Universidad Nacional de Ingeniería
Escuela Profesional de Matemática
Ciclo 2020-1

[Cálculo de Probabilidades - CM1H2]
[Los profesores] UNI, 31 de agosto de 2020

Práctica Calificada 4

1. (4pts) Demostrar que si X tiene una distribución exponencial con media λ, entonces la variable aleatoria $Y = [X]$ (máximo entero de X) es una distribución geométrica de parámetro p. Determine p en términos de λ.

> **Solución:** $f_X(x) = \frac{1}{\lambda}e^{-x/\lambda}$, cuando $x > 0$. Sea $k = 0, 1, 2, ...$ se tiene que:
>
> $$P(Y = k) = P([X] = k) = P(k \leq X < k+1) = \int_k^{k+1} \frac{1}{\lambda}e^{-\frac{x}{\lambda}}dx = -e^{-x/\lambda}\Big|_k^{k+1} = e^{-k/\lambda}(1-e^{-1/\lambda}),$$
>
> es decir, $P(Y = k) = p(1-p)^k$ con $p = (1 - e^{-1/\lambda})$.

2. (5pts) Un examen de admisión de opción múltiple tiene 200 preguntas, cada una con 4 opciones, de las que sólo una es la correcta. Si se pudiera utilizar el Teorema de Límite Central (TLC), entonces responda: ¿Cuál es la probabilidad de que solamente adivinando se obtengan de 25 a 30 respuestas correctas para 80 de los 200 problemas sobre los que el estudiante no tiene conocimientos?. (Sug. Si F es la función acumulada de la Normal estándar, entonces $F(1{,}16) = 0{,}8770$ y $F(2{,}71) = 0{,}9966$ y una aproximación por continuidad de 0,5)

> **Solución:** La probabilidad de adivinar la respuesta es $1/4$. Como adivina sólo un grupo de 80 de las 200 preguntas. Definimos X como el número de respuestas correctas de las 80 que adivinó. Luego, se tiene que $X \sim Bi(80, 1/4)$. Entonces, nos piden: $\mathbb{P}[25 \leq X \leq 30]$.
> Pero, $X = \sum_{i=1}^{80} X_i$ con $X_i \sim Be(1/4)$ para $i = 1, \cdots, 80$, independientes. Como el número de variables de Bernoulli es mayor a 30 pordemos utilizar TLC y aproximar la estandarizada de X a una Normal estándar. Hallamos la media y la desviación estándar: $\mu = np = 80(\frac{1}{4}) = 20$ y $\sigma = \sqrt{npq} = \sqrt{80\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)} = \sqrt{15} \approx 3,873$. Usando la aproximación por continuidad, hallamos X entre $24,5$ y $30,5$. Entonces:
>
> $$z_1 = \frac{24{,}5 - 20}{3,873} = 1,16; \quad z_2 = \frac{30,5 - 20}{3,873} = 2,71$$
>
> Luego,
>
> $$\begin{aligned} \mathbb{P}[24{,}5 \leq X \leq 30{,}5] &= \mathbb{P}[1{,}16 < Z < 2,71] \\ &= \mathbb{P}[Z < 2,71] - \mathbb{P}[Z < 1,16] \\ &= 0,9966 - 0,8770 = 0,1196. \end{aligned}$$

3. Un mecánico mantiene un gran número de arandelas en un depósito. El 50 % de éstas son de 1/4 de pulgada de diámetro, el 30 % son de 1/8 de pulgada de diámetro, el 20 % restante con 3/8 de pulgada de diámetro. Se eligen 10 arandelas al azar.

a) (1pt) ¿Cuál es la probabilidad de que haya exactamente 5 arandelas de 1/4, 4 de 1/8, y 1 de 3/8 de pulgada?

b) (2pts) ¿Cuál es la probabilidad de que sólo haya 2 clases de arandelas entre las elegidas?

c) (1.5pts) ¿Cuál es la probabilidad de que las 3 clases de arandelas estén entre las elegidas?

d) (1.5pts) ¿Cuál es la probabilidad de que haya 3 de una clase, 3 de otra clase, y 4 de la tercera clase en la muestra?

Solución: Sean los eventos:
$A_1 = \{\text{Se elige 1 arandela de } 1/4\}$
$A_2 = \{\text{Se elige 1 arandela de } 1/8\}$
$A_3 = \{\text{Se elige 1 arandela de } 3/8\}$
Sean las v.a. X_i:"número de veces que ocurre A_i en las 10 extracciones", $i = 1, 2, 3$.

a)

$$\mathbb{P}[X_1 = 5, X_2 = 4, X_3 = 1] = \frac{10!}{5!\,4!\,1!}(0{,}5)^5(0{,}3)^4(0{,}2) \approx 0{,}064\,.$$

b) Puede ocurrir los sgtes casos: $A_1 \cap A_2$, $A_1 \cap A_3$ y $A_2 \cap A_3$.

i) Sea $X_1 = y$, $X_2 = 10 - y$, $X_3 = 0$, con y de 1 a 9, entonces

$$\sum_{y=1}^{9} \mathbb{P}[X_1 = 9, X_2 = 10 - y, X_3 = 0] = \sum_{y=1}^{9} \frac{10!}{y!\,(10-y)!}\,(0{,}5)^y\,(0{,}3)^{10-y} \approx 0{,}1064\,.$$

ii) Sea $X_1 = y$, $X_2 = 0$, $X_3 = 10 - y$, con y de 1 a 9, luego

$$\sum_{y=1}^{9} \mathbb{P}[X_1 = y, X_2 = 0, X_3 = 10 - y] = \sum_{y=1}^{9} \frac{10!}{y!\,(10-y)!}\,(0{,}5)^y\,(0{,}2)^{10-y} \approx 0{,}02727\,.$$

iii) Sea $X_1 = 0$, $X_2 = y$, $X_3 = 10 - y$, con y de 1 a 9, entonces

$$\sum_{y=1}^{9} \mathbb{P}[X_1 = 0, X_2 = y, X_3 = 10 - y] = \sum_{y=1}^{9} \frac{10!}{y!\,(10-y)!}\,(0{,}3)^y\,(0{,}2)^{10-y} \approx 0{,}00097\,.$$

De (i), (ii) y (iii) la probabilidad pedida es

$$\mathbb{P}[\text{solo hay 2 clases de arandelas}] = 0{,}1064 + 0{,}02727 + 0{,}00097 = 0{,}1346\,.$$

c) $\mathbb{P}[\text{Las 3 clases de arandelas estén}] = 1 - \mathbb{P}[\text{sólo haya 2 clases}] - \mathbb{P}[\text{haya sólo 1 clase}]$.

$$\begin{aligned}\mathbb{P}[\text{haya solo 1 clase}] &= \mathbb{P}[X_1 = 10, X_2 = 0, X_3 = 0] \\ &\quad + \mathbb{P}[X_1 = 0, X_2 = 10, X_3 = 0] + \mathbb{P}[X_1 = 0, X_2 = 0, X_3 = 10] \\ &= (0{,}5)^{10} + (0{,}3)^{10} + (0{,}2)^{1} \approx 0{,}00098\,.\end{aligned}$$

Entonces, $\mathbb{P}[\text{Las 3 clases de arandelas esten}] = 1 - 0{,}1346 - 0{,}00098 = 0{,}8644$

d) Sea $A = \{\text{haya 3 de una clase, 3 de otra y 4 de la tercera}\}$. Entonces

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}[X_1 = 3, X_2 = 3, X_3 = 4] \\ &\quad + \mathbb{P}[X_1 = 3, X_2 = 4, X_3 = 3] + \mathbb{P}[X_1 = 4, X_2 = 3, X_3 = 3] \\ &= \frac{10!}{3!\,3!\,4!}\,(0{,}5)^3\,(0{,}3)^3\,(0{,}2)^4 \\ &\quad + \frac{10!}{3!\,4!\,3!}\,(0{,}5)^3\,(0{,}3)^4\,(0{,}2)^3 + \frac{10!}{4!\,3!\,3!}\,(0{,}5)^4\,(0{,}3)^3\,(0{,}2)^3 \\ &= 0{,}1134\,.\end{aligned}$$

4. De una muestra de 8 observaciones conjuntas de valores de dos variables aleatorias X e Y, se obtiene la siguiente información:

$$\sum x_i = 24, \quad \sum x_i y_i = 64, \quad , \sum y_i = 40, S_Y^2 = 12, S_X^2 = 6.$$

Calcule:

a) (2pts) La recta de regresión de Y sobre X. Explique el significado de los parámetros.

b) (2pts) El coeficiente de determinación. Comente el resultado e indique el tanto por ciento de la variación de Y que no está explicada por el modelo lineal de regresión.

c) (1pt) Si el modelo es adecuado ¿cuál es la predicción para $x = 4$?

Solución: En primer lugar calculamos las medias y la covarianza entre ambas variables:

$$\overline{x} = \sum x_i/n = 24/8 = 3. \quad \overline{y} = \sum Y_i/n = 40/8 = 5.$$

$$S_{XY} = \frac{\sum X_i Y_i}{n} - \overline{x}\,\overline{y} = 64/8 - 3 \times 5 = -7.$$

Ahora determinamos:

$$b = \frac{S_{XY}}{S_X^2} = \frac{-7}{6} = -1,1667$$

Como es negativa, si X aumenta Y disminuye.

$$a = \overline{y} - b\overline{x} = 5 - \frac{-7}{6} \times 3 = 8,5.$$

Entonces $\widehat{Y} = -1,1667.X + 8,5$.

El grado de bondad del ajuste lo obtenemos por el coeficiente de determinación (r^2):

$$r^2 = \left(\frac{S_{XY}}{S_X.S_Y}\right)^2 = \frac{(-7)^2}{6 \times 12} = 0,6805 = 68,05\,\%.$$

El modelo de regresión linel explica el 68,05 % de la variabilidad de Y en función de la de X. Por tanto queda un 32 % de variabilidad no explicada.

La predicción que realiza el modelo para $x = 4$ es:

$$\widehat{y} = -1,1667(4) + 8,5 = 3,833.$$

lo cual hay que considerar con ciertas reservas pues hay un 32 % de variabilidad que no es explicada por el modelo.

Universidad Nacional de Ingeniería
Escuela Profesional de Matemática
Ciclo 2020-1

[Cálculo de Probabilidades - CM1H2]
[Los profesores] UNI, 10 de agosto de 2020

Práctica Calificada 3

1. (6 ptos) Sea X una variable aleatoria con la siguiente función de densidad:

$$f(x) = \begin{cases} \dfrac{kx^2}{9} & , \ 0 \leq x \leq 6 \\ 0 & , \ \text{en otro caso} \end{cases}$$

   a) Indique si se trata de una variable discreta o continua.
   b) Determine el valor de k para que f sea función de densidad. Represente dicha función.
   c) Obtenga la función de distribución de X y represéntela gráficamente.
   d) Halle la media, la mediana y la moda de la distribución. (Recuerde que la media es la esperanza de X, la mediana es el punto que divide en partes iguales a la distribución bajo la densidad, es decir $\mathbb{P}(X \leq x_{med}) = 0{,}5$, y la moda es la mayor frecuencia relativa de la densidad, si consideramos que la densidad representa la gráfica de las frecuencias relativas)
   e) Obtenga la varianza, la desviación estándar.

Solución:

a) La variable X es continua puesto que su función de densidad toma valores distintos de cero en un intervalo de la recta real: $0 \leq x \leq 6$.

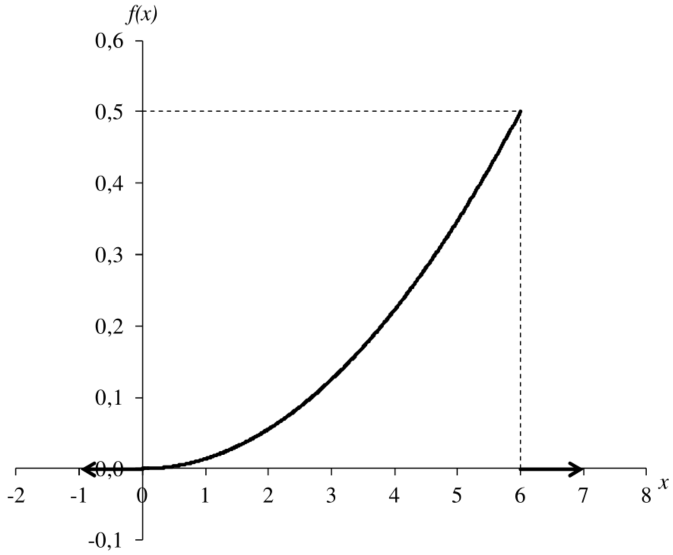
b) Deduciremos el valor de k teniendo en cuenta las propiedades que debe cumplir f para ser función de densidad:

   1) $f(x) = \frac{kx^2}{9} \geq 0 \Rightarrow k \geq 0$.
   2) $1 = \displaystyle\int_{\mathbb{R}} f = \int_0^6 \frac{kx^2}{9}\,dx = \frac{k}{9}\left[\frac{x^3}{3}\right]_0^6 = \frac{k}{9}\left[\frac{6^3}{3} - \frac{0^3}{3}\right] = 8k \Rightarrow k = \frac{1}{8}$.

   Por tanto, la función de densidad es:

$$f(x) = \begin{cases} \dfrac{x^2}{72} & , \ 0 \leq x \leq 6 \\ 0 & , \ \text{en otro caso} \end{cases}$$

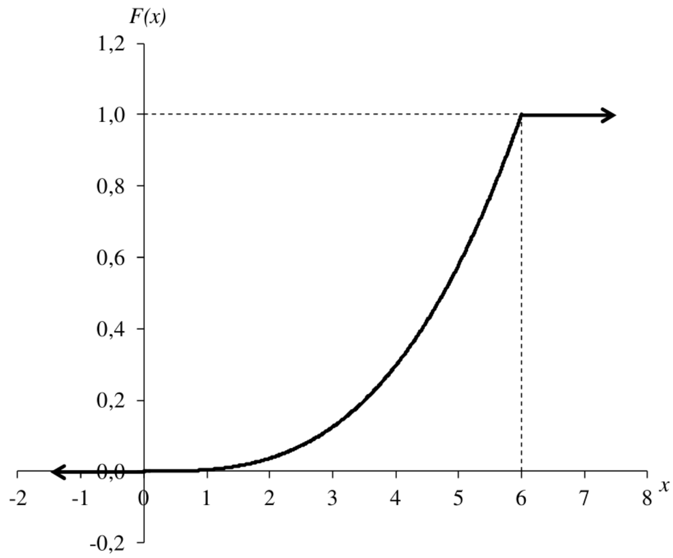
   Y cuya representación gráfica es



c) Para $0 \leq x \leq 6$, se tiene $\displaystyle\int_0^x f = \int_0^x \frac{t^2}{72}\,dt = \frac{1}{72}\left[\frac{t^3}{3}\right]_0^x = \frac{x^3}{216}$.

$$F(x) = \mathbb{P}(X \leq x) = \begin{cases} 0 & , \ x < 0 \\ \dfrac{x^3}{216} & , \ 0 \leq x \leq 6 \\ 1 & , \ x > 6. \end{cases}$$

Y su representación gráfica es



d) Media (μ):

$$\mu = \mathbb{E}(X) = \int_{\mathbb{R}} x f(x)\,dx = \int_0^6 x\frac{x^2}{72}\,dx = \frac{1}{72}\left[\frac{x^4}{4}\right]_0^6 = 4{,}5.$$

Mediana (x_{med}):

$$F(x_{med}) = \mathbb{P}[X \leq x_{med}] = 0{,}5$$

luego,

$$\frac{(x_{medd})^3}{216} = 0{,}5 \Rightarrow x_{med} = 4{,}76.$$

Moda (x_{mod}): $\quad x_{mod} = 6$.
Debido a que es el valor de la variable para el que la función de densidad alcanza su valor máximo y se observa directamente de su representación gráfica.

e) Varianza (σ^2):

$$\sigma^2 = Var(X) = \mathbb{E}[X^2] - \mu^2 = \int_0^6 x^2 \frac{x^2}{72}\, dx - (4{,}5)^2 = \frac{1}{72}\left[\frac{x^5}{5}\right]_0^6 - (4{,}5)^2 = 1{,}35.$$

Desviación estándar (σ):

$$\sigma = \sqrt{Var(X)} = \sqrt{1{,}35} = 1{,}162.$$

2. (6 ptos) En el tiempo 0, una moneda ha salido cara con probabilidad p, luego se lanza y cae al suelo. Elegimos intervalos de tiempo de acuerdo con un proceso de Poisson con tasa λ, durante el cual la moneda se recoge y se lanza. (Entre estos momentos, la moneda permanece en el suelo). ¿Cuál es la probabilidad de que la moneda salga cara en el momento t?
Sugerencia: ¿Cuál sería la probabilidad condicional si no hubiera giros adicionales en el tiempo t, y cuál sería si hubiera giros adicionales en el tiempo t?

Solución: Los tiempos que tomamos para mantener la moneda girando en el aire sigue un proceso de Poisson con tasa λ. Cada envento de Poisson en un intervalo de tiempo t será el número de giros que da la moneda durante dicho tiempo, la cual representaremos por la v.a N.
Sea A_t la abreviatura de "Algo es cierto en el momento t". En nuestro caso, A_t es "la moneda muestra cara en el momento t". Entonces, por probabilidad total respecto a $N = 0$ y $N \geq 1$, se tiene

$$\begin{aligned}\mathbb{P}(A_t) &= \sum_{N=0}^{\infty} \mathbb{P}[A_t | N \,\text{eventos han ocurrido en el tiempo}\, t] \cdot \mathbb{P}[N \,\text{eventos han ocurrido en el tiempo}\, t] \\ &= \sum_{N=0}^{\infty} \mathbb{P}[A_t | N \,\text{eventos han ocurrido en el tiempo}\, t] \cdot e^{-\lambda t}\frac{(\lambda t)^N}{N!}\end{aligned}$$

En nuestro caso, si la moneda es cara o no, solo depende del lanzamiento más reciente. En el primer lanzamiento suponemos que no hubo giros por lo que la probabilidad condicional respecto a $N = 0$ es segura "cara". Además, la probabilidad de obtener cara, no depende de cuántos giros se hayan producido, siempre será p.
Entonces la probabilidad pedida es:

$$\mathbb{P}(A_t) = e^{-\lambda t} + \sum_{N=1}^{\infty} p \cdot e^{-\lambda t}\frac{(\lambda t)^N}{N!} = e^{-\lambda t} + p(1 - e^{-\lambda t}).$$

3. (4 ptos) Determine la función de densidad de $Y = X^2$ siendo $X \sim U[-1, 2]$.

Solución: X tiene función de distribución:

$$F_X(x) = \begin{cases} 1 & x > 2 \\ \frac{x+1}{3} & -1 \leq x \leq 2 \\ 0 & x < -1 \end{cases}$$

como $\sqrt{Y} \in [-1, 2]$, entonces $Y \in [0, 4]$. Hallamos la función acumulada de Y:
$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}) = \frac{\sqrt{y}+1}{3} - \frac{-\sqrt{y}+1}{3} = \frac{2\sqrt{y}}{3}$. Por tanto $f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{1}{3\sqrt{y}}$. Por tanto, la función de densidad de Y será:

$$f_Y(y) = \begin{cases} \frac{1}{3\sqrt{y}} & 0 \leq y \leq 4 \\ 0 & \text{en otro caso} \end{cases}$$

4. (4 ptos) El senado de la facultad de una determinada universidad tiene 20 miembros. Suponga que hay 12 hombres y 8 mujeres. Se selecciona al azar un comité de 10 senadores.

   a) ¿Cuál es la probabilidad de que haya 6 hombres y 4 mujeres en el comité?

   b) ¿Cuál es el número esperado de hombres en este comité?

Solución: Sea X el número de hombres del comité de 10 seleccionados al azar. Entonces X es una variable aleatoria hipergeométrica con $N = 20$, $r = 12$ y $n = 10$. Con función de masa:

$$f(x) = \frac{\binom{12}{x}\binom{8}{10-x}}{\binom{20}{10}}.$$

a) $\mathbb{P}[X = 6] = \dfrac{\binom{12}{6}\binom{8}{4}}{\binom{20}{10}} \approx 0{,}3501$

b) $\mathbb{E}(X) = \dfrac{nr}{N} = \dfrac{12 \times 10}{20} = 6.$

Universidad Nacional de Ingeniería
Escuela Profesional de Matemática
Ciclo 2020-1

[Cálculo de Probabilidades - CM1H2]
[Los profesores] UNI, 13 de julio de 2020

Práctica Calificada 2

1. (5 ptos) Supongamos que la probabilidad de que un jurado, seleccionado para un juicio de un caso criminal, llegue al veredicto correcto es del 95 %. La policía estima que el 99 % de los individuos que llegan a un juicio son realmente culpables. Calcular la probabilidad de que un individuo sea realmente inocente dado que el jurado ha dictaminado que es inocente.

Solución:

- A_1 : El individuo es culpable. Dato policial: $\mathbb{P}(A_1) = 0{,}99$.
- A_2 : El individuo es inocente. Luego:

$$\mathbb{P}(A_2) = 1 - \mathbb{P}(A_1) = 1 - 0,99 = 0,01.$$

- B : El veredicto del jurado es correcto. Se necesita un veredicto.
- Veredicto Correcto: "INOCENTE". Luego, $\mathbb{P}(B|A_2) = 0,95$.
- Veredicto Incorrecto: "INOCENTE", es decir era culpable. Luego,

$$\mathbb{P}(B|A_1) = 1 - \mathbb{P}(B|A_2) = 1 - 0,95 = 0{,}05.$$

- Por el teorema de Bayes, se tiene que la probabilidad de que un individuo sea realmente inocente sabiendo que el jurado lo halló inocente es

$$\begin{aligned}\mathbb{P}(A_2|B) &= \frac{\mathbb{P}(B|A_2)\mathbb{P}(A_2)}{\mathbb{P}(B|A_1)\mathbb{P}(A_1) + \mathbb{P}(B|A_2)\mathbb{P}(A_2)}\\ &= \frac{0,95 \times 0,01}{0,05 \times 0,99 + 0,95 \times 0,01}\\ &= \frac{0,0095}{0,059} = 0,1610.\end{aligned}$$

2. (5 ptos) Un dado justo es lanzado 3 veces y los siguientes eventos son considerados:
   - A : los lanzamiento 1 y 2 producen diferentes resultados.
   - B : los lanzamiento 2 y 3 producen diferentes resultados.
   - C : los lanzamiento 3 y 1 producen diferentes resultados.

   Diga el valor de verdad de las siguientes proposiciones:

   a) Los eventos A, B y C son independientes dos a dos.
   b) Los eventos A, B y C son independientes.

Solución:

Definición 1 *Un conjunto de eventos es mutuamente independiente si cada evento es independiente de cualquier intersección de los otros eventos. Es decir, si y sólo si para cada $k \leq n$ y para cada subconjunto de k eventos $\{B_i\}_{i=1}^{k}$ de $\{A_i\}_{i=1}^{n}$, se cumple:*

$$\mathbb{P}\left(\bigcap_{i=1}^{k} B_i\right) = \prod_{i=1}^{k} \mathbb{P}(B_i).$$

Esta es llamada la regla de multiplicación para eventos independientes.

- Espacio muestral: $\Omega = \{CCC, CCS, CSC, CSS, SCC, SCS, SSC, SSS\}$.
- Los eventos son:

$$\begin{aligned} A_1 &:= A = \{CSC, CSS, SCC, SCS\}, \\ A_2 &:= B = \{CCS, CSC, SCS, SSC\}, \\ A_3 &:= C = \{CCS, CSS, SCC, SSC\}. \end{aligned}$$

- Hallamos las intersecciones:

$$\begin{aligned} A \cap B &= \{CSC, SCS\}, \quad B \cap C = \{CCS, SSC\} \\ A \cap C &= \{CSS, SCC\}, \quad A \cap B \cap C = \emptyset. \end{aligned}$$

- Verificamos la independencia dos a dos:

$$\mathbb{P}(A_i \cap A_j) = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} = \mathbb{P}(A_i)\mathbb{P}(A_j) \quad i \neq j.$$

- Verificamos que no son mutuamente excluyentes (aunque cumplan parte de la definición):

$$\mathbb{P}(A \cap B \cap C) = 0 \neq \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C).$$

- Respuesta: V F.

3. (5 ptos) Un operador elige al azar entre n chips de una caja. La probabilidad de que sea defectuoso es $0,2$.

   1. Si $n = 7$, ¿cuál es la probabilidad de que al menos 3 chips sean defectuosos?
   2. Si $n = 50$, ¿cuál es la probabilidad de tener entre 9 y 12 chips defectuosos?
   3. ¿Cuántos chips hay en la caja si la varianza es 32?

Solución:

1. $n = 7$, X : Chips defectuosos $\sim Bi(7; 0,2)$. Luego,

$$\begin{aligned} \mathbb{P}(X \geq 3) &= 1 - \mathbb{P}(X \leq 2) \\ &= 1 - [\mathbb{P}(0) + \mathbb{P}(1) + \mathbb{P}(2)] \\ &= 1 - \left[\binom{7}{0} 0,2^0 0,8^7 + \binom{7}{1} 0,2^1 0,8^6 + \binom{7}{2} 0,2^2 0,8^5\right] \\ &= 1 - 0,852 = 0,148. \end{aligned}$$

2. $n = 50$, $X \sim Bi(50; 0,2)$, $P(9 \leq X \leq 12) = 0,3948$ calcular por tabla o por calculadora.
3. Como $Var(X) = n.p.q = 32$, la probabilidad de éxito $p = 0{,}2$, y por tanto $q = 1 - p = 0{,}8$. Se obtiene que $n = 200$ chips.

4. (5 ptos) Determine el valor de verdad de las siguientes preguntas:

a) Si $\mathbb{P}(B \cap C) \neq 0$, entonces $\mathbb{P}(A|B \cap C)\mathbb{P}(B|C) = \mathbb{P}(A \cap B|C)$.

b) $\mathbb{P}(A) \neq 0$ y $\mathbb{P}(B) \neq 0$, entonces $\mathbb{P}(A|B) \geq \mathbb{P}(A)$ es equivalente a $\mathbb{P}(B|A) \geq \mathbb{P}(B)$.

Solución:

- Desde que $\mathbb{P}(B \cap C) \neq 0$, se tiene $0 < \mathbb{P}(B \cap C) \leq \mathbb{P}(C)$, luego $\mathbb{P}(C) \neq 0$ y por lo tanto está bien definido las probabilidades condicionadas: $\mathbb{P}(B|C)$ y $\mathbb{P}(A \cap B|C)$. Se cumple:

$$\mathbb{P}(A|B \cap C)\mathbb{P}(B|C) = \frac{\mathbb{P}(A \cap B \cap C)}{\cancel{\mathbb{P}(B \cap C)}} \times \frac{\cancel{\mathbb{P}(B \cap C)}}{\mathbb{P}(C)} = \mathbb{P}(A \cap B|C).$$

- Como $\mathbb{P}(A) \neq 0$ y $\mathbb{P}(B) \neq 0$ entonces están bien definidas las probabilidades condicionales. Entonces,

$$\begin{aligned} \mathbb{P}(A|B) \geq \mathbb{P}(A) &\iff \mathbb{P}(A \cap B) \geq \mathbb{P}(A)\mathbb{P}(B) \\ &\iff \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \geq \mathbb{P}(B) \\ &\iff \mathbb{P}(B|A) \geq \mathbb{P}(B). \end{aligned}$$

- Respuesta: V V.

Universidad Nacional de Ingeniería
Escuela Profesional de Matemática
Ciclo 2020-1

[Cálculo de Probabilidades - CM1H2]
[Los profesores] UNI, 22 de junio de 2020

Práctica Calificada 1

1. En una fábrica que produce tornillos se clasifican las unidades de un lote como defectuosos (D) o no defectuosos (N). Se empieza a clasificar y anotar los resultados hasta obtener 2 tornillos defectuosos consecutivos o hasta que se verifique 4 de ellos, cualquiera que ocurra primero. Describa el espacio muestral para este experimento. (Sug. utilice un diagrama de árbol) [5 ptos]

Solución: El espacio muestral es

$$\Omega = \{DD, DNDD, DNDN, DNND, DNNN, NDD, NDND, NDNN, NNDD, NNDN, NNND, NNNN\}.$$

2. Sea Ω un espacio muestral de un experimento aleatorio. Responder V o F según corresponda:

   a) El número de **eventos posibles** es calculado siempre por $2^{|\Omega|} - 1$. [1 ptos]

   b) Siempre hay por lo menos un evento posible. [1 ptos]

   c) Para dos eventos cualesquiera A y B de Ω se cumple que $A \cap B^c$ y $A^c \cap B$ son eventos mutuamente excluyentes. [1 ptos]

   d) Siempre es posible definir una secuencia infinita de eventos $E_1 \subset E_2 \subset E_3 \subset ....$ [2 ptos]

Solución:

1. (F) $\emptyset$ es el único evento que no es posible. Cumple en el caso finito. En el caso infinito no.

2. (V) Para el conjunto no vacío $\Omega \neq \emptyset$ (experimento aleatorio), es evento posible, $2^n - 1 \geq 1$, $\forall n \in \mathbb{N}$, $n = |\Omega|$.

3. (V) Probamos que $[A \cap B^c] \cap [A^c \cap B] = [A \cap A^c] \cap [B \cap B^c] = \emptyset \cap \emptyset = \emptyset$ son disjuntos.

4. (V) Si Ω es un conjunto finito podemos tener la familia $\{E_i\}_{i=1}^n$, definimos $[\Omega - E_1] \subset [\Omega - E_1 \cap E_2] \subset .... \subset [\Omega - E_1 \cap E_2 \cap ... \cap E_n] \subset F_{n+1} \subset F_{n+2} \subset ....$, donde los siguientes $F_j = F_n = \Omega - E_1 \cap ... \cap E_n$, $j \in \mathbb{N}$, $j > n$. Si es infinito, definimos de la familia infinita $\{A_i\}_{i \in \mathbb{N}}$ arbitraria de eventos, la familia: $[\Omega - A_1] \subset [\Omega - A_1 \cap A_2] \subset [\Omega - A_1 \cap A_2 \cap A_3] \subset ....$, determina una secuencia infinita de eventos monótonos.

3. Dada la secuencia de números reales distintos $x_1, x_2, \cdots, x_m$ con $m = n^2 + 1$. Si los números de esta secuencia son ordenados en un conjunto de columnas según el siguiente proceso:

   i) Colocamos a x_1 en la primera columna y, para $i > 1$,

   ii) si x_i es mayor o igual al valor que está encima de alguna de las columnas formadas, ponemos x_i encima de la primera de esta columnas que cumpla dicha condición.

iii) De lo contrario, comience una nueva columna con x_i.

Determine el valor de verdad de las siguientes proposiciones:

a) Se puede formar más de n columnas ordenadas. [0.5 ptos]

b) Se puede obtener una columna de altura mayor que n. [0.5 ptos]

c) Si se obtuvieron k columnas ordenadas, entonces se puede observar una subsecuencia estrictamente decreciente. [1.5 ptos]

d) Se debe obtener o más de n columnas o alguna columna que tenga una altura mayor que n. [1.5 ptos]

e) Siempre es posible encontrar una subsecuencia monótona de longitud $n+1$. [1 ptos]

Solución: Esta es una simple prueba geométrica del Teorema de Erdös y Szekeres, usando el principio del palomar. Todas las proposiciones son ciertas, describen el camino que se debe seguir para la demostración. Piense que necesariamente primero debemos llenar cada una de las n columnas con n números de la secuencia, para ello se utiliza los primeros n^2 números de dicha secuencia. Por último, podemos o crear una nueva columna o aumentar el tamaño de alguna preexistente.

4. Los coeficientes de la ecuación cuadrática $ax^2+bx+c=0$ se determina lanzando un dado tres veces (el primer resultado es a, el segundo b y el tercero c). Encuentra la probabilidad de que la ecuación no tenga raíces reales. [5 ptos]

Solución: Para que no tenga raíces reales debe cumplirse que $b^2-4ac<0$:

- $1^2<4ac$, cumplen todos los 36 pares.
- $2^2<4ac$ cumplen todos excepto $\{(1,1)\}$. 35 pares.
- $3^2<4ac$ cumplen todos excepto $\{(1,1),(1,2),(2,1)\}$. 33 pares.
- $4^2<4ac$ cumplen todos excepto $\{(1,1),(1,2),(1,3),(1,4),(2,1),$ $(2,2),(3,1),(4,1)\}$. 28 pares.
- $5^2<4ac$ cumplen todos excepto $\{(1,1),(1,2),(1,3),(1,4),(1,5),$ $(1,6),(2,1),(2,2)(2,3)(3,1),(3,2),(4,1),(5,1),(6,1)$. 22 pares.
- $6^2<4ac$ cumplen todos excepto $\{(1,1),(1,2),(1,3),(1,4),(1,5),$ $(1,6),(2,1),(2,2),(2,3),(2,4),(3,1),(3,2),(3,3),(4,1),(4,2),(5,1),(6,1)\}$. 19 pares.

Contamos: 173 pares, que serán 173 tripletes del total que son 6^3.

$$P(\text{Ecuación no tenga raíces reales})=\frac{173}{216}.$$

Universidad Nacional de Ingeniería
Escuela Profesional de Matemática
Ciclo 2020-1

[Cálculo de Probabilidades - CM1H2]
[Los profesores] UNI, 28 de setiembre de 2020

Examen Sustitutorio

1. (5 pts) Cierto tipo de motor eléctrico falla por obstrucción de los cojinetes, por combustión del embobinado o por desgaste de las escobillas. Supóngase que la probabilidad de la obstrucción es el doble de la combustión, la cual es 4 veces más probable que la inutilización de las escobillas. ¿Cuál es la probabilidad de fallo de cada uno de esos 3 mecanismos? (Sug. Defina eventos mutuamente excluyentes)

Solución: Definimos los sucesos:
$A = \{$ el motor falla por obstrucción de los cojinetes$\}$
$B = \{$ el motor falla por combustión del embobinado$\}$
$C = \{$ el motor falla por desgaste de las escobillas$\}$. Si $\mathbb{P}(C) = x$, entonces $\mathbb{P}(B) = 4x$ y $\mathbb{P}(A) = 8x$.
Suponiendo que los sucesos A, B y C son mutuamente excluyentes, y $\Omega = A \cup B \cup C$.

$$\begin{aligned} \mathbb{P}(\Omega) &= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) \\ &= 8x + 4x + x = 1, \end{aligned}$$

luego $x = 1/13$ y $\mathbb{P}(A) = 8/13$, $\mathbb{P}(B) = 4/13$ y $\mathbb{P}(C) = 1/13$.

2. (5 pts) Se dice que la variable aleatoria X es una variable aleatoria uniforme discreta en los enteros $1, 2, \ldots, n$ si $\mathbb{P}(X = i) = 1/n$ para $i = 1, 2, \ldots, n$. Para cualquier número real no negativo x, sea $\lfloor x \rfloor$ el entero más grande que es menor o igual a x. Si U es una variable uniforme $]0, 1[$, muestra que $X = \lfloor nU \rfloor + 1$ es una variable aleatoria uniforme discreta en $1, \ldots, n$.

Solución: Hallamos el rango de X, para esto considere para $0 < u < 1$:

$$\begin{aligned} \frac{i}{n} \leq u < \frac{i+1}{n}, \quad i &= 0, 1, \cdots, n-1 \\ \Rightarrow i \leq nu &< i+1 \\ \Rightarrow \lfloor nU \rfloor &= i. \end{aligned}$$

Entonces el rango de X es $R_X = \{1, 2, \cdots, n\}$. Y su función de masa es

$$\begin{aligned} \mathbb{P}[X = x] &= \mathbb{P}[\lfloor nU \rfloor + 1 = x] \\ &= \mathbb{P}[x - 1 \leq nU < x] \\ &= \mathbb{P}\left[\frac{x-1}{n} \leq U < \frac{x}{n}\right] \\ &= \frac{x}{n} - \frac{x-1}{n} = \frac{1}{n}. \end{aligned}$$

3. (5 pts) Suponga que el tiempo de viaje desde su hogar a su oficina está normalmente distribuido con media de 40 minutos y desviación estándar de 7 minutos. Si desea estar 95 % seguro de que no llegará tarde a una cita en su oficina a la 1 p.m., ¿cuál es lo más tarde que puede salir de su casa?

Solución: Sea X: "tiempo de viaje de mi casa a la oficina". Esta es por dato $X \sim N(\mu = 40, \sigma^2 = 49)$.
En promedio se demora 40 min así que hallaremos el 95 % de la distribución normal alrededor del valor promedio. Esto es si $\mathbb{P}[-z \leq Z \leq z] = 1 - 2\alpha = 0{,}95$, donde $\mathbb{P}[Z \leq z] = 1 - \alpha$. Es fácil ver que $1 - \alpha = 0{,}975$. Luego por tablas se tiene que $z = 1{,}96$. Entonces, los valores correspondientes a X son 26,28 y 53,72. Por tanto, hay un 0,95 de probabilidad de tardarse entre 26,28 min y 53,72 min. Entonces para asegurarme llegar a la hora debería salir a 1pm menos 53,72min, es decir aproximadamente $12:06$ pm.
También puede ser válido en cierto sentido, considerar: $\mathbb{P}[Z \leq z] = 0{,}95$.
De las tablas, se tiene $\mathbb{P}[Z \leq 1{,}64] = 0{,}9495$ y $\mathbb{P}[Z \leq 1{,}65] = 0{,}9505$, se podria aproximar tomando la media de estos valores, es decir 1,645, por lo tanto el valor de x correspondientes será $x = 1{,}645 \times 7 + 40 = 51{,}515$. Por lo tanto debería salir a más tardar aproximadamente $12:08$ pm.

4. (5 pts) Se sabe que un contenedor de 5 transistores contiene 2 que son defectuosos. Los transistores deben ser probados, uno a la vez, hasta que se identifiquen los defectuosos. Denote por N_1 el número de pruebas realizadas hasta que se detecta el primer defectuoso y por N_2 el número de pruebas adicionales hasta que se detecta el segundo defectuoso. Encuentre la función de masa de probabilidad conjunta de N_1 y N_2.

Solución: Sean
N_1:"N° de pruebas hasta que se detecta el 1er defectuoso"
N_2:"N° de pruebas adicionales hasta que se detecta el 2do defectuoso"
Es fácil ver que la función de masa conjunta es

$$p_{N_1,N_2}(x,y) = p_{N_2|N_1}(y|x) \times p_{N_1}(x),$$

y que N_1 y $[N_2|N_1 = x]$ son variables aleatorias del mismo tipo de distribución. Entonces, hallamos la función de masa de N_1:

$$\mathbb{P}[N_1 = x] = \frac{V_{x-1}^m}{V_{x-1}^n} \times \frac{n-m}{n-(x-1)} \quad , \quad x = 1, 2, \cdots, (n-1)\,, \tag{1}$$

donde $n = 5$ es el número total de transmisores y $m = 3$ es el número de transmisores no defectuosos.

Puede hacerse una idea hallando la masa para cada x y dándose cuenta que a medida que hace las pruebas, las probabilidades de ser defectuoso de los transmisores que quedan van cambiando.

Observación 1 *Note que el primer factor de* (1) *representa la probabilidad de que en las primeras $x-1$ pruebas no hubo transmisores defectuosos y el segundo factor representa la probabilidad de que en la x-ésima prueba se obtiene el defectuoso. No se considera la prueba n-ésima porque en la $(n-1)$-ésima prueba se obtendrá el 1er defectuoso con seguridad.*

Ahora, hallamos la función de masa de $[N_2|N_1 = x]$, con $x = 1, 2, 3, 4$, usando (1) para $n = 5 - x$ y $m = 3$:

$$p_{N_2|N_1}(y|x) = \frac{V_{y-1}^3}{V_{y-1}^{5-x}} \times \frac{(5-x)-3}{(5-x)-(y-1)} \quad , \quad y = 1, 2, \cdots, (4-x)\,. \tag{2}$$

La función de masa conjunta se deduce de multiplicar (1) y (2).

Universidad Nacional de Ingeniería
Escuela Profesional de Matemática
Ciclo 2020-1

[Cálculo de Probabilidades - CM1H2]
[Los profesores] UNI, 27 de julio de 2020

Examen Parcial

1. Una urna contiene igual número de bolas negras, blancas, rojas y verdes. Cuatro bolas son extraídas independientemente, con reemplazamiento. Encontrar la probabilidad $p(k)$ que exactamente k colores aparecerán en la muestra, $k = 1, 2, 3, 4$. [6ptos]

Solución:

- Este se ajusta a un problema multinomial. Es decir $n = 4$, n_1, n_2, n_3 n_4 el números de bolas de un tipo de color. Y la probabilidad es dada por la fórmula multinomial:

$$p(n_1, n_2, n_3, n_4) = \frac{n!}{n_1!n_2!n_3!n_4!} p_1^{n_1} p_2^{n_2} p_3^{n_3} p_4^{n_4},$$

  donde p_i, $i = 1, 2, 3, 4$ son las probabilidades de cada tipo.

- El espacio muestral consiste de todas las muestras ordenadas de tamaño 4, con reemplazamiento, de los siguientes tipos:
  - N: Negras,
  - B: Blancas,
  - R: Rojas,
  - V Verdes.

  Se considera que cada suceso ordenado es equiprobable, por ejemplo no tendrán la misma probabilidad los eventos "4 bolas blancas" y "tres blancos y una negra".

- $k = 4$: La probabilidad que todas sean de un mismo color, es dada por

$$\frac{4!}{1!1!1!1!} \left(\frac{1}{4}\right)^4 = \frac{6}{64} = p(4).$$

- $k = 3$: La probabilidad de obtener 2 negras, 1 blanca y 1 roja es traducido a tomar $n_1 = 2$, $n_2 = n_3 = 1$, $n_4 = 0$; esto es,

$$\frac{4!}{2!1!1!0!} \left(\frac{1}{4}\right)^4 = \frac{3}{64}.$$

  Los sucesos de este tipo son equiprobables, solo debemos encontrar el número de estos.

- Para encontrar la probabilidad total de obtener exactamente tres colores, multiplique por el número de formas de seleccionar tres colores de cuatro, $\binom{4}{3} = 4$, y el número de formas de seleccionar uno de los tres colores que se repetirán $\binom{3}{1} = 3$. Así, obtenemos

$$p(3) = \frac{36}{64}.$$

- $k=2$: La probabilidad de obtener por ejemplo, 2 negras y 2 blancas es

$$\frac{4!}{2!2!0!0!}\left(\frac{1}{4}\right)^4=\frac{3}{128}.$$

Luego, el número de obtener 2 bolas de un color y 2 de otro es $2/128$ veces el número de formas de seleccionar 2 colores de 4, $\binom{4}{2}=6$, es decir $\frac{9}{64}$. La otra posibilidad es obtener 3 de un color y 1 de otro color, esto es

$$\frac{4!}{3!1!0!0!}\left(\frac{1}{4}\right)^4(4)(3)=\frac{12}{64}.$$

Aquí si importa el orden de cómo elegir los dos colores, por ejemplo no es lo mismo elegir 3 negras y 1 blanca que 1 negra y 3 blancas. Por eso multiplicamos por el factor (4)(3) en vez de $\binom{4}{2}$. Por tanto, sumamos ambos casos excluyentes:

$$p(2)=\frac{9}{64}+\frac{12}{64}=\frac{21}{64}.$$

- $k=1$: La probabilidad de que todas las bolas sean del mismo color es

$$p(1)=\frac{4!}{4!0!0!0!}\left(\frac{1}{4}\right)^4(4)=\frac{1}{64},$$

4 representa las 4 formas de elegir un color de los 4 disponibles.

2. En cierta industria química, se sabe que el 5 % de todos los trabajadores están expuestos a un nivel de concentración diaria alta (H) de un potencial carcinógeno potencial (es decir, son miembros del Grupo H), que el 15 % de todos los trabajadores están expuestos a un nivel de concentración diario intermedio (I) (es decir, son miembros del Grupo I), que el 20 % de todos los trabajadores están expuestos a un nivel de concentración diario bajo (L) (es decir, son miembros del Grupo L) y que los restantes 60 % de todos los trabajadores no están expuestos (U) a este potencial carcinógeno (es decir, son miembros del Grupo U). Supongamos que cuatro trabajadores son elegidos al azar de una gran población de trabajadores de la industria química.

   a) ¿Cuál es la probabilidad de que los cuatro trabajadores elegidos al azar sean miembros del mismo grupo?. [1pt]

   b) Dado que los cuatro trabajadores elegidos al azar están expuestos a niveles no nulos de este carcinógeno potencial, ¿cuál es la probabilidad de que exactamente dos de estos cuatro trabajadores sean miembros del Grupo H?. [2ptos]

   c) Supongamos que C es el evento que un trabajador en esta industria qímica desarrolla cáncer. Sea $\pi_H = P(C|H) = 0,002$, la probabilidad condicional de que un trabajador del Grupo H desarrolle cáncer. De manera similar, sea $\pi_I = P(C|I) = 0,001$, $\pi_L = P(C|L) = 0,0001$, y $\pi_U = P(C|U) = 0,00001$. Si un trabajador en esta industria química desarrolla cáncer, ¿cuál es la probabilidad de que este trabajador sea miembro del Grupo H o del Grupo I? [4ptos]

Solución:

a) $\mathbb{P}$((los cuatro trabajadores elegidos al azar son miembros del mismo grupo)) =
$\mathbb{P}$(los cuatro trabajadores en el Grupo H) + $\mathbb{P}$(los cuatro trabajadores en el Grupo I)
+ $\mathbb{P}$(los cuatro trabajadores en el Grupo L) + $\mathbb{P}$(los cuatro trabajadores en el Grupo U)
$= (0{,}05)^4 + (0{,}15)^4 + (0{,}20)^4 + (0{,}60)^4 = 0{,}132.$

b) Sea A el evento que 'dos de los cuatro trabajadores son miembros del Grupo H', y que B sea el evento de que 'los cuatro trabajadores están expuestos a niveles distintos de cero del potencial carcinógeno'. Entonces,

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{C_2^4 (0{,}05)^2 (0{,}35)^2}{(0{,}40)^4} = 0{,}0718.$$

El numerador se obtiene como el producto de formas de dividir la muestra de 4 en dos subconjuntos uno correspondiente al grupo H, por lo que se obtiene $0{,}05^2$ y el otro grupo que puede ser de los tipos I o L los cuales cuentan con una probabilidad de $0{,}15+0{,}20 = 0{,}35$, por lo cual se obtiene $0{,}35^2$. El denominador corresponde a la probabilidad de estar por lo menos expuesto a una intensidad de carcinógeno es decir: $0{,}05 + 0{,}15 + 0{,}20 = 0{,}40$. Por tanto el valor de $0{,}40^4$.

c) Sea B el caso de que un trabajador sea miembro del Grupo H o del Grupo I. Por lo tanto,

$$\begin{aligned}\mathbb{P}(B|C) &= \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}[(H \cup I) \cap C]}{\mathbb{P}(C)} = \frac{\mathbb{P}(H \cap C) + \mathbb{P}(I \cap C)}{\mathbb{P}(C)} \\ &= \frac{\mathbb{P}(C|H)\mathbb{P}(H) + \mathbb{P}(C|I)\mathbb{P}(I)}{\mathbb{P}(C)} = \frac{(0{,}002)(0{,}05) + (0{,}001)(0{,}15)}{\mathbb{P}(C)} \\ &= \frac{0{,}000250}{\mathbb{P}(C)}.\end{aligned}$$

Desde que

$$\mathbb{P}(C) = \mathbb{P}(C|H)\mathbb{P}(H) + \mathbb{P}(C|I)\mathbb{P}(I) + \mathbb{P}(C|L)\mathbb{P}(L) + \mathbb{P}(C|U)\mathbb{P}(U)$$

$$= (0{,}002)(0{,}05) + (0{,}001)(0{,}15) + (0{,}0001)(0{,}20) + (0{,}00001)(0{,}60) = 0{,}000276.$$

Luego se sigue que $\mathbb{P}(B|C) = 0{,}000250/0{,}000276 = 0{,}906$.

3. Usted está rindiendo una prueba de opción múltiple para la cual ha estudiado el 70 % del curso. Suponga que esto significa que tiene una probabilidad de 0,7 de saber la respuesta a una pregunta de prueba aleatoria, y que si no sabe la respuesta a una pregunta, entonces selecciona aleatoriamente entre las cuatro opciones de respuesta. Finalmente, suponga que esto es válido para cada pregunta, independientemente de las demás.

a) ¿Cuál cree usted que será su puntaje (en porcentaje) a dicho examen? [3ptos]

Sea p la probabilidad de obtener una pregunta aleatoria correcta. Probablemente haya encontrado el valor de p en la parte (a), pero en cualquier caso debe suponer $0{,}7 < p < 0{,}9$. En las partes (b), (c) y (d) puede usar la letra p para esta probabilidad.

b) Si el examen tiene 10 preguntas, ¿cuál es la probabilidad de que obtenga un 90 % o más? (No simplifique la expresión que obtenga). [2ptos]

c) ¿Cuál es la probabilidad de que las primeras 6 preguntas en el examen sean correctas? (Nuevamente, no simplifique la expresión que obtenga). [1pt]

d) Supongamos que necesita un 90 % para mantener su beca de estudios. ¿Prefiere tener un examen con 10 o 100 preguntas? ¿Por qué? [1pt]

Solución:

a) Observamos que tenemos una probabilidad de 0,7 de encontrarnos con una pregunta que sabemos resolver, es decir podemos considerar responder correctamente con una probabilidad de 1, en caso contrario tenemos 0,3 de encontrarnos con una que no podemos resolver y por tanto una probabilidad de 0,25 de adivinar la respuesta correcta.

- Cada pregunta puede ser considerada como una v.a. de Bernoulli, X_j, cuya probabilida de éxito (responder correctamente) es

$$p = 0{,}7 + (0{,}25)(0{,}3) = 0{,}775,$$

donde $X_j \sim Be(p)$, $j = 1, \cdots, n$, con n preguntas.

- La nota que el alumno espera obtener sería un promedio de cada probabilidad obtenida, que se puede ver como la esperanza de la v.a.

$$X = \frac{X_1 + \cdots + X_n}{n}.$$

Por la linealidad de la esperanza, se obtiene que $\mathbb{E}(X) = p$. Es decir la nota que espera obtener es de 77,5 %.

b) Sea la v.a. Y : "Número de respuestas correctas". $Y \sim Bi(10, p)$. Luego,

$$\mathbb{P}(Y \geq 9) = \binom{10}{9} p^9(1-p) + p^{10} = \binom{10}{9}(0{,}775)^9(0{,}225) + (0{,}775)^{10}.$$

c) $p^6 = (0{,}775)^6$.

d) El alumno preferirá tener un examen con 10 preguntas, ya que cuantas más preguntas, más probabilidades tendrá de obtener el valor medio de 77,5 % que pronóstica la media después de hacer un ensayo muchas veces. Así que como sabe que es muy probable obtener 77,5 %, es posible que con menos preguntas, tenga la suerte de obtener un puntaje mayor.

Universidad Nacional de Ingeniería
Escuela Profesional de Matemática
Ciclo 2020-1

[Cálculo de Probabilidades - CM1H2]
[Los profesores] UNI, 14 de setiembre de 2020

Examen Final

Elija 3 preguntas

1. (5 pts) Cierta enfermedad tiene una probabilidad muy baja de ocurrir, $p = 10^{-5}$. Calcular la probabilidad de que en una ciudad con 500, 000 habitantes haya más de 3 personas con dicha enfermedad. Calcular el número esperado de habitantes que la padecen. Utilice la distribución de los eventos raros para aproximar los cálculos y justifique por qué es posible usar dicha distribución.

Solución: Sea X : número de personas que padecen la enfermedad. Esta distribución sigue un modelo binomial, es decir $X \sim Bi(5 \times 10^5, 10^{-5})$. Esta se puede aproximar por un modelo de Poisson debido a que $np = 5$ es pequeõ y será considerado como la frecuencia media de habitantes que pueden padecer dicha enfermedad. Es decir, aproximamos X a una variable aleatoria $\approx Poi(\lambda = 5)$. Recordemos su función de densidad:

$$\mathbb{P}[X = k] = e^{-\lambda}\frac{\lambda^k}{k!} \quad , \quad k = 0, 1, 2\cdots$$

Luego, el número esperado de enfermos es $\mathbb{E}(X) = 5$. Ahora hallemos la probabilidad de encontrar más de 3 enfermos:

$$\begin{aligned}\mathbb{P}[X > 3] &= 1 - \mathbb{P}[X \leq 3] \\ &= 1 - e^{-5}\left(1 + 5 + \frac{25}{2} + \frac{125}{6}\right) \\ &= 1 - 0{,}265 = 0{,}735\,.\end{aligned}$$

Es decir es muy probable de encontrar más de 3 personas enfermas.

2. (5 pts) Se lanza una moneda 50 veces. En caso sea posible (justifique!), utilice el teorema del límite central (TLC) para estimar la probabilidad de que menos de 20 de esos lanzamientos salgan cara. (Sug. Utilice el ajuste por continuidad de 0,5 y dado $Z \sim N(0, 1)$, se tiene que $f_Z(1{,}56) = 0{,}9406$.)

Solución:

- El número de lanzamientos que salen cara es una variable aleatoria X binomial, que se puede escribir como una suma de $n = 50$ variables aleatorias independientes de Bernoulli con probabilidad de éxito de obtener cara igual a $p = 1/2$.
- Dado que 50 es un número razonablemente grande, tiene sentido usar TLC y aproximar X (el número de caras en 50 lanzamientos) por una variable aleatoria normal de media $np = 50 \cdot \frac{1}{2} = 25$ y varianza $np(1-p) = 50 \cdot \frac{1}{2} \cdot \frac{1}{2} = 12{,}5$.
- Por continuidad, se tiene

$$\mathbb{P}(X < 20) = \mathbb{P}(X \leq 19) = \mathbb{P}(X \leq 19{,}5),$$

luego, utilizando la estandarización $Z = \frac{X-\mu_X}{\sigma_X}$, se tiene

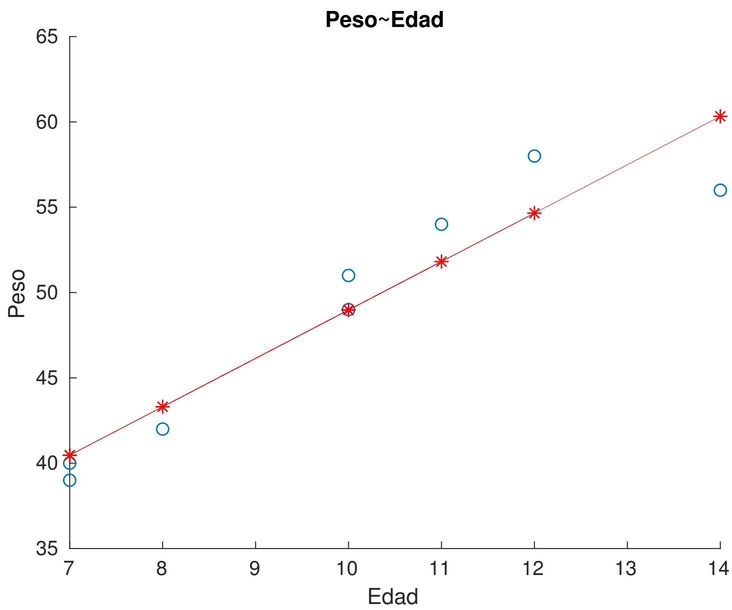
$$\begin{aligned}\mathbb{P}(X \leq 19{,}5) &= \mathbb{P}\big(Z \leq \frac{19{,}5-25}{\sqrt{12{,}5}}\big) \\ &= \mathbb{P}(Z \leq -1{,}56) \\ &= 1-\mathbb{P}(Z \leq 1{,}56) \\ &= 1-0{,}9406 = 0{,}0594\,.\end{aligned}$$

3. (8 pts) Se miden peso y edad en un grupo de ocho pacientes en un año específico, obteniéndose

X : *edad*	12	8	10	11	7	7	10	14
Y : *peso*	58	42	51	54	40	39	49	56

¿Existe una relación lineal importante entre ambas variables? Calcular la recta de regresión de la edad en función del peso y la del peso en función de la edad. Calcular la bondad del ajuste. ¿En qué medida, por término medio, varía el peso cada año? ¿En cuánto aumenta la edad por cada kilo de peso?.

Solución: El siguiente diagrama de dispersión nos da el indicio de que es razonable la regresión lineal.



a) Deseamos hallar la recta de regresión de $\bar{Y} = a_1 + b_1 X$ con respecto a X, considerando la tabla de abajo

$$\bar{x} = \frac{79}{8} = 9{,}875\,, \quad \bar{y} = \frac{389}{8} = 48{,}625\,, \quad b_1 = \frac{Cov(X,Y)}{Var(X)},$$

donde $Var(X) = \frac{\sum x^2}{n} - \bar{x}^2 = \frac{823}{8} - \left(\frac{79}{8}\right)^2 = \frac{343}{64} = 5{,}3594$ y

$$Cov(X,Y) = \frac{\sum xy}{n} - \bar{x}\bar{y} = \frac{3963}{8} - \frac{79}{8} \times \frac{389}{8} = \frac{973}{64} = 15{,}2031.$$

Entonces, $b_1 = \frac{973}{343} \approx 2{,}8367$, además $a_1 = \bar{y} - b_1\bar{x} \approx 20{,}6122$ y por lo tanto, la recta de regresión es

$$y = 20{,}6122 + 2{,}8367x\,.$$

b) Hallamos el coeficiente de correlación y determinación:

$$\rho = Corr(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{973}{\sqrt{343 \times 3103}} \approx 0{,}9431,$$

donde $Var(Y) = \frac{\sum y^2}{n} - \bar{y}^2 = \frac{19303}{8} - \left(\frac{389}{8}\right)^2 = \frac{3103}{64} \approx 48{,}4844$. Y el coeficiente de terminación es $\rho^2 \approx 0{,}8895$.

c) La recta de regresión de X en función de Y, $\bar{X} = a_2 + b_2 Y$, es $\bar{X} = -5{,}3722 + 0{,}3136Y$.

d) Por tanto podemos decir que el $88,95\,\%$ de la variabilidad del peso en función de la edad es explicada mediante la recta de regresión correspondiente. Lo mismo podemos decir en cuanto a la variabilidad de la edad en función del peso. Del mismo modo puede decirse que hay un $11,05\,\%$ de varianza que no es explicada por las rectas de regresión.

e) Por último la cantidad en que varía el peso de un paciente cada año es, según la recta de regresión del peso en función de la edad, la pendiente de esta recta, es decir, $b_1 = 2,8367$ Kg/año. Y cuando dos personas difieren en peso, en promedio la diferencia de edades se rige por la cantidad $b_2 = 0,3136$ años/Kg de diferencia.

x	y	xy	x^2	y^2
12	58	696	144	3364
8	42	336	64	1764
10	51	510	100	2601
11	54	594	121	2916
7	40	280	49	1600
7	39	273	49	1521
10	49	490	100	2401
14	56	784	196	3136
79	389	3963	823	19303

4. (7 pts) Sean X, Y variables aleatorias continuas con función de densidad conjunta

$$f(x,y) = \begin{cases} 1/x & , \quad 0 < y < x < 1 \\ 0 & , \quad \text{en otro caso.} \end{cases}$$

a) (1 pt) Haga un bosquejo del dominio de f donde no se anula.

b) (1 pt) Calcular $\mathbb{E}(X)$ y $\mathbb{E}(Y)$.

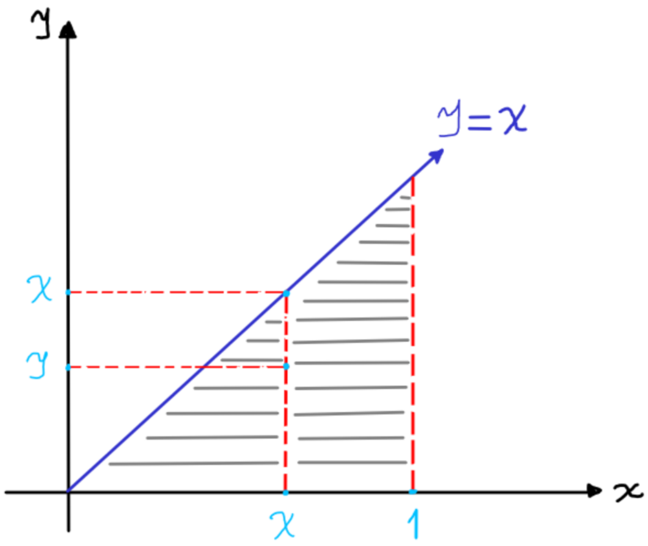
c) (2 pts) Calcular la función de densidad condicional de Y dado $X = x$, para todo $0 < x < 1$.

d) (1 pt) Calcular $\mathbb{E}[Y|X = x]$ para todo $0 < x < 1$.

e) (2 pts) Calcular $Cov(X,Y)$.

Solución:

a) El dominio donde no se anula f es el de la región sombreada:



b) $\mathbb{E}(X) = \iint_{\mathbb{R}^2} x f(x,y)\, d(x,y) = \int_0^1 \int_0^x x \cdot \frac{1}{x}\, dy\, dx = \int_0^1 x\, dx = \frac{1}{2}$,
$\mathbb{E}(Y) = \iint_{\mathbb{R}^2} y f(x,y)\, d(x,y) = \int_0^1 \int_0^x y \cdot \frac{1}{x}\, dy\, dx = \int_0^1 \frac{x}{2}\, dx = \frac{1}{4}$.

c) La función de densidad marginal para X se calcula como $f_X(x) = \int_{\mathbb{R}} f(x,y)\, dy$. Entonces, si $0 < x < 1$, se tiene

$$f(x) = \cancel{\int_{-\infty}^{0} f(x,y)\, dy} + \int_0^x f(x,y)\, dy + \cancel{\int_0^{\infty} f(x,y)\, dy}$$
$$= \int_0^x \frac{1}{x}\, dy = 1\,,$$

y es cero en otro caso, es decir, X es una distribución uniforme en $(0,1)$

$$f_X(x) = \begin{cases} 1 & , \quad 0 < x < 1 \\ 0 & , \quad \text{en otro caso.} \end{cases}$$

Por tanto la función de densidad condicional es

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{1}{x}\,, \quad 0 < y < x,$$

y 0 en otro caso. Es decir, para x fijo la variable aleatoria $[Y|X = x]$ posee una distribución uniforme en $(0, x)$.

d) Como $[Y|X = x]$ es uniforme, entonce podemos según la teoría saber que su valor esperado se encuentra en el punto medio del intervalo, es decir su valor medio es $x/2$, o podemos usar la definición de esperanza condicional:

$$\mathbb{E}[Y|X = x] = \int_{\mathbb{R}} y \cdot f_{Y|X}(y|x)\, dy = \int_0^x \frac{y}{x}\, dy = \frac{x}{2}\,.$$

e) Primero hallamos la esperanza del producto

$$\mathbb{E}(XY) = \int_0^1 \int_0^x xy \cdot \frac{1}{x}\, dy\, dx = \int_0^1 \frac{x^2}{2}\, dx = \frac{1}{6},$$

Asi, por la siguiente identidad, se tiene

$$Cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \frac{1}{6} - \frac{1}{2} \times \frac{1}{4} = \frac{1}{24}\,.$$