $f(x) = x$ (Patterson y Gibson, 2017, pág. 66).

b) Función de paso

Castillo (1999 pág. 10), indica que esta función da una salida binaria ya sea positiva o negativa, queda representada mediante la función signo

$$\text{signo}(x) \begin{cases} -1, & \text{si } x < 0 \\ 1, & \text{si } x > 0 \end{cases}$$

c) Función sigmoidea

Tomando en cuenta a Patterson y Gibson (2017, pág. 66-67) y Castillo (1999, pág. 10), esta función contiene funciones continuas no lineales que a continuación se detalla:

c1) Función logística

Transforma valores grandes a valores simples entre cero y uno, una representación de esta es

$$f(x) = \frac{1}{1 + e^{-x}}$$

c2) Función tangente hiperbólica

Es una función trigonométrica hiperbólica y viene a significar la relación entre el seno hiperbólico y el coseno hiperbólico, al igual que la función logística transforma valores grandes a valores pequeños entre -1 y 1, una representación de dicha función es

$$f(x) = \tanh(x) = \frac{\sinh(x)}{\cosh(x)}$$

Asimismo, cabe precisar que el uso de la mencionada función es muy útil debido a que el entrenamiento se extiende en forma uniforme sobre cada capa de la red neuronal artificial (Kalman y Kwasny, 1992, pág. IV-578, IV-581).

2.1.8.1.3.2 Arquitectura de la red neuronal

Sobre el tema en cuestión se tiene las siguientes (Mehrotra, Mohan y Ranka, 1996):

a) Redes completamente conectadas

Se caracteriza porque cada nodo está unido a cada nodo, dichas uniones pueden ser de tipo excitatorio cuando se tiene pesos positivos, inhibitorio si presenta pesos negativos y fútil con pesos casi nulos; asimismo, se tiene dentro de estas las redes simétricas, como se puede apreciar en la figura 3.

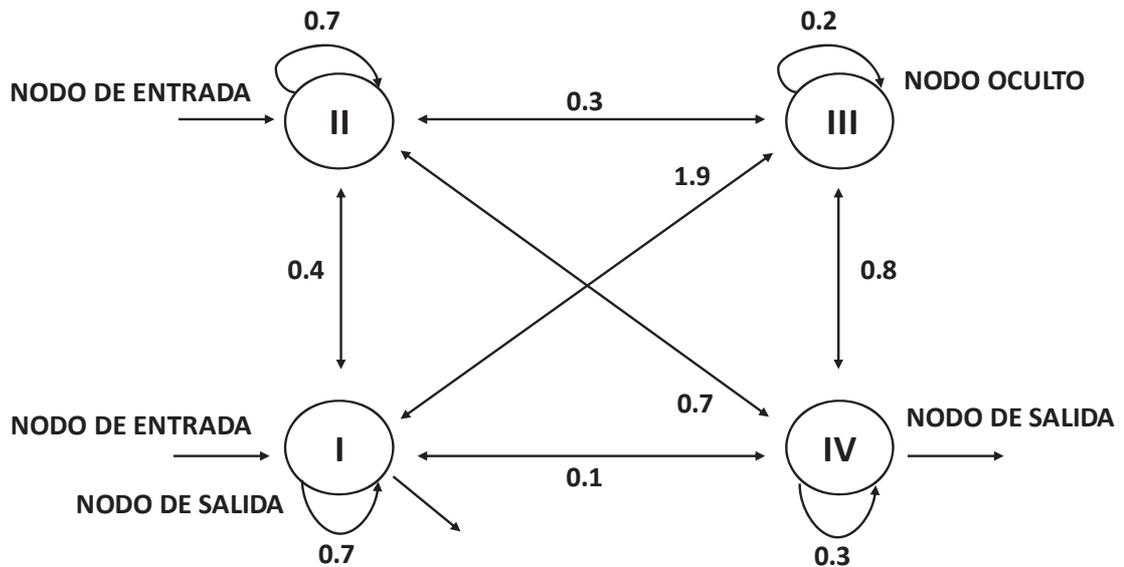


Figura 3. Red simétrica totalmente conectada

Fuente: Mehrotra, Mohan y Ranka (1996, pág. 18).

b) Redes en capas

En dicha red los nodos están particionados en subconjuntos denominados capas, una única señal llega a uno de los nodos de la capa de entrada y luego la señal atraviesa la capa 1, la capa 2 y finalmente se obtiene una salida en la capa 3, además en las capas ocultas existe conexiones intracapas, lo cual se puede observar en la figura 4.

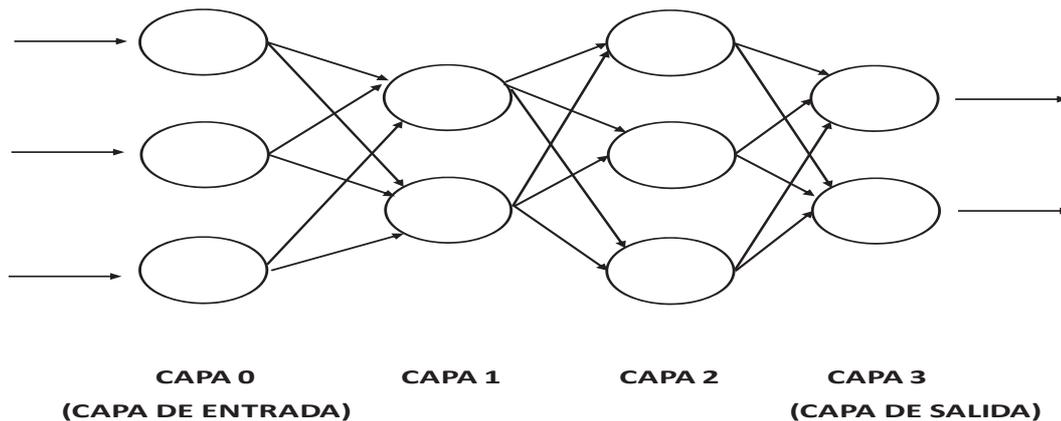


Figura 6. Red de alimentación hacia adelante

Fuente: Mehrotra, Mohan y Ranka (1996, pág. 20).

2.1.8.2 Red Neuronal Recurrente (RNN)

La red neuronal recurrente simple conocida también como la red de Elman (Goldberg, 2017, pág. 177).

Esta se caracteriza básicamente porque las conexiones entre las neuronas forman un ciclo, asimismo pueden utilizar una memoria interna para su procesamiento (Ciaburro y Venkateswaran, 2017, pág. 185).

Cuando se pretende modelar las dependencias de largo alcance en la estructura de un grupo de datos de entrada mediante el uso de redes neuronales recurrentes, se presenta un inconveniente denominado gradiente evanescente el cual se produce cuando las gradientes son muy grandes o muy pequeños (Patterson y Gibson, 2017, pág. 311).

2.1.8.2.1 Estructura

En la siguiente figura N° 07 se aprecia una representación de una red neuronal recurrente simple, donde se observa cuatro capas, de entrada, oculta, asociativa y de salida en las cuales se encuentran las unidades o nodos de entrada, ocultas, contexto y salida respectivamente (Pham y Xing, 1996, pág. 222). Por otro lado k , y , x , u y x_c representan el momento, vector de salida de los nodos de salida, vector de salida de los nodos ocultos, vector de entrada hacia los nodos entrada,

y vector de salida de los nodos de contexto respectivamente; W^1, W^2, W^3 representan la matriz de pesos que tienen la función de mantener unidos la capa oculta con la capa de salida, la capa de entrada con la capa oculta y la capa de asociación con la capa oculta respectivamente, finalmente indicar que en la capa oculta se encuentra una función de transferencia no lineal y en la capa de salida una función de transferencia lineal (Ren et al., 2009, pág. 56).

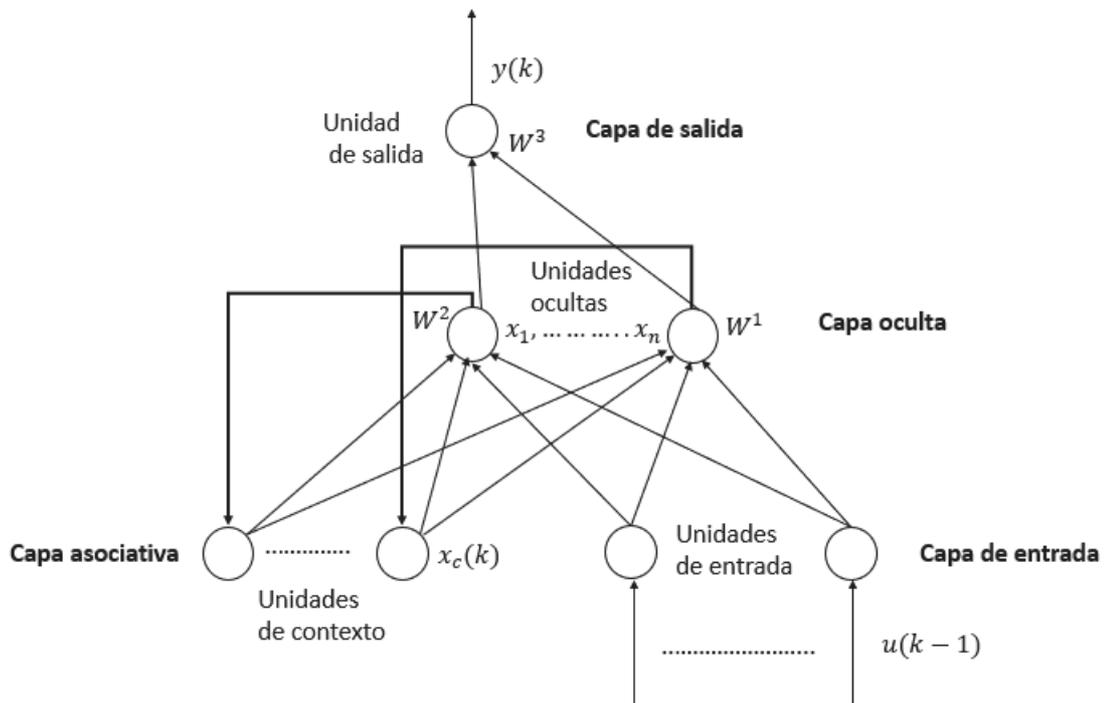


Figura 7. Red neuronal recurrente

Fuente: Ren et al. (2009, pág. 56) y Pham y Xing (1995, pág 49)

Considerando a Pham y Xing (1996, pág 222), He y Tian (2007, pág. 552) y Ren et al. (2009, pág. 56) una RNN se modela de la siguiente forma

$$v_i(k) = \sum_{j=1}^n W_{i,j}^x(k-1)x_j^c(k) + W_i^u(k-1)u(k)$$

$$x_i(k) = f(v_i)$$

$$x_j^c(k) = x_j(k-1)$$

$$y(k) = g\left(\sum_{i=1}^n W_i^y(k-1)x_i(k)\right) \quad (2.49)$$

donde

k es el momento de tiempo dado

$u(k)$ es la entrada hacia la red neuronal

$y(k)$ es la salida de la red neuronal

$v_i(k)$ es la entrada total del i -ésimo nodo oculto

$x_i(k)$ es la salida del i -ésimo nodo oculto

$x_j^c(k)$ es la salida del j -ésimo nodo de contexto

$W_{i,j}^x$ $ij = 1,2, \dots n$ es la matriz de pesos que sirve de conexión entre los nodos de contexto y nodos ocultos

W_i^u $i = 1,2, \dots n$ es la matriz de pesos que sirve de conexión entre los nodos de entrada y nodos ocultos

W_i^y $i = 1,2, \dots n$ es la matriz de pesos que sirve de conexión entre los nodos ocultos y los nodos de salida

$f(v_i)$ es la función de transferencia (función sigmoideal) de los nodos de la capa oculta

$g(\sum_{i=1}^n W_i^y (k-1)x_i(k))$ es la función de transferencia (función lineal) de los nodos de la capa de salida

2.1.8.2.2 Entrenamiento de la red

El entrenamiento de la red se realiza con el fin de que pueda aprender una determinada tarea, dicho aprendizaje se llama propagación hacia atrás; el procedimiento de aprendizaje estándar se apoya en la optimización del gradiente la cual consiste en establecer una relación entre los parámetros del modelo y la función de pérdida, lo que permite propagar la información de gradiente calculada en la función de pérdida a los parámetros del modelo con el fin de modificarlos (Bianchi et al., 2017, pág. 11)

2.1.8.2.2.1 Propagación hacia atrás dinámica

El vector de retroalimentación $x^c(k) = \{x_i^c(k)\}$ es $x(k-1)$ el cual está en función de

$$W^x(k-2)x(k-2) + W^u(k-2)u(k-2)$$

luego $x^c(k-1) = \{x_i^c(k)\}$ es $x(k-2)$ el cual esta en función de

$$W^x(k-3)x(k-3) + W^u(k-3)u(k-3)$$

de lo anterior implica que $x_c(k)$ depende de los pesos un paso de tiempo antes $x(k-1)$, dicha dependencia debe ser considerada al realizar la propagación hacia atrás dinámica (Pham y Liu, 1996, pág. 222).

La propagación hacia atrás es la actualización de los pesos de las matrices en base a la diferencia entre el dato real o deseado y el dato de salida de la red, lo antes mencionado se realiza hasta que dicha diferencia sea minima; las actualizaciones de pesos de las matrices se resumen con la siguiente expresión (Bokka et al., 2019, pág. 133)

$$W_{nuevo} = W_{anterior} + \Delta W \quad (2.50)$$

donde

W_{nuevo} es la matriz de pesos actualizada

$W_{anterior}$ es la matriz de pesos no actualizada

ΔW es el gradiente (error) que se define como $-\eta \frac{\partial E_k}{\partial W}$, η es el grado de aprendizaje y E_k es el error al cuadrado

A continuación, se desarrolla la propagación hacia atrás dinámica tomando en cuenta a Pham y Liu (1996, pág. 222)

Partimos de dos datos $u(k)$ que es el dato de entrada y $y_d(k)$ el dato deseado a conseguir

$$(u(k), y_d(k)), \quad k = 1, 2, \dots, N$$

Cuando una entrada y salida de datos es presentada a la RNN en un tiempo k , el error al cuadrado en la salida de la red se define como

$$E_k = \frac{1}{2} (y_d(k) - y(k))^2 \quad (2.51)$$

Para que la red aprenda los pesos de las matrices deben actualizarse cada paso de tiempo k en base al E_k

El gradiente $\Delta W_i^y(k)$ para la matriz $W_i^y(k-1)$ es

$$\begin{aligned}\Delta W_i^y(k) &= -\eta \frac{\partial E_k}{\partial W_i^y(k-1)} = \eta ((y_d(k) - y(k)) \frac{\partial y(k)}{\partial W_i^y(k-1)}) \\ &= \eta (y_d(k) - y(k)) x_i(k)\end{aligned}$$

El gradiente $\Delta W_i^u(k)$ para la matriz $W_i^u(k-1)$ es

$$\begin{aligned}\Delta W_i^u(k) &= -\eta \frac{\partial E_k}{\partial W_i^u(k-1)} = \eta \frac{\partial E_k}{\partial y(k)} \frac{\partial y(k)}{\partial x_i(k)} \frac{\partial x_i(k)}{\partial v_i(k)} \frac{\partial v_i(k)}{\partial W_i^u(k-1)} \\ &= \eta (y_d(k) - y(k)) W_i^y(k-1) f_{v_i} u(k)\end{aligned}$$

f_{v_i} es la derivada de f respecto a v_i

El gradiente $\Delta W_{i,j}^x(k)$ para la matriz $W_{i,j}^x(k-1)$ es

$$\begin{aligned}\Delta W_{i,j}^x(k) &= -\eta \frac{\partial E_k}{\partial W_{i,j}^x(k-1)} = \eta \frac{\partial E_k}{\partial y(k)} \frac{\partial y(k)}{\partial x_i(k)} \frac{\partial x_i(k)}{\partial W_{i,j}^x(k-1)} \\ &= \eta (y_d(k) - y(k)) W_i^y(k-1) \frac{\partial x_i(k)}{\partial W_{i,j}^x(k-1)}\end{aligned}$$

En cuanto a la actualización de los pesos esta se realiza en forma iterativa, con la siguiente expresión

$$\begin{aligned}\frac{\partial x_i(k)}{\partial W_{i,j}^x(k-1)} &= \frac{\partial x_i(k)}{\partial v_i(k)} \frac{\partial v_i(k)}{\partial W_{i,j}^x(k-1)} = f_{v_i} \frac{\partial v_i(k)}{\partial W_{i,j}^x(k-1)} \\ &= f_{v_i} \{x_j(k-1) + \sum_{l=1}^n W_{i,l}^x(k-1) \frac{\partial x_l(k-1)}{\partial W_{i,j}^x(k-2)}\}\end{aligned}$$

donde

$$f_{v_i} = \frac{\partial f}{\partial v_i} = (1 - f^2(v_i))$$

$$\frac{\partial x_l(k-1)}{\partial W_{l,j}^x(k-2)} = f_{v_l} \sum_{m=1}^n W_{j,m}^x(k-2) \frac{\partial x_m(k-2)}{\partial W_{l,j}^x(k-3)}, l \neq i$$

2.1.8.2.3 Gradientes explosivos y gradientes desvanecidos

Los gradientes explosivos es un problema que se presenta después de un tiempo de pasos, es decir las derivadas parciales y la multiplicación del gradiente dan un valor mayor a uno, lo cual hace que la red no pueda aprender correctamente; en cuanto a los gradientes evanescentes es lo opuesto a los gradientes desvanecidos, es decir en las derivadas parciales y la multiplicación del gradiente se obtiene valores muy bajos por lo que el gradiente se desvanece y en consecuencia la actualización de los pesos no se produce (Bokka et al., 2019, pág. 142).

2.1.8.3 Redes Neuronales Recurrentes de Memoria a corto y largo plazo (RNNLSTM)

La RNNLSTM fue una propuesta de Hochreiter y Schmidhuber en 1997, con el fin de superar el problema de los gradientes explosivos y desvanecidos, que presentan las redes neuronales recurrentes.

La unidad básica de una RNNLSTM es una memoria bloque con uno o más tipos de celdas memoria y tres multiplicaciones adaptativas llamadas puerta de entrada, puerta de olvido y puerta de salida (Zhuge, Xu y Zhang, 2017, pág. 171).

2.1.8.3.1 Arquitectura de la RNNLSTM

Queda expresada en la figura 8, donde se aprecia una RNNLSTM con tres capas, la capa de entrada, una capa oculta simple (solo una capa) y la capa de salida, asimismo se observa que tanto el número de datos de entrada x_n como de salida y_n son iguales y que el número de bloques de memoria de la capa oculta no necesariamente es igual al número de datos de entrada y finalmente se ve que hay una recurrencia al mismo bloque de memoria antes de la salida de los datos.

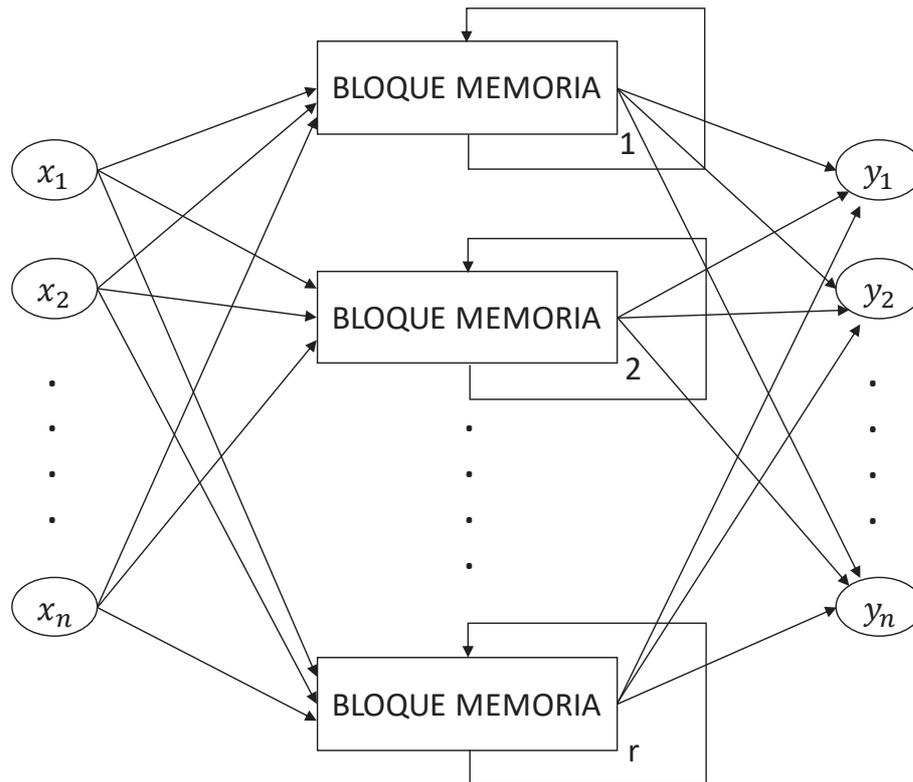


Figura 8. Arquitectura de una RNNLSTM simple

Fuente. González (2019, pág. 28) y Zhu et al, (2016, pág. 3698)

La arquitectura del RNNLSTM de la figura 1 se expresaría matemáticamente de la manera siguiente

$$y_{nx1}^t = \sum_{p=1}^r h_{\text{Bloque memoria}_p}^t \quad (2.52)$$

donde

y_{nx1}^t es un vector que contiene los datos de salida de la red neuronal

$h_{\text{Bloque memoria}}^t$ es un vector que contiene datos de salida del bloque memoria

r indica el número de bloques memoria de la red

p indica que se empieza con un bloque de memoria

t indica el tiempo

2.1.8.3.2 Arquitectura del bloque memoria

Queda expresada de la forma siguiente

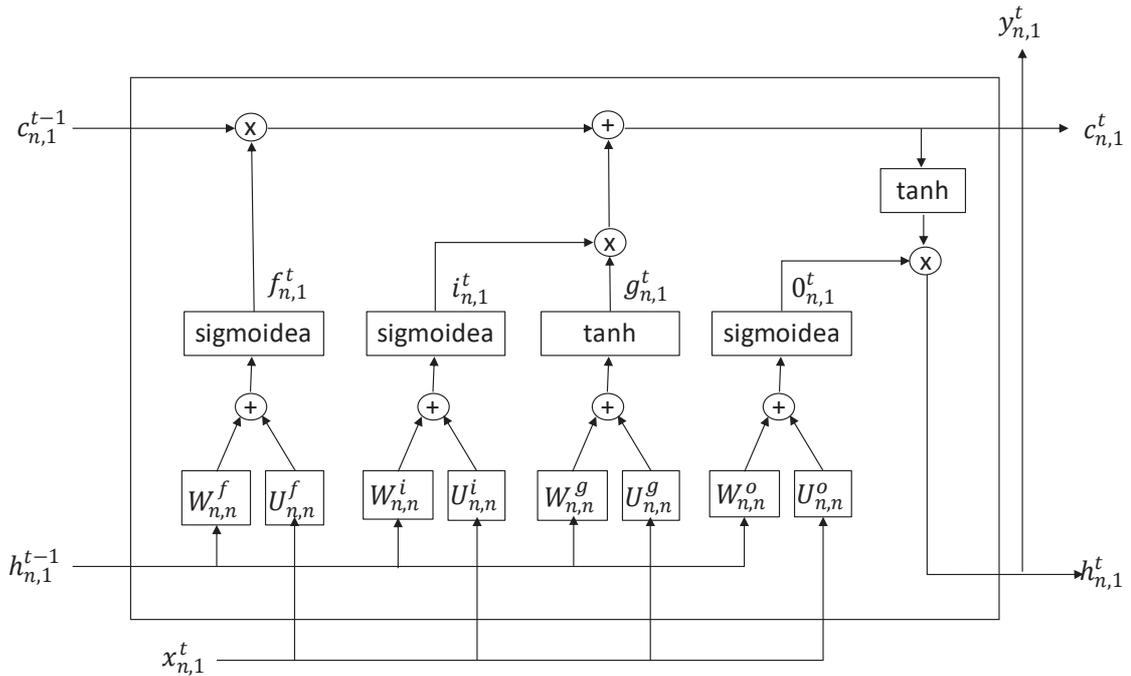


Figura 9. Propagación hacia adelante del bloque memoria de una RNNLSTM
Fuente. Fichou (2019)

A continuación, se describe la figura 9 considerando a Patterson y Gibson (2017, pág. 155) y Smagulova y James (2019, pág. 142, 143, 146), donde:

$f_{nx1}^t, i_{nx1}^t, o_{nx1}^t, g_{nx1}^t, c_{nx1}^t, c_{nx1}^{t-1}, h_{nx1}^{t-1}, h_{nx1}^t, y_{nx1}^t, x_{nx1}^t$ son vectores, todos con la misma dimensionalidad en el tiempo t .

$W_{n \times n}^f, U_{n \times n}^f$ son matrices con la misma dimensión.

$i_{nx1}^t, f_{nx1}^t, o_{nx1}^t$ son las puertas de entrada, de olvido y de salida respectivamente.

x_{nx1}^t es el vector de entrada.

y_{nx1}^t es el vector de salida

h_{nx1}^t y h_{nx1}^{t-1} son los vectores que representan el estado del bloque memoria en un tiempo t y $t - 1$ respectivamente.

c_{nx1}^t, c_{nx1}^{t-1} son los vectores que representan el estado de la celda memoria en un tiempo t y $t - 1$ respectivamente.

g_{nx1}^t es el vector candidato a ser incluido en la celda memoria actual, también se le denomina activador de la celda memoria

$\text{sigmoidea}(\cdot)$ es la función sigmoidea logística $\sigma(x) = \frac{1}{1+e^{-x}}$

$\tanh(\cdot)$ es la función tangente hiperbólica $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

en consecuencia el bloque memoria de la RNNLSTM se puede modelar relacionando las anteriores expresiones matemáticas de la manera siguiente:

Las puertas $f_{n,x}^t, i_{nx1}^t, o_{nx1}^t$ son las puertas de olvido, entrada y salida respectivamente.

$$f_{nx1}^t = \text{sigmoidea}(W_{n,n}^f * h_{nx1}^{t-1} + U_{n,n}^f * x_{nx1}^t) \quad (2.53)$$

$$i_{nx1}^t = \text{sigmoidea}(W_{n,n}^i * h_{nx1}^{t-1} + U_{n,n}^i * x_{nx1}^t) \quad (2.54)$$

$$o_{nx1}^t = \text{sigmoidea}(W_{n,n}^o * h_{nx1}^{t-1} + U_{n,n}^o * x_{nx1}^t) \quad (2.55)$$

$g_{n,1}^t$ celda memoria candidata a ser incluida en la celda memoria actual

$$g_{nx1}^t = \tanh(W_{n,n}^g * h_{nx1}^{t-1} + U_{n,n}^g * x_{nx1}^t) \quad (2.56)$$

$c_{n,1}^t$ es el estado de la celda memoria actual y \odot es el producto de Hadamard

$$c_{nx1}^t = c_{nx1}^{t-1} \odot f_{nx1}^t + i_{nx1}^t \odot g_{nx1}^t \quad (2.57)$$

$y_{n,1}^t$ es la salida de la red neuronal

$$y_{nx1}^t = h_{nx1}^t = \tanh(c_{nx1}^t) \odot o_{nx1}^t \quad (2.58)$$

Juntando las expresiones matemáticas indicadas anteriormente y asumiendo que la RNNLSTM tiene un solo bloque memoria, la ecuación 2.58 se expresa de la siguiente manera

$$y_{nx1}^t = h_{nx1}^t = \tanh(c_{nx1}^t) \odot o_{nx1}^t$$

$$y_{nx1}^t = h_{nx1}^t = \tanh(c_{nx1}^{t-1} \odot f_{nx1}^t + i_{nx1}^t \odot g_{nx1}^t) \odot \text{sigmoidea}(W_{n,n}^o * h_{nx1}^{t-1} + U_{n,n}^o * x_{nx1}^t)$$

$$y_{nx1}^t = h_{nx1}^t = \tanh(c_{nx1}^{t-1} \odot \text{sigmoidea}(W_{n,n}^f * h_{nx1}^{t-1} + U_{n,n}^f * x_{nx1}^t) + \text{sigmoidea}(W_{n,n}^i * h_{nx1}^{t-1} + U_{n,n}^i * x_{nx1}^t) \odot \tanh(W_{n,n}^g * h_{nx1}^{t-1} + U_{n,n}^g * x_{nx1}^t)) \odot \text{sigmoidea}(W_{n,n}^o * h_{nx1}^{t-1} + U_{n,n}^o * x_{nx1}^t)$$

2.1.8.3.2.1 Componentes del bloque memoria

Observando la figura 9, del bloque memoria de la RNNLSTM se tiene los siguientes (Géron, 2017, pág. 402 y Patterson y Gibson (2017, pág. 154):

a) Tres puertas

a.1 Puerta de entrada, solo ingresa datos relevantes, asimismo dicha función controla que partes de $g_{(t)}$ deben agregarse a la memoria de largo plazo.

a.2 Puerta de olvido, controla que partes de la memoria a largo plazo debe ser borrado.

a.3 Puerta de salida, controla que partes de la memoria a largo plazo debe leerse y transmitirse en un paso de tiempo hacia $h_{(t)}$ y $y_{(t)}$.

b) Celda memoria (carrusel de error constante)

Esta permite a la red mantenerse en el tiempo.

c) Funciones de activación de las puertas, de salida y de entrada

c.1) Función de activación de las puertas, mediante esta función se controla las puertas de olvido por $(f_{(t)})$, de entrada $(i_{(t)})$ y de salida por $(o_{(t)})$, dichas funciones son de tipo sigmoidea.

c.2) Función de activación de entrada y de salida, mediante estas funciones se controla que dato sale o entra, en la celda básica estas funciones son de tipo tangente hiperbólica.

2.1.8.3.3 Funcionamiento de la red

Tomando en cuenta a Géron (2017, pág. 402) y la figura 9, el funcionamiento de la RNNLSTM es como sigue:

Primero indicar que la celda básica se divide en dos estados, $h_{(t)}$ como el estado a corto plazo y $c_{(t)}$ como el estado a largo plazo; el estado a largo plazo $c_{(t-1)}$ se dirige de izquierda a derecha y que en su recorrido pasa por la puerta de olvido donde elimina todo lo que no sirve es decir deja y agrega algunos recuerdos lo cual se realiza gracias a la operación de adición, después el

resultado $c_{(t)}$ sale sin ninguna alteración, cabe señalar que después de la operación adición el estado a largo plazo se copia y pasa a través de la función tangente donde es filtrado por la puerta de salida lo cual genera el estado a corto plazo; ahora se pasa a explicar de dónde provienen los nuevos recuerdos, para lo cual primero se asume que se suministra datos de entrada al vector $x_{(t)}$ y al estado previo de corto plazo $h_{(t-1)}$ la cual está conectada con cuatro diferentes capas, la capa principal es la que genera a $g_{(t)}$, esta analiza las entradas de $x_{(t)}$ y el estado anterior de corto plazo de $h_{(t-1)}$ además a la salida de esta capa se almacena parcialmente en el estado a largo plazo, dicha capa se activa mediante la función tangente; la otras tres capas funcionan como controladoras de las puertas utilizan la función logística para activarse.

2.1.8.3.4 Entrenamiento de la red

Para que la RNNLSTM pueda entrenarse, primeramente, los datos de entrada deben de fluir hacia adelante y luego se evalúa el dato de salida mediante la función cuadrática del error, luego se realiza la propagación hacia atrás con el fin de distribuir el error en las matrices que contienen los pesos ponderados esta acción se repite hasta que la función cuadrática del error sea mínima lo cual nos indica que la RNNLSTM logro aprender (Yan et al., 2019, pág 3).

2.1.8.3.4.1 Propagación hacia adelante

Se refiere a que los datos ingresan a la red neuronal y avanzan hacia adelante cabe precisar que los pesos ponderados de las matrices se asignan en forma aleatoria, solo en la primera iteración luego se evalúa la salida de los datos con la función del error cuadrático medio (Ghatak, 2019, pág. 26, Ayyadevara, 2019, pág. 13).

Para este caso los datos ingresan a la red neuronal a través de la siguiente ecuación matemática la cual se obtiene al realizar los reemplazos respectivos realizados en la ecuación (2.58)

$$\begin{aligned}
y_{nx1}^t = h_{nx1}^t = & \tanh(c_{nx1}^{t-1} \odot \text{sigmoidea}(W_{n,n}^f * h_{nx1}^{t-1} + U_{n,n}^f * x_{nx1}^t) + \\
& \text{sigmoidea}(W_{n,n}^i * h_{nx1}^{t-1} + U_{n,n}^i * x_{nx1}^t) \odot \\
& \tanh(W_{n,n}^g * h_{nx1}^{t-1} + U_{n,n}^g * x_{nx1}^t)) \odot \\
& \text{sigmoidea}(W_{n,n}^o * h_{nx1}^{t-1} + U_{n,n}^o * x_{nx1}^t)
\end{aligned}$$

2.1.8.3.4.2 Función del error cuadrático medio

También conocida como función de coste, con la cual se evalúa en qué magnitud el valor de salida de la red se acerca al valor objetivo (Qing y Niu, 2018, pág. 465).

La función de coste se expresa con la siguiente ecuación matemática (Chen, 2018, pág. 7)

$$\mathcal{L}(x, \theta) = \min \sum_1^N (y_o - y_s)^2 \quad (2.59)$$

donde

y_o es el dato objetivo

y_s es el dato de salida de la red neuronal quien depende de θ y x

$\theta = \{W_{n,n}^f, W_{n,n}^i, W_{n,n}^o, W_{n,n}^g, U_{n,n}^f, U_{n,n}^i, U_{n,n}^o, U_{n,n}^g\}$ son las matrices que contienen los pesos ponderados los cuales deben ser estimados

x son los datos de entrada

2.1.8.3.4.3 Propagación hacia atrás

Es un algoritmo que permite operar en forma recursiva capa tras capa hacia atrás hasta la primera capa de tal forma se distribuye el error total de la función de coste a todas las capas, para tal fin se se hace uso de las derivadas parciales, la regla de la cadena y el descenso del gradiente (Dot CSV, 2018, 8m36s).

2.1.8.3.4.4 Descenso del gradiente

Es un método por el cual se permite optimizar la función de coste lo cual se traduce en obtener un valor de salida de la red neuronal mas cercano al valor objetivo (Goldberg, 2017, pág. 30).

Con fines de una mayor comprensión se usa las siguientes funciones (Chen, 2018, pag. 7, 8), la función de coste como

$$\mathcal{L}(t) = \frac{1}{2}(y_o - y_t)^2 \quad (2.60)$$

y la función de descenso del gradiente como

$$\mathcal{L}(t - 1) = \mathcal{L}(t) - \eta \partial\mathcal{L}(t) \quad (2.61)$$

donde

y_o es el valor objetivo

y_t es el valor de salida de la red neuronal

$\mathcal{L}(t - 1)$ es un valor más cerca al mínimo global

$\mathcal{L}(t)$ es un valor más lejos del mínimo global

η es el grado de aprendizaje

$\partial\mathcal{L}(t)$ es la derivada de la función de la función de coste que nos da la ubicación en la superficie de la misma función

2.1.8.3.4.5 Actualización de los pesos

La actualización de los pesos se realiza con la siguiente ecuación (Bokka et al., 2019, pág. 133)

$$W_{nuevo} = W_{anterior} + \Delta W \quad (2.62)$$

donde

W_{nuevo} es la matriz de pesos actualizada

$W_{anterior}$ es la matriz de pesos no actualizada

ΔW es el gradiente que se define como $-\eta \frac{\partial \mathcal{L}(t)}{\partial W}$, η es el grado de aprendizaje y $\mathcal{L}(t)$ es la función de coste.

2.1.8.3.5 Propagacion hacia atrás

Para abordar el tema en cuestión es bastante complejo sin embargo para tal fin se utiliza la siguiente figura 10.

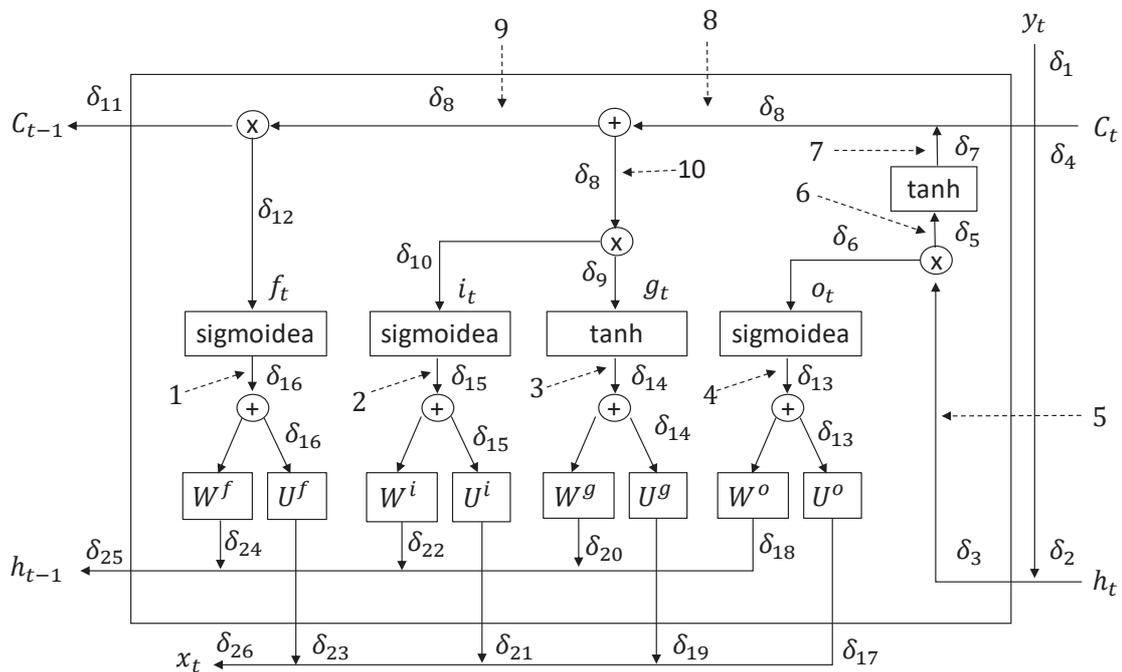


Figura 10. Propagación hacia atrás del bloque memoria de una RNNLSTM
Fuente. Fichou (2019)

Teniendo en cuenta a Fichou(2019), en la figura 1 se aprecia la propagación hacia atrás de un bloque de memoria de una RNNLSTM donde la propagación hacia atrás primeramente proviene de los deltas δ_1 dato que sale fuera de la red neuronal, δ_4 es el estado de la celda memoria y δ_2 es el estado de la celda memoria, el punto (.) y el asterisco (*) significa multiplicación y multiplicación matricial respectivamente, la propagación hacia atrás resulta relacionando las siguientes expresiones:

$$\delta_3 = \delta_1 + \delta_2$$

$$\delta_5 = \delta_3 \cdot o_t = \delta_3 \cdot o_t$$

$$\begin{aligned}
\delta_6 &= \delta_3 \cdot 5 = \delta_3 \cdot \tanh(c_t) \\
\delta_7 &= \delta_5 \cdot f'(5) = \delta_5 \cdot \tanh'(\tanh(c_t)) \\
\delta_8 &= \delta_7 \cdot \delta_4 \\
\delta_9 &= \delta_8 \cdot 10 = \delta_8 \cdot i_t \\
\delta_{10} &= \delta_8 \cdot 9 = \delta_8 \cdot g_t \\
\delta_{11} &= \delta_8 \cdot 12 = \delta_8 \cdot f_t \\
\delta_{12} &= \delta_8 \cdot 11 = \delta_8 \cdot c_{t-1} \\
\delta_{13} &= \delta_6 \cdot f'(6) = \delta_6 \cdot \text{sigmoidea}'(o_t) \\
\delta_{14} &= \delta_9 \cdot f'(9) = \delta_9 \cdot \tanh'(g_t) \\
\delta_{15} &= \delta_{10} \cdot f'(10) = \delta_{10} \cdot \text{sigmoidea}'(i_t) \\
\delta_{16} &= \delta_{12} \cdot f'(12) = \delta_{12} \cdot \text{sigmoidea}'(f_t) \\
\delta_{17} &= \delta_{13} * U_o^T \\
\delta_{19} &= \delta_{14} * U_g^T \\
\delta_{21} &= \delta_{15} * U_i^T \\
\delta_{23} &= \delta_{16} * W_f^T \\
\delta_{18} &= \delta_{13} * W_o^T \\
\delta_{20} &= \delta_{14} * W_g^T \\
\delta_{22} &= \delta_{16} * W_i^T \\
\delta_{24} &= \delta_{16} * W_f^T \\
\delta_{25} &= \delta_{18} + \delta_{20} + \delta_{22} + \delta_{24} \\
\delta_{26} &= \delta_{17} + \delta_{19} + \delta_{21} + \delta_{23}
\end{aligned}$$

Los errores δ_{11} , δ_{25} y δ_{26} son usados para la siguiente capa, una vez que todos los errores están disponibles recién se realiza la actualización de los pesos

$$\begin{aligned}
\delta W_f &= \delta W_f + h_{t-1}^T * \delta_{16} \\
\delta U_f &= \delta U_f + x_t^T * \delta_{16} \\
\delta W_i &= \delta W_i + h_{t-1}^T * \delta_{15} \\
\delta U_i &= \delta U_i + x_t^T * \delta_{15} \\
\delta W_g &= \delta W_g + h_{t-1}^T * \delta_{14} \\
\delta U_g &= \delta U_g + x_t^T * \delta_{14} \\
\delta W_o &= \delta W_o + h_{t-1}^T * \delta_{13} \\
\delta U_o &= \delta U_o + x_t^T * \delta_{13}
\end{aligned}$$

Otra manera más simple de la propagación hacia atrás sería de la forma siguiente (Smagulova y James, 2019, pág. 142):

La información se transmite hacia atrás con el fin de que la red neuronal aprenda y asu vez se produce la actualización de los pesos de las matrices de conección con las expresiones matemáticas

$$gate S_t = \begin{bmatrix} g_t \\ i_t \\ f_t \\ o_t \end{bmatrix}, \quad W = \begin{bmatrix} W^g \\ W^i \\ W^f \\ W^o \end{bmatrix}, \quad U = \begin{bmatrix} U^g \\ U^i \\ U^f \\ U^o \end{bmatrix}, \quad b = \begin{bmatrix} b^g \\ b^i \\ b^f \\ b^o \end{bmatrix}$$

$$\delta h_t = \Delta_t + \Delta h_t$$

$$\delta C_t = \delta h_t \odot o_t \odot (1 - \tanh^2(C_t)) + \delta C_{t+1} \odot f_{t+1}$$

$$\delta g_t = \delta C_t \odot i_t \odot (1 - g_t^2)$$

$$\delta i_t = \delta C_t \odot g_t \odot (1 - i_t)$$

$$\delta f_t = \delta C_t \odot C_{t-1} \odot f_t \odot (1 - f_t)$$

$$\delta o_t = \delta h_t \odot \tanh(C_t) \odot o_t \odot (1 - o_t)$$

$$\delta x_t = W^t \delta gate S_t$$

2.2 MARCO CONCEPTUAL

2.2.1 Series de tiempo

Conjuntos de datos estadísticos referentes a distintos momentos o intervalos de un periodo y a un mismo indicador, presentado en secuencia cronológica (INEGI, 2012, pág. 20).

Conocida también como serie cronológica. Es la sucesión de observaciones cuantitativas ordenadas en el tiempo de un fenómeno. Los datos tienen un orden que no es posible variar. La información puede ser mensual, trimestral, anual o de cualquier otro intervalo temporal (INEI, 2006, pág.57).

2.2.2 Red neuronal

Según el diccionario Cambridge se define como un sistema de computadora o un tipo de programa de computadora que está diseñado para copiar la forma en que opera el cerebro humano; las redes neuronales pueden aprender soluciones a problemas difíciles.

2.2.3 Demanda de medicamentos

Necesidad de comprar productos farmacéuticos con el fin de reestablecer la salud o tener una calidad de vida adecuada, dicha necesidad es originada por la prescripción del personal asistencial competente o por la misma persona solo en el caso que el producto farmacéutico sea de venta libre (Vásquez, et. al., 2013).

2.2.4 Tiempo

Según el diccionario Oxford se define como un período determinado durante el que se realiza una acción o se desarrolla un acontecimiento.

CAPÍTULO III METODOLOGÍA

3.1 ÁMBITO DE ESTUDIO

La presente investigación se realizó en el área de farmacia del Hospital Regional del Cusco.

3.1.1 Localización política

Pais: Perú

Región: Cusco

Departamento: Cusco

Provincia: Cusco

Distrito: Wanchaq

3.1.2 Localización geográfica

Latitud Sur: 13.30° (referencia Meridiano de Greenwich))

Longitud Oeste: 72° (referencia Meridiano de Greenwich)

Altitud: 3350 msnm.

La información antes señalada se obtuvo de la página web de la institución².

3.2 TIPO Y NIVEL DE INVESTIGACIÓN

3.2.1 Tipo de investigación

La presente investigación es netamente aplicada, debido a que se utilizan los conocimientos de la Ciencia de la Estadística con la finalidad de aplicarlas en el

² <http://www.hospitalregionalcusco.gob.pe/descripcion-de-principales-caracteristicas-del-hospital/>

proceso de la adquisición de medicamentos a futuro, con un adecuado pronóstico de la cantidad de medicamentos a requerir por parte del establecimiento de salud.

3.2.2 Nivel de investigación

La presente investigación es de tipo descriptivo, longitudinal y predictivo por los siguientes motivos:

- Descriptivo porque se determina rasgos característicos de la demanda de los medicamentos.
- Longitudinal porque el estudio se realiza entre los años 2013 y 2017
- Predictivo porque se pronostica el comportamiento de la demanda de los medicamentos para el año 2018.

3.3 IDENTIFICACIÓN Y OPERACIONALIZACIÓN DE VARIABLES E INDICADORES

3.3.1 Identificación de variables

- **Demanda de medicamentos**

Adquisición necesaria de productos farmacéuticos por parte de los usuarios de los establecimientos de salud con el fin de recuperar su salud o mantener una calidad de vida; dicha adquisición necesaria es generada por la prescripción del personal asistencial competente o por el mismo usuario cuando el producto farmacéutico es de venta libre (Vásquez, et al., 2013, pág. 151, 153).

- **Tiempo (Año 2018)**

Duración de las cosas sujeta a cambio; magnitud física la cual permite ordenar los sucesos en secuencias es decir estableciendo un pasado, un presente y un futuro (González, 2006, pág. 7). Según el diccionario de la Real Academia

Española es una magnitud física que permite ordenar la secuencia de los sucesos, estableciendo un pasado, un presente y un futuro.

3.3.2 Operacionalización de variables

VARIABLE	DIMENSIÓN	TIPO	ESCALA	CRITERIO	TÉCNICA
Demanda de medicamentos	Número	Cuantitativo	Razón	Cantidad de medicamentos que requiere el paciente	Modelo autoregresivo integrado de media móvil
					Modelo de red neuronal recurrente de memoria a corto y largo plazo

VARIABLE	DIMENSIÓN	TIPO	ESCALA	CRITERIO	TÉCNICA
Tiempo (Año 2018)	Meses	Cuantitativo	Razón	Enero - 2013 a Diciembre - 2017	Registro de la base de datos

3.4 UNIDAD DE ANÁLISIS

Medicamentos requeridos por los pacientes que pertenece a una determinada unidad de servicio del Hospital Regional del Cusco.

3.5 POBLACIÓN DE ESTUDIO

Todos los medicamentos que se usa en el Hospital Regional del Cusco para la atención de sus pacientes.

3.6 TAMAÑO DE LA MUESTRA

Se selecciona datos de los cinco medicamentos con mayor demanda a partir de enero 2013 a diciembre 2017, para tal fin se realiza un muestreo no probabilístico por conveniencia.

3.7 TÉCNICA DE RECOLECCIÓN DE LA INFORMACIÓN

Se recurrió al departamento de Farmacia del Hospital Regional del Cusco, debido a que es ahí donde se registra y centraliza la demanda de los medicamentos requeridos por todos los servicios.

3.8 VALIDACIÓN Y PREDICCIÓN

Para tal fin se actuó de la siguiente manera:

a) Procesamiento estadístico

Primeramente, se extrajo cinco medicamentos con mayor demanda (2013 a 2017), haciendo uso de la estadística descriptiva mediante un grafico circular de porcentajes.

Luego de identificar los cinco medicamentos se determino sus estadígrafos así como también se evaluo si había datos faltantes y datos atípicos.

Después se describe los rasgos característicos respecto a la demanda de los medicamentos (2013-2017) utilizando graficos temporales, funciones de autocorrelación simple y parcial así como también pruebas de estacionariedad y normalidad.

Lo antes indicado se realizo utilizándose el programa Excell versión 10.0 y software libre R.

b) Diseño de modelos

b1) Metodología autorregresivo integrado de media móvil

Siguiendo la metodología Box Jenkins se procedió de la siguiente manera (Gujarati y Porter, pág. 778):

1. Identificación del modelo (selección tentativa de p, d, q)

2. Estimacion de los parámetros del modelo elegido.
3. Examen diagnostico de los residuales
6. Pronóstico

b2) Metodología de las Redes neuronales artificiales

Para abordar el tema en cuestión se toma en cuenta a Lewis (2017, pág. 104) y se procedió de la forma siguiente:

1. Crear una matriz para los datos de entrada (entrenamiento)
3. Especificar los parámetros del modelo
4. Obtener el modelo ingresando los datos de entrada
5. Crear una matriz para los datos a pronosticar
6. Obtener el pronóstico con el modelo ya obtenido
7. Validar el pronóstico

3.9 ANÁLISIS COMPARATIVO PARA DEMOSTRAR LA VERDAD O FALSEDAD DE LAS HIPÓTESIS PLANTEADAS

Para tal fin se utilizó la raíz del error cuadrático medio (RMSE) y el error porcentual absoluto medio (MAPE).

3.10 PROCEDIMIENTO

Se solicitó la base de datos sobre la demanda de medicamentos del año 2013 – 2018 al Área de Almacenamiento y Gestión de la Programación de productos farmacéuticos del Hospital Regional del Cusco.

Después una vez obtenida la base de datos que fue entregada en un archivo Excel, se ordenó bajo el criterio de mayor a menor demanda de medicamentos y se seleccionó los cinco primeros medicamentos.

Finalmente, las demandas de los cinco medicamentos fueron sometidos a un procesamiento con las funciones del software R.

3.11 PROCESAMIENTO DE DATOS

3.11.1 Modelo de serie de tiempo autorregresivo integrado de media móvil

Para abordar el tema referente al modelo ARIMA se tomó como referencia a Hyndman y Athanasopoulos (2018), a continuación, se detalla el procesamiento de los datos.

1. Cargar los datos

Se utiliza las funciones *ts* y *window* de la librería *stats* (software R, 2019a), con los cuales se convierte un vector numérico a un objeto de serie de tiempo y se extrae los datos específicos, respectivamente.

```
datostotal<-ts(data=c(6852,.....,10854),start=c(2013,1), frequency=12)  
datos<-window(datostotal, start=c(2013,1), end=c(2017,12))
```

2. Verificar presencia de datos faltantes

Se verifica visualmente si hay datos faltantes (son datos de cinco años por lo que es manejable en forma visual).

3. Imputar datos faltantes

Se utiliza la función en R *na.interpolation* de la librería *imputeTS* de autoría de Moritz (2018), que por defecto realiza una interpolación lineal toma los datos que se encuentran al lado derecho e izquierdo del dato faltante y calcula la media para la serie de datos.

```
library (imputeTS)  
na.interpolation (datos)
```

4. Detectar datos atípicos

Se utiliza la función *tsoutliers* de la librería *forecast* de autoría de Hyndman et. al. (2019), que emplea la función *supsmu* para series no estacionales y la función *stl* para series estacionales dichas funciones reconocen datos atípicos y estiman datos para reemplazar el dato atípico.

```
library (forecast)
```

```
tsoutliers (datos)
```

5. Hallar estadísticos descriptivos

Se utiliza las funciones *summary* de la librería *base* (software R, 2019b), *window* y *sd* de la librería *stats* (software R, 2019a), con los cuales se calcula la media, la mediana, cuartiles dato máximo y mínimo, extracción de datos de la serie y calcula la desviación estándar respectivamente; el objeto *coefvar* contiene una secuencia lógica para determinar el coeficiente de variación. Todas las funciones y el objeto mencionado anteriormente son introducidos en un bucle *for* para hallar los estadísticos descriptivos de la serie total, así como de las series por años.

a) Para hallar las series por año

```
for (i in 2013:2017)
```

```
{
```

```
  x<-summary(window(datos,start=c(i,1),end=c(i,12)))
```

```
  y<-sd(window(datos,start=c(i,1),end=c(i,12)))
```

```
  z<-coefvar<-((sd(window(datos,start=c(i,1),end=c(i,12))))*100/(mean(window(  
    (datos,start=c(i,1),end=c(i,12))))
```

```
  cat("Año: ");
```

```
  print(i)
```

```
  print(x)
```

```
  cat("Desviación estandar: ");
```

```
  print(y)
```

```
  cat("Coeficiente de variación: ");
```

```
  print(z)
```

```
}
```

b) Para hallar la serie total

```
summary(datos)
```

```
sd(datos)
```

```
coefvar<-((sd(datos))*100/(mean(datos)))
```

```
coefvar
```

6. Obtener el gráfico de la serie

Se emplea la función *autoplot* de la librería *ggplot2*, de autoría de Wickham et. al. (2018), con el cual se obtiene un gráfico y con los argumentos *ggtitle* coloca el título, *xlab* pone etiqueta al eje x, *ylab* pone etiqueta al eje y.

```
autoplot(datos)+ggtitle("")+xlab("Año")+ylab("Unidades")
```

7. Obtener el gráfico de la función de autocorrelación de la serie

Se utiliza la función *ggAcf* de la librería *forecast* de autoría de Hyndman et. al. (2019), con el cual se obtiene un gráfico de la función de autocorrelación de la serie, además con los argumentos *ggtitle* coloca el título, *xlab* pone etiqueta al eje x, *ylab* pone etiqueta al eje y.

```
ggAcf(datos, lag=48)+ ylab("Autocorrelación")+ xlab("Retraso")+ggtitle("")
```

8. Obtener el diagrama de caja de la serie

Se utiliza las funciones *ggplot*, *geom_boxplot* de la librería *ggplot2*, de autoría de Wickham et. al. (2018), con los cuales se traza un gráfico y se establece de que tipo, en este caso un diagrama de caja y con los argumentos *xlab* pone etiqueta al eje x, *ylab* pone etiqueta al eje y, *scale_x_discrete(breaks=NULL)* pone la escala en el eje x, y *coord_flip* muestra el diagrama de caja en forma horizontal.

```
ggplot(datos)+geom_boxplot(mapping=aes("var",datos))+xlab("")+  
ylab("Unidades")+scale_x_discrete(breaks=NULL)+coord_flip()
```

9. Obtener el diagrama de caja de la serie por meses

Se emplea las funciones *cycle* de la librería *stats* (software R, 2019a), *ggplot* y *geom_boxplot* de la librería *ggplot2*, de autoría de Wickham et. al. (2018), con los cuales se obtiene las posiciones en el ciclo de cada observación, traza un gráfico de diagrama de caja respectivamente, y con los argumentos *factor* se señala que los datos se presenten en forma ordenada, *labs* se pone etiquetas al eje x y y, *scale_x_discrete* pone la escala al eje x y *labels* cambia la escala con etiquetas.

```

tsboxplot<- function(x = x) {
season <- cycle(x)
season.factor <- factor(season)
ggplot() + geom_boxplot(mapping = aes(x = season.factor, y = x)) +
labs(x = "Mes", y = "Unidades")+
scale_x_discrete(breaks=c("1", "2", "3", "4", "5", "6", "7", "8", "9", "10", "11", "12"),
labels=c("Enero", "Febrero", "Marzo", "Abril", "Mayo", "Junio", "Julio", "Agosto",
"Setiembre", "Octubre", "Noviembre", "Diciembre"))
}

```

10. Descomponer la serie

Se utiliza la función *seas* de la librería *seasonal* de autoría de Sax y Eddelbuette (2018), con el cual se extrae los componentes de la serie por el método x11 y con los argumentos *autoplot* se traza un gráfico, *ggtitle* pone el título principal y *xlab* pone etiqueta al eje x.

```

datos %>% seas(x11="")->fit
autoplot(fit) + ggtitle("")+ xlab("Año")

```

11. Evaluar estacionariedad, tendencia y estacionalidad de la serie

a) Estacionariedad, se utiliza la función *ur.kpss* (prueba KPSS) de la librería *urca* de autoría de Pfaff, Zivot y Stigler (2016).

```

datos %>% ur.kpss() %>% summary()

```

b) Tendencia, se utiliza la función *mk.test* (prueba de Mann-Kendall) de la librería *trend*, de autoría de Pohlert, T. (2018).

```

mk.test(datos)

```

c) Estacionalidad, se emplean la funciones *kruskal.test* (prueba de Kruskal-Wallis) y *window* de la librería *stats* (software R, 2019a).

```

a2013<-window(datos,start=c(2013,1),end=c(2013,12))

```

```
b2014<-window(datos,start=c(2014,1),end=c(2014,12))
c2015<-window(datos,start=c(2015,1),end=c(2015,12))
d2016<-window(datos,start=c(2016,1),end=c(2016,12))
e2017<-window(datos,start=c(2017,1),end=c(2017,12))
```

```
kruskal.test(list(a2013,b2014,c2015,d2016,e2017))
```

```
kruskal.test(list(a2013,b2014))
```

```
kruskal.test(list(a2013,c2015))
```

```
kruskal.test(list(a2013,d2016))
```

```
kruskal.test(list(a2013,e2017))
```

```
kruskal.test(list(b2014,c2015))
```

```
kruskal.test(list(b2014,d2016))
```

```
kruskal.test(list(b2014,e2017))
```

```
kruskal.test(list(c2015,d2016))
```

```
kruskal.test(list(c2015,e2017))
```

```
kruskal.test(list(d2016,e2017))
```

12. Evaluar normalidad

Se utiliza varias pruebas para evaluar la normalidad entre ellas:

a) La funciones de *shapiro.test* de la librería stats (software R, 2019a).

```
library(stats)
```

```
shapiro.test(datos)
```

13. Gráficos sobre la normalidad

Se utilizan las funciones *par*, *hist*, *lines* y *boxplot* de la librería *graphics* (Software R, 2019c); y *qqnorm* y *qqline* de la librería *stats* (Software R, 2019a), con los cuales se realiza lo siguiente *par(mfrow=c(1,3))* para especificar que en una fila se tiene tres gráficos, *hist()* para obtener un gráfico de histograma, *lines()* para obtener y sobreponer una distribución de campana de Gauss de color rojo sobre el histograma, *qqnorm()* para obtener un gráfico de distribución de normalidad de los cuantiles, *qqline()* para obtener y sobreponer una línea ascendente recta

de color rojo sobre la distribución de la normalidad y *boxplot()* para obtener un gráfico de diagrama de caja en base a los cuartiles e identificar sesgo.

```
par(mfrow= c(1,3))
hist(datos, probability =T, main="", ylab='Densidad', xlab='Unidades')
lines(density(datos), col="red")
qqnorm(datos, ylab='Cuantiles de muestra',xlab='Cuantiles teóricos', main=' ')
qqline(datos, col = 2)
boxplot(datos, horizontal=T, ylab=' ', xlab='Unidades')
```

14. Transformar los datos si no hay normalidad

Se utiliza la función *BoxCox.lambda* de la librería *forecast* de autoría de Hyndman et. al. (2019) que por defecto utiliza el método Guerrero (1993), dicha función obtiene subseries de la serie principal luego se realiza una estimación de la media y varianza de cada una de ellas; el coeficiente lambda de la transformación potencia es estimado mediante mínimos cuadrados.

```
library (forecast)
lambda<-BoxCox.lambda(datos)
```

15. Diferenciar si existe tendencia

Para tal fin se emplea la función *diff* de la librería base (Software R, 2019b), que consiste en restar a cada dato el dato anterior con lo cual se elimina la tendencia.

```
datbcox<-(datos^lambda)
diff(datbcox)
```

16. Gráficos de las funciones de autocorrelación simple y parcial

Se utilizan las funciones *ggAcf* y *ggPacf* de la librería *forecast* de autoría de Hyndman et. al. (2019), con las cuales se calcula las funciones de autocorrelacion simple y parcial; y para obtener el gráfico se hace uso de la función *grid.arrange* de la librería *gridExtra* de autoría de Auguie y Antonov (2017); de la interpretación de estas gráficas surgen los modelos ARIMA candidatos a proponer.

```

library(gridExtra)
fas<-ggAcf((diff(datbcox)),lag=48)+ylab("Autocorrelación simple")+
xlab("Retraso")+ggtitle("")
fap<-ggPacf((diff(datbcox)),lag=48)+ylab("Autocorrelación parcial")+
xlab("Retraso")+ggtitle("")
grid.arrange(fas,fap,ncol=2)

```

17. Obtener el mejor modelo

Se utiliza la función *Arima* de la librería *forecast* de autoría de Hyndman et. al. (2019), mediante la cual se obtiene el modelo indicando el orden de las componentes del modelo (Hyndman y Khandakar, 2008).

```
library(forecast)
```

```
###ARIMA regular
```

```
fit1<-Arima(datos, order=c(1,1,1)) ##ARIMA regular
```

```
###ARIMA con dos componentes (regular y estacional)
```

```
fit1<-Arima(datos, order=c(1,1,1), seasonal=c(0,0,1))
```

```
###ARIMA con dos componentes (regular y estacional) y el coeficiente lambda
```

```
fit1<-Arima(datos, order=c(1,1,1), seasonal=c(0,0,1), lambda=lambda)
```

18. Pronóstico para el año 2018

Se utiliza la función *forecast* de la librería *forecast* de autoría de Hyndman et. al. (2019) con la cual se obtiene un pronóstico específico en base al modelo ya obtenido, en este caso se argumentó un pronóstico para el año 2018 con un intervalo de confianza de 95%;

```
fc1<-forecast(fit1, h=12,level=95)
```

```
summary(fc1)
```

19. Cálculo de los errores del pronóstico 2018

Se utiliza las funciones *mape* y *rmse* de la librería *Metrics* de autoria de Hammer, Frasco y LeDell (2018), con los cuales se obtiene el error porcentual absoluto medio y el error cuadrático medio respectivamente.

```
library(Metrics)
```

```
fc1$mean ## se extrae la media del pronóstico
```

```
año2018<-window(datostotal,start=c(2018,1),end=c(2018,12))
```

```
mape((as.vector(año2018)),(as.vector(fc1$mean)))
```

```
rmse((as.vector(año2018)),(as.vector(fc1$mean)))
```

20. Gráficos de: datos originales (2013-2018), entrenamiento (2013-2017) y pronóstico (2018).

Se utiliza las funciones de *autoplot* y *autolayer* de la librería *ggplot2*, de autoría de Wickham et. al. (2018), con los cuales se obtiene un gráfico de la serie original, además permite seguir añadiendo gráficos sobre este y con los argumentos *ggtitle* coloca el título, *xlab* pone etiqueta al eje x, *ylab* pone etiqueta al eje y, y *guides* coloca el título a la leyenda.

```
library(ggplot2)
```

```
autoplot(fc1,series="Pronóstico")+
```

```
autolayer(año2018,series="Pronóstico MAPE (4.50%")+
```

```
autolayer(datostotal,series="Datos 2013-2018")+
```

```
autolayer((fitted(fc1)),series="Entrenamiento MAPE(5.11%")+
```

```
ggtitle(")+xlab('Año')+ylab('Unidades')+
```

```
guides(colour=guide_legend(title='Series'))
```

21. Evaluación de los residuales del modelo

Se utiliza la función *checkresiduals* de la librería *forecast* de autoría de Hyndman et. al. (2019), con la cual se verifica si los residuos del modelo de serie de tiempo tienen un comportamiento de ruido blanco y la independencia, básicamente con dicha función se obtiene tres gráficos (serie de tiempo, función de

autocorrelación e histograma) y la prueba de Ljung-Box para evaluar la independencia.

```
checkresiduals(fit1)
```

```
summary(fit1)
```

22. Normalidad de los residuales

Se utiliza las mismas funciones indicadas en el punto 12).

3.11.2 Modelo de red neuronal recurrente de memoria a corto y largo plazo

Para tratar el tema concerniente al modelo de la red neuronal recurrente de memoria a corto y largo plazo (RNNLSTM) se tomó como referencia a Lewis (2017), cabe precisar que para explicar como se realiza el procesamiento de los datos se toma como ejemplo a la demanda del Cloruro de sodio a continuación se indica los pasos que se siguieron.

1. Preparación y creación de los datos de entrada.

Se utilizan las funciones *log* y *cbind* de la librería *base* (software R, 2019b), *Lag* de la librería *quantmod* de autoría de Ryan et al. (2019) y *as.zoo* de la librería *zoo* de autoría de Zeileis et al. (2019); a continuación, se detalla al respecto. Los datos fueron llevados a una escala logarítmica con el fin de reducir los números, luego se utiliza la función *as.zoo* para coaccionar a los datos (vector, matriz), luego se emplea la función *Lag* con el fin de crear 12 retrasos, después se crea el vector *x* juntando todos los retrasos con la función *cbind* (Anexo N° 05).

```
data<-log(data)
```

```
data
```

```
library(quantmod)
```

```
library(zoo)
```

```
data<-as.zoo(data)
```

```
x1<-Lag(data, k = 1)
```

```
x2<-Lag(data, k = 2)
```

```

x3<-Lag(data, k = 3)
x4<-Lag(data, k = 4)
x5<-Lag(data, k = 5)
x6<-Lag(data, k = 6)
x7<-Lag(data, k = 7)
x8<-Lag(data, k = 8)
x9<-Lag(data, k = 9)
x10<-Lag(data, k = 10)
x11<-Lag(data, k = 11)
x12<-Lag(data, k = 12)
x<-cbind(x1,x2,x3,x4,x5,x6,x7,x8,x9,x10,x11,x12, data)

```

2. Extracción de los datos faltantes.

Se extrae los datos faltantes generados al crear los datos de entrada y se maneja el objeto *x* como una matriz, seguidamente los datos se pusieron en una escala entre cero y uno, es decir se normaliza mediante la función *range_data* definida de la siguiente manera *range_data<-function(x){(x-min(x))/(max(x)- min(x))}*.

```

head(round(x,3))
x <- x [-(1:12), ]
x<-data.matrix(x)
x
min_data<-min( x )
min_data
max_data<-max( x )
max_data
x <-range_data ( x )

```

3. Datos para el entrenamiento y la prueba.

Se utiliza la función *as.matrix* de la librería *base* (software R, 2019b), con dicha función se crean matrices y se especifica los datos de entrada para el entrenamiento en este caso con los 48 primeros datos.

```

x1<-as.matrix ( x [, 1 ] )

```

```

x2<-as.matrix ( x [, 2 ] )
x3<-as.matrix ( x [, 3 ] )
x4<-as.matrix ( x [, 4 ] )
x5<-as.matrix ( x [, 5 ] )
x6<-as.matrix ( x [, 6 ] )
x7<-as.matrix ( x [, 7 ] )
x8<-as.matrix ( x [, 8 ] )
x9<-as.matrix ( x [, 9 ] )
x10<-as.matrix ( x [, 10 ] )
x11<-as.matrix ( x [, 11 ] )
x12<-as.matrix ( x [, 12 ] )
y<-as.matrix ( x [, 13 ] )
y
n_train <- 48
y_train<-as.matrix ( y [ 1 : n_train ] )
x1_train<-as.matrix ( t ( x1 [ 1 : n_train , ] ) )
x2_train<-as.matrix ( t ( x2 [ 1 : n_train , ] ) )
x3_train<-as.matrix ( t ( x3 [ 1 : n_train , ] ) )
x4_train<-as.matrix ( t ( x4 [ 1 : n_train , ] ) )
x5_train<-as.matrix ( t ( x5 [ 1 : n_train , ] ) )
x6_train<-as.matrix ( t ( x6 [ 1 : n_train , ] ) )
x7_train<-as.matrix ( t ( x7 [ 1 : n_train , ] ) )
x8_train<-as.matrix ( t ( x8 [ 1 : n_train , ] ) )
x9_train<-as.matrix ( t ( x9 [ 1 : n_train , ] ) )
x10_train<-as.matrix ( t ( x10 [ 1 : n_train , ] ) )
x11_train<-as.matrix ( t ( x11 [ 1 : n_train , ] ) )
x12_train<-as.matrix ( t ( x12 [ 1 : n_train , ] ) )

```

4. Agrupación de las matrices (datos de entrada)

Se utiliza la función `array` de la librería `base` (software R, 2019b), con la cual se juntan las matrices en el objeto `x_train`, el cual contiene los datos de entrada.

```

x_train <-array(c(x1_train,x2_train,x3_train,x4_train,x5_train,x6_train,
                x7_train,x8_train,x9_train,x10_train,x11_train,x12_train),
              dim=c(dim( x1_train ),12))

```

dim(x_train)

5. Especificación del modelo

Se utiliza la función *trainr* de la librería *rnn*, de autoría de Quast y Fichou (2019), donde se especifica el modelo el cual tiene un grado de aprendizaje de 0.05, con 7 nodos, 1000 épocas, el tipo de la red neuronal es LSTM y la función de activación es la tangente hiperbólica (Tabla N° 07). Asimismo se utiliza la función *set.seed* de la librería *base* (software R, 2019b), con la cual se genera números aleatorios con el fin de evitar la varibilidad de los datos.

library(rnn)

set.seed(2018) #se utiliza una semilla para evitar la variabilidad de los datos

model1 <- trainr(Y=t(y_train) ,

X = x_train ,

learningrate = 0.05,

hidden_dim = 7,

numepochs = 1000 ,

network_type = "lstm" ,

sigmoid = "tanh")

6. Pronóstico para los años 2014 a 2017

Se obtiene los datos utilizando el modelo de la red neuronal obtenida y la función *predictr* de la librería *rnn* de autoría de Quast y Fichou (2019), con el cual se obtiene los datos de salida de la red neuronal.

pred1_train <- t(predictr(model1, x_train))

pred1_train

7. Pronóstico para el año 2018

Se utiliza las dimensiones del objeto *x1* hasta *x12* desde las filas 49 hasta 60 luego son colocados en los objetos *x1_test* hasta *x12_test*, los mencionados objetos se juntan haciendo uso de la función *array* en el objeto *x_test*. Asimismo, se extrae los datos originales de la demanda del medicamento para el año 2018

del objeto *y* y es colocado en el objeto *y_test*. Finalmente, el pronóstico se realiza utilizando la función *predictr*, y se procede igual que en el punto 6).

```
pronost<-48
```

```
x1_test<-as.matrix (t(x1[( pronost+1): 60 , ] ) )  
x2_test<-as.matrix (t(x2[( pronost+1) : 60 , ] ) )  
x3_test<-as.matrix (t(x3[( pronost+1) : 60 , ] ) )  
x4_test<-as.matrix (t(x4[( pronost+1) : 60 , ] ) )  
x5_test<-as.matrix (t(x5[( pronost+1) : 60 , ] ) )  
x6_test<-as.matrix (t(x6[( pronost+1) : 60 , ] ) )  
x7_test<-as.matrix (t(x7[( pronost+1) : 60 , ] ) )  
x8_test<-as.matrix (t(x8[( pronost+1) : 60 , ] ) )  
x9_test<-as.matrix (t(x9[( pronost+1) : 60 , ] ) )  
x10_test<-as.matrix (t(x10[( pronost+1) : 60 , ] ) )  
x11_test<-as.matrix (t(x11[( pronost+1) : 60 , ] ) )  
x12_test<-as.matrix (t(x12[( pronost+1) : 60 , ] ) )
```

```
y_test<-as.matrix ( y [ 49 : 60 ] )
```

```
y_test
```

```
x_test <- array(c(x1_test,x2_test,x3_test,x4_test,x5_test,x6_test,  
                x7_test,x8_test,x9_test,x10_test,x11_test,x12_test),  
              dim=c(dim(x1_test),12))
```

```
pred1_test <-t(predictr(model1,x_test))
```

8. Retirando la escala.

Se retira la escala a los datos de los objetos *pred1_train*, *pred1_actual* y *y_actual*, mediante la función *unscale_data* definida de la siguiente forma *unscale_data<-function(x ,max_x,min_x){x *(max_x-min_x)+min_x}*, seguidamente se realiza la operación exponencial y finalmente se convierte al formato de series de tiempo con la función *ts* de la librería *stats* (software R, 2019a).

```
unscale_data<-function(x ,max_x,min_x)  
  {x *(max_x-min_x)+min_x}
```

```

pred1_train<-unscale_data(pred1_train,max_data,min_data )
pred1_train<-exp(pred1_train)
pred1_train<-ts(pred1_train,start=c(2014,1),frequency=12)

```

```

pred1_actual<-unscale_data ( pred1_test,max_data,min_data )
pred1_actual<-exp(pred1_actual)
pred1_actual<-ts(matrix(pred1_actual),end=c(2018,12),frequency=12)

```

```

y_actual<-unscale_data(y_test,max_data,min_data)
y_actual<-exp(y_actual)
y_actual<-ts(matrix(y_actual),end=c(2018 ,12),frequency=12)

```

9. Inspección visual de los datos originales y el pronóstico año 2018

Se utiliza la función *cbind* para unir los datos originales y el pronóstico de la serie para el año 2018, después se coloca etiquetas a las columnas mediante la función *colnames* de la librería *base* (software R, 2019b) y se presenta los resultados en el objeto *result_all* .

```

result_all <-cbind(y_actual,round(pred1_actual,2))
colnames(result_all)<-c("Actual","RNNLSTM")
result_all

```

10. Medición de los errores.

Se calcula y se evalúa los errores mediante las funciones *mape* y *rmse* de la librería *Metrics* de autoría de Hammer, Frasco y LeDell (2018), con las cuales se obtiene el error porcentual absoluto medio y la raíz del error cuadrático medio respectivamente. Asimismo, con la función *window* de la librería *stats* (software R, 2019a), se extrae datos del objeto *datos* a partir de enero del año 2014 a diciembre del año 2017.

```

library(Metrics)
datos<-window(datos,start=c(2014,1),end=c(2017,12))
mape(datos,pred1_train)

```

```
rmse(datos,pred1_train)
```

```
mape(y_actual,pred1_actual)
```

```
rmse(y_actual,pred1_actual)
```

11. Gráficos de: datos originales (2013-2018), entrenamiento (2014-2017) y pronóstico (2018).

Se utiliza las funciones de *autoplot* y *autolayer* de la librería *ggplot2*, de autoría de Wickham et. al. (2018), con los cuales se obtiene un gráfico de la serie original, además permite seguir añadiendo gráficos sobre este y con los argumentos *ggtitle* coloca el título, *xlab* pone etiqueta al eje x, *ylab* pone etiqueta al eje y, y *guides* coloca el título a la leyenda.

```
autoplot(datos,series="Datos 2013 - 2018")+
```

```
autolayer(pred1_train,series="Entrenamiento MAPE (12.76%")+
```

```
autolayer((ts(pred1_actual,start=c(2018,1),frequency=12)),series="Pronóstico  
2018 MAPE (23.22%))+xlab('Año')+ylab('Unidades')+
```

```
guides(colour=guide_legend(title='Series'))
```

CAPÍTULO IV RESULTADOS

IDENTIFICACIÓN DE LOS MEDICAMENTOS CON MAYOR DEMANDA

Tabla 1
Demanda de medicamentos del año 2013 al 2017

Nro.	Medicamentos	Cantidad	Porcentaje
1	METAMIZOL SODICO - 1 G - 2 ML	639490	4.18
2	SEVOFLURANO 100ML/100ML 250ML	584005	3.82
3	PARACETAMOL 500 MG - TAB	524099	3.43
4	SODIO CLORU 900MG/100ML(0.9%)1L	475537	3.11
5	OMEPRAZOL - 20 MG - TAB	429440	2.81
6	IBUPROFENO - 400 MG - TAB	414213	2.71
7	VALPROATO SOD. 500 MG - TAB	401706	2.63
8	ENALAPRIL - 10 MG - TAB	397430	2.6
9	METFORMINA CLORHIDRATO 850MG TAB	392683	2.57
10	NAPROXENO 550 MG - TAB	391460	2.56
11	CIPROFLOXACINO 500 MG - TAB	359946	2.35
12	ACIDO ALFA LIPOICO 600MG -TAB	339533	2.22
13	RANITIDINA 25 MG/ML- 2ML - INY	309664	2.02
14	ATORVASTATINA 40MG - TAB	242375	1.58
15	FENITOINA 100 MG - TAB	239275	1.56
16	DEXAMETASONA 4MG/2ML - INY	226725	1.48
17	CEFUROXIMA 250MG/5ML X 50ML	214373	1.4
18	LAMIVUDINA 150 MG - TAB	198706	1.3
19	CEFAZOLINA SODICA - 1 G - INY	184594	1.21
20	CARBAMAZEPINA 100MG/5ML 120ML	180264	1.18
21	LOSARTAN 50MG TAB	175941	1.15
22	CLONAZEPAN 0,5MG TAB	169341	1.11
23	CLINDAMICINA 600MG/4 ML - INY	160532	1.05
24	LOPINAVIR + RITONAVIR 250MG TAB	158176	1.03
25	LAMIVUDINA+ZIDOVUDINA 300+150MG TAB	153714	1
26 - 457	OTROS	7333350	47.94
Total		15296572	100

Fuente. Departamento de Farmacia del Hospital Regional del Cusco (Anexo N° 01)

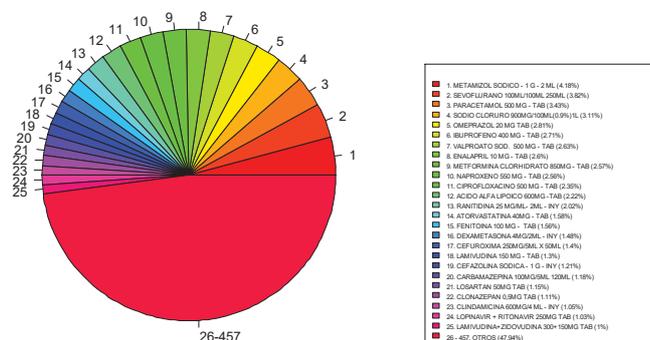


Figura 11. Porcentaje de la demanda de medicamentos del año 2013 al 2017

INTERPRETACIÓN

De 457 medicamentos que representan el 100%, los que tuvieron mayor demanda durante los años 2013 al 2017 son el metamizol sódico 1 g – 2ml, el sevoflurano 100ml/100ml 250 ml, el paracetamol 500 mg tab, el cloruro de sodio, 900 mg/100ml (0.9%)1L, el omeprazol 20 mg tab y el ibuprofeno 400 mg tab con un porcentaje de 4.18%, 3.82%, 3.43%, 3.11%, 2.81% y 2.71% respetivamente.

Por otro lado, los medicamentos que tuvieron una demanda menor al 1% (otros) representan el 47.94%.

Cabe precisar que el quinto medicamento con mayor demanda es el omeprazol sin embargo se descarto este debido a que presenta varios datos atípicos, motivo por el cual se trabajo con el siguiente medicamento con mayor demanda que es el ibuprofeno.

DEMANDA DE CLORURO DE SODIO (DCS) 900 MG/1000 ML INYECTABLE

Descripción del comportamiento temporal de la DCS

Tabla 2

Estadísticos descriptivos de la DCS por año

Año	2013	2014	2015	2016	2017
n (meses)	12	12	12	12	12
Media	7091	7234	7714	8244	9346
Desviación estandar	357.1068	380.4365	585.3706	560.0799	507.0064
Quartil 1	6837	6974	7520	7812	8962
Mediana	7160	7181	7860	8538	9394
Quartil 3	7376	7433	8133	8656	9507
Coefficiente de variación	5.0362	5.2591	7.5887	6.7941	5.4248

Fuente. Departamento de Farmacia del Hospital Regional del Cusco (Anexo N° 01)

INTERPRETACIÓN

Se aprecia que no hay mucha variación en cuanto a las medidas de tendencia central como la media y la mediana entre los años 2013 y 2015, pero si una variación ascendente entre los años 2016 y 2017. Por otro lado, el coeficiente de variación indica que los datos dentro de cada año no presentan mucha variación respecto a la media, dicho coeficiente de 2013 al 2017 no sobrepasan el 7.59%.

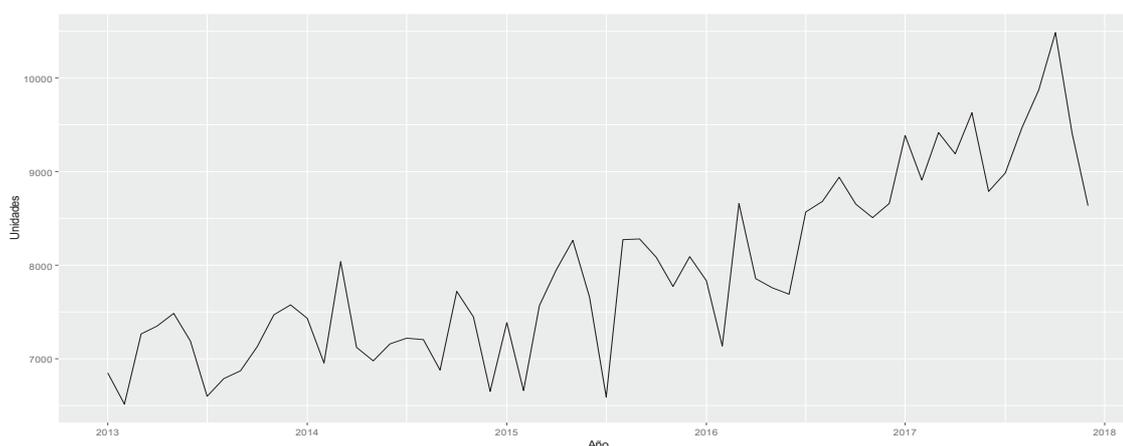


Figura 12. DCS mensual de enero 2013 a diciembre de 2017

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se percibe que la serie es no estacionaria, presenta tendencia positiva y una probable estacionalidad.

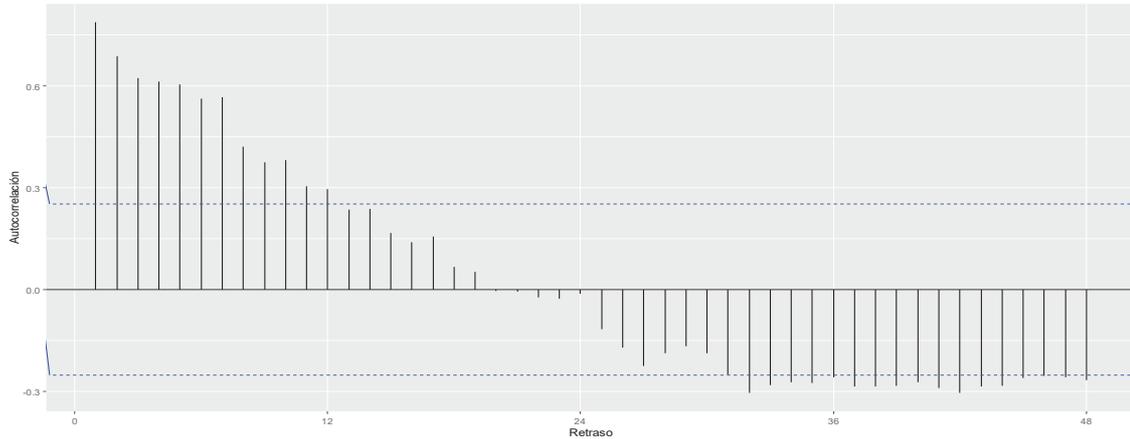


Figura 13. Función de autocorrelación de la DCS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la autocorrelación una disminución lenta, forma festoneada irregular y picos significativos lo cual indica que la serie presenta tendencia positiva, no estacionalidad y no estacionariedad, respectivamente.

Tabla 3

Estadísticos descriptivos de toda la serie de la DCS

Año	2013 - 2017
n (meses)	60
Media	7926
Desviación estandar	948.7621
Quartil 1	7183
Mediana	7740
Quartil 3	8653
Coefficiente de variación	11.97083

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la serie entre los años 2013 al 2017 una media, mediana y coeficiente de variación de 7926, 7740 y 11.97% respectivamente.

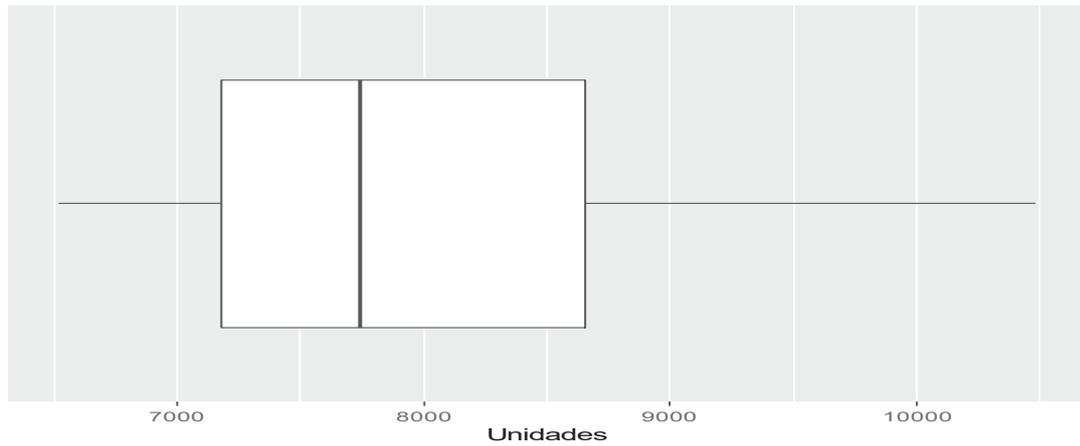


Figura 14. Diagrama de caja de la DCS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa en la serie una distribución asimétrica con un sesgo a la izquierda y ausencia de datos atípicos.

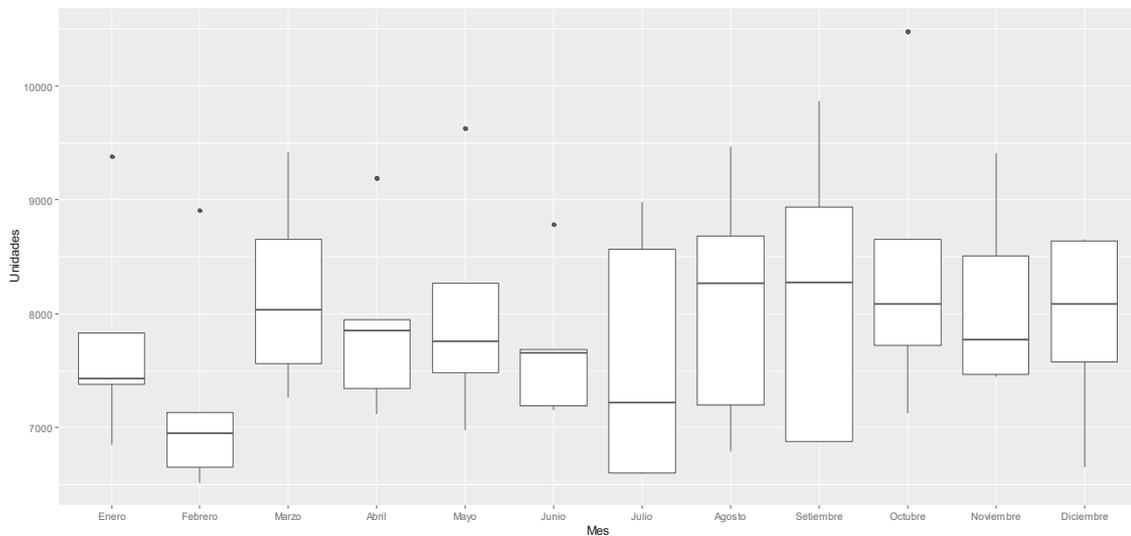


Figura 15. Diagrama de caja por mes de la DCS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la serie datos atípicos en los meses de enero, febrero, abril, mayo, junio y octubre; una distribución simétrica débil en los meses de marzo y diciembre lo cual indica estacionalidad, por otro lado, en los meses de julio y setiembre se tiene una mayor dispersión de la demanda, el mismo que se confirma con el rango intercuartílico (altura de las cajas). Asimismo, una menor

y mayor demanda del medicamento durante los meses de febrero y setiembre, respectivamente.

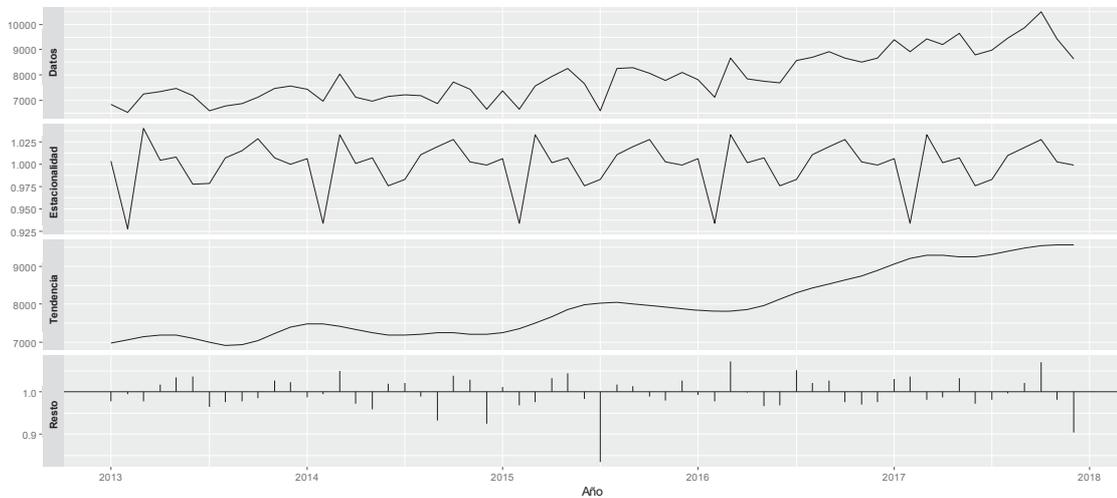


Figura 16 Descomposición de la DCS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

En la parte superior la serie original, luego la componente de estacionalidad donde se aprecia una aparente estacionalidad, después la componente de la tendencia donde se distingue una tendencia positiva y finalmente el componente del resto.

Tabla 4
Pruebas de estacionariedad, tendencia, estacionalidad y normalidad de la DCS

KPSS		Mann-Kendall	Kruskal-Wallis		Shapiro - Wilk
p	valor crítico 5%	p	Año	p	p
			2013 a 2018	1.846E-08	
			2013 - 2014	0.4884	
			2013 - 2015	0.006657	
			2013 - 2016	0.0001387	
			2013 - 2017	3.226E-05	
1.3883	0.463	1.752E-13	2014 - 2015	0.02434	0.02439
			2014 - 2016	0.000532	
			2014 - 2017	3.226E-05	
			2015 - 2016	0.03767	
			2015 - 2017	3.226E-05	
			2016 - 2017	0.0001749	

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se percibe que la serie presenta no estacionariedad, tendencia y no estacionalidad lo cual se verifica con las pruebas de KPSS ($p > 0.463$), Mann-Kendall ($p < 0.05$) y Kruskal-Wallis ($p < 0.05$), respectivamente.

En cuanto a la estacionalidad si bien la prueba de Kruskal-Wallis indica que la serie para los años 2013 a 2017 no presenta estacionalidad, pero si presenta estacionalidad para los años 2013 y 2014.

Respecto a la prueba de normalidad de Shapiro Wilk se observa que la prueba de normalidad de Shapiro-Wilk es significativa ($p < 0.05$) señal de heterogeneidad de la serie.

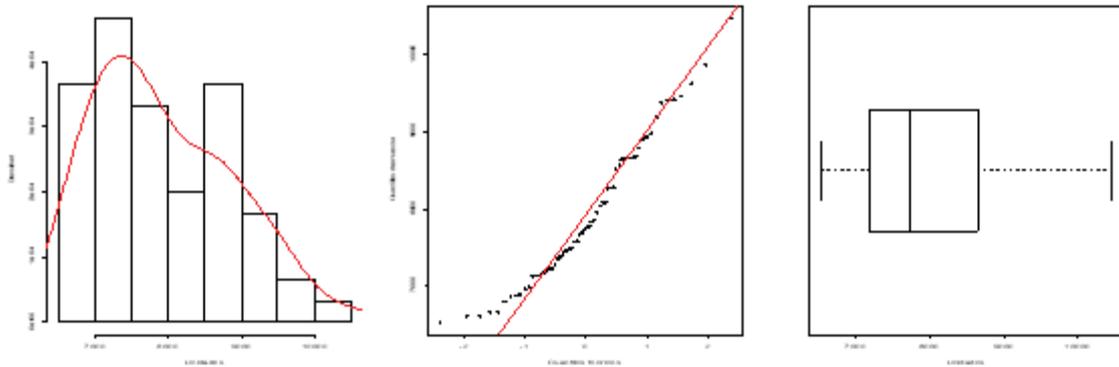


Figura 17. Evaluación de la normalidad de la DCS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Los gráficos señalan que los datos de la serie no siguen un perfil de campana de Gauss, lo cual se verifica en el gráfico del histograma una distribución asimétrica con un sesgo a la izquierda, el gráfico de la distribución de probabilidad línea recta se verifica que no sigue una línea recta y el diagrama de la caja es asimétrico con un sesgo a la izquierda.

Identificación del modelo (selección tentativa de p,d,q)

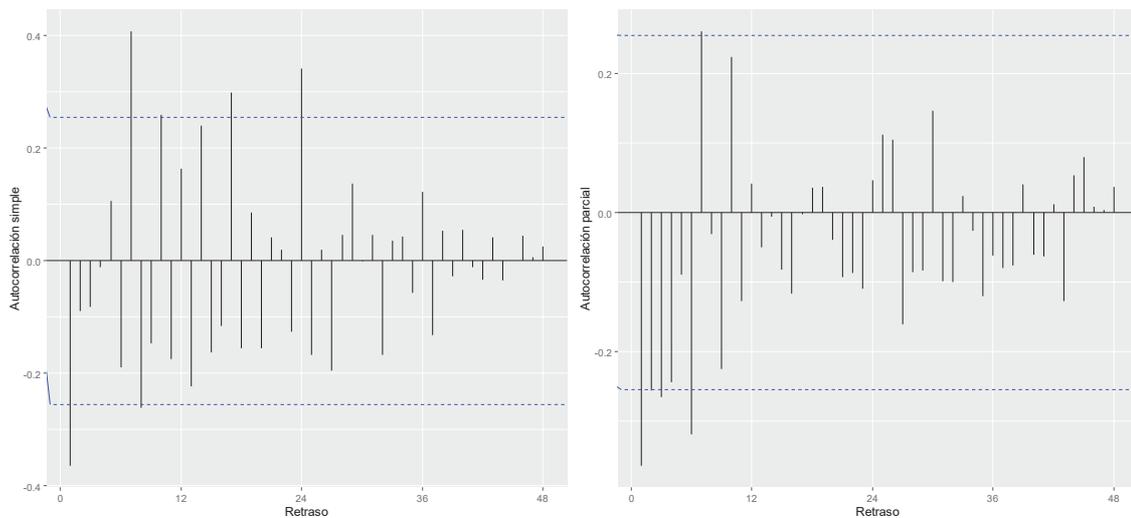


Figura 18. Función de autocorrelación simple y parcial de la serie con transformación BoxCox y diferenciación de la DCS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la función de autocorrelación simple el primer rezago por defecto es siempre altamente significativo, por otro lado, se tiene tres picos significativos en la banda superior y un pico ligeramente saliendo en la banda inferior, asimismo se observa que los rezagos no siguen un patrón de descenso; respecto a la función de autocorrelación parcial se tiene que el primer rezago por defecto es siempre altamente significativo, a su vez se observa dos picos el primer pico ligeramente significativo y el segundo altamente significativo en la banda inferior y un pico ligeramente significativo en la banda superior, además se percibe que los rezagos no siguen un patrón de descenso. Del gráfico se tiene indicios de proponer un modelo con componentes AR(1) y MA(3).

Tabla 5
Identificación del modelo para la DCS

Modelos	Medidas de ajuste					Pruebas	
	AIC	Entrenamiento		Pronostico		Ljung-Box	Shapiro-Wilk
		RMSE	MAPE	RMSE	MAPE		
ARIMA(3,1,3)	-391.12	465.6491	4.712674	630.3925	4.687721	0.004421	0.9671
ARIMA(2,1,3)	-392.99	465.4832	4.689311	631.1726	4.68868	0.01207	0.9582
ARIMA(1,1,3)	-389.27	514.2756	5.24489	725.6501	5.663716	0.02025	0.4442
ARIMA(1,1,2)	-391.27	514.3181	5.246921	724.9312	5.659444	0.0345	0.4275
ARIMA(1,1,1)	-392.71	515.4788	5.212838	735.9054	5.71353	0.02876	0.5552
ARIMA(0,1,1)	-398.65	485.23	5.110848	477.6973	4.501506	0.1825	0.07479

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que de los diferentes modelos propuestos el mejor modelo es el ARIMA(0,1,1) debido a que presenta el menor AIC (-398.65), los residuales del modelo cumplen los requisitos de normalidad e independencia, asimismo el MAPE del pronóstico es el más bajo.

Estimación de los parámetros del modelo elegido

Tabla 6
Modelo ARIMA para la DCS

Modelo	Parámetros						Medición del error			
	Coeficiente	Estimado	Error estándar	AIC	BIC	lambda	Entrenamiento		Pronóstico	
							RMSE	MAPE	RMSE	MAPE
ARIMA(0,1,1)	ma1	-0.8432	0.0729	-398.65	-392.42	-0.2316016	485.23	5.110848	477.6972	4.501505
	constante	0.0007	0.0002							

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que el modelo para la DCS, es un ARIMA compuesto por una parte regular la cual contiene un componente diferencial y un componente MA de orden 1.

En cuanto a los parámetros se tiene una ma1 con -0.8432, una constante con 0.0007, un AIC de -398.65, un BIC de -392.42 y una lambda de -0.2316016.

En relación a la medición del error se tiene que en el entrenamiento y pronóstico un RMSE de 485.23 y 477.6972 respectivamente y un MAPE de 5.11 y 4.50 respectivamente.

Respecto al modelo estimado se tiene un modelo ARIMA (0,1,1) cuya ecuación general es $Y_t = c + Y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$ (Anexo N° 02) al reemplazar los parámetros estimados $c = 0.0007$ y $\theta_1 = -0.8432$ se obtiene la siguiente ecuación del modelo ARIMA como sigue $Y_t = 0.0007 + Y_{t-1} + \varepsilon_t - 0.8432\varepsilon_{t-1}$.

Validación del modelo

Tabla 7

Pruebas de normalidad de los residuos y comportamiento de ruido blanco del modelo ARIMA (0,1,1) de la DCS

Prueba	p
Shapiro-Wilk	0.07479
Ljung-Box	0.1825

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se percibe que las pruebas de normalidad de Shapiro-Wilk no es significativa ($p > 0.05$), con lo cual se evidencia la normalidad de los residuos. En cuanto a la prueba de Ljung-Box se aprecia que no es significativa ($p > 0.05$), señal de que los residuos son independientes, es decir tienen un comportamiento de ruido blanco.

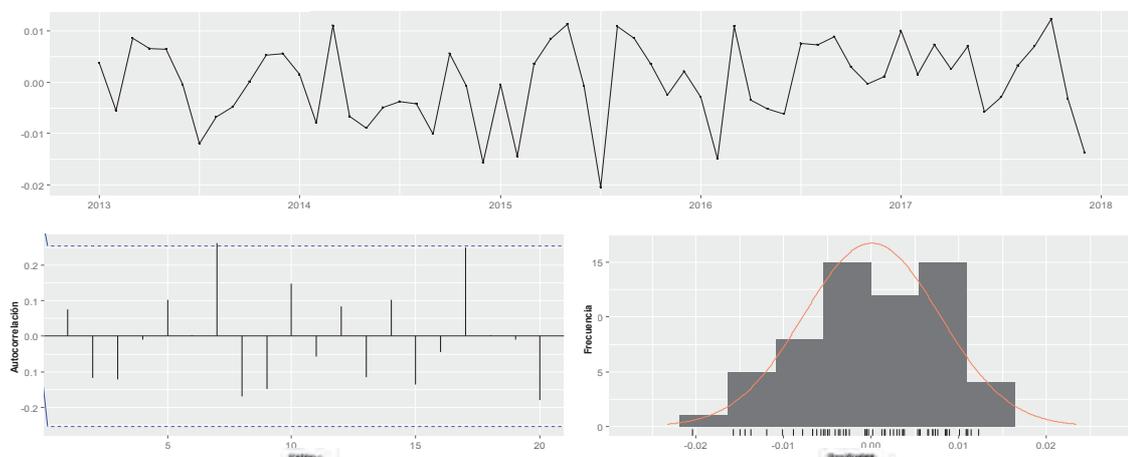


Figura 19. Evaluación de la normalidad y comportamiento de ruido blanco de los residuos del modelo ARIMA (0,1,1) de la DCS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

En la parte superior la serie de los residuos y en la parte inferior la función de autocorrelación donde un pico es significativo de lo cual se verificó que no afecta al comportamiento de ruido blanco (prueba de Ljung-Box, $p > 0.05$) y el histograma que sigue un perfil de campana de Gauss.

Pronóstico

Tabla 8
Pronóstico del modelo ARIMA para la DCS

	Modelo	RMSE	MAPE
ARIMA(0,1,1)	Entrenamiento	485.23	5.110848
	Pronóstico	477.6972	4.501505

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que la medición del error del modelo ARIMA(0,1,1) del entrenamiento con un RMSE y MAPE de 485.23 y 5.11% respectivamente; del pronóstico con RMSE y MAPE de 477.69 y 4.50% respectivamente, cuyos pronóstico se detalla en la tabla 6.

Tabla 9
Valores reales y pronosticados de la DCS

Meses	Año (2018)		Límite de confianza 95%	
	Real	Pronóstico	Inferior	Superior
Enero	8834	9555.696	8410.866	10898.63
Febrero	9171	9610.994	8445.112	10981.60
Marzo	9755	9666.688	8479.686	11065.12
Abril	10271	9722.779	8514.586	11149.22
Mayo	9998	9779.271	8549.809	11233.90
Junio	9361	9836.169	8585.355	11319.18
Julio	9298	9893.474	8621.221	11405.06
Agosto	10468	9951.192	8657.407	11491.55
Setiembre	9880	10009.324	8693.912	11578.67
Octubre	9646	10067.876	8730.734	11666.41
Noviembre	9703	10126.85	8767.873	11754.80
Diciembre	10854	10186.25	8805.329	11843.84

Fuente. Elaboración propia 2019 (Anexo N° 03)

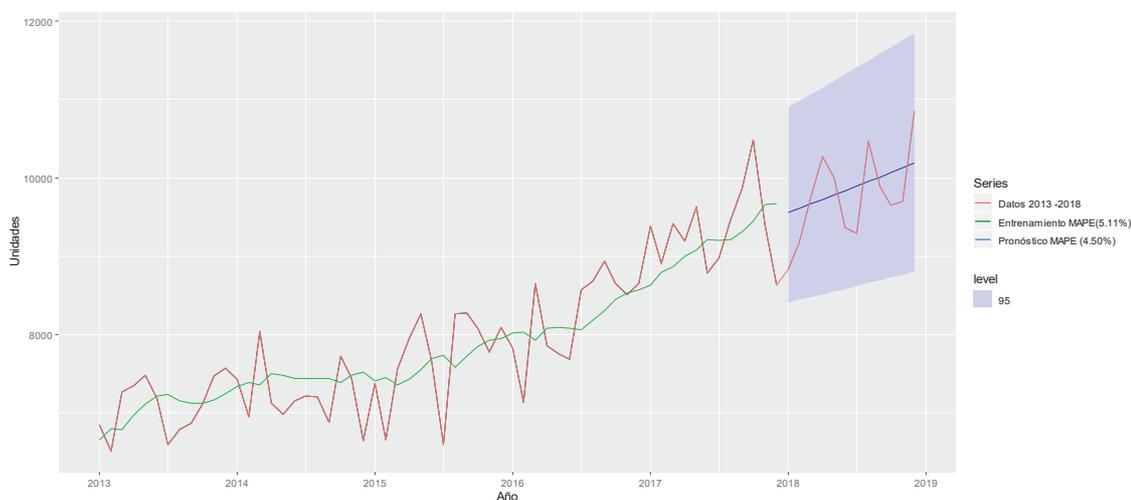


Figura 20. Valores del entrenamiento y pronóstico del modelo ARIMA (0,1,1) para la DCS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se aprecia las series de los datos 2013 a 2018, del entrenamiento y del pronóstico con un MAPE de 5.11% y 4.50% respectivamente, el pronóstico tiene un nivel de confianza de 95%.

Tabla 10
Estimación del modelo de la RNNLSTM de la DCS

Modelo	Parámetros					Medición del error			
	Retraso	Grado de aprendizaje	Función de activación	Número de época	Nodo	Entrenamiento		Pronóstico	
						RMSE	MAPE	RMSE	MAPE
RNNLSTM	12	0.05	Sigmoidea	1000	1	965.9134	10.425	1469.519	13.69942
					2	929.9725	9.812215	1590.826	14.97124
					3	625.7362	6.530046	947.2069	7.692918
					4	702.7854	7.081001	949.8443	8.0897
					5	764.7543	8.131129	816.3543	6.663667
					6	923.7946	9.911896	1322.007	11.7493
					7	582.4484	6.270912	642.4783	4.78815
					8	890.9118	9.267342	1822.067	17.35436
					9	924.269	9.530383	1762.359	16.93505
					10	930.4101	9.569321	1769.675	16.89535

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se tiene diferentes modelos de la RNNLSTM, donde difieren únicamente en el número de nodos, es así que el mejor modelo presenta los siguientes parámetros retrasos con 12, grado de aprendizaje con 0.05, función de activación es de sigmoidea y tangente hiperbólica, número de épocas con 1000 y número de nodos con 7, con el mencionado modelo en cuanto a la medición de los errores se tiene se tiene que en el entrenamiento y pronóstico un RMSE de 582.4484 y 642.4783 respectivamente y un MAPE de 6.27% y 4.78% respectivamente. El modelo estimado queda expresado como $y_{12x1} = \sum_{p=1}^7 h_{Bloque\ memoria_p}$ (Anexo N° 04).

Tabla 11
Valores reales y pronósticos de la DCS

Año (2018)		
Meses	Real	Pronóstico
Enero	8834	8940.78
Febrero	9171	9104.74
Marzo	9755	9392.35
Abril	10271	9358.65
Mayo	9998	9246.57
Junio	9361	9114.84
Julio	9298	9221.76
Agosto	10468	9487.96
Setiembre	9880	9476.25
Octubre	9646	9458.64
Noviembre	9703	9385.53
Diciembre	10854	9410.67

Fuente. Elaboración propia 2019 (Anexo N° 03)

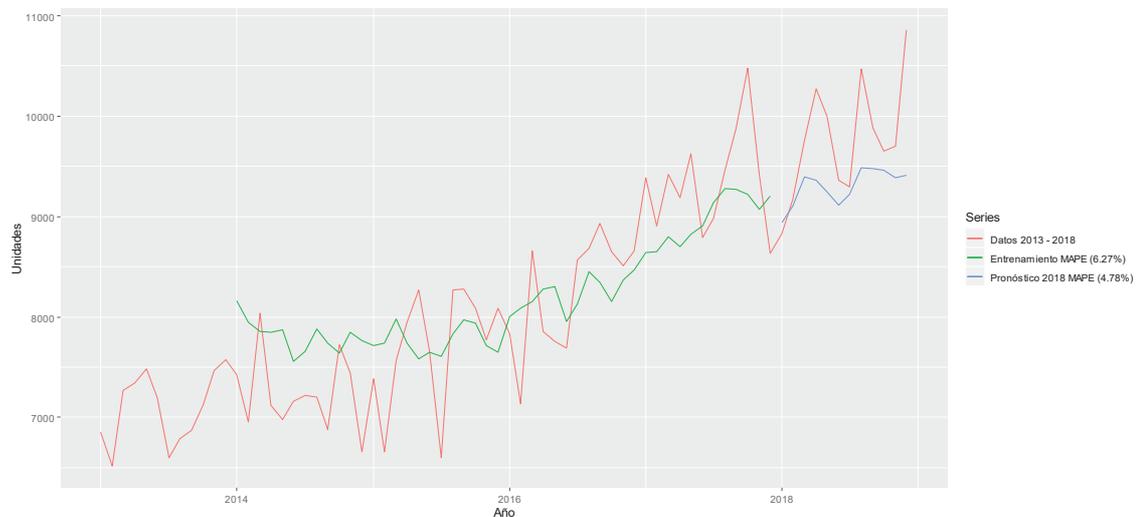


Figura 21. Valores del entrenamiento y pronóstico del modelo de la RNNLSTM para la DCS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se aprecia las series de los datos 2013 a 2018, del entrenamiento y del pronóstico con un MAPE de 6.27% y 4.78% respectivamente.

Tabla 12
Pronóstico de los modelos para la DCS

Medición del error		Modelo	
		ARIMA	RNNLSTM
RMSE	Entrenamiento	485.23	582.4484
	Pronóstico	477.6972	642.4783
MAPE	Entrenamiento	5.110848	6.270912
	Pronóstico	4.501505	4.78815

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que la medición del error de los modelos ARIMA y RNNLSTM del entrenamiento con un RMSE de 485.23 y 582.4484 y con un MAPE de 5.11% y 6.27% respectivamente; del pronóstico con un RMSE de 477.6972 y 642.4783 y con un MAPE de 4.50% y 4.78%, respectivamente.



Figura 22. Serie con datos reales, entrenamiento y pronóstico de los modelos ARIMA y RNNLSTM

Fuente. Elaboración propia 2019 (Anexo N° 03)

DEMANDA DE IBUPROFENO (DI) 400 MG TABLETA

Descripción del comportamiento temporal de la DI

Tabla 13

Estadísticos descriptivos de la DI por año

Año	2013	2014	2015	2016	2017
n (meses)	12	12	12	12	12
Media	6647	6498	6879	7300	7193
Desviación estandar	1136.647	932.7412	699.2781	925.1938	1755.088
Quartil 1	6020	5985	6279	6462	6602
Mediana	6600	6418	7211	7188	7049
Quartil 3	7125	7073	7344	8013	7994
Coefficiente de variación	17.10058	14.3541	10.16503	12.67302	24.39966

Fuente. Departamento de Farmacia del Hospital Regional del Cusco (Anexo N° 01)

INTERPRETACIÓN

Se aprecia que no hay mucha variación en cuanto a las medidas de tendencia central como la media y la mediana entre los años 2013 y 2015, pero si una variación que se mantienen estable entre los años 2016 y 2017. Por otro lado, el coeficiente de variación indica que los datos dentro de cada año si presentan una variación respecto a la media, dicho coeficiente de 2013 al 2107 no sobrepasan el 24.40%.

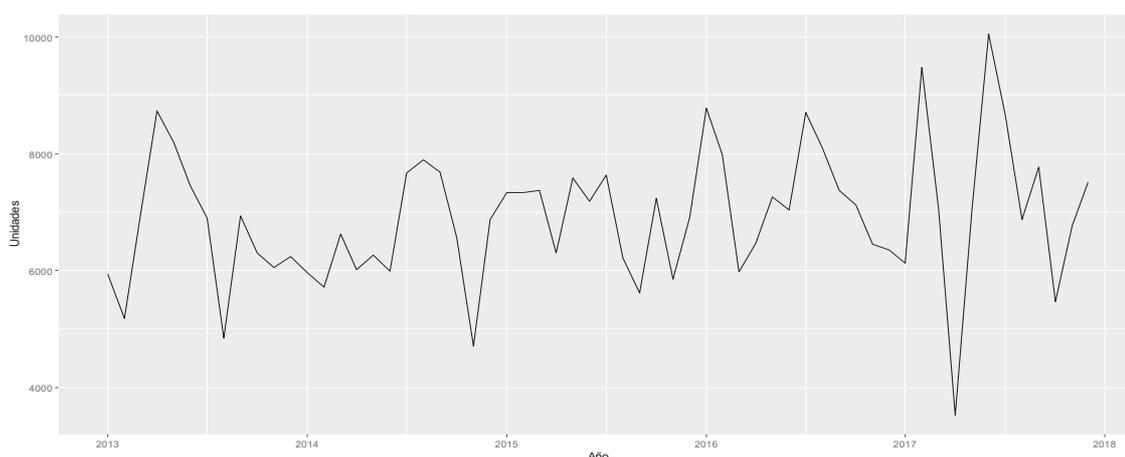


Figura 23. DI mensual de enero 2013 a diciembre de 2017

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se percibe que la serie es estacionaria, no presenta tendencia y una probable ausencia de estacionalidad.

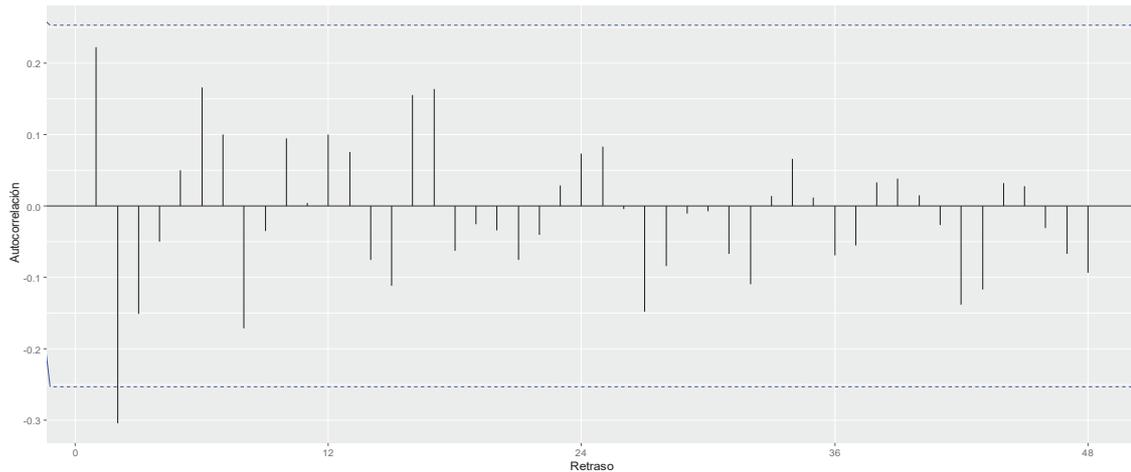


Figura 24. Función de autocorrelación de la DI

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la autocorrelación picos no significativos a excepción de uno solo y una distribución irregular lo cual sugiere que la serie es estacionaria y ausencia de estacionalidad.

Tabla 14

Estadísticos descriptivos de toda la serie de la DI

Año	2013 - 2017
n (meses)	60
Media	6940
Desviación estandar	1150.738
Quartil 1	6186
Mediana	6922
Quartil 3	7529
Coefficiente de variación	16.66878

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la serie entre los años 2013 al 2017 una media, mediana y coeficiente de variación de 6904, 6922 y 16.67% respectivamente.

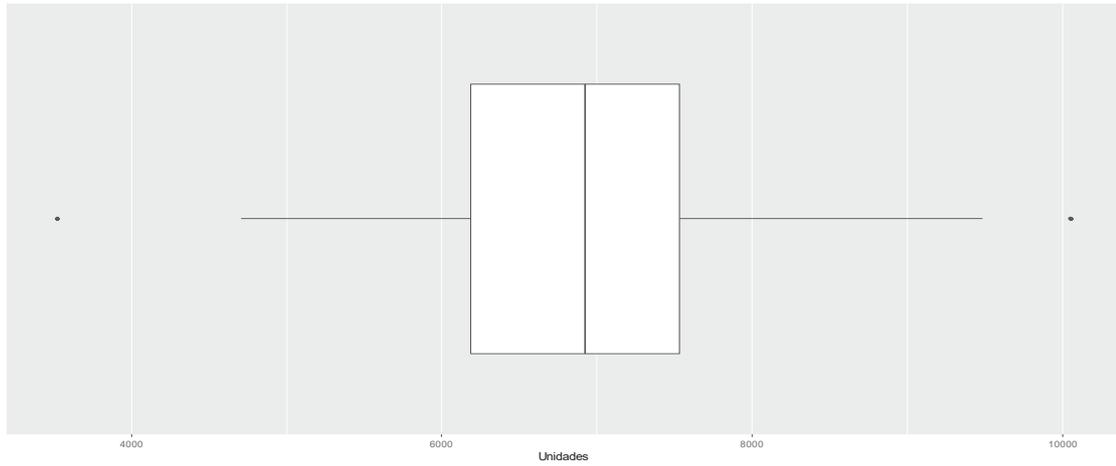


Figura 25. Diagrama de caja de la DI

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa en la serie una distribución casi simétrica con un sesgo ligero a la derecha y presencia de dos atípicos.

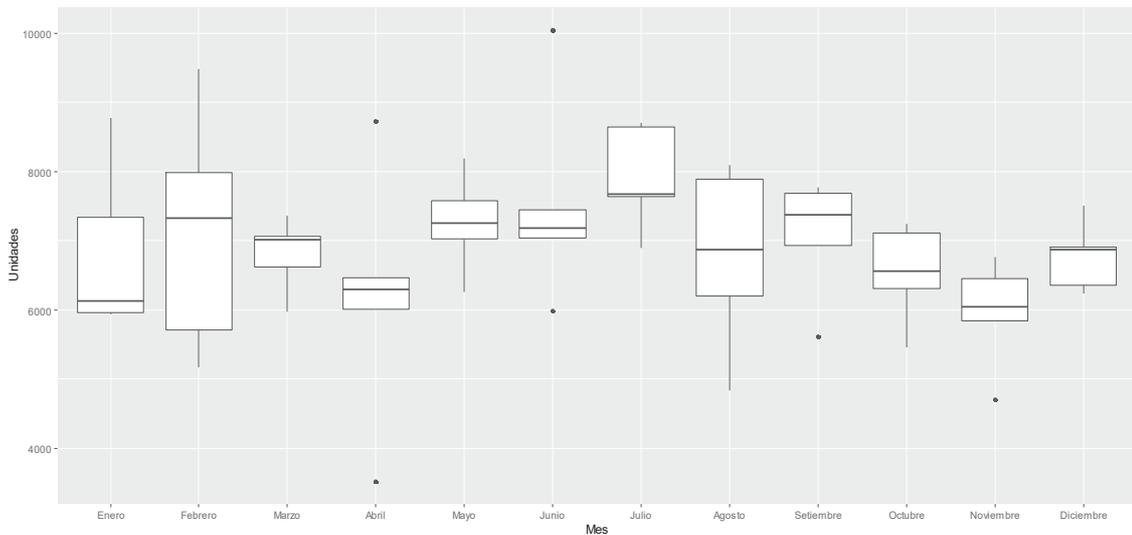


Figura 26. Diagrama de caja por mes de la DI

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la serie datos atípicos en los meses de abril, junio, setiembre y noviembre; una distribución concentrada en los meses de marzo, mayo y diciembre lo cual indica estacionalidad por otro lado, en los meses enero, febrero y agosto se tiene una mayor dispersión de la demanda, el mismo que se confirma

con el rango intercuartílico (altura de las cajas). Asimismo, una menor y mayor demanda del medicamento durante los meses de abril y julio, respectivamente.

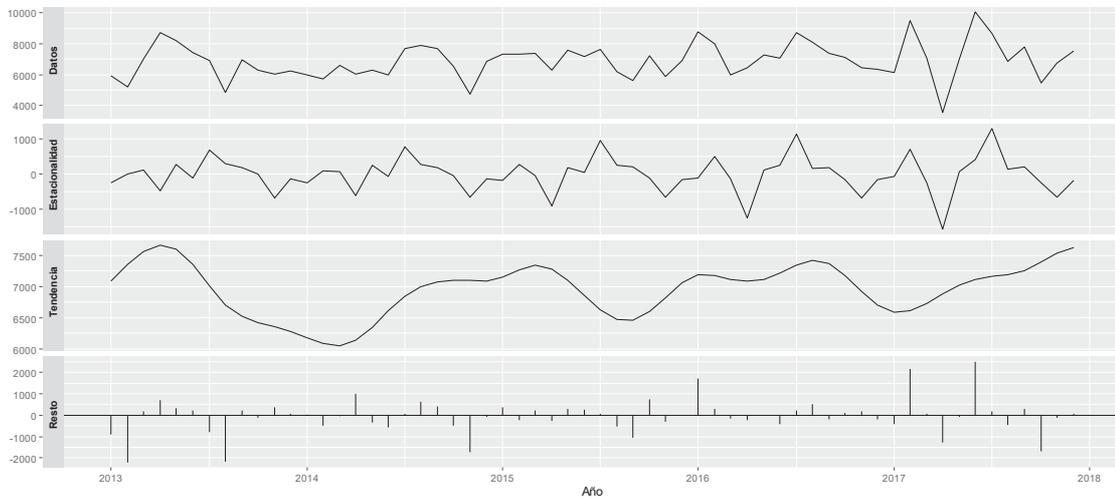


Figura 27 Descomposición de la DI

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

En la parte superior la serie original, luego la componente de estacionalidad donde se aprecia una aparente estacionalidad, después la componente de la tendencia donde no se distingue una tendencia ni positiva ni negativa y finalmente el componente del resto.

Tabla 15
Pruebas de estacionariedad, tendencia, estacionalidad y normalidad de la DI

KPSS		Mann-Kendall	Kruskal-Wallis		Shapiro-Wilk
p	valor crítico 5%	p	Año	p	p
			2013 a 2018	0.297	
			2013 - 2014	0.6442	
			2013 - 2015	0.4529	
			2013 - 2016	0.106	
			2013 - 2017	0.2727	
0.3366	0.463	0.1152	2014 - 2015	0.4189	0.6548
			2014 - 2016	0.05675	
			2014 - 2017	0.1489	
			2015 - 2016	0.3263	
			2015 - 2017	0.729	
			2016 - 2017	0.8174	

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se percibe que la serie presenta estacionariedad, ausencia de tendencia y estacionalidad lo cual se verifica con las pruebas de KPSS ($p < 0.463$), Mann-Kendall ($p > 0.05$) y Kruskal-Wallis ($p > 0.05$), respectivamente.

Con respecto a la estacionalidad al realizar la prueba Kruskal-Wallis pareada por años, se verifica también estacionalidad para todos.

En cuanto a la prueba de normalidad de Shapiro-Wilk se percibe que la prueba no es significativa ($p > 0.05$) lo cual evidencia homogeneidad de la serie.

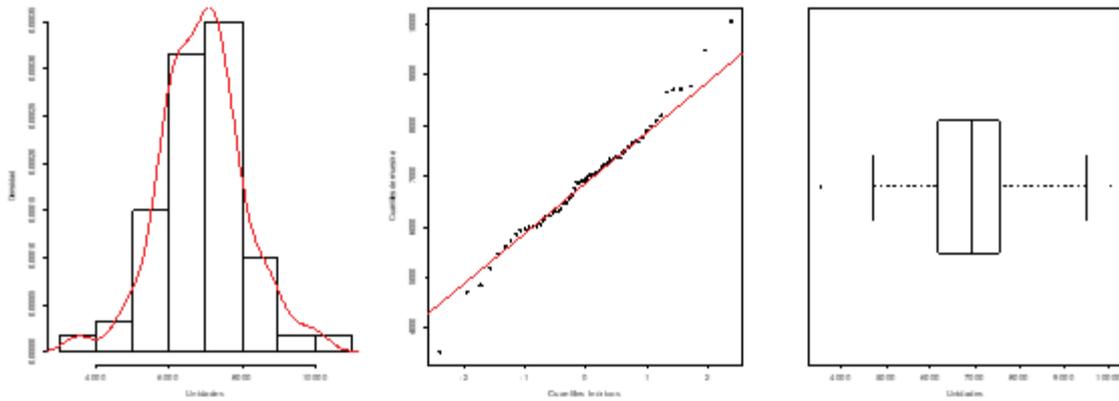


Figura 28. Evaluación de la normalidad de la DI

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Los gráficos señalan que los datos de la serie siguen un perfil de campana de Gauss, lo cual se verifica en el gráfico del histograma una distribución simétrica, el gráfico de la distribución de probabilidad línea recta se verifica que sigue una línea recta y el diagrama de la caja es simétrico.

Identificación de modelo (selección tentativa de p,d,q)

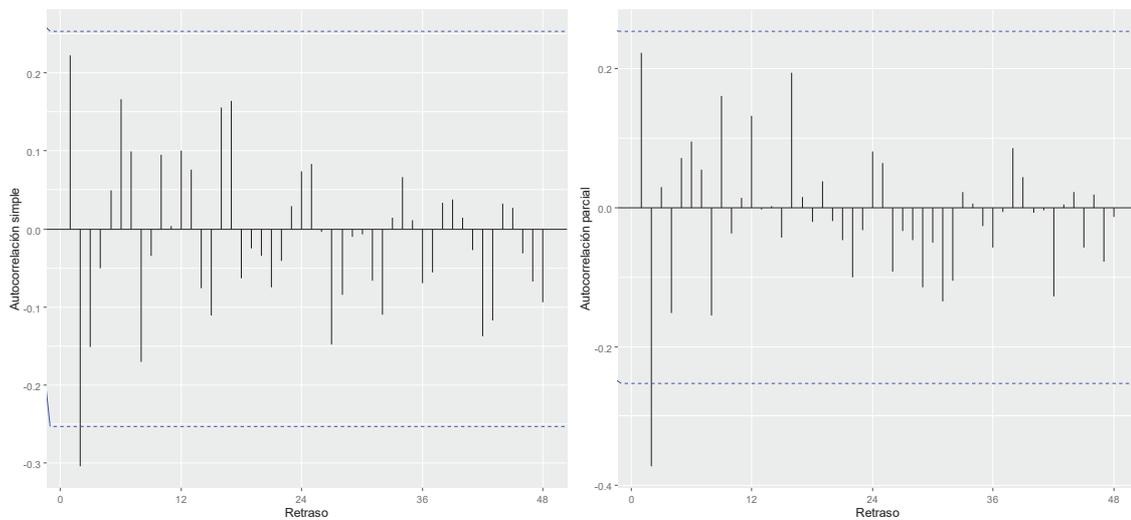


Figura 29. Función de autocorrelación simple y parcial de la serie de la DI
Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa en la función de autocorrelación simple que el primer rezago por defecto es siempre altamente significativo, por otro lado, se tiene un pico altamente significativo en la banda inferior que correspondiente al segundo retraso, asimismo se observa que los rezagos no siguen un patrón de descenso; respecto a la función de autocorrelación parcial se tiene que el primer rezago por defecto es siempre altamente significativo, a su vez se observa un pico altamente significativo en la banda inferior que corresponde al segundo retraso, además se percibe que los rezagos no siguen un patrón de descenso. Del gráfico se tiene indicios de plantear un modelo con componentes AR(1) y MA(1).

Tabla 16
Identificación del modelo para la DI

Modelos	Medidas de ajuste					Pruebas	
	AIC	Entrenamiento		Pronóstico		Ljung-Box	Shapiro-Wilk
		RMSE	MAPE	RMSE	MAPE		
ARIMA(1,0,1)	1012.72	1043.714	12.71329	1778.328	27.57768	0.4266	0.7553
ARIMA(0,0,1)	1013.34	1067.754	13.09545	1773.94	27.26661	0.2829	0.9816
ARIMA(1,0,0)	1018	1112.094	13.34755	1778.251	27.37043	0.09545	0.1882

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se ve claramente que de los diferentes modelos propuestos el mejor modelo es el ARIMA(1,0,1) debido a que presenta el menor AIC (1012.72) y además cumple los requisitos de normalidad e independencia de los residuales del modelo, a pesar que el MAPE del pronóstico es ligeramente alto.

Estimación de los parámetros del modelo elegido

Tabla 17
Modelo ARIMA para la DI

Modelo	Parámetros				Medición del error			
	Coeficiente	Estimado	Error estándar	AIC	Entrenamiento		Pronóstico	
					RMSE	MAPE	RMSE	MAPE
ARIMA(1,0,1)	ar1	-0.3699	0.2037	1012.72	1043.714	12.71329	1778.328	27.57768
	ma1	0.7742	0.1412					
	constante	6902.0165	174.0292					

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que el modelo para la DI, es un ARIMA compuesto por una parte regular la cual contiene componentes AR de orden 1 y MA de orden 1.

En cuanto a los parámetros se tiene un ar1 con -0.3699, una ma1 con 0.7742, una constante con 6902.0165, y un AIC de 1012.72.

En relación a la medición del error se tiene que en el entrenamiento y pronóstico un RMSE de 1043.714 y 1778.328 respectivamente y un MAPE de 12.71 y 27.58 respectivamente.

Respecto al modelo estimado se tiene un modelo ARIMA (1,0,1) cuya ecuación general es $Y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \varphi_1 Y_{t-1}$ (Anexo N° 02) al reemplazar los parámetros estimados $c = 6902.0165$, $\varphi_1 = -0.3699$ y $\theta_1 = 0.7742$ y se obtiene la siguiente ecuación del modelo ARIMA como sigue $Y_t = 6902.0165 + \varepsilon_t - 0.3699\varepsilon_{t-1} + 0.7742Y_{t-1}$.

Validación del modelo

Tabla 18
Pruebas de normalidad de los residuos y comportamiento de ruido blanco del modelo ARIMA (1,0,1) de la DI

Prueba	p
Shapiro-Wilk	0.7553
Ljung-Box	0.4266

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se aprecia que la prueba de normalidad de Shapiro-Wilk no es significativa ($p > 0.05$), con lo cual se evidencia la normalidad de los residuos. En cuanto a la prueba de Ljung-Box se aprecia que no es significativa ($p > 0.05$), señal de que los residuos son independientes, es decir tienen un comportamiento de ruido blanco.

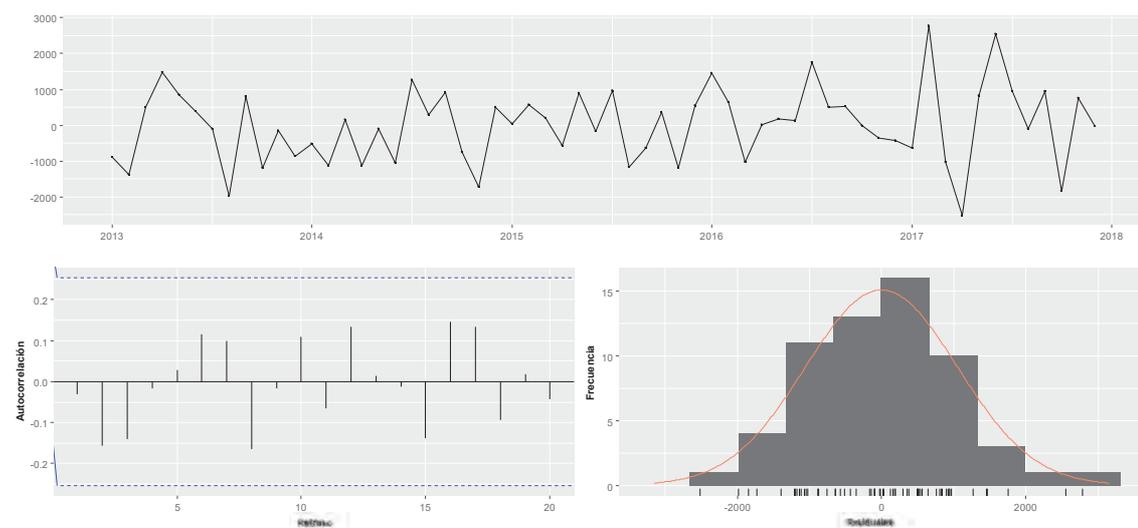


Figura 30. Evaluación de la normalidad y comportamiento de ruido blanco de los residuos del modelo ARIMA (1,0,1) de la DI

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

En la parte superior la serie de los residuos y en la parte inferior la función de autocorrelación donde no se tiene picos significativos con lo cual se verifica que el modelo sigue un comportamiento de ruido blanco (prueba de Ljung-Box, $p > 0.05$) y el histograma que sigue un perfil de campana de Gauss.

Pronóstico

Tabla 19
Pronóstico del modelo ARIMA para la DI

	Modelo	RMSE	MAPE
ARIMA(1,0,1)	Entrenamiento	1043.714	12.71329
	Pronóstico	1778.328	27.57768

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que la medición del error del modelo ARIMA(1,0,1) del entrenamiento con un RMSE y MAPE de 1043.714 y 12.71% respectivamente; del pronóstico con RMSE y MAPE de 1778.328 y 27.58% respectivamente.

Tabla 20
Valores reales y pronosticados de la DI

Meses	Año (2018)		Límite de confianza 95%	
	Real	Pronóstico	Inferior	Superior
Enero	6931	6655.855	4610.213	8701.497
Febrero	4795	6993.076	4786.571	9199.58
Marzo	2827	6868.332	4640.719	9095.946
Abril	6492	6914.477	4683.99	9144.963
Mayo	7325	6897.407	4666.528	9128.287
Junio	9530	6903.722	4672.788	9134.655
Julio	5036	6901.386	4670.445	9132.326
Agosto	6370	6902.25	4671.308	9133.191
Setiembre	7445	6901.93	4670.989	9132.872
Octubre	7526	6902.048	4671.107	9132.99
Noviembre	8513	6902.005	4671.063	9132.946
Diciembre	8454	6902.021	4671.079	9132.962

Fuente. Elaboración propia 2019 (Anexo N° 03)



Figura 31. Valores del entrenamiento y pronóstico del modelo ARIMA (1,0,1) para la DI

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se aprecia las series de los datos 2013 a 2018, del entrenamiento y del pronóstico con un MAPE de 12.71% y 27.58% respectivamente, el pronóstico tiene un nivel de confianza de 95%.

Tabla 21
Estimación del modelo de la RNNLSTM de la DI

Modelo	Parámetros					Medición del error			
	Retraso	Grado de aprendizaje	Función de activación	Número de época	Nodo	Entrenamiento		Pronóstico	
						RMSE	MAPE	RMSE	MAPE
RNNLSTM	12	0.05	Sigmoidea	1000	1	1998.963	23.82898	2284.197	30.54214
					2	1649.533	18.56706	1953.134	27.73048
					3	1150.287	12.75557	1871.926	27.06213
					4	1121.634	12.80117	1722.321	26.61708
					5	1126.439	12.48708	1670.095	25.45653
			6		1138.295	12.90447	1821.811	28.51988	
			7		1111.551	12.7195	1646.159	25.00375	
			8		1678.857	19.71014	2291.604	37.86992	
			9		1142.043	13.10045	1770.097	27.60552	
			10		1139.134	13.00075	1771.488	27.71806	

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se tiene diferentes modelos de la red neuronal tipo LSTM, donde difieren únicamente en el número de nodos, es así que el mejor modelo presenta los siguientes parámetros retrasos con 12, grado de aprendizaje con 0.05, función de activación es de sigmoidea y tangente hiperbólica, número de épocas con 1000 y número de nodos con 7, con el mencionado modelo en cuanto a la medición de los errores se tiene que en el entrenamiento y pronóstico un RMSE de 1111.551 y 1646.159 respectivamente y un MAPE de 12.71% y 25.00% respectivamente. El modelo estimado queda expresado como $y_{12x1} = \sum_{p=1}^7 h_{\text{Bloque memoria}_p}$ (Anexo N° 04).

Tabla 22
Valores reales y pronosticados de la DI

Año (2018)		
Meses	Real	Pronóstico
Enero	6931	6580.02
Febrero	4795	6765.55
Marzo	2827	6502.05
Abril	6492	6854.72
Mayo	7325	7035.66
Junio	9530	6895.99
Julio	5036	6577.03
Agosto	6370	6490.62
Setiembre	7445	6825.97
Octubre	7526	6700.58
Noviembre	8513	6976.08
Diciembre	8454	7026.91

Fuente. Elaboración propia 2019 (Anexo N° 03)



Figura 32. Valores del entrenamiento y pronóstico del modelo de la RNNLSTM para la DI

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se aprecia las series de los datos 2013 a 2018, del entrenamiento y del pronóstico con un MAPE de 12.71% y 25.00% respectivamente.

Tabla 23
Pronóstico de los modelos para la DI

Medición del error		Modelo	
		ARIMA	RNNLSTM
RMSE	Entrenamiento	1043.714	1111.551
	Pronóstico	1778.328	1646.159
MAPE	Entrenamiento	12.71329	12.7195
	Pronóstico	27.57768	25.00

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que la medición del error de los modelos ARIMA y RNNLSTM del entrenamiento con un RMSE de 1043.714 y 1111.551 y con un MAPE de 12.71% y 12.72% respectivamente; del pronóstico con un RMSE de 1778.328 y 1646.159 y con un MAPE de 27.58% y 25.00%, respectivamente.

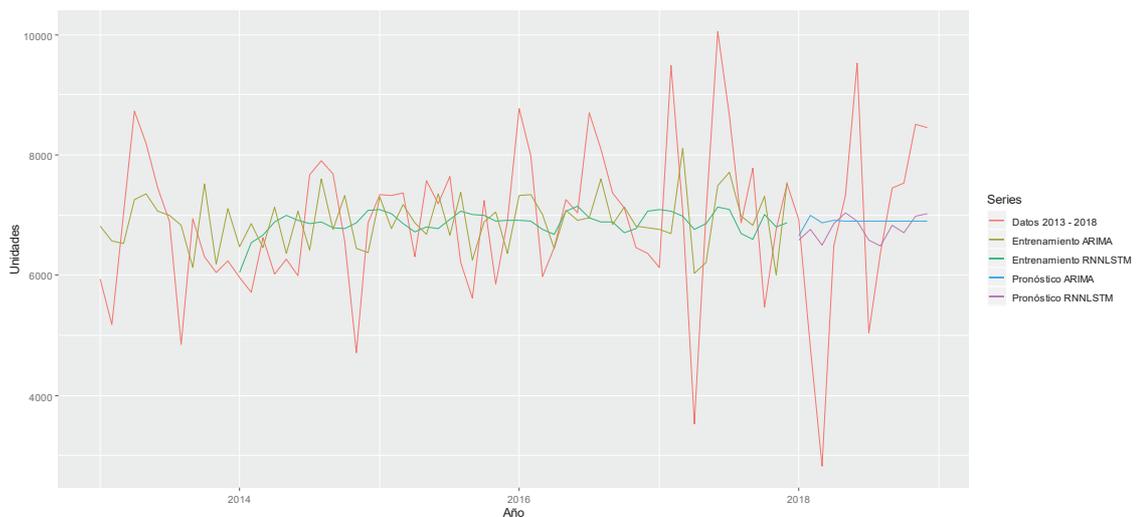


Figura 33. Serie con datos reales, entrenamiento y pronóstico de los modelos ARIMA y RNNLSTM

Fuente. Elaboración propia 2019 (Anexo N° 03)

DEMANDA DE METAMIZOL SODICO (DMS) 1 G - 2 ML INYECTABLE

Descripción del comportamiento temporal de la DMS

Tabla 24

Estadísticos descriptivos de la DMS por año

Año	2013	2014	2015	2016	2017
n (meses)	12	12	12	12	12
Media	9413	9671	10914	12125	12183
Desviación estandar	1244.828	700.115	873.3692	1184.947	1255.881
Quartil 1	8841	9274	10121	11355	11702
Mediana	9684	9700	10807	11662	12356
Quartil 3	10213	9981	11540	13332	12944
Coefficiente de variación	13.22444	7.239449	8.002406	9.772556	10.30875

Fuente. Departamento de Farmacia del Hospital Regional del Cusco (Anexo N° 01)

INTERPRETACIÓN

Se aprecia que no hay mucha variación en cuanto a las medidas de tendencia central como la media y la mediana entre los años 2013 y 2014, pero si una variación que se mantienen estable entre los años 2015 y 2017. Por otro lado, el coeficiente de variación indica que los datos dentro de cada año no presentan mucha variación respecto a la media, dicho coeficiente de 2013 al 2107 no sobrepasan el 13.22%.

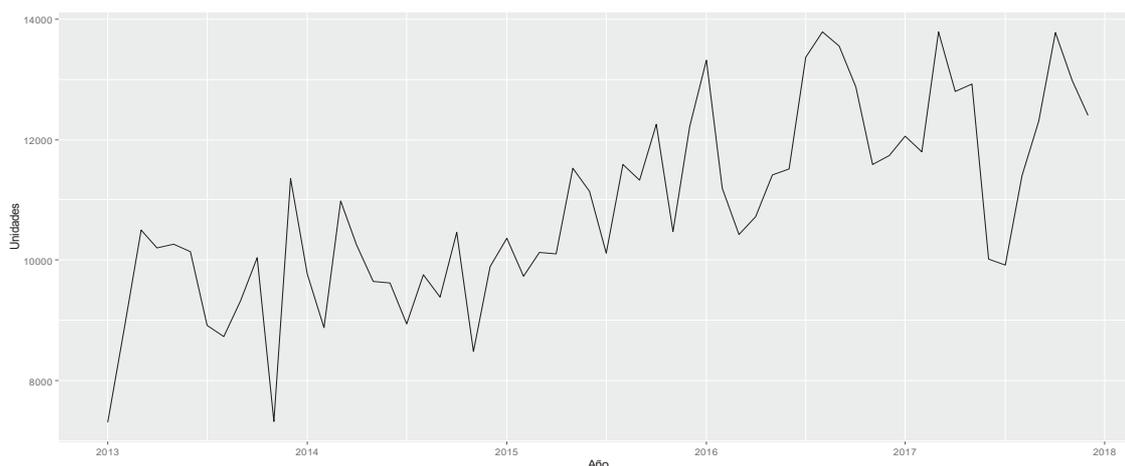


Figura 34. DMS mensual de enero 2013 a diciembre de 2017

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se percibe que la serie es no estacionaria, presenta tendencia positiva y una probable estacionalidad.

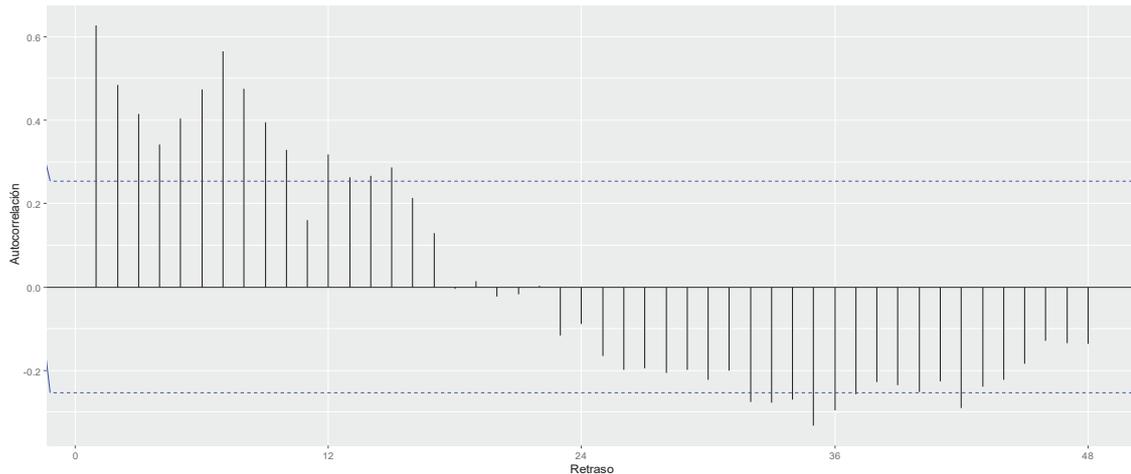


Figura 35. Función de autocorrelación de la DMS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la autocorrelación una disminución lenta, forma festoneada regular y picos significativos lo cual indica que la serie presenta tendencia positiva, estacionalidad y no estacionariedad, respectivamente.

Tabla 25

Estadísticos descriptivos de toda la serie de la DMS

Año	2013 - 2017
n (meses)	60
Media	10861
Desviación estandar	1572.826
Quartil 1	9863
Mediana	10488
Quartil 3	11865
Coefficiente de variación	14.48123

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la serie entre los años 2013 al 2017 una media, mediana y coeficiente de variación de 10861, 10488 y 14.48% respectivamente.

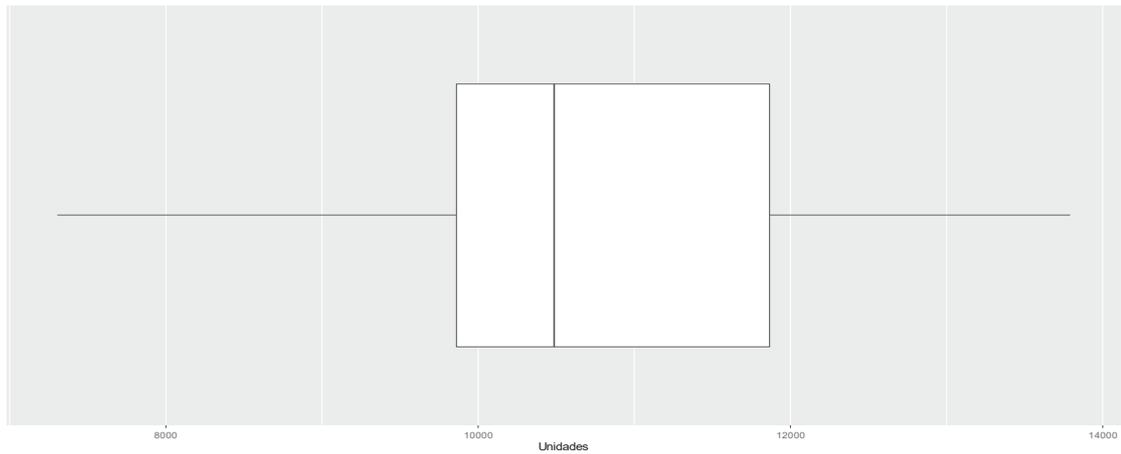


Figura 36. Diagrama de caja de la DMS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa en la serie una distribución asimétrica con un sesgo a la izquierda y ausencia de datos atípicos.

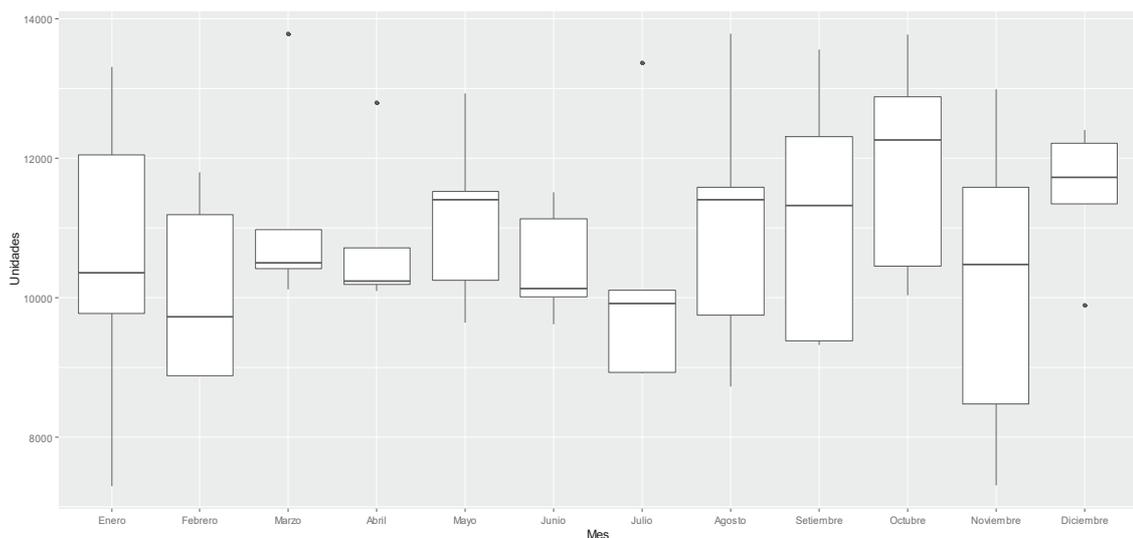


Figura 37. Diagrama de caja por mes de la DMS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la serie datos atípicos en los meses de marzo, abril, julio y diciembre; una distribución concentrada en los meses de marzo, abril, mayo, junio, julio y diciembre lo cual indica estacionalidad, por otro lado en los meses enero, febrero, agosto, setiembre, octubre y noviembre se tiene una mayor dispersión de la demanda, el mismo que se confirma con el rango intercuartílico

(altura de las cajas). Asimismo, una menor y mayor demanda del medicamento durante los meses de julio y octubre, respectivamente.

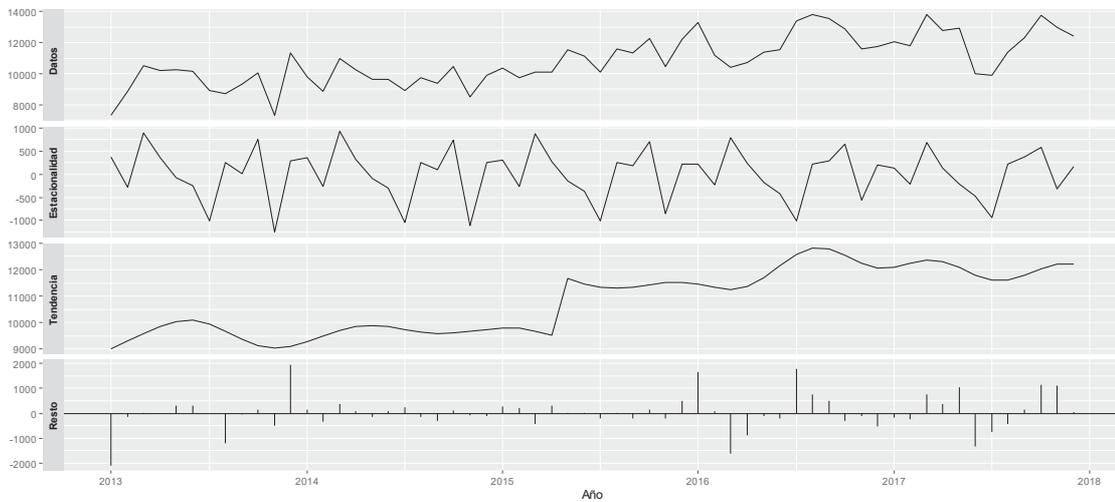


Figura 38. Descomposición de la DMS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

En la parte superior la serie original, luego la componente de estacionalidad donde se aprecia una aparente estacionalidad, después la componente de la tendencia donde se distingue una aparente tendencia positiva y finalmente el componente del resto.

TABLA 26***Pruebas de estacionariedad, tendencia, estacionalidad y normalidad de la DMS***

Prueba					
KPSS		Mann-Kendall	Kruskal-Wallis		Shapiro-Wilk
p	valor crítico 5%	p	Año	p	p
			2013 a 2018	6.113E-07	
			2013 - 2014	0.7728	
			2013 - 2015	0.009375	
			2013 - 2016	8.638E-05	
			2013 - 2017	0.000532	
1.3181	0.463	7.824E-10	2014 - 2015	0.001823	0.2553
			2014 - 2016	6.785E-05	
			2014 - 2017	0.0001387	
			2015 - 2016	0.01531	
			2015 - 2017	0.01531	
			2016 - 2017	0.795	

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se percibe que la serie presenta no estacionariedad, tendencia y ausencia estacionalidad lo cual se verifica con las pruebas de KPSS ($p > 0.463$), Mann-Kendall ($p < 0.05$) y Kruskal-Wallis ($p < 0.05$), respectivamente.

Respecto a la estacionalidad si bien la prueba de Kruskal-Wallis indica que la serie para los años 2103 a 2017 no presenta estacionalidad, pero si presenta estacionalidad para los años 2013 y 2014; 2016 y 2017.

Se observa que la prueba de normalidad de Shapiro-Wilk no es significativa ($p > 0.05$) lo cual evidencia homogeneidad de la serie.

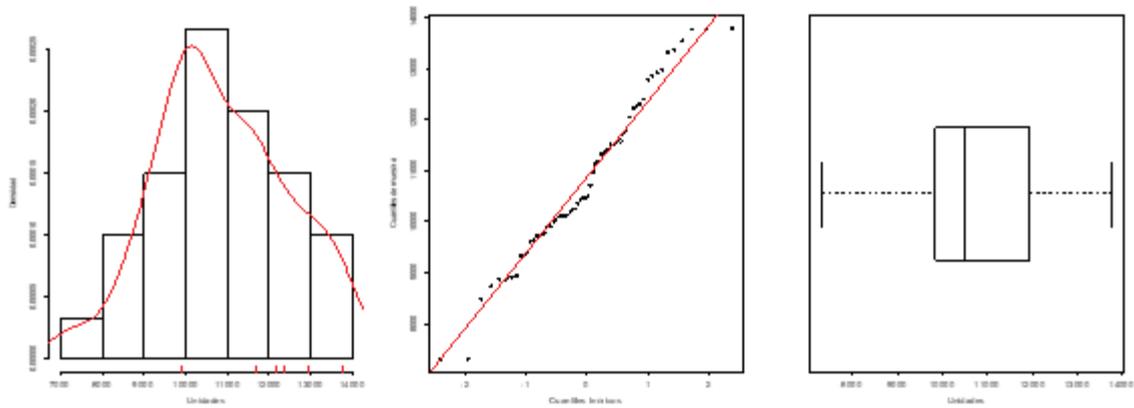


Figura 39. Evaluación de la normalidad de la DMS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Los gráficos señalan que los datos de la serie siguen un perfil de campana de Gauss, lo cual se verifica en el gráfico del histograma una distribución simétrica, el gráfico de la distribución de probabilidad línea recta se verifica que sigue una línea recta y el diagrama de la caja es asimétrico con un sesgo a la izquierda.

Identificación del modelo (selección tentativa de p,d,q)

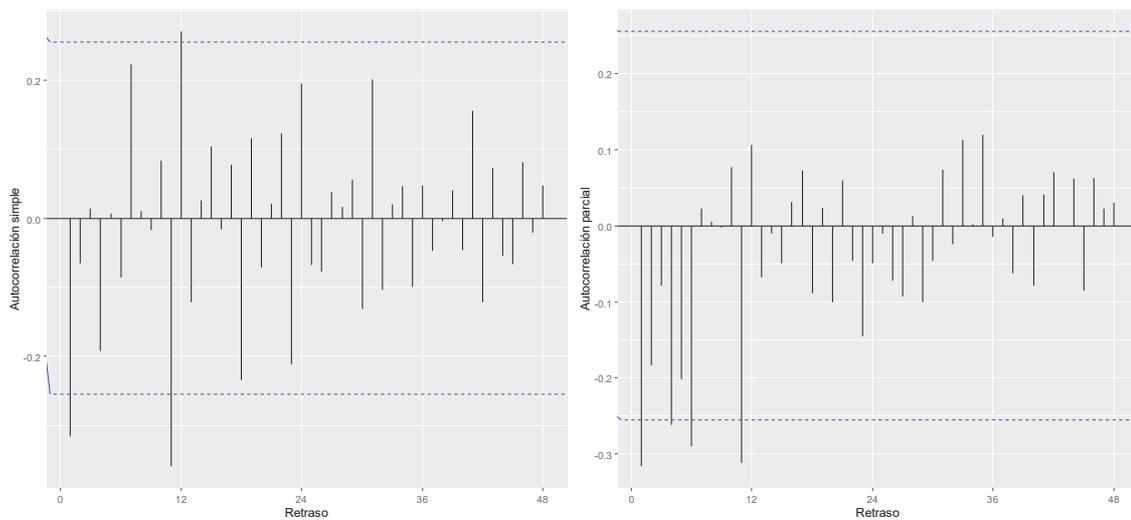


Figura 40. Función de autocorrelación simple y parcial de la serie con diferenciación de la DMS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la función de autocorrelación simple el primer rezago por defecto es siempre altamente significativo, por otro lado, se tiene dos picos significativos en la banda superior e inferior en los rezagos 12 y 11 respectivamente; respecto a la función de autocorrelación parcial se tiene que el primer rezago por defecto es siempre altamente significativo, a su vez se observa tres picos el primer pico ligeramente significativo y el segundo y tercero altamente significativos en la banda inferior, además se percibe que los rezagos no siguen un patrón de descenso. Del gráfico se propone un modelo con los componentes $AR(3)$, $MA(1)$ y $MA(1)_{12}$.

Tabla 27
Identificación del modelo para la DMS

Modelos	Medidas de ajuste					Pruebas	
	AIC	Entrenamiento		Pronostico		Ljung-Box	Shapiro-Wilk
		RMSE	MAPE	RMSE	MAPE		
ARIMA(3,1,1)(0,0,1) ₁₂	1002.7	1045.169	7.09161	664.4076	4.662146	0.06647	0.1066
ARIMA(1,1,1)(0,0,1) ₁₂	1000.17	1063.178	7.257202	781.3016	5.335493	0.04344	0.3147
ARIMA(2,1,1)(0,0,1) ₁₂	1001.79	1057.997	7.194991	762.1944	5.275552	0.03884	0.1729

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que de los diferentes modelos propuestos el mejor modelo es el ARIMA(3,1,1)(0,0,1)₁₂ debido a que los residuales del modelo cumplen con los requisitos de normalidad e independencia ($p > 0.05$) y además el MAPE del pronostico es el mas bajo, a pesar de que el AIC (1002.7) no sea el más bajo.

Estimación de los parámetros del modelo elegido

Tabla 28
Modelo ARIMA para la DMS

Modelo	Parámetros				Medición del error			
	Coeficiente	Estimado	Error estándar	AIC	Entrenamiento		Pronóstico	
					RMSE	MAPE	RMSE	MAPE
ARIMA(3,1,1)(0,0,1) ₁₂	ar1	0.3464	0.1603	1002.7	1045.169	7.09161	664.4076	4.662146
	ar2	-0.0662	0.1423					
	ar3	-0.1559	0.1456					
	ma1	-0.8132	0.0894					
	sma1	0.2864	0.1459					

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que el modelo para la DMS, es un ARIMA compuesto por una parte regular que contiene una diferencial de orden 1, un AR de orden 3 y un MA de orden 1; asimismo se tiene una parte estacional que contiene un MA de orden 1.

En cuanto a los parámetros se tiene un ar1 con 0.3464, un ar2 con -0.0662, un ar3 con -0.1559, una ma1 con -0.8132, una sma1 con 0.2864 y un AIC de 1002.7.

En relación a la medición del error se tiene que en el entrenamiento y pronóstico un RMSE de 1045.169 y 664.4076 respectivamente y un MAPE de 7.09 y 4.66 respectivamente.

Respecto al modelo estimado se tiene un modelo ARIMA (3,1,1)(0,0,1)₁₂ cuya ecuación general es $Y_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + (\Theta_1 + \theta_1\Theta_1)\varepsilon_{t-12} + (\varphi_1 + 1)Y_{t-1} + (\varphi_2 - \varphi_1)Y_{t-2} + (\varphi_3 - \varphi_2)Y_{t-3} - \varphi_3Y_{t-4}$ (Anexo N° 02) al reemplazar los parámetros estimados $\varphi_1 = 0.3464$, $\varphi_2 = -0.0662$, $\varphi_3 = -0.1559$, $\theta_1 = -0.8132$ y $\Theta_1 = 0.2864$ se obtiene la siguiente ecuación del modelo ARIMA como sigue $Y_t = \varepsilon_t - 0.8132\varepsilon_{t-1} + 0.0535\varepsilon_{t-12} + 1.3464Y_{t-1} - 0.4126Y_{t-2} - 0.0897Y_{t-3} + 0.1559Y_{t-4}$

Validación del modelo

Tabla 29

Pruebas de normalidad de los residuos y comportamiento de ruido blanco del modelo ARIMA (3,1,1)(0,0,1)₁₂ de la DMS

Prueba	p
Shapiro-Wilk	0.1066
Ljung-Box	0.06647

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se percibe que las pruebas de normalidad de Shapiro-Wilk no es significativa ($p > 0.05$), con lo cual se evidencia la normalidad de los residuos. En cuanto a la prueba de Ljung-Box se aprecia que no es significativa ($p > 0.05$), señal de que los residuos son independientes, es decir tienen un comportamiento de ruido blanco.

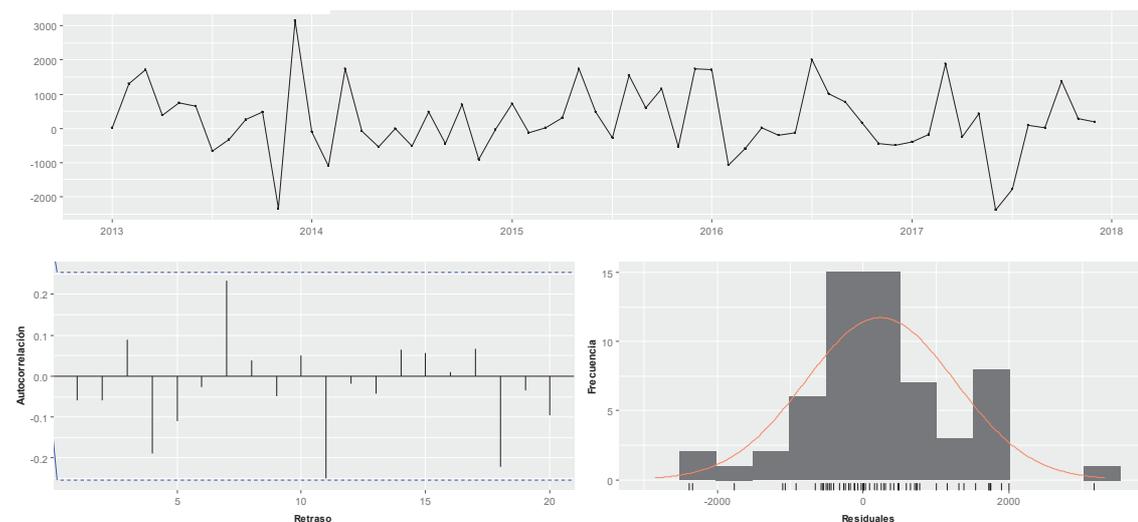


Figura 41. Evaluación de la normalidad y comportamiento de ruido blanco de los residuos del modelo ARIMA (3,1,1)(0,0,1)₁₂ de la DMS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

En la parte superior la serie de los residuos y en la parte inferior la función de autocorrelación donde no se tiene picos significativos con lo cual se verifica que el modelo sigue un comportamiento de ruido blanco y el histograma que sigue un perfil de campana de Gauss.

Pronóstico

Tabla 30
Pronóstico del modelo ARIMA para la DMS

Modelo	RMSE	MAPE	
ARIMA(3,1,1)(0,0,1) ₁₂	Entrenamiento	1045.169	7.09161
	Pronóstico	664.4076	4.662146

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que la medición del error del modelo ARIMA(3,1,1)(0,0,1)₁₂ del entrenamiento con un RMSE y MAPE de 1045.169 y 7.09% respectivamente; del pronóstico con RMSE y MAPE de 664.4076 y 4.66% respectivamente.

Tabla 31
Valores reales y pronosticados de la DMS

Meses	Año (2018)		Límite de confianza 95%	
	Real	Pronóstico	Inferior	Superior
Enero	10667	11866.84	9801.059	13932.62
Febrero	11194	11885.28	9544.18	14226.39
Marzo	11960	12600.93	10176.356	15025.51
Abril	11355	12420.9	9987.293	14854.51
Mayo	11428	12486.48	10040.568	14932.380
Junio	11867	11629.76	9157.716	14101.81
Julio	11642	11403.42	8888.727	13918.12
Agosto	11832	11808.24	9247.444	14369.03
Setiembre	12115	12080.07	9477.042	14683.1
Octubre	12311	12575.79	9935.213	15216.37
Noviembre	11741	12426.63	9750.746	15102.520
Diciembre	12049	12288.28	9577.462	14999.1

Fuente. Elaboración propia 2019 (Anexo N° 03)



Figura 42. Valores del entrenamiento y pronóstico del modelo ARIMA (3,1,1)(0,0,1)₁₂ para la DMS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se aprecia las series de los datos 2013 a 2018, del entrenamiento y del pronóstico con un MAPE de 7.09% y 4.66% respectivamente, el pronóstico tiene un nivel de confianza de 95%.

Tabla 32
Estimación del modelo de la RRNNLSTM de la DMS

Modelo	Parámetros				Medición del error				
	Retraso	Grado de aprendizaje	Función de activación	Número de época	Nodo	Entrenamiento		Pronóstico	
						RMSE	MAPE	RMSE	MAPE
RRNNLSTM	12	0.05	Sigmoidea	1000	1	1851.915	12.04143	1695.411	13.88813
					2	1701.652	11.84045	1461.907	11.92875
					3	1003.436	7.160659	526.819	3.740168
					4	1269.804	9.44636	473.8777	3.384836
					5	1265.447	9.401383	643.1575	4.866628
			6		1115.495	7.994922	293.7089	1.8546	
			7		1043.818	7.58407	296.6202	2.212114	
			8		1383.573	10.23482	624.723	4.523433	
			9		1333.47	9.916054	552.4916	3.916852	
			10		1417.815	10.77745	677.5174	4.906891	

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se tiene diferentes modelos de la red neuronal tipo LSTM, donde difieren únicamente en el número de nodos, es así que el mejor modelo presenta los siguientes parámetros retrasos con 12, grado de aprendizaje con 0.05, función de activación es de tangente hiperbólica, número de épocas con 1000 y número de nodos con 6, con el mencionado modelo en cuanto a la medición de los errores se tiene que en el entrenamiento y pronóstico un RMSE de 1115.495 y 293.7089 respectivamente y un MAPE de 7.99% y 1.85% respectivamente. El modelo estimado queda expresado como $y_{12x1} = \sum_{p=1}^6 h_{\text{Bloque memoria}_p}$ (Anexo N° 04).

Tabla 33
Valores reales y pronosticados de la DMS

Año (2018)		
Meses	Real	Pronóstico
Enero	10667	10601.01
Febrero	11194	10864.45
Marzo	11960	11279.86
Abril	11355	11371.98
Mayo	11428	11665.89
Junio	11867	11880.56
Julio	11642	12015.48
Agosto	11832	12069.34
Setiembre	12115	12087.37
Octubre	12311	12079.74
Noviembre	11741	12131.51
Diciembre	12049	12047.35

Fuente. Elaboración propia 2019 (Anexo N° 03)

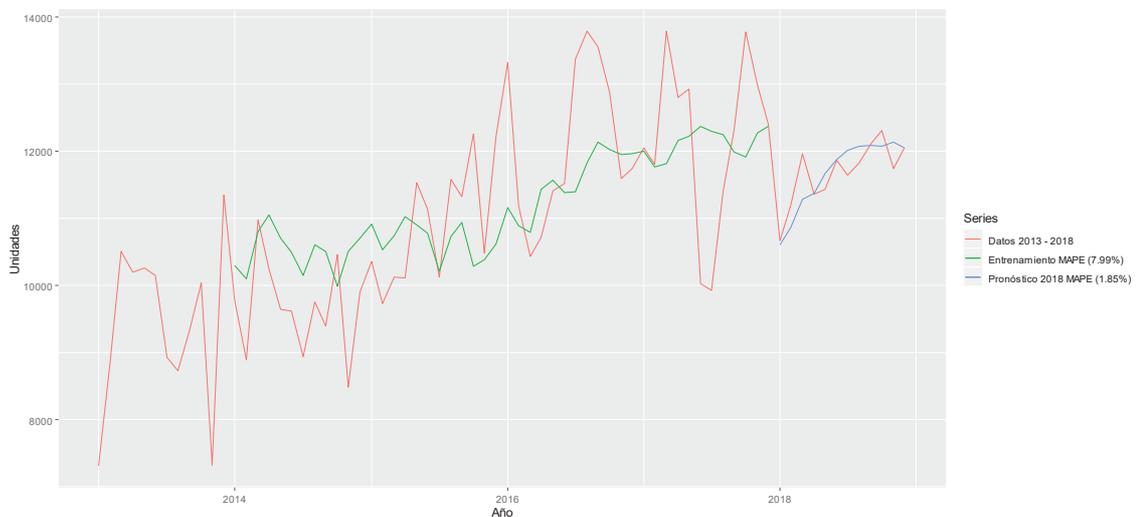


Figura 43. Valores del entrenamiento y pronóstico del modelo de la RNNLSTM para la DMS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se aprecia las series de los datos 2013 a 2018, del entrenamiento y del pronóstico con un MAPE de 7.99% y 1.85% respectivamente.

Tabla 34
Pronóstico de los modelos para la DMS

Medición del error		Modelo	
		ARIMA	RNNLSTM
RMSE	Entrenamiento	1045.169	1115.495
	Pronóstico	664.4076	293.7089
MAPE	Entrenamiento	7.09161	7.994922
	Pronóstico	4.662146	1.8546

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que la medición del error de los modelos ARIMA y RNNLSTM del entrenamiento con un RMSE de 1045.169 y 1115.495 y con un MAPE de 7.09% y 7.99% respectivamente; del pronóstico con un RMSE de 664.4076 y 293.7089 y con un MAPE de 4.66% y 1.85% respectivamente.

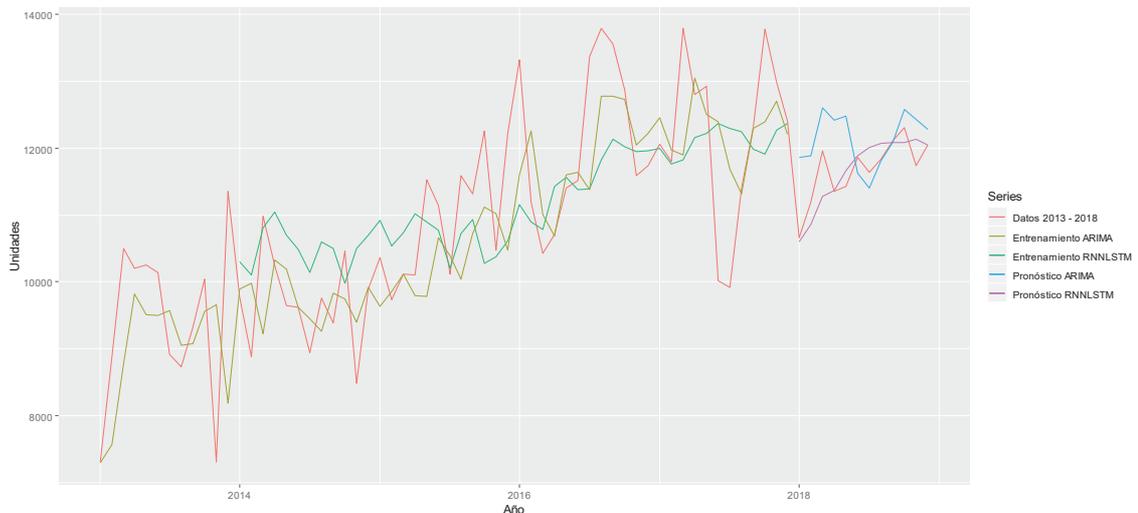


Figura 44. Serie con datos reales, entrenamiento y pronóstico de los modelos ARIMA y RNNLSTM

Fuente. Elaboración propia 2019 (Anexo N° 03)

DEMANDA DE PARACETAMOL (DP) 500 MG TABLETA

Descripción del comportamiento temporal de la DP

Tabla 35

Estadísticos descriptivos de la DP por año

Año	2013	2014	2015	2016	2017
n (meses)	12	12	12	12	12
Media	7413	7185	8176	11404	10368
Desviación estandar	615.8693	1361.812	2679.458	1576.077	2242.449
Quartil 1	7152	7208	7700	11086	7841
Mediana	7668	7547	8888	11404	11052
Quartil 3	7818	7942	9864	12222	12024
Coefficiente de variación	8.307496	18.95333	32.77291	13.82029	21.62787

Fuente. Departamento de Farmacia del Hospital Regional del Cusco (Anexo N° 01)

INTERPRETACIÓN

Se aprecia que no hay mucha variación en cuanto a las medidas de tendencia central como la media y la mediana entre los años 2013 y 2014, pero si una ligera variación ascendente entre los años 2015 al 2017. Por otro lado, el coeficiente de variación indica que los datos dentro de cada año si presentan mucha variación respecto a la media, dicho coeficiente de 2013 al 2107 no sobrepasan el 32.77%.

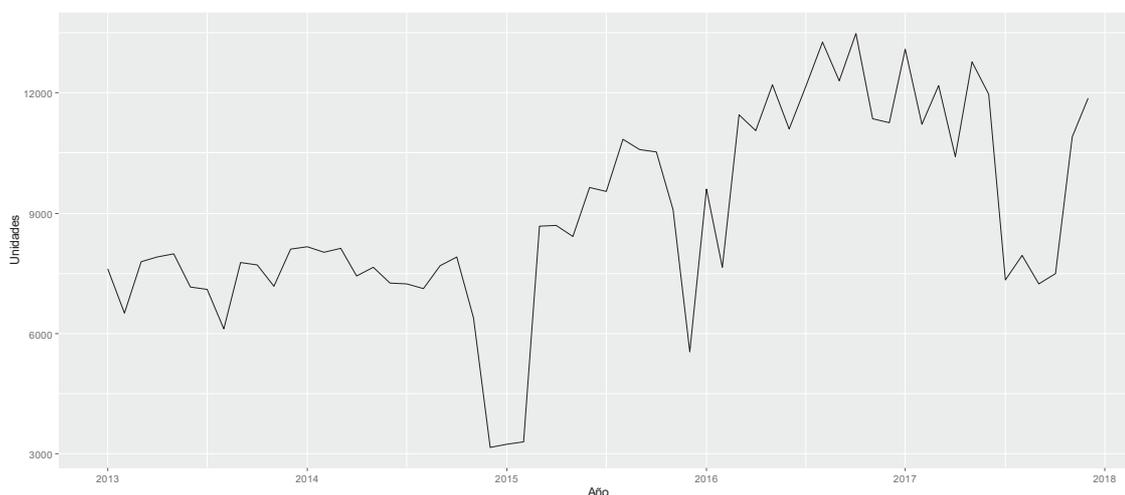


Figura 45. DP mensual de enero 2013 a diciembre de 2017

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se percibe que la serie presente probablemente una estacionariedad, tendencia positiva y estacionalidad.

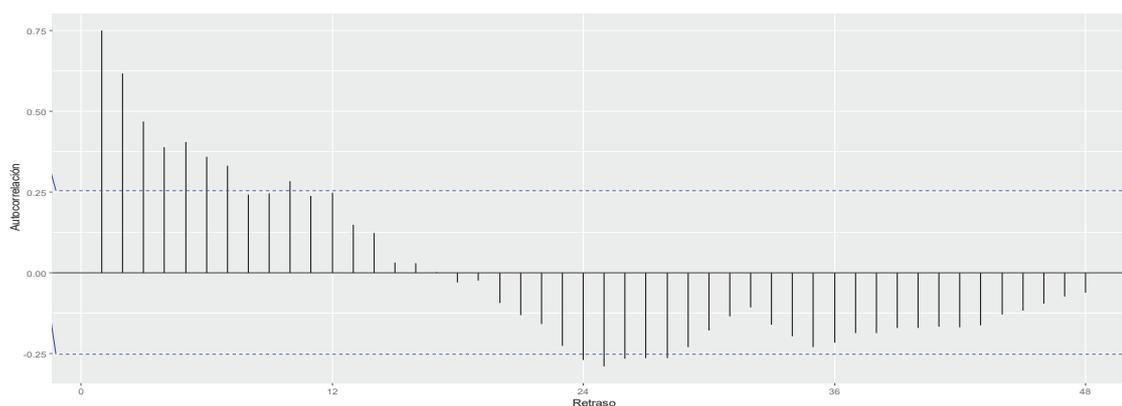


Figura 46. Función de autocorrelación de la DP

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la autocorrelación una disminución lenta, forma festoneada casi regular y picos significativos lo cual indica que la serie presenta tendencia positiva, estacionalidad y no estacionariedad, respectivamente.

Tabla 36

Estadísticos descriptivos de toda la serie de la DP

Año	2013 - 2017
n (meses)	60
Media	8909
Desviación estandar	2454.331
Quartil 1	7414
Mediana	8115
Quartil 3	11070
Coefficiente de variación	27.54781

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la serie entre los años 2013 al 2017 una media, mediana y coeficiente de variación de 8909, 8115 y 27.55% respectivamente.

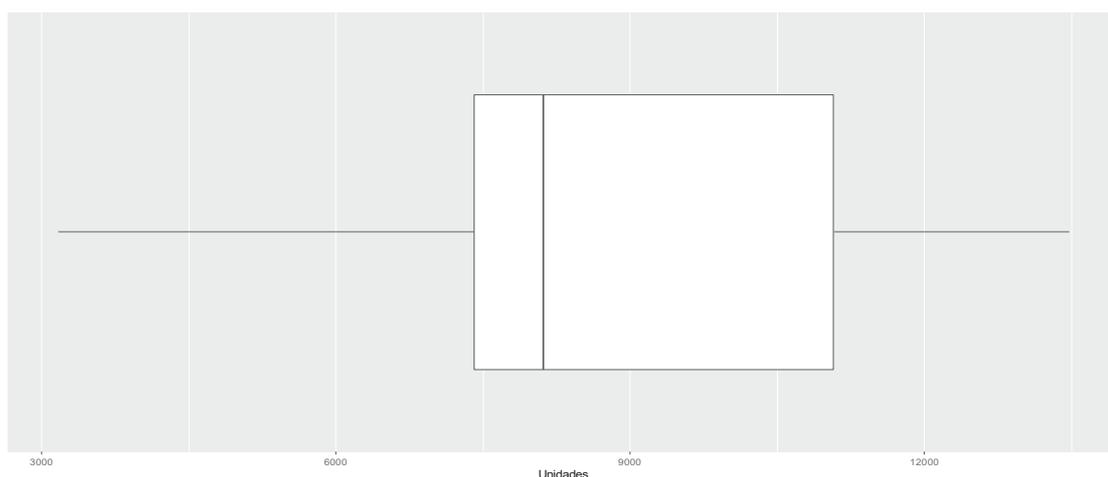


Figura 47. Diagrama de caja de la DP

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa en la serie una distribución asimétrica con un sesgo a la izquierda y ausencia de datos atípicos.

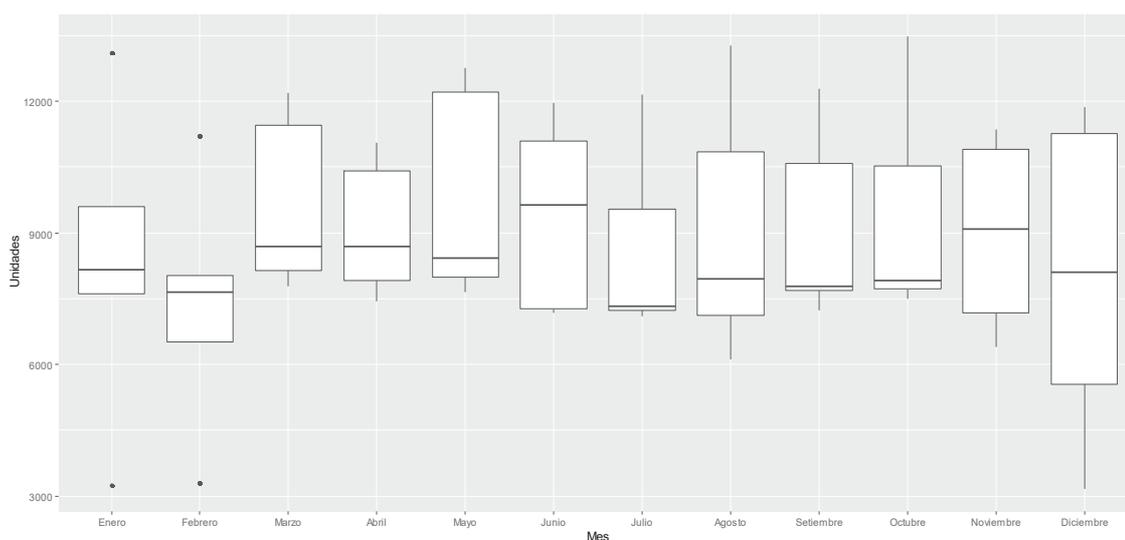


Figura 48. Diagrama de caja por mes de la DP

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la serie datos atípicos en los meses de enero y febrero; una distribución concentrada en la mayoría de los meses a excepción del mes de diciembre lo cual indica estacionalidad. Asimismo, una menor y mayor demanda del medicamento durante los meses de febrero y mayo, respectivamente.

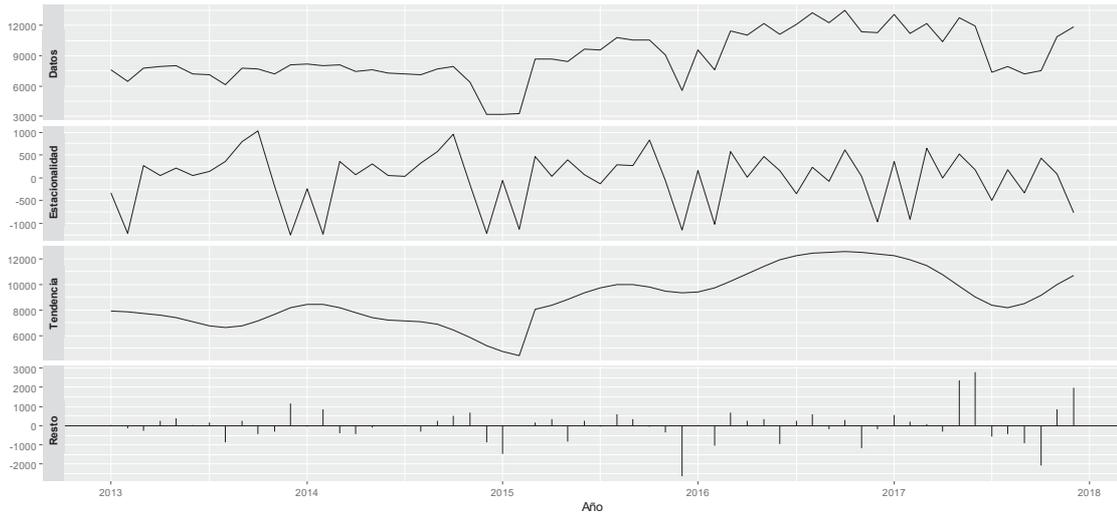


Figura 49 Descomposición de la DP

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

En la parte superior la serie original, luego la componente de estacionalidad donde se aprecia una aparente estacionalidad, después la componente de la tendencia donde se distingue una aparente tendencia positiva y finalmente el componente del resto.

Tabla 37**Pruebas de estacionariedad, tendencia, estacionalidad y normalidad de la DP**

Prueba					
KPSS		Mann-Kendall	Kruskal-Wallis		Shapiro-Wilk
p	valor crítico 5%	p	Año	p	p
			2013 a 2018	1.118E-05	
			2013 - 2014	0.8174	
			2013 - 2015	0.03767	
			2013 - 2016	0.0001387	
			2013 - 2017	0.004669	
0.8462	0.463	1.566E-06	2014 - 2015	0.02434	0.01027
			2014 - 2016	0.0001387	
			2014 - 2017	0.005584	
			2015 - 2016	0.0006583	
			2015 - 2017	0.05675	
			2016 - 2017	0.2253	

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se percibe que la serie presenta no estacionariedad, tendencia y no estacionalidad lo cual se verifica con las pruebas de KPSS ($p > 0.463$), Mann-Kendall ($p < 0.05$) y Kruskal-Wallis ($p < 0.05$), respectivamente.

En relación a la estacionalidad si bien la prueba de Kruskal-Wallis indica que la serie para los años 2103 a 2017 no presenta estacionalidad, pero si presenta estacionalidad para los años 2013 y 2014; 2015 y 2017; 2016 y 2017.

Se observa que las prueba de Shapiro-Wilk es significativa ($p < 0.05$) lo cual evidencia heterogeneidad de la serie.

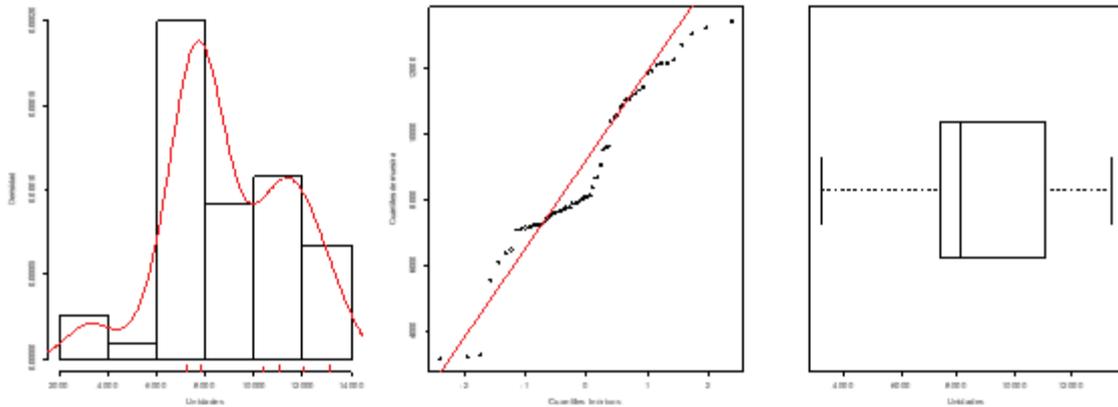


Figura 50. Evaluación de la normalidad de la DP
 Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Los gráficos señalan que los datos de la serie no siguen un perfil de campana de Gauss, lo cual se verifica en el gráfico del histograma una distribución asimétrica, el gráfico de la distribución de probabilidad línea recta se verifica que no sigue una línea recta y el diagrama de la caja es asimétrico con un sesgo a la izquierda.

Identificación del modelo (selección tentativa de p,d,q)

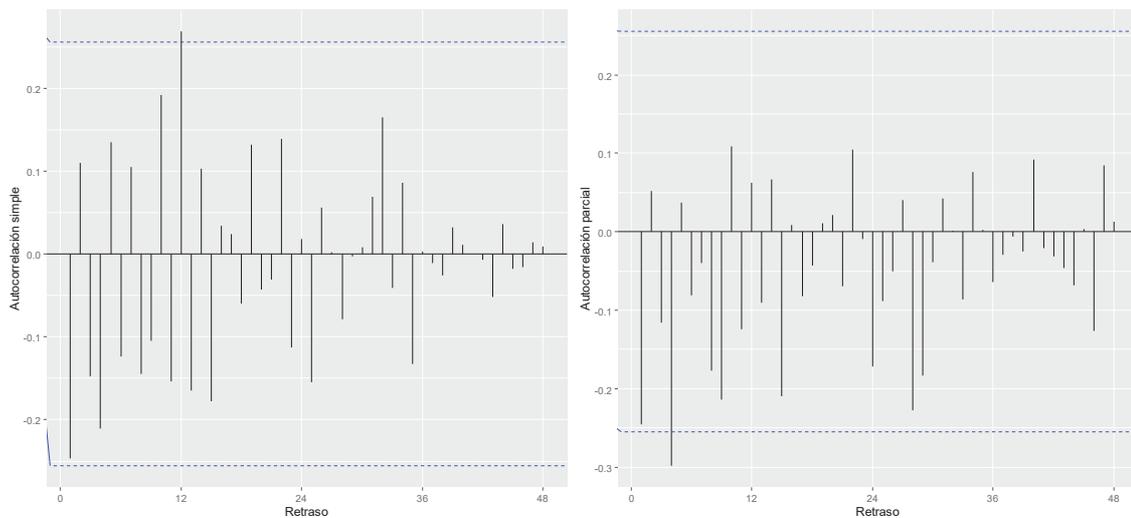


Figura 51. Función de autocorrelación simple y parcial de la serie con transformación BoxCox y diferenciación de la DP

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la función de autocorrelación simple el primer rezago por defecto es siempre altamente significativo, por otro lado, se tiene un pico significativo en la banda superior que corresponde al rezago 12, asimismo se observa que los rezagos no siguen un patrón de descenso; respecto a la función de autocorrelación parcial se tiene que el primer rezago por defecto es siempre altamente significativo, a su vez se observa un pico significativo en la banda inferior, además se percibe que los rezagos no siguen un patrón de descenso. Del gráfico se propone un modelo con componentes $AR(1)$ y $MA(1)_{12}$.

Tabla 38
Identificación del modelo para la DP

Modelos	Medidas de ajuste					Pruebas	
	AIC	Entrenamiento		Pronostico		Ljung-Box	Shapiro-Wilk
		RMSE	MAPE	RMSE	MAPE		
ARIMA(1,1,0)(0,0,1) ₁₂	697.43	1522.119	12.46366	2818.81	23.47212	0.5983	8.78E-06
ARIMA(2,1,0)(0,0,1) ₁₂	699.43	1522.061	12.46253	2820.987	23.48403	0.5031	8.86E-06
ARIMA(3,1,0)(0,0,1) ₁₂	700.59	1525.645	12.80218	2918.657	23.96733	0.412	4.21E-05

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que de los diferentes modelos propuestos el mejor modelo es el ARIMA(1,1,0)(0,0,1)₁₂ debido a que presenta el menor AIC (697.43), los residuales del modelo son independientes ($p > 0.05$). y el MAPE del pronóstico es el más bajo (23.47%); sin embargo, los residuales no siguen una distribución normal.

Estimación de los parámetros del modelo elegido

Tabla 39
Modelo ARIMA para la DP

Modelo	Parámetros					Medición del error			
	Coeficiente	Estimado	Error estándar	AIC	lambda	Entrenamiento		Pronóstico	
						RMSE	MAPE	RMSE	MAPE
ARIMA(1,1,0)(0,0,1) ₁₂	ar1	-0.1719	0.1282	697.43	0.6740248	1522.119	12.46366	2818.81	23.47212
	sma1	0.3436	0.1522						

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que el modelo para la DP, es un ARIMA compuesto por una parte regular que contiene una diferencial de orden 1 y un AR de orden 1; y la parte estacional que contiene un componente MA de orden 1.

En cuanto a los parámetros se tiene una ar1 con -0.1719, una sma1 con 0.3436, y un AIC de 697.43.

En relación a la medición del error se tiene que en el entrenamiento y pronóstico un RMSE de 1522.119 y 2818.81 respectivamente y un MAPE de 12.46 y 23.47 respectivamente.

Respecto al modelo estimado se tiene un modelo ARIMA (1,1,0)(0,0,1)₁₂ cuya ecuación general es $Y_t = \varepsilon_t + Y_{t-1} + \varphi_1 Y_{t-1} - \varphi_1 Y_{t-2} + \Theta_1 \varepsilon_{t-12}$ (Anexo N° 02) al reemplazar los parámetros estimados $\varphi_1 = -0.1719$ y $\Theta_1 = 0.3436$ se obtiene la siguiente ecuación del modelo ARIMA como sigue $Y_t = \varepsilon_t + Y_{t-1} - 0.1719 Y_{t-1} + 0.1719 Y_{t-2} + 0.3436 \varepsilon_{t-12}$.

Validación del modelo

Tabla 40
Pruebas de normalidad de los residuos y comportamiento de ruido blanco del modelo ARIMA (1,1,0)(0,0,1)₁₂ de la DP

Prueba	p
Shapiro-Wilk	8.78E-06
Ljung-Box	0.5983

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se percibe que las pruebas de normalidad de Shapiro-Wilk es significativa ($p < 0.05$), con lo cual se evidencia la no normalidad de los residuos. En cuanto a la prueba de Ljung-Box se aprecia que no es significativa ($p > 0.05$), señal de que los residuos son independientes, es decir tienen un comportamiento de ruido blanco.

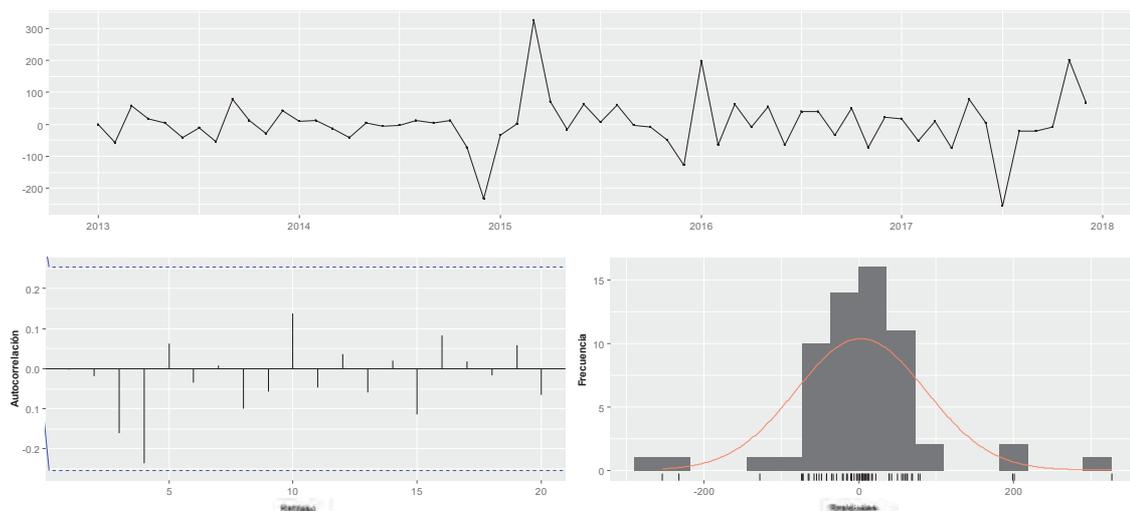


Figura 52. Evaluación de la normalidad y comportamiento de ruido blanco de los residuos del modelo ARIMA (1,1,0)(0,0,1)₁₂ de la DP

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

En la parte superior la serie de los residuos y en la parte inferior la función de autocorrelación donde no se tiene picos significativos con lo cual se verifica que el modelo sigue un comportamiento de ruido blanco (Ljung-Box, $p > 0.05$) y el histograma no sigue un perfil de campana de Gauss.

Pronóstico

Tabla 41
Pronóstico del modelo ARIMA para la DP

Modelo	RMSE	MAPE	
ARIMA(1,1,0)(0,0,1) ₁₂	Entrenamiento	1522.119	12.46366
	Pronóstico	2818.81	23.47212

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que la medición del error del modelo ARIMA(1,1,0)(0,0,1)₁₂ del entrenamiento con un RMSE y MAPE de 1522.119 y 12.46% respectivamente; del pronóstico con RMSE y MAPE de 2818.81 y 23.47% respectivamente.

Tabla 42
Valores reales y pronosticados de la DP

Meses	Año (2018)		Límite de confianza 95%	
	Real	Pronóstico	Inferior	Superior
Enero	10218	11813.785	8445.05	15531.02
Febrero	6520	11452.087	7202.058	16297.21
Marzo	8493	11583.986	6547.076	17474
Abril	11906	11030.91	5476.06	17717.13
Mayo	12924	11684.959	5454.257	19283.040
Junio	13343	11604.324	4912.337	19928.2
Julio	13397	9822.492	3208.921	18452.01
Agosto	13271	9984.545	2970.008	19282.83
Setiembre	13444	9811.932	2539.186	19667.8
Octubre	13634	9778.45	2231.29	20198.70
Noviembre	11721	11194.394	2843.7	22543.090
Diciembre	10138	11436.334	2721.933	23394.86

Fuente. Elaboración propia 2019 (Anexo N° 03)

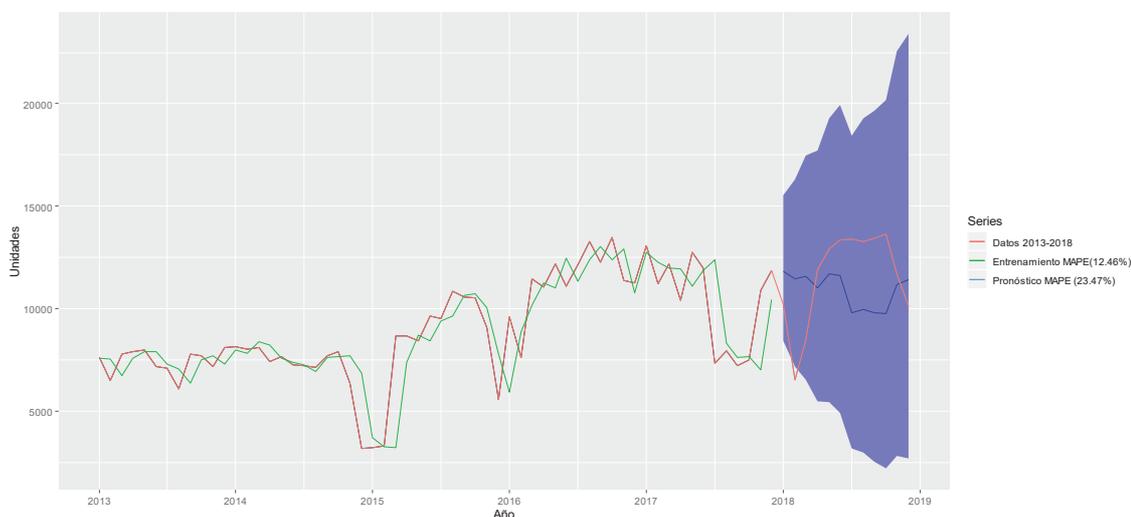


Figura 53. Valores del entrenamiento y pronóstico del modelo ARIMA (1,1,0)(0,0,1)₁₂ para la DP

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se aprecia las series de los datos 2013 a 2018, del entrenamiento y del pronóstico con un MAPE de 12.46% y 23.47% respectivamente, el pronóstico tiene un nivel de confianza de 95%.

Tabla 43
Estimación del modelo de la RNNLSTM de la DP

Modelo	Parámetros				Medición del error				
	Retraso	Grado de aprendizaje	Función de activación	Número de época	Nodo	Entrenamiento		Pronóstico	
						RMSE	MAPE	RMSE	MAPE
RNNLSTM	12	0.05	Sigmoidea	1000	1	3717.186	33.9524	5390.369	39.99071
					2	3385.995	30.7291	5070.929	37.69172
					3	2172.313	24.64014	3116.06	22.676
					4	2675.782	29.46248	3826.042	28.35353
					5	2192.155	24.96277	4452.86	33.08766
					6	2585.66	28.10933	3637.234	27.83573
					7	2122.731	22.76413	2870.549	22.19628
					8	3736.001	34.20005	5470.777	40.58866
					9	2569.323	28.53345	3748.463	29.47212
					10	2605.696	29.06795	3522.207	26.8137

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se tiene diferentes modelos de la red neuronal tipo LSTM, donde difieren únicamente en el número de nodos, es así que el mejor modelo presenta los siguientes parámetros retrasos con 12, grado de aprendizaje con 0.05, función de activación es de tangente hiperbólica, número de épocas con 1000 y número de nodos con 7, con el mencionado modelo en cuanto a la medición de los errores se tiene se tiene que en el entrenamiento y pronóstico un RMSE de 2122.731 y 2870.549 respectivamente y un MAPE de 22.76% y 22.19% respectivamente. El modelo estimado queda expresado como $y_{12x1} = \sum_{p=1}^7 h_{Bloque\ memoria_p}$ (Anexo N° 04).

Tabla 44
Valores reales y pronósticos de la DP

Año (2018)		
Meses	Real	Pronóstico
Enero	10218	10032.95
Febrero	6520	9960.67
Marzo	8493	8693.25
Abril	11906	8718.61
Mayo	12924	8900.15
Junio	13343	9803.51
Julio	13397	9399.86
Agosto	13271	10332.63
Setiembre	13444	10938.58
Octubre	13634	10224.25
Noviembre	11721	9709.94
Diciembre	10138	8900.92

Fuente. Elaboración propia 2019 (Anexo N° 03)



Figura 54. Valores del entrenamiento y pronóstico del modelo de la RNNLSTM para la DP

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se aprecia las series de los datos 2013 a 2018, del entrenamiento y del pronóstico con un MAPE de 22.76% y 22.19% respectivamente.

Tabla 45
Pronóstico de los modelos para la DP

Medición del error		Modelo	
		ARIMA	RNNLSTM
RMSE	Entrenamiento	1522.119	2122.731
	Pronóstico	2818.81	2870.549
MAPE	Entrenamiento	12.46366	22.76413
	Pronóstico	23.47212	22.19628

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que la medición del error de los modelos ARIMA y RNNLSTM del entrenamiento con un RMSE de 1522.119 y 2122.731 y con un MAPE de 12.46% y 22.76% respectivamente; del pronóstico con un RMSE de 2818.81 y 2870.549 y con un MAPE de 23.47% y 22.19% respectivamente.

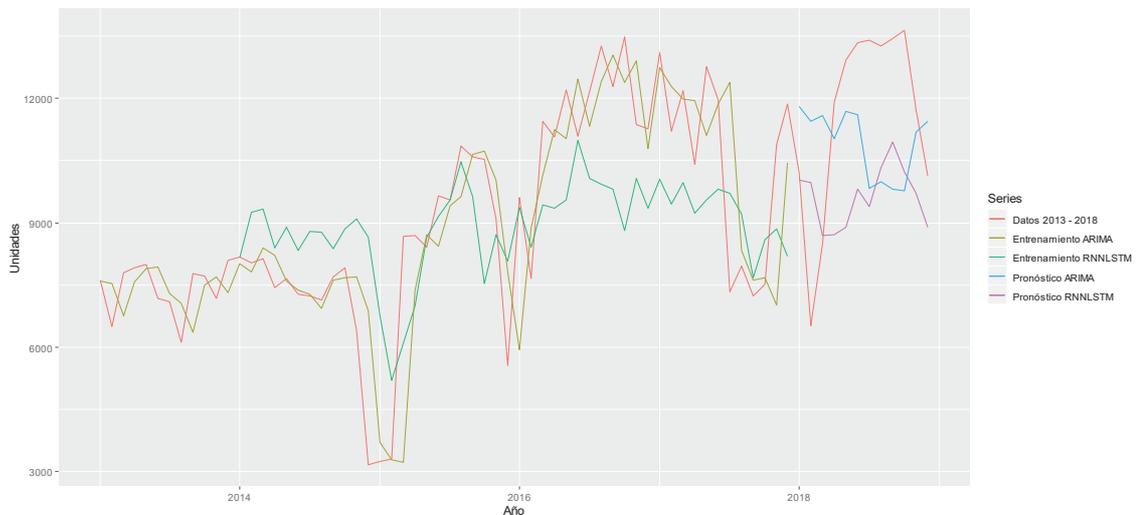


Figura 55. Serie con datos reales, entrenamiento y pronóstico de los modelos ARIMA y RNNLSTM

Fuente. Elaboración propia 2019 (Anexo N° 03)

DEMANDA DE SEVOFLURANO (DS) 250 ML SOLUCIÓN

Descripción del comportamiento temporal de la DS

Tabla 46

Estadísticos descriptivos de la DS por año

Año	2013	2014	2015	2016	2017
n (meses)	12	12	12	12	12
Media	8542	7395	10783	11235	10712
Desviación estandar	1367.796	1370.521	1379.514	2168.535	1969.667
Quartil 1	7830	6976	9982	10246	9094
Mediana	8762	7664	10455	11112	10660
Quartil 3	9678	8243	11175	11638	11874
Coefficiente de variación	16.01259	18.53328	12.79332	19.30232	18.38663

Fuente. Departamento de Farmacia del Hospital Regional del Cusco (Anexo N° 01)

INTERPRETACIÓN

Se aprecia que no hay mucha variación en cuanto a las medidas de tendencia central como la media y la mediana entre los años 2013 y 2014, pero si una variación que se mantiene estable entre los años 2015 y 2017. Por otro lado, el coeficiente de variación indica que los datos dentro de cada año si presentan una relativa variación respecto a la media, dicho coeficiente de 2013 al 2107 no sobrepasan el 19.30%.

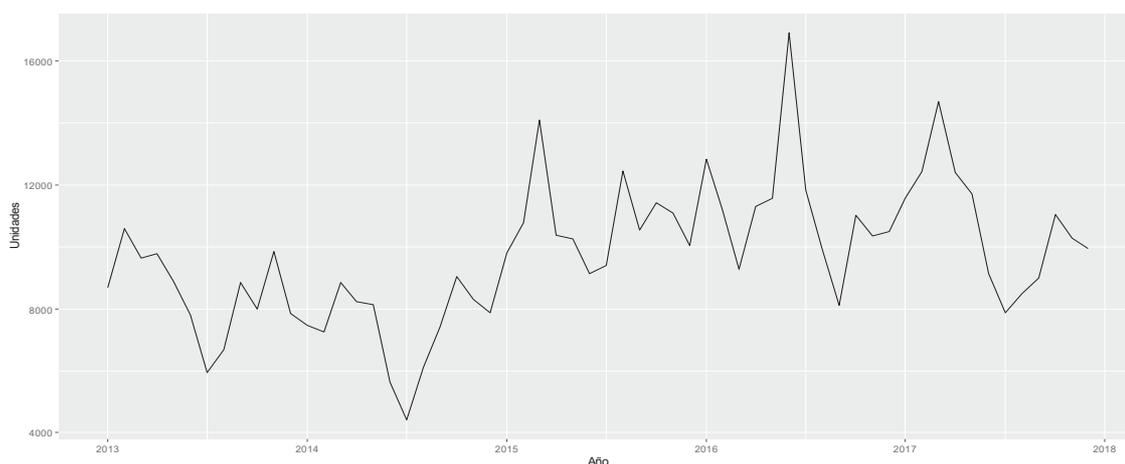


Figura 56. DS mensual de enero 2013 a diciembre de 2017

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se percibe que la serie presenta una estacionariedad, ausencia de tendencia positiva o negativa y probable estacionalidad.

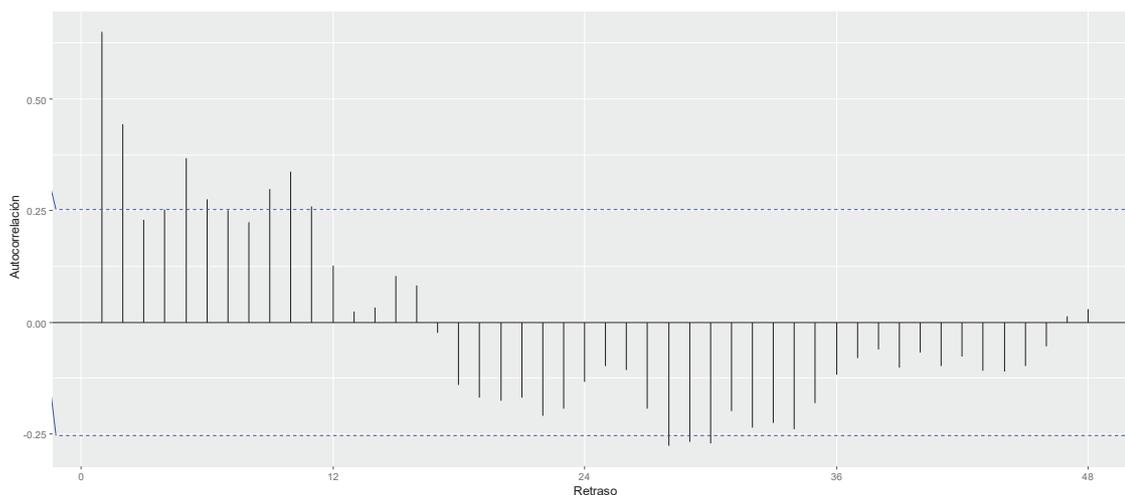


Figura 57. Función de autocorrelación de la DS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la autocorrelación una disminución lenta, forma festoneada irregular y picos significativos lo cual indica que la serie presenta tendencia positiva, estacionalidad y no estacionariedad, respectivamente.

Tabla 47

Estadísticos descriptivos de toda la serie de la DS

Año	2013 - 2017
n (meses)	60
Media	9733
Desviación estandar	2220.863
Quartil 1	8200
Mediana	9800
Quartil 3	11053
Coefficiente de variación	22.81689

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la serie entre los años 2013 al 2017 una media, mediana y coeficiente de variación de 9733, 9800 y 22.82% respectivamente.

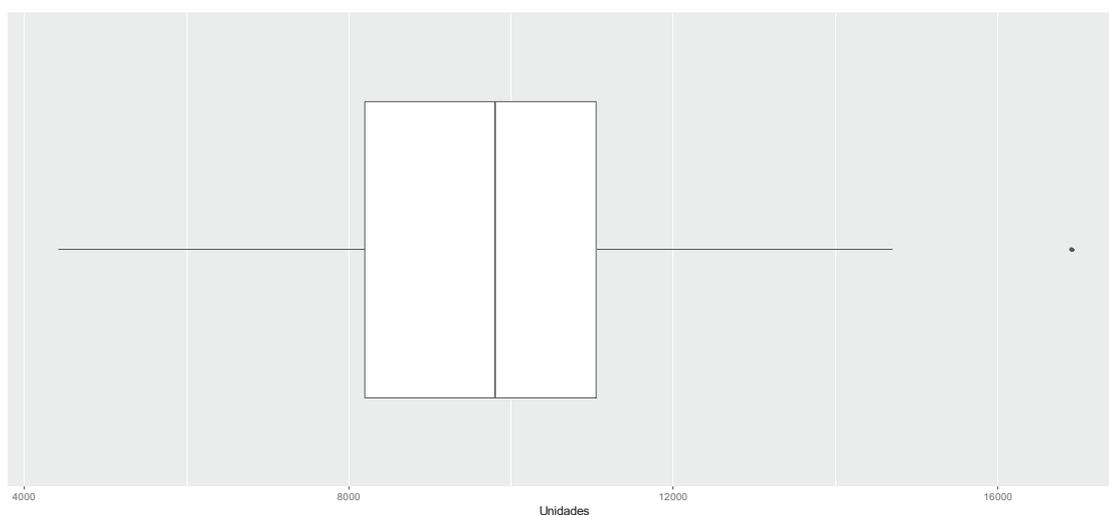


Figura 58. Diagrama de caja de la DS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa en la serie una distribución asimétrica con un sesgo ligero a la derecha y ausencia de datos atípicos.

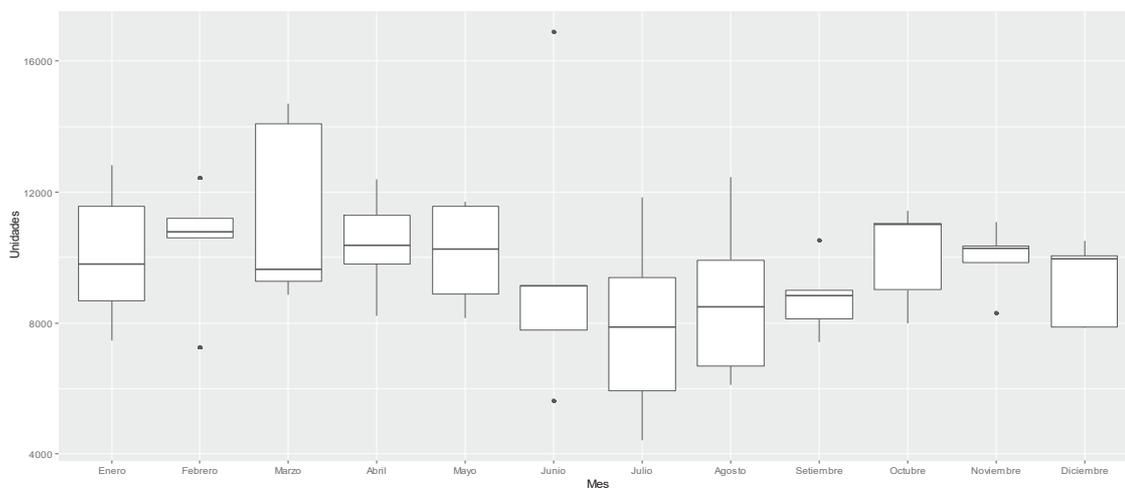


Figura 59. Diagrama de caja por mes de la DS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la serie datos atípicos en los meses de marzo, junio, setiembre y noviembre; una distribución concentrada en los meses de febrero, abril, mayo, junio, setiembre, octubre, noviembre y diciembre lo cual indica estacionalidad. Asimismo, una menor y mayor demanda del medicamento durante los meses de junio marzo, respectivamente.

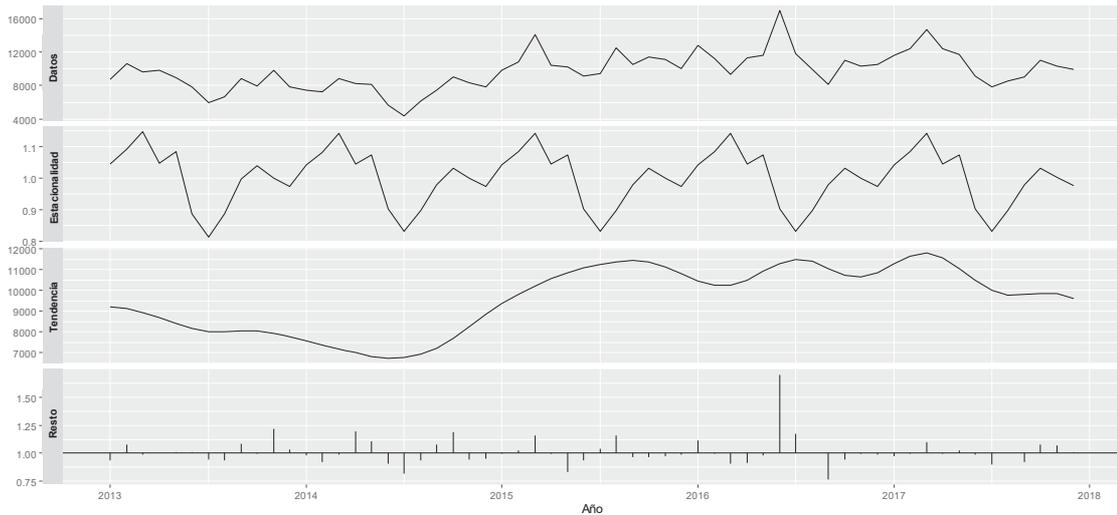


Figura 60 Descomposición de la DS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

En la parte superior la serie original, luego la componente de estacionalidad donde se aprecia una aparente estacionalidad, después la componente de la tendencia donde se distingue una aparente tendencia positiva y finalmente el componente del resto.

Tabla 48**Pruebas de estacionariedad, tendencia, estacionalidad y normalidad de la DS**

Prueba					
KPSS		Mann-Kendall	Kruskal-Wallis		Shapiro-Wilk
p	valor crítico 5%	p	Año	p	p
			2013 a 2018	1.864E-06	
			2013 - 2014	0.07349	
			2013 - 2015	0.0008121	
			2013 - 2016	0.0008121	
			2013 - 2017	0.006657	
0.7228	0.463	9.242E-05	2014 - 2015	3.226E-05	0.484
			2014 - 2016	0.0001096	
			2014 - 2017	0.0002755	
			2015 - 2016	0.4884	
			2015 - 2017	0.8174	
			2016 - 2017	0.6033	

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se percibe que la serie presenta no estacionariedad, tendencia y no estacionalidad lo cual se verifica con las pruebas de KPSS ($p > 0.463$), Mann-Kendall ($p < 0.05$) y Kruskal-Wallis ($p < 0.05$), respectivamente.

Respecto a la estacionalidad si bien la prueba de Kruskal-Wallis indica que la serie para los años 2103 a 2017 no presenta estacionalidad, pero si presenta estacionalidad para los años 2013 y 2014; 2015 y 2016; 2015 y 2017; 2016 y 2017.

Se observa que la prueba de normalidad de Shapiro-Wilk no es significativa ($p > 0.05$) lo cual evidencia homogeneidad de la serie.

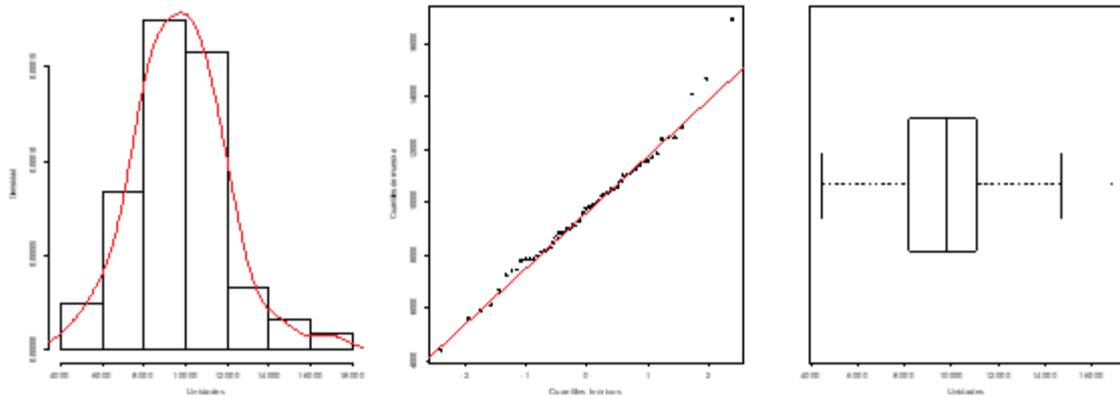


Figura 61. Evaluación de la normalidad de la DS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Los gráficos señalan que los datos de la serie siguen un perfil de campana de Gauss, lo cual se verifica en el gráfico del histograma una distribución simétrica, el gráfico de la distribución de probabilidad línea recta se verifica que sigue una línea recta y el diagrama de la caja es asimétrico con un ligero sesgo a la derecha.

Identificación del modelo (selección tentativa de p,d,q)

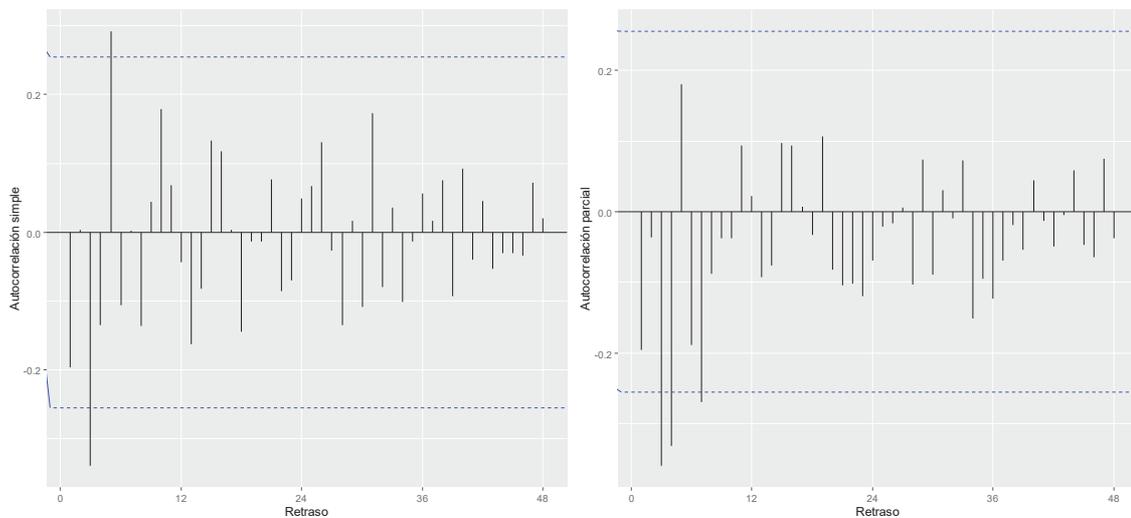


Figura 62. Función de autocorrelación simple y parcial de la serie con diferenciación de la DS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se advierte en la función de autocorrelación simple el primer rezago por defecto es siempre altamente significativo, por otro lado, se tiene dos picos significativos en la banda superior e inferior, asimismo se observa que los rezagos no siguen un patrón de descenso; respecto a la función de autocorrelación parcial se tiene que el primer rezago por defecto es siempre altamente significativo, a su vez se observa tres picos significativos, además se percibe que los rezagos no siguen un patrón de descenso. Del gráfico se propone un modelo con componentes AR(3) y MA(2).

Tabla 49
Identificación del modelo para la DS

Modelos	Medidas de ajuste					Pruebas	
	AIC	Entrenamiento		Pronóstico		Ljung-Box	Shapiro-Wilk
		RMSE	MAPE	RMSE	MAPE		
ARIMA(3,1,2)	1052.09	1603.227	13.04459	1225.411	10.54432	0.1224	0.07156
ARIMA(3,1,1)	1050.09	1603.26	13.04734	1227.051	10.56798	0.1806	0.07371
ARIMA(1,1,2)	1053.14	1675.749	13.23943	1266.069	9.7614	0.1273	0.01538
ARIMA(2,1,2)	1053.2	1645.426	13.10762	1226.735	9.395984	0.1209	0.0161

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que de los diferentes modelos propuestos el mejor modelo es el ARIMA(3,1,2) debido a que satisface los requisitos de normalidad e independencia de los residuales del modelo ($p > 0.05$), el MAPE del pronóstico es el más bajo; a pesar de que el AIC no es el más bajo (1052.09).

Estimación de los parámetros del modelo elegido

Tabla 50
Modelo ARIMA para la DS

Modelo	Parámetros				Medición del error			
	Coeficiente	Estimado	Error estándar	AIC	Entrenamiento		Pronóstico	
					RMSE	MAPE	RMSE	MAPE
ARIMA(3,1,2)	ar1	0.3168	0.2661	1052.09	1603.227	13.04459	1225.411	10.54432
	ar2	0.0088	0.1907					
	ar3	-0.3333	0.1379					
	ma1	-0.6659	0.2724					
	ma2	-0.0081	0.2185					

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que el modelo para la DS, es un ARIMA compuesto por una parte regular que contiene una diferencial de orden uno, una AR de orden tres y un MA de orden dos.

En cuanto a los parámetros se tiene un ar1 con 0.3168, un ar2 con 0.0088, un ar3 con -0.33333, una ma1 con -0.6659, una ma2 con -0.0081 y un AIC de 1052.09.

En relación a la medición del error se tiene que en el entrenamiento y pronóstico un RMSE de 1603.227 y 1225.411 respectivamente y un MAPE de 13.04459 y 10.54432 respectivamente.

Respecto al modelo estimado se tiene un modelo ARIMA (3,1,2) cuya ecuación general es $Y_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + (\varphi_1 + 1)Y_{t-1} + (\varphi_2 - \varphi_1)Y_{t-2} + (\varphi_3 - \varphi_2)Y_{t-3} - \varphi_3Y_{t-4}$ (Anexo N° 02) al reemplazar los parámetros estimados $\varphi_1 = 0.3168$, $\varphi_2 = 0.0088$, $\varphi_3 = -0.3333$, $\theta_1 = -0.6659$ y $\theta_2 = -0.0081$ se obtiene la siguiente ecuación del modelo ARIMA como sigue $Y_t = \varepsilon_t - 0.6659\varepsilon_{t-1} - 0.0081\varepsilon_{t-2} + 1.3168Y_{t-1} - 0.308Y_{t-2} - 0.3421Y_{t-3} + 0.3333Y_{t-4}$

Validación del modelo

Tabla 51
Pruebas de normalidad de los residuos y comportamiento de ruido blanco del modelo ARIMA (3,1,2) de la DS

Prueba	p
Shapiro-Wilk	0.07156
Ljung-Box	0.1224

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se aprecia que las pruebas de normalidad de Shapiro-Wilk no es significativa ($p > 0.05$), con lo cual se evidencia la normalidad de los residuos. Respecto a la prueba de Ljung-Box se aprecia que no es significativa ($p > 0.05$), señal de que los residuos son independientes, es decir tienen un comportamiento de ruido blanco.

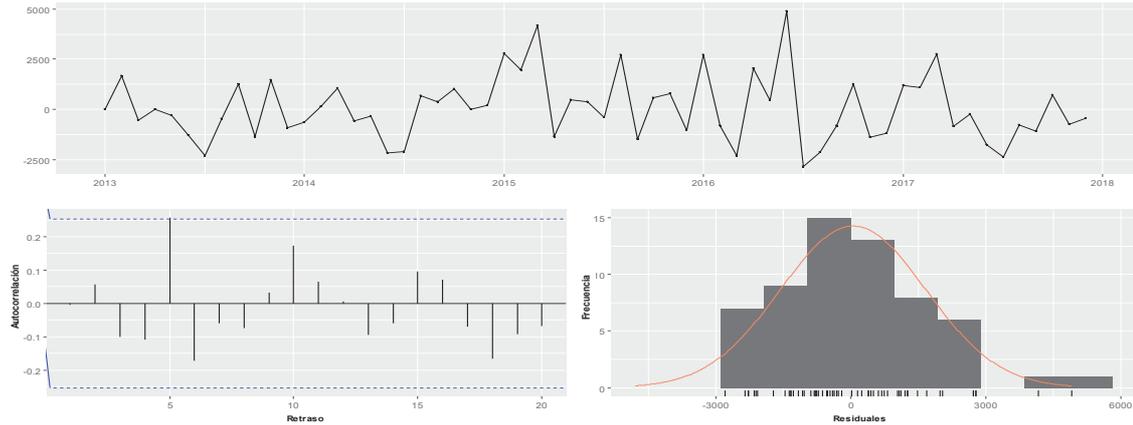


Figura 63. Evaluación de la normalidad y comportamiento de ruido blanco de los residuos del modelo ARIMA (3,1,2) de la DS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

En la parte superior la serie de los residuos y en la parte inferior la función de autocorrelación donde se tiene un pico significativo de lo cual se verifica que no afecta al comportamiento de ruido blanco (prueba Ljung-Box, $p > 0.05$) y el histograma sigue un perfil de campana de Gauss.

Pronóstico

Tabla 52
Pronóstico del modelo ARIMA para la DS

	Modelo	RMSE	MAPE
ARIMA(3,1,2)	Entrenamiento	1603.227	13.04459
	Pronóstico	1225.411	10.54432

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que la medición del error del modelo ARIMA(3,1,2) del entrenamiento con un RMSE y MAPE de 1603.227 y 13.04% respectivamente; del pronóstico con RMSE y MAPE de 1225.411 y 10.54% respectivamente.

Tabla 53
Valores reales y pronosticados de la DS

Meses	Año (2018)		Límite de confianza 95%	
	Real	Pronóstico	Inferior	Superior
Enero	10289	9441.938	6273.154	12610.72
Febrero	10331	9533.797	5752.886	13314.71
Marzo	10123	9667.043	5515.555	13818.53
Abril	11444	9880.739	5694.493	14066.98
Mayo	12380	9918.991	5699.154	14138.830
Junio	8348	9888.579	5621.259	14155.9
Julio	7957	9808.055	5411.158	14204.95
Agosto	8611	9769.528	5212.52	14326.54
Setiembre	9445	9766.751	5051.77	14481.73
Octubre	10380	9792.371	4956.62	14628.12
Noviembre	10292	9813.304	4878.252	14748.360
Diciembre	10193	9821.086	4795.179	14846.99

Fuente. Elaboración propia 2019 (Anexo N° 03)

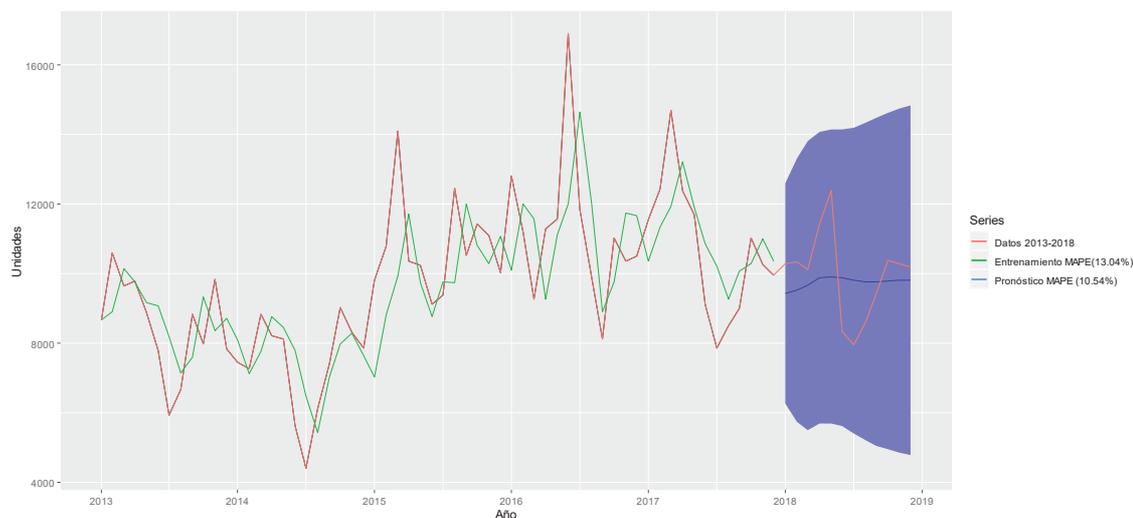


Figura 64. Valores del entrenamiento y pronóstico del modelo ARIMA (3,1,2) para la DS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se aprecia las series de los datos 2013 a 2018, del entrenamiento y del pronóstico con un MAPE de 13.04% y 10.54% respectivamente, el pronóstico tiene un nivel de confianza de 95%.

Tabla 54
Estimación del modelo de la RNNLSTM de la DS

Modelo	Parámetros					Medición del error			
	Retraso	Grado de aprendizaje	Función de activación	Número de época	Nodo	Entrenamiento		Pronóstico	
						RMSE	MAPE	RMSE	MAPE
RNNLSTM	12	0.05	Sigmoidea	1000	1	2659.37	20.47254	1781.917	14.11608
					2	2411.128	19.73141	1421.486	11.92229
					3	2097.627	17.69697	1332.115	11.3578
					4	2209.8	18.57506	1209.269	10.19059
					5	2054.603	16.98197	1462.687	12.50975
			6		2125.694	17.97049	1228.552	9.539383	
			7		2068.807	17.68227	1361.402	10.69559	
			8		2097.385	16.74341	2751.693	23.69987	
			9		2285.429	19.23876	1233.989	10.36628	
			10		2271.769	18.64828	1360.615	11.76676	

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se tiene diferentes modelos de la red neuronal tipo LSTM, donde difieren únicamente en el número de nodos, es así que el mejor modelo presenta los siguientes parámetros retrasos con 12, grado de aprendizaje con 0.05, función de activación es de tangente hiperbólica, número de épocas con 1000 y número de nodos con 6, con el mencionado modelo en cuanto a la medición de los errores se tiene que en el entrenamiento y pronóstico un RMSE de 2125.694 y 11228.552 respectivamente y un MAPE de 17.97% y 9.53% respectivamente. El modelo estimado queda expresado como $y_{12x1} = \sum_{p=1}^6 h_{\text{Bloque memoria}_p}$ (Anexo N° 04).

Tabla 55
Valores reales y pronósticos de la DS

Meses	Año (2018)	
	Real	Pronóstico
Enero	10289	9840.27
Febrero	10331	10090.98
Marzo	10123	10099.57
Abril	11444	9923.14
Mayo	12380	9964.29
Junio	8348	10146.4
Julio	7957	10067.25
Agosto	8611	9792.32
Setiembre	9445	9345.13
Octubre	10380	9775.73
Noviembre	10292	9820.83
Diciembre	10193	10038.61

Fuente. Elaboración propia 2019 (Anexo N° 03)

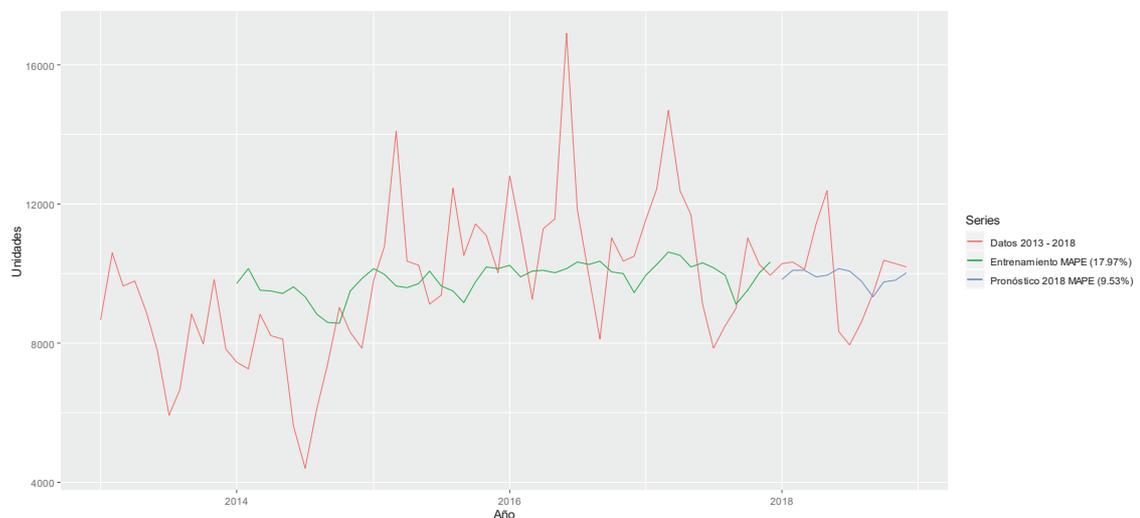


Figura 65. Valores del entrenamiento y pronóstico del modelo de la RNNLSTM para la DS

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se aprecia las series de los datos 2013 a 2018, del entrenamiento y del pronóstico con un MAPE de 17.97% y 9.53% respectivamente.

Tabla 56
Pronóstico de los modelos para la DS

Medición del error		Modelo	
		ARIMA	RNNLSTM
RMSE	Entrenamiento	1603.227	2125.694
	Pronóstico	1225.411	1228.552
MAPE	Entrenamiento	13.04459	17.97049
	Pronóstico	10.54432	9.53938

Fuente. Elaboración propia 2019 (Anexo N° 03)

INTERPRETACIÓN

Se observa que la medición del error de los modelos ARIMA y RNNLSTM del entrenamiento con un RMSE de 1603.227 y 2125.694 y con un MAPE de 13.04% y 17.97% respectivamente; del pronóstico con un RMSE de 1225.411 y 1228.552 y con un MAPE de 10.54% y 9.53% respectivamente.

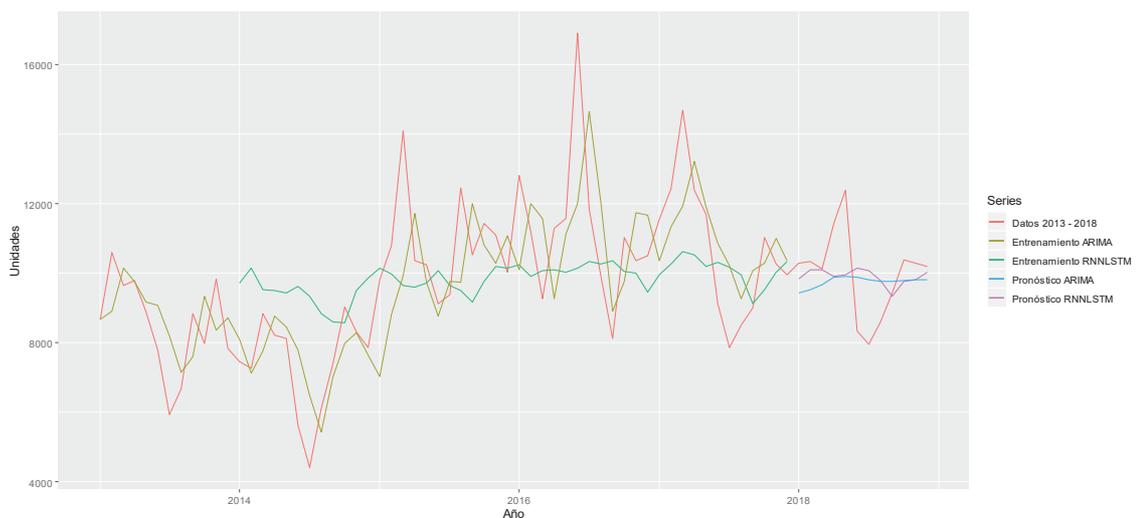


Figura 66. Serie con datos reales, entrenamiento y pronóstico de los modelos ARIMA y RNNLSTM

Fuente. Elaboración propia 2019 (Anexo N° 03)

Tabla 57***Pronóstico de los modelos ARIMA y RNNLSTM en función del MAPE***

MEDICAMENTO	MAPE			
	ARIMA	RNNLSTM	REDUCCIÓN	PROMEDIO
DCS	4.5	4.78	0.28	0.28
DI	27.57	25.00	2.57	
DMS	4.66	1.85	2.81	1.92
DP	23.47	22.19	1.28	
DS	10.54	9.53	1.01	

INTERPRETACIÓN

Se percibe que, para el pronóstico de la DCS el mejor modelo es el ARIMA con un MAPE de 4.50; y para el pronóstico de la DI, DMS, DP y DS el mejor modelo es la RNNLSTM con un MAPE de 25.00, 1.85, 22.19 y 9.53 respectivamente.

Asimismo, se observa que el mejor modelo con capacidad de pronóstico de la demanda de medicamentos es la RNNLSTM frente al modelo ARIMA con una reducción del error en promedio de 1.92 y 0.28 respectivamente.

Tabla 58***Pronóstico de los modelos ARIMA y RNNLSTM en función del RMSE***

MEDICAMENTO	RMSE			PROMEDIO
	ARIMA	RNNLSTM	REDUCCIÓN	
DCS	477.6972	642.4783	164.78	
DP	2818.81	2870.549	51.74	73.22
DS	1225.411	1228.552	3.14	
DI	1778.328	1646.159	132.17	
DMS	664.4076	293.709	370.70	251.43

INTERPRETACIÓN

Se percibe que, para el pronóstico de la DCS, DP y DS el mejor modelo es el ARIMA con un RMSE de 477.69, 2818.81, 1225.41; y para el pronóstico de la DI, y DMS el mejor modelo es la RNNLSTM con un RMSE de 1646.159 y 293.709 respectivamente.

Asimismo, se observa que el mejor modelo con capacidad de pronóstico de la demanda de medicamentos es la RNNLSTM frente al modelo ARIMA con una reducción del error en promedio de 251.43 y 73.22 respectivamente.

DISCUSIÓN

En cuanto al Cloruro de sodio 900 mg/1000 ml inyectable la serie presenta no estacionariedad, tendencia, no estacionalidad y heterogeneidad (Tabla 4); a su vez se tiene los modelos ARIMA (0,1,1) con una RMSE y MAPE en el entrenamiento con 485.23 y 5.11 y en el pronóstico con 477.6972 y 4.50 respectivamente; y el modelo de la RNNLSTM con una RMSE y MAPE en el entrenamiento con 582.4484 y 6.27 y en el pronóstico con 642.4783 y 4.78 respectivamente (Tabla 12). De lo antes mencionado se tiene que el mejor modelo es el ARIMA (RMSE 477.6972 y MAPE 4.50) frente al modelo de la RNNLSTM (RMSE 642.4783 y MAPE 4.78).

En relación al Ibuprofeno 400 mg tableta la serie presenta estacionariedad, sin tendencia, estacionalidad y homogeneidad (Tabla 15); sin embargo, del análisis del gráfico de la función de autocorrelación simple y parcial indica un modelo ARIMA solo con un componente regular (Figura 29); es así que se tiene los modelos ARIMA (1,0,1) con una RMSE y MAPE en el entrenamiento con 1043.714 y 12.71 y en el pronóstico con 1778.328 y 27.57 respectivamente; y el modelo de la RNNLSTM con una RMSE y MAPE en el entrenamiento con 1111.551 y 12.71 y en el pronóstico con 1646.159 y 25.00 respectivamente (Tabla 23). Asimismo, se evidencia que el mejor modelo es la RNNLSTM (RMSE 1646.159 y MAPE 25.00) frente al modelo ARIMA (RMSE 1778.328 y MAPE 27.57), se puede apreciar que el MAPE y el RMSE en los dos modelos es bastante alto lo cual se puede deber a que los registros sobre la demanda del medicamento no sean reales, es decir debido a un desabastecimiento del medicamento o a un acceso restringido hacia el paciente ambulatorio en beneficio del suministro del medicamento a los pacientes internados.

A cerca del Metamizol sódico 1 g/2ml inyectable la serie presenta no estacionariedad, tendencia, sin estacionalidad y homogeneidad (Tabla 26); sin embargo, del análisis del gráfico de la función de autocorrelación simple y parcial

indica un modelo ARIMA con un componente regular y estacional (Figura 18); es así que se tiene los modelos ARIMA $(3,1,1)(0,0,1)_{12}$ con un RMSE y MAPE en el entrenamiento con 1045.16 y 7.09 y en el pronóstico con 664.40 y 4.66 respectivamente; y el modelo de la RNNLSTM con un RMSE y MAPE en el entrenamiento con 1115.495 y 7.99 y en el pronóstico con 293.7089 y 1.85 respectivamente (Tabla N° 34). Se verifica que el mejor modelo es la RNNLSTM (RMSE 293.7089 y MAPE 1.85) frente al modelo ARIMA (RMSE 664.40 y MAPE 4.66).

Con respecto al Paracetamol 500 mg tableta la serie presenta no estacionariedad, tendencia, sin estacionalidad y heterogeneidad (Tabla 37); sin embargo, del análisis del gráfico de la función de autocorrelación simple y parcial indica un modelo ARIMA con un componente regular y estacional (Figura 51); por lo que se tiene los modelos ARIMA $(1,1,0)(0,0,1)_{12}$ con un RMSE y MAPE en el entrenamiento con 1522.119 y 12.46 y en el pronóstico con 2818.81 y 23.47 respectivamente; y el modelo de la RNNLSTM con un RMSE y MAPE en el entrenamiento con 2122.731 y 22.76 y en el pronóstico con 2870.549 y 22.19 respectivamente (Tabla 45). Se confirma que el mejor modelo es la RNNLSTM (MAPE 22.19) frente al modelo ARIMA (MAPE 23.47); se puede percibir que el MAPE en los dos modelos es bastante alto lo cual se puede deber a que los registros sobre la demanda del medicamento no sean reales, es decir debido a un desabastecimiento del medicamento o a un acceso restringido hacia el paciente ambulatorio en beneficio del suministro del medicamento a los pacientes internados; sin embargo, tomando en cuenta la RMSE se tiene que el mejor modelo es el ARIMA (RMSE 2818.81) frente a la RNNLSTM (RMSE 2870.549).

Referente al Sevoflurano 250 ml solución inhalatoria la serie presenta no estacionariedad, tendencia, sin estacionalidad y homogeneidad (Tabla 48), a su vez se tiene los modelos ARIMA $(3,1,2)$ con un RMSE y MAPE en el entrenamiento con 1603.227 y 13.04 y en el pronóstico con 1225.411 y 10.54 respectivamente; y el modelo de la RNNLSTM con un RMSE y MAPE en el entrenamiento con 2125.694 y 17.97 y en el pronóstico con 1228.552 y 9.53 respectivamente (Tabla 56). Se constata que el mejor modelo es la RNNLSTM

(MAPE 9.53) frente al modelo ARIMA (MAPE 10.54); se observa que el MAPE en los dos modelos es bastante alto es probable que ocurra debido a que los registros sobre la demanda del medicamento no sean reales, es decir debido a un desabastecimiento del medicamento o a un acceso restringido hacia el paciente ambulatorio en beneficio del suministro del medicamento a los pacientes internados; no obstante, considerando la RMSE se tiene que el mejor modelo es el ARIMA (RMSE 1225.411) frente a la RNNLSTM (RMSE 1228.552).

Por otro lado, en los medicamentos como el paracetamol y el sevoflurano para el modelo ARIMA se tiene un MAPE alto y una RMSE baja y para el modelo de la RNNLSTM se tiene un MAPE bajo y una RMSE alta, lo cual no es coherente es decir si el MAPE es alto la RMSE también debe ser alta o viceversa, este inconveniente no se presenta en los otros medicamentos, al respecto Armstrong (2001, pág. 717), indica que no se debe utilizar la RMSE para realizar comparaciones entre métodos de pronóstico debido a que es poco confiable, motivo por el cual se toma en cuenta solo al MAPE.

Asimismo, se tienen los estudios realizados por Kaushik et al. (2017), cuyo objetivo fue predecir el gasto mensual en medicamentos en los EE.UU. entre los años 2011 y 2015, donde se concluyó que el mejor modelo es la RNNLSTM apilado, seguidamente de la RNNLSTM estándar y finalmente el ARIMA; el estudio realizado por Li y Cao (2018), donde el objetivo fue predecir el flujo de turismo del 2013 al 2015 del museo de Xi'an de la ciudad de Shaanxi China, se concluyó que el mejor modelo es la RNNLSTM simple frente a los modelos ARIMA, red neuronal BPN y RNNLSTM apilado; y el estudio efectuado por Li et al. (2017), cuyo objetivo fue proponer un nuevo modelo de RNNLSTM extendido que considera intrínsecamente las correlaciones espacio temporales para la predicción de la concentración de contaminantes en el aire de la ciudad de la ciudad de Beijing de China del 01/01/2014 al 28/05/2016, se concluyó que el mejor modelo es la RNNLSTM extendida frente a los modelos RNNLSTM, ARMA, STDL, TDNN Y SVR.

Considerando los resultados del presente estudio y los antecedentes antes indicados se evidencia una similitud en cuanto a los resultados obtenidos, que el mejor pronóstico lo tiene el modelo de la red neuronal.

Finalmente se resuelve que el modelo de la RNNLSTM permite predecir adecuadamente la demanda de medicamentos para el año 2018 frente al modelo ARIMA con una reducción del MAPE en promedio de 1.92 y 0.28 respectivamente (Tabla 57).

CONCLUSIONES

La red neuronal recurrente tipo de memoria a corto y largo plazo permite predecir adecuadamente la demanda de medicamentos en el Hospital Regional del Cusco frente al modelo ARIMA con una reducción del MAPE en promedio de 1.92 y 0.28 respectivamente.

Los modelos autorregresivo integrado de media móvil que proporcionan el mejor ajuste para el pronóstico de la demanda de medicamentos son los siguientes: Cloruro de sodio 900 mg/1000 ml inyectable con un ARIMA (0,1,1) y un MAPE de 4.50; Ibuprofeno 400 mg tableta con un ARIMA (1,0,1) y un MAPE de 27.57; Metamizol sódico 1g/2 ml inyectable con un ARIMA (3,1,1)(0,0,1)₁₂ y un MAPE de 4.66; Paracetamol 500 mg tableta con un ARIMA (1,1,0)(0,0,1)₁₂ y un MAPE de 23.47; y Sevoflurano 250 ml solución con un ARIMA (3,1,2) y un MAPE de 10.54.

Los modelos red neuronal recurrente tipo de memoria a corto y largo plazo que proporcionan el mejor ajuste para el pronóstico de la demanda de medicamentos son los siguientes: Cloruro de sodio 900 mg/1000 ml inyectable con $y_{12x1} = \sum_{p=1}^7 h_{\text{Bloque memoria}_p}$ y un MAPE de 4.78; Ibuprofeno 400 mg tableta con $y_{12x1} = \sum_{p=1}^7 h_{\text{Bloque memoria}_p}$ y un MAPE de 25.00; Metamizol sódico 1g/2 ml inyectable con $y_{12x1} = \sum_{p=1}^6 h_{\text{Bloque memoria}_p}$ y un MAPE de 1.85; Paracetamol 500 mg tableta con $y_{12x1} = \sum_{p=1}^7 h_{\text{Bloque memoria}_p}$ y un MAPE de 22.19; y Sevoflurano 250 ml solución con $y_{12x1} = \sum_{p=1}^6 h_{\text{Bloque memoria}_p}$ y un MAPE de 9.53.

RECOMENDACIONES

Al Director de la Escuela de Posgrado de la UNSAAC incluir en el plan de estudios de la maestría en Estadística un curso relacionado a redes neuronales artificiales para afianzar conocimientos e incursionar en nuevas líneas de investigación por parte de los estudiantes.

Al Jefe del departamento de Farmacia del Hospital Regional del Cusco implementar un programa piloto para realizar el pronóstico de medicamentos utilizando los modelos ARIMA y red neuronal recurrente con la finalidad de disminuir el desabastecimiento de medicamentos y que los pacientes tengan acceso a los mismos en el momento oportuno.

A los estudiantes de la Escuela de Posgrado de la UNSAAC continuar el seguimiento de estudios con la misma metodología comparativa con ARIMA y red neuronal recurrente, de los siguientes años para tener una visión clara del comportamiento del pronóstico de la demanda de medicamentos en el Hospital Regional del Cusco con el fin de que el responsable del departamento de Farmacia tome decisiones acertadas.

REFERENCIAS

Ayyadevara, V. (2019). *Neural Networks with Keras Cookbook* (1st ed.). Birmingham, UK: Packt Publishing Ltd. ISBN 978-1-78934-664-0.

Armstrong, J. Scott (2001). *Principles of Forecasting: A Handbook for Researchers and Practitioners* (1st ed.). New York, USA: Published by Kluwer Academic. ISBN 0792379306.

Adamowski Jan, Chan Hiu Fung, Prasher Shiv O., Ozga-Zielinski Bogdan, Sliusarieva Anna (2012). *Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban wáter demand forecasting in Montreal, Canada*. American Geophysycal Union, 48 1-14. DOI: 10.1029/2010WR009945.

Agirrezabala, J. R., Aizpurua, I., Albizuri, M., Alfonso, I., Armendáriz, M., Barrondo, S., Barroso, J., Domingo, S., Etxeberria, A., Fernández, J., Giménez, A. I., Gómez, N., Ibarra, E., Iglesias, J. J., Iribar, J., Jaio, N., Lekue, I., López, J., Martínez, J., Mendizabal, A., Mozo, C., Olloquiegi, E., Ruiz de Velasco, E., Sainz de Rozas, R., Valverde, E. (2015). *Desabastecimiento de medicamentos un problema sin resolver*. Información Farmacoterapéutica de la Comarca, 23 (7), 47-52. ISSN:1575054-X.

Arnau, J. (2001). *Diseño de Series Temporales Técnica de Análisis*. España: Ediciones Universidad de Barcelona. ISBN 84-8338-250-4.

Aguirre, A. (1994). *Introducción al Tratamiento de Series Temporales: Aplicación a las Ciencias de la Salud*. Madrid, España: Ediciones Díaz de Santos S.A..ISBN 84-7978-153-X.

Auguie, B., Antonov, B. (2017). *Package 'gridExtra'. Versión 2.3*. Recuperado de <https://cran.r-project.org/web/packages/gridExtra/gridExtra.pdf>.

Basu, A., Basu S. (2016). *A User's Guide to Business Analytics* (1st ed.). Boca Raton, Florida, USA: Published by CRC Press Taylor & Francis Group. ISBN 13: 978-1-4665-9165-3.

Bisgaard, S., Kulahci, M. (2011). *Times series analysis and forecasting by example*. Hoboken, New Jersey, USA: John Wiley & Sons Inc. Publishing. ISBN 978-1-118-05694-3.

Bianchi, F.M., Maiorino, E., Kampffmeyer, M.C., Rizzi, A., Jenssen, R. (2017). *Recurrent Neural Networks for Short-Term Load Forecasting: An Overview and Comparative Analysis* (1st ed.). Cham, Switzerland: Published by Springer Nature. ISBN 978-3-319-70337-4.

Box G., Jenkins G. (1976). *Time Series Analysis Forecasting and Control* (1st ed.). California, USA: Published by HOLDEN-DAY. ISBN 0-8162-1104-3.

Box, G., Jenkins, G., Reinsel, G., Ljung, G. (2016). *Time Series Analysis Forecasting and Control* (5th ed.). Hoboken, New Jersey, USA: John Wiley & Sons Publishing. ISBN 9781118675021.

Bokka, K. R., Hora, S., Jain, T., Wambugu, M. (2019). *Deep Learning for Natural Language Processing: Solve your natural language processing problems with smart deep neural networks* (1st ed.). Birmingham, UK: Published by Packt Publishing Ltd.. ISBN 98-1-83855-029-5.

Brooks, C. (2008). *Introductory Econometrics for Finance* (2nd ed.). New York, USA: Cambridge University Press Publishing. ISBN 13 978-0-521-69468-1.

Brockwell, P., Davis, R. (2016). *Introduction to Time Series and Forecasting* (3rd ed.). Switzerland: Springer International Publishing. ISBN 978-3-319-29854-2.

Cajiao, S., Melo, L.F., Parra, D. (2014). *Pronósticos para una economía menos volátil: el caso colombiano*. Coyuntura Económica: Investigación Económica y Social 44 (2) 35-59. Recuperado de https://www.researchgate.net/profile/Luis_Melo-Velandia/publication/281650848_Pronosticos_para_una_economia_meno

s_volatil_el_caso_colombiano/links/55f3034f08ae7a10cf876cb7/Pronosticos-para-una-economia-menos-volatil-el-caso-colombiano.pdf

Cordero, M. (2006). *Políticas que promueven el acceso a los medicamentos en América Latina Informe sobre Bolivia, Colombia, Chile, Ecuador y Perú*. Lima, Perú: Editorial Nova Print SAC. ISBN 9972-804-61-5.

Constitución Política del Perú 1993. Diario Oficial de la República de Perú, Lima, Perú, 30 de diciembre de 1993.

Ciaburro, G., Venkateswaran, B. (2017). *Neural Network with R*. (1ts ed.). Birmingham, UK: Pack Publishing Ltd. ISBN 978-1-78839-787-2.

Chatfield, C. (2003). *The Analysis of Time Series An Introduction* (6th ed.). Boca Raton, London New York Washington, D.C., USA: Chapman and Hall/CRC. ISBN 1584883170.

Chen, G. (2018). *A Gentle Tutorial of Recurrent Neural Network with Error Backpropagation*. arXiv: 1610.02583v3. Recuperado de <https://arxiv.org/abs/1610.02583v3>.

Dirección General de Medicamentos Insumos y Drogas (DIGEMID) (2018a). *Evaluación de Indicadores de Disponibilidad de Medicamentos a Nivel Nacional*. Recuperado de <http://www.digemid.minsa.gob.pe/Main.asp?Seccion=447&SeccionCategoria=17>.

Dirección General de Medicamentos Insumos y Drogas (DIGEMID) (2018b). *Indicadores de Disponibilidad de Medicamentos Esenciales en los EESS del 1° Nivel de Atención (C.S, P.S, y hospitales no ejecutores) a Nivel Nacional*. Recuperado de <http://www.digemid.minsa.gob.pe/Main.asp?Seccion=447&SeccionCategoria=17>.

Dirección General de Medicamentos Insumos y Drogas (DIGEMID) (2018c). *Indicadores de Disponibilidad de Medicamentos Esenciales en los EESS del 3°*

Nivel de Atención (Hospitales, Instituto ejecutores) a Nivel nacional. Recuperado de <http://www.digemid.minsa.gob.pe/main.asp?Seccion=447&SeccionCtaegoria=33&pag=4>.

Dot CSV. (3 de octubre 2018). ¿Qué es una Red Neuronal? Parte 3 : Backpropagation | DotCSV [Video]. Youtube. https://www.youtube.com/watch?v=eNlqz_noix8

Fox, E.R., Sweet, B.V., Jensen, V. (2014). *Drugs shortages: a complex health care crisis.* Mayo Clinic Proceedings, 89 (3) 361-373. DOI: <https://doi.org/10.1016/j.mayocp.2013.11.014>.

Fichou, D. (2019). *LSTM units.* Obtenido de: *The R Project for Statistical Computing*: https://cran.r-project.org/web/packages/rnn/vignettes/LSTM_units.html.

Fu, R., Zhang, Z., Li, Li. (2016). *Predicción del flujo de tráfico utilizando los métodos de red neuronal LSTM y GRU.* 31ra Conferencia Anual Académica Juvenil de la Asociación China de Automatización, Wuhan, China. Recuperado de <https://sci-hub.tw/https://ieeexplore.ieee.org/document/7804912/>.

Florescu, I. (2015). *Probability and Stochastic Processes.* Hoboken, New Jersey, USA: Published by Jhon Wiley & Sons. ISBN 978-0-470-62455-5.

García, P.L., Río, J.C., Coto, J. (2000). *Análisis y Predicción de la Demanda en Líneas de Distribución Utilizando Redes Neuronales Artificiales.* Centro de Información Tecnológica (CIT) Chile, 11 (4) 67-73. ISSN 0716-8756.

Géron, A. (2017). *Hands-On Machine Learning with Scikit-Learn & Tensor Flow* (1st ed.). USA: Published by O'Reilly Media Inc. ISBN 978-1-491-96229-9

González, M. J. (2006). *Gestión Eficaz del Tiempo.* Málaga, España: Editorial Innovación y Cualificación S.L... ISBN 10:84-96493-12-1.

González, M.P. (2009). *Análisis de Series Temporales: Modelos ARIMA*. España: Editorial Sarrico-On. ISBN 978-84-692-3814-1.

González, E.Y. (2019). *Clásificación de Visualizaciones Motoras Utilizando Señales de un EEG, Basado en Algoritmos de Aprendizaje Profundo* (Tesis de Maestría). Universidad Autónoma de Chihuahua Facultad de Ingeniería, México.

Goldberg, Y. (2017). *Neural Network Methods for Natural Language Processing*. Morgan & Claypool Publishers. ISBN 9781627052986.

Gujarati, D., Porter, D. (2010). *Econometría* (5ta ed.). México: Editorial McGraw-Hill Interamericana. ISBN 978-607-15-0294-0.

Guidolin M., Pedio M. (2018). *Essential of Time Series for Financial Applications* (1st ed.). Chennai, India: Published by Elseiver. ISBN 978-0-12-813409-2.

Guerrero, V.M. (1993) *Time-series analysis supported by power transformations*. Journal of Forecasting, 12(1), 37–48.

Graupe, D. (2007). *Principles of Artificial Neural Networks* (2nd ed.). Singapore: World Scientific Publishing Co. Pte. Ltd. ISBN 10-981-270-624-0.

Ghatak, A. (2019). *Deep Learning with R* (1st ed.). Singapore: Springer Nature. ISBN 978-981-13-5850-0.

Hang, N. (2010). *Time Series Applications to Finance with R and Splus* (2nd ed.). Hoboken, New Jersey, USA: Wiley & Sons, Inc. Publishing. ISBN 978-0-470-58362-3.

Hanke, J. E., Wichern, D. W. (2006). *Pronósticos en los Negocios* (8va ed.). México: Editorial Pearson Educación. ISBN 970-26-0759-0.

Haykin, S. (2009). *Neural networks and Learning Machines* (3rd ed.). New Jersey, USA: Pearson Educations Inc. Publishing. ISBN 13: 978-0-13-147139-9.

Hammer, B., Frasco, M. LeDell, E. (2018). *Package Metrics. Versión 0.1.4*. Recuperado de <https://cran.r-project.org/web/packages/Metrics/Metrics.pdf>.

He, H., Tian, X. (2007). *An Improved Elman Network and Its Application in Flatness Prediction Modeling*. IEEE Second International Conference on Innovative Computing, Information and Control, 552- 555. DOI: 10.1109/ICICIC.2007.149.

Hipel, K.W., McLeod A.I. (1994). *Time Series Modelling of Water Resources and Environmental Systems* (1st ed.). Amsterdam, The Netherlands: Published by Elsevier Science. ISBN 0444892702.

Hyndman, R.J., Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). Melbourne, Australia: OTexts Publishing. ISBN 978-0-9875071-1-2.

Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., Yasmeeen, F., Ihaka, R., Reid, D., Shaub, D., Tang, Y., Zhou, Z. (2019). *Package 'forecast'*. Versión 8.5. Recuperado de <https://cran.r-project.org/web/packages/forecast/forecast.pdf>.

Hyndman, R., Khandakar, Y. (2008). *Automatic Time Series Forecastinf: The forecast Package for R*. Journal of Statistical Software, 27(3), 1-22.

Hernández, J., Zúñiga J. (2013). *Modelos Económicos para el Análisis Económico*. Madrid, España: Editorial ESIC. ISBN 9788473568920.

Instituto Nacional de Estadística e Informática (INEI) (2006). *Glosario básico de términos estadísticos*. Perú. Recuperado de https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitaless/Est/Lib0900/Libro.pdf.

Instituto Nacional de Estadística y Geografía (INEGI) (2012). *Glosario de Estadística Básica*. México. ISBN 978-607-494-348-1.

Kalman, B. L., Kwasny S.C. (1992). *Why Tanh: Choosing a Sigmoidal Function*. Baltimore, Maryland, USA: Published in International Joint Conference on Neural Networks. DOI: 10.1109/IJCNN.1992.227257.

Kaushik, S., Choudhury, A., Dasgupta, N., Natarajan, S., Pickett, L.A., Dutt, V. (2017). *Uso de LSTMs para predecir el gasto del paciente en Medicamentos*. Conferencia Internacional sobre Aprendizaje Automático y Ciencia de los Datos, Noida, India. Recuperado de <https://ieeexplore.ieee.org/document/8320269/>.

Kirchgässner, G., Wolters, J., Hassler, U. (2013). *Introduction to Modern Time Series Analysis* (2nd ed.). New York Dordrecht London: Springer-Verlag Berlin Heidelberg Publishing. ISBN 978-3-642-33436-8.

Khurram, M. (2010). *The effect of optimization of error metrics* (Master's thesis in informatics). School of business and informatics, University of Boras, Sweden.

Ley N° 26842 – Ley General de Salud. Diario Oficial de la República de Perú, Lima, Perú 15 de julio de 1997.

Lewis, N.D. (2017). *Neural Networks for Time Series Forecasting with R*. ISBN 13: 978-1544752952.

Lindsey, J.K.(2004). *Statistical Analysis of Stochastic Processes in Time*. New York, USA: Board, Cambridge University Press Publishing. ISBN 978-0-511-21194-2.

Li., Y., Cao, H. (2018). *Predicción del Flujo de Turismo basado en la Red Neuronal LSTM*. Procedia Computer Science. 129, 277-283. Recuperado de <https://www.sciencedirect.com/science/article/pii/S1877050918303016>.

Li, X., Peng, L., Yao, X., Cui, S., Hu, Y., You, C., Chi, T. (2017). *Red neuronal de memoria larga a corto plazo para contaminantes del aire predicciones de concentración: desarrollo y evaluación de métodos*. Environmental Pollution. 231, 997-1004. Recuperado de <https://www.sciencedirect.com/science/article/pii>

/S0269749117307534.

Mehrotra, K., Mohan, C.K., Ranka, S. (1996). *Element of Artificial Neural Networks*. USA: Massachusetts Institute of Technology Publishing. ISBN 0-262-13328-8.

Madsen, H. (2008). *Time Series Analysis* (1st ed.). New York, USA: Chapman and Hall/CRC Publishing. ISBN 13:978-1-4200-5967-0.

Mills, T.C. (2019). *Applied Time Series Analysis: A Practical Guide to Modeling and Forecasting* (1st ed.). San Diego, California, USA: Published by Academic Press. ISBN 978-0-12-813117-6.

Montgomery, D.C., Jennings, C.L., Kulahci, M. (2008). *Introduction to Time Series Analysis and Forecasting* (1st ed.). Hoboken, New Jersey, USA: Published by John Wiley & Sons. ISBN 978-0-471-65397-4.

Moritz, S. (2018). *Package imputeTS*. Versión 2.7. Recuperado de <https://cran.r-project.org/web/packages/imputeTS/imputeTS.pdf>.

Natinga, D. (2017). *Data Science Algorithms in a Week*. (1st ed.). Birmingham, UK: Packt Publishing Ltd. ISBN 978-1-78728-458-6.

Pattanayak, S. (2017). *Pro Deep Learning with TensorFlow a Mathematical Approach to Advanced Artificial Intelligence in Python* (1st ed.). Bangalore, Karnataka, India: Publishing Apress. ISBN 978-1-4842-3096-1.

Patterson, J., Gibson, A. (2017). *Deep learning a Practitioner's Approach*. (1st ed.). USA: O'Reilly Media Inc. Publishing. ISBN 978-1-491-91425-0.

Pham, D.T., Xing, L. (1995). *Neural Networks for Identification, Prediction and Control* (1st ed.). London, England: Published by Springer-Verlag London. ISBN 13:978-1-4471-3246-2.

Pham, D.T., Liu, X. (1996). *Training of Elman Networks and Dynamic System Modelling*. International Journal of Systems Science, 27(2)221-226. DOI 10.1080/00207729608929207.

Pfaff, B., Zivot, E., Stigler, M. (2016). *Package urca*. Versión 1.3-0. Recuperado de <https://cran.r-project.org/web/packages/urca/urca.pdf>.

Pérez, C., Santín, D. (2007). *Minería de Datos Técnicas y Herramientas* (2da ed.). Madrid, España: Editorial Paraninfo S.A. ISBN 978-84-9732-492-2.

Pérez, F. O. (2007). *Introducción a las Series de Tiempo Métodos Paramétricos* (1ra ed.). Colombia: Editorial Universidad de Medellín. ISBN 978-958-98010-7-9.

Pérez, C. (2006). *Problemas Resueltos de Econometría*. Madrid, España: Editorial Paraninfo S.A. ISBN 84-9732-376-9.

Pohlert, T. (2018). *Package trend*. Versión 1.1.1. Recuperado de <https://cran.r-project.org/web/packages/trend/trend.pdf>

Qing, X., Niu, Y. (2018). *Hourly day-ahead solar irradiance prediction using weather forecast by LSTM*. Energy, 148, 461-468. DOI: <https://doi.org/10.1016/j.energy.2018.01.177>.

Quijano, O. Í. (2016). *La salud: Derecho Constitucional de carácter programático y operativo*. Derecho y Sociedad. 47, 307-319. Recuperado de <http://revistas.pucp.edu.pe/index.php/derechoysociedad/article/view/18893>.

Quast, B., Fichou, D. (2019). *Package rnn*. Versión 0.9.8. Recuperado de <https://cran.r-project.org/web/packages/rnn/rnn.pdf>.

Ramón, N., López, J.J. (2016). *Econometría Series Temporales y Modelos de Ecuaciones Simultáneas*. España: Editorial Universidad Miguel Hernández de Elche. ISBN 978-84-16024-42-1.

Ren, L., Liu, Y., Rui, Z., Li, H., Feng, R. (2009). *Application of Elman Neural Network and MATLAB to Load Forecasting*. IEEE International Conference on Information Technology and Computer Science, 55-59. DOI: 10.1109/ITCS.2009.20.

Ryan, J.M., Ulrich J.M., Thielen W., Teetor, P., Bronder, S. (2019). Package quantmod. Versión 0.4-14. Recuperado de <https://cran.r-project.org/web/packages/quantmod/quantmod.pdf>.

Sahai, H., Ageel, M.I. (2000). *The Analysis of Variance* (1st ed.). New York, USA: Springer Science Business Media. ISBN 978-1-4612-7104-8.

Sax, C., Eddelbuecke, D. (2018). *Package seasonal*. Versión 1.7.0. Recuperado de <https://cran.r-project.org/web/packages/seasonal/seasonal.pdf>.

Smagulova, K., James, P. A. (2019). *Deep Learning Classifiers with Memrisitive Networks*. Modeling and Optimization in Science and Technologies, 14 139-153. <https://doi.org/10.1007/978-3-030-14524-8>.

Software R (2019a). *Package stats*. Versión 3.6.0. Recuperado de <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/00Index.html>.

Software R (2019b). *Package base*. Versión 3.7.0. Recuperado de <https://stat.ethz.ch/R-manual/R-devel/library/base/html/00Index.html>.

Software R (2019c). *Package graphics*. Versión 3.7.0. Recuperado de <https://stat.ethz.ch/R-manual/R-devel/library/base/html/00Index.html>.

Shumway, R.H., Stoffer, D.S. (2011). *Time Series Analysis and Its Applications* (3rd ed.). New York, USA: Springer. ISBN 978-1-4419-7864-6.

Tsay, R. S. (2013). *An Introduction to Analysis of Financial Data with R*. Hoboken, New Jersey, USA: John Wiley & sons, Inc. Publishing. ISBN 978-0-470-89081-3.

Uriel E. (1985). *Análisis de Series Temporales Modelos ARIMA* (1ra ed.). Valencia, España: Editorial Paraninfo. ISBN 84-283-1398-9.

Vásquez, F. M. (2017). *Modelo de Box Jenkins y Redes Neuronales para pronosticar el precio del dólar del sistema bancario en moneda nacional año 2015* (Tesis de pregrado). Escuela Profesional de Estadística, Facultad de Ciencias Físicas y Matemáticas, Universidad Nacional Pedro Ruiz Gallo, Lambayeque, Perú.

Vásquez, J., Gómez, K., Castaño, E., Cadavid, J.V., Ramírez, A. (2013). *Elasticidad de la demanda por medicamentos en el mercado farmacéutico privado en Colombia*. Ecos de Economía. 17 (36) 147-172. ISSN 1657-4206.

Verzani J. (2014). *Using R for Introductory Statistics* (2nd ed.). Boca Raton, Florida, USA: CRC Press Taylor & Francis Group. ISBN 13: 978-1-4665-9074-8.

Wickham, H., Chang, W., Henry, L., Lin Pedersen, T., Takahashi, K., Wilke, C., Woo, K. (2019). *Package ggplot2. Versión 3.1.0*. Recuperado de <https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf>.

Woodward, W.A., Gray, H.L., Elliott A.C. (2016). *Applied Time Series Analysis with R* (2nd. ed.). Boca Raton, Florida, USA: CRC Press. ISBN 13:978-1-4987-3422-6.

World Human Organization (WHO) (2016). *Medicines shortages. Drug Information*. 30 (2). 180-185. Recuperado de https://www.who.int/medicines/publications/druginformation/WHO_DI_30-2_Medicines.pdf?ua=1.

Yan, J., Chen, S., Deng, S. (2019). *A EEG-based emotion recognition model with rhythm and time characteristics*. Brain Informatics 6, 7 (1-8). DOI:<https://doi.org/10.1186/s40708-019-0100-y>.

Zeileis, A., Grothendieck, G., Ryan, J.A., Ulrich, J.M., Andrews F. (2019). *Package zoo*. Versión 1.8-6. Recuperado de <https://cran.r-project.org/web/packages/zoo/zoo.pdf>.

Zivot, E., Wang J. (2005). *Modeling Financial Times Series with S-PLUS* (2 nd ed.). New York USA: Springer Science & Business Media Publishing. ISBN 13: 978-0387-27965-7.

Zhu, W., Lan, C., Xing, J., Zeng, W., Li, Y., Shen, L., Xie, X. (2016). *Co-occurrence Feature Learning for Skeleton based Action Recognition using Regularized Deep LSTM Networks*. Proceedings Thirtieth Association for the Advanced of Artificial Intelligence Conference on Artificial Intelligence 2, 3697 – 3703. ISBN 1577357620.

Zhuge, Q., Xu, L., Zhang, G. (2017). *LSTM Neural Network with Emotional Analysis for Prediction of Stock Price*. Engineering Letters 25(2) 167-175.

ANEXOS

ANEXO N° 01

RECOPIACIÓN DE LOS DATOS



PERÚ
Ministerio
de Salud

HOSPITAL REGIONAL DEL
CUSCO

Departamento de Apoyo al
Tratamiento



- "Cusco, Capital Histórica del Perú" -
"Río de la lucha contra la corrupción y la impunidad"

Cusco, 17 de Enero de 2019

Q.F. Néstor Arzubialde Zamalloa
JEFE DEL DEPARTAMENTO DE FARMACIA DEL HOSPITAL REGIONAL DEL
CUSCO

Presente.-

Por medio de la presente Yo, NÉSTOR ARZUBIALDE ZAMALLOA, manifiesto que se hace entrega del consolidado de la demanda de PRODUCTOS FARMACÉUTICOS Y AFINES DE LOS AÑOS 2013 AL 2018 DEL HOSPITAL REGIONAL DEL CUSCO, al Sr. AMILCAR PUMA CÁRDENAS, identificado con DNI N° 40232314, en las fechas que a continuación se detalla:

PRODUCTOS FARMACÉUTICOS Y AFINES	FECHA DE ENTREGA
Consolidado 2013 al 2017	04 - Octubre -2018
Consolidado 2018	17 - Enero - 2019

Asimismo, indicar que dicha información se entrega al interesado con fines de desarrollar el proyecto de tesis intitulado "COMPARACIÓN DE MODELOS AUTORREGRESIVO INTEGRADO DE MEDIA MÓVIL Y RED NEURONAL RECURRENTE DE MEMORIA A CORTO Y LARGO PLAZO SOBRE LA DEMANDA DE MEDICAMENTOS EN EL HOSPITAL REGIONAL DEL CUSCO AÑO 2018".

Atentamente.

HOSPITAL REGIONAL DEL CUSCO
DEPARTAMENTO DE FARMACIA
Q.F. Néstor Arzubialde Zamalloa
JEFE DEL DEPARTAMENTO DE APOYO AL TRATAMIENTO
C.Q.F.P. 10438

Cc/Archivo/maz

ANEXO N° 02

ECUACIONES MATEMÁTICAS DE LA DEMANDA DE LOS MEDICAMENTOS CON EL MODELO ARIMA

Cloruro de sodio

ARIMA (0,1,1)

$$(1 - B)^1 Y_t = (1 + \theta_1 B^1) \varepsilon_t$$

$$(1 - B)^1 Y_t = (1 + \theta_1 B^1) \varepsilon_t$$

$$Y_t - BY_t = \varepsilon_t + \theta_1 B^1 \varepsilon_t$$

$$Y_t - Y_{t-1} = \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

$$Y_t = Y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

Ibuprofeno

ARIMA(1,0,1)

$$(1 - \varphi_1 B^1) Y_t = (1 + \theta_1 B^1) \varepsilon_t$$

$$Y_t - \varphi_1 B^1 Y_t = \varepsilon_t + \theta_1 B^1 \varepsilon_t$$

$$Y_t - \varphi_1 Y_{t-1} = \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \varphi_1 Y_{t-1}$$

Metamizol

ARIMA(3,1,1)(0,0,1)₁₂

$$\begin{aligned} (1 - \varphi_1 B^1 - \varphi_2 B^2 - \varphi_3 B^3 - B^1 + \varphi_1 B^2 + \varphi_2 B^3 + \varphi_3 B^4) Y_t \\ = (1 + \theta_1 B^1 + \theta_1 B^{12} + \theta_1 \theta_1 B^{12}) \varepsilon_t \end{aligned}$$

$$\begin{aligned}
Y_t - \varphi_1 B^1 Y_t - \varphi_2 B^2 Y_t - \varphi_3 B^3 Y_t - B^1 Y_t + \varphi_1 B^2 Y_t + \varphi_2 B^3 Y_t + \varphi_3 B^4 \\
= \varepsilon_t + \theta_1 B^1 \varepsilon_t + \Theta_1 B^{12} \varepsilon_t + \theta_1 \Theta_1 B^{12} \varepsilon_t
\end{aligned}$$

$$\begin{aligned}
Y_t - \varphi_1 Y_{t-1} - \varphi_2 Y_{t-2} - \varphi_3 Y_{t-3} - Y_{t-1} + \varphi_1 Y_{t-2} + \varphi_2 Y_{t-3} + \varphi_3 Y_{t-4} \\
= \varepsilon_t + \theta_1 \varepsilon_{t-1} + \Theta_1 \varepsilon_{t-12} + \theta_1 \Theta_1 \varepsilon_{t-12}
\end{aligned}$$

$$\begin{aligned}
Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \Theta_1 \varepsilon_{t-12} + \theta_1 \Theta_1 \varepsilon_{t-12} + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varphi_3 Y_{t-3} + Y_{t-1} \\
- \varphi_1 Y_{t-2} - \varphi_2 Y_{t-3} - \varphi_3 Y_{t-4}
\end{aligned}$$

$$\begin{aligned}
Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + (\Theta_1 + \theta_1 \Theta_1) \varepsilon_{t-12} + (\varphi_1 + 1) Y_{t-1} + (\varphi_2 - \varphi_1) Y_{t-2} + (\varphi_3 \\
- \varphi_2) Y_{t-3} - \varphi_3 Y_{t-4}
\end{aligned}$$

Paracetamol

ARIMA(1,1,0)(0,0,1)₁₂

$$(1 - \varphi_1 B^1)(1 - B)Y_t = (1 + \Theta_1 B^{12})\varepsilon_t$$

$$(1 - \varphi_1 B^1)(1 - B)Y_t = (1 + \Theta_1 B^{12})\varepsilon_t$$

$$(1 - \varphi_1 B^1 - B + \varphi_1 B^2)Y_t = (1 + \Theta_1 B^{12})\varepsilon_t$$

$$Y_t - \varphi_1 B^1 Y_t - B Y_t + \varphi_1 B^2 Y_t = \varepsilon_t + \Theta_1 B^{12} \varepsilon_t$$

$$Y_t - \varphi_1 Y_{t-1} - Y_{t-1} + \varphi_1 Y_{t-2} = \varepsilon_t + \Theta_1 \varepsilon_{t-12}$$

$$Y_t = \varepsilon_t + Y_{t-1} + \varphi_1 Y_{t-1} - \varphi_1 Y_{t-2} + \Theta_1 \varepsilon_{t-12}$$

Sevoflurano

ARIMA (3,1,2)

$$(1 - \varphi_1 B^1 - \varphi_2 B^2 - \varphi_3 B^3)(1 - B)Y_t = (1 + \theta_1 B^1 + \theta_2 B^2)\varepsilon_t$$

$$(1 - \varphi_1 B^1 - \varphi_2 B^2 - \varphi_3 B^3 - B + \varphi_1 B^2 + \varphi_2 B^3 + \varphi_3 B^4)Y_t = (1 + \theta_1 B^1 + \theta_2 B^2)\varepsilon_t$$

$$\begin{aligned} Y_t - \varphi_1 B^1 Y_t - \varphi_2 B^2 Y_t - \varphi_3 B^3 Y_t - B Y_t + \varphi_1 B^2 Y_t + \varphi_2 B^3 Y_t + \varphi_3 B^4 Y_t \\ = \varepsilon_t + \theta_1 B^1 \varepsilon_t + \theta_2 B^2 \varepsilon_t \end{aligned}$$

$$\begin{aligned} Y_t - \varphi_1 Y_{t-1} - \varphi_2 Y_{t-2} - \varphi_3 Y_{t-3} - Y_{t-1} + \varphi_1 Y_{t-2} + \varphi_2 Y_{t-3} + \varphi_3 Y_{t-4} \\ = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \end{aligned}$$

$$\begin{aligned} Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varphi_3 Y_{t-3} + Y_{t-1} - \varphi_1 Y_{t-2} - \varphi_2 Y_{t-3} \\ - \varphi_3 Y_{t-4} \end{aligned}$$

$$\begin{aligned} Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + (\varphi_1 + 1)Y_{t-1} + (\varphi_2 - \varphi_1)Y_{t-2} + (\varphi_3 - \varphi_2)Y_{t-3} \\ - \varphi_3 Y_{t-4} \end{aligned}$$

ANEXO N° 03

CODIGO DEL SOFTWARE LIBRE R PARA EL ANÁLISIS ESTADÍSTICO Y CÁLCULO DE LOS MODELOS ARIMA Y RNNLSTM

a) Modelo ARIMA

```
#####  
### Cloruro Sodio 900MG/100ML(0.9%1L) inyectable 2013-2018 H.R. Cusco #####  
##### ARIMA #####  
#####  
#####  
##### cargando librerías necesarias #####  
  
library(stats)  
library(imputeTS)  
library(forecast)  
library(ggplot2)  
library(seasonal)  
library(urca)  
library(trend)  
library(tseries)  
library(Metrics)  
  
#####  
##### datos enero 2013 a diciembre 2018 #####  
  
CINa<-ts(data=c(6852,6518,7266,7346, 7482,7191,6600,6791, 6870,7130, 7468,  
7576,7429,6954,8038,7123,6981,7159,7220,7203, 6880,7723, 7444,  
6653,7383,6657,7566,7946,8267,7657,6595,8269,8277,8084,7775,8088,  
7831,7135,8657,7855,7757,7687,8570,8684, 8934,8652, 8506,8656,  
9381,8909,9416,9190,9627,8788,8980,9467, 9871,10479, 9408,8636,  
8834,9171,9755,10271, 9998,9361,9298,10468, 9880,9646, 9703,10854),  
start=c(2013,1), frequency=12)  
  
datostotal<-CINa  
  
#####  
##### extracción de la serie enero 2013 a diciembre 2017 #####  
  
datos<-window(datostotal,start=c(2013,1),end=c(2017,12))  
  
datos  
  
#####  
##### verificar si hay datos faltantes #####  
#se realizó visualmente, debido a que son datos de cinco años es manejable #####  
  
#####  
### imputación de datos faltantes #####
```

```

na.interpolation (datos)

#####
### verificar si hay datos atípicos #####

tsoutliers(datos)

#####
### estadísticos descriptivos de los datos enero 2013 a diciembre 2017 por año #####

for (i in 2013:2017)
{
  x<-summary(window(datos,start=c(i,1),end=c(i,12)))
  y<-sd(window(datos,start=c(i,1),end=c(i,12)))
  z<-coefvar<-
(sd(window(datos,start=c(i,1),end=c(i,12))))*100/(mean(window(datos,start=c(i,1),end=
c(i,12))))
  cat("Año: ");
  print(i)
  print(x)
  cat("Desviación estandar: ");
  print(y)
  cat("Coeficiente de variación: ");
  print(z)
}

#####
### estadísticos descriptivos de los datos enero 2013 a diciembre 2017 #####

summary(datos)
sd(datos)
coefvar<-((sd(datos))*100/(mean(datos)))
coefvar

#####
### gráfico de la serie enero 2013 a diciembre 2017 #####

autoplot(datos)+ggtitle("")+xlab("Año")+ylab("Unidades")

#####
### gráfico de la función de autocorrelación #####

ggAcf(datos, lag=48)+ ylab("Autocorrelación")+ xlab("Retraso")+ggtitle("")

#####
### diagrama de caja de la series enero 2013 a diciembre 2017 #####

ggplot(datos)+geom_boxplot(mapping=aes("var",datos))+xlab("")+
ylab("Unidades")+scale_x_discrete(breaks=NULL)+coord_flip()

```

```
#####
### diagrama de caja de la series enero 2013 a diciembre 2017 por meses #####

tsboxplot<- function(x = x) {
season <- cycle(x)
season.factor <- factor(season)
ggplot() +
  geom_boxplot(mapping = aes(x = season.factor,
                             y = x)) +
  labs(x = "Mes", y = "Unidades")+
  scale_x_discrete(breaks=c("1", "2", "3", "4", "5", "6",
                             "7", "8", "9", "10", "11", "12"),
                  labels=c("Enero", "Febrero", "Marzo", "Abril", "Mayo", "Junio", "Julio",
                             "Agosto", "Setiembre", "Octubre", "Noviembre", "Diciembre"))
}

tsboxplot(datos)

#####
### descomposición de la serie enero 2013 a diciembre 2017 #####

datos %>% seas(x11="")->fit
autoplot(fit) + ggtitle("")+ xlab("Año")

#####
### pruebas de estacionariedad, tendencia y estacionalidad de la serie #####
### enero 2013 a diciembre 2017 #####

datos %>% ur.kpss() %>% summary() ### prueba KPSS evaluar estacionariedad
mk.test(m2) ### prueba de Mann-Kendall evaluar tendencia

##### pruebas Kruskal-Wallis evaluar estacionalidad #####

a2013<-window(datos,start=c(2013,1),end=c(2013,12))
b2014<-window(datos,start=c(2014,1),end=c(2014,12))
c2015<-window(datos,start=c(2015,1),end=c(2015,12))
d2016<-window(datos,start=c(2016,1),end=c(2016,12))
e2017<-window(datos,start=c(2017,1),end=c(2017,12))

##### pruebas Kruskal-Wallis 2013 a 2017 #####

kruskal.test(list(a2013,b2014,c2015,d2016,e2017))

##### pruebas Kruskal-Wallis en forma pareada por años #####

kruskal.test(list(a2013,b2014))
kruskal.test(list(a2013,c2015))
kruskal.test(list(a2013,d2016))
kruskal.test(list(a2013,e2017))
kruskal.test(list(b2014,c2015))
kruskal.test(list(b2014,d2016))
kruskal.test(list(b2014,e2017))
kruskal.test(list(c2015,d2016))
kruskal.test(list(c2015,e2017))
```

```

kruskal.test(list(d2016,e2017))

#####
##### pruebas de normalidad de la serie enero 2013 a diciembre 2017 #####

##### Prueba Shapiro-Wilk #####

shapiro.test(datos)

#####
##### evaluación de la normalidad en forma gráfica #####

par(mfrow= c(1,3))
hist(datos, probability =T, main="",ylab='Densidad',xlab='Unidades')
lines(density(datos), col="red")
qqnorm(datos,ylab='Cuantiles de muestra',xlab='Cuantiles teóricos',main=")
qqline(datos, col = 2)
boxplot(datos, horizontal=T,ylab="",xlab='Unidades')

#####
##### obtener el modelo ARIMA adecuado #####

## si los datos presentan heterogeneidad use transformación Box Cox ###
## ausencia de normalidad #####

lambda<-BoxCox.lambda(datos)
lambda

##si los datos presentan tendencia diferenciar #####

datbcox<-(datos^lambda)
diff(datbcox)

##Gráficos de las funciones de autocorrelación simple y parcial###

library(gridExtra)
fas<-ggAcf((diff(datbcox)),lag=48)+ylab("Autocorrelación simple")+
xlab("Retraso")+ggtitle("")
fap<-ggPacf((diff(datbcox)),lag=48)+ylab("Autocorrelación parcial")+
xlab("Retraso")+ggtitle("")
grid.arrange(fas,fap,ncol=2)

##Obtener el mejor modelo#####

library(forecast)
###ARIMA regular
fit1<-Arima(datos, order=c(1,1,1)) ##ARIMA regular

```

```

###ARIMA con dos componentes (regular y estacional)
fit1<-Arima(datos, order=c(1,1,1),seasonal=c(0,0,1))
###ARIMA con dos componentes (regular y estacional) y el coeficiente lambda
fit1<-Arima(datos, order=c(1,1,1),seasonal=c(0,0,1),lambda=lambda)

#####
##### pronóstico año 2018 #####

fc1<-forecast(fit1, h=12,level=95)
summary(fc1)

#####
##### calcular RMSE y MAPE del pronóstico año 2018 #####

fc1$mean
año2018<-window(datostotal,start=c(2018,1),end=c(2018,12))
mape((as.vector(año2018)),(as.vector(fc1$mean)))
rmse((as.vector(año2018)),(as.vector(fc1$mean)))

#####
##### gráficos dato total, entrenamiento y pronóstico #####

autoplot(fc1,series="Pronóstico")+
autolayer(año2018,series="Pronóstico MAPE (4.50%)" )+
autolayer(datostotal,series="Datos 2013-2018")+
autolayer((fitted(fc1)),series="Entrenamiento MAPE(5.11%)" )+
ggtitle("")+xlab('Año')+ylab('Unidades')+
guides(colour=guide_legend(title='Series'))

#####
### evaluación de los residuales del modelo en forma gráfica #####
### y prueba de Ljung-Box #####

checkresiduals(fit1)

#####
##### pruebas de normalidad de los residuales #####

res1<-residuals(fit1)

##### Prueba Shapiro-Wilk #####

shapiro.test(res1)

#####
##### Fin #####
#####

```

b) Modelo RNNLSTM

```
#####  
##### Cloruro sodio inyectable 2013-2018 H.R. Cusco #####  
##### RNNLSTM#####  
#####  
  
library(fpp2)  
library(forecast)  
  
CINa<-ts(data=c(6852,6518, 7266, 7346, 7482,7191,6600,6791,6870,7130,7468,7576,  
7429, 6954, 8038, 7123, 6981,7159, 7220, 7203, 6880,7723, 7444, 6653,  
7383, 6657,7566, 7946, 8267,7657, 6595, 8269, 8277,8084, 7775, 8088,  
7831, 7135,8657, 7855, 7757,7687, 8570, 8684, 8934,8652, 8506, 8656,  
9381, 8909,9416, 9190, 9627,8788, 8980, 9467, 9871,10479, 9408, 8636,  
8834, 9171,9755, 10271, 9998,9361, 9298, 10468, 9880, 9646, 9703, 10854),  
start=2013 , frequency=12)  
  
#####  
#### se aplica logaritmo para reducir los números #####  
  
data<-log(CINa)  
data  
  
#####  
##### se crea 12 datos de entrada #####  
  
library(quantmod) #####permite usar el comando Lag para los retrasos  
data<-as.zoo(data) ## permite que el vector sea número, matriz o factor  
x1<-Lag(data, k = 1)  
x2<-Lag(data, k = 2)  
x3<-Lag(data, k = 3)  
x4<-Lag(data, k = 4)  
x5<-Lag(data, k = 5)  
x6<-Lag(data, k = 6)  
x7<-Lag(data, k = 7)  
x8<-Lag(data, k = 8)  
x9<-Lag(data, k = 9)  
x10<-Lag(data, k = 10)  
x11<-Lag(data, k = 11)  
x12<-Lag(data, k = 12)  
  
#### todo se une en un objeto x  
  
x<-cbind(x1,x2,x3,x4,x5,x6,x7,x8,x9,x10,x11,x12, data)  
x  
  
#####  
##### escala de los datos y sacando observaciones faltantes #####  
  
head(round(x,3)) ### permite observar los datos faltantes  
x <- x [-(1:12), ] ## quitamos los datos faltantes  
  
x<-data.matrix(x) ## se convierte los datos del objeto x en una matrix
```

```

x
dim(x)
range_data<-function(x){(x-min(x))/(max(x)- min(x))}
min_data<-min( x )
min_data
max_data<-max( x )
max_data
x <-range_data ( x ) ## los datos son puestos en una escala 0 y 1
x
head(round(x,3))

#####
##### creando matriz de entrenamiento y la prueba #####

x1<-as.matrix ( x [ , 1 ] )## se extrae la columna 1 de la matrix x dim(60x1)
x1
dim(x1)
class(x1)
x2<-as.matrix ( x [ , 2 ] )## se extrae la columna 2 de la matrix x dim(60x1)
x3<-as.matrix ( x [ , 3 ] )
x4<-as.matrix ( x [ , 4 ] )
x5<-as.matrix ( x [ , 5 ] )
x6<-as.matrix ( x [ , 6 ] )
x7<-as.matrix ( x [ , 7 ] )
x8<-as.matrix ( x [ , 8 ] )
x9<-as.matrix ( x [ , 9 ] )
x10<-as.matrix ( x [ , 10 ] )
x11<-as.matrix ( x [ , 11 ] )
x12<-as.matrix ( x [ , 12 ] )

y<-as.matrix ( x [ , 13 ] )## se extrae la columna 13 de la matrix x dim(60x1)
y

n_train <- 48 ## número de filas para el entrenamiento

y_train<-as.matrix ( y [ 1 : n_train ] ) ## se pone los datos reales 2013 a 2016
dim(y_train)## dim(48x1)
y_train

x1_train<-as.matrix ( t ( x1 [ 1 : n_train , ] ) )## se pone los datos de entrada
dim(x1_train)## dim(1x48)
x1_train
dim(x1)
x2_train<-as.matrix ( t ( x2 [ 1 : n_train , ] ) )
x3_train<-as.matrix ( t ( x3 [ 1 : n_train , ] ) )
x4_train<-as.matrix ( t ( x4 [ 1 : n_train , ] ) )
x5_train<-as.matrix ( t ( x5 [ 1 : n_train , ] ) )
x6_train<-as.matrix ( t ( x6 [ 1 : n_train , ] ) )
x7_train<-as.matrix ( t ( x7 [ 1 : n_train , ] ) )
x8_train<-as.matrix ( t ( x8 [ 1 : n_train , ] ) )
x9_train<-as.matrix ( t ( x9 [ 1 : n_train , ] ) )
x10_train<-as.matrix ( t ( x10 [ 1 : n_train , ] ) )
x11_train<-as.matrix ( t ( x11 [ 1 : n_train , ] ) )
x12_train<-as.matrix ( t ( x12 [ 1 : n_train , ] ) )
x12_train

```

```

x12

ncol(x1_train)
ncol(x2_train)
ncol(x3_train)
ncol(x4_train)

length(y_train)

#####
##### juntando las matrices #####

x_train <-array(c(x1_train,x2_train,x3_train,x4_train,
                 x5_train,x6_train,x7_train,x8_train,
                 x9_train,x10_train,x11_train,x12_train),
dim=c(dim( x1_train ),12)) ## dim(48x12)
dim(x_train) ## (1x48x12)##contiene 12 matrices de dim(1x48)
x_train

class(x_train)

#####
##### especificando el modelo red neuronal #####
##requiere(rnn)

library(rnn)
set.seed(2018)
model1 <- trainr(Y=t( y_train ) ,
                 X = x_train ,
                 #learningrate = 0.06 ,
                 learningrate = 0.05,
                 hidden_dim = 7,
                 numepochs = 1000 ,
                 network_type = "lstm" ,
                 sigmoid = "tanh")
summary(model1)

#####
### obteniendo el modelo en base a valores pronosticados 2014 a 2017 #####

pred1_train <- t(predictr(model1, x_train))
pred1_train
#####
##### evaluar el rendimiento del conjunto de pruebas #####
##se utiliza las dimensiones del objeto x1 a x12 desde las filas 49 a 60 #####
##se juntan en el objeto x_test en este mismo se tendrá los resultados del#####
##del pronóstico año 2018 #####

pronost<-48
x1_test<-as.matrix (t(x1[( pronost+1): 60 , ] ) )
x1_test
x2_test<-as.matrix (t(x2[( pronost+1) : 60 , ] ) )
x3_test<-as.matrix (t(x3[( pronost+1) : 60 , ] ) )
x4_test<-as.matrix (t(x4[( pronost+1) : 60 , ] ) )
x5_test<-as.matrix (t(x5[( pronost+1) : 60 , ] ) )

```

```

x6_test<-as.matrix (t(x6[( pronost+1) : 60 , ] )
x7_test<-as.matrix (t(x7[( pronost+1) : 60 , ] )
x8_test<-as.matrix (t(x8[( pronost+1) : 60 , ] )
x9_test<-as.matrix (t(x9[( pronost+1) : 60 , ] )
x10_test<-as.matrix (t(x10[( pronost+1) : 60 , ] )
x11_test<-as.matrix (t(x11[( pronost+1) : 60 , ] )
x12_test<-as.matrix (t(x12[( pronost+1) : 60 , ] )

## se extrae datos del objeto y desde las filas 49 hasta 60 ####

y_test<-as.matrix ( y [ 49 : 60 ] )
y_test
x_test <- array(c(x1_test,x2_test,x3_test,x4_test,
                 x5_test,x6_test,x7_test,x8_test,
                 x9_test,x10_test,x11_test,x12_test),
              dim=c(dim(x1_test),12))
dim(x_test) ## contiene 12 matrices de dim(1x12)

#####
##### se obtiene el pronóstico año 2018 #####

pred1_test <-t(predictr(model1,x_test))

#####
##### quitando la escala a los datos #####
##### función que quita la escala #####

unscale_data<-function(x ,max_x,min_x)
  {x *(max_x-min_x)+min_x}

##### quita la escala a los datos entrenamiento año 2014 a 2016 #####

pred1_train<-unscale_data(pred1_train,max_data,min_data )
pred1_train<-exp(pred1_train)
pred1_train<-ts(pred1_train,start=c(2014,1),frequency=12)

##### quita la escala a los datos pronóstico año 2018 #####

pred1_actual<-unscale_data ( pred1_test,max_data,min_data )
pred1_actual<-exp(pred1_actual)
pred1_actual<-ts(matrix(pred1_actual),end=c(2018,12),frequency=12)

##### quita la escala a los datos reales año 2018 #####

y_actual<-unscale_data(y_test,max_data,min_data)
y_actual<-exp(y_actual)
y_actual<-ts(matrix(y_actual),end=c(2018 ,12),frequency=12)
y_actual

#####
##### inspección visual dato real y pronóstico año 2018 #####

result_all <-cbind(y_actual,round(pred1_actual,2))
colnames(result_all)<-c("Actual","NNETLSTM")
result_all

```

```
#####
##### medida de los errores rmse y mape #####

##### rmse y mape del modelo red nueronal #####

library(Metrics)

datos<-window(m2,start=c(2014,1),end=c(2017,12))
mae(m2,pred1_train)
mape(m2,pred1_train)
rmse(m2,pred1_train)

##### rmse y mape del pronóstico año 2018 #####

mae(y_actual,pred1_actual)
mape(y_actual,pred1_actual)
rmse(y_actual,pred1_actual)
#####
##### gráfico final #####

autoplot(m2,series="Datos 2013 - 2018")+
autolayer(pred1_train,series="Entrenamiento MAPE (6.27%)" )+
autolayer((ts(pred1_actual,start=c(2018,1),frequency=12)),series="Pronóstico    2018
MAPE (8.55%)" )+
xlab('Año')+ylab('Unidades')+
guides(colour=guide_legend(title='Series'))

#####
##### Fin #####
```

ANEXO N° 04

BLOQUE MEMORIA DE LA RNNLSTM

El bloque memoria de la RNNLSTM queda expresado de la forma siguiente:

$$y_{nx1}^t = h_{nx1}^t = \tanh(c_{nx1}^{t-1} \odot \text{sigmoidea}(W_{nxn}^f * h_{nx1}^{t-1} + U_{nxn}^f * x_{nx1}^t) + \text{sigmoidea}(W_{nxn}^i * h_{nx1}^{t-1} + U_{nxn}^i * x_{nx1}^t) \odot \tanh(W_{nxn}^g * h_{nx1}^{t-1} + U_{n,n}^g * x_{nx1}^t)) \odot \text{sigmoidea}(W_{n,n}^o * h_{n,1}^{t-1} + U_{n,n}^o * x_{nx1}^t)$$

ANEXO N° 05

MANEJO DE LOS DATOS DE ENTRADA PARA OBTENER EL MODELO DE LA RNNLSTM

Nota	N°	Año	N°	Lag1	Lag2	Lag3	Lag4	Lag5	Lag6	Lag7	Lag8	Lag9	Lag10	Lag11	Lag12	Data total		
Datos que quedan fuera del Lag1.....Lag12		2013	1	na	na	na	1											
			2	1	na	na	na	na	2									
			3	2	1	na	na	na	na	3								
			4	3	2	1	na	na	na	na	na	4						
			5	4	3	2	1	na	na	na	na	na	na	na	na	na	na	5
			6	5	4	3	2	1	na	na	na	na	na	na	na	na	na	6
			7	6	5	4	3	2	1	na	na	na	na	na	na	na	na	7
			8	7	6	5	4	3	2	1	na	na	na	na	na	na	na	8
			9	8	7	6	5	4	3	2	1	na	na	na	na	na	na	9
			10	9	8	7	6	5	4	3	2	1	na	na	na	na	na	10
			11	10	9	8	7	6	5	4	3	2	1	na	na	na	na	11
			12	11	10	9	8	7	6	5	4	3	2	1	na	na	na	12
Datos para el entrenamiento 2014 a 2017 Lag 1.....Lag 12 (fila 1 hasta la fila 48)	1	2014	1	12	11	10	9	8	7	6	5	4	3	2	1	13		
	2		2	13	12	11	10	9	8	7	6	5	4	3	2	14		
	3		3	14	13	12	11	10	9	8	7	6	5	4	3	15		
	4		4	15	14	13	12	11	10	9	8	7	6	5	4	16		
	5		5	16	15	14	13	12	11	10	9	8	7	6	5	17		
	6		6	17	16	15	14	13	12	11	10	9	8	7	6	18		
	7		7	18	17	16	15	14	13	12	11	10	9	8	7	19		
	8		8	19	18	17	16	15	14	13	12	11	10	9	8	20		
	9		9	20	19	18	17	16	15	14	13	12	11	10	9	21		
	10		10	21	20	19	18	17	16	15	14	13	12	11	10	22		
	11		11	22	21	20	19	18	17	16	15	14	13	12	11	23		
	12		12	23	22	21	20	19	18	17	16	15	14	13	12	24		
	13	1	24	23	22	21	20	19	18	17	16	15	14	13	25			
	14	2	25	24	23	22	21	20	19	18	17	16	15	14	26			
	15	3	26	25	24	23	22	21	20	19	18	17	16	15	27			
	16	4	27	26	25	24	23	22	21	20	19	18	17	16	28			
	17	5	28	27	26	25	24	23	22	21	20	19	18	17	29			
	18	6	29	28	27	26	25	24	23	22	21	20	19	18	30			
	19	7	30	29	28	27	26	25	24	23	22	21	20	19	31			
	20	8	31	30	29	28	27	26	25	24	23	22	21	20	32			
	21	9	32	31	30	29	28	27	26	25	24	23	22	21	33			
	22	10	33	32	31	30	29	28	27	26	25	24	23	22	34			
	23	11	34	33	32	31	30	29	28	27	26	25	24	23	35			
	24	12	35	34	33	32	31	30	29	28	27	26	25	24	36			
	25	1	36	35	34	33	32	31	30	29	28	27	26	25	37			
	26	2	37	36	35	34	33	32	31	30	29	28	27	26	38			
	27	3	38	37	36	35	34	33	32	31	30	29	28	27	39			
	28	4	39	38	37	36	35	34	33	32	31	30	29	28	40			
	29	5	40	39	38	37	36	35	34	33	32	31	30	29	41			
	30	6	41	40	39	38	37	36	35	34	33	32	31	30	42			
	31	7	42	41	40	39	38	37	36	35	34	33	32	31	43			
	32	8	43	42	41	40	39	38	37	36	35	34	33	32	44			
	33	9	44	43	42	41	40	39	38	37	36	35	34	33	45			
	34	10	45	44	43	42	41	40	39	38	37	36	35	34	46			
	35	11	46	45	44	43	42	41	40	39	38	37	36	35	47			
	36	12	47	46	45	44	43	42	41	40	39	38	37	36	48			
	37	1	48	47	46	45	44	43	42	41	40	39	38	37	49			
	38	2	49	48	47	46	45	44	43	42	41	40	39	38	50			
	39	3	50	49	48	47	46	45	44	43	42	41	40	39	51			
	40	4	51	50	49	48	47	46	45	44	43	42	41	40	52			
	41	5	52	51	50	49	48	47	46	45	44	43	42	41	53			
	42	6	53	52	51	50	49	48	47	46	45	44	43	42	54			
	43	7	54	53	52	51	50	49	48	47	46	45	44	43	55			
	44	8	55	54	53	52	51	50	49	48	47	46	45	44	56			
	45	9	56	55	54	53	52	51	50	49	48	47	46	45	57			
	46	10	57	56	55	54	53	52	51	50	49	48	47	46	58			
	47	11	58	57	56	55	54	53	52	51	50	49	48	47	59			
	48	12	59	58	57	56	55	54	53	52	51	50	49	48	60			
49	1	60	59	58	57	56	55	54	53	52	51	50	49	61				
Datos no utilizados	50	2018	2	61	60	59	58	57	56	55	54	53	52	51	50	62		
	51		3	62	61	60	59	58	57	56	55	54	53	52	51	63		
	52		4	63	62	61	60	59	58	57	56	55	54	53	52	64		
	53		5	64	63	62	61	60	59	58	57	56	55	54	53	65		
	54		6	65	64	63	62	61	60	59	58	57	56	55	54	66		
	55		7	66	65	64	63	62	61	60	59	58	57	56	55	67		
	56		8	67	66	65	64	63	62	61	60	59	58	57	56	68		
	57		9	68	67	66	65	64	63	62	61	60	59	58	57	69		
	58		10	69	68	67	66	65	64	63	62	61	60	59	58	70		
	59		11	70	69	68	67	66	65	64	63	62	61	60	59	71		
	60		12	71	70	69	68	67	66	65	64	63	62	61	60	72		
	Datos que quedan fuera del (Lag1.....Lag12)						72	71	70	69	68	67	66	65	64	63	62	
						72	71	70	69	68	67	66	65	64	63			
							72	71	70	69	68	67	66	65	64			
								72	71	70	69	68	67	66	65			
									72	71	70	69	68	67	66			
										72	71	70	69	68	67			
											72	71	70	69	68			
												72	71	70	69			
													72	71	70			
														72	71			
															72			

ANEXO N° 06

MATRIZ DE CONSISTENCIA

PROBLEMAS	OBJETIVOS	HIPÓTESIS
<p>Problema general</p> <p>¿Cuál de los modelos de pronóstico, entre el modelo autorregresivo integrado de media móvil y la red neuronal recurrente permite pronosticar adecuadamente la demanda de los medicamentos en el Hospital Regional del Cusco?</p>	<p>Objetivo general</p> <p>Analizar los modelos autorregresivo integrado de media móvil o la red neuronal recurrente para predecir adecuadamente la demanda de medicamentos en el Hospital Regional del Cusco.</p>	<p>Hipótesis general</p> <p>El modelo de la red neuronal recurrente permite pronosticar adecuadamente la demanda de los medicamentos en el Hospital Regional del Cusco.</p>
<p>Problemas específicos</p> <ul style="list-style-type: none"> • ¿Qué modelos autorregresivo integrado de media móvil proporcionan un mejor ajuste para el pronóstico de la demanda de los medicamentos? • ¿Qué modelos red neuronal recurrente proporcionan un mejor ajuste para el pronóstico de la demanda de los medicamentos? 	<p>Objetivos específicos</p> <ul style="list-style-type: none"> • Estimar los modelos autorregresivo integrado de media móvil que proporcionen el mejor ajuste para el pronóstico de la demanda de medicamentos. • Estimar los modelos red neuronal recurrente que proporcionen el mejor ajuste para el pronóstico de la demanda de medicamentos. 	<p>Hipótesis específicas</p> <ul style="list-style-type: none"> • Los modelos autorregresivo integrado de media móvil proporcionan un mejor ajuste para el pronóstico de la demanda de los medicamentos. • Los modelos red neuronal recurrente proporcionan un mejor ajuste para el pronóstico de la demanda de los medicamentos.

VARIABLE	DIMENSIÓN	TIPO	ESCALA	CRITERIO	TÉCNICA
Demanda de medicamentos	Número	Cuantitativo	Razón	Cantidad de medicamentos que requiere el paciente	Modelo autoregresivo integrado de media móvil
					Modelo de red neuronal recurrente de memoria a corto y largo plazo
Tiempo (Año 2018)	Meses	Cuantitativo	Razón	Enero - 2013 a Diciembre - 2017	Registro de la base de datos