Segundo parcial Probabilidad y Estadstica, 28 de junio de 2021.
Letra y solución.

Ejercicio 1

La cantidad de usuarios que ingresan por minuto a un canal web es una V.A. $X \sim \text{Poisson}(\lambda)$, con $1 \leq \lambda \leq 2$ (definida en miles de usuarios). Se desea estimar λ a partir del promedio $\overline{X}_n$ de una muestra $X_1, X_2, ..., X_n$ i.i.d. de modo que con probabilidad de al menos 0.93 la distancia entre $\overline{X}_n$ y λ no supere los 100 usuarios. A partir de la desigualdad de Chevishoff, indicar el menor n natural que satisfaga la condición deseada.

Ejercicio 2

Si $1, 16, 9, 16, 25$ es una muestra iid (muestra aleatoria simple) de la variable X tal que $\text{Rec}(X) = \{0, 1, 4, 9, 16, 25, ...\}$ siendo $P(X = x) = \frac{e^{-\lambda}\lambda^{\sqrt{x}}}{\sqrt{x}!}$ siendo $\lambda > 0$.

Hallar la estimación máximo verosímil de λ.

Ejercicio 3

Dada la siguiente muestra iid

1.91, 2.8, 1.72, 2.52, 2.53, 2.68, 2.48, -1.28, 2.23, 2.6.

Calcular el estadístico de Kolmogorov y Smirnov para la hipótesis nula en la que los datos siguen una distribución normal de parámetros $\mu = 3$ y desvío $\sigma = 2$.

Ejercicio 4

Los datos

9.6, 7.4, 6.6, 13.1, 1.5, 4.6, 2.8, 6.8, 5.0 son una MAS (observaciones iid) correspondientes a una variable aleatoria X cuya función de densidad viene dada por

$$f_X(x) = \begin{cases} \frac{\beta}{x^{\beta+1}} & \text{si} \quad x \geq 1 \\ 0 & \text{si} \quad x < 1 \end{cases} \quad \text{siendo } \beta > 1.$$

Ejercicio 5

Se realiza un estudio de mercado sobre 2 productos A y B. Para ello se obtiene una muestra aleatoria simple de posibles consumidores de dichos productos. La misma arrojó los siguientes resultados: 27 personas comprarían solo el artículo A, 22 solo el artículo B, 8 ambos artículos y 20 ninguno. Se desea estimar el porcentaje de individuos que compraría los dos artículos mediante un intervalo de confianza aproximado de la forma $\left[\overline{X}_n - k, \overline{X}_n + k\right]$ al 90 %.

Calcular el valor de k.

Ejercicio 6

Un barco realiza un viaje con 400 personas. En dicho barco, debido a restricciones por COVID, se encuentra un único puesto de venta que vende remeras, gorros y hamburguesas a 8 dólares cada uno. La probabilidad de que una persona compre una remera es $1/4$, la probabilidad de que compre un gorro es $1/4$ y la probabilidad de que compre una hamburguesa es $1/4$, elecciones de compra independientes entre sí. Sabiendo que ninguna persona compra más de una remera, ni más de un gorro, ni más de una hamburguesa, calcule la probabilidad aproximada de que el barco tenga ventas superiores a 2430 dólares.

Ejercicio 7

Se instalará un nuevo sistema de atención al cliente solamente si la proporción p de clientes insatisfechos con el sistema actual supera el valor 0.150. Para la toma de la decisión se plantea la prueba de hipótesis

$H_0 : p \leq 0.15$ versus $H_1 : p > 0.15$ al 5%, para la cual se elegirán al azar e independientemente 1000 clientes y se contabilizará la cantidad de ellos que se declaren insatisfechos con el actual sistema.

Si el verdadero porcentaje de clientes insatisfechos fuera 0.17, calcular aproximadamente la potencia de la prueba.

Ejercicio 8

El número de vehículos que atraviesa un peaje en una cierta franja horaria es una variable aleatoria con distribución Poisson(λ). A partir de una muestra aleatoria en dicha franja horaria a lo largo de 100 días se obtiene un valor promedio de 1606 vehículos. Se considera la prueba de hipótesis $H_0 : \lambda = 1600$ versus $H_1 : \lambda < 1600$. Hallar aproximadamente el p-valor de la misma.

Solución

Ejercicio 1

Planteamos (a partir de Chevishoff)

$$P\left(\left|\overline{X}_n-\lambda\right|\leq 0.1\right)\geq 1-\frac{V\left(\overline{X}_n\right)}{0.1^2 n}=1-\frac{\lambda}{0.1^2 n}\overset{\text{porque }\lambda\leq 2}{\geq}1-\frac{2}{0.01n}\geq 0.93.$$

Despejando el valor de n nos queda $n\geq\frac{200}{0.07}=2857.1$, entonces el menor n que podemos lograr a partir de la desigualdad de Chevishoff es

$$n=2858.$$

Ejercicio 2

Planteamos el logaritmo de la función de verosimilitud.

$$h\left(\lambda\right)=\sum_{i=1}^{n}\log P(X=x_i)=\sum_{i=1}^{n}\log\left(\frac{e^{-\lambda}\lambda^{\sqrt{x_i}}}{\sqrt{x_i}!}\right)=$$

$$\sum_{i=1}^{n}\left(-\lambda+\sqrt{x_i}\log\left(\lambda\right)-\sqrt{x_i}!\right)=-n\lambda+\log\left(\lambda\right)\sum_{i=1}^{n}\sqrt{x_i}-\sum_{i=1}^{n}\sqrt{x_i}!.$$

$$h'(\lambda)=-n+\frac{1}{\lambda}\sum_{i=1}^{n}\sqrt{x_i}=0 \text{ si y sólo si } \lambda=\frac{1}{n}\sum_{i=1}^{n}\sqrt{x_i}.$$

Verificando con el signo de la derivada que el valor obtenido es máximo de la función h, se tiene que el estimador máximo verosímil de λ es $\widehat{\lambda}=\frac{1}{n}\sum_{i=1}^{n}\sqrt{x_i}$.

Con los datos de nuestra muestra, concluímos que la estimación máximo verosímil es

$$\widehat{\lambda}=\frac{1}{5}\left(1+4+3+4+5\right)=\frac{17}{5}=3.4.$$

Ejercicio 3

i	X_i^*	$\left\|i/10-\phi\left(\frac{X_i^*-3}{2}\right)\right\|$	$\left\|(i-1)/10-\phi\left(\frac{X_i^*-3}{2}\right)\right\|$
1	-1.28	0.084	0.016
2	1.72	0.061	0.161
3	1.91	0.007	0.093
4	2.23	0.050	0.050
5	2.48	0.103	0.003
6	2.52	0.195	0.095
7	2.53	0.293	0.193
8	2.6	0.379	0.279
9	2.68	0.464	0.364
10	2.8	0.540	0.440

Por lo tanto el estadístico de Kolmogorov y Smirnov en este caso es 0.54.

Ejercicio 4

Hallar la estimación de β por el método de los momentos.

$E(X) = \int_{-\infty}^{+\infty} x f_X(x) = \beta \int_1^{+\infty} x^{-\beta} dx = \frac{\beta}{\beta-1}$.
Por otro lado $\overline{X}_n = \frac{9.6+7.4+6.6+13.1+1.5+4.6+2.8+6.8+5.0}{9} = 6.377\,8$.
Entonces $\frac{\beta}{\beta-1} = 6.3778$, despejando β, obtenemos que la estimación queda

$$\widehat{\beta} = 1.1859.$$

Ejercicio 5

Aplicando el intervalo de confianza al nivel $1-\alpha$ para el porcentaje de éxito en una población Ber(p) es aproximado mediante el TCL para n grande, se obtiene que $k = \frac{\sqrt{\overline{X}_n(1-\overline{X}_n)}}{\sqrt{n}} \phi^{-1}(1-\alpha/2)$ donde $\alpha = 0.1$, $n = 27+22+8+20 = 77$ y $\overline{X}_n = \frac{8}{77} = 0.103\,90$, obtenemos

$$k = 0.057.$$

Ejercicio 6

Definimos la variable X =gasto de un pasajero en el barco.
Tenemos que $\mathrm{Rec}(X) = \{0, 8, 16, 24\}$.
$P(X=0) = \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} = \frac{27}{64}$, $P(X=8) = 3 \times \frac{1}{4} \times \frac{3}{4} \times \frac{3}{4} = \frac{27}{64}$,
$P(X=16) = 3 \times \frac{1}{4} \times \frac{1}{4} \times \frac{3}{4} = \frac{9}{64}$, $P(X=24) = \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$.
$\mu = E(X) = 8 \times \frac{27}{64} + 16 \times \frac{9}{64} + 24 \times \frac{1}{64} = 6$.
$\sigma^2 = E(X^2) - \mu^2 = 8^2 \times \frac{27}{64} + 16^2 \times \frac{9}{64} + 24^2 \times \frac{1}{64} - 36 = 36$.
Tenemos $X_1, X_2, ..., X_n$ MAS de X con $n = 400$. Por TCL la distribución de X es aproximadamente $N(\mu, \sigma^2/n) = N(6, 0.3)$. Entonces $\frac{2430}{400} = \frac{243}{40} = 6.075$

$$P\left(\overline{X}_{400} > \frac{2430}{400}\right) = P\left(\overline{X}_{400} > 6.075\right) \overset{\text{aprox}}{=} 1 - \phi\left(\frac{6.075-6}{0.3}\right) = 0.401.$$

Ejercicio 7

La región crítica de nivel 5% para este tipo de pruebas es de la forma $\left\{\overline{X}_n \geq 0.15 + \frac{\sqrt{0.15 \times 0.85}}{\sqrt{1000}} \times 1.645\right\} = \left\{\overline{X}_n \geq 0.16857\right\}$.
Aplicando el TCL sabemos que la distribución de $\overline{X}_n$ es aproximadamente

$$N\left(\mu, \frac{\sigma^2}{n}\right) = N\left(p, \frac{p(1-p)}{n}\right) \overset{p=0.17}{=} N(0.17, 0.0001411)$$

por lo que la potencia es

$$P\left(\overline{X}_n \geq 0.16857\right) = 1 - \phi\left(\frac{0.16887 - 0.17}{\sqrt{0.0001411}}\right) = 0.548.$$

Ejercicio 8

Dado que en esta variable tenemos que $\mu = \lambda$, se tiene que la región crítica aproximada es $\left\{\overline{X}_n \leq cte\right\}$. La distribución aproximada de $\overline{X}_n$ es $\mathrm{N}\left(\lambda, \frac{\lambda}{n}\right) \overset{\text{bajo } H_0}{=} N\left(1600, 16\right)$ y dado que en esta muestra se obtuvo $\overline{X}_n = 1606$ se deduce que

$$p - valor = P_{H_0}\left(\overline{X}_n \leq 1606\right) \overset{\text{aprox}}{=} \phi\left(\frac{1606 - 1600}{4}\right) = 0.933.$$