



UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO
DIVISIÓN ACADÉMICA DE CIENCIAS BÁSICAS
LICENCIATURA EN ACTUARÍA



ECONOMETRÍA

**MODELO ECONOMETRICO DE LA DEMANDA DEL SEGURO DE GASTOS
MÉDICOS MAYORES EN TABASCO 2005-2019.**

MI-LAI YAMAMOTO JOO



21 DE JULIO DE 2021

Contenido

Introducción	2
Datos y variables	3
Variable dependiente	4
Variables explicativas	4
Media y Desviación estándar.....	5
Correlación	6
Autocorrelación.....	7
Prueba Durbin-Watson.....	7
Prueba Breusch Godfrey	8
Pruebas de raíz unitaria (test de Dickey-Fuller)	9
Heterocedasticidad	16
Método Gráfico.....	16
Histograma de los errores estimados	17
Q-norm de errores estimados	18
Prueba de Shapiro-Wilk	19
Prueba de Shapiro-Francia	19
Asimetría y Curtosis.....	19
Prueba Jarque – Bera	20
Prueba White.....	21
Prueba Breusch-Pagan.....	21
Criterio de Akaike.....	22
Pruebas para variables omitidas (Prueba de Ramsey)	23
Estimación por mínimos cuadrados ordinarios (MCO)	23
Explicación de la salida de un modelo de regresión lineal múltiple estimado en Stata	24
Pruebas de Hipótesis	25
Prueba de significancia individual	25
Prueba de significancia global	26
Conclusión	27
Base de datos	28

Introducción

Ante la pandemia muchas personas se vieron afectadas al desembolsar enormes cantidades de dinero por causa de Covid-19 en los hospitales y medicinas.

Contar con un seguro de gastos médicos mayores te brinda la protección financiera en caso de que tú o tus dependientes económicos necesiten atención médica por accidente o enfermedad.

El objetivo general de este proyecto es analizar e identificar cuáles son los principales factores macroeconómicos que determinan la demanda de los seguros de gastos médicos mayores en Tabasco para el periodo de 2005-2019.

El periodo de análisis se da por el acceso a los datos, del cual se logra recopilar información de la emisión de primas (variable dependiente) en el ramo gastos médicos mayores por la Comisión Nacional de Seguros y Fianzas (CNSF) y nuestras otras variables como son: siniestros ocurridos, siniestros pagados ,el Ingreso laboral per cápita a pesos corrientes (nominal) anual, defunciones, esperanza de vida, Índice de la Tendencia Laboral de la Pobreza (ITLP) ,nivel de ingresos: Más de 5 salarios mínimos, PIB, población ocupada, unidades médicas privadas, personas alfabetas femeninas 15-24, mediana del ingreso nominal de la población de 15 a 29 años ocupada que reportan ingreso, población económicamente activa, población en edad de trabajar, tasa de participación y el índice de sobremortalidad masculina, mediante el Instituto Nacional de Estadística y Geografía (INEGI) y El Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL).

Datos y variables

Variable	Definición
PE	Prima emitida
SIN_OC	Siniestros ocurridos
SIN_PAG	Siniestros pagados
ILPA	Ingreso laboral per cápita a pesos corrientes (nominal) anual
DEF	Defunciones
EV	Esperanza de vida
ITLP	Índice de la Tendencia Laboral de la Pobreza (ITLP)
T_NI5	Nivel de ingresos: Más de 5 salarios mínimos
PIB	PIB (Total millones de pesos)
POB_OCUPADA	Población Ocupada
UMP	Unidades médicas privadas
PFALF_15-24	Personas alfabetas femeninas 15-24
MINPO_15_29	Mediana del ingreso nominal de la población de 15 a 29 años ocupada que reportan ingreso
PEA_15	Población económicamente activa - 15 años y más
POET	Población en edad de trabajar (15 años y más)
TP_15	Tasa de participación - 15 años y más
ISM	Índice de sobremortalidad masculina

AÑO	PE	SIN_OC	SIN_PAG	ILPA	DEF	EV	ITLP	PIB	POB_OCUPADA	T_NI5
2005.0	68120390.0	31194696.0	30760775.0	1322.2	9050.0	75.3	1.1	421079.7	759149.5	71024.0
2006.0	57255631.0	38611631.0	38243274.0	1536.3	9244.0	75.4	0.9	445309.6	777830.5	99418.0
2007.0	68444386.0	93044278.0	95551681.0	1601.6	9403.0	75.3	1.0	454079.1	822573.8	86949.5
2008.0	95290504.0	47462663.0	48486861.0	1720.7	9809.0	75.2	0.9	475202.8	835810.0	91801.0
2009.0	132686000.8	54531601.2	56506856.7	1689.1	10432.0	75.1	1.0	495944.2	844133.3	93286.0
2010.0	154568656.0	79988066.0	81932810.0	1689.0	11511.0	74.7	1.0	525011.9	845015.3	72378.0
2011.0	153197552.9	77056600.8	84204247.8	1674.9	11067.0	74.8	1.0	549751.1	872238.3	71332.7
2012.0	252305034.6	83237701.0	81207323.7	1833.6	11189.0	74.8	1.1	564003.8	871406.3	80527.0
2013.0	235885051.7	85563136.7	93132585.1	1878.9	12145.0	74.9	1.1	553628.2	899111.0	81697.0
2014.0	272414589.5	124984553.1	133486367.4	1932.3	12548.0	74.8	1.0	564794.2	909862.0	77405.0
2015.0	268184684.0	150276680.0	139769670.0	1948.6	13160.0	74.5	1.0	562212.3	907599.3	53147.0
2016.0	283003061.0	141220440.0	129965108.0	1869.1	13690.0	74.7	1.1	534685.5	928967.3	34838.0
2017.0	346092764.0	127239522.0	131685185.0	1843.1	13369.0	74.7	1.3	509388.4	919710.0	28198.0
2018.0	342291361.0	111672078.0	117828226.0	1904.3	13965.0	74.8	1.3	466935.1	935431.3	30847.0
2019.0	79417077.0	35103127.0	37093226.0	2144.2	15147.0	74.9	1.2	448555.3	986906.5	31193.5

UMP	PFALF_15-24	MINPO_15_29	PEA_15	POET	TP_15	ISM
53.0	212965.0	2386.3	785104.5	1439087.3	54.6	138.8
52.0	213850.0	2585.0	806015.5	1459973.3	55.2	140.5
53.0	214607.0	2793.8	860673.3	1499773.8	57.4	146.2
50.0	215194.0	2988.5	875981.0	1530339.0	57.3	139.2
48.0	216959.0	3031.5	895718.0	1547496.5	57.9	141.0
47.0	217093.0	3082.8	913980.5	1567080.0	58.3	142.7
47.0	217044.0	3186.5	933016.3	1595473.0	58.5	142.4
49.0	216830.0	3440.0	923909.0	1614445.5	57.2	138.7
49.0	216473.0	3570.0	963094.3	1653401.5	58.2	136.5
49.0	216686.0	3902.5	972619.3	1682592.5	57.8	134.7
50.0	215969.0	4096.3	973243.8	1693388.0	57.5	135.3
50.0	215172.0	4150.0	1003588.5	1718577.5	58.4	136.5
48.0	214347.0	3935.0	988887.5	1748287.3	56.6	131.4
48.0	213613.0	4152.8	1008350.5	1776367.5	56.8	132.9
49.0	212913.0	4822.5	1062143.5	1804249.3	58.9	133.0

Variable dependiente

Primas Emitidas: son los valores cobrados por la compañía sobre riesgos asumidos, en un período determinado. Dichos valores corresponden a la sumatoria de primas que recibe la compañía por cualquier tipo de contrato de seguro.

Variables explicativas

- Siniestros ocurridos: son los siniestros que ocurren antes del cierre de la información financiera.
- Siniestros pagados: aquellos cuyas consecuencias económicas han sido completamente indemnizadas o reparadas por la entidad aseguradora.
- Ingreso laboral per cápita a pesos corrientes: hace referencia a todas las entradas económicas que recibe una persona, una familia, una empresa, una organización, etc.
- Defunciones: cantidad de personas fallecidas.
- Esperanza de Vida: cálculo demográfico sobre la edad media que pueden alcanzar los individuos de una población concreta en una época determinada.
- Índice de la Tendencia Laboral de la Pobreza (ITLP): que muestra trimestralmente la tendencia de la proporción de personas que no pueden adquirir la canasta alimentaria con el ingreso de su trabajo.
- Nivel de ingresos: Más de 5 salarios mínimos: es una variable que resulta fundamental para establecer la capacidad adquisitiva de la demanda potencial, así como para determinar desequilibrios de mercado o indicadores de esfuerzo financiero necesario para asumir el teórico coste de un préstamo hipotecario.
- El producto interior bruto (PIB): es un indicador económico que refleja el valor monetario de todos los bienes y servicios finales producidos por un país.
- Población Ocupada: está compuesta por las personas que tienen un empleo remunerado o ejercen una actividad independiente y han trabajado al menos una hora durante la semana.
- Unidades médicas privadas: aquéllas que no pertenecen a una institución del Sector Salud y que cuente con (y haya utilizado) camas de hospitalización censables.
- Personas alfabetas femeninas 15-24: personas del sexo femenino con capacidad de leer y escribir de 15 a 24 años.
- Mediana del ingreso nominal de la población de 15 a 29 años ocupada que reportan ingreso.
- Población económicamente activa: comprende a las personas de 15 años y más que durante el período de referencia estaban trabajando (ocupados) o buscando activamente un trabajo (desempleados).
- Población en edad de trabajar (PET). - Es aquella población definida por las normas internacionales (OIT), como apta en cuanto a edad para ejercer funciones productivas (de 15 años y más de edad).
- Tasa de participación - 15 años y más. es un indicador de la proporción de la población en edad de trabajar de un país que participa activamente en el mercado de trabajo, ya sea trabajando o buscando empleo.
- Índice de sobremortalidad masculina: es la relación de las defunciones masculinas por cada cien defunciones femeninas. Las cifras se refieren a las defunciones por entidad federativa de residencia habitual de la persona fallecida.

Media y Desviación estándar.

. sum

Variable	Obs	Mean	Std. Dev.	Min	Max
ao	15	2012	4.472136	2005	2019
pe	15	1.87e+08	1.03e+08	5.73e+07	3.46e+08
sin_oc	15	8.54e+07	3.92e+07	3.12e+07	1.50e+08
sin_pag	15	8.67e+07	3.81e+07	3.08e+07	1.40e+08
ilpa	15	1772.532	199.1712	1322.16	2144.23
def	15	11715.27	1908.904	9050	15147
ev	15	74.92733	.2694304	74.5	75.43
itlp	15	1.058667	.1026691	.93	1.26
pib	15	504705.4	49857.22	421079.7	564794.3
pob_ocupada	15	874382.9	61504.93	759150	986907
t_ni5	15	66936.11	24852.51	28198	99418
ump	15	49.46667	1.9223	47	53
lpalf_1524	15	215314.3	1526.095	212913	217093
lminpo_15_29	15	3474.883	688.338	2386.25	4822.5
lpea_15	15	931088.3	76912.19	785104.5	1062144
lpoet	15	1622035	114783	1439087	1804249
ltp_15	15	57.355	1.204204	54.55	58.875
lism	15	137.9867	4.172678	131.4	146.2

Figura 1.0

. sum lpe lump lpalf_1524 lminpo_15_29 lpea_15 lpoet

Variable	Obs	Mean	Std. Dev.	Min	Max
lpe	15	18.87602	.6384853	17.86304	19.66222
lump	15	3.900605	.0384172	3.850147	3.970292
lpalf_1524	15	12.27983	.0070946	12.26864	12.28808
lminpo_15_29	15	8.13477	.200347	7.777478	8.481048
lpea_15	15	13.74084	.0843036	13.57357	13.8758
lpoet	15	14.29684	.0710444	14.17952	14.40565

Figura 1.05

Correlación

Necesitamos saber si existe una relación entre nuestra variable dependiente (Las primas emitidas) respecto a todas las otras variables como son: siniestros ocurridos, los siniestros pagados, el Ingreso laboral per cápita a pesos corrientes (nominal) anual, las defunciones, la esperanza de vida, Índice de la Tendencia Laboral de la Pobreza (ITLP), las unidades médicas privadas, etc. Por ello realizamos una correlación múltiple.

En la figura 1.1 se muestra la correlación múltiple de todas nuestras variables y en la figura 1.2 aplicamos logaritmo a cada una.

```
. correlate pe sin_oc sin_pag ilpa def ev itlp pib pob_ocupada t_ni5 ump pfalf_1524 minpo_15_29 pea_15 poet tp_15 ism
(obs=15)
```

	pe	sin_oc	sin_pag	ilpa	def	ev	itlp	pib	pob_oc~a	t_ni5	ump	pfa~1524	minpo~29
pe	1.0000												
sin_oc	0.8329	1.0000											
sin_pag	0.8526	0.9890	1.0000										
ilpa	0.5481	0.4636	0.4744	1.0000									
def	0.6495	0.5291	0.5333	0.8916	1.0000								
ev	-0.7763	-0.7554	-0.7494	-0.6838	-0.7652	1.0000							
itlp	0.6062	0.2574	0.3017	0.4834	0.7296	-0.4660	1.0000						
pib	0.6175	0.6398	0.6485	0.4491	0.3250	-0.7465	-0.0171	1.0000					
pob_ocupada	0.6200	0.5151	0.5334	0.9507	0.9620	-0.7353	0.6536	0.3961	1.0000				
t_ni5	-0.5769	-0.4749	-0.4606	-0.5705	-0.8521	0.6260	-0.8265	-0.0202	-0.7507	1.0000			
ump	-0.4976	-0.2403	-0.2988	-0.5093	-0.5202	0.6618	-0.4237	-0.5560	-0.5568	0.3015	1.0000		
pfalf_1524	0.2317	0.2882	0.3103	0.1006	-0.0983	-0.4020	-0.3636	0.8255	0.0126	0.4117	-0.5410	1.0000	
minpo_15_29	0.6073	0.5195	0.5184	0.9449	0.9783	-0.7205	0.6437	0.3347	0.9785	-0.7910	-0.4465	-0.0955	1.0000
pea_15	0.6431	0.5395	0.5582	0.9389	0.9677	-0.7756	0.6581	0.4304	0.9961	-0.7573	-0.6042	0.0582	0.9695
poet	0.7118	0.5627	0.5805	0.9096	0.9821	-0.7553	0.7555	0.3546	0.9790	-0.8308	-0.5419	-0.0746	0.9768
tp_15	0.1849	0.2778	0.2958	0.6695	0.5274	-0.5572	0.0610	0.5476	0.6604	-0.1938	-0.6120	0.5313	0.5523
ism	-0.6535	-0.3700	-0.3801	-0.6659	-0.7837	0.4912	-0.7485	-0.1327	-0.7080	0.7295	0.2769	0.3335	-0.7761
	pea_15	poet	tp_15	ism									
pea_15	1.0000												
poet	0.9778	1.0000											
tp_15	0.6828	0.5153	1.0000										
ism	-0.6887	-0.8029	-0.0096	1.0000									

Figura 1.1

```
. correlate lpe lsin_oc lsin_pag lilpa ldef lev litlp lpob_ocupada lt_ni5 lpib lump lpfalf_1524 lminpo_15_29 lpea_15 lpoet ltp_15 lism
(obs=15)
```

	lpe	lsin_oc	lsin_pag	lilpa	ldef	lev	litlp	lpob_o~a	lt_ni5	lpib	lump	lpf~1524	lminp~29	lpea_15	lpoet	ltp_15
lpe	1.0000															
lsin_oc	0.8375	1.0000														
lsin_pag	0.8430	0.9951	1.0000													
lilpa	0.5978	0.4897	0.5074	1.0000												
ldef	0.6804	0.5156	0.5262	0.8852	1.0000											
lev	-0.8376	-0.7499	-0.7473	-0.6952	-0.8000	1.0000										
litlp	0.5401	0.2379	0.2643	0.4623	0.7266	-0.4827	1.0000									
lpob_ocupada	0.6495	0.5195	0.5415	0.9480	0.9612	-0.7487	0.6468	1.0000								
lt_ni5	-0.4719	-0.3430	-0.3370	-0.5438	-0.8177	0.5570	-0.8375	-0.7299	1.0000							
lpib	0.7444	0.7077	0.7135	0.4974	0.3935	-0.7510	0.0118	0.4349	0.0395	1.0000						
lump	-0.5992	-0.3309	-0.3784	-0.5274	-0.5534	0.6563	-0.4301	-0.5653	0.2854	-0.5683	1.0000					
lpfalf_1524	0.4037	0.4155	0.4328	0.1545	-0.0324	-0.4019	-0.3432	0.0476	0.4661	0.8310	-0.5403	1.0000				
lminpo_15_29	0.6553	0.5246	0.5336	0.9477	0.9756	-0.7608	0.6349	0.9845	-0.7561	0.4171	-0.4892	-0.0118	1.0000			
lpea_15	0.6803	0.5497	0.5721	0.9383	0.9679	-0.7889	0.6493	0.9959	-0.7272	0.4728	-0.6157	0.0985	0.9771	1.0000		
lpoet	0.7081	0.5429	0.5601	0.9063	0.9833	-0.7687	0.7477	0.9795	-0.8153	0.3907	-0.5483	-0.0440	0.9830	0.9770	1.0000	
ltp_15	0.3309	0.3636	0.3949	0.6944	0.5543	-0.5588	0.0776	0.6776	-0.1619	0.5659	-0.6100	0.5345	0.5918	0.7024	0.5345	1.0000
lism	-0.5761	-0.2698	-0.2732	-0.6374	-0.7754	0.4920	-0.7511	-0.6982	0.7504	-0.1282	0.2658	0.3380	-0.7613	-0.6747	-0.7976	-0.0121
	lism															
lism	1.0000															

Figura 1.2

```
. correlate lpe lump lpfalf_1524 lminpo_15_29 lpea_15 lpoet
(obs=15)
```

	lpe	lump	lpf~1524	lminp~29	lpea_15	lpoet
lpe	1.0000					
lump	-0.5992	1.0000				
lpfalf_1524	0.4037	-0.5403	1.0000			
lminpo_15_29	0.6553	-0.4892	-0.0118	1.0000		
lpea_15	0.6803	-0.6157	0.0985	0.9771	1.0000	
lpoet	0.7081	-0.5483	-0.0440	0.9830	0.9770	1.0000

Figura 1.3

Dado que existe una correlación positiva entre nuestras variables respecto a la variable dependiente, luego de varias pruebas, elegimos el modelo más adecuado. La figura 1.3 muestra la correlación de nuestro modelo de regresión.

Autocorrelación

Si existe un patrón sistemático de comportamiento en dicha gráfica, se presume que hay indicios de autocorrelación en los errores. Para generar los residuos rezagados en un periodo, en Stata hacemos uso del operador L, por lo tanto, $e_{-1} = L.e$

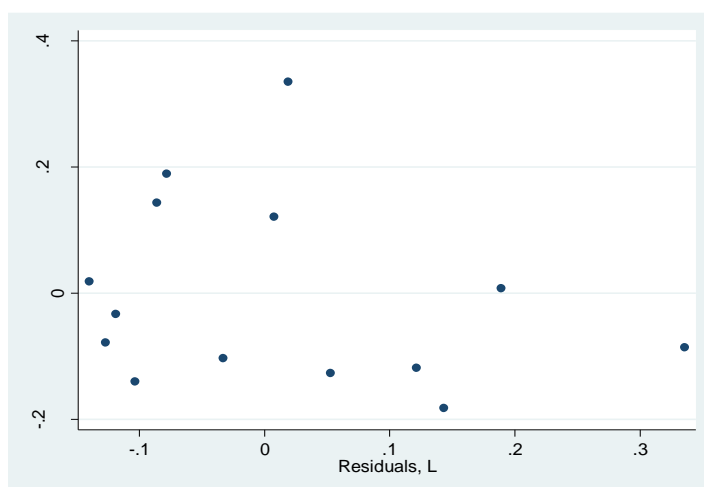


Figura 2.1

En la Figura 2.1 de residuos versus los residuos rezagados muestra un comportamiento no tan sistemático, por consiguiente, se puede pensar que no existe autocorrelación en los errores del modelo. Obsérvese que a medida que los residuos rezagados aumentan, los residuos disminuyen. No obstante, el método gráfico es muy subjetivo y por ello debe ser validado con pruebas formales, como los que se presentan a continuación.

Prueba Durbin-Watson

```
. estat dwatson
```

```
Durbin-watson d-statistic( 6, 15) = 2.32165
```

Se tiene las siguientes hipótesis:

H_0 : no existe autocorrelación

H_a : existe autocorrelación

Como el estadístico de prueba dw es mayor a 2, entonces no se rechaza H_0 , por lo que no existe autocorrelación

Prueba Breusch Godfrey

```
. estat bgodfrey, lags(1)
```

Breusch-Godfrey LM test for autocorrelation

lags(p)	chi2	df	Prob > chi2
1	1.371	1	0.2417

H0: no serial correlation

Figura 2.2

En la figura 2.2 se muestra un p-value igual a $0.2417 > 0.05$ entonces no rechazamos la hipótesis nula. Por lo que no hay problemas de autocorrelación.

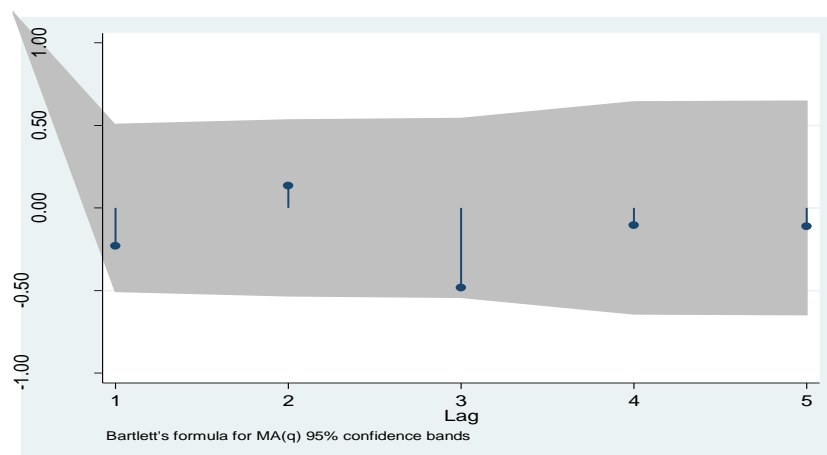


Figura 2.3 Autocorrelación de los errores.

La figura 2.3 muestra como se distribuyen los errores

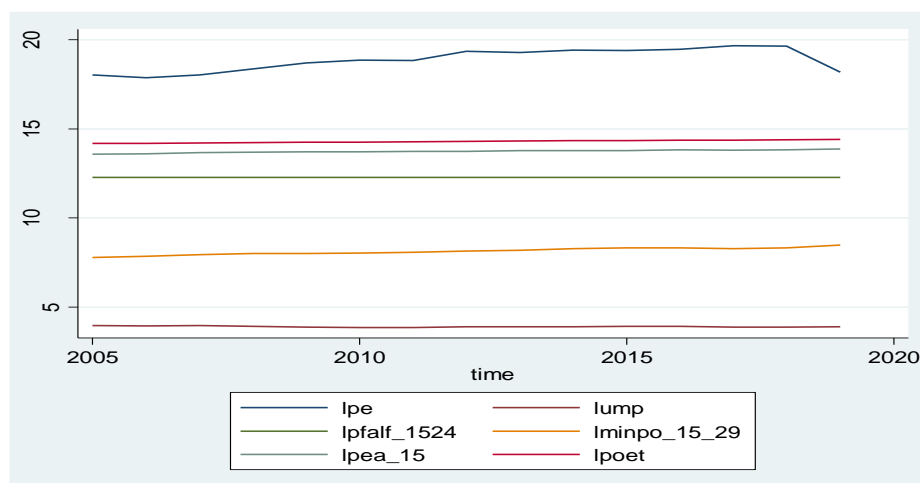


Figura 2.4

Pruebas de raíz unitaria (test de Dickey-Fuller)

Las hipótesis en esta prueba son $H_0: \delta = 0$ vs $H_a: \delta < 0$. Si no se rechaza la hipótesis nula, entonces la serie posee raíz unitaria, es decir, no es estacionaria. Si se rechaza, la serie es estacionaria.

Graficamos la tendencia de todas nuestras variables



```
. dfuller pe, drift regress lags(3)
```

Augmented Dickey-Fuller test for unit root

Number of obs = 11

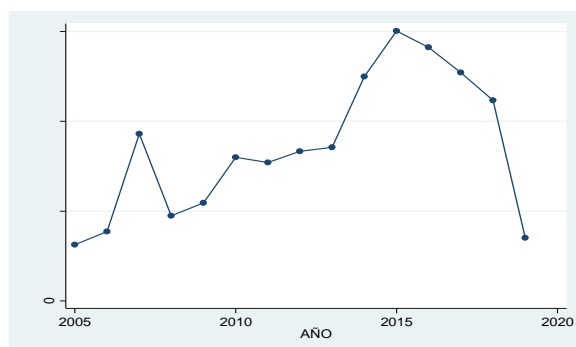
	Test Statistic	1% Critical Value	Z(t) has t-distribution	5% Critical Value	10% Critical Value
Z(t)	-1.434	-3.143		-1.943	-1.440

p-value for Z(t) = 0.1008

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
pe					
L1.	-.5210827	.3634282	-1.43	0.202	-1.410359 .368194
L2D.	-.5539816	1.208438	-0.46	0.663	-3.510922 2.402958
L2D.	-1.336065	1.432815	-0.93	0.387	-4.842037 2.169907
L3D.	-.3717382	1.146226	-0.32	0.757	-3.176451 2.432975
_cons	1.75e+08	1.04e+08	1.68	0.143	-7.94e+07 4.29e+08

Figura 3.1 Primas Emitidas 2005-2019

La figura 3.1 presenta una tendencia clara positiva. Además, escogemos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p = 0.1008 > 0.05$ nivel de significancia del 5%, por tanto, no se rechaza la hipótesis nula y se concluye que la serie no es estacionaria.



```
. dfuller sin_oc, drift regress lags(3)
```

Augmented Dickey-Fuller test for unit root

Number of obs = 11

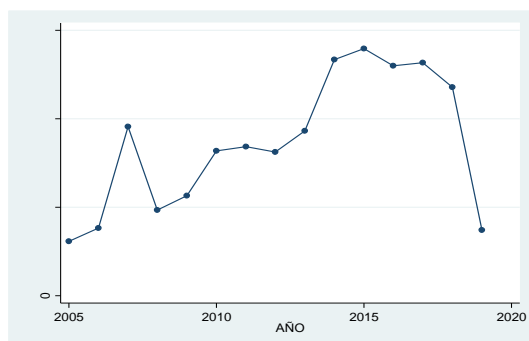
	Test Statistic	1% Critical Value	Z(t) has t-distribution	5% Critical Value	10% Critical Value
Z(t)	-2.701	-3.143		-1.943	-1.440

p-value for Z(t) = 0.0178

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
sin_oc					
L1.	-.6513241	.2411053	-2.70	0.036	-1.241288 -.0613607
L2D.	1.066075	.3855701	2.76	0.033	.1226193 2.009531
L2D.	.5585997	.3164239	1.77	0.128	-.2156617 1.332861
L3D.	.6299972	.3179039	1.98	0.095	-.1478856 1.40788
_cons	5.04e+07	2.36e+07	2.14	0.076	-7227652 1.08e+08

Figura 3.2 Siniestros Ocultos 2005-2019

La figura 3.2 no presenta una tendencia clara. Por lo que incorporamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p = 0.0178 < 0.05$ nivel de significancia del 5%, por tanto, se rechaza la hipótesis nula y se concluye que la serie es estacionaria.



```
. dfuller sin_pag, drift regress lags(3)
```

Augmented Dickey-Fuller test for unit root

Number of obs = 11

	Test Statistic	1% Critical Value	Z(t) has t-distribution	5% Critical Value	10% Critical Value
Z(t)	-2.278	-3.143		-1.943	-1.440

p-value for Z(t) = 0.0315

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
sin_pag					
L1.	-.627179	.2753716	-2.28	0.063	-1.300989 .0466311
L2D.	1.123469	.4909157	2.29	0.062	-.0777579 2.324697
L2D.	.5019385	.3907803	1.28	0.246	-.4542665 1.458143
L3D.	.6866321	.3629623	1.89	0.107	-.2015047 1.574769
_cons	4.88e+07	2.76e+07	1.77	0.127	-1.87e+07 1.16e+08

Figura 3.3 Siniestros Pagados 2005-2019

La figura 3.3 no presenta una tendencia clara. Incorporamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p = 0.0315 < 0.05$ nivel de significancia del 5%, por tanto, se rechaza la hipótesis nula y se concluye que la serie es estacionaria.

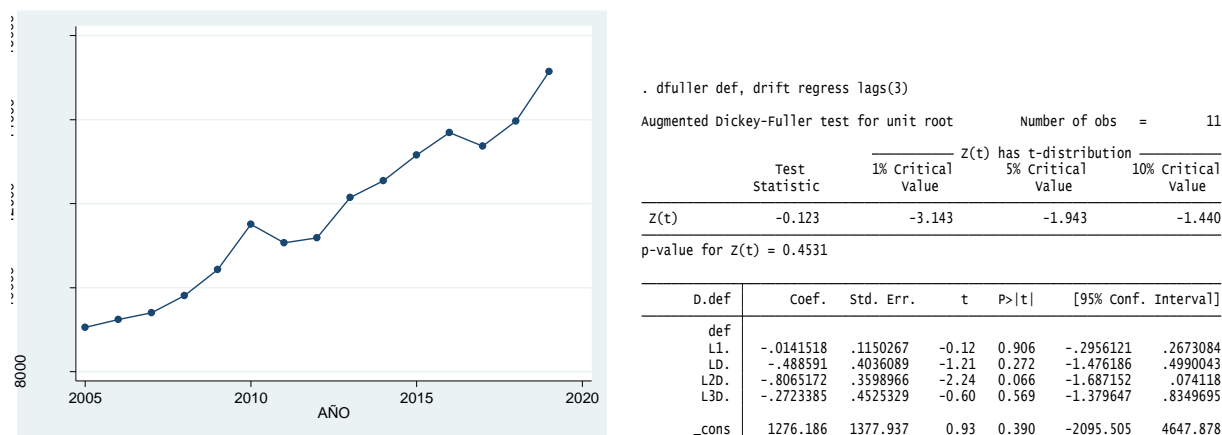


Figura 3.4 Defunciones 2005-2019

La figura 3.4 muestra una tendencia clara de las defunciones. Incorporamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p = 0.4531 > 0.05$ nivel de significancia del 5%, por tanto, no se rechaza la hipótesis nula y se concluye que la serie no es estacionaria.

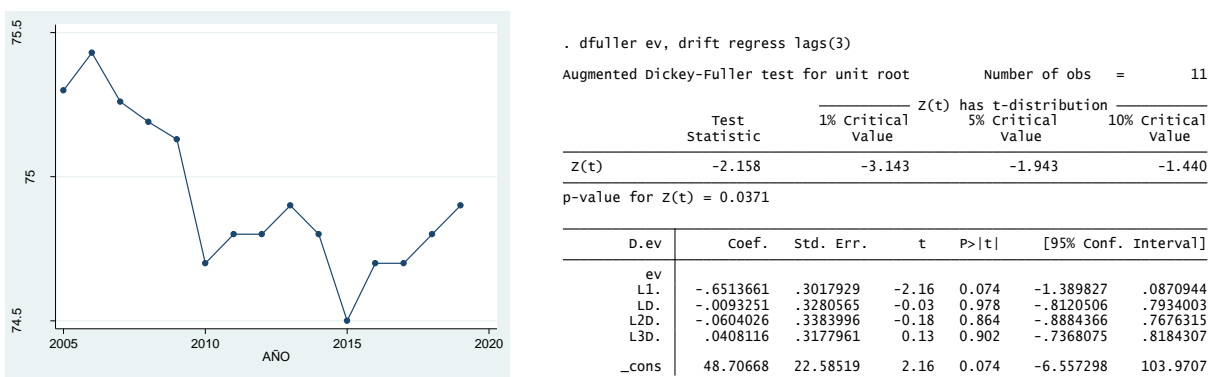


Figura 3.5 Esperanza de Vida 2005-2019

La figura 3.5 no muestra una tendencia clara. Incorporamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p = 0.0371 < 0.05$ nivel de significancia del 5%, por tanto, se rechaza la hipótesis nula y se concluye que la serie es estacionaria.

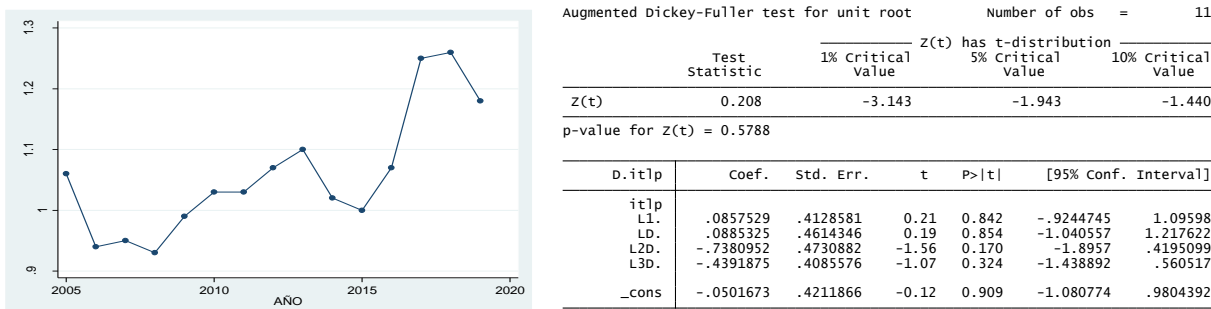


Figura 3.6 índice de Tendencia Laboral de Pobreza 2005-2019

La figura 3.6 parece una tendencia positiva, aunque no muy precisa. Incorporamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p = 0.5788 > 0.05$ nivel de significancia del 5%, por tanto, no se rechaza la hipótesis nula y se concluye que la serie no es estacionaria.

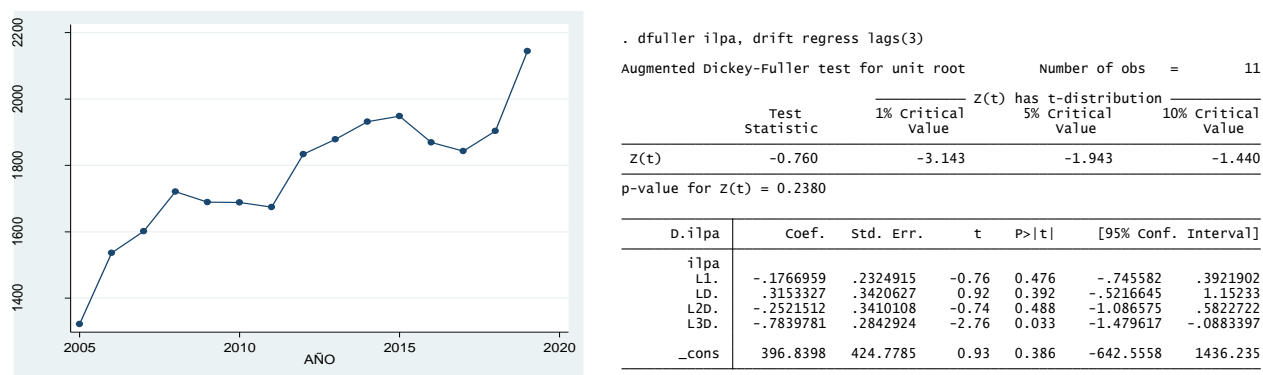


Figura 3.7 Ingreso Laboral per cápita 2005-2019

La figura 3.7 se observa el Ingreso Laboral per cápita 2005-2019. Incorporamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p = 0.2380 > 0.05$ nivel de significancia del 5%, por tanto, no se rechaza la hipótesis nula y se concluye que la serie no es estacionaria.

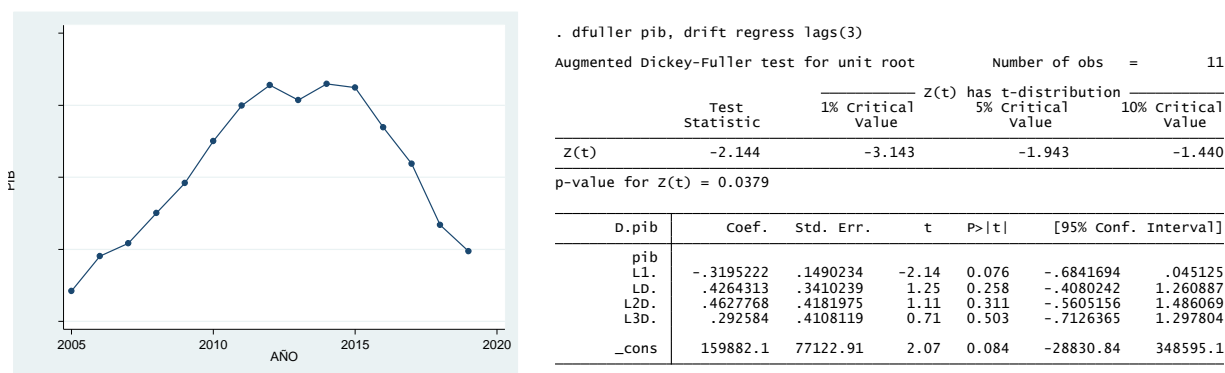


Figura 3.8 Producto Interno Bruto 2005-2019

La figura 3.8 no muestra una tendencia clara, por lo que implementamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p = 0.0379 < 0.05$ nivel de significancia del 5%, por tanto, se rechaza la hipótesis nula y se concluye que la serie es estacionaria.

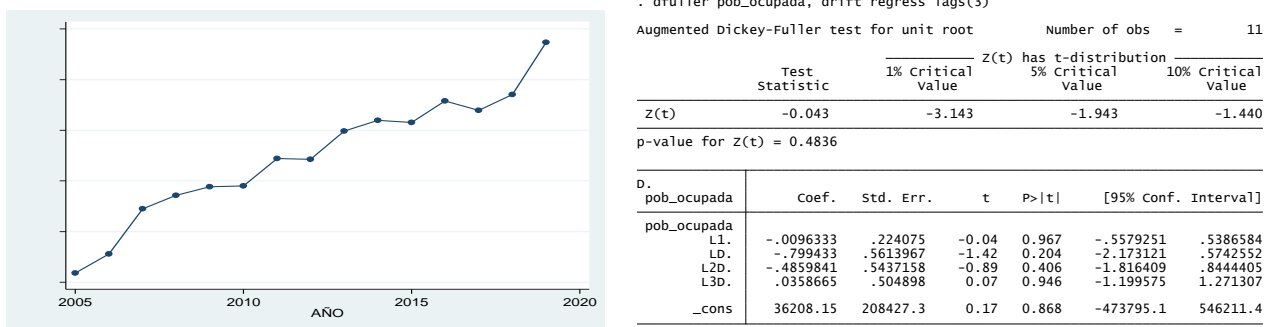
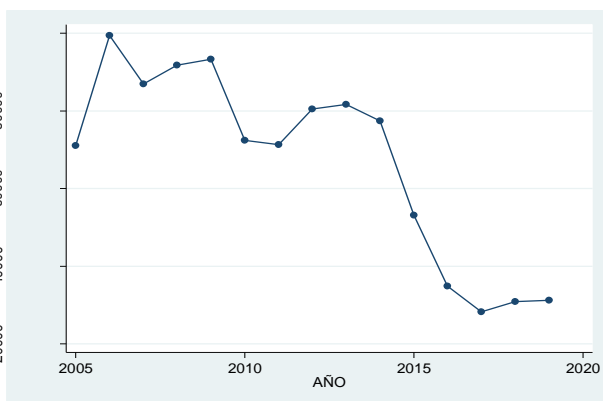


Figura 3.9 Población Ocupada 2005-2019

La figura 3.9 muestra que la Población Ocupada va aumentando respecto aumenta el tiempo. Además, implementamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p = 0.4836 > 0.05$ nivel de significancia del 5%, por tanto, no se rechaza la hipótesis nula y se concluye que la serie no es estacionaria.



```
. dfuller t_ni5, drift regress lags(3)
```

Augmented Dickey-Fuller test for unit root Number of obs = 11

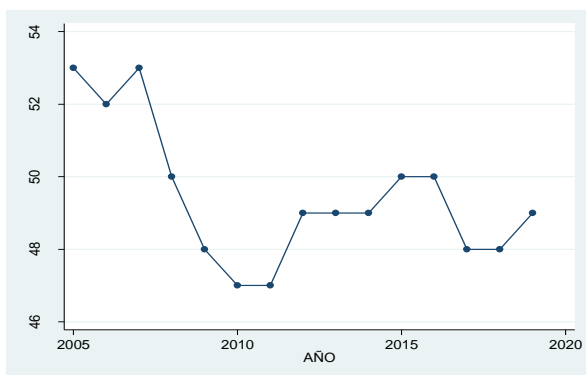
	Test Statistic	1% Critical Value	Z(t) has t-distribution	5% Critical Value	10% Critical Value
Z(t)	-0.091	-3.143		-1.943	-1.440

p-value for Z(t) = 0.4651

D.t_ni5	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
t_ni5					
L1.	-.0250571	.2744109	-0.09	0.930	-.6965164 .6464022
LD.	.3782561	.3996641	0.95	0.380	-.5996866 1.356199
L2D.	-.5137925	.4351042	-1.18	0.282	-1.578454 .5508692
L3D.	-.021362	.3345923	-0.06	0.951	-.8400799 .797356
_cons	-5347.673	22154.7	-0.24	0.817	-59558.28 48862.94

Figura 3.10 Nivel de Ingresos: más de 5 salarios mínimos. 2005-2019

La figura 3.10 muestra una tendencia decreciente. Además, implementamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p = 0.4651 > 0.05$ nivel de significancia del 5%, por tanto, no se rechaza la hipótesis nula y se concluye que la serie no es estacionaria



```
. dfuller ump, drift regress lags(3)
```

Augmented Dickey-Fuller test for unit root Number of obs = 11

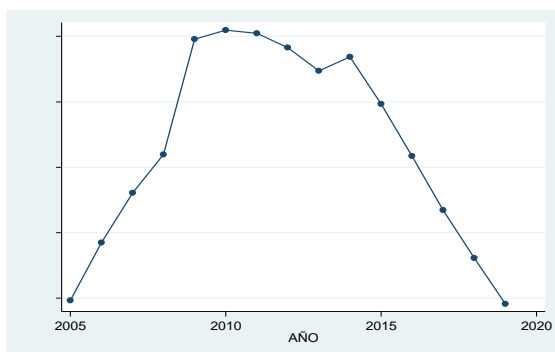
	Test Statistic	1% Critical Value	Z(t) has t-distribution	5% Critical Value	10% Critical Value
Z(t)	-3.015	-3.143		-1.943	-1.440

p-value for Z(t) = 0.0118

D.ump	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ump					
L1.	-1.048901	.3479318	-3.01	0.024	-1.900259 -.1975424
LD.	.4595125	.18433	2.49	0.047	.0084732 .9105517
L2D.	.202658	.2426484	0.84	0.436	-.3910813 .7963973
L3D.	.2458837	.2019125	1.22	0.269	-.2481783 .7399457
_cons	51.27343	17.01356	3.01	0.024	9.642754 92.90411

Figura 3.11 Unidades Médicas Privadas 2005-2019

La figura 3.11 no muestra una tendencia clara. Implementamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p = 0.0118 < 0.05$ nivel de significancia del 5%, por tanto, se rechaza la hipótesis nula y se concluye que la serie es estacionaria.



```
. dfuller pfalf_1524, drift regress lags(3)
```

Augmented Dickey-Fuller test for unit root Number of obs = 11

	Test Statistic	1% Critical Value	Z(t) has t-distribution	5% Critical Value	10% Critical Value
Z(t)	-1.582	-3.143		-1.943	-1.440

p-value for Z(t) = 0.0824

D.pfalf_1524	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
pfalf_1524					
L1.	-.3618723	.2288041	-1.58	0.165	-.9217359 .1979913
LD.	.376826	.3381965	1.11	0.308	-.4507111 1.204363
L2D.	.4185021	.3517316	1.19	0.279	-.4421541 1.279158
L3D.	.4082856	.3602008	1.13	0.300	-.473094 1.289665
_cons	77869.88	49384.9	1.58	0.166	-42970.61 198710.4

Figura 3.12 P. Alfabeta Fem. de 15-24 años 2005-2019

En la figura 3.12 se observa una tendencia creciente y luego decreciente de la población femenina que sabe leer y escribir de 15 a 24 años. Por eso implementamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p = 0.0824 > 0.05$ nivel de significancia del 5%, por tanto, no se rechaza la hipótesis nula, y la serie no estacionaria.

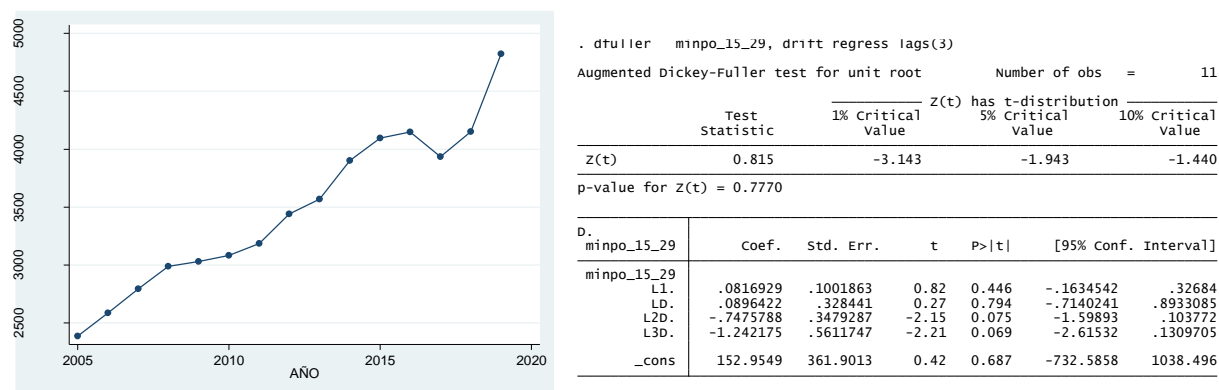


Figura 3.13 Mediana de Ingreso nominal de la población de 15-29 años. 2005-2019

En la figura 3.13 se observa una tendencia positiva, también implementamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p = 0.7770 > 0.05$ nivel de significancia del 5%, por tanto, no se rechaza la hipótesis nula, y la serie no estacionaria.

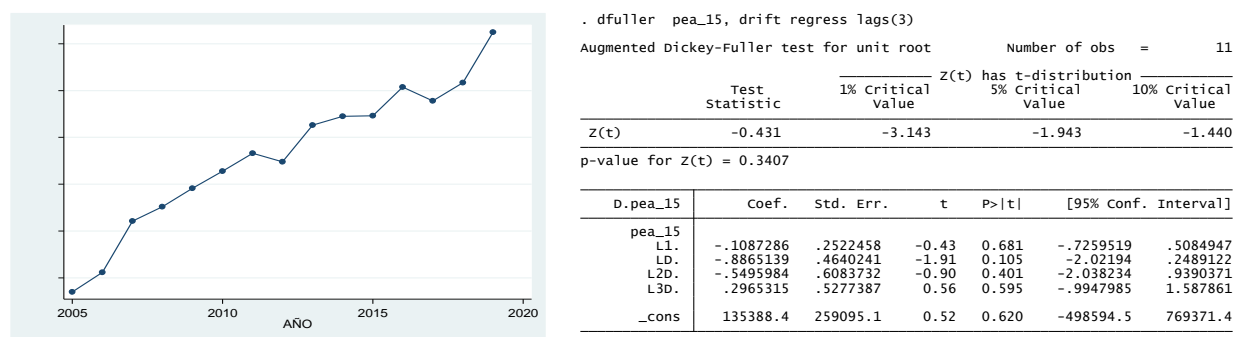


Figura 3.14 Población Económicamente activa 15+ años. 2005-2019

En la figura 3.14 se observa una tendencia positiva, también implementamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p = 0.7770 > 0.05$ nivel de significancia del 5%, por tanto, no se rechaza la hipótesis nula, y la serie no estacionaria.

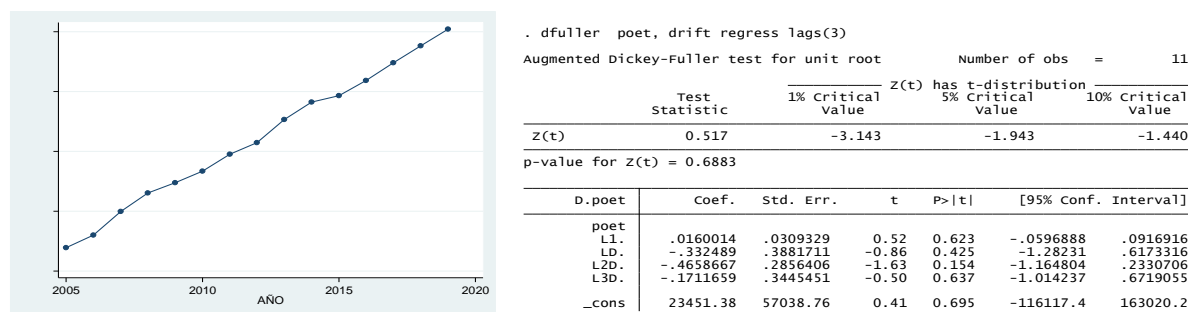


Figura 3.15 Población en edad de trabajar. 2005-2019

En la figura 3.15 se observa una tendencia positiva, también implementamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p=0.6883 > 0.05$ nivel de significancia del 5%, por tanto, no se rechaza la hipótesis nula, y la serie no estacionaria.

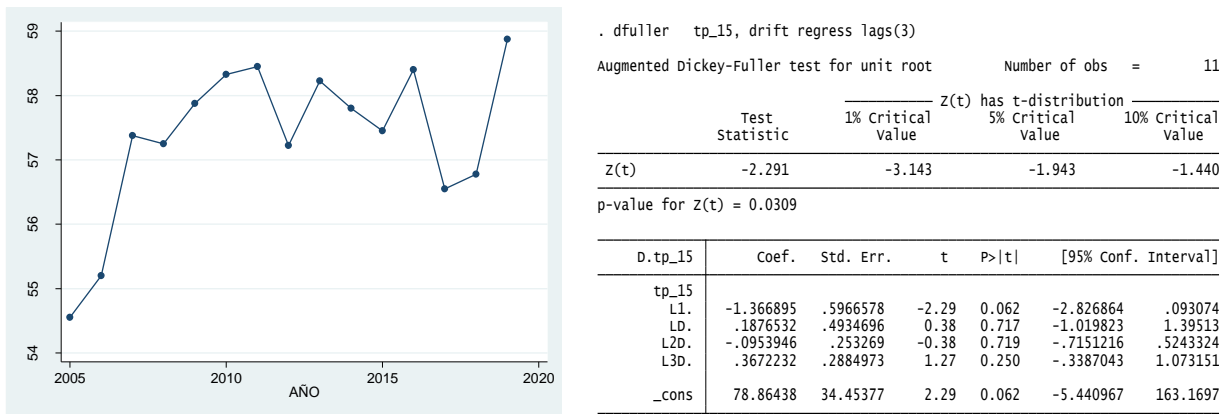


Figura 3.16 Tasa de participación 15+. 2005-2019

En la figura 3.16 se observa una tendencia positiva, también implementamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p=0.0309 < 0.05$ nivel de significancia del 5%, por tanto, se rechaza la hipótesis nula, y se concluye que la serie es estacionaria.

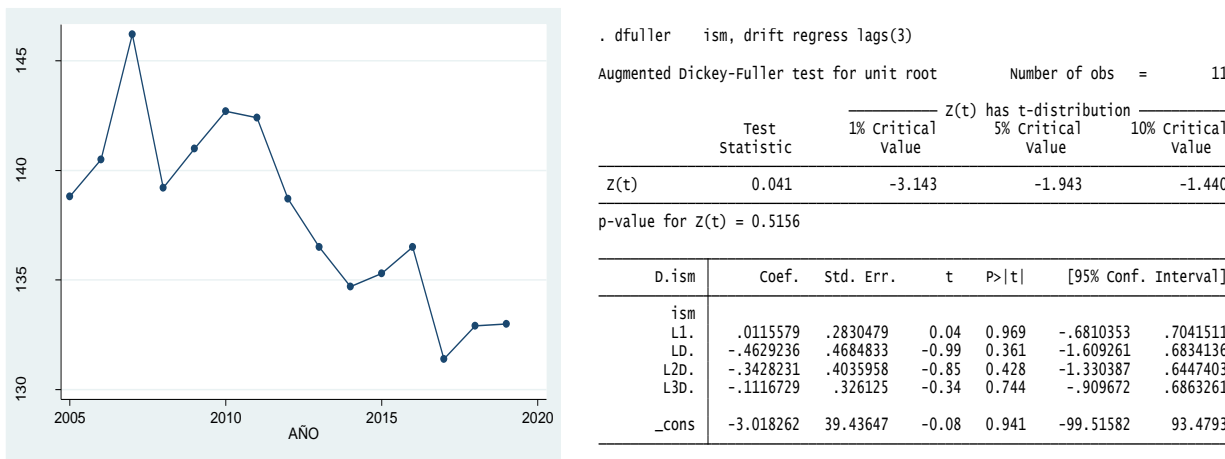


Figura 3.17 índice de Sobremortalidad Masculina 2005-2019

En la figura 3.17 se observa una tendencia negativa, además implementamos la prueba de Dickey-Fuller aumentado. Los resultados muestran un valor $p=0.5156 > 0.05$ nivel de significancia del 5%, por tanto, no se rechaza la hipótesis nula, y se concluye que la serie no es estacionaria.

Modelo Econométrico de la demanda de los Seguros de Gastos Médicos Mayores.

Se realizaron distintos modelos de regresión hasta llegar a uno adecuado. Usaremos series de tiempo donde la variable dependiente será la emisión de primas en los seguros de gastos médicos mayores, esta se encuentra en función de las unidades médicas privadas, personas alfabetas femeninas 15-24, mediana del ingreso nominal de la población de 15 a 29 años ocupada que reportan ingreso, población económicamente activa, y población en edad de trabajar (15 años y más).

El modelo para estimar la demanda de los seguros de Gastos Médicos Mayores en Tabasco para el periodo de 2005-2019 se plantea de la siguiente forma:

$$\text{Log}(pe) = \beta_0 + \beta_1 \log(ump) + \beta_2 \log(pfalf_1524) + \beta_3 \log(minpo_1529) + \beta_4 \log(pea15) + \beta_5 \log(poet)$$

En donde: β es el coeficiente estimado de los regresores y ϵ es el error.

Gráficos de correlación de las variables independientes con la variable dependiente:

Figura 4.1 $\log(ump)$

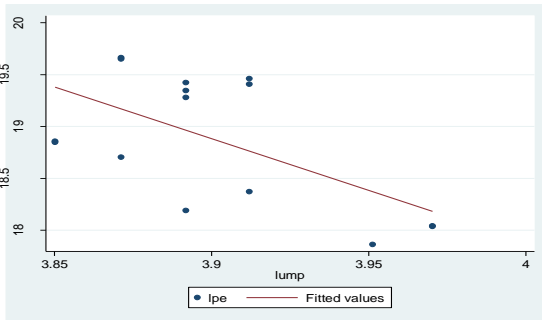


Figura 4.2 $\log(minpo_1529)$

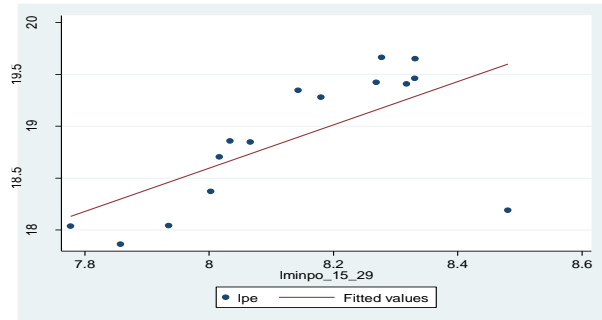


Figura 4.3 $\log(pfalf_1524)$

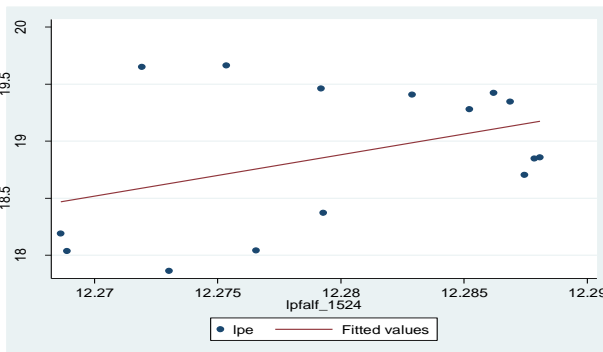


Figura 4.4 $\log(pea_15)$

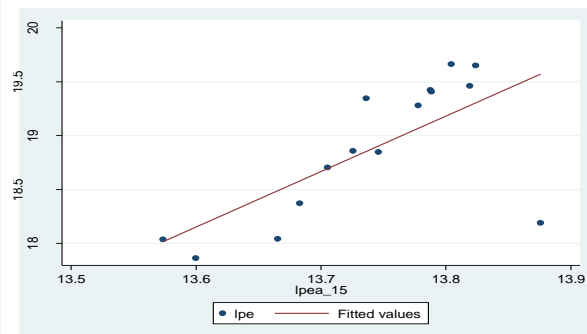
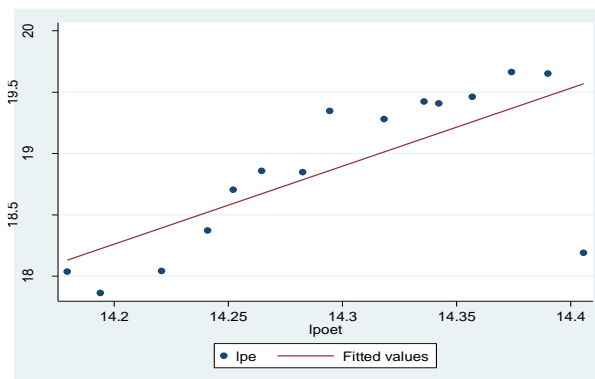


Figura 4.5 $\log(poet)$



Heterocedasticidad

A continuación, realizaremos las pruebas de detección de heterocedasticidad de los errores mediante diferentes métodos.

Método Gráfico

En la figura 5.1 se observan cambios en la variabilidad de los errores asociados a distintos momentos del tiempo.

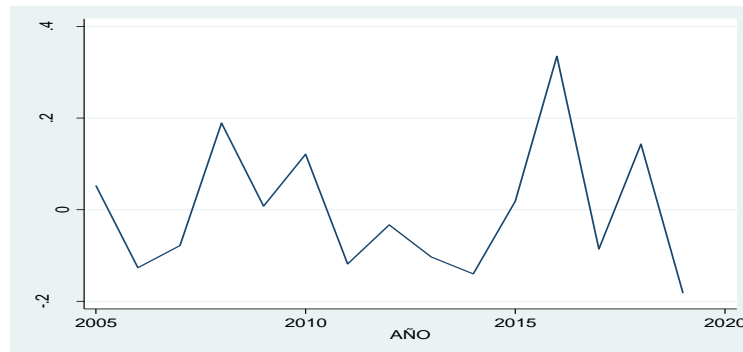


Figura 5.1 Errores

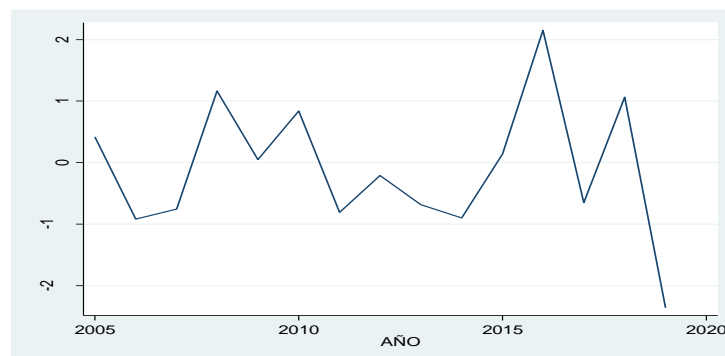


Figura 5.2 Residuos estudentizados

En la figura 5.2 Observamos a nuestros residuos estudentizados, dado que son superiores a 2 en valor absoluto se considera que la observación asociada es atípica, lo que implica un cuidado especial en el análisis.

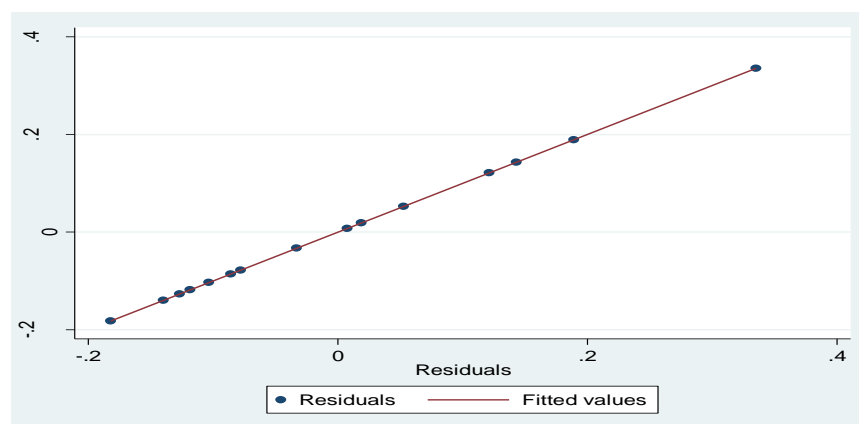


Figura 5.3

En la figura 5.3 Observamos como los errores siguen una distribución normal.

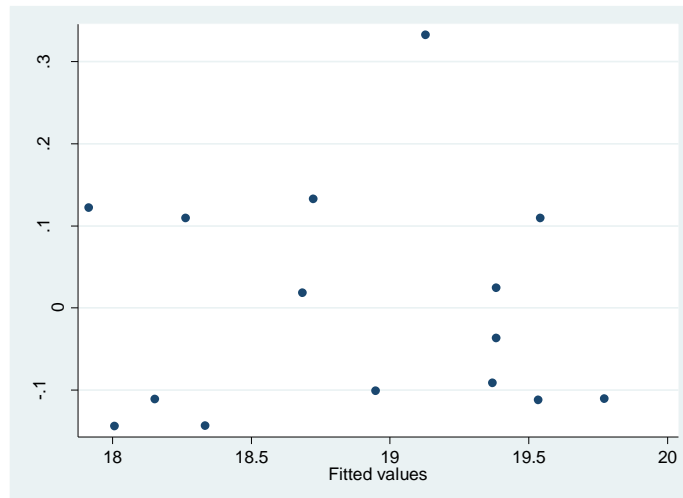


Figura 5.4 Residuals versus fitted values

Histograma de los errores estimados

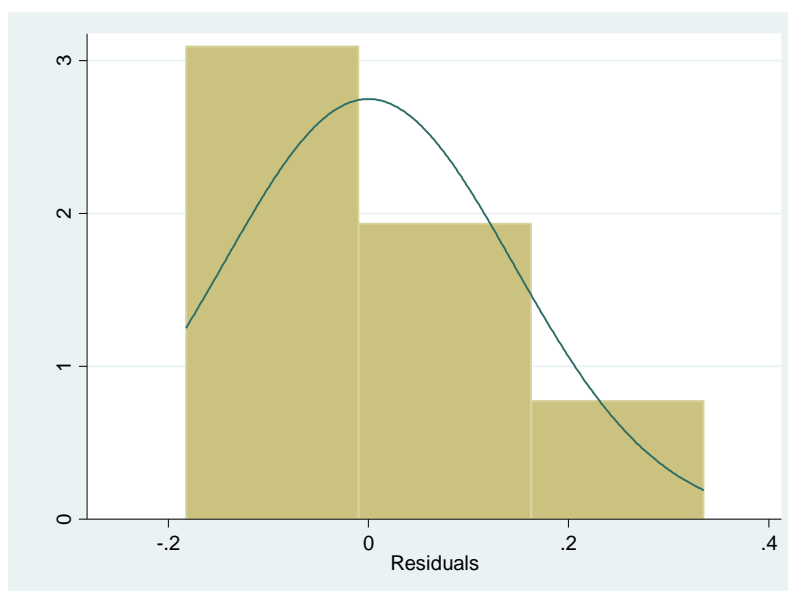


Figura 5.5

```
. hist error,normal
(bin=3, start=-.18229726, width=.17255204)
```

En la figura 5.5, observamos el histograma de los errores estimados, aparentemente se distribuyen normal.

Q-norm de errores estimados

La figura 5.6 representa la distribución normal acumulada en forma de recta. Sobre ella se grafican ordenados los errores estimados de forma creciente, como estos se aproximan a la línea de la distribución normal, podemos decir que los errores estimados tienen distribución normal.

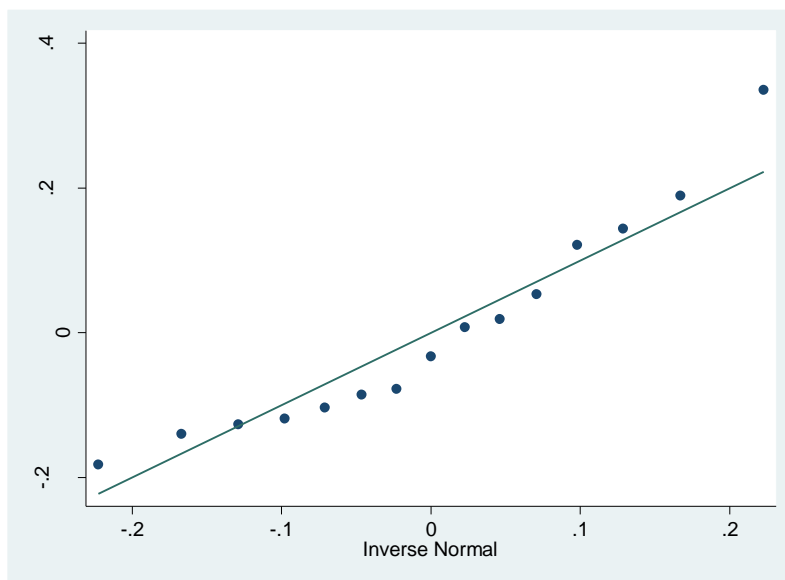


Figura 5.6.

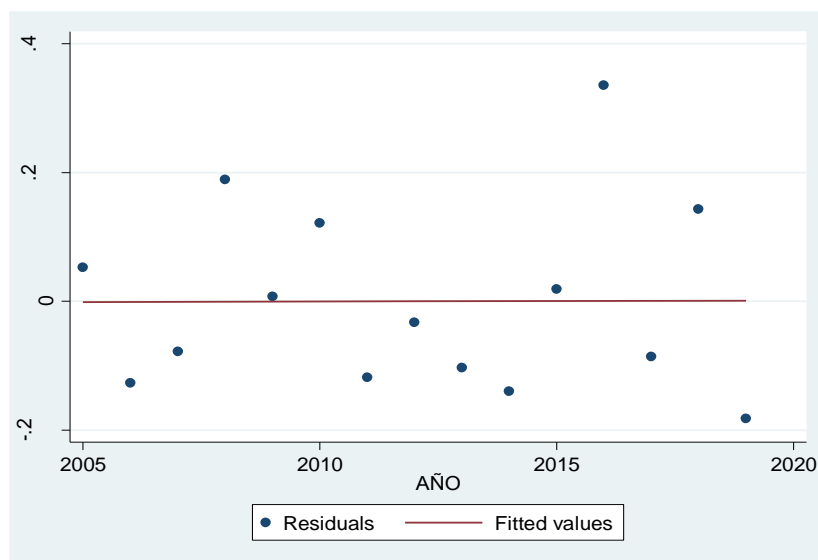


Figura 5.7

La figura 5.7 presenta una gráfica de los errores estimados en función del tiempo para comprobar la aleatoriedad de estos y que no presenten secuencias temporales provenientes de la posible temporalidad de los datos.

Prueba de Shapiro-Wilk

Las hipótesis para contrastar son: H0: los errores tienen distribución normal vs. Ha: los errores no son normales.

```
. swilk error
```

Shapiro-wilk w test for normal data					
variable	Obs	w	V	z	Prob>z
error	15	0.92365	1.480	0.776	0.21889

Figura 5.8

En la tabla que muestra la figura 3.8 se observa un valor p de 0.21889 mayor que el nivel de significancia del 5%, por lo que H0 no se rechaza.

Prueba de Shapiro-Francia

```
. sfrancia error
```

Shapiro-Francia w' test for normal data					
Variable	Obs	w'	V'	z	Prob>z
error	15	0.92517	1.609	0.827	0.20424

Figura 5.9

En esta prueba obtenemos un valor p de 0.20424 mayor que el nivel de significancia del 5%, entonces no rechazamos la hipótesis nula, lo que nos indica que si existe normalidad. (véase figura 5.9)

Asimetría y Curtosis

```
. sktest error
```

Skewness/Kurtosis tests for Normality						
variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj	joint	
error	15	0.1051	0.5760	chi2(2)	Prob>chi2	
				3.44	0.1788	

Figura 6.1

Los resultados de la figura 6.1 muestran un valor p de 0.1788 mayor que el nivel de significancia del 5% podemos concluir que efectivamente existe normalidad de los errores.

En la figura 6.2 se observa un valor de 0.8364844 para la asimetría y 2.884218 Curtosis.

```
. scalar list
      m3 = .00230621
      m2 = .01966197
      m4 = .00111502

. scalar q1= m3/m2^(2/3)      Asimetría = .8364844
. scalar q2= m4/m2^(2)        Curtosis = 2.884218

. dis q2
2.884218
```

```
. sum error, d
```

Residuals				
	Percentiles	Smallest		
1%	-.1822973	-.1822973		
5%	-.1822973	-.1399013		
10%	-.1399013	-.1269562	Obs	15
25%	-.1186612	-.1186612	Sum of wgt.	15
50%	-.032987		Mean	3.10e-11
		Largest	Std. Dev.	.1451427
75%	.1213062	.1213062		
90%	.1890614	.1433331	Variance	.0210664
95%	.3353589	.1890614	Skewness	.8364844
99%	.3353589	.3353589	Kurtosis	2.884218

Figura 6.2

Prueba Jarque – Bera

La hipótesis para contrastar es la siguiente:

H0: $\gamma_1 = 0$ y $\gamma_2 = 3$ son asintóticamente independientes y siguen una distribución $N(0, 1)$ vs. Ha: no son normales

Bajo la hipótesis nula, el estadístico de prueba JB se distribuye aproximadamente como una v.a. χ^2 con dos grados de libertad. Con nivel de significancia 5%.

```
. scalar jb=(_N)*(0.8364844^2/6+((2.884218-3)^2)/24)

. display jb
1.7576438

. display "valor critico= " invchi2(2,0.95)
valor critico= 5.9914645

. display "p-value= " chi2tail(2,jb)
p-value= .41527186
```

Figura 6.3

Según la figura 6.3, dado que el p-value $0.4152 > 0.05$ no se rechaza H0. Además, el valor obtenido para el estadístico de prueba JB* es menor que el valor crítico tabulado ($\chi^2 = 5.9914$ para un nivel de significación del 5%), por lo tanto, los errores son asintóticamente independientes y siguen una distribución $N(0, 1)$.

Prueba White

Es una prueba general para probar heterocedasticidad, donde la hipótesis nula es que no hay heterocedasticidad versus la hipótesis alterna, que considera que hay heterocedasticidad.

$H_0: \text{Var}(\varepsilon)=0$ vs $H_a: \text{Var}(\varepsilon) \neq 0$

```
. quietly reg lpe lump lpfal_f_1524 lminpo_15_29 lpea_15 lpoet
. imtest, white

White's test for Ho: homoskedasticity
  against Ha: unrestricted heteroskedasticity

      chi2(14)    =    15.00
      Prob > chi2  =    0.3782

Cameron & Trivedi's decomposition of IM-test
```

Source	chi2	df	p
Heteroskedasticity	15.00	14	0.3782
Skewness	6.51	5	0.2597
Kurtosis	0.04	1	0.8502
Total	21.55	20	0.3657

Figura 6.4

Los resultados de la figura 6.4 arrojan un valor $p = 0.3782 > 0.05$, por lo tanto no se rechaza H_0 . Es decir, la varianza de los errores es constante, por lo tanto, no hay presencia de heterocedasticidad.

Prueba Breusch-Pagan

La hipótesis para contrastar es $H_0 : \alpha_2 = \alpha_3 = \dots = \alpha_p = 0$, vs. $H_a : \alpha_i \neq 0$ para alguna $i = 2, \dots, p$.

H_0 : no hay heterocedasticidad vs H_a : hay heterocedasticidad.

```
. display e(rss)
.29492958

. gen e2=error^2

. gen e2_bp= e2 / (e(rss)/e(N))
```

```
. reg e2_bp error_est
```

Source	SS	df	MS
Model	.005493965	1	.005493965
Residual	28.2577714	13	2.17367472
Total	28.2632654	14	2.01880467

Number of obs = 15
 F(1, 13) = 0.00
 Prob > F = 0.9607
 R-squared = 0.0002
 Adj R-squared = -0.0767
 Root MSE = 1.4743

e2_bp	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
error_est	.0318603	.6337297	0.05	0.961	-1.33723 1.40095
_cons	.3986045	11.96835	0.03	0.974	-25.45744 26.25465

Figura 6.5

```

. display e(mss)
.00549396

. display "Estadístico de prueba = " e(mss) / 2
Estadístico de prueba = .00274698

. scalar est_prb=e(mss) / 2

. display "p-value= " chi2tail(1,e(mss) / 2)
p-value= .95820068

. display "valor critico= " invchi2(1,0.05)
valor critico= .00393214

. quietly reg lpe lump lpfalf_1524 lminpo_15_29 lpea_15 lpoet

. hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: fitted values of lpe

chi2(1)      =      0.00
Prob > chi2   =      0.9582

```

Figura 6.6 Prueba de Breusch-Pagan-Godfrey.

Los resultados presentados en la figura 6.6 muestran un valor $p=0.9582$, lo que implica que los errores del modelo sobre la demanda de las primas emitidas no presentan heterocedasticidad, pues no se rechaza la hipótesis nula a un nivel de significancia del 5%.

Criterio de Akaike

$$AIC = 2(-L(\text{likelihood}) + k)$$

Donde k es el número de regresores incluidos, n el número de observaciones y $L(\text{likelihood})$ es el logaritmo de la verosimilitud del modelo máximo.

$$L(\text{likelihood}) = \ln(SSR/n) = \ln(0.294929585/15) = \ln(.01966197) = -3.929069$$

$$AIC = 2(-(-3.929069) + 5) = 17.858138$$

```

. dis .294929585/15
.01966197

. dis log(.01966197)
-3.929069

. dis 2*(-log(.01966197)+5)
17.858138

```

Figura 7.1

La figura 7.1 muestra el coeficiente de AIC igual a 17.8581, por ende se concluye que nuestro modelo es relativamente bueno.

Pruebas para variables omitidas (Prueba de Ramsey)

H0: el modelo está bien especificado, es decir, que los coeficientes asociados a los términos adicionados son simultáneamente iguales a 0.

Ha: es que el modelo está mal especificado, es decir, que al menos uno de los coeficientes asociados a las variables adicionadas es diferente de 0.

p-value > α

```
. ovtest
```

```
Ramsey RESET test using powers of the fitted values of lpe
Ho: model has no omitted variables
      F(3, 6) =      1.17
      Prob > F =      0.3950
```

Figura 7.2

Los resultados de la figura 7.2 indican un valor p de 0.3950, por lo que no se rechaza la hipótesis nula, al nivel de significancia del 5%. Por lo que el modelo está bien especificado.

Estimación por mínimos cuadrados ordinarios (MCO)

Incorporamos un modelo lineal en su forma funcional log-log. Usamos estimación por mínimos cuadrados ordinarios (MCO). El modelo ajustado es:

$$\begin{aligned} \log(pe) = & -1459.601 + 6.616\log(ump) + 90.457\log(pfalf_{1524}) - 5.501\log(minpo_{1529}) \\ & - 15.498\log(pea_{15}) + 41.9383\log(poet) \end{aligned}$$

El modelo estimado y la explicación detallada de la información retornada por el Stata se muestra en la Figura 8.1

```
. reg lpe lump lptalf_1524 lminpo_15_29 lpea_15 lpoet
```

Source	SS	df	MS	Number of obs = 15		
Model	5.41235894	5	1.08247179	F(5, 9) = 33.03		
Residual	.294929585	9	.032769954	Prob > F = 0.0000		
Total	5.70728852	14	.407663466	R-squared = 0.9483		
				Adj R-squared = 0.9196		
				Root MSE = .18102		

lpe	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lump	6.615647	2.752422	2.40	0.040	.3892357	12.84206
lpfalf_1524	90.45656	11.62881	7.78	0.000	64.15037	116.7627
lminpo_15_29	-5.500498	1.990338	-2.76	0.022	-10.00296	-.998041
lpea_15	-15.49826	4.261584	-3.64	0.005	-25.13863	-5.857883
lpoet	41.9383	5.506481	7.62	0.000	29.48178	54.39483
_cons	-1459.601	191.1468	-7.64	0.000	-1892.006	-1027.197

Figura 8.1 Estimación MCO en Stata.

Explicación de la salida de un modelo de regresión lineal múltiple estimado en Stata

La regresión estimada en la cual figura la variable dependiente primas emitidas seguida por las variables independientes: unidades médicas privadas, personas alfabetas femeninas 15-24, mediana del ingreso nominal de la población de 15 a 29 años ocupada que reportan ingreso, población económicamente activa y población en edad de trabajar.

Model-df=k= 5: representa los grados de libertad asociados a la suma explicada de cuadrados.

Residual-df= 9: representa los grados de libertad asociados a la suma residual de cuadrados.

Total-df= 14: representa los grados de libertad asociados a la suma total de cuadrados o STC.

n=15: representa el número de observaciones.

$$SSR = 0.294929585$$

$$SSM = 5.41235894$$

$$MSR = \frac{SSR}{n-p} = \frac{0.294929585}{15-6} = 0.03276995$$

$$\begin{aligned} & . \text{dis } .294929585 / (15-6) \\ & .03276995 \end{aligned}$$

$$MSM = \frac{SSM}{k} = \frac{5.41235894}{5} = 1.0824718$$

$$\begin{aligned} & . \text{dis } 5.41235894 / 5 \\ & 1.0824718 \end{aligned}$$

$$F0 = \frac{MSM}{MSR} = \frac{1.0824718}{0.03276995} = 33.032448$$

$$\begin{aligned} & . \text{dis } 1.0824718 / .032769954 \\ & 33.032448 \end{aligned}$$

$$STC = SSM + SSR = 5.41235894 + 0.294929585 = 5.70728852$$

Prob >f= 0.0000: es el valor p asociado al estadístico F.

R-squared: representa el coeficiente de determinación o el R², definido como:

$$R^2 = \frac{SSM}{STC} = \frac{5.41235894}{5.70728852} = 0.9483$$

$$\begin{aligned} & . \text{dis } 5.41235894 / 5.70728852 \\ & .94832405 \end{aligned}$$

Es decir, el 94.83 % de la variabilidad de las primas emitidas es explicada por el modelo propuesto.

El **R2 ajustado** está dado por

$$R_{adj}^2 = 1 - \frac{SCE/(n-p)}{SCT/(n-1)} = 1 - \frac{0.294929585/(9)}{5.70728852/(14)} = 0.9196$$

$$\begin{aligned} & . \text{dis } (0.294929585/9)/(5.70728852/14) \\ & .08038482 \end{aligned}$$

$$\begin{aligned} & . \text{dis } 1-0.08038482 \\ & .91961518 \end{aligned}$$

Root-RSE= 0.18102: representa la desviación estándar estimada del error, es decir, la raíz cuadrada de la varianza estimada del error.

$$\hat{Y} = -1459.601 + 6.616\log(ump) + 90.457\log(pfalf_1524) - 5.501\log(minpo_1529) - 15.498\log(pea15) + 41.9383\log(poet)$$

```
. dis 90.45656 /100
.9045656

. dis -5.500498 /100
-.05500498

. dis -15.49826/100
-.1549826

. dis 41.9383/100
.419383

. dis 6.615647 /100
.06615647
```

Nuestro modelo explica que las primas emitidas aumentan 0.06615647% cuando hay un incremento de 1% en las unidades médicas privadas. Análogamente, cuando la población alfabetizada femenina de 15-24 años crece 1%, el porcentaje de primas emitidas del seguro de gastos médicos mayores aumenta 0.9045656%. Por otro lado, disminuyen 0.05500498% cuando la mediana del ingreso nominal de la población de 15 a 29 años ocupada que reportan ingreso aumenta 1%, así mismo cuando la población económicamente activa crece 1%, las primas emitidas disminuyen 0.1549826%. Finalmente, si la población en edad de trabajar aumenta 1%, existe un cambio de 0.419383 en la variación porcentual de primas emitidas del Seguro de GMM.

Pruebas de Hipótesis

Prueba de significancia individual

Con relación a la inferencia estadística alrededor de los coeficientes de regresión poblacionales, el Stata incorpora en la salida mostrada en la figura 8.1, lo que se conoce como pruebas de significancia individual, cuya estructura es:

$H_0: \beta_j = 0$ vs $H_a: \beta_j \neq 0$ con $j = 1, 2, 3, 4, 5$

```
. * lpoet Pob. en edad de trabajar

. display "t-value      " = ( 41.9383 / 5.506481 )
t-value      7.6161708

. *t tabla

. display "t(15, 0.975)=  " invttail(15,0.025)
t(15, 0.975)=  2.1314495
```

```

. *Significancia individual de la regresión
. * lump: Unidades médicas privadas
. display "t-value" = ( 6.615647/ 2.752422)
t-value2.4035729

. *Significancia individual de la regresión
. * lump: Unidades médicas privadas
. display "t-value      " = ( 6.615647/ 2.752422)
t-value      2.4035729

. * pfalf_1524 Pob.femenina alfabeto
. display "t-value      " = ( 90.45656/11.62881)
t-value      7.7786601

. * lminpo_15_29 Mediana de ingreso
. display "t-value      " = ( -5.500498 / 1.990338 )
t-value      -2.7636

. * lpea_15 Pob. Económicamente activa
. display "t-value      " = ( -15.49826/ 4.261584 )
t-value      -3.636737

```

Para el estadístico de prueba de todas nuestras variables explicativas $|t^*| > t(160, 0.975)$, entonces la hipótesis nula de no significancia individual se rechaza, es decir, $\beta_1, \beta_2, \beta_3, \beta_4$ y β_5 poblacional son significativas, lo que implica que las variables asociadas al mismo contribuyen a la explicación de la variable dependiente (primas emitidas).

Prueba de significancia global

```

. display ( 5.41235894/5) /( .294929585 /9)
33.032448

. display Ftail(5,9,33.032448)
.00001572

```

Dado que el valor del estadístico de prueba $|F^*| = 33.032448$ es mayor al valor p asociado 0.00001572, esto implica que se rechaza la hipótesis nula de no significancia global, es decir, al menos uno de los coeficientes de regresión que acompañan a las variables explicativas es diferente de 0, por lo que nuestro modelo es significativo globalmente.

Conclusión

Es importante mencionar que en Tabasco el 68.5% de la población se encuentra afiliada a servicios de salud (INEGI). Por esta razón, las aseguradoras aún tienen un campo de crecimiento bastante amplio y se encuentran en la obligación de implementar diferentes métodos de venta que las lleven a conseguir más clientes y aumentar el porcentaje de la población asegurada, lo que se traduce a una mayor emisión de primas y mayor utilidad para las aseguradoras.

Anteriormente el Covid-19 como nueva enfermedad no era un riesgo asegurable, sin embargo, como se trata de una enfermedad respiratoria muchas compañías de seguros ya lo toman en cuenta.

Bajo la intención de conocer la dependencia de la demanda de los seguros de GMM, se realizó este proyecto. En un principio se pensaba que los siniestros ocurridos, los siniestros pagados, el Ingreso laboral per cápita a pesos corrientes (nominal) anual, las defunciones, la esperanza de vida, y el Índice de la Tendencia Laboral de la Pobreza (ITLP) eran factores de suma importancia, pues se creía que había una relación entre las defunciones y las primas emitidas ya que es consecuencia de cualquier enfermedad mortal y eso hace que las personas se vean en la necesidad de comprar un seguro de GMM. Además, sí el Ingreso laboral per cápita a pesos corrientes aumentaba las personas tendrían mayor posibilidad de adquirirlo, entre otras.

Después de modelar con distintos grupos de variables y establecer nuestro modelo más adecuado, nos dimos cuenta de que cuando hay un incremento de 1% en las unidades médicas privadas, las primas emitidas aumentan 0.06615647%. Esto se debe a que, al haber más espacio y personal médico, los asegurados obtienen una mejor calidad de servicio en la salud. Después sí la población femenina alfabetizada de 15-24 años crecía 1%, el porcentaje de primas emitidas del seguro de gastos médicos mayores aumentaba 0.9045656%. Por esto la educación es un factor muy importante, pues así los jóvenes tienen oportunidad de tener una mejor calidad de vida. Según la Condusef en promedio una póliza de seguro de gastos médicos mayores tiene un valor de \$11,696 para hombres, y \$15,261 para Mujeres de entre 20 y 30 años con una suma asegurada con cobertura básica de 40 millones de pesos.

La estabilidad económica durante los últimos 15 años en Tabasco no ha contribuido al crecimiento de la emisión de primas en este ramo de seguro, pues estas disminuyen 0.05500498% cuando la mediana del ingreso nominal de la población de 15 a 29 años ocupada que reportan ingreso aumenta 1% y cuando la población económicamente activa crece 1%, las primas emitidas disminuyen 0.1549826%. No obstante, dicho crecimiento económico ha llevado a una reducción en la pobreza y a mayor cobertura de educación, lo que lleva a que las personas entiendan el funcionamiento de un seguro, logrando así mitigar riesgos futuros y que esto los motive a comprar este tipo de servicios.

Finalmente, muchas empresas ofrecen a sus colaboradores el seguro de GMM, como son Western Unión, FedEx Express, Steren y Mars México, esto explica que exista un cambio de 0.419383% en la variación porcentual de las primas emitidas sí la población en edad de trabajar de Tabasco aumenta 1%.

Atravesamos por una situación nunca vista, la pandemia se considera una de las mayores catástrofes de estos tiempos. El sector asegurador tiene grandes retos, una de ellas promover la cultura del seguro en el país.

Los resultados como se explican en el análisis son útiles para poder cuantificar el impacto que tiene el cambio en políticas monetarias sobre la demanda de seguros, en este caso expresado por ciertas variables macroeconómicas que influyen finalmente en la decisión de los consumidores, sobre si compran o no compran un seguro de gastos médicos mayores. Se demuestra que variables como unidades médicas privadas, personas alfabetizadas femeninas 15 a 24, mediana del ingreso nominal de la población de 15 a 29 años ocupada que reportan ingreso, población económicamente activa y población en edad de trabajar influyen en la demanda de seguros de GMM.

Base de datos

CNSF. Entidades Supervisadas. Seguro de Gastos Médicos Mayores. 2005-2019. Recuperado de: <https://www.cnsf.gob.mx/EntidadesSupervisadas/InstitucionesSociedadesMutualistas/Paginas/InformacionConsolidada.aspx>

Esperanza de Vida en Tabasco. Recuperado de: <https://datosmacro.expansion.com/demografia/esperanza-vida/mexico>

INEGI. https://www.inegi.org.mx/sistemas/olap/consulta/general_ver4/MDXQueryDatos.asp?#Regreso&c=11144

INEGI. Mortalidad. Defunciones Generales 2005-2019. Consultado en:

https://www.inegi.org.mx/app/tabulados/interactivos/?px=Mortalidad_09&bd=Mortalidad

Índice de la Tendencia Laboral de la Pobreza (ITLP) en Tabasco. 2005-2019

Ingreso laboral per cápita a pesos corrientes (nominal) anual. Tabasco. Base de datos 2005-2019.

<https://www.coneval.org.mx/coordinacion/entidades/Tabasco/Paginas/itlp.aspx>

Producto Interno Bruto. Recuperado de: [PIB TABASCO](#)

INEGI. Población Ocupada. <https://www.inegi.org.mx/programas/pibent/2013/#Tabulados>

https://www.inegi.org.mx/sistemas/olap/consulta/general_ver4/MDXQueryDatos_colores.asp?c=