

# ExAMEN FINAL

## Problema 1

Se ha tomado una muestra aleatoria de 40 baterías y se ha registrado su duración en años. Estos resultados se los ha agrupado en 7 clases o intervalos en la siguiente tabla de frecuencia:

X: Duración en años	Frec. Absoluta ( $O_i$ )
1.45 --- 1.95	2
1.95 --- 2.45	1
2.45 --- 2.95	4
2.95 --- 3.45	15
3.45 --- 3.95	10
3.95 --- 4.45	5
4.45 --- 4.95	3
TOTAL: 40	

Verificar con 5% de significancia que la duración en años de las baterías producidas por este fabricante tienen una distribuida normalmente con media 3.5 y desviación estándar 0.7

- a) ( 1pto) **Definir la variable estadística y el tipo a la que pertenece. Plantear las hipótesis:**

X: duración en años de baterías

Tipo: cuantitativa-continua

Hipótesis:

H0: la duración de baterías están distribuidas normalmente

H1: la duración de baterías no están distribuidas normalmente

- b) ( 1pto) **Calcule los valores esperados**

X: Duración en años	Frec. Absoluta ( $O_i$ )	probabilidad	frec. Esperada
1.45 --- 1.95	2	0.012037	0.48148
1.95 --- 2.45	1	0.052904	2.11616
2.45 --- 2.95	4	0.145048	5.80192
2.95 --- 3.45	15	0.248317	9.93268
3.45 --- 3.95	10	0.265575	10.623
3.95 --- 4.45	5	0.193496	7.73984
4.45 --- 4.95	3	0.062893	2.51572
n=	40		39.2108

- c) ( 1pto) **Calcular el valor del estadístico de prueba  $\chi^2$  usando la fórmula:**

c.a)  $(1.45 < x < 1.95)$

$$\frac{1.95 - 3.5}{0.7} - \frac{1.45 - 3.5}{0.7}$$

$$-2.2 - -2.9 = 0.013903 - 0.001866 = 0.012037$$

c.b)  $(1.95 < x < 2.45)$

$$\frac{2.45-3.5}{0.7} - \frac{1.95-3.5}{0.7}$$

$$-1.5 - -2.2 = 0.066807 - 0.013903 = 0.052904$$

c.c)  $(2.45 < x < 2.95)$

$$\frac{2.95-3.5}{0.7} - \frac{2.45-3.5}{0.7}$$

$$-0.8 - -1.5 = 0.211855 - 0.066807 = 0.145048$$

c.d)  $(2.95 < x < 3.45)$

$$\frac{3.45-3.5}{0.7} - \frac{2.95-3.5}{0.7}$$

$$-0.1 - -0.8 = 0.460172 - 0.211855 = 0.248317$$

c.e)  $(3.45 < x < 3.95)$

$$\frac{3.95-3.5}{0.7} - \frac{3.45-3.5}{0.7}$$

$$0.6 - -0.1 = 0.725747 - 0.460172 = 0.265575$$

c.f)  $(3.95 < x < 4.45)$

$$\frac{4.45-3.5}{0.7} - \frac{3.95-3.5}{0.7}$$

$$1.4 - 0.6 = 0.919243 - 0.725747 = 0.193496$$

c.g)  $(4.45 < x < 4.95)$

$$\frac{4.95-3.5}{0.7} - \frac{4.45-3.5}{0.7}$$

$$2.1 - 1.4 = 0.982136 - 0.919243 = 0.062893$$

c) ( 1pto) **Calcular el valor crítico de la tabla  $\chi^2$  usando el nivel de significación:  $\alpha$**

$$X^2_{cal} = \frac{(2-0.48)^2}{0.48} + \frac{(1-2.11)^2}{2.11} + \frac{(4-5.80)^2}{5.80} + \frac{(15-9.93)^2}{9.93} + \frac{(10-10.62)^2}{10.62} + \frac{(5-7.73)^2}{7.73} + \frac{(3-2.52)^2}{2.52} = 9.62$$

$$X_{1-\alpha}(7-0-1)=6 \quad X^2_{TAB(0.95)}(6)=12.95$$

d) ( 1pto) **Criterio de decisión y Conclusión al problema.**

Dado que 9.62 es  $<$  a 12.95 no se rechaza la hipótesis nula

Conclusión:

La duración de baterías están distribuidas normalmente

## Problema 2

Suponga que usted sea dueño de una cafetería en la UNIVERSIDAD PRIVADA “ALAS PERUANAS” con sede en Ica y que para la venta de sus hamburguesas, ofrezca una salsa especial que acompaña a los demás ingredientes. Usted cuenta con clientes regulares que conforman la comunidad universitaria en todos los días. La mayoría de sus clientes son docentes, estudiantes y empleados de oficina y de no oficina. Algunos de ellos casi siempre piden la salsa especial, pero usted tiene la impresión de que a los estudiantes y a los docentes les gusta más esta salsa que a los empleados. Basándose en 10 días representativos para cada grupo, puede confirmar esta sospecha par un nivel de significación de  $\alpha = 5\%$

Número de clientes piden la salsa especial			
Día	Docentes	Empleados	Estudiantes
1	25	15	25
2	10	18	29
3	14	13	28
4	14	15	27
5	20	14	25
6	27	9	28
7	19	9	29
8	22	10	27
9	19	11	26
10	14	13	28

- a) ( 0.5 pto. ) Identifique la variable de estudio o de respuesta y el tipo a la que pertenece.

**Y: preferencias de salsa**

**X: Numero de clientes que piden salsa especial**

**Tipo: cuantitativa discretas**

- b) ( 0.5 pto. ) Plantee las hipótesis correspondientes.

$$H_0 : \mu_1 = \mu_2 = \mu_3$$

$H_1$  : Algún  $\mu_i$  de las medias difiere de las demás

- c) ( 2.0 pto. ) Realice el análisis de varianza para ver si existe alguna diferencia significativa entre estos grupos.

### ANÁLISIS DE VARIANZA

Origen de las	Suma de	Grados de	Promedio de los	F	Probabilidad	Valor crítico
---------------	---------	-----------	-----------------	---	--------------	---------------

<i>variaciones</i>	<i>cuadrados</i>	<i>libertad</i>	<i>cuadrados</i>	<i>para F</i>		
Entre grupos	31.9	9	3.54444444	0.0362418	0.9999925	2.210697
Dentro de los grupos	2934	30	97.8			
Total	2965.9	39				

d) ( 1.0 pto. ) Decisión y haga conclusiones finales.

Rechazo  $H_0$  Si  $F_{cal} > 2.21$

No Rechazo  $H_0$  Si  $F_{cal} \leq 2.21$

Como  $F_{cal} = 0.036 < 2.21$

$\therefore$  Rechazamos  $H_1 \Rightarrow$  Aceptamos  $H_0$

Conclusión:

En La preferencia de la salsa tanto para los docentes, alumnos, y trabajadores no se ve la diferencia de la aceptación.

### Problema 3

Supóngase que el administrador de una planta industrial conjetura que el rendimiento (en número de artículos producidos por turno de 8 horas) de una línea de producción depende de dos variables cualitativas: El primero depende del supervisor de la línea (de los cuales hay dos, digamos  $A_1$  y  $A_2$ ) y el segundo depende del turno para el cual se mide la producción ( al que dividiremos en tres turnos, de 8:00 am a 4:00 pm, de 4:00 pm a 12:00 am y de 12:00 am a 8:00 am. A los que denotaremos por:  $B_1$ ;  $B_2$  y  $B_3$ :

El administrador desea establecer si existen diferencias entre las producciones de los distintos turnos, con cada supervisor y si existe interacción entre los dos factores, para ello efectuó  $r = 3$  réplicas de un experimento factorial de  $2 \times 3$  para investigar el efecto de “**supervisor**” ( con dos niveles ) y .turno.(con tres niveles). Las observaciones se dan en la tabla a continuación:

Supervisor (Factor A)	Turno (Factor B)		
	B <sub>1</sub> (8:00 am - 4:00 pm)	B <sub>2</sub> (4:00 pm - 12:00 am)	B <sub>3</sub> (12:00 am - 8:00 am)
A <sub>1</sub>	570	480	470
	610	475	430
	625	540	450
A <sub>2</sub>	480	625	630
	515	600	680
	465	580	660

- a) ( 0.5 pto. ) Identifique la variable de estudio o de respuesta y el tipo a la que pertenece.

**X: El rendimiento de una línea de producción**

**Variable:cuantitativa-continua**

- b) ( 0.5 pto. ) Plantee las hipótesis correspondientes.

**X1 : intervalo de producción cualitativa nominal**

**X2 : superficie de producción cualitativa nominal**

$H_0$  : no hay interacción entre las investigaciones

$H_1$  : si hay interacción entre las investigaciones

- c) ( 2.0 pto. ) Realice el análisis de varianza para ver si existe alguna relación significativa entre estos factores.

	B1 (8:00 am -4:00 pm)	B2 (4:00 pm -12:00 am)	B1 (12:00 am -8:00 am)
A1	570	480	470
	610	475	430
A2	515	600	680
	465	580	660

Análisis de varianza de dos factores con varias muestras por grupo				
RESUMEN	b	c	D	Total
A1				
Cuenta	2	2	2	6
Suma	1180	955	900	3035
Promedio	590	477.5	450	505.8333333
Varianza	800	12.5	800	4724.166667

A2				
Cuenta	2	2	2	6
Suma	980	1180	1340	3500
Promedio	490	590	670	583.3333333
Varianza	1250	200	200	6836.666667

<i>Total</i>			
Cuenta	4	4	4
Suma	2160	2135	2240
Promedio	540	533.75	560
Varianza	4016.666667	4289.583333	16466.66667

ANÁLISIS DE VARIANZA						
<i>Origen de las variaciones</i>	<i>Suma de cuadrados</i>	<i>Grados de libertad</i>	<i>Promedio de los cuadrados</i>	<i>F</i>	<i>Probabilidad</i>	<i>Valor crítico para F</i>
Muestra	18018.75	1	18018.75	33.13793103	0.001197326	5.987377607
Columnas	1504.166667	2	752.0833333	1.383141762	0.320632257	5.14325285
Interacción	53037.5	2	26518.75	48.77011494	0.000194593	5.14325285
Dentro del grupo	3262.5	6	543.75			
Total	75822.91667	11				

d) ( 1.0 pto. ) Decisión y haga conclusiones finales.

**Nota:** Recuerde que si no hay interacción, se realiza un análisis de varianza por factor como si fuese de una sola vía

## Problema 4

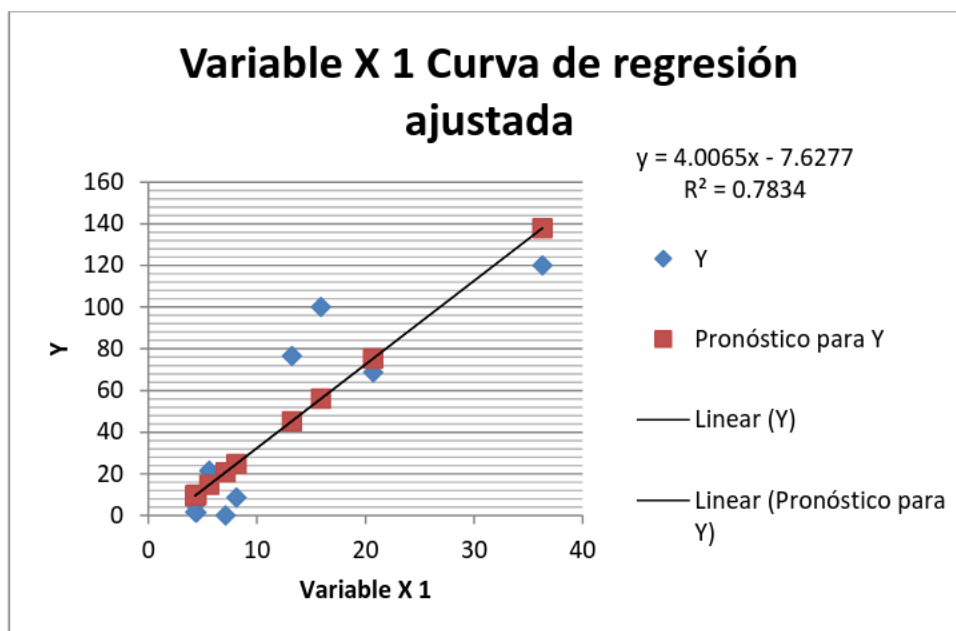
Se presume que hay una relación lineal entre las Ventas de cerveza y los gastos de publicidad en la que realiza cada una de las marcas de cerveza en el Perú. El siguiente cuadro muestra los ventas (en millones de dólares) y los Gastos en publicidad (en millones de dólares) para 10 marcas principales de cerveza en el Perú:

<i>Marca</i>	<b>Ventas Y</b>	<b>Gastos en publicidad X</b>
Cristal	120	36,3
Pilsen callao	68,7	20,7
Cusqueña	100,1	15,9
Brahma	76,6	13,2
Arequipeña	8,7	8,1
Zenda	0,1	7,1
Franca	21,5	5,6
Club	1,4	4,4
Pilsen trujillo	5,3	4,3
Barena	1,7	4,3

a) ( 0.5 pto. ) Identifique las variables de estudio y el tipo a la que pertenecen.

**X,Y: Ambas son continuas cuantitativas continuas**

b) ( 0.5 pto. ) Haga un gráfico de dispersión. ¿Cuál es el modelo de estimación a usarse?





Estadísticas de la regresión								
Coeficiente de correlación	0.8850711							
Coeficiente de determinación	0.7833508							
R <sup>2</sup> ajustado	0.7562696							
Error típico	22.831353							
Observaciones	10							
ANÁLISIS DE VARIANZA								
	Grados de libertad	Suma de cuadrados	Cuadrado de los cuadrados	F	Valor crítico de F			
Regresión	1	15078.3034	15078.3034	28.9260522	0.00066299			
Residuos	8	4170.16558	521.270698					
Total	9	19248.469						
	Coeficientes	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%	Inferior 95.0%	Superior 95.0%
Intercepción	-7.6276699	11.4849257	-0.66414621	0.52527213	-34.1119561	18.8566163	-34.1119561	18.8566163
Variable X 1	4.0064779	0.74493464	5.37829454	0.00066299	2.28865554	5.72430024	2.28865554	5.72430024
Análisis de los residuales								
Observación	Predicción	Residuos						
1	137.80748	-17.8074775						
2	75.306422	-6.60642243						
3	56.075329	44.0246714						
4	45.257838	31.3421618						
5	24.824801	-16.124801						
6	20.818323	-20.7183231						
7	14.808606	6.69139372						
8	10.000833	-8.60083281						
9	9.600185	-4.30018502						
10	9.600185	-7.90018502						

El modelo a usarse es el modelo lineal

c) ( 2.0 pto. ) Según su modelo de estimación y de regresión, responda a las siguientes preguntas:

- ¿Se incrementarán las ventas de cerveza en un próximo período al aumentar los gastos de publicidad? (Solo observe su tendencia del gráfico de dispersión para dar su respuesta)

Si se incrementa

- ¿Es adecuado suponer que el ajuste entre estas variables es efectivamente lineal teniendo en cuenta los valores de las variables? Ajuste el modelo lineal e interprete el coeficiente de correlación y los coeficientes del modelo lineal.



Pues si es adecuado suponer que el ajuste entre las variables es efectivamente lineal

---

*Estadísticas de la regresión*

---

Coeficiente de correlación múltiple	0.8850711
Coeficiente de determinación $R^2$	0.7833508

$R = 1$  Entonces correlación entre X e Y es perfecta positiva.

- ¿Qué porcentaje de las ventas de cerveza son explicadas por las variaciones de los gastos de publicidad?

37.8%

- Si se presupuesta para un próximo período en 12.6 millones de dólares los gastos en publicidad ¿Cuál serían las ventas de cerveza para ese próximo período?

Seria de 42.85%

- La asociación de “Cervezas del Perú” estiman que para un próximo período las ventas de cerveza alcanzaran la cifra de 48.5 millones de dólares ¿cuál sería el monto de gasto en publicidad que tendrían que llevar a cabo; para lograr dicha estimación?

- c) ( 1 pto. ) En base a sus resultados, está en lo correcto suponer que entre las ventas de cerveza y los gastos de publicidad; hay una relación lineal.

Seria 13.9 millones de dolares

## Problema 5

Supongamos que las ventas de una empresa en la ciudad de Ica sean como se muestra en el cuadro. La *variable dependiente* para el análisis será “el nivel anual de ventas” en miles de soles de los diferentes periodos anuales. Se aprecia en el también que los Gastos en publicidad varían enormemente de un periodo a otro en cuanto a las comisiones cobradas por los vendedores también son muy variables y este es un factor que seguramente debe afectar el nivel de las Ventas anuales.

Nuestro problema consistirá, por tanto, en encontrar aquellos coeficientes que nos permitan explicar esta variación observada; por lo que una mejor aproximación de estimación a usarse es la siguiente regresión múltiple:

$$\hat{Y} = a + b \cdot X_1 + c \cdot X_2$$

	Y	X1	X2
AÑO	VENTAS	GASTOS DE PUBLICIDAD	COMISIONES DE VENDEDORES

2000	264000	550	15840
2001	384000	590	19250
2002	400200	680	26013
2003	422400	700	16896
2004	543000	750	16290
2005	548000	760	16380
2006	556000	750	16200
2007	543000	740	16500
2008	538000	750	16400
2009	558000	780	16900
2010	562000	800	17300
2011	553000	820	17500
2012	567000	840	17800

a) ( 0.5 pto. ) Identifique las variables de estudio y el tipo a la que pertenecen.

**Y:ventas de una empresa en la ciudad de ica**  
**Variable: cuanlitativa –continua**

b) ( 0.25 pto. ) ¿Cuál es el modelo de estimación?

Es el modelo multineal  $Y = A + BX_1 + CX_2$

c) ( 2.25 pto. ) Según su modelo de estimación y de regresión, responda a las siguientes preguntas:

Resumen

---

*Estadísticas de la regresión*

Coficiente

de

correlación      0.9400786

múltiple

Coeficiente

de

determinación 0.8837477

$R^2$  8

0.8604973

$R^2$  ajustado 3

35748.869

Error típico 3

Observaciones

13

#### ANÁLISIS DE VARIANZA

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	2	9.7152E+1	4857599326	38.00993	2.1233E-05
Residuos	10	1.278E+10	1277981656		
Total	12	1.0993E+1	1		

	Coeficiente	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%	Inferior 95.0%	Superior 95.0%
Intercepción	177544.80	124323.62	-	0.183745	-454555.11	99465.5	-	99465.5
Variable X 1	1031.7857	124.50795	8.28690594	8.64E-06	754.36471	1309.2	754.36	1309.20
Variable X 2	4.6478152	3.9291616	1.182902551	0.264208	-13.402533	4.1069	-	4.10690

- ¿Es adecuado suponer que el ajuste entre estas variables es efectivamente multilíneal teniendo en cuenta los valores de las variables? Ajuste el modelo multilíneal e interprete el coeficiente de correlación que responderá a esta pregunta.

Pues si es de acuerdo suponer que estas variables son efectivamente multilíneales

#### Estadísticas de la regresión

Coeficiente de correlación múltiple	0.9400786
Coeficiente de determinación $R^2$	0.88374778
$R^2$ ajustado	0.86049733
Error típico	35748.8693
Observaciones	13

$R = 1$  Entonces correlación entre X e Y es perfecta positiva.

- ¿Qué porcentaje de la variabilidad de las ventas **son explicadas** por las variaciones de los Gastos en publicidad y Las comisiones cobradas?

63.68%

- Si se incrementa el gasto en publicidad para este año en 880 mil soles y la comisión que se espera pagar es de 18200 ¿Cuál será la predicción de las Ventas anual en este año?

Será de 620012

- d) ( 1 pto. ) En base a sus resultados, está en lo correcto la suposición de que el nivel de ventas anuales aumentan al aumentar los gastos en publicidad y las comisiones de los vendedores; haga sus conclusiones que le ayudarán a tomar decisiones y sugerencias.

Entonces también aumentaran las ventas también para la empresa por lo que será favorable para el siguiente año .