PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FACULTAD DE CIENCIAS SOCIALES
ESPECIALIDAD DE ECONOMÍA

EXAMEN PARCIAL
CURSO: EST 241 Estadística Inferencial. HORARIO: 0621. PROFESOR: Arturo Calderón G.
FECHA: 18 de mayo de 2019. SEMESTRE: 2019-1. DURACIÓN DE LA PRUEBA: 2.5 horas.

Salvo formulario y calculadora, no está permitido el uso de otro material de consulta durante el desarrollo de la prueba, ni el uso de otros dispositivos electrónicos (celulares, tablets, laptops, iPods, etc.). Cualquier material relativo al curso, así como celulares y dispositivos electrónicos, deben permanecer guardados durante la prueba.

Problema 1 (8 puntos)

a) Si el número de trabajadores eventuales que tiene cualquier taller de mecánica de un distrito es una variable aleatoria discreta X con función de probabilidad $P_X(x)$ dada en la tabla de abajo. Si se sabe que la probabilidad de que X tome su valor máximo es la mitad de la probabilidad de que X tome su valor mínimo:

X	0	1	2	3	(1)Halle a, b y el número esperado μ_X de trabajadores eventuales por taller. (1p.)
$P_X(x)$	a	0.4	0.3	b	
(2)Si en general un taller tiene 1 trabajador formal que gana 950 u.m. mensuales y cada trabajador eventual recibe 750 u.m. mensuales de pago ¿Cuál sería el costo promedio de la planilla de un taller? Si son 40 los talleres en ese distrito ¿Cuánto se espera que pagaran en total a todos sus trabajadores? (2p.)					

b) En una región el valor total Y de la producción de la región es función de la inversión X del gobierno regional vía $Y = 100 + 0.7X + U$, donde X e Y están medidas en millones de unidades monetarias y U es una v.a. continua que representa una perturbación aleatoria que puede afectar el ingreso regional, donde U tiene función de densidad: $f_U(u) = \frac{1}{2a} \quad -a \le u \le a$. Se sabe que si la inversión es de $X = 300$ millones de unidades monetarias, entonces el valor Y de la producción será inferior a 340 millones con 2/3 de probabilidad. ¿Cuál es el valor de a? (2p.)

c) Un comerciante compra S unidades de un bien al precio unitario de 6 soles para revender el producto al precio unitario de 9 soles. Pasada la temporada de ventas, cualquier sobrante se desecha. La cantidad de producto que le pueden demandar a este mayorista es una v.a. continua $X \sim N(\mu_X = 650, \sigma_X^2)$ y se sabe que con 67% de probabilidad esta demanda no pasará de 694 unidades. Halle σ_X y el stock óptimo S del producto que debiera tener por temporada para maximizar la utilidad esperada con la venta de este bien. (3p.)

Solución:

a) En (1) Como $1 = \sum_{x=0}^{3} P_X(x) = a + 0.4 + 0.3 + b = a + b + 0.7 \Rightarrow a + b = 0.3$ (I); además $b = P(X = 3)$ es la mitad de $P(X = 0) = a$, o sea "b es la mitad de a": $b = P(X = 3) = \frac{1}{2}P(X = 0) = \frac{a}{2} \Rightarrow a = 2b$ (II) $\Rightarrow$ reemplazando en (I): $2b + b = 0.3 \Rightarrow b = 0.1, a = 0.2$ y la función de probabilidad $P_X(x)$ de X es la de la tabla 1. Para hallar $\mu_X = \sum_{x=0}^{3} xP_X(x)$ mejor ordenamos datos como figura en la tabla 2:

Tabla 1

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

$\Rightarrow$

Tabla 2

X	0	1	2	3	Total
$P_X(x)$	0.2	0.4	0.3	0.1	1
$x\,P_X(x)$	0	0.4	0.6	0.3	1.3= μ

El valor esperado de X es $E(X) = \mu_X = 1.3$

(2) Sea CT = Costo de la planilla de un taller cualquiera del distrito, entonces $CT = 950 + 750X$ y el "costo promedio de la planilla" es $E(CT) = E(950 + 750X) = E(950) + E(750X) = 950 + 750E(X) = 950 + 750 \times 1.3 = 1{,}925$ u.m., o sea, en promedio un taller paga 1,925 u.m. a todos sus trabajadores (el formal más los eventuales que pudiera tener).

Si son 40 los talleres en el distrito, en total se espera que paguen $40 \times 1{,}925 = 77{,}000$ u.m. a todos los trabajadores del distrito, entre formales y eventuales.

b) Por dato: Si $X = 300 \Rightarrow P(Y \le 340) = \frac{2}{3} \Leftrightarrow P(100 + 0.7 \times 300 + U \le 340) = P(310 + U \le 340) = P(U \le 30) = \int_{-a}^{30} \frac{1}{2a} du = \frac{1}{2a}[u]_{-a}^{30} = \frac{30+a}{2a} = \frac{2}{3}$ que genera la ecuación $90 + 3a = 4a \Rightarrow a = 90$ y $f_U(u) = \frac{1}{180} \quad -90 \le u \le 90$ es la función de densidad específica de U.

c) $X \sim N(\mu_X = 650, \sigma_X^2)$, entonces para hallar σ_X, usamos el dato $0.67 = P(X \le 694) = P\left(Z \le \frac{694-650}{\sigma_X}\right) \Rightarrow \frac{694-650}{\sigma_X} = 0.44$ $\Rightarrow 44 = 0.44\sigma_X \Rightarrow \sigma_X = 100$.

La utilidad del comerciante, por temporada, es $U(X,S) = \begin{cases} 9X - 6S & si\ X < S \\ 3S & si\ X \ge S \end{cases}$ (pues si le demandan su costo siempre es $6S$; si le demandan $X < S$ unidades sólo venderá las X unidades y el sobrante $(S - X)$ se pierde, no se vende ni remata. En cambio si le demandan $X \ge S$, sólo puede vender el stock completo S y nada más. Tomando valor esperado:

$E[U(X,S)] = \int_{-\infty}^{\infty} U(X,S) f_X(x)dx = \int_{-\infty}^{S} (9x - 6S) f_X(x)dx + \int_{S}^{\infty} 3S f_X(x)dx =$

$9\int_{-\infty}^{S} x f_X(x)dx - 6S \underbrace{\int_{-\infty}^{S} f_X(x)dx}_{P(X\le S)=F_X(S)} + 3S \underbrace{\int_{S}^{\infty} f_X(x)dx}_{P(X>S)=1-F_X(S)} = 9\int_{-\infty}^{S} x f_X(x)dx - 6SF_X(S) + 3S\big(1 - F_X(S)\big) =$

$9\int_{-\infty}^{S} x f_X(x)dx - 6SF_X(S) + 3S - 3SF_X(S) = 9\int_{-\infty}^{S} x f_X(x)dx - 9SF_X(S) + 3S =: \varphi(S)$ que es función diferenciable de S y que podemos derivar para maximizar:

$\frac{d}{ds}\varphi(S) = \frac{d}{ds}\left(9\int_{-\infty}^{S} x f_X(x)dx\right) - \frac{d}{ds}\big(9SF_X(S)\big) + 3 = 9Sf_X(S) - 9F_X(S) - 9S\overbrace{\frac{d}{ds}F_X(S)}^{f_X(S)} + 3 = -9F_X(S) + 3 = 0 \Rightarrow S$ satisface $F_X(S) = \frac{1}{3} = 0.33 = P(X \le S) = P\left(Z \le \frac{S-650}{100}\right) \Rightarrow \frac{S-650}{100} = -0.44 \Rightarrow S = 606$ es el punto crítico. Como $\varphi''(S) = -9f_X(S) < 0$ se trata de un máximo. El stock $S = 606$ es el "óptimo", el que maximiza la utilidad esperada del comerciante.

Problema 2 (5 puntos)
Sea (X,Y) vector aleatorio discreto donde X=# de créditos que toma un cliente de un banco e Y=# de créditos, de los tomados, que el mismo cliente tiene que refinanciar. Si (X,Y) tienen f. de p. conjunta $P_{XY}(x,y) = c(2x+y)$ $\ x = 1,2,3; y = 0,1,\dots x$. donde $c > 0$ es constante.

a) Halle c y la probabilidad de que refinancie cada crédito tomado. (2p.)
b) Halle $P_X(x)$ y $E(Y|X)$ ¿Se podría explicar y predecir la cantidad de créditos refinanciados a partir de la cantidad de créditos tomados? Justifique. (3p.)

Solución:
a) Aplicamos $\sum_{(x,y)} P_{XY}(x,y) = 1$. Dado el rango R_{XY} con pocos valores es más sencillo usar una tabla (**recordando que, por el enunciado, $y = 0, 1, \dots x$ o sea se debe cumplir $y \le x$.**)

$P_{XY}(x,y)$	x			
y	1	2	3	Total
0	2c	4c	6c	12c
1	3c	5c	7c	15c
2	0	6c	8c	14c
3	0	0	9c	9c
Total	5c	15c	30c	50c

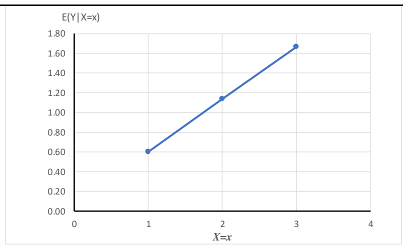
$50c = 1 \Rightarrow c = \frac{1}{50}$ y multiplicando por c tenemos $P_{XY}(x,y)$ y sumando obtenemos $P_X(x)$ y $P_Y(Y)$ en los márgenes inferior y derecho

$P_{XY}(x,y)$	x			
y	1	2	3	$P_Y(y)$
0	0.04	0.08	0.12	0.24
1	0.06	0.10	0.14	0.30
2	0	0.12	0.16	0.28
3	0	0	0.18	0.18
$P_X(x)$	0.10	0.30	0.60	1.00

Si A = "Refinancia todos sus créditos" $\Rightarrow A = \{(x,y)|x = y)\} \Rightarrow P(A) = P_{XY}(1,1) + P_{XY}(2,2) + P_{XY}(3,3) = 0.06 + 0.12 + 0.18 = 0.36$

b) $P_X(x)$ ya figura en el margen inferior de la tabla de $P_{XY}(x,y)$. Para $E(Y|X)$ tenemos que tabular cada $P_{Y|X}(y|x) = \frac{P_{XY}(x,y)}{P_X(x)}$, $y = 1,2,\dots,x;\ x = 1,2,3.$ y calcular el valor esperado $E(Y|X)$ correspondiente:

$P_{XY}(y\|x)$	$X = 1$	$X = 2$	$X = 3$
y	$P_{Y\|X}(y\|1)$	$P_{Y\|X}(y\|2)$	$P_{Y\|X}(y\|3)$
0	0.400	0.267	0.200
1	0.600	0.333	0.233
2	0.000	0.400	0.267
3	0.000	0.000	0.300
$E(Y\|X = x)$	0.60	1.13	1.67



Los puntos son colineales y es sencillo calcular la ecuación de la recta: $E(Y|X = x) = \frac{8x+1}{15}$ $\ x = 1,2,3$

Entre la cantidad X de créditos tomados y la cantidad Y de créditos refinanciados, hay una relación lineal directa. En efecto, sí es posible explicar y predecir Y a partir de X.

Nota: Otra manera de resolver el problema, más analítica, pero que implica conocer ciertas identidades, es:
$1 = \sum_{(x,y)} P_{XY}(x,y) = \sum_{x=1}^{3}\sum_{y=0}^{x} c(2x+y) = c\sum_{x=1}^{3}\left[\sum_{y=0}^{x}(2x+y)\right] = c\sum_{x=1}^{3}\left[\sum_{y=0}^{x} 2x + \sum_{y=0}^{x} y\right] = c\sum_{x=1}^{3}\left[x\sum_{y=0}^{x} 2\right] + c\sum_{x=1}^{3}\left[\sum_{y=0}^{x} y\right] = c\sum_{x=1}^{3}\left[2x(x+1) + \frac{x(x+1)}{2}\right] = c\frac{5}{2}\sum_{x=1}^{3}[x(x+1)] = c\frac{5}{2}[2+6+12] = 50c \Rightarrow c = \frac{1}{50}$ y $P_{XY}(x,y) = \frac{1}{50}(2x+y)$ $\ x = 1,2,3; y = 0,1,\dots x$. De ahí, es directa la probabilidad $P(A) = P(Y = X) = P_{XY}(1,1) + P_{XY}(2,2) + P_{XY}(3,3) = 0.06 + 0.12 + 0.18 = 0.36$; También $P_X(x) = \sum_{y=0}^{x}\frac{1}{50}(2x+y) = \frac{1}{50}\left[\frac{5}{2}x(x+1)\right] = \frac{x(x+1)}{20}$ $\ x = 1,2,3$ y
$P_{Y|X}(y|x) = \frac{\frac{1}{50}(2x+y)}{\frac{x(x+1)}{20}} = \frac{2}{5}\frac{(2x+y)}{x(x+1)}$ $\ y = 0,1,\dots,x; x = 1,2,3\ (fijo) \Rightarrow E(Y|X) = \sum_{y=0}^{x} y\frac{2}{5}\frac{(2x+y)}{x(x+1)} = \frac{2}{5x(x+1)}\sum_{y=0}^{x} y(2x+y) =$

$\frac{2}{5x(x+1)}\left[2x\sum_{y=0}^{x} y+\sum_{y=0}^{x} y^2\right]=\frac{2}{5x(x+1)}\left[2x(\frac{x(x+1)}{2})+(\frac{x(x+1)(2x+1)}{6})\right]=\frac{2}{5}\left[x+(\frac{2x+1}{6})\right]=\frac{2}{5}\left[(\frac{8x+1}{6})\right]=\frac{(8x+1)}{15}$, y es claro que $E(Y|X)$ es una función (lineal) de X, o sea sí se podría explicar y predecir la cantidad Y de créditos refinanciados a partir de la cantidad X de créditos tomados.

Para resolver esta parte, usando fórmulas cerradas, es necesario conocer las identidades: $\sum_{k=0}^{N} k=\sum_{k=1}^{N} k=\frac{N(N+1)}{2}$ y $\sum_{k=0}^{N} k^2=\sum_{k=1}^{N} k^2=\frac{N(N+1)(2N+1)}{2}$.

Problema 3 (7 puntos)

Sea (X,Y) v.a. continuo que representa las rentabilidades diarias (en puntos porcentuales) de las acciones A y B respectivamente. Si X tiene función de densidad $f_X(x)=\frac{3x^2+1}{4}$, $-1\le x\le 1$ y $f_{Y|X}(y|x)=\frac{3(x^2+y^2)}{(3x^2+1)}$ $0\le y\le 1; x=valor\ dado$, con $-1\le x\le 1$.

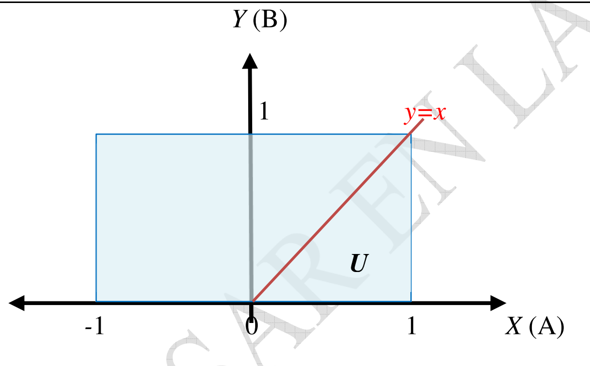
a) Halle la f. de densidad conjunta $f_{XY}(x,y)$ y la f. densidad marginal $f_Y(y)$. No olvide especificar los rangos correspondientes. (1p.)
b) Una persona comenta que B será menos rentable que A ¿Qué opina usted? Justifique con probabilidades. (1p.)
c) ¿Cuál sería la rentabilidad esperada Y de B, si A tuvo rentabilidad de x puntos porcentuales? (1p.)
d) ¿En promedio cuáles son las rentabilidades esperadas de X y de Y? (1.5p.)
e) Calcule la covarianza σ_{XY} ¿No habría ninguna relación entre estas rentabilidades? (1p.)
f) Le dicen que se puede predecir la rentabilidad de B conociendo la de A y viceversa. Analice la afirmación. (1.5p.)

Solución:

a) $f_{Y|X}(y|x)=\frac{f_{XY}(x,y)}{f_X(x)}\Rightarrow f_{XY}(x,y)=f_{Y|X}(y|x)f_X(x)=\frac{3(x^2+y^2)}{(3x^2+1)}\frac{3x^2+1}{4}=\frac{3}{4}(x^2+y^2)$ $-1\le x\le 1; 0\le y\le 1$ es la función de densidad conjunta de (X,Y).

Para hallar $f_Y(y)$ tenemos que por definición $f_Y(y)=\int_{-\infty}^{+\infty} f_{XY}(x,y)dx=\int_{-1}^{1}\frac{3}{4}(x^2+y^2)dx=\frac{3}{4}\left[\frac{x^3}{3}+y^2x\right]_{-1}^{1}=$ $\frac{3}{4}\left(\left[\frac{1}{3}+y^2\right]-\left[-\frac{1}{3}-y^2\right]\right)=\frac{3}{4}\left(2y^2+\frac{2}{3}\right)=\frac{1}{2}(3y^2+1)\Rightarrow f_Y(y)=\frac{1}{2}(3y^2+1)$ $0\le y\le 1$ es la f.d. marginal de Y.
En este caso no ha sido necesario graficar la zona de integración porque los rangos de X y de Y no se acotan mutuamente.

b) Sea $U=$ "B *será menos rentable que* A", tenemos que hallar $P(U)$ y ver si es mayor que 0.5, y si así fuera, opinaríamos que la predicción se cumplirá. Hay que aplicar integral doble, es mejor graficar U:



$U=\{(x,y)\in R_{XY}|y<x\}$ es la zona triangular (el triángulo debajo de la recta $y=x$) en el gráfico de la izquierda. Integrando

$$P(U)=\int_0^1\left[\int_0^x\frac{3}{4}(x^2+y^2)dy\right]dx=\frac{3}{4}\int_0^1\left[x^2y+\frac{y^3}{3}\right]_0^x dx$$

$=\frac{3}{4}\int_0^1\left(x^3+\frac{x^3}{3}\right)dx=\frac{3}{4}\int_0^1\left(\frac{4}{3}x^3\right)dx=\left(\frac{x^4}{4}\right)_0^1=\frac{1}{5}=0.25$ y

$P(U)=0.25<0.5$: Este evento no ocurrirá

c) Si ocurrió $X=x\Rightarrow$ La rentabilidad esperada de B sería $E(Y|X=x)=\int_{-\infty}^{\infty} yf_{Y|X}(y|x)dy=\int_0^1 y\frac{3(x^2+y^2)}{(3x^2+1)}dy=$ $\left[\frac{3\left(x^2\frac{y^2}{2}+\frac{y^4}{4}\right)}{(3x^2+1)}\right]_0^1=\frac{3\left(x^2\frac{1}{2}+\frac{1}{4}\right)}{(3x^2+1)}=\frac{3\left(\frac{2x^2}{4}+\frac{1}{4}\right)}{(3x^2+1)}=\frac{(6x^2+3)}{4(3x^2+1)}=\frac{(6x^2+3)}{(12x^2+4)}=\frac{3(2x^2+1)}{4(3x^2+1)}$ $-1\le x\le 1\Rightarrow$

$\boldsymbol{E(Y|X=x)=\frac{(6x^2+3)}{(12x^2+4)}}$ $\boldsymbol{-1\le x\le 1}$ es la rentabilidad esperada de B cuando A tiene rentabilidad de x puntos.

d) $\boldsymbol{E(X)=\mu_X=\int_{-1}^{1} xf_X(x)dx=\int_{-1}^{1} x\frac{3x^2+1}{4}dx=\frac{1}{4}\int_{-1}^{1}(3x^3+x)dx=\frac{1}{4}\left(3\frac{x^4}{4}+\frac{x^2}{2}\right)_{-1}^{1}=\frac{1}{4}\left(\frac{3}{4}+\frac{1}{2}-\frac{3}{4}-\frac{1}{2}\right)=0}$

$$\boldsymbol{E(Y)=\mu_Y=\int_0^1 yf_Y(y)dy=\int_0^1 y\frac{1}{2}(3y^2+1)dy=\frac{1}{2}\int_0^1(3y^3+y)dy=\frac{1}{2}\left(\frac{3y^4}{4}+\frac{y^2}{2}\right)_0^1=\frac{5}{8}=0.625}$$

e) $\sigma_{XY}=E(XY)-E(X)E(Y)=E(XY)-0\times\frac{5}{8}=E(XY)=\int_{-1}^{1}\left[\int_0^1 xy\frac{3}{4}(x^2+y^2)dy\right]dx=$ $\frac{3}{4}\int_{-1}^{1} x\left[\int_0^1 y(x^2+y^2)dy\right]dx=\frac{3}{4}\int_{-1}^{1} x\left[\left(x^2y^2+\frac{y^3}{3}\right)\right]_0^1 dx=\frac{3}{4}\int_{-1}^{1} x\left(x^2+\frac{1}{3}\right)dx=\frac{3}{4}\left[\left(\frac{x^4}{4}+\frac{x^2}{6}\right)\right]_{-1}^{1}=$

$\frac{3}{4}\left[\left(\frac{1}{4}+\frac{1}{6}\right)-\left(\frac{1}{4}+\frac{1}{6}\right)\right]_{-1}^{1} = 0 \Rightarrow \boldsymbol{\sigma_{XY} = 0}$. No hay relación lineal entre X e Y, pero como ya vimos en c)
$\boldsymbol{E(Y|X=x) = \frac{3(2x^2+1)}{4(3x^2+1)} \quad -1 \leq x \leq 1}$ muestra que entre Y y X hay una relación no lineal: Sí hay relación entre estas variables, pero es no lineal. El gráfico de abajo ilustra la relación no lineal.

f) Como $\boldsymbol{E(Y|X=x) = \frac{3(2x^2+1)}{4(3x^2+1)} \quad -1 \leq x \leq 1}$, sí se puede predecir la rentabilidad Y de B, conociendo la rentabilidad X de A (como ya se dijo, la figura de abajo ilustra este hecho, aunque no es necesario graficar).

Pasando a $E(X|Y=y)$:
$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{\frac{3}{4}(x^2+y^2)}{\frac{1}{2}(3y^2+1)} = \frac{3}{4}\frac{(x^2+y^2)}{(3y^2+1)} \quad -1 \leq x \leq 1;\ y = valor\ dado$, con $0 \leq y \leq 1 \Rightarrow$
$\boldsymbol{E(X|Y=y) = \int_{-1}^{1} x \frac{3}{4}\frac{(x^2+y^2)}{(3y^2+1)} dx = \frac{3}{4}\frac{1}{(3y^2+1)} \int_{-1}^{1} x(x^2+y^2) dx = \frac{3}{4}\frac{1}{(3y^2+1)}\left(\frac{x^4}{4} + y^2\frac{x^2}{2}\right)_{-1}^{1} = 0}$ y se ve que en este caso, conocer el valor de $\boldsymbol{Y}$ no permite predecir el valor de $\boldsymbol{X}$, así que esta parte de la afirmación no es cierta.