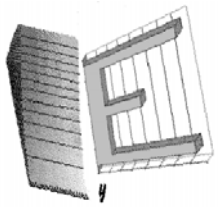
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE INGENIERÍA
DIVISIÓN DE CIENCIAS BÁSICAS
COORDINACIÓN DE CIENCIAS APLICADAS
DEPARTAMENTO DE PROBABILIDAD Y ESTADÍSTICA
PRIMER EXAMEN FINAL
RESOLUCIÓN



SEMESTRE 2009-2 **TIPO 1**
DURACIÓN MÁXIMA 2.5 HORAS **JUNIO 11 DE 2009**

NOMBRE_____

1. "Scram" es el término que utilizan los ingenieros nucleares para describir un rápido cierre de emergencia de un reactor nuclear. La industria nuclear ha hecho esfuerzos por reducir significativamente el número de cierres no planeados. La tabla muestra el número de "Scrams" en cada uno de 56 reactores nucleares en E.U.A.; 2004
   a) Agrupar los datos en ocho intervalos, con un ancho de clase de dos y comenzando en la frontera inferior -1.5
   b) Calcular la media del número de "Scrams" para datos agrupados.
   c) Obtener el segundo cuartil (la mediana) para datos agrupados.

Marcas de clase	Frecuencia absoluta
-0.5	6
1.5	13
3.5	18
5.5	6
7.5	8
9.5	3
11.5	1
13.5	1

15 Puntos
Resolución

a) La tabla de distribución de frecuencias es

Límites	Fronteras	Marcas de clase	f_i	f_i^*	F_i	F_i^*
-1 - 0	-1.5 - 0.5	-0.5	6	0.107	6	0.107
1 - 2	0.5 - 2.5	1.5	13	0.232	19	0.339
3 - 4	2.5 - 4.5	3.5	18	0.321	37	0.660
5 - 6	4.5 - 6.5	5.5	6	0.107	43	0.767
7 - 8	6.5 - 8.5	7.5	8	0.142	51	0.910
9 - 10	8.5 - 10.5	9.5	3	0.053	54	0.964
11 - 12	10.5 - 12.5	11.5	1	0.017	55	0.982
13 - 14	12.5 - 14.5	13.5	1	0.017	56	1.000
			56			

b) El promedio de "Scrams" para datos agrupados está dada por la expresión

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{m} x_i f_i = \sum_{i=1}^{m} x_i f_i^*$$

sustituyendo

$$\bar{x} = \frac{1}{56}\sum_{i=1}^{8} x_i f_i = \frac{1}{56}\left[(-0.5)(6) + \ldots + (13.5)(1)\right] = \frac{226}{56} = 4.036$$

c) El segundo cuartil es un valor que divide al conjunto de datos en dos subconjuntos de igual tamaño, se obtiene mediante una interpolación lineal, entre las fronteras y la frecuencia acumulada, de tal manera que se acumulen 28 datos, el valor de la frontera resultado de la interpolación, será el segundo cuartil, mismo que coincide con la mediana.

Fronteras	**Fi**
2.5	19
$\tilde{x}$	28
4.5	37

$$m = \frac{37-19}{4.5-2.5} = \frac{18}{2} = 9$$

con la ecuación de una recta dado un punto y la pendiente, se tiene:

$$y - 19 = 9(x - 2.5)$$

con $(\tilde{x}\ ,\ 28)$, sustituyendo:

$$28 - 19 = 9(\tilde{x} - 2.5)$$

$$\tilde{x} = 1 + 2.5$$

$$\tilde{x} = 3.5$$

Se observa que en este caso, la mediana coincide con el punto medio de las fronteras de la clase en cuestión.

2. En una planta de artículos electrónicos, se sabe por experiencia que la probabilidad de que un nuevo trabajador que haya asistido al programa de capacitación conozca la cuota de producción es 0.86 y que la probabilidad correspondiente es 0.35 para otro que no haya asistido. Si el 80 % de todos los empleados de ingreso reciente asisten al programa, determinar la probabilidad de que un nuevo empleado conozca la cuota de producción.

10 Puntos

Resolución

Sean los eventos

A: *El trabajador asistió al programa de capacitación.*

B: *El trabajador conoce la cuota de producción.*

Se pide calcular la probabilidad de que un nuevo empleado conozca la cuota de producción, $P(B)$.

De los datos se tiene: $P(A) = 0.8$, $P(\bar{A}) = 0.2$, $P(B|A) = 0.86$ y $P(B|\bar{A}) = 0.35$

Del teorema de probabilidad total

$$P(B) = P(A \cap B) + P(\bar{A} \cap B)$$

$$P(B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A})$$

sustituyendo

$$P(B) = 0.8(0.86) + 0.2(0.35) = 0.758$$

3. Una tienda de computadoras adquirió tres computadoras Dell a \$4000.00 cada una. Las venderá a \$5500.00 cada una. El fabricante se comprometió a adquirir cualquier computadora que no se haya vendido después de un periodo especificado a \$2000.00 cada una. Sea X la variable aleatoria que representa el número de computadoras vendidas y supóngase que $P(X=0)=0.1$, $P(X=1)=0.2$, $P(X=2)=0.3$ y $P(X=3)=0.4$. Donde $U(X)$ denota la utilidad asociada con la venta de X unidades, la información dada implica que $U(X)=Ingresos-Costos=5500X+2000(3-X)-12000=3500X-6000$.

   a) Calcular la utilidad esperada.
   b) Determinar la desviación estándar de la utilidad.

15 Puntos

Resolución

La función de probabilidad de X se puede escribir, como

x	0	1	2	3
$f_X(x)$	0.1	0.2	0.3	0.4

a) La utilidad está dada por

$U(X)=Ingresos-Costos=5500X+2000(3-X)-12000=3500X-6000$

entonces la utilidad esperada es

$E[U(X)]=E(3500X-6000)=3500E(X)-6000$

El valor esperado del número de computadoras vendidas está definido por

$$E(X)=\sum_{\forall x} x\, f_X(x)$$

sustituyendo

$E(X)=(0)(0.1)+(1)(0.2)+(2)(0.3)+(3)(0.4)$

$E(X)=2$

sustituyendo en la utilidad esperada

$E[U(X)]=3500(2)-6000=1000$

b) La variancia de la utilidad esperada es el segundo momento con respecto de la media, o bien, el segundo momento con respecto del origen menos el primero al cuadrado

$Var[U(X)]=Var(3500X-6000)=(3500)^2\ Var(X)$

Se sabe que la variancia está definida por

$Var(X)=E(X^2)-[E(X)]^2$

Calculando el segundo momento con respecto del origen

$$E(X^2)=\sum_{\forall x} x^2 f_X(x)$$

sustituyendo

$E(X^2)=(0)^2(0.1)+(1)^2(0.2)+(2)^2(0.3)+(3)^2(0.4)$

$E(X^2)=5$

La variancia es

$Var(X)=5-[2]^2=1$

La variancia de la utilidad es

$Var[U(X)]=(3500)^2$

La desviación estándar de la utilidad es la raíz de la variancia

$\sigma_U=\sqrt{Var(U(X))}=\sqrt{(3500)^2}=3500$

4. El chef de un restaurante de comida cantonesa prepara una ensalada que contiene, en promedio, cinco vegetales. Obtener la probabilidad de que la ensalada contenga más de cinco vegetales:
   a) en un día dado,
   b) por primera vez en abril, el quinto día.

15 Puntos
Resolución

Sea X la variable aleatoria que representa el número de vegetales que contiene la ensalada.

$$X \sim Poisson\left(\lambda = 5\frac{vegetales}{ensalada}\right)$$

a) La probabilidad de que la ensalada tenga más de cinco vegetales en un día dado, es

$$P(X>5) = 1 - \sum_{x=0}^{5} \frac{5^x e^{-5}}{x!}$$

desarrollando

$$P(X>5) = 1 - \left[\frac{5^0 e^{-5}}{1} + \frac{5^1 e^{-5}}{1} + \frac{5^2 e^{-5}}{2} + \frac{5^3 e^{-5}}{6} + \frac{5^4 e^{-5}}{24} + \frac{5^5 e^{-5}}{120}\right]$$

$$P(X>5) = 1 - e^{-5}\left[1 + 5 + \frac{25}{2} + \frac{125}{6} + \frac{625}{24} + \frac{3125}{120}\right]$$

$$P(X>5) = 1 - \frac{1097}{12}e^{-5} \approx 0.384$$

b) Sea Y la variable aleatoria que representa el quinto día de abril es el primer día que la ensalada contiene más de cinco vegetales.

$$Y \sim Geométrica(p = 0.384)$$

entonces

$$P(Y=5) = (0.616)^4(0.384) \approx 0.055$$

5. Sea la función de densidad conjunta

$$f_{XY}(x,y) = \begin{cases} \dfrac{xy}{96} & ; \quad 0 \le x \le 4 \ , \ 1 \le y \le 5 \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

a) Obtener las funciones de densidad marginal.
b) ¿Son X y Y variables aleatorias conjuntas independientes?
c) Obtener el coeficiente de correlación.

15 Puntos
Resolución

a) Las funciones marginales están definidas por:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy \qquad \text{y} \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx$$

sustituyendo se tiene

$$f_X(x) = \int_1^5 \frac{1}{96}xy\,dy = \frac{1}{96}x\left(\frac{1}{2}y^2\right)\Big|_1^5 = \frac{24}{192}x = \frac{1}{8}x \quad ; \quad 0 < x < 4$$

por lo tanto

$$f_X(x) = \begin{cases} \dfrac{1}{8}x & ; \quad 0 < x < 4 \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

para la función marginal de Y

$$f_Y(y) = \int_0^4 \frac{1}{96} xy\, dx = \frac{1}{96} y \left(\frac{1}{2} x^2 \right) \Big|_0^4 = \frac{1}{192} y(16-0) = \frac{16}{192} y \quad ; \quad 1 < y < 5$$

$$f_Y(y) = \begin{cases} \frac{1}{12} y & ; \quad 1 < y < 5 \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

b) Para que las vv.aa. sean independientes, debe cumplir

$$f_{XY}(x,y) = f_X(x)\, f_Y(y)$$

sustituyendo

$$\frac{1}{96} xy = \left(\frac{1}{8} x \right) \left(\frac{1}{12} y \right)$$

$$\frac{1}{96} xy = \frac{1}{96} xy$$

por lo tanto, sí son variables aleatorias conjuntas estadísticamente independientes.

c) El coeficiente de correlación está definido por

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

por independencia, se tiene

$$\rho(X,Y) = 0$$

Las variables aleatorias, no tienen asociación lineal.

6. Un antropólogo quiere estimar la estatura promedio de los hombres de cierta raza. Si se supone que la desviación estándar de la población es de 2.5 [cm] y se seleccionan al azar a 100 hombres. Calcular la probabilidad de que la diferencia entre la media de la muestra y la media verdadera de la población no exceda de 0.5 [cm]

15 Puntos

Resolución

Sea X la v.a. que representa la estatura de los hombres.

$$X \sim Normal\left(\mu_X \ , \ \sigma_X^2 = (2.5)^2 \right)$$

Sea X_i una muestra aleatoria de 100 hombres. $i = 1, 2, ..., 100$

$$X_i \sim Normal\left(\mu_X \ , \ \sigma_X^2 = (2.5)^2 \right) \quad ; \quad i = 1, 2, ..., 100$$

La muestra es grande y se conoce la variancia de la población, entonces por el Teorema del Límite Central

$$\bar{X} \sim Normal\left(\mu_{\bar{X}} = \mu_X \ , \ \sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} \right)$$

sustituyendo

$$\bar{X} \sim Normal\left(\mu_{\bar{X}} = \mu_X \ , \ \sigma_{\bar{X}}^2 = \frac{(2.5)^2}{100} \right)$$

se pide calcular

$$P\left(\left| \bar{X} - \mu_X \right| < 0.5 \right) = P\left(-0.5 < \bar{X} - \mu_X < 0.5 \right) \approx P\left(\frac{-0.5}{\frac{2.5}{10}} < \frac{\bar{X} - \mu_X}{\frac{\sigma_X}{\sqrt{n}}} < \frac{0.5}{\frac{2.5}{10}} \right) = P(-2 < Z < 2)$$

De tablas de la distribución acumulativa normal estándar:

$$P(-2 < Z < 2) = F_Z(Z = 2) - F_Z(Z = -2) = 0.9772 - 0.0228 = 0.9544$$

7. En la producción de herramientas de acero, se ha considerado ilustrar la relación entre la deformación (x) y la dureza Brinell (y).
   a) Obtener la ecuación de la recta de regresión.
   b) Calcular el coeficiente de determinación.
   c) La dureza cuando la deformación es de 25 [mm]

$x\ [mm]$	6	9	11	22	26	28	33	35
$y\left[\frac{kg}{mm^2}\right]$	68	67	65	44	40	37	34	32

Usar los cálculos siguientes

	x	y	x^2	y^2	xy
Suma:	170	387	4496	20423	7012

15 Puntos
Resolución

a) El ajuste de los datos a un modelo lineal de regresión por el criterio de mínimos cuadrados está dado por

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

donde

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

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

realizando los productos y las sumas, se tiene

x	y	x^2	y^2	xy
6	68	36	4624	408
9	67	81	4489	603
11	65	121	4225	715
22	44	484	1936	968
26	40	676	1600	1040
28	37	784	1369	1036
33	34	1089	1156	1122
35	32	1225	1024	1120
Suma:	**170**	**387**	**4496**	**20423**	**7012**

de donde

$$SS_{xx} = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n} = 4496 - \frac{(170)^2}{8} = 883.5$$

$$SS_{xy} = \sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n} = 7012 - \frac{(170)(387)}{8} = -1211.75$$

sustituyendo

$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}} = \frac{-1211.75}{883.5} = -1.372$$

para calcular los promedios, se tiene

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{8}(170) = 21.25$$

y

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{8}(387) = 48.375$$

sustituyendo

$$\hat{\beta}_0 = 48.375 - (-1.372)(21.25) = 77.53$$

por lo tanto el modelo lineal de regresión es

$$\hat{y} = 77.53 - 1.372\,x$$

b) El coeficiente de correlación está definido por

$$r = \frac{SS_{xy}}{\sqrt{SS_{xx}\,SS_{yy}}}$$

calculando

$$SS_{yy} = \sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n} = 20423 - \frac{(387)^2}{8} = 1701.875$$

por lo que el coeficiente de correlación es

$$r = \frac{-1211.75}{\sqrt{(883.5)(1701.875)}} = -0.988$$

El coeficiente de determinación está definido por

$$R^2 = (r)^2 = (-0.988)^2 = 0.976$$

Se concluye que las variables tienen una buena relación lineal, puesto que: $R^2 = 0.976$

c) Si la deformación es de 25 [mm], entonces

$$\hat{y} = 77.53 - 1.372(25) = 43.23$$