

EXAMEN DE ECONOMETRÍA II
29 de julio de 2015

EJERCICIO 1 (35 puntos)

Un grupo de ecologista quiere investigar la población histórica de los lince en Canadá. Para ello cuentan con datos de las pieles de lince vendidas anualmente por la compañía Hudson's Bay entre 1857 y 1911¹. A continuación en la figura 1 se grafica la serie de datos en niveles y a raíz de lo sugerido por un investigador se decide aplicar logaritmos cuyo gráfico se presenta en la figura 2. En las figuras 3 y 4 se presentan los respectivos correlogramas de dichas series.

Figura 1

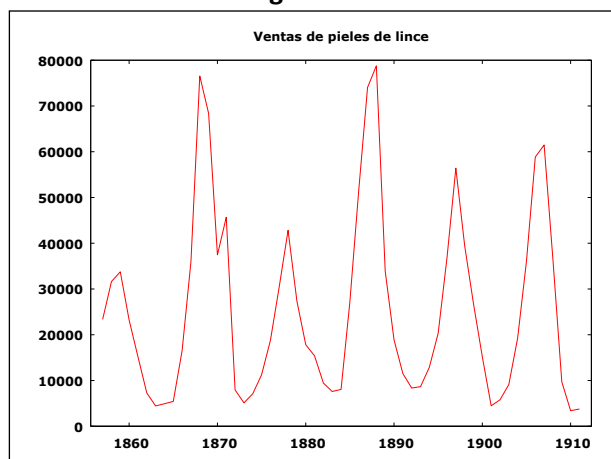


Figura 2

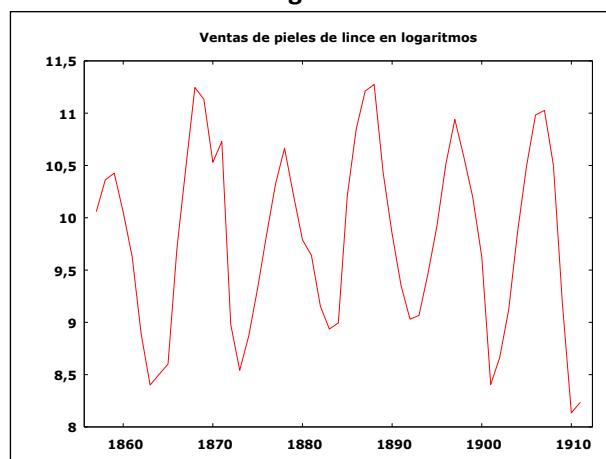


Figura 3

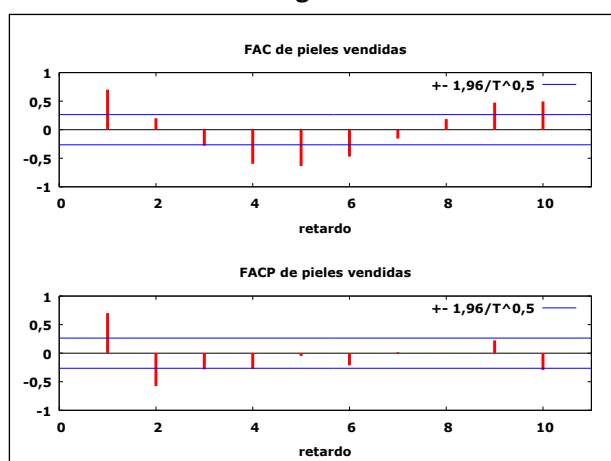
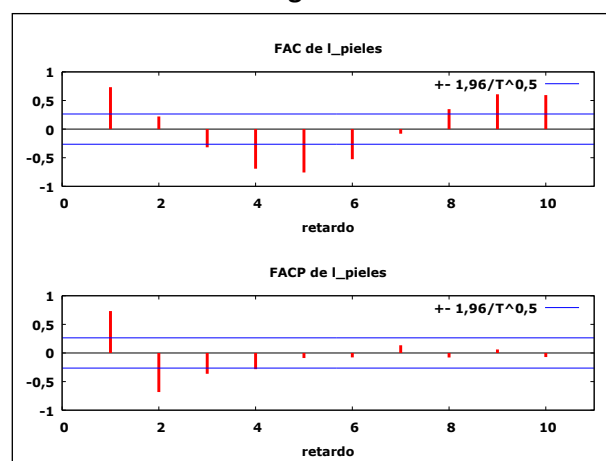


Figura 4



¹ Datos provenientes del libro "Análisis de Series de Tiempo: métodos univariantes y multivariantes" de William W.S. Wei, Segunda Edición.

- (1) Responda las siguientes preguntas: Defina el concepto de estacionariedad en sentido débil ¿Por qué un investigador sugirió la aplicación de la transformación logarítmica? ¿Cree usted que se logró el objetivo que buscaba? ¿Considera que se debe realizar alguna transformación adicional?

Una propuesta manejada en el equipo de investigación es el Modelo 1 con su correspondiente correlograma y gráfico de los residuos:

Modelo 1 usando las observaciones 1859-1911 (T = 53)
Variable dependiente: \ln_{pieles}

	<i>Coefficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	4,31485	0,704124	6,1280	<0,0001	***
phi_1	1,31389	0,0967163	13,5850	<0,0001	***
phi_2	-0,754229	0,100458	-7,5079	<0,0001	***
Media de la vble. dep.	9,787674	D.T. de la vble. dep.		0,894564	
media innovaciones	3,59e-15	D.T. innovaciones		0,409017	
Log-verosimilitud	-26,27769	Criterio de Akaike		58,55539	
Criterio de Schwarz	64,46626	Crit. de Hannan-Quinn		60,82843	

	<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR				
Raíz 1	0,8710	-0,7531	1,1515	-0,1135
Raíz 2	0,8710	0,7531	1,1515	0,1135

Figura 5

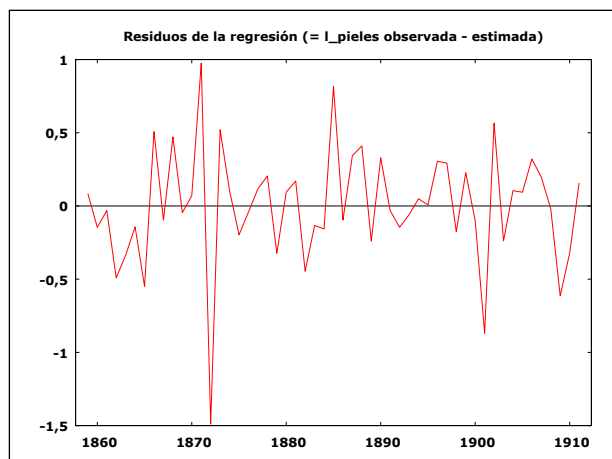


Figura 6

Modelo 1: Función de autocorrelación de los residuos				
RETARDO	FAC	FACP	Estad-Q. [valor p]	
1	-0,3443 **	-0,3443 **		
2	0,0735	-0,0511		
3	0,1144	0,1403	7,7180	[0,005]
4	-0,2214	-0,1537	10,6347	[0,005]
5	0,0447	-0,1049	10,7563	[0,013]
6	-0,2828 **	-0,3572 ***	15,7155	[0,003]
7	0,2047	0,0535	18,3714	[0,003]
8	-0,0710	0,0167	18,6976	[0,005]
9	0,0521	0,0949	18,8772	[0,009]
10	0,1636	0,0742	20,6923	[0,008]

- (2) Escriba el modelo sugerido para la serie de ventas de pieles en logaritmo en la notación ARIMA(p,d,q). ¿Qué elementos considera motivaron esta especificación? En base a la información presentada, ¿considera es un modelo adecuado? Comente las salidas y justifique utilizando la metodología Box-Jenkins.

En el grupo de investigación se analiza también el siguiente modelo:

Modelo 2, usando las observaciones 1860-1911 (T = 52)
Variable dependiente: I_{pieles}

	<i>Coefficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	6,40444	0,843047	7,5968	<0,0001	***
phi_1	0,96588	0,127175	7,5949	<0,0001	***
phi_2	-0,122632	0,191302	-0,6410	0,5215	
phi_3	-0,496333	0,13274	-3,7391	0,0002	***
Media de la vble. dep.	9,775380	D.T. de la vble. dep.		0,898759	
media innovaciones	-1,01e-15	D.T. innovaciones		0,367205	
Log-verosimilitud	-19,60823	Criterio de Akaike		47,21647	
Criterio de Schwarz	55,02144	Crit. de Hannan-Quinn		50,20871	

	<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR				
Raíz 1	0,8101	0,6503	1,0388	0,1077
Raíz 2	0,8101	-0,6503	1,0388	-0,1077
Raíz 3	-1,8672	0,0000	1,8672	0,5000

Figura 7

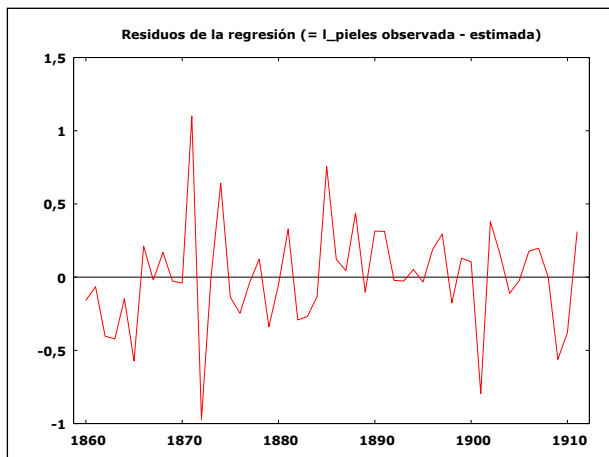
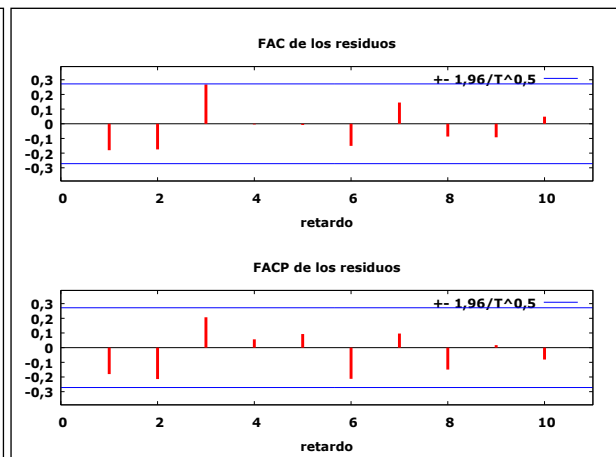


Figura 8



- (3) ¿Por qué considera usted que este modelo fue propuesto? Responda considerando la información disponible hasta el momento. Comente las salidas de acuerdo a la metodología Box-Jenkins.

Alternativamente otro investigador del equipo propone el siguiente modelo para el logaritmo de la serie de ventas de pieles de linco:

Modelo 3, usando las observaciones 1859-1911 (T = 53)
Variable dependiente: l_pieles

	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	3,90124	0,458136	8,5154	<0,0001	***
phi_1	1,54455	0,0871084	17,7314	<0,0001	***
phi_2	-0,942178	0,083352	-11,3036	<0,0001	***
theta_1	-0,591011	0,137741	-4,2907	<0,0001	***
Media de la vble. dep.	9,787674	D.T. de la vble. dep.		0,894564	
media innovaciones	-0,003819	D.T. innovaciones		0,341550	
Log-verosimilitud	-18,26790	Criterio de Akaike		46,53581	
Criterio de Schwarz	56,38727	Crit. de Hannan-Quinn		50,32420	

	<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR				
Raíz 1	0,8197	-0,6241	1,0302	-0,1036
Raíz 2	0,8197	0,6241	1,0302	0,1036
MA				
Raíz 1	1,6920	0,0000	1,6920	0,0000

Figura 9

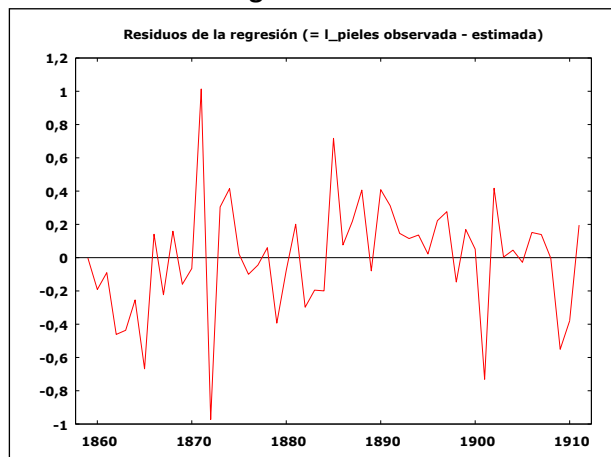
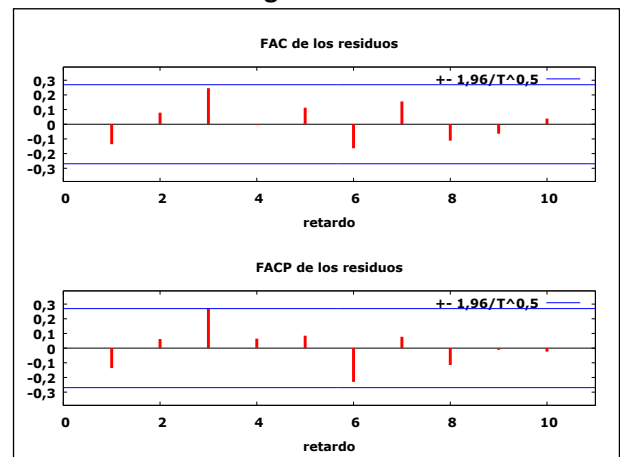


Figura 10



- (4) Comente la estimación de este modelo de acuerdo a la metodología Box-Jenkins.
- (5) En virtud de los 3 modelos sugeridos hasta ahora, ¿cuál considera usted es el modelo más adecuado para el logaritmo de las ventas de pieles de linco? Justifique.
- (6) Finalmente un miembro del equipo estimó el siguiente Modelo 4 y lo presenta al grupo:

$$(1 - 0.8L)(\ln Y_t + 0.2\ln Y_{t-1} - 0.03\ln Y_{t-2}) = \varepsilon_t - 0.1\varepsilon_{t-1} - 0.12\varepsilon_{t-2}$$

¿A qué notación ARIMA(p,d,q) corresponde este modelo? Luego de analizar esta ecuación por unos minutos se descartó dicho modelo sin siquiera considerar otra información del mismo. ¿Cuál(es) podrían ser las razones para descartarlo?

EJERCICIO 2 (29 puntos)

Se poseen datos de los aportes a la Seguridad Social de 827 individuos inscritos en el BPS de los cuales 452 aportaron y 375 no aportaron en el mes de octubre de 2004. Se quiere estimar el efecto de la edad en las contribuciones. Para ello se estiman dos modelos, en ambas la variable de interés principal es el monto de contribuciones de cada individuo en octubre de 2004 (denominada APORTES):

1. La estimación 1 es un modelo Tobit, que utiliza todas las observaciones, pero trata de forma diferencial las observaciones en las que el aporte es cero, siendo los regresores el conjunto X1 definido más adelante.
2. La estimación 2 considera una estimación siguiendo el modelo de selección de Heckman estimado en dos etapas. En la primera etapa se estima un Probit para una variable dependiente (denominada COTIZA) que asume el valor 1 si el individuo aportó en octubre 2004 (siendo los regresores el conjunto X1); a partir de esta estimación se obtiene el valor estimado de la variable λ (que es la inversa del ratio de Mills). En la segunda etapa se estima por MCO una ecuación en la que la variable dependiente es el monto de contribuciones de cada individuo en octubre de 2004 y los regresores el conjunto X1 más la variable λ

El conjunto X1 incluye: SEXO: el sexo del individuo (1=mujer, 0=hombre); EDAD: la edad del individuo en octubre de 2004; PUBLICO: una variable dummy que adopta el valor 1 si el individuo es funcionario público y 0 en otros casos.

Los resultados obtenidos son:

Estimación 1: REGRESION TOBIT

					Number of obs	=	827
					LR chi2(3)	=	132.39
					Prob > chi2	=	0.0000
Log likelihood = -4358.7171					Pseudo R2	=	0.0150
-----+-----							
APORTES	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]		
-----+-----							
sexo	-515.9782	170.9039	-3.02	0.003	-851.4364	-180.5201	
edad	27.61204	6.970924	3.96	0.000	13.92918	41.2949	
publico	2315.908	232.95	9.94	0.000	1858.662	2773.153	
_cons	-702.6031	383.6336	-1.83	0.067	-1455.617	50.41091	
-----+-----							
/sigma	2209.662	78.01832			2056.524	2362.8	
-----+-----							
Obs. summary:	375	left-censored observations at contrib<=0					
	452	uncensored observations					

Estimación 2: HECKMAN SELECTION MODEL

```
Heckman selection model -- two-step estimates      Number of obs      =      827
(regression model with sample selection)          Censored obs       =      375
                                                Uncensored obs     =      452
```

```
Wald chi2(6)      =      66.32
Prob > chi2       =      0.0000
```

		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
-----+-----							
COTIZA (Modelo Probit)							
sexo		-.0807356	.0906688	-0.89	0.373	-.2584432	.096972
edad		.0054876	.0036357	1.51	0.131	-.0016383	.0126135
publico		1.284687	.1657662	7.75	0.000	.959791	1.609583
_cons		-.1228815	.1990194	-0.62	0.537	-.5129523	.2671894
-----+-----							
APORTES (Ecuación lineal que incorpora la estimación de las inversas del ratio de Mills estimada por MCO)							
sexo		-1131.163	1992.859	-0.57	0.570	-5037.095	2774.768
edad		69.46866	130.2752	0.53	0.594	-185.8661	324.8034
publico		8265.101	23541.88	0.35	0.726	-37876.14	54406.34
lambda		11947.89	38908.98	0.31	0.759	-64312.3	88208.08
_cons		-9588.518	34092.69	-0.28	0.779	-76408.96	57231.93
-----+-----							

Se pide:

- Es la variable “contribuciones a la seguridad social” una variable a) Censurada, b) Truncada c) Cualquiera de las dos anteriores (Justifique su opinión)
- Obtenga las conclusiones que, respecto al efecto de la edad en la contribución a la seguridad social, se puedan extraer de las estimaciones realizadas.
- Sabiendo que $\Phi(\bar{x}'\hat{\beta}_{Tobit}) = 0.61$ y $\phi(\bar{x}'\hat{\beta}_{Probit}) = 0.384$ calcule los efectos parciales asociados a la variable edad, referidos a los siguientes momentos:
 - $E(\text{Aportes}|x)$
 - $\Pr(\text{Cotiza}=1|x)$
 - $E(\text{Aportes}|x, \text{COTIZA}=1)$
- Indique como interpreta los resultados obtenidos en 3).
- ¿Está correctamente identificado el modelo de Heckman? Justifique su respuesta. En caso negativo señale bajo qué condiciones el modelo estaría correctamente identificado.

Pregunta 1 (12 puntos) (extraído de Wooldridge 15.1)

Consideremos un modelo simple para estimar el efecto de tener un computador personal (PC) sobre el promedio de calificaciones de los estudiantes de una gran universidad pública:

$$Nota_i = \beta_0 + \beta_1 PC_i + u_i$$

Donde $Nota$ es el promedio de las calificaciones del alumno i y PC es una variable binaria que indica si el alumno tiene o no un PC.

- 1) ¿Por qué podría estar PC correlacionada con el error?
- 2) Suponga que PC es endógena, ¿sería la variable “ingresos de los padres del alumno” un buen instrumento? Razone ofreciendo argumentos a favor y en contra, concluya.
- 3) Supongamos que, hace 4 años, la universidad concedió becas para comprar computadoras para aproximadamente la mitad de los estudiantes de primer año, y los alumnos que la recibieron fueron elegidos al azar. Explicar con detalle como utilizaría esta información para obtener una estimación consistente de β_1 .

Pregunta 2 (12 puntos)

Se desea estudiar la influencia de los gastos en I+D de las empresas en la obtención de patentes por parte de las mismas. Se dispone de información sobre la cantidad de patentes obtenidas en el año por parte de cada empresa, el gasto en I+D, las ventas y el personal ocupado. Un investigador propone estimar el modelo suponiendo que la distribución de probabilidad de obtener patentes sigue una distribución Poisson. El modelo para la esperanza condicional es:

$$E(y_i | x_i) = \exp(x_i' \beta)$$

donde $y_i = \text{patentes}_i$; $x_i = (I+D_i, \text{ventas}_i, \text{personal_ocupado}_i, \text{constante})'$

La función de cuantía está dada por $f(y_i | x_i) = \exp(-\mu_i) \frac{\mu_i^{y_i}}{y_i!}$ con $\mu_i = \exp(x_i' \beta)$

- 1) Plantee la función de verosimilitud, para una muestra de tamaño N de observaciones i.i.d. Formule la log-verosimilitud, plantee el problema de maximización y encuentre las condiciones de primer orden para hallar el estimador máximo verosímil de β .
- 2) Derive el efecto parcial asociado al gasto en I+d en el modelo e indique como lo evaluaría.
- 3) ¿Qué estimador alternativo al máximo verosímil (MV) conoce para el modelo propuesto? Explique en qué consiste dicho estimador y cuáles son las condiciones de momentos que dan lugar al mismo. ¿Es este estimador alternativo preferible al MV? ¿Por qué?

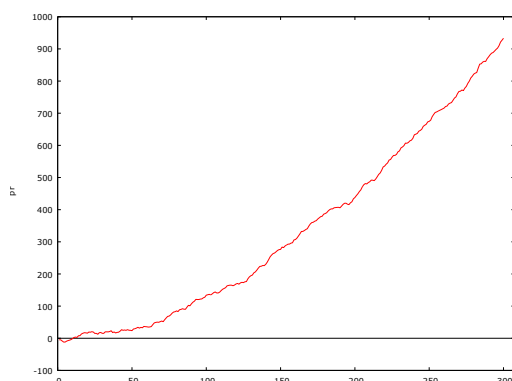
Pregunta 3 (12 puntos)

Considere el siguiente paseo aleatorio con deriva:

Modelo 1 $y_t = 0.3 + y_{t-1} + e_t$, con $t=1,2,\dots,T$ y $e_t \sim iid(0, \sigma^2)$

Se pide 1) Muestre que la media de este proceso es creciente en el tiempo e indique la magnitud esperada del cambio de y en cada período t .

El investigador A afirma que el modelo 1 es el adecuado para explicar la trayectoria del precio de un determinado activo, a partir de lo que se puede observar en la siguiente gráfica (300 observaciones) y el resultado del test de Dickey-Fuller Aumentado. Utilizando el criterio BIC se concluye que la cantidad de rezagos más apropiada es 0. Los resultados de la estimación para contrastar la existencia de una raíz unitaria en un modelo con constante:



Contraste de Dickey-Fuller
modelo: $(1-L)y = b_0 + (a-1)y(-1) + e$
valor estimado de $(a-1)$: 0,00627289
Estadístico de contraste: $\tau_c(1) = 3,85012$
Valor $p = 1$

Se pide 2) A partir de la evidencia disponible hasta este punto, ¿comparte usted la opinión del investigador A? Fundamente.

Un segundo investigador B considera que esta serie está caracterizada por una tendencia temporal determinística, pero considera necesario someter a prueba la existencia de una raíz unitaria en un modelo con tendencia determinística (**Modelo 2**). Para ello realiza un test de Dickey-Fuller. A continuación los resultados del test DF con constante y tendencia

Estadístico de contraste: $\tau_{ct}(1) = -1,01218$
Valor p 0,9396
Regresión de Dickey-Fuller
MCO, usando las observaciones 2-300 ($T = 299$)
Variable dependiente: d_{pr}

	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	0,0563986	0,534388	0,1055	0,9160
pr_1	-0,002883	0,002848	-1,012	0,9396
time	0,0267303	0,009229	2,896	0,0041 ***

Se pide 3) Concluya a partir de toda la evidencia disponible cuál es el mejor modelo para caracterizar la evolución del precio del activo. Justifique rigurosamente.