UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE INGENIERÍA
DIVISIÓN DE CIENCIAS BÁSICAS
COORDINACIÓN DE CIENCIAS APLICADAS
DEPARTAMENTO DE PROBABILIDAD Y ESTADÍSTICA
SOLUCIÓN SEGUNDO EXAMEN FINAL "A"

SEMESTRE 2016 -1 | **TIPO A**
DURACIÓN MÁXIMA 2.0 HORAS | **11 DE DICIEMBRE DE 2015**

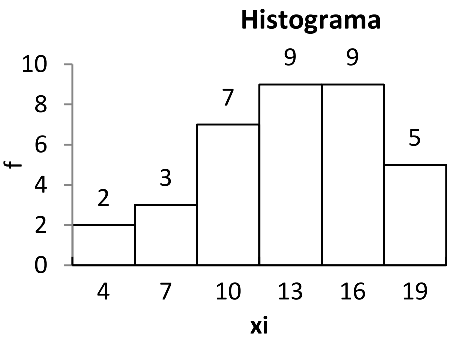
1. Estudios recientes han indicado la presencia de partículas dañinas para el ser humano, la secretaría de salud hizo un estudio preliminar durante 35 días, y los resultados encontrados fueron agrupados en la siguiente tabla de distribución de frecuencias:

Intervalo	Front. Inferior	Front. Superior	x_i	f	f*	F	F*
1	2.5	5.5	4	2	0.0571	2	0.0571
2	5.5	8.5	7	3	0.0857	5	0.1429
3	8.5	11.5	10	7	0.2000	12	0.3429
4	11.5	14.5	13	9	0.2571	21	0.6000
5	14.5	17.5	16	9	0.2571	30	0.8571
6	17.5	20.5	19	5	0.1429	35	1.0000
			Suma	35	1.0000		

Determine si la distribución de los datos es sesgada y justifique su respuesta. **(15 ptos.)**

SOLUCIÓN:

Al realizar el histograma de frecuencias, se obtiene:



En la gráfica se observa que la distribución de los datos no es simétrica y que los datos se encuentran sesgados hacia la izquierda. Por lo tanto la distribución de los datos es sesgada.

Si se emplean las medidas de tendencia central $\bar{x}, \tilde{x}, x_{mo}$, se sabe que la distribución de un conjunto de datos es sesgada si: $\bar{x} \neq \tilde{x} \neq x_{mo}$

En nuestro caso:

$$\bar{x} = \frac{1}{n}\sum x_i f_i = \frac{1}{35}[4(2) + 7(3) + 10(7) + 13(9) + 16(9) + 19(5)] = 13$$

$\tilde{x}$ es el valor para el cual la frecuencia acumulada es $\frac{n+1}{2} = 18$, el cual se obtiene interpolando los siguientes datos:

Frontera de clase superior	Frecuencia acumulada
11.5	12
$\tilde{x}$	18
14.5	21

Al realizar la interpolación se obtiene: $\tilde{x} = \left(\frac{18-12}{21-12}\right)(14.5 - 11.5) + 11.5 = 13.5$

x_{mo} = valor de la marca de clase con mayor frecuencia, en este caso existen 2 modas, las cuales son: 13 y 16.

Al comparar: $\bar{x} = 13, \tilde{x} = 13.5,\ x_{mo} = 13, 16$, se observa que estas tres medidas de tendencia central no son exactamente iguales, por lo tanto, la distribución de los datos analizados no es simétrica, por lo que puede decirse que la distribución de los datos es sesgada.

Si se calcula el coeficiente de sesgo a_3, se sabe que si $a_3 \neq 0$, entonces la distribución de los datos analizados es sesgada.

$a_3 = \frac{m_3}{(s)^3}$ donde: m_3= 3er momento respecto a la media y s^3 = $(desviación\ estándar)^3$

$m_3 = \frac{1}{n}\sum_{i=i}^{m}(xi - \bar{x})^3 fi$

$=\frac{1}{35}[(4-13)^3(2) + (7-13)^3(3) + (10-13)^3(7) + (13-13)^3(9) + (16-13)^3(9) + (19-13)^3(5)]$

= - 27.77

Desviación estándar: s = $\sqrt{s^2}$

$s^2 = \frac{1}{n-1}\sum_{i=1}^{m}(x_i - \bar{x}\ \ )^2 f_i$

$=\frac{1}{34}[(4-13)^2(2) + (7-13)^2(3) + (10-13)^2(7) + (13-13)^2(9) + (16-13)^2(9) + (19-13)^2(5)]$

= 17.47

∴ : s = $\sqrt{s^2}$ = $\sqrt{17.47} \approx 4.18$

Por lo que: $a_3 = \frac{m_3}{(s)^3} = \frac{-27.77}{(4.18)^3} \approx -0.38$, como $a_3 < 0$, entonces la distribución de los datos analizados es sesgada hacia la izquierda.

2. En un departamento de control de calidad, se deben inspeccionar losetas para piso para determinar si tienen defectos en la superficie. El porcentaje de losetas con defectos es de 4%. La probabilidad de que le toque inspección a una loseta es de 5%. La probabilidad de tener una loseta inspeccionada y defectuosa es de 0.2%
Si a un cliente le llega una loseta defectuosa, ¿Cuál es la probabilidad de que haya sido inspeccionada?

(15 ptos.)

SOLUCIÓN:

Datos: Sean los eventos: D= la loseta está defectuosa
I = la loseta fue inspeccionada

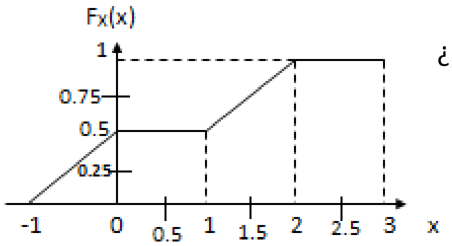
Del enunciado se sabe que: P(D)=0.04 , P(I)=0.05 y P(D∩I)= 0.2% = 0.002

La probabilidad de que una loseta defectuosa haya sido inspeccionada es: P(I|D)

$P(I|D)= \frac{P(I\cap D)}{P(D)} = \frac{0.002}{0.04} = 0.05$

∴ La probabilidad de que una loseta defectuosa haya sido inspeccionada es 0.05= 5%

3. Considere la siguiente figura:



¿Cuál es la probabilidad de que la variable aleatoria X sea ≤1.5?

(15 ptos.)

SOLUCIÓN:

La probabilidad de que la variable aleatoria X sea ≤ 1.5 es: P(X ≤ 1.5)

P(X ≤ 1.5) = $F_x(1.5)$, la gráfica muestra la distribución de probabilidad acumulada de la variable aleatoria X. En ella se observa que para X=1.5 el valor de F(x) es 0.75, por lo tanto: P(X ≤ 1.5) es 0.75

4. Suponga que el 70% de los ingenieros civiles que presentan el examen para obtener la licencia de perito de construcción, lo aprueban. Si 50 ingenieros civiles presentarán el próximo sábado el siguiente examen de este tipo, encuentre la probabilidad de que de 37 a 39 ingenieros civiles lo aprueben. **(15 ptos.)**

SOLUCIÓN:

Sea X la variable aleatoria discreta que representa el número de ingenieros civiles que aprueban el examen.

X tiene una distribución binomial con parámetros n=50 y p=0.7, por lo tanto: f(x)=$\binom{n}{x} p^x q^{n-x}$

La probabilidad de que de 37 a 39 ingenieros aprueben el examen es: P(37$\leq X \leq 39$)

P(37$\leq X \leq 39$)=$\binom{50}{37}(0.7)^{37}(0.3)^{13} + \binom{50}{38}(0.7)^{38}(0.3)^{12} + \binom{50}{39}(0.7)^{39}(0.3)^{11}$

= 0.1050 + 0.0838 + 0.0601 $\approx$ 0.2489

O bien como p>0.5 y np= (50)(0.7)= 35 >5

Entonces la variable aleatoria X binomial, puede aproximarse a una variable aleatoria continua X normal con parámetros: μ_x =np=35 y $\sigma^2{}_x = npq = (35)(0.3) = 10.5$

$\therefore$ X ~ N(μ_x =35, $\sigma^2{}_x = 10.5$)

Entonces, la probabilidad de tener de 37 a 39 ingenieros que aprueben el examen es :

P(37$\leq X \leq 39$)= P($\frac{37-0.5-39}{\sqrt{10.5}} \leq z \leq \frac{39+0.5-39}{\sqrt{10.5}}$) = P(-0.7715$\leq z \leq 0.1543$) $\approx F_z(0.15) - F_z(-0.77)$

= 0.5596 - 0.2206 = 0.3390

5. Una población de profesionales tiene un salario promedio de $25,000 con desviación estándar de $2,135. Si se toma una muestra de tamaño 100, ¿cuál es la probabilidad de que el promedio de sus salarios sea menor a $24,500? **(15 ptos.)**

SOLUCIÓN:

Sea X la variable aleatoria continua que representa el valor del salario de los profesionales, con $\mu_x = \$25{,}000$ y $\sigma_x = \$2135$

Sea $\overline{X}$ la variable aleatoria continua que representa el valor promedio de una muestra de 100 salarios.

Por el teorema del límite central: $\bar{X} \sim N(\mu_{\bar{X}} = \mu_x, \sigma^2_{\bar{x}} = \frac{\sigma^2_x}{n})$

Por lo tanto: $\bar{X} \sim N(\mu_{\bar{X}} = 25{,}000,\ \sigma^2_{\bar{x}} = \frac{(2135)^2}{100})$

La probabilidad de que el promedio de los salarios muestreados sea menor a $24,500 es:
P($\bar{X}$<24500)

P($\bar{X}$<24500)= P(Z<$\frac{24500-25000}{\frac{2135}{\sqrt{100}}}$) = P(Z<-2.3419)$\approx F_z(-2.34)$= 0.0096

6. Dos variables aleatorias tienen la siguiente función de densidad conjunta

$$f(x,y) = \begin{cases} \frac{1}{2}(x + y) & ; \quad 0 \le x \le 1 \quad 1 \le y \le 2 \\ 0 & ; \quad en\ cualquier\ otro\ caso \end{cases}$$

a) Calcule la probabilidad de que Y>1.5 dado que X<1.
b) Calcule la probabilidad de que X< 0.8 **(15 ptos.)**

<u>SOLUCIÓN:</u>

a) Se pide calcular: P(Y>1.5|X<1)

P(Y>1.5|X<1) = $\frac{P(Y > 1.5,\ X < 1)}{P(X<1)}$

P(Y>1.5, X<1) = $\int_{1.5}^{\infty}\int_{-\infty}^{1} f(x,y)dxdy = \frac{1}{2}\int_{1.5}^{2}\int_{0}^{1}(x+y)dxdy = \frac{1}{2}\int_{1.5}^{2}\left(\frac{x^2}{2}+xy\right)\Big|_0^1 dy$

= $\frac{1}{2}\int_{1.5}^{2}\left(\frac{1}{2}+y\right)dy$= $\frac{1}{2}\left(\frac{1}{2}y + \frac{y^2}{2}\right)\Big|_{1.5}^{2}$ = $\frac{1}{2}\left[\frac{1}{2}(2)+\frac{2^2}{2}-\left(\frac{1}{2}(1.5)+\frac{1.5^2}{2}\right)\right]$ = $\frac{1}{2}(1.125)$ = 0.5625 = $\frac{9}{16}$

P(x<1) = $\int_{-\infty}^{1} f(x)dx$

f(x) = $\int_{-\infty}^{\infty} f(x,y)dy$ = $\frac{1}{2}\int_{1}^{2}(x+y)\,dy$= $\frac{1}{2}$ (xy+$\frac{y^2}{2}$)$|_1^2$ = $\frac{1}{2}$ (2x+$\frac{2^2}{2}$ - (x+$\frac{1^2}{2}$)

= $\frac{1}{2}$ (2x+2-x - $\frac{1}{2}$) = x+1 - $\frac{x}{2}$ - $\frac{1}{4}$ = $\frac{x}{2}+\frac{3}{4}$ = $\frac{1}{2}(x+\frac{3}{2})$

$$\therefore f(x) = \begin{cases} \frac{1}{2}\left(x+\frac{3}{2}\right) \ ; \ 0 \le x \le 1 \\ 0 \ ; \quad en\ otro\ caso \end{cases}$$

Entonces: $P(X<1) = \frac{1}{2}\int_0^1(x+\frac{3}{2})$ dx = $\frac{1}{2}\left(\frac{x^2}{2}+\frac{3}{2}x\right)\Big|_0^1 = \frac{1}{2}(\frac{1^2}{2}+\frac{3}{2}(1)-0)$= $\frac{1}{2}(\frac{1}{2}+\frac{3}{2}) = \frac{1}{4}+\frac{3}{4}$ = 1

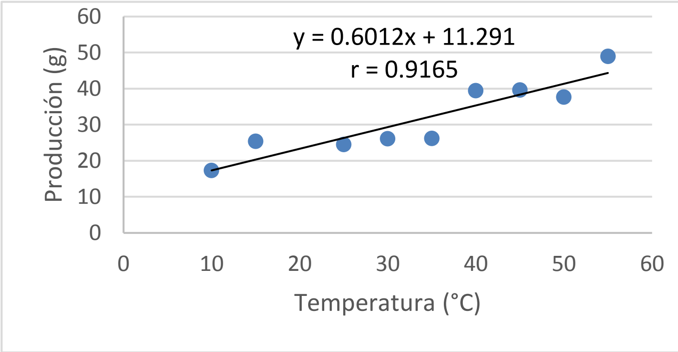
$\therefore$ P(Y>1.5| X<1) = $\frac{0.5625}{1}$ = 0.5625 = $\frac{9}{16}$

b) Piden P(X < 0.8)

P(X < 0.8) = $\int_{-\infty}^{0.8} f(x)dx$ = $\int_0^{0.8}\frac{1}{2}\left(x+\frac{3}{2}\right)$ dx = $\frac{1}{2}\left(\frac{x^2}{2}+\frac{3}{2}x\right)\Big|_0^{0.8} = \frac{1}{2}\left(\frac{0.8^2}{2}+\frac{3}{2}(0.8)-0\right) = 0.76$

7. Los siguientes datos corresponden a un experimento químico, para el cual se quiere determinar si existe alguna relación entre la temperatura ambiente (°C) y la producción de la reacción (gr).

	X [°C]	Y [gr]	XY
	10	17.3	173
	15	25.4	381
	25	24.5	612.5
	30	26.1	783
	35	26.2	917
	40	39.4	1576
	45	39.6	1782
	50	37.6	1880
	55	48.9	2689.5
Sumas:	*305*	*285*	*10794*
	$\bar{X}$= 33.89	**$\bar{Y}$ = 31.67**	



Con los datos proporcionados, conteste la siguiente pregunta: ¿Cuál de las siguientes afirmaciones es incorrecta?

a. Existe una relación lineal entre la temperatura ambiente y la producción.
b. La recta de regresión pasa por el punto (33.89, 31.67).
c. La covarianza de X y Y es positiva.
d. El coeficiente de determinación es 0.9165. **(10 ptos.)**

SOLUCIÓN:

De acuerdo a la información proporcionada el coeficiente de correlación r tiene un valor de 0.9165, al ser este valor cercano a 1, puede decirse que sí existe una relación lineal entre la temperatura ambiente y la producción de la reacción, por lo que la afirmación a. es correcta.

Al sustituir el valor de x= 33.89 en la ecuación de la recta de regresión obtenida: y = 0.6012x + 11.291
Se obtiene el valor de y= 31.6656, por lo que puede confirmarse que la recta de regresión pasa por el punto: (33.89, 31.67), por lo tanto la afirmación b. es correcta.

Al calcular el valor de la covarianza de X y Y se tiene:

$$\text{Cov (X, Y)}= E(XY) - E(X)E(Y) = \frac{10794}{9} - (33.89)(31.67) = 126.037$$

Se observa que el valor obtenido es positivo, por lo tanto, la afirmación c. es correcta.

El valor del coeficiente d determinación r^2 es: $r^2 = (0.9165)^2 = 0.8399$, por lo tanto la afirmación d. es incorrecta.

De acuerdo a lo anterior, la respuesta a la pregunta es: ***el inciso d****, ya que ésta es la única afirmación incorrecta.*