Curso: Introducción a la Probabilidad y Estadística Inferencial
CM-274
Tarea 3

Lista de ejercicios

---

1. Se espera que una gran cantidad de insectos sean atraídos por una cierta variedad de plantas de rosas. Un insecticida comercial anuncia que es 99% efectivo. Supongamos que 2000 insectos infectan una rosaleda donde se ha aplicado el insecticida, y sea $X =$ el número de insectos supervivientes.

   - ¿ Qué distribución de probabilidad deberia proporcionar un modelo razonable para este experimento?.
   - Expresa, pero no evalua, una expresión para la probabilidad de que menos de 100 insectos sobreviven, usando el modelo del ítem anterior.
   - Evalua una aproximación para la probabilidad en el ítem anterior.

2. La distribución hipergeométrica puede ser aproximada por una distribución binomial o de Poisson (aunque puede ser aproximada por otras distribuciones). Sea X que tiene una distribución hipergeométrica,

$$\mathbb{P}(X = x|N, M, K) = \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}, \ x = 0, 1, \ldots, K.$$

   - Muestra que cuando $N \to \infty, M \to \infty$ y $M/N \to p$,

$$\mathbb{P}(X = x|N, M, K) \to \binom{K}{x} p^x (1-p)^{K-x}, \ x = 0, 1, \ldots, K.$$

   - Usando el hecho que la distribución binomial puede ser aproximada por la distribución de Poisson para mostrar que si $N \to \infty, M \to \infty, K \to \infty, M/N \to 0$ y $KM/N \to \lambda$, entonces

$$\mathbb{P}(X = x|N, M, K) \to \frac{e^{-\lambda}\lambda^x}{x!}, \ \ x = 0, 1, \ldots$$

   - Verifica que la aproximación en el ítem anterior directamente, sin usar la aproximación de Poisson a la binomial.

3. Una distribución discreta truncada es aquella en la que una clase particular que no se puede observar se elimina del espacio muestral. En particular, si X tiene rango $0, 1, 2, \ldots$ y la clase 0 no puede ser observada , la variable aleatoria $0-$ truncada X_T tiene PMF

$$\mathbb{P}(X_T = x) = \frac{\mathbb{P}(X = x)}{\mathbb{P}(X > 0)}, \ x = 1, 2, \ldots$$

   Encuentra el PMF, media y varianza de la variable aleatoria $0-$ truncada empezando desde

   - $X \sim \text{Poisson}(\lambda)$.
   - $X \sim \text{binomial negativa}(r, p)$.

4. Una fábrica produce un determinado tipo de instrumento médico, y su calidad de funcionamiento depende de la suma de las excentricidades X_1 y X_2, $S = X_1 + X_2$ de dos ruedas de forma similar que se utilizan en la construcción de dicho instrumento médico. La excentricidad es una medida matemática de una desviación de la circularidad (es decir, una medida de `redondez`).

Para cualquier rueda de este tipo, la excentricidad X tiene una distribución modelada por la función de densidad exponencial negativa

$$f_X(x) = e^{-x},\ 0 < x < \infty.$$

Se sugieren dos métodos alternativos para elegir dos ruedas para la construcción de cualquiera de estos instrumentos médicos:

Ensamblado aleatorio (denotado RA): dos ruedas se eligen al azar.

Emsamblado por capas (denotado SA): las ruedas se dividen primero en dos capas de acuerdo a que si sus excentricidades están por encima de la mediana de $f_X(x)$ o por debajo de la mediana de $f_X(x)$ y luego se elige una rueda aleatoriamente de cada uno de estas dos capas.

¿Qué método de ensamblaje (RA o SA) produce la menor variabilidad en la suma $S = (X_1 + X_2)$ de las excentricidades de dos ruedas elegidas para la construcción de cualquiera de estos instrumentos médicos?.

5. En cada uno de los siguientes ejemplos encuentra el PDF de Y.

   - $Y = X^2$ y $f_X(x) = 1,\ 0 < x < 1$
   - $Y = -\log X$ y X tiene PDF

   $$f_X(x) = \frac{(n+m+1)!}{n!m!} x^n (1-x)^m,\ 0 < x < 1,\ \ m, n \text{ enteros positivos}$$

   - $Y = e^X$ y X tiene un PDF

   $$f_X(x) = \frac{1}{\sigma^2} x e^{-(x/\sigma)^2/2},\ 0 < x < \infty,\ \sigma^2 \text{ es una constante positiva}$$

6. Considera una variable aleatoria Y que puede ser discreta o continua. Sea A un evento arbitrario (un conjunto de Borel) definido a través de Y. Definimos una nueva variable aleatoria $X = I(A)$, la variable indicador del conjunto A, definido como

   $$X = \begin{cases} 0 & \text{si } A^c \text{ ocurre} \\ 1 & \text{si } A \text{ ocurre} \end{cases}$$

   Evalua μ y σ.

7. Los aviones de la fuerza área de un país se numeran de 1 a N. En una guerra este país pierde n aviones aleatorios al enemigo. que descubre que los aviones capturados están numerados. Si $X_1, X_2, \ldots, X_n$ son los números de aviones capturados, ¿ qué es $\mathbb{E}(\max X_i)$?. ¿ Cómo puede el enemigo usar $\mathbb{E}(\max X_i)$ para encontrar una estimación de N, el número total de aviones de ese país?.

8. Shuster (1991) describe una serie de cálculos de probabilidad que se hizo para un caso judicial relacionado con la venta de cocaína. Un departamento de policía de Florida incautó 496 paquetes sospechosos de cocaína, de los cuales cuatro fueron seleccionados y analizados al azar y se descubri"o que en realidad eran cocaína.

   La policía luego eligió dos paquetes más al azar y haciéndose pasar por traficantes de drogas, vendió los paquetes al acusado. Estos últimos dos paquetes se perdieron antes de que pudieran ser probados para verificar que fueran, de hecho, cocaína.

   - Si los 496 paquetes originales estaban compuestos de N paquetes de cocaína y $M = 496 - N$ no de cocaína, muestra que la probabilidad de seleccionar 4 paquetes de cocaína y 2 paquetes no de cocaína, que es la probabilidad de que el acusado sea inocente de comprar cocaína, es

   $$\frac{\binom{N}{4}\binom{M}{2}}{\binom{N+M}{4}\binom{N+M-4}{2}}.$$

- Maximizando (en M y N) la probabilidad en el ítem anterior, maximiza la `probabilidad de inocencia` del acusado. Demuestra que esta probabilidad es .022, que se alcanza en $M = 165$ y $N = 331$.

9. La función generadora de momentos factorial de X es definida como $\mathbb{E}t^X$, si la esperanza X existe. El nombre viene del hecho que la función satisface

$$\frac{d^r}{dt^r}\mathbb{E}t^X\bigg|_{t=1} = \mathbb{E}[X(X-1)\dots(X-r+1)],$$

donde el lado derecho es el momento factorial. Si X es una variable aleatoria discreta, podemos escribir

$$\mathbb{E}t^X = \sum_x t^x \mathbb{P}(X = x)$$

y la función generadora de momentos factorial es llamada función generadora de probabilidad, desde que los coeficientes de la serie de potencias dan probabilidades. Esto es, para obtener la probabilidad que $X = k$, se calcula

$$\frac{1}{k!}\frac{d^k}{dt^k}\mathbb{E}t^X\bigg|_{t=1} = \mathbb{P}(X = k).$$

Para calcular momentos de distribuciones discretas, es útil trabajar con momentos factoriales.

- Calcula el momento factorial $\mathbb{E}[X(X-1)]$ para las distribuciones binomial y de Poisson.
- Usa el resultado anterior para calcular las varianzas de las distribuciones binomial y de Poisson.
- Una particulas distribución es la beta-binomial, con PMF

$$\mathbb{P}(Y = y) = a(y+a)\frac{\binom{n}{a}\binom{a+b-1}{a}}{\binom{n+a+b-1}{y+a}},$$

donde n, a y b son enteros y $y = 0, 1, 2, \dots, n$. Usa el momento factorial para calcular la varianza de la distribución beta-binomial.

10. Muchas otras distribuciones son casos especiales de las más comunes distribuciones ya mencionadas. Para cada una de esas otras distribuciones deriva la forma de su PDF, verifica que es un PDF y calcula la media y la varianza.

- Si $X \sim$ Exponencial(β), entonces $Y = X^\gamma$ tiene la distribucíon de Weibull(γ, β), donde $\gamma > 0$ es una constante.
- Si $X \sim$ Exponencial(β), entonces $Y = (2X/\beta)^{1/2}$ tiene una distribucíon de Rayleigh.
- Si $X \sim$ Gamma(a, b), entonces $Y = 1/X$ tiene la distribuciíon gamma invertida IG(a, b). Esta distribución es útil en la estimación Bayesiana de la varianza.
- Si $X \sim$ Gamma$\left(\frac{3}{2}, \beta\right)$, entonces $Y = (X/\beta)^{1/2}$ tiene la distribucíon de Maxwell.
- Si $X \sim$ Exponencial(1), entonces $Y = \alpha - \gamma \log X$ tiene la distribución Gumbel(α, γ) donde $-\infty < \alpha < \alpha$ y $\gamma > 0$. (La distribución de Gumbel también es conocida como la distribución de valor extremo).