UNIVERSIDAD NACIONAL DEL LITORAL

FACULTAD DE INGENIERÍA Y CIENCIAS HÍDRICAS

ESTADÍSTICA

PROBABILIDAD Y ESTADÍSTICA

NOTAS DE INTRODUCCIÓN A LA TEORÍA

UNIDAD 8 - SERIES CRONOLÓGICAS

Responsable de cátedra: Prof. Juan Pablo Taulamet

Equipo de cátedra: ***Auxiliares:*** Ing. Ana Lisa Eusebi (JTP) - Prof. Fátima Bolatti (JTP) - Lic. Denis Lizazo Torres (Ay. 1°) ***Ayudantes:*** AIA Cristian Bottazzi - Téc. Eliana García

Carreras: Ingeniería en Recursos Hídricos - Ingeniería en Informática - Ingeniería Ambiental - Ingeniería en Agrimensura

AÑO ACADÉMICO 2025 - PRIMER CUATRIMESTRE

ANÁLISIS DE SERIES CRONOLÓGICAS

El análisis de series cronológicas, en el contexto de esta cátedra, incluye un conjunto de métodos cuantitativos analíticos y gráficos orientados al estudio de muestras de datos provenientes de una variable aleatoria vinculada asociada a períodos regulares de tiempo. En el presente material se realizará una introducción a la teoría que abordará la estructura que deben poseer los datos, la representación gráfica de una serie, la suavización y el análisis completo de las componentes, incluyendo la realización de estimaciones o proyecciones a futuro.

ESTRUCTURA DE LOS DATOS

Para realizar el análisis de una serie temporal es necesario contar con datos que hayan sido recabados en períodos de tiempo regulares y consecutivos en forma completa. A continuación se presenta un ejemplo.

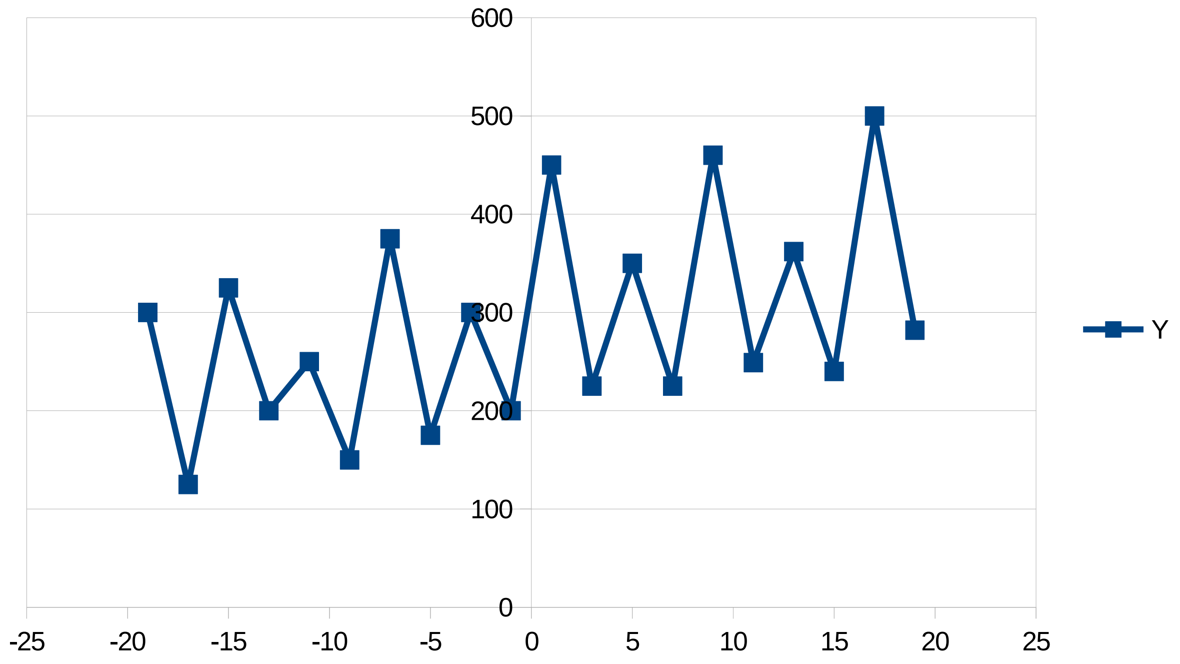
Se dispone de los datos tabulados del número de clientes (en miles) que el hotel *Los Silos,* ha recibido durante cada estación del año entre 2010 y 2014:

	Verano	Otoño	Invierno	Primavera
2010	300	150	450	249
2011	125	375	225	362
2012	325	175	350	240
2013	200	300	225	500
2014	250	200	460	282

En el ejemplo anterior se observan cuatro períodos por año, si bien podría discutirse que el verano comience en enero del año 2010, lo importante es que las referencias corresponden a períodos de tiempo que tienen la misma longitud temporal y se dispone de los datos en forma consecutiva, es decir que no falta ningún dato intermedio.

REPRESENTACIÓN GRÁFICA

La representación gráfica de una serie se puede realizar de manera similar a la construcción de un diagrama de dispersión, con trazos rectos que unen los puntos, de la siguiente forma:



SUAVIZACIÓN

La suavización de la serie tiene por objetivo obtener una versión que tenga menores variaciones o saltos menos pronunciados los valores de la serie. En el contexto de nuestra asignatura se ofrecerán dos métodos, el suavizado exponencial y la suavización por promedios móviles.

SUAVIZACIÓN POR PROMEDIOS MÓVILES

Se trata de un proceso mediante el cual el valor de de la serie suavizada de cada período se obtiene a partir del promedio de un conjunto de valores consecutivos de la serie original. La cantidad de períodos utilizados dependen del orden del promedio móvil. Por ejemplo si se toma un PM de orden 3, cada valor de la serie suavizada se calcula con el promedio de tres períodos consecutivos de, de tal forma que el valor de la serie suavizada en el período temporal 2, representado por S_2, se obtendrá promediando los primeros tres valores de la serie observada: Y_1, Y_2 e Y_3, es decir:

$$S_2 = \frac{Y_1 + Y_2 + Y_3}{3}$$

A continuación se presenta una tabla con el promedio movil de orden 3 (PM_3), del ejemplo del hotel:

Estación	Año	P	Y	PM3
Verano	2010	1	300	
Otoño	2010	2	125	250
Invierno	2010	3	325	216.67
Primavera	2010	4	200	258.33
Verano	2011	5	250	200
Otoño	2011	6	150	258.33
Invierno	2011	7	375	233.33
Primavera	2011	8	175	283.33
Verano	2012	9	300	225
Otoño	2012	10	200	316.67
Invierno	2012	11	450	291.67
Primavera	2012	12	225	341.67
Verano	2013	13	350	266.67
Otoño	2013	14	225	345
Invierno	2013	15	460	311.33
Primavera	2013	16	249	357
Verano	2014	17	362	283.67
Otoño	2014	18	240	367.33
Invierno	2014	19	500	340.67
Primavera	2014	20	282	

Pérdida de n-1 datos

El método de suavización por PM, posee algunas desventajas, la primera de las cuales tiene que ver con la cantidad de datos de la serie. Tal como puede verse en el ejemplo anterior, no puede obtenerse los valores para el primer período ni para el último, ya que no se cuentan con los valores observados consecutivos para realizar el promedio. De esta forma, un PM de orden n, implicará que la nueva serie suavizada pierda n-1 datos. Es por ello que a mayor orden de suavización, más datos se perderán.

Centrado de un PM de orden par

Otra de las desventajas del método de suavización por PM, se verifica cuando el orden es par. Tal como se vió en el ejemplo anterior para un orden 3, se pierden dos datos, y basta con ubicar la serie de tal forma de perder un dato al principio y un dato al final para que la misma quede centrada temporalmente. Esto significa que en el renglón correspondiente al período 2, se encontrará un dato que ha sido calculado a partir del período 2, el inmediato anterior y el inmediato posterior, de una manera equilibrada. En el caso de un PM de orden par, esto no será posible, ya que se perderán un conjunto impar de datos y no hay forma de que se pierda la misma cantidad de datos al inicio y al final. A contuación se presenta una tabla en la que se puede apreciar la falta de centralidad de los PM de orden par:

Estación	Año	P	Y	PM3	PM4	PM5
Verano	2010	1	300			
Otoño	2010	2	125	250	237.5	
Invierno	2010	3	325	216.67	225	240
Primavera	2010	4	200	258.33	231.25	210
Verano	2011	5	250	200	243.75	260
Otoño	2011	6	150	258.33	237.5	230
Invierno	2011	7	375	233.33	250	250
Primavera	2011	8	175	283.33	262.5	240
Verano	2012	9	300	225	281.25	300
Otoño	2012	10	200	316.67	293.75	270
Invierno	2012	11	450	291.67	306.25	305
Primavera	2012	12	225	341.67	312.5	290
Verano	2013	13	350	266.67	315	342
Otoño	2013	14	225	345	321	301.8
Invierno	2013	15	460	311.33	324	329.2
Primavera	2013	16	249	357	327.75	307.2
Verano	2014	17	362	283.67	337.75	362.2
Otoño	2014	18	240	367.33	346	326.6
Invierno	2014	19	500	340.67		
Primavera	2014	20	282			

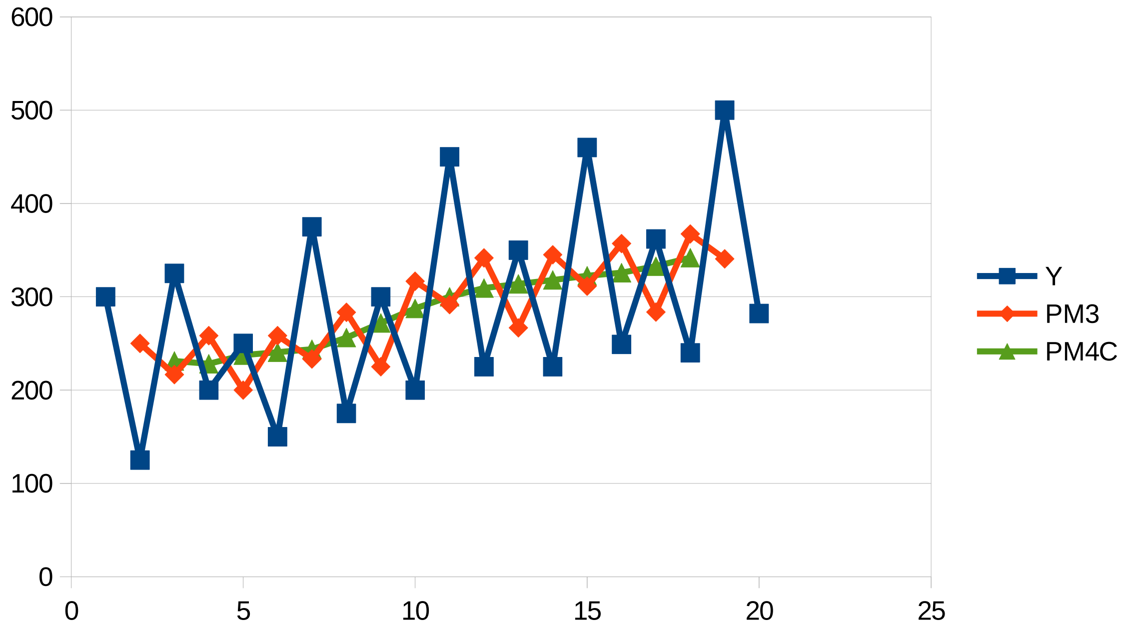
Para solucionar este problema, luego de obtener un PM de orden par, debe realizarse un nuevo PM de orden 2, para centralizar los datos. De esta forma si se desea hacer un PM de orden 4, deberá realizarse un nuevo PM de orden 2, que dará origen a un PM de orden 4 centrado, en una serie que habrá perdido 3 datos en el primer PM y 1 uno el centrado. A continuación se muestra un ejemplo para el cálculo de un PM_4C.

Estación	Año	P	Y	PM3	PM4	PM4C
Verano	2010	1	300			
Otoño	2010	2	125	250	237.5	
Invierno	2010	3	325	216.67	225	231.25
Primavera	2010	4	200	258.33	231.25	228.13
Verano	2011	5	250	200	243.75	237.5
Otoño	2011	6	150	258.33	237.5	240.63
Invierno	2011	7	375	233.33	250	243.75
Primavera	2011	8	175	283.33	262.5	256.25
Verano	2012	9	300	225	281.25	271.88

Otoño	2012	10	200	316.67	293.75	287.5
Invierno	2012	11	450	291.67	306.25	300
Primavera	2012	12	225	341.67	312.5	309.38
Verano	2013	13	350	266.67	315	313.75
Otoño	2013	14	225	345	321	318
Invierno	2013	15	460	311.33	324	322.5
Primavera	2013	16	249	357	327.75	325.88
Verano	2014	17	362	283.67	337.75	332.75
Otoño	2014	18	240	367.33	346	341.88
Invierno	2014	19	500	340.67		
Primavera	2014	20	282			

Por lo anterior, cuando se pueda elegir, será preferible tomar un PM de orden impar por sobre uno de orden par.

A continuación se presenta una representación gráfica de la serie del ejemplo del hotel y algunas versiones suavizadas por PM:



SUAVIZACIÓN EXPONENCIAL

El método de Suavizado Exponencial (SE) es una serie en la que cada período se calcula mediante un promedio ponderado entre el elemento del período actual de la serie observada y el período anterior de la serie suavizada. Existe un parámetro representado por el símbolo ω, que establece el peso que tendrá en dicha ponderación la serie observada y qué peso la suavizada a través de su complemento. De esta forma si se quiere obtener el valor suavizado de la serie para el período i, denotado por S_i, se lo podrá calcular haciendo:

$$S_i = Y_i * \omega + S_{i-1} * (1 - \omega)$$

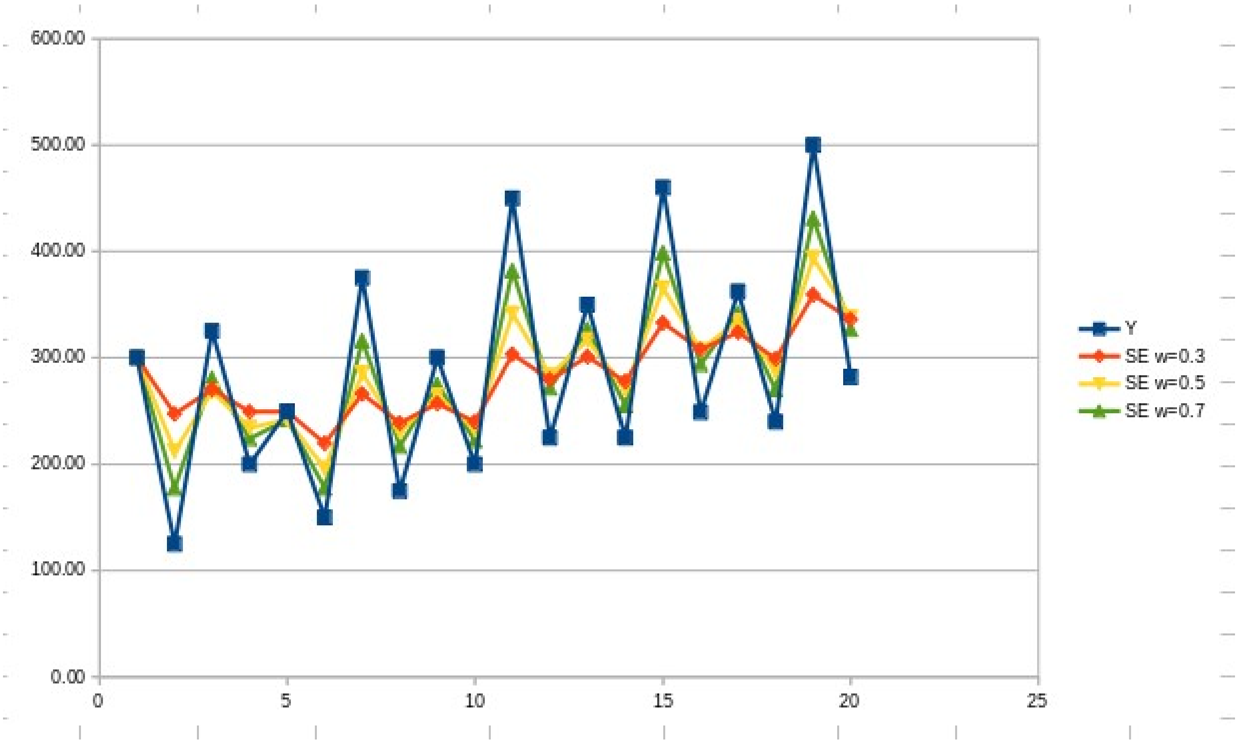
Debido a que para el cálculo del primer período suavizado, no se cuenta con un período anterior, excepcionalmente se le otorga todo el peso a la serie observada, es decir:

$$S_1 = Y_1$$

A continuación se presenta una tabla con el SE de la serie del ejemplo del hotel con diferentes valores de ω.

Estación	Año	P	Y	SE w=0.3	SE w=0.5	SE w=0.7
Verano	2010	1	300	300.00	300.00	300.00
Otoño	2010	2	125	247.50	212.50	177.50
Invierno	2010	3	325	270.75	268.75	280.75
Primavera	2010	4	200	249.52	234.38	224.23
Verano	2011	5	250	249.67	242.19	242.27
Otoño	2011	6	150	219.77	196.09	177.68
Invierno	2011	7	375	266.34	285.55	315.80
Primavera	2011	8	175	238.94	230.27	217.24
Verano	2012	9	300	257.26	265.14	275.17
Otoño	2012	10	200	240.08	232.57	222.55
Invierno	2012	11	450	303.06	341.28	381.77
Primavera	2012	12	225	279.64	283.14	272.03
Verano	2013	13	350	300.75	316.57	326.61
Otoño	2013	14	225	278.02	270.79	255.48
Invierno	2013	15	460	332.62	365.39	398.64
Primavera	2013	16	249	307.53	307.20	293.89
Verano	2014	17	362	323.87	334.60	341.57
Otoño	2014	18	240	298.71	287.30	270.47
Invierno	2014	19	500	359.10	393.65	431.14
Primavera	2014	20	282	335.97	337.82	326.74

En la imagen que se presenta a continuación puede verse el impacto que el valor de ω en la magnitud de la suavización: cuanto mayor es el valor del parámetro, más peso tendrá la serie observada y por tanto será menos suavizada.



Pronóstico mediante SE

El último valor de la serie suavizada por este método provee un mecanismo de pronóstico para el período siguiente de la serie.

Acerca del valor ω

El método de elección del valor del parámetro dependerá de cuestiones subjetivas vinculadas a la necesidad de una mayor o menor suavización, algo que podría juzgarse a partir de la representación gráfica de la serie observada y la serie suavizada. A su vez, existen diversos métodos para encontrar el valor óptimo del parámetro que minimice el promedio de las desviaciones al cuadrado entre cada valor de la serie observada en el período i y su valor pronosticado, es decir entre Y_i y S_{i-1}.

ANÁLISIS COMPLETO DE UNA SERIE DE TIEMPO

El análisis completo de una serie cronológica supone que la misma posee un conjunto de elementos diferentes e intenta estudiarlos por separado. En este apartado se definirán dichos elementos bajo el nombre ***componentes*** y se propondrá un modelo para relacionarlas. Posteriormente se mostrarán los mecanismos que permitirán aislar dichas componentes (es decir obtener sus magnitudes en forma aislada) y eliminarlas de las serie. Por último se realizarán representaciones gráficas de cada componente y se exploraran los mecanismos para estimar el comportamiento a futuro de la serie.

ACERCA DE LOS COMPONENTES DE UNA SERIE

ÍNDICE ESTACIONAL (IE)

Se trata de un índice que representa las variaciones que la serie realiza en períodos que se encuentran dentro del año. El nombre Estacional refiere a que inicialmente representaba el impacto de las estaciones del año: Verano, Primavera, Otoño e Invierno. Al margen de las estaciones del año, este índice puede utilizar para estudiar cualquier variación dentro del año. De esta forma si se posee una muestra de datos que incluye n datos por año, deberán determinarse los n valores correspondientes al IE, es decir existirá un valor para cada período.

La suma de todos los IE debe arrojar el valor n, de tal forma que su media es 1. Es por ello que los mismos, tomarán valores que deberán pensarse en referencia al número 1 (o vincularse al 100%), de tal forma que, si en una serie de datos la variación estacional es la misma para todos los períodos del año, todos los IE tomarán el valor 1.

A partir de lo anterior, si fuera el caso de que para un período en particular el IE tomara un valor menor a 1, por ejemplo 0.8, podrá pensarse que en dicho período la variación estacional implica una disminución del valor de la serie en 20% con respecto al promedio. Si por el contrario, para un período el IE tomara el valor 1.5, se podrá interpretar que en dicho período la estacionalidad marca un aumento del 50% que los otros períodos en promedio. A continuación se presentan los Índices Estacionales correspondientes al ejemplo del hotel:

Estación	IE
Verano	1,09
Otoño	0,70
Invierno	1,46
Primavera	0,75
Suma	4,00

En invierno las visitas aumentan un 46% en comparación con otras estaciones del año en promedio.

Se observa un descenso del 30% menos en Otoño con respecto al promedio

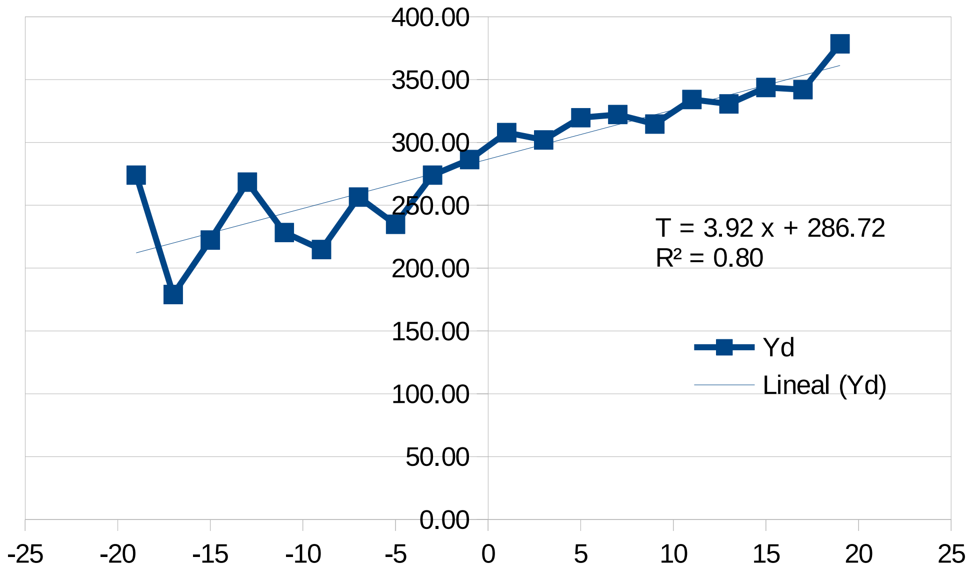
TENDENCIA SECULAR

La tendencia es el componente que representa la dirección a largo plazo de la serie. Para el estudio de la misma se realizará un análisis de regresión y correlación de manera análogo a lo estudiado en la unidad 7.

De esta forma, si se cuenta con un análisis de correlación que garantiza un buen ajuste del modelo de regresión, se utilizará el mismo para estimar la tendencia de la serie. Esta componente es fundamental para estudiar el comportamiento futuro de la serie, es decir realizar proyecciones de los valores que podrá tomar a futuro y podrá combinarse con la componente estacional para enriquecer dichas proyecciones.

En el presente material se propondrá la estimación de la tendencia a partir de la relación existente entre los valores del tiempo codificados X_c como variable explicativa y los valores de la serie sin estacionalidad Y_d.

A continuación presentamos una gráfica con la tendencia estimada con una regresión lineal para el ejemplo del hotel.



MOVIMIENTOS CÍCLICOS E IRREGULARES

La componente cíclica de la serie representa los movimientos que se realizan con respecto a la tendencia en períodos mayores a un año. Esta componente no puede aislarse completamente y se la suele estimar en conjunto con las variaciones irregulares. A su vez la componente irregular representa una variación que no se intentará estudiar en forma matemática aunque sí podría intentar interpretarse. Por ejemplo si al estudiar una serie de tiempo se encuentra un cambio en el año 2020, podría atribuirse la pandemia del Covid-19.

EL MODELO MULTIPLICATIVO Y MODELO ADITIVO

Los modelos se basan en la suposición de que la variación de la serie a lo largo del tiempo, se vincula con las cuatro componentes mediante el modelo multiplicativo o el modelo aditivo:

$$Y = IE * T * C * I$$

$$Y = IE + T + C + I$$

En el modelo anterior, la letra Y, representa el valor de la serie cronológica, es decir el valor de la serie observado en un período de tiempo en particular, y en ese contexto las letras IE, T, C e I representan al Índice Estacional, la Tendencia, la componente Cíclica y la componente Irregular de la serie, que abordaremos en breve en forma separada.

A lo largo de este material se utilizará el modelo multiplicativo que opcionalmente incluye las componentes cíclicas y las irregulares en forma unida y una referencia al período de tiempo con el subíndice i, de la siguiente forma:

$$Y_i = IE_i * T_i * CI_i$$

CÁLCULO DE LA COMPONENTE ESTACIONAL

Para obtener las magnitudes correspondientes a las variaciones estacionales en forma aislada, se determinarán los IE. Para ello se propone inicialmente evaluar la cantidad de períodos por año de datos disponibles y se denota dicha cantidad con la letra n. Naturalmente el presente cálculo tiene sentido cuando existe más de un período por año, es decir, con $n > 1$. El primer paso para el estudio de los IE consiste en obtener un promedio móvil de orden n. Dado que este promedio le otorga el mismo peso a cada una de los períodos considerados, es razonable suponer que posee todas las componentes menos la estacional, es decir, según el modelo propuesto:

$$PM_n = T * CI$$

Por lo tanto si se plantea el cociente entre la serie observada y este promedio móvil se obtendrá una primera medida para la estacionalidad, en símbolos:

$$\frac{Y}{PM_n} = \frac{IE * T * CI}{T * CI} = IE$$

Si bien lo anterior es cierto desde el punto de vista conceptual, es posible (y de hecho deseable) que exista más de un valor observado de la serie para cada período del año y aunque se pierdan algunos(en rigor de verdad $n-1$), también han de existir varios PM para cada período. Es por ello que se establecerá un mecanismo (a través del cálculo de la mediana) que permita agrupar todos los cocientes $\frac{Y}{PM_n}$ correspondientes a cada período y obtener con ellos un único valor, es decir un IE para cada período. Por último, se garantizará que el promedio de todos los IE posean un promedio 1, lo cual es lo mismo que establecer que la suma de todos los IE coincida con n.

Para ilustrar el proceso utilizaremos el ejemplo del hotel. Se presenta a continuación una tabla con los valores de la serie observada, y dado que se dispone de datos con cuatro períodos por año, se trabajará con un promedio móvil de orden 4 centrado y el cociente entre el valor de la serie y el PM.

Estación	Año	P	Y	PM4	PM4C	E = Y/ PM4C
Verano	2010	1	300			
Otoño	2010	2	125	237,5		
Invierno	2010	3	325	225	231,25	1,41
Primavera	2010	4	200	231,25	228,125	0,88
Verano	2011	5	250	243,75	237,5	1,05
Otoño	2011	6	150	237,5	240,625	0,62
Invierno	2011	7	375	250	243,75	1,54
Primavera	2011	8	175	262,5	256,25	0,68
Verano	2012	9	300	281,25	271,875	1,10
Otoño	2012	10	200	293,75	287,5	0,70
Invierno	2012	11	450	306,25	300	1,50
Primavera	2012	12	225	312,5	309,375	0,73
Verano	2013	13	350	315	313,75	1,12
Otoño	2013	14	225	321	318	0,71
Invierno	2013	15	460	324	322,5	1,43
Primavera	2013	16	249	327,75	325,875	0,76
Verano	2014	17	362	337,75	332,75	1,09
Otoño	2014	18	240	346	341,875	0,70
Invierno	2014	19	500			
Primavera	2014	20	282			

A partir de lo anterior, se realiza una agrupación por período de los cocientes $\frac{Y}{PMC_{12}}$ y se calcula para cada uno de ellos la mediana, tal como se muestra a continuación:

	2010	2011	2012	2013	2014	Mediana
Verano		1,05	1,10	1,12	1,09	1,10
Otoño		0,62	0,70	0,71	0,70	0,70
Invierno	1,41	1,54	1,50	1,43		1,46
Primavera	0,88	0,68	0,73	0,76		0,75

Si bien es cierto que dichas medianas representan la estacionalidad para cada período, debe verificarse ahora si la suma de todos estos valores coincide con n. A menudo esto no sucede y es por tanto necesario ajustarlos, para que coincidan. Para ello se suele utilizar un factor de corrección que se obtiene haciendo un cociente entre n y el valor de la suma. A continuación se presenta una tabla con el proceso de corrección y verificación que concluye con el valor de los IE.

	2010	2011	2012	2013	2014	Mediana	IE
Verano		1,05	1,10	1,12	1,09	1,10	1,09
Otoño		0,62	0,70	0,71	0,70	0,70	0,70
Invierno	1,41	1,54	1,50	1,43		1,46	1,46
Primavera	0,88	0,68	0,73	0,76		0,75	0,75
					Σ	4,003	4,000
				FC = 4/Σ		0,999	

Por último, ya que se disponen los valores correspondientes a los IE para cada período, será posible obtener la versión desestacionalizada de la serie, haciendo:

$$Y_d = \frac{Y}{IE}$$

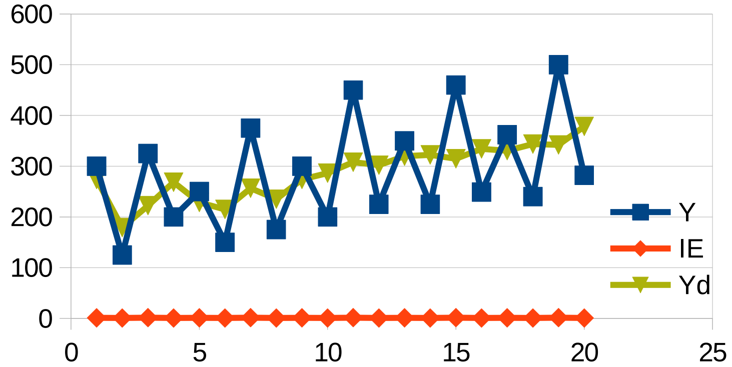
A partir de los supuestos del modelo multiplicativo, la serie desestacionalizada no posee estacionalidad es decir, sólo quedan las otras componentes:

$$Y_d = T * CI$$

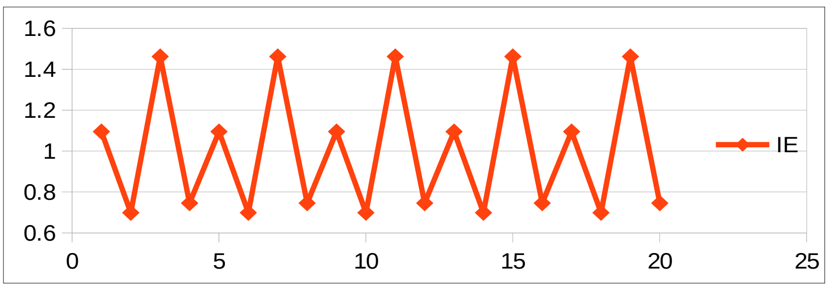
A continuación se presenta una tabla con el cálculo de la Y_d.

Estación	Año	P	Y	PM4	PM4C	E = Y/ PM4C	IE	Yd
Verano	2010	1	300				1,09	274,03
Otoño	2010	2	125	237,5			0,70	179,02
Invierno	2010	3	325	225	231,25	1,41	1,46	222,31
Primavera	2010	4	200	231,25	228,125	0,88	0,75	268,44
Verano	2011	5	250	243,75	237,5	1,05	1,09	228,36
Otoño	2011	6	150	237,5	240,625	0,62	0,70	214,82
Invierno	2011	7	375	250	243,75	1,54	1,46	256,51
Primavera	2011	8	175	262,5	256,25	0,68	0,75	234,88
Verano	2012	9	300	281,25	271,875	1,10	1,09	274,03
Otoño	2012	10	200	293,75	287,5	0,70	0,70	286,43
Invierno	2012	11	450	306,25	300	1,50	1,46	307,81
Primavera	2012	12	225	312,5	309,375	0,73	0,75	301,99
Verano	2013	13	350	315	313,75	1,12	1,09	319,71
Otoño	2013	14	225	321	318	0,71	0,70	322,24
Invierno	2013	15	460	324	322,5	1,43	1,46	314,65
Primavera	2013	16	249	327,75	325,875	0,76	0,75	334,20
Verano	2014	17	362	337,75	332,75	1,09	1,09	330,67
Otoño	2014	18	240	346	341,875	0,70	0,70	343,72
Invierno	2014	19	500				1,46	342,01
Primavera	2014	20	282				0,75	378,49

Un vez aislada y eliminada la estacionalidad, es posible representar gráficamente la serie observada, los índices estacionales y la serie sin estacionalidad.



Por una cuestión de escala, es preferible representar los IE en una gráfica aparte:



CÁLCULO DE LA TENDENCIA

Para el cálculo de tendencia se propone un análisis de regresión y correlación entre los valores del tiempo codificados X_c como variable explicativa y los valores de la serie sin estacionalidad Y_d.

Se propone codificar la X, para contar un un número que nos permita cuantificar cada período temporal, si bien es cierto que numerar los períodos es una buena forma de comenzar, existe una codificación que posee algunas virtudes a partir de poseer una media centrada en cero, por lo que se propone para codificar la variable el siguiente método:

$$X_c = (X_i - \overline{X}) * 2$$

De esta forma si se tienen valores de X que corresponden a años, por ejemplo desde el 2000 y hasta el 2015, la codificación sería de la siguiente forma:

X	Xc
2000	-15
2001	-13
2002	-11
2003	-9
2004	-7
2005	-5
2006	-3
2007	-1
2008	1
2009	3
2010	5
2011	7
2012	9
2013	11
2014	13
2015	15

$$\overline{X} = 2007.5$$

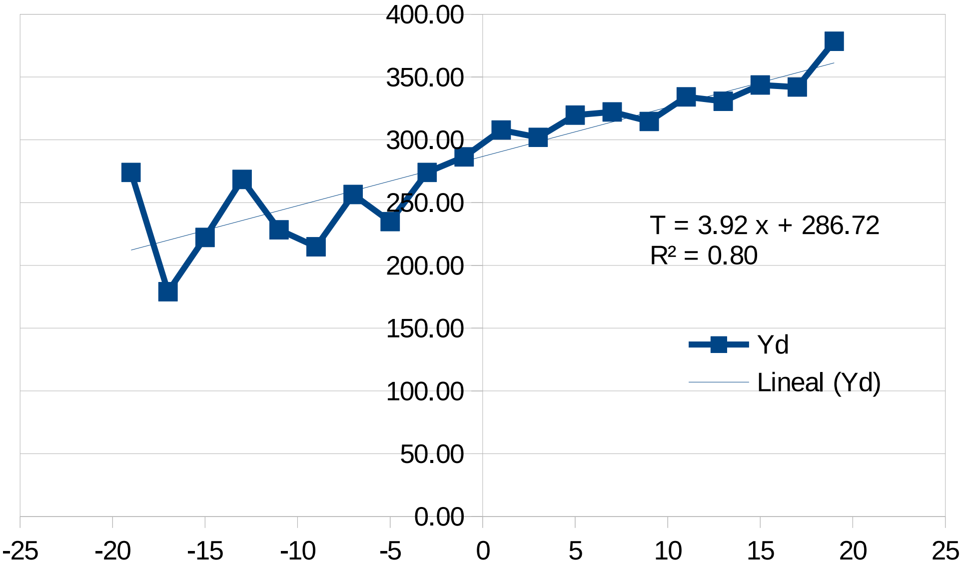
$$\overline{X}_c = 0$$

$$(2005 - 2007.5) \times 2 = -5$$

$$(2011 - 2007.5) \times 2 = 7$$

Si no se cuenta con números vinculados a la X, tal como sucede en el ejemplo del Hotel, que posee períodos con nombres, se podrá tomar la numeración de los mismos desde 1 en adelante como valores de X para con ellos calcular la versión codificada.

Luego de calcular la versión codificada de la X, y la versión desestacionalizada de la Y, se propone estimar la tendencia utilizando una ecuación de regresión que se ajuste con una buena correlación. De esta forma, retomando el ejemplo del hotel presentado anteriormente, se obtuvo la siguiente regresión lineal:



Por lo tanto podrá calcularse la tendencia haciendo para cada período[1]:

$$T = \hat{Y}_d = 3.92 * X_c + 286.72$$

De esta forma, puede completarse la tabla tal como se ve a continuación:

Año	P	Y	PM4	PM4C	E = Y/ PM4C	IE	Xc	Yd	T
2010	1	300				1,09	-19	274,03	212,24
2010	2	125	237,5			0,70	-17	179,02	220,08
2010	3	325	225	231,25	1,41	1,46	-15	222,31	227,92
2010	4	200	231,25	228,125	0,88	0,75	-13	268,44	235,76
2011	5	250	243,75	237,5	1,05	1,09	-11	228,36	243,60
2011	6	150	237,5	240,625	0,62	0,70	-9	214,82	251,44
2011	7	375	250	243,75	1,54	1,46	-7	256,51	259,28
2011	8	175	262,5	256,25	0,68	0,75	-5	234,88	267,12
2012	9	300	281,25	271,875	1,10	1,09	-3	274,03	274,96
2012	10	200	293,75	287,5	0,70	0,70	-1	286,43	282,80
2012	11	450	306,25	300	1,50	1,46	1	307,81	290,64
2012	12	225	312,5	309,375	0,73	0,75	3	301,99	298,48
2013	13	350	315	313,75	1,12	1,09	5	319,71	306,32
2013	14	225	321	318	0,71	0,70	7	322,24	314,16

1 Esto sólo tiene sentido si se cuenta con un coeficiente de correlación adecuado.

Año	P	Y	PM4	PM4C	E = Y/ PM4C	IE	Xc	Yd	T
2013	15	460	324	322,5	1,43	1,46	9	314,65	322,00
2013	16	249	327,75	325,875	0,76	0,75	11	334,20	329,84
2014	17	362	337,75	332,75	1,09	1,09	13	330,67	337,68
2014	18	240	346	341,875	0,70	0,70	15	343,72	345,52
2014	19	500				1,46	17	342,01	353,36
2014	20	282				0,75	19	378,49	361,20

A partir de lo anterior, y considerando una vez más el modelo multiplicativo es posible aislar la componente Cíclica e Irregular, es decir, si se considera que:

$$Y_d = T * CI$$

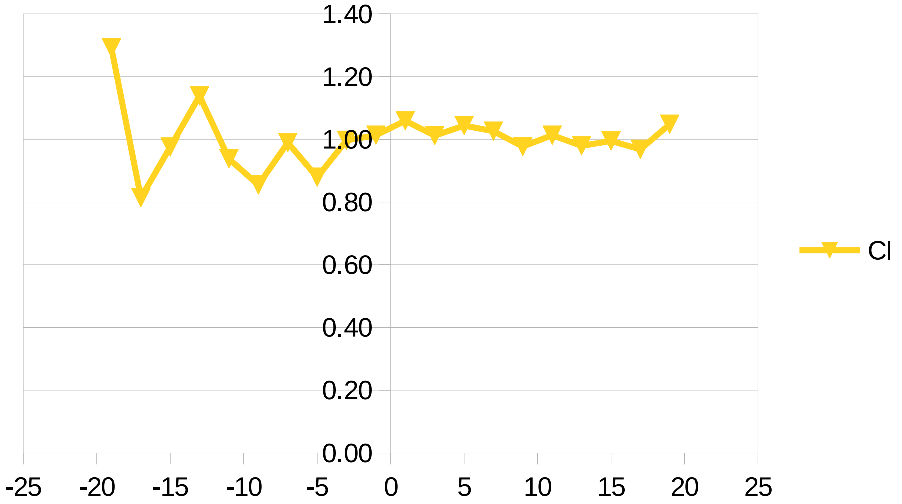
Será posible calcular para cada período el cociente $\dfrac{Y_d}{T}$ de tal forma que:

$$\frac{Y_d}{T} = CI$$

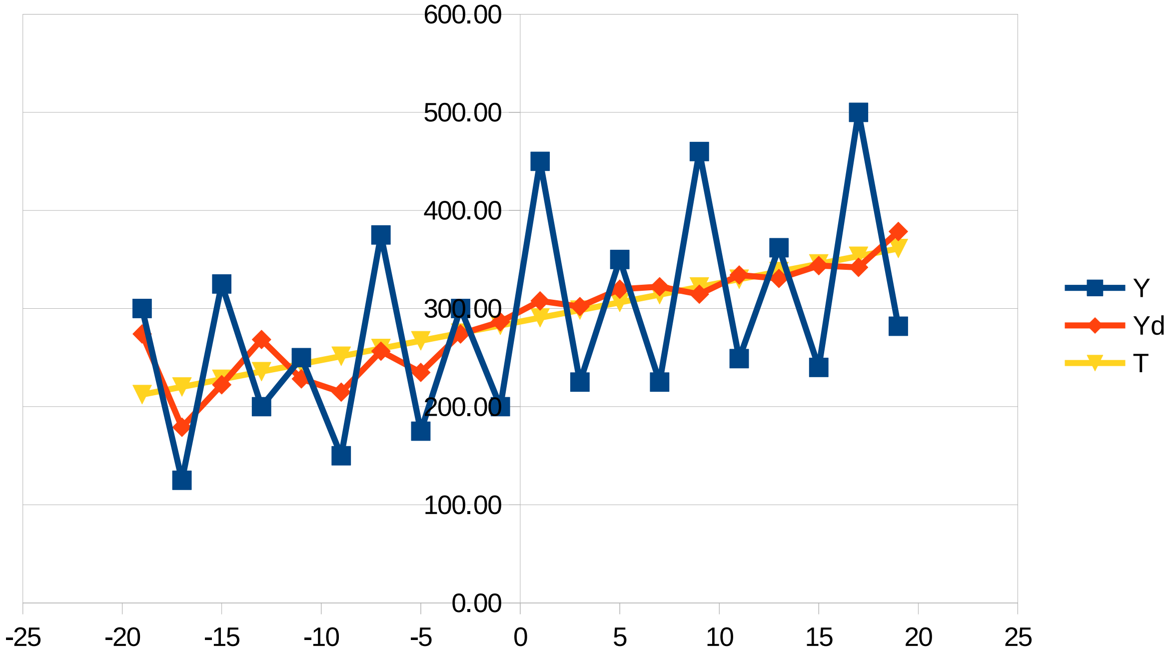
A continuación se presenta una tabla con los valores de CI.

Xc	Yd	T	CI
-19	274.03	212.24	1.29
-17	179.02	220.08	0.81
-15	222.31	227.92	0.98
-13	268.44	235.76	1.14
-11	228.36	243.60	0.94
-9	214.82	251.44	0.85
-7	256.51	259.28	0.99
-5	234.88	267.12	0.88
-3	274.03	274.96	1.00
-1	286.43	282.80	1.01
1	307.81	290.64	1.06
3	301.99	298.48	1.01
5	319.71	306.32	1.04
7	322.24	314.16	1.03
9	314.65	322.00	0.98
11	334.20	329.84	1.01
13	330.67	337.68	0.98
15	343.72	345.52	0.99
17	342.01	353.36	0.97
19	378.49	361.20	1.05

De manera similar al caso de la estacionalidad, por una cuestión de escala se representa gráficamente CI en forma separada:



A su vez pueden representarse en un mismo gráfico la serie observada, la serie desestacionalizada y la tendencia secular:



PROYECCIONES A FUTURO

Para realizar una proyección de un valor a futuro de la serie se utilizará la componente estimada de la tendencia y el valor codificado correspondiente al período temporal que se quiere estimar. A su vez, se hará una distinción entre dos tipos diferentes de estimaciones: predicción y pronóstico. La predicción y el pronóstico a nivel puntual toman el mismo valor, porque se calculan utilizando la misma ecuación de regresión[2] y es por ello que la diferencia puede parecer sutil, y se vincula con el valor de la serie que se desea estimar. A su vez luego de obtener la predicción y/o el pronóstico se podrá optar por agregar el efecto de la estacionalidad correspondiente al X_h.[3]

PREDICCIÓN

Se considera una predicción a la obtención de un valor estimado de $\overline{Y}$ en el período X_h. Para obtener dicha predicción se obtendrá utilizando la ecuación de regresión utilizada para estimar la tendencia. La predicción arrojará el valor estimado de Yd para el período a estimar X_h y no contendrá estacionalidad.

La expresión simbólica correspondiente a la predicción es la siguiente:

$$\hat{Y}_h = a + b * x_h$$

PRONÓSTICO

Se considera un pronóstico a la obtención de un valor estimado de Y en el período X_h. Para obtener el pronóstico se obtendrá utilizando la ecuación de regresión utilizada para estimar la tendencia. El pronóstico arrojará el valor estimado de Yd para el período a estimar X_h y no contendrá estacionalidad.

La expresión simbólica correspondiente al pronóstico es la siguiente:

$$\tilde{Y}_h = a + b * x_h$$

AGREGANDO LA ESTACIONALIDAD A UNA ESTIMACIÓN

Para agregar el efecto de la estacionalidad a una estimación basta con multiplicar el valor estimado por el IE correspondiente al período Xh.

A modo de ejemplo se propone retomar el caso del Hotel y realizar una predicción del número de clientes esperado para el otoño de 2015. Para ello se tomará el valor codificado a dicho período, es decir $X_h = 23$ y se lo utilizará para evaluar de dicho valor la ecuación de la tendencia, es decir haciendo:

---

2 Esto sólo tiene sentido si se cuenta con un coeficiente de correlación adecuado

3 Esto sólo tiene sentido si la serie posee estacionalidad.

$$\tilde{Y_h} = 3.92 * 23 + 286.72 = 376.97$$

De esta forma se espera que haya aproximadamente 377 clientes en el otoño de 2015, sin considerar la estacionalidad. Si se la desea agregar, basta con multiplicar por el IE del otoño:

$$376.97 * 0.70 = 263.22$$

De esta forma, dado que durante el otoño existe una estacionalidad que marca un descenso de la cantidad de clientes de un 30% en promedio, si agregamos la estacionalidad a la predicción el valor se hace un 30% menor, es decir en vez de 377 personas, serán aproximadamente 263.

INFERENCIAS SOBRE UNA ESTIMACIÓN PUNTUAL

A partir de una estimación puntual y contando con la magnitud de varianza de la regresión será posible obtener una estimación por intervalos de confianza utilizando las siguientes expresiones:

PREDICCIÓN	PRONÓSTICO
$\left(\hat{Y}_h \pm t_{n-2;1-\frac{\alpha}{2}} \hat{\sigma}_{pred}\right)$	$\left(\tilde{Y}_h \pm t_{n-2;1-\frac{\alpha}{2}} \hat{\sigma}_{pron}\right)$
$\hat{\sigma}_{pred} = S_{y/x} \times \sqrt{\frac{1}{n} + \frac{(X_h - \overline{X})^2}{\sum_{\forall i}(X_i - \overline{X})}}$	$\hat{\sigma}_{pron} = S_{y/x} \times \sqrt{1 + \frac{1}{n} + \frac{(X_h - \overline{X})^2}{\sum_{\forall i}(X_i - \overline{X})^2}}$
$S^2_{y/x} = \frac{\sum_{\forall i}(Y_i - \hat{Y}_i)^2}{n-2}$	$S^2_{y/x} = \frac{\sum_{\forall i}(Y_i - \tilde{Y}_i)^2}{n-2}$

A partir del ejemplo del hotel, se propone construir un IC del 95% para la predicción obtenida. Para ello se considera la varianza regresión lineal entre X_c e Y_d igual a:

$$S^2_{y/x} = 574,58$$

Se toma el valor del cuantil $t_{18,97.5\%} = 2.10$ para un 95% de confianza.

De esta forma para un X_h = 23 se tiene,

$$\hat{\sigma}_{y/x} \approx 26.79$$

Tener en cuenta para lo anterior que se trabaja con un X_c que posee $\overline{X} = 0$.

De esta forma con $1 - \alpha = 95\%$, el IC sin estacionalidad será $(320.69; 433.25)$ y considerando la estacionalidad del período $IE = 0.70$, el IC será $(206.94; 319.49)$.

SOBRE LA VERSIÓN DE ESTE MATERIAL

Este documento es parte del material de cátedra para el cursado del primer cuatrimestre de 2025 de la asignatura Estadística perteneciente a los planes de estudios de las carreras: Ingeniería en Recursos Hídricos, Ingeniería Ambiental, Ingeniería en Agrimensura, Ingeniería en Informática; así como para la asignatura Probabilidad y Estadística perteneciente a los planes de estudios de las carreras: Ingeniería en Recursos Hídricos e Ingeniería en Inteligencia Artificial.

Ya que este documento se actualiza periódicamente, al momento de su lectura se recomienda revisar si no se encuentra disponible alguna versión más actualizada. Ante cualquier duda sobre este material se solicita enviar un correo a *taulamet@unl.edu.ar*.

Este documento ha sido actualizado por última vez el día 07/06/25 a las 09:34:35 hs.