

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICAS
DEPARTAMENTO DE MATEMÁTICA
INGENIERÍA EN AUDITORIA Y CONTADURÍA PÚBLICA AUTORIZADA

AÑO LECTIVO: 2019-2020	PERIODO: Segundo Terminó
MATERIA: Técnicas de Muestreo Y Análisis Multivariado	PROFESORES:
EVALUACIÓN: Tercera	Msc. Sandra Gonzalez C.
TIEMPO DE DURACIÓN: 2 horas	FECHA: 13/02/2020

COMPROMISO DE HONOR

Yo, al firmar este compromiso, reconozco que el presente examen está diseñado para ser resuelto de manera individual, que puedo usar una calculadora ordinaria para cálculos aritméticos, un lápiz o esferográfico; que solo puedo comunicarme con la persona responsable de la recepción del examen; y, cualquier instrumento de comunicación que hubiere traído, debo apagarlo y depositarlo en la parte anterior del aula, junto con algún otro material que se encuentre acompañándolo. No debo, además, consultar libros, notas, ni apuntes adicionales a las que se entreguen en esta evaluación. Los temas debo desarrollarlos de manera ordenada.

Firmo al pie del presente compromiso, como constancia de haber leído y aceptar la declaración anterior.

"Como estudiante de ESPOL me comprometo a combatir la mediocridad y actuar con honestidad, por eso no copio ni dejo copiar".

FIRMA: _____

NÚMERO DE MATRÍCULA: _____

PARALELO: 1

Nota 1: Es válido utilizar aproximaciones en caso de ser necesario; use cuatro decimales de aproximación.

Nota 2: Trabaje con el 95% de confianza para sus cálculos

Tema 1 (20 puntos)

La playa de “CostaNova”, cuenta con un total de 10000 (diez mil) apartamentos, se desea conocer la proporción de apartamentos que los dueños están dispuestos a alquilarlos al menos un mes al año. En la playa “CostaNova” hay tres urbanizaciones. Se selecciono en cada urbanización una muestra aleatoria de apartamentos en número proporcional al total de apartamentos de la urbanización. La información recogida se muestra en la siguiente tabla:

Urbanización	N	n	Alquilaron	Costo
Solares	2877	1050	800	10
Vistabella	2466	900	600	12
AguaClara	4657	1700	1300	15

- Estime la proporción de apartamentos que estarían dispuestos a ser alquilados al menos una vez al año, el límite de error de estimación y el intervalo de confianza de la estimación.
- El organizador de la inmobiliaria desea realizar un nuevo estudio en este 2020, e incluir nuevas variables importantes para el tema turístico. Se le solicita utilizando afijación que minimiza el costo encuentre el tamaño de la muestra con un límite de error de estimación de 0.05. Utilice las proporciones calculadas en el literal a) para el cálculo de las varianzas poblacionales requeridas.

Tema 2 (30 puntos)

Mediante dos modelos lineales simples se estudia el consumo para dos modelos diferentes de automóviles, según su potencia y en función de los siguientes datos:

	Marca A			Marca B		
Consumo (1/100 Km)	7.5	6.2	9.6	8.1	7.6	11.4
Potencia (CV)	9.7	7	12	9.8	7.5	13

- Para cada marca determine los estimadores de Mínimos cuadrados para un modelo de Regresión Lineal Simple y escriba la ecuación que mejor modele los datos
- Realice la tabla ANOVA para cada modelo y la prueba global para cada modelo. Comente los resultados.
- Determine el R^2 para cada modelo y comente los resultados.
- Determine el valor estimado de consumo para una potencia de 10 CV y un intervalo de confianza para la estimación para cada marca.
- ¿Se puede decir que el consumo de la marca A es menor al de la marca B, para un mismo valor de la potencia CV? Justifique su respuesta.

Tema 3 (20 puntos)

Determine el Sesgo de la varianza muestral $s_n^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$, de un M.A.S, como un estimador de la varianza poblacional.

Tema 4 (30 puntos)

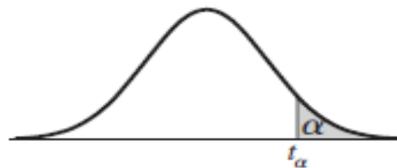
Para estimar el modelo $y_t = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon_t$ se ha obtenido una muestra de $n=14$, de la cual ha resultado:

$$(X^t X)^{-1} = \begin{bmatrix} 1.3214 & -0.5 & -1 \\ -0.5 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \quad X^t Y = \begin{pmatrix} 10 \\ 6 \\ 12 \end{pmatrix}, \quad SCE = 4.9998 \quad SCT = 6.8569$$

- Determine un modelo de regresión que mejor represente los datos
- Realice la prueba global del modelo, utilizando un 95% de confianza.
- ¿Cuáles β_i son significativos en el modelo? Justifique su respuesta
- Realice un intervalo de confianza para cada β_i significativo del modelo planteado
- Determine el coeficiente de determinación R^2 del modelo e interprete su resultado
- Utilizando el modelo del literal a). Estime el valor y_0 sabiendo que $x_1 = 5$ y $x_2 = 7$
- Calcular un intervalo de confianza para el estimador obtenido en el literal f)

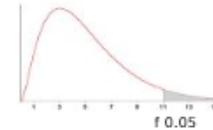
MUESTREO ALEATORIO ESTRATIFICADO	
Estimador de la proporción poblacional \hat{p}	
$\hat{p}_{st} = \frac{1}{N} \sum_{i=1}^L N_i \hat{p}_i \qquad \hat{V}_{(\hat{p}_{st})} = \frac{1}{N^2} \sum_{i=1}^L N_i^2 \left(\frac{\hat{p}_i \hat{q}_i}{n_i - 1} \right) \left(\frac{N_i - n_i}{N_i} \right)$	
Para estimar P: Afijación Proporcional	Para estimar P: Afijación de Neymann:
$n = \frac{\sum_{i=1}^L N_i p_i q_i}{ND + \frac{1}{N} \sum_{i=1}^L N_i p_i q_i} \qquad n_i = n \left(\frac{N_i}{N} \right) \qquad D = \frac{B^2}{4}$	$n_i = n \left(\frac{N_i \sqrt{p_i q_i}}{\sum_{i=1}^L N_i \sqrt{p_i q_i}} \right) \qquad n = \frac{(\sum_{i=1}^L N_i \sqrt{p_i q_i})^2}{N^2 D + \sum_{i=1}^L N_i p_i q_i} \qquad D = \frac{B^2}{4}$
Para estimar P: Afijación que minimiza el costo	
$n_i = n \left(\frac{N_i \sqrt{\frac{p_i q_i}{c_i}}}{\sum_{i=1}^L N_i \sqrt{\frac{p_i q_i}{c_i}}} \right) \qquad n = \frac{(\sum_{i=1}^L \frac{N_i \sigma_i}{\sqrt{c_i}}) (\sum_{i=1}^L N_i \sigma_i \sqrt{c_i})}{N^2 D + \sum_{i=1}^L N_i \sigma_i^2} \qquad D = \frac{B^2}{4} \qquad \sigma_i^2 = p_i q_i$	

Tabla 5 Puntos porcentuales de las distribuciones t



$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	gl
3.078	6.314	12.706	31.821	63.657	1
1.886	2.920	4.303	6.965	9.925	2
1.638	2.353	3.182	4.541	5.841	3
1.533	2.132	2.776	3.747	4.604	4
1.476	2.015	2.571	3.365	4.032	5
1.440	1.943	2.447	3.143	3.707	6
1.415	1.895	2.365	2.998	3.499	7
1.397	1.860	2.306	2.896	3.355	8
1.383	1.833	2.262	2.821	3.250	9
1.372	1.812	2.228	2.764	3.169	10
1.363	1.796	2.201	2.718	3.106	11
1.356	1.782	2.179	2.681	3.055	12
1.350	1.771	2.160	2.650	3.012	13
1.345	1.761	2.145	2.624	2.977	14
1.341	1.753	2.131	2.602	2.947	15

Tabla D.9: VALORES CRITICOS DE LA DISTRIBUCION F (0,05)



g.d.l.	Grados de libertad del Numerador														g.d.l.	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14		
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,0	243,9	244,7	245,4	245,9	1
2	18,513	19,000	19,164	19,247	19,296	19,330	19,353	19,371	19,385	19,396	19,405	19,413	19,419	19,424	19,429	2
3	10,128	9,552	9,277	9,117	9,013	8,941	8,887	8,845	8,812	8,786	8,763	8,745	8,729	8,715	8,703	3
4	7,709	6,944	6,591	6,388	6,256	6,163	6,094	6,041	5,999	5,964	5,936	5,912	5,891	5,873	5,858	4
5	6,608	5,786	5,409	5,192	5,050	4,950	4,876	4,818	4,772	4,735	4,704	4,678	4,655	4,636	4,619	5
6	5,987	5,143	4,757	4,534	4,387	4,284	4,207	4,147	4,099	4,060	4,027	4,000	3,976	3,956	3,938	6
7	5,591	4,737	4,347	4,120	3,972	3,866	3,787	3,726	3,677	3,637	3,603	3,575	3,550	3,529	3,511	7
8	5,318	4,459	4,066	3,838	3,687	3,581	3,500	3,438	3,388	3,347	3,313	3,284	3,259	3,237	3,218	8
9	5,117	4,256	3,863	3,633	3,482	3,374	3,293	3,230	3,179	3,137	3,102	3,073	3,048	3,025	3,006	9
10	4,965	4,103	3,708	3,478	3,326	3,217	3,135	3,072	3,020	2,978	2,943	2,913	2,887	2,865	2,845	10
11	4,844	3,982	3,587	3,357	3,204	3,095	3,012	2,948	2,896	2,854	2,818	2,788	2,761	2,739	2,719	11
12	4,747	3,885	3,490	3,259	3,106	2,996	2,913	2,849	2,796	2,753	2,717	2,687	2,660	2,637	2,617	12
13	4,667	3,806	3,411	3,179	3,025	2,915	2,832	2,767	2,714	2,671	2,635	2,604	2,577	2,554	2,533	13
14	4,600	3,739	3,344	3,112	2,958	2,848	2,764	2,699	2,646	2,602	2,565	2,534	2,507	2,484	2,463	14
15	4,543	3,682	3,287	3,056	2,901	2,790	2,707	2,641	2,588	2,544	2,507	2,475	2,448	2,424	2,403	15
16	4,494	3,634	3,239	3,007	2,852	2,741	2,657	2,591	2,538	2,494	2,456	2,425	2,397	2,373	2,352	16
17	4,451	3,592	3,197	2,965	2,810	2,699	2,614	2,548	2,494	2,450	2,413	2,381	2,353	2,329	2,308	17
18	4,414	3,555	3,160	2,928	2,773	2,661	2,577	2,510	2,456	2,412	2,374	2,342	2,314	2,290	2,269	18
19	4,381	3,522	3,127	2,895	2,740	2,628	2,544	2,477	2,423	2,378	2,340	2,308	2,280	2,256	2,234	19
20	4,351	3,493	3,098	2,866	2,711	2,599	2,514	2,447	2,393	2,348	2,310	2,278	2,250	2,225	2,203	20
21	4,325	3,467	3,072	2,840	2,685	2,573	2,488	2,420	2,366	2,321	2,283	2,250	2,222	2,197	2,176	21
22	4,301	3,443	3,048	2,817	2,661	2,549	2,464	2,397	2,342	2,297	2,259	2,226	2,198	2,173	2,151	22
23	4,279	3,422	3,027	2,796	2,640	2,528	2,442	2,375	2,320	2,275	2,236	2,204	2,175	2,150	2,128	23
24	4,260	3,403	3,008	2,776	2,619	2,507	2,421	2,354	2,300	2,255	2,216	2,183	2,155	2,130	2,108	24
25	4,242	3,385	2,991	2,759	2,602	2,490	2,404	2,337	2,282	2,236	2,198	2,165	2,136	2,111	2,089	25

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$$

$$S_{xx} = \frac{\sum(x_i - \bar{x})^2}{n} = \frac{\sum x_i^2}{n} - \bar{x}^2$$

$$S_{yy} = \frac{\sum(y_i - \bar{y})^2}{n} = \frac{\sum y_i^2}{n} - \bar{y}^2$$

$$S_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{n} = \frac{\sum x_i y_i}{n} - \bar{x} \bar{y}$$

$$R^2 = \frac{SCR}{SCT}$$

$$SCR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

G de Libertad: p-1

$$SCE = \sum_{i=1}^n (y_i - \hat{y})^2$$

Grados de Libertad: n-p

Estadístico prueba global del modelo:

$$F_o = \frac{MCR}{MCE} > F_{(\alpha, p-1, n-p)}$$

$$SCT = \sum_{i=1}^n (y_i - \bar{y})^2$$

Grados de Libertad: n-1

Estimación modelo de regresión lineal simple

$$S_e = \sqrt{MCE}$$

El intervalo de (1-α)100% de confianza para la estimación individual es:

$$\hat{y}_o \pm t_{\frac{\alpha}{2}, n-2} S_e \sqrt{\left[1 + \frac{1}{n} + \frac{(x_o - \bar{x})^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right]}$$