

NO SE RESPONDERÁN PREGUNTAS DURANTE EL EXAMEN.

- Un estudiante de Ingeniería que hace su práctica en una Agencia para Protección Ambiental, ha sido encargado de resolver el problema de establecer normas para la cantidad permisible de ciertos productos químicos tóxicos en lagos y ríos de agua dulce. Una medida común de la toxicidad para cualquier contaminante es la concentración del contaminante que matará la mitad de los especímenes de prueba en un intervalo de tiempo dado (normalmente 96 horas para los peces). Esta medida se denomina CL50 (concentración letal que mata el 50% de los especímenes de prueba). En muchos estudios las mediciones de  $\ln(\text{CL50})$  tienen una distribución normal y por lo tanto los análisis se basan en los datos de  $\ln(\text{CL50})$ . Los estudios de los efectos del cobre en cierta especie de peces (digamos especie A) revelan que la varianza de las mediciones de  $\ln(\text{CL50})$  tiene un valor alrededor de 0.4, con las mediciones de las concentraciones en miligramos por litro.
  - Si hay que efectuar  $n = 10$  estudios de CL50 para cobre, encuentre la probabilidad de que la media de la muestra de  $\ln(\text{CL50})$  difiera de la verdadera media de la población en no más de 0.5.
  - Si se requiere que la media muestral difiera a lo más en 0.5 de la media de la población, con una probabilidad de 0.95 ¿cuántas pruebas deben efectuarse? (Nota: en el muestreo siempre se aproxima por exceso)
- Se prueba una operación de línea de producción para ver la exactitud del peso de llenado (en onzas por envase) con las siguientes hipótesis:
 

<u>Hipótesis</u>	<u>Conclusión y acción</u>
$H_0: \mu = 16$	Llenado correcto; seguir trabajando.
$H_1: \mu \neq 16$	Llenado fuera de la norma; parar y ajustar la máquina.

 El tamaño de la muestra es 30 y la desviación estándar de la población es  $\sigma = 0.8$ . Utilizando un  $\alpha = 0.05$ 
  - ¿Cuál es la probabilidad de cometer un error de tipo II cuando la máquina llena los envases con 0.5 onzas por debajo del valor de llenado correcto?
  - ¿Cuál es la potencia de la prueba estadística cuando la máquina llena los envases con un exceso de 0.5 onzas?
- Se recolectan datos para determinar si hay diferencia entre las varianzas de los ingresos diarios (\$000) en dos tiendas. Dados los siguientes datos y con una significancia del 5% ¿considera usted que existe suficiente evidencia para afirmar que las varianzas de ambas tiendas son diferentes? Justifique.

Tienda 1	Tienda 2
45,78	67,89
34,66	76,45
65,89	87,12
54,78	98,65
98,66	65,87
12,55	34,91
37,77	56,88
21,87	45,99
23,45	
56,98	

- En febrero de 1995, la media de los costos de un viaje redondo en avión era de US \$258 (USA TODAY, 1995). En una muestra aleatoria de 15 boletos de viaje redondo en el mes de marzo se obtuvieron los siguientes datos

310	260	265	255	300	310	230	250
265	280	290	240	285	250	260	

Suponiendo que los costos se distribuyen normalmente:

- Construir un intervalo de confianza del 95% para estimar la verdadera media poblacional e identificar y escribir de cuánto es el margen de error o error de estimación alcanzado por dicha muestra.
- ¿Permiten los datos verificar que el costo medio de un viaje redondo es diferente de 258? Justifique su respuesta utilizando  $\alpha=0.05$

5. TRANSWORLD AIRWAYS desea determinar si existe una relación sistemática entre el número de vuelos que las personas toman y su ingreso. Con base en los siguientes datos correspondientes a 100 viajeros, ¿cuál es su conclusión respecto a la relación de las dos variables con un nivel de significancia del 5%?

Ingreso	Frecuencia de vuelos			Total
	2 o menos veces al año	Entre 2 y 12 veces al año	Más de 12 veces al año	
Menos de US\$30,000	20	15	2	37
US\$30,000-50,000	8	5	1	14
US\$50,000-70,000	7	8	12	27
Más de US\$70,000	2	5	15	22
<b>Total</b>	37	33	30	100

6. Un administrador que trabaja en un departamento de investigación está preparando un estudio sobre el comportamiento del consumidor. Con los datos recolectados por él, quiso determinar si existe una relación causa-efecto entre el ingreso del consumidor y los niveles de consumo. Para ello construyó un modelo de regresión lineal simple mediante el método de mínimos cuadrados, tomando como variable independiente el nivel de Ingresos X (\$000) y como variable de respuesta el nivel de consumo Y (\$000). Con una confiabilidad del 95% entre muchos de los reportes que obtuvo, se destaca el siguiente usando un software estadístico.

Predicted Values

X	Predicted Y	95,00% Prediction Limits		95,00% Confidence Limits	
		Lower	Upper	Lower	Upper
20,0	12,9413	7,71954	18,1631	11,4853	14,3973
25,0	15,7322	10,418	21,0464	13,9733	17,491
30,0	18,523	12,9881	24,058	16,1802	20,8658

Con base en la anterior tabla usted debe completar las siguientes frases.

- Con un 95% de confiabilidad, un consumidor con ingreso de 20 tendrá un consumo entre \$\_\_\_\_\_ y \$\_\_\_\_\_.
  - Con 95% de confiabilidad, un grupo de consumidores, todos con ingreso de 30, tendrán un consumo promedio entre \$\_\_\_\_\_ y \$\_\_\_\_\_.
7. Cada año, los estudiantes de noveno grado en el suroeste del estado de Ohio deben presentar un examen de aptitud. El archivo anexo contiene datos de 47 distritos escolares en ese estado para 1994-1995. Las variables en el conjunto de datos son: “% aprobados” es el % de estudiantes que pasa la prueba de aptitud; “% asistencia” es el promedio de los porcentajes diarios de estudiantes que asisten a clase. Suponga que se construye un modelo de regresión lineal por mínimos cuadrados donde “% aprobados” es la variable dependiente y “% asistencia” es la variable independiente.
- Complete los espacios en blanco de las tablas anexas donde se necesite.
  - Plantee y resuelva la hipótesis correspondiente a la tabla de análisis de varianza anexa.
  - Escriba el modelo de regresión lineal ajustado a esta serie de datos e interprete la pendiente del modelo.
  - Con base en este modelo, calcule el porcentaje medio de aprobados para un porcentaje promedio diario del 95% de estudiantes que asisten a clase en una de las instituciones del distrito escolar.
  - Plantee, resuelva e interprete el resultado de la prueba de hipótesis inherente a  $\beta_1$  con un  $\alpha=0,05$ .
  - Calcule e interprete el coeficiente de determinación del modelo.
  - Plantee y resuelva las hipótesis correspondientes a la verificación de cada uno de los supuestos del modelo con un  $\alpha=0,05$ . ¿Qué puede decir acerca de la homoscedasticidad de los residuales?
  - Plantee y resuelva la hipótesis usual correspondiente al coeficiente de correlación lineal entre las dos variables con un  $\alpha=0,05$

## TABLA DE DATOS PARA LA REGRESIÓN

<b>Distrito de Aprobados</b>	<b>% Aprobados</b>	<b>% Asistencia</b>
Edgewood	70	93,84
Fairfield	67	94,76
Hamilton	49	92,22
Lakota	82	96,07
Madison	60	93,93
Middletown	42	92,21
New Miami	22	93,22
Ross	55	94,83
Talawanda	66	94,54
Batavia	67	94,03
Bethel-Tate	59	93,96
Clermont	38	91,93
Felicity-Franklin	49	94,2
Goshen	43	93,05
Milford	68	94,42
New Richmond	48	93,06
West Clermont	53	92,82
Williamsburg	71	94,16
Cincinnati	26	88,67
Deer Park	65	94,4
Finneytown	78	94,36
Forest Hill	78	95,93
Indian Hill	92	96,13
Lockland	60	92,08
Loveland	38	94,93
Madeira	88	95,66
Mariemont	83	96,33
Mount Healthy	40	92,87
North College Hill	44	94,04
Northwest	59	93,95
Norwood	58	92,63
Oak Hills	72	95,77
Reading	69	94,4
Southwest	60	94,5
St.Bernard-Elmwood Place	58	93,71
Sycamore	84	95,76
Three Rivers	45	93,67
Winton Woods	46	94,25
Wyoming	93	96,84
Carlisle	58	94,97
Franklin	59	93,39
Kings	66	94,71
Lebanon	69	94,66
Little Miami	62	94,72
Mason	84	95,56
Springboro	74	95,76
Wayne	79	95,68

## FÓRMULAS

Intervalo de confianza para el promedio muestral con varianza conocida:

$$\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Prueba F para la comparación de varianzas de dos muestras independientes.

$$F = \frac{S_1^2}{S_2^2}, S_i^2 \text{ Varianza muestral del grupo } i\text{-ésimo.}$$

Intervalo de confianza para el promedio muestral con varianza desconocida

$$\bar{X} \pm t_{n-1; \alpha/2} \frac{s}{\sqrt{n}}, \text{ donde } s \text{ es la desviación estándar muestral.}$$

Estadístico de prueba para Independencia de dos variables

$$\chi^2 = \sum_{\text{todas las celdas}} \frac{(f_o - f_e)^2}{f_e}, \text{ donde } f_e = \frac{\text{total fila } X \text{ total columna}}{n}$$

Con  $(r-1) \times (c-1)$  grados de libertad,  $r$  y  $c$  número de filas y columnas respectivamente.

Estadístico de prueba para el coeficiente de correlación.

$$t = \frac{r - \rho}{\sqrt{\frac{1-r^2}{n-2}}}, \text{ donde el estadístico } t \text{ tiene una distribución t-student con } n-2 \text{ grados}$$

de libertad

Valores porcentuales de los puntos del examen.

1<sup>er</sup> punto: 15%

2<sup>o</sup> punto: 15%

3<sup>er</sup> punto: 5%

4<sup>o</sup> punto: 5%

5<sup>o</sup> punto: 15%

6<sup>o</sup> punto: 10%

7<sup>o</sup> punto: 35%

Regression Analysis

<i>Estadísticas de la regresión</i>	
Coeficiente de correlación múltiple	0,776171075
Coeficiente de determinación R <sup>2</sup>	0,602441537
R <sup>2</sup> ajustado	0,593606905
Error típico	10,57871589
Observaciones	47

ANÁLISIS DE VARIANZA

	<i>Grados de libertad</i>	<i>Suma de cuadrados</i>	<i>Promedio de los cuadrados</i>	<i>F</i>
Regresión	1		7631,191042	
Residuos		5035,915341		
Total	46	12667,10638		

	<i>Coefficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>	<i>Inferior 95%</i>	<i>Superior 95%</i>
Intercepción	-771,5868596	100,9110801	-7,646205536	974,8322051	568,3415141
% Asistencia		1,071074095	8,257778122	6,687438293	11,00194616

**Durbin-Watson Calculations**

Sum of Squared Difference of Residuals	10047,60837
Sum of Squared Residuals	5035,915341

<b>Durbin-Watson Statistic</b>	<b>1,995190088</b>
--------------------------------	--------------------

PRUEBA DE LOS RESIDUALES

$e_i$	$e_i$ orden	$e_i$	FREC	FREC ACUM	$F_s(x)$	$F_t(x)$	$F_s - F_t$	$ F_s - F_t $
11,60094	-30,91535	-30,9153496	1	1	0,0212766	0,00156485	0,01971175	0,019711746
0,463824	-30,039773	-30,0397733	1	2	0,04255319	0,00204574	0,04050745	0,040507449
4,929343	-16,167997	-16,1679972	1	3	0,06382979	0,06114412	0,00268567	0,002685672
3,877278	-16,025383	-16,0253826	1	4	0,08510638	0,06280939	0,022297	0,022296996
0,804919	-12,583148	-12,583148	1	5	0,10638298	0,11456139	-0,00817841	0,008178408
-1,98221	-12,155304	-12,1553041	1	6	0,12765957	0,1226724	0,00498718	0,004987175
-30,91535	-11,895461	-11,8954611	1	7	0,14893617	0,12779075	0,02114542	0,021145424
-12,1553	-10,393561	-10,393561	1	8	0,17021277	0,16026872	0,00994405	0,009944047
1,409657	-9,8197073	-9,81970733	1	9	0,19148936	0,17399155	0,01749781	0,017497808
6,92045	-8,4117519	-8,41175193	1	10	0,21276596	0,21071446	0,00205149	0,002051493
-0,460422	-4,2365557	-4,23655565	1	11	0,23404255	0,34277385	-0,1087313	0,108731296
-3,505697	-4,1823879	-4,18238794	1	12	0,25531915	0,34467862	-0,08935947	0,089359471
-12,58315	-3,5056966	-3,50569664	1	13	0,27659574	0,3687924	-0,09219666	0,092196655
-8,411752	-3,5001989	-3,50019885	1	14	0,29787234	0,3689906	-0,07111826	0,071118257
4,47102	-3,4693148	-3,46931478	1	15	0,31914894	0,37010463	-0,05095569	0,050955693
-3,500199	-1,9822105	-1,98221046	1	16	0,34042553	0,42487095	-0,08444542	0,084445417
3,622527	-1,3808679	-1,38086786	1	17	0,36170213	0,44750201	-0,08579989	0,085799887
9,77064	-0,4604219	-0,46042185	1	18	0,38297872	0,48245047	-0,09947174	0,099471744
13,328	-0,3719749	-0,37197493	1	19	0,40425532	0,48582014	-0,08156482	0,08156482
1,647914	-0,093941	-0,09394102	1	20	0,42553191	0,49641822	-0,0708863	0,070886303
15,0017	0,4638244	0,46382437	1	21	0,44680851	0,51767914	-0,07087063	0,070870628
1,115534	0,7507512	0,7507512	1	22	0,46808511	0,52860048	-0,06051537	0,060515371
13,3466	0,8049189	0,80491891	1	23	0,4893617	0,53066011	-0,04129841	0,041298408
17,1676	1,1155345	1,11553446	1	24	0,5106383	0,54245322	-0,03181492	0,03181492
-30,03977	1,4096567	1,40965666	1	25	0,53191489	0,55358594	-0,02167105	0,021671049
13,5036	1,6479136	1,64791357	1	26	0,55319149	0,56257368	-0,00938219	0,009382189
2,577658	2,5776576	2,57765757	1	27	0,57446809	0,59729706	-0,02282897	0,022828975
-9,819707	3,3482936	3,34829359	1	28	0,59574468	0,62551954	-0,02977486	0,029774857
-16,168	3,6225273	3,62252728	1	29	0,61702128	0,63541107	-0,0183898	0,018389797
-0,371975	3,8772776	3,87727755	1	30	0,63829787	0,64451992	-0,00622205	0,006222049
10,30302	4,3267075	4,32670752	1	31	0,65957447	0,66038742	-0,00081295	0,000812949
-3,469315	4,4710197	4,47101972	1	32	0,68085106	0,66542438	0,01542668	0,01542668
5,647914	4,5810527	4,58105272	1	33	0,70212766	0,66924505	0,03288261	0,032882611
-4,236556	4,9293426	4,92934262	1	34	0,72340426	0,68122138	0,04218287	0,042182873
0,750751	5,6479136	5,64791357	1	35	0,74468085	0,70533032	0,03935053	0,039350535
8,619132	6,9204497	6,92044969	1	36	0,76595745	0,74582689	0,02013056	0,02013056
-11,89546	8,0668645	8,06686454	1	37	0,78723404	0,77964132	0,00759272	0,007592721
-16,02538	8,6191321	8,61913214	1	38	0,80851064	0,7949634	0,01354724	0,013547237
8,066865	9,7706397	9,7706397	1	39	0,82978723	0,82480137	0,00498587	0,004985868
-10,39356	10,303019	10,3030188	1	40	0,85106383	0,83761442	0,01344941	0,013449411
4,581053	10,388071	10,3880706	1	41	0,87234043	0,83960343	0,03273699	0,032736993
-0,093941	11,600941	11,6009412	1	42	0,89361702	0,86623059	0,02738643	0,02738643
3,348294	13,328	13,328	1	43	0,91489362	0,89863464	0,01625898	0,016258978
-4,182388	13,346596	13,346596	1	44	0,93617021	0,89894929	0,03722092	0,037220918
10,38807	13,503601	13,5036014	1	45	0,95744681	0,90157758	0,05586923	0,055869226
-1,380868	15,001701	15,0017013	1	46	0,9787234	0,92418139	0,05454201	0,054542015
4,326708	17,1676	17,1675995	1	47	1	0,94957803	0,05042197	0,050421967

47

Estadístico de Kolmogorov-Smirnov **D = 0,10873130**

