Examen Febrero
Miércoles 31 de enero de 2018.

Número de Examen	Cédula	Nombre y Apellido

Para uso docente:

Ej. 1	Ej. 2	Ej.3	Ej.4	Ej.5	Ej.6/7	TOTAL

El examen es con material y dura 4 horas. Se deberá entregar la resolución de un máximo de 6 ejercicios.

Ejercicio 1 (10 puntos)

Se lanzan dos dados justos de 6 caras y se definen los siguientes sucesos:

- $A = \{\text{la suma de los dos dados es } 3\}$,
- $B = \{\text{la suma de los dos dados es } 7\}$,
- $C = \{\text{en al menos uno de los dos dados se obtuvo 1 como resultado}\}$.

1. Hallar $P(A|C)$ y $P(B|C)$.
2. ¿A y C son independientes? ¿B y C son independientes? Justifique su respuesta.

Ejercicio 2 (10 puntos)

Sea X el resultado de lanzar un dado equilibrado de 4 caras (numeradas del 1 al 4). Sea Y el resultado de lanzar un dado equilibrado de 6 caras (numeradas del 1 al 6). Se juega un juego en el que usted gana $2X$ miles de pesos si $X > Y$ y pierde 1000 pesos en otro caso. Después de jugar el juego 60 veces indique cuál es la ganancia esperada del mismo.

Ejercicio 3 (15 puntos)

Considere un "círculo aleatorio" centrado en el origen tal que el radio R de dicho círculo distribuye como una variable aleatoria uniforme en el intervalo $[0, 1]$.

1. Calcular el valor esperado del área del círculo.
2. Hallar la función de distribución de la variable aleatoria A, siendo A el área de dicho círculo.

Ejercicio 4 (20 puntos)

Se mide la tasa de éxito de un tratamiento médico en 12 pacientes en cada uno de los 60 centros que participan de la medición. Sea p la probabilidad de que el tratamiento tenga éxito en un paciente. Se asume que los resultados del tratamiento para los diferentes pacientes son independientes y que la probabilidad p es igual para todos los centros.

1. Indicar la distribución del número de éxitos por centro así como su valor esperado.
2. Se tiene la siguiente muestra sobre el número de éxitos:

Número de Éxitos	0	1	2	3	4	5
Número de Centros	4	15	17	10	8	6

De la tabla se tiene que por ejemplo en 15 centros hubo un solo paciente exitoso.

(a) Dado $X_1, X_2, \dots X_n$ una muestra independiente e idénticamente distribuida del número de pacientes exitosos, hallar el estimador de máxima verosimilitud de p. Calcularlo para los datos de la muestra.

(b) Hallar un intervalo de confianza aproximado para p al nivel 0.95.

Ejercicio 5 (25 puntos)

Sea $X_1, \ldots, X_n$ una sucesión de variables aleatorias independientes con distribución Bernoulli de parámetro p. Sea $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ el promedio de dichas variables. Dado $t > 0$ se define el intervalo

$$I_t = \left[\bar{X}_n - t\frac{\sqrt{p(1-p)}}{\sqrt{n}}, \bar{X}_n + t\frac{\sqrt{p(1-p)}}{\sqrt{n}}\right].$$

1. (a) Probar que $P(p \in I_t) \underset{n\to\infty}{\to} 2\Phi(t) - 1$, donde Φ es la función de distribución de una variable normal estándar.
   (b) Hallar t tal que $P(p \in I_t) \underset{n\to\infty}{\to} 0.98$.
2. Se define el intervalo $I = \left[\bar{X}_n - z_{\alpha/2}\frac{0.5}{\sqrt{n}}, \bar{X}_n + z_{\alpha/2}\frac{0.5}{\sqrt{n}}\right]$. Probar que $\mathrm{Var}(\bar{X}_n) \leq \frac{1}{4n}$ para todo valor de p y deducir que I es un intervalo de confianza aproximado conservativo al nivel α para p, esto es $\lim_{n\to\infty} P(p \in I) \geq 1 - \alpha$. Recordar que $P(Z \leq z_\alpha) = 1 - \alpha$ siendo $Z \sim \mathcal{N}(0, 1)$.
3. En una ciudad se elige alcalde entre dos candidatos: Juan y Juana. En una encuesta a 1000 personas se obtiene que 450 personas votan por Juana, 400 votan por Juan y 150 están indecisas. Hallar un intervalo de confianza aproximado al nivel 0.98 para la proporción de votantes de Juana y compararlo con el intervalo conservativo definido en la parte 2.

Deberán realizar uno y solo uno de los ejercicios que se enuncian a continuación.

Ejercicio 6 (20 puntos)

Sea X una variable aleatoria con densidad $f(x) = \frac{\lambda}{2}e^{-\lambda|x|}$ para $x \in \mathbb{R}$ y $\lambda > 0$.

1. Hallar la función generatriz de momentos de X. Indicar su dominio de definición.
2. Usando la parte anterior, hallar valor esperado y varianza de X.
3. Considere una sucesión de variables aleatorias independientes $X_1, X_2, \ldots, X_n$ todas con la distribución de X para $\lambda = 1$. Usando la desigualdad de Tchebychev (o Markov), determinar el mínimo valor de n tal que $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ diste del origen en más de $0,01$ con una probabilidad de a lo sumo $0,1$.

Ejercicio 7 (20 puntos)

Sea X una variable aleatoria con la siguiente densidad:

$$f(x) = \begin{cases} 2x/b^2 & \text{si } x \in [0, b], \\ 0 & \text{en otro caso,} \end{cases}$$

donde $b > 0$.

1. Hallar el estimador por momentos de b.
2. Hallar el estimador por máxima verosimilitud de b y hallar la función de distribución de dicho estimador.
3. Calcular el sesgo de los estimadores de las partes anteriores.

Solución

Ejercicio 1 Consideramos como modelo probabilístico para la situación el espacio muestral Ω que consiste en los pares ordenados de números enteros entre 1 y 6 (Ω tiene 36 elementos) con la medida de probabilidad $\mathbb{P}$ para la cual cada evento elemental tiene probabilidad $\frac{1}{36}$.
En este modelo se obtiene:

- $A = \{(1,2),(2,1)\}$
- $B = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$
- $C = \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),(3,1),(4,1),(5,1),(6,1)\}$

1. Se calcula $\mathbb{P}(A|C) = \frac{\mathbb{P}(A\cap C)}{\mathbb{P}(C)} = \frac{2/36}{11/36} = \frac{2}{11}$ y $\mathbb{P}(B|C) = \frac{\mathbb{P}(B\cap C)}{P(C)} = \frac{2/36}{11/36} = \frac{2}{11}$.
2. Calculamos $\mathbb{P}(A) = \frac{2}{36} \neq \frac{2}{11} = \mathbb{P}(A|C)$ por lo cual A y C no son independientes. Del mismo modo $\mathbb{P}(B) = 6/36 \neq 2/11 = \mathbb{P}(B|C)$ por lo cual B y C no son independientes.

Ejercicio 2 Sean $Z_1,\dots,Z_{60}$ variables aleatorias i.i.d. tales que Z_i modela la cantidad ganada en la i-ésima ronda del juego. Se obtiene que:

- $\mathbb{P}(Z_i = 4000) = \mathbb{P}(X=2, Y=1) = 1/24$
- $\mathbb{P}(Z_i = 6000) = \mathbb{P}(X=3, Y=1) + \mathbb{P}(X=3, Y=2) = 2/24$
- $\mathbb{P}(Z_i = 8000) = \mathbb{P}(X=4, Y=1) + \mathbb{P}(X=4, Y=2) + \mathbb{P}(X=4, Y=3) = 3/24$
- $\mathbb{P}(Z_i = -1000) = 18/24$.

Usando esto calculamos la ganancia esperada luego de 60 rondas como sigue:

$$\begin{aligned}\mathbb{E}(Z_1+\cdots+Z_{60}) = 60\mathbb{E}(Z_1) &= 60\left(4000\cdot\frac{1}{24} + 6000\cdot\frac{2}{24} + 8000\cdot\frac{3}{24} - 1000\cdot\frac{18}{24}\right)\\ &= 60\cdot\frac{4000+12000+24000-18000}{24} = 55000.\end{aligned}$$

Ejercicio 3
1. Sea $A = \pi R^2$ el área del círculo, entonces $\mathbb{E}(A) = \pi\int_0^1 x^2dx = \frac{\pi}{3}$.
2. $F_A(a) = \mathbb{P}(A \leq a) = \mathbb{P}(R^2 \leq \frac{a}{\pi}) = \mathbb{P}(R \leq \sqrt{\frac{a}{\pi}})$. Entonces:

$$F_A(a) = \begin{cases} 0 \text{ si } a < 0;\\ \sqrt{\frac{a}{\pi}} \text{ si } 0 \leq a \leq \pi;\\ 1 \text{ si } a > \pi.\end{cases}$$

Ejercicio 4
1. El número de éxitos en un centro se puede modelar como una variable Binomial de parmetros 12 y p, entonces su valor esperado es $12p$.
2. a) Las variables $X_1,\dots,X_n$ son Binomiales $(12,p)$ i.i.d.. La verosimilitud de $X_1 = k_1,\dots,X_n = k_n$ se calcula como:

$$L(p) = \binom{12}{x_1}p^{x_1}(1-p)^{12-x_1}\cdots\binom{12}{x_n}p^{x_n}(1-p)^{12-x_n}.$$

A los efectos de encontrar el $p\in[0,1]$ que maximiza $V(p)$ calculamo:

$$\partial_p\log(L(p)) = (x_1+\cdots+x_n)\frac{1}{p} - (12n-(x_1+\cdots+x_n))\frac{1}{1-p}.$$

Igualando a 0 obtenemos que el estimador de máxima verosimilitud es:

$$\hat{p} = \frac{x_1+\cdots+x_n}{12n},$$

(el punto crítico debe ser máximo de $L(p)$ porque $L(p)$ vale 0 en los extremos del intervalo $[0,1]$).
En la muestra obtenemos $\hat{p} = 141/720 \approx 0.1958$.

b) Utilizando el Teorema Central del Límite, tenemos que un intervalo de confianza aproximado a nivel $1-\alpha$ para $\mu = \mathbb{E}(X) = 12p$ es:

$$I_{1-\alpha}(\mu) = [\bar{x}_n \pm k] \text{ con } k = \frac{s_n}{\sqrt{n}}z_{\frac{\alpha}{2}}.$$

En este caso: $\bar{x}_n = 2.35$, $s_n = 1.41$, $z_{\frac{\alpha}{2}} = 1.96$, entonces $k = 0.357$ y

$$I_{0.95}(\mu) = [1.993,\ 2.707] = [A,B]$$

Como $\mu = 12p$ y $\mathbb{P}(A \leq \mu \leq B) \approx 0.95$, conseguimos un intervalo de confianza aproximado para p así:

$$0.95 \approx \mathbb{P}(A \leq \mu \leq B) = \mathbb{P}\left(\frac{A}{12} \leq p \leq \frac{B}{12}\right).$$

Entonces:

$$I_{0.95}(p) \approx \left[\frac{A}{12}, \frac{B}{12}\right] = [0.166,\, 0.226]$$

Ejercicio 5

1. a) Por el Teorema Central del Límite:

$$\mathbb{P}(p \in I_t) = \mathbb{P}\left(-t \leq \frac{\sqrt{n}}{\sqrt{p(1-p)}}(\bar{X}_n - p) \leq t\right) \to \Phi(t) - \Phi(-t) = 2\Phi(t) - 1,$$

   b) $2\Phi(t) - 1 = 0.98 \leftrightarrow \Phi(t) = 0.99 \leftrightarrow t \approx 2.33$.

2. Se calcula $\mathrm{Var}(\overline{X}_n) = \frac{1}{n^2}\mathrm{Var}(X_1 + \cdots + X_n) = \frac{1}{n}\mathrm{Var}(X_1) = p(1-p)/n$. El máximo valor posible para $p \in [0,1]$ se da en el punto crítico $p = 1/2$ (la función en cuestión es un polinomio de grado 2 con concavidad negativa) y el valor es $1/4$. Por lo tanto $\mathrm{Var}(\overline{X}_n) \leq \frac{1}{4n}$ como se pedía demostrar.

   Si en la parte anterior tomamos $t = z_\alpha$, entonces:

$$1 - \alpha \leftarrow \mathbb{P}(p \in I_{z_\alpha}) = \mathbb{P}\left(|\bar{X}_n - p| \leq \frac{z_\alpha\sqrt{p(1-p)}}{\sqrt{n}}\right) \leq \mathbb{P}\left(|\bar{X}_n - p| \leq \frac{z_\alpha 0.5}{\sqrt{n}}\right)$$

3. Sea $X_i = \begin{cases} 1 \text{ si el votante i vota por Juana} \\ 0 \text{ si no} \end{cases}$. Se tiene que $n = 1000$ y $\bar{x}_n = \frac{450}{1000}$.

   Utilizando la primera parte,

$$I_{0.98}(p) \approx \left[\bar{x_n} \pm 2.33\frac{\sqrt{\bar{x}_n(1-\bar{x}_n)}}{\sqrt{n}}\right] \approx [0.413,\, 0.487]$$

   Utilizando el intervalo conservativo,

$$I_c(p) = \left[\bar{x}_n \pm \frac{0.5z_{0.02}}{\sqrt{n}}\right] \approx [0.417,\, 0.483]$$

   Observación: También se considerará como correcto cuando no se tomen a los indecisos como parte de la muestra, es decir, si se toma $n = 450 + 400$ y $\bar{x}_n = \frac{450}{850}$ (con los intervalos de confianza correspondientes).

Ejercicio 6

1. La función generatriz se calcula como sigue

$$\begin{aligned} f(t) = \mathbb{E}(e^{tX}) &= \int_{-\infty}^{+\infty} e^{tx}\frac{\lambda}{2}e^{-\lambda|x|}dx = \frac{\lambda}{2}\int_{-\infty}^{+\infty} e^{tx-\lambda|x|}dx \\ &= \frac{\lambda}{2}\int_{-\infty}^{0} e^{(t+\lambda)x}dx + \frac{\lambda}{2}\int_{0}^{+\infty} e^{(t-\lambda)x}dx \\ &= \frac{\lambda}{2}\frac{1}{\lambda+t} + \frac{\lambda}{2}\frac{1}{\lambda-t} \\ &= \frac{\lambda}{2}\frac{1}{\lambda^2-t^2}. \end{aligned}$$

   Está definida para $t \in (-\lambda, \lambda)$ (de otro modo las integrales impropias involucradas no convergen).

2. Calculamos

$$f'(t) = -\frac{\lambda}{2}\frac{1}{(\lambda+t)^2} + \frac{\lambda}{2}\frac{1}{(\lambda-t)^2},$$

$$f''(t) = \lambda\frac{1}{(\lambda+t)^3} + \lambda\frac{1}{(\lambda-t)^3}.$$

   Se obtiene $\mathbb{E}(X) = f'(0) = 0$ y $\mathbb{E}(X^2) = f''(0) = \frac{2}{\lambda^2}$ por lo cual $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{2}{\lambda^2}$.

3. De la parte anterior se deduce que $\mathbb{E}(\overline{X}_n) = 0$ y $\sigma_n^2 = \mathrm{Var}(\overline{X}_n) = 2/n = \sigma_n$. La desigualdad de Tchebychev dice que para todo $t \geq 0$ y n tenemos:

$$\mathbb{P}(|\overline{X}_n - \mathbb{E}(\overline{X}_n)| \geq t\sigma_n) \leq \frac{1}{t^2}.$$

Tomando $t = \sqrt{10}$ obtenemos:

$$\mathbb{P}(|\overline{X}_n| \geq \sqrt{20/n}) \leq 0.1.$$

Por lo tanto alcanza con tomar $n \geq 200000$ para que

$$\mathbb{P}(|\overline{X}_n| \geq 0.01) \leq 0.1.$$

Ejercicio 7 En lo que sigue suponemos que $X, X_1, \ldots, X_n$ son variables i.i.d. con densidad f y $\overline{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$.

1. Se calcula $\mathbb{E}(X) = \int_0^b 2x^2/b^2 dx = 2b/3$ por lo cual un estimador por momentos del parámetro b es $\bar{b} = 3\overline{X}_n/2$.
2. La verosimilitud de $X_1 = x_1, \ldots, X_n = x_n$ es:

$$L(b) = \begin{cases} 2^n x_1 \cdots x_n b^{-2n} \text{ si } b \geq \max\{x_1, \ldots, x_n\} \\ 0 \text{ en otro caso.} \end{cases}$$

Como $L(b)$ decrece con b si $b \geq \max\{x_1, \ldots, x_n\}$ el estimador de máxima verosimilitud es $\hat{b} = \max(x_1, \ldots, x_n)$. Llamemos $F(t)$ a la función de distribución de $\hat{b}$. Se obtiene $F(t) = 0$ si $t \leq 0$, $F(t) = 1$ si $t \geq b$ y para $t \in [0, b]$

$$F(t) = \mathbb{P}(X_1 \leq t, \ldots, X_n \leq t) = \mathbb{P}(X_1 \leq t)^n = \left(\int_0^t 2x/b^2 dx\right)^n = \frac{t^2 n}{b^{2n}}.$$

3. Calculamos $\mathbb{E}(\bar{b}) = 3\mathbb{E}(\overline{X}_n)/2 = b$ por lo cual el estimador de momentos es insesgado.

Calculamos $\mathbb{E}(\hat{b} = \int_0^b 1 - \frac{t^2 n}{b^{2n}} dt = b - b/(2n+1)$ por lo cual el estimador de máxima verosimilitud es sesgado. El sesgo es:

$$sesgo(\hat{b}) = \mathbb{E}(\hat{b}) - b = -\frac{b}{2n+1}.$$