Guía 9

9.1 Sea X_1, X_2, X_3 una muestra aleatoria de la distribución Bernoulli(p).

(**a**) Verificar que $T = X_1 + X_2 + X_3$ es un estadístico suficiente para p.

(**b**) ¿$T = X_1 + 2X_2 + X_3$ es un estadístico suficiente para p?. :*no es difícil analizar los 8 casos.*

9.2 Sea $\mathbf{X}_n$ una muestra aleatoria de tamaño n de una distribución con densidad

$$f_\theta(x) = e^{-(x-\theta)}\mathbf{1}\{x > \theta\}.$$

Verificar que $T = \text{mín}(X_1, \dots, X_n)$ es un estadístico suficiente para θ.

9.3 Sea $\mathbf{X}_n$ una muestra aleatoria de tamaño n de la distribución Poisson(λ) y sea $T = \sum_{i=1}^n X_i$. Hallar la distribución de $\mathbf{X}_n | T = t$ y deducir que T es un estadístico suficiente para λ.

9.4 Mostrar que las siguientes familias de distribuciones son familias exponenciales a 1 parámetro: (**a**) Bernoulli(p); (**b**) Pascal$(4, p)$; (**c**) Poisson(λ); (**d**) Exponencial(λ).

9.5 Una moneda tiene una probabilidad de cara p, $p \in \{2/5, 4/5\}$. En 10 lanzamientos de la moneda se observaron exactamente 3 caras. Estimar por máxima verosimilitud la probabilidad de que en otros tres lanzamientos se observe exactamente una cara.

9.6 La cantidad de accidentes de tránsito por semana en la intersección de Paseo Colón y Estados Unidos tiene una distribución de Poisson de media λ.

(**a**) Hallar el estimador de máxima verosimilitud de λ basado en una muestra aleatoria de la cantidad de accidentes durante n semanas. Mostrar que se trata de un estimador insesgado para λ y hallar la expresión de su error cuadrático medio.

(**b**) En una muestra de 100 semanas se observaron las frecuencias:

Cantidad de accidentes	0	1	2	3	4	5
Frecuencia	10	29	25	17	13	6

En virtud de la información muestral calcular el estimador de máxima verosimilitud de λ y estimar la probabilidad de que en la semana del 24 de diciembre de 2017 no ocurra ningún accidente en la mencionada esquina.

9.7 Sea $\mathbf{X}_n$ una muestra aleatoria de tamaño n de la distribución uniforme sobre el intervalo $[0, \theta]$.

(**a**) Hallar un estadístico suficiente para θ basado en $\mathbf{X}_n$.

(**b**) Hallar el estimador de máxima verosimilitud de θ basado en $\mathbf{X}_n$.

(**c**) Sea $\hat{\theta}_n$ el estimador de máxima verosimilitud de θ hallado en el inciso anterior. Mostrar que $\mathbf{E}_\theta[\hat{\theta}_n] = \frac{n}{n+1}\theta$ y $\mathbf{var}_\theta(\hat{\theta}_n) = \frac{n\theta^2}{(n+1)^2(n+2)}$ para concluir que $\hat{\theta}_n$ converge en media cuadrática a θ cuando $n \to \infty$.

9.8 El tamaño, X (en GB), de ciertos archivos es una variable aleatoria cuya densidad es

$$f_\theta(x) = 3\theta^3 x^{-4}\mathbf{1}\{x \geqslant \theta\}, \qquad \theta > 0.$$

(**a**) Hallar el estimador de máxima verosimilitud de θ basado en una muestra aleatoria de los tamaños de n archivos.

(**b**) Hallar expresiones para la esperanza y la varianza del estimador de máxima verosimilitud de θ.

(**c**) Mostrar que el estimador de máxima verosimilitud de θ converge en media cuadrática al verdadero valor de θ.

9.9 La duración, X, en años de ciertos discos rígidos tiene la distribución Pareto con densidad

$$f_\theta(x) = \theta x^{-(\theta+1)}\mathbf{1}\{x > 1\}, \qquad \theta > 1.$$

(**a**) Usar el *criterio de factorización de Neyman-Fisher* para hallar un estadístico suficiente para θ, basado en una muestra aleatoria de la duración de n discos.

(**b**) Mostrar que las distribuciones f_θ, $\theta > 1$, pertenecen a una familia exponencial y usar esa propiedad para hallar un estadístico suficiente para θ. ¿Cuál es su distribución? :*notar que* $\log X$ ∼Exponencial(θ).

(**c**) Hallar el estimador de máxima verosimilitud de θ basado en una muestra aleatoria de la duración de n discos rígidos. Mostrar que se trata de un estimador asintóticamente insesgado y cuya varianza tiende a 0 cuando $n \to \infty$.

(**d**) Hallar la distribución asintótica del estimador de máxima verosimilitud de θ. : *es fácil ver que* $I(\theta) = \theta^{-2}$.

9.10 La duración en años de cierto tipo de dispositivos es una variable aleatoria X con función intensidad $\lambda(x) = 3\theta^{-3}x^2\mathbf{1}\{x > 0\}$.

(**a**) Hallar un estadístico suficiente para θ basado en una muestra aleatoria de la duración de n dispositivos.

(**b**) Hallar el estimador de máxima verosimilitud de θ basado en una muestra aleatoria de la duración de n dispositivos.

(**c**) Usando los números aleatorios

$$0.186,\ 0.178,\ 0.488,\ 0.255,\ 0.392,\ 0.234,\ 0.597,\ 0.205,\ 0.611,\ 0.651.$$

simular 10 valores de X cuando $\theta = 1$ y en base a esa información muestral calcular el estimador de máxima verosimilitud de θ.

(**d**) Se pusieron a prueba 10 de esas máquinas y se obtuvieron los siguientes tiempos:

$$2.00,\ 5.48,\ 2.43,\ 1.90,\ 5.85,\ 1.58,\ 2.30,\ 2.87,\ 3.62,\ 2.71.$$

Basándose en la información muestral calcular el estimador de máxima verosimilitud de la probabilidad de que una máquina del mismo tipo funcione sin fallas más de dos años y medio.

9.11 En una mesa electoral votaron 129 ciudadanos. Se extrajeron (sin reposición) 7 sobres al azar de la urna, se examinaron y resultó que el candidato verde obtuvo exactamente 3 votos. Estimar por máxima verosimilitud la cantidad de votos por el candidato verde que había en la urna.

9.12 Mostrar que la familia de distribuciones $\Gamma(\nu, \lambda)$ es una familia exponencial a 2 parámetros. Hallar un estadístico suficiente para (ν, λ) basado en una muestra aleatoria de tamaño n

---

9.13

(**a**) Mostrar que la familia de distribuciones $\mathcal{N}(\mu, \sigma^2)$ puede expresarse en la forma

$$f_\theta(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \exp\left(\frac{\mu}{\sigma^2}x - \frac{1}{2\sigma^2}x^2\right),$$

donde $\theta = (\mu, \sigma^2)$.

(**b**) Sea $X_1, \dots, X_n$ una muestra aleatoria de tamaño n de la distribución $\mathcal{N}(\mu, \sigma^2)$. Hallar la expresión de la densidad conjunta y mostrar que $T = \left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right)$ es un estadístico suficiente para θ.

(**c**) Sea $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$. Mostrar que $\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2$ y deducir que $T' = (\bar{X}, S^2)$, donde

$$S^2 := \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2,$$

es un estadístico suficiente para θ.

(**d**) Hallar el estimador de máxima verosimilitud para θ basado en la muestra aleatoria $X_1, \dots, X_n$.

9.14 Se arroja un dado piramidal n veces. El dado tiene las caras numeradas $1, 2, 3, 4$ y están cargadas con probabilidades p_1, p_2, p_3, p_4, respectivamente. Sean X_1, X_2, X_3, X_4 la cantidad de lanzamientos en los que el dado cae en la cara $1, 2, 3, 4$, respectivamente. Mostrar que la distribución de (X_1, X_2, X_3, X_4) pertenece a una familia exponencial a 3 parámetros.

9.15 Al finalizar el primer semestre de gobierno se realizó una encuesta entre 1200 ciudadanos, 414 de los cuales declararon ser oficialistas, 196 declararon no ser ni oficialistas ni opositores y el resto declaró ser opositor. En base a esa información muestral estimar por máxima verosimilitud (p_1, p_2), donde p_1 es la probabilidad de que un ciudadano elegido al azar sea oficialista y p_2 la de que sea opositor.
