

Inferencia Estadística: Tarea 2

Estimación puntual

Fecha de entrega: 11 de octubre

1. Sea X una variable aleatoria con distribución $t_{(n)}$.
 - (a) Encuentre la media y varianza de X . (*Hint: Mostrar que la distribución de X , se puede obtener cuando se asume que $X|\lambda \sim N(0, \lambda^{-1})$ con $\lambda \sim Ga(n/2, n/2)$.*)
 - (b) Demuestre que X^2 tiene una distribución $F_{(1,n)}$.
 - (c) Utilizando la fórmula de Stirling, demuestre que

$$\lim_{n \rightarrow \infty} f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

2. Sea X una variable aleatoria con distribución $F_{(m,n)}$.
 - (a) Demuestre que $1/X$ tiene distribución $F_{(n,m)}$.
 - (b) Demuestre que $W = (m/n)X/[1 + (m/n)X]$ tiene distribución $Be(m/2, n/2)$.
 - (c) Utilizando el inciso anterior, encuentre la media y varianza de X . (*Hint: Encuentre los primeros dos momentos de $mX/n = W/(1 - W)$*).
3. Sea X_1, \dots, X_n una muestra aleatoria con densidad $f(x | \theta)$. Encuentre el estimador de momentos del parámetro θ de las siguientes funciones de densidad.
 - (a) $f(x | \theta) = \theta x^{\theta-1}$, $0 < x < 1$, $0 < \theta < \infty$
 - (b) $f(x | \theta) = \frac{1}{2\theta}$, $-\theta < x < \theta$, $0 < \theta < \infty$
 - (c) $f(x | \theta) = \theta x^{-2}$, $0 < \theta < x < \infty$

4. La distribución Pareto, es una distribución de probabilidad continua nombrada a partir del economista italiano Vilfredo Pareto, el cual introdujo un principio matemático conocido como la regla 80/20 para medir la desigualdad de la distribución de la riqueza. La función de densidad está dada por:

$$f(x | \theta) = \frac{\theta \alpha^\theta}{x^{\theta+1}}, \quad 0 < \alpha \leq x < \infty, \quad 0 < \theta < \infty$$

donde α es un parámetro de escala y θ es un parámetro de forma.

- (a) Asumiendo que α es conocido, encuentre el estimador máximo verosímil de θ a partir de una muestra aleatoria X_1, \dots, X_n .
 - (b) Asumiendo que los dos parámetros son desconocidos, encuentre el estimador máximo verosímil de α y θ .
 - (c) Exhiba las estadísticas suficientes para α y θ .
5. Considera la base de datos *precioCasas.txt*, la cual detalla el precio de venta (en miles de dólares) de un conjunto de casas con ciertas características. Considerando la variable precio realiza lo siguiente.
- (a) Suponiendo que los precios de venta siguen una distribución normal. Encuentre los estimadores máximo verosímiles de μ y σ .
 - (b) Suponiendo ahora que los precios de venta siguen una distribución log-normal. Encuentre los estimadores máximo verosímiles de los parámetros del modelo.
 - (c) En una misma gráfica muestre el histograma de los datos y grafique las funciones de densidad estimadas bajo ambos modelos. ¿Qué distribución es más apropiada?
6. Considera los datos *tiemposOperacion.txt*, los cuales son registros de 120 tiempos de operación en una empresa de manufactura de gimnasios y realiza lo siguiente.
- (a) Ajusta a los datos una distribución exponencial de parámetro (desconocido) λ , el cual debes encontrar mediante el método de máxima verosimilitud. Una vez que se tenga $\hat{\lambda}$ dibuja el histograma y añade en la misma gráfica la curva de la densidad exponencial con parámetro $\hat{\lambda}$. ¿Se ajusta bien a los datos?
 - (b) Ahora considera una distribución gamma de parámetros α, β donde α se asume que es igual a 2 y β es desconocida. Encuentra el estimador máximo verosímil de β y realiza el mismo análisis que en el inciso (a).
 - (c) Finalmente, considera una distribución gamma con parámetros α y β desconocidos. Encuentra de forma analítica (o numérica en caso de ser necesario) los estimadores máximo verosímiles y realiza el mismo análisis que en el inciso (a).
 - (d) ¿Qué distribución se ajusta mejor a los datos?
7. Considera que la colección de variables aleatorias X_1, \dots, X_n son condicionalmente independientes dado un parámetro λ de tal forma que

$$f(x_1, \dots, x_n | \lambda) = \prod_{i=1}^n f(x_i | \lambda).$$

Suponiendo que $\lambda \sim Ga(\alpha, \beta)$ y $X_i | \lambda \sim Exp(\lambda)$.

- (a) Obtén la distribución posterior de λ . ¿Qué puedes decir sobre la elección de este modelo y de esta distribución inicial?
- (b) Obtén la esperanza posterior y demuestra que se puede ver como un promedio ponderado de la media inicial y el estimador máximo verosímil de λ .

8. Sea X_1, \dots, X_n una muestra aleatoria proveniente de una población con densidad,

$$f(x | \theta) = \frac{e^{-|x-\theta|}}{2} \mathbf{I}_{(-\infty, \infty)}^{(x)}.$$

Encuentra el estimador máximo verosímil de θ .

9. Considera que la colección de variables aleatorias X_1, \dots, X_n son condicionalmente independientes dado un parámetro μ . Suponiendo que $\mu \sim N(\eta, \tau^2)$ y $X_i | \mu \sim N(\mu, \sigma^2)$, donde σ^2 es conocida.

- (a) Obtén la distribución posterior de μ . ¿Qué puedes decir sobre la elección de este modelo y de esta distribución inicial?
- (b) Obtén la esperanza posterior de μ .

10. Sean $\hat{\theta}_1$ y $\hat{\theta}_2$ dos estimadores insesgados de θ , con $\text{Var}(\hat{\theta}_1) = \sigma_1^2$ y $\text{Var}(\hat{\theta}_2) = \sigma_2^2$. Considere un tercer estimador dado por,

$$\hat{\theta}_3 = \alpha \hat{\theta}_1 + (1 - \alpha) \hat{\theta}_2, \quad 0 < \alpha < 1.$$

- (a) ¿Es $\hat{\theta}_3$ un estimador insesgado de θ ?
- (b) Si $\hat{\theta}_1$ y $\hat{\theta}_2$ son independientes. ¿Para qué valor de α se minimiza la varianza de $\hat{\theta}_3$?
- (c) Si $\hat{\theta}_1$ y $\hat{\theta}_2$ no son independientes y $\text{cov}(\hat{\theta}_1, \hat{\theta}_2) = c$. ¿Para qué valor de α se minimiza la varianza de $\hat{\theta}_3$?

11. Sea X_1, \dots, X_n los tiempos de supervivencia de una muestra aleatoria de n individuos a los que se les dio un tratamiento médico. Asumiendo que siguen una distribución exponencial de media θ .

- (a) Demuestra que $\hat{\theta} = \bar{X}$ es un estimador insesgado de θ . Obtenga su varianza.
- (b) ¿ $\hat{\theta}$ alcanza la cota inferior de Cramér - Rao para estimadores insesgados?
- (c) Sea $Y = \min\{X_1, \dots, X_n\}$. Encuentre $\mathbb{P}(Y > y)$ y deduzca la función de densidad de Y . Obtén además su media y varianza.
- (d) Proponga un estimador insesgado de θ basado en Y . De los estimadores obtenidos ¿Cuál es eficiente? ¿Cuál es consistente?

12. Sea X_1, \dots, X_n una muestra aleatoria de una distribución Bernoulli(p). En este problema, el interés es estimar $\nu = p(1 - p)$.

(a) Muestre que el estimador $\tilde{\nu} = X_1(1 - X_2)$ es insesgado para ν .

(b) Utilizando los resultados vistos en clase, obtén un mejor estimador de $\hat{\nu}$ para ν .

(c) ¿Son el estimador máximo verosímil de ν y el obtenido en el inciso anterior similares?

13. Sea X_1, \dots, X_n una muestra aleatoria con distribución uniforme en el intervalo $(0, \theta)$.

(a) Encuentre la cota inferior de Cramér - Rao de la varianza de un estimador insesgado de θ .

(b) Ahora considere $\hat{\theta} = X_{(n)}$ el estimador máximo verosímil de θ , encuentre una constante c tal que $\tilde{\theta} = c\hat{\theta}$ es un estimador insesgado de θ .

(c) Calcule la varianza de $\tilde{\theta}$. ¿Por qué la varianza de este estimador es menor que la cota inferior de Cramér - Rao?

14. Sea X_1, \dots, X_n una muestra aleatoria de la población con distribución uniforme $U(\theta, \theta + 1)$. Demuestre que $S = (Y_1, Y_n)$, donde $Y_1 = \min\{X_i\}$ y $Y_n = \max\{X_i\}$ es una estadística suficiente minimal pero no es completa.

15. Sea X_1, \dots, X_n una muestra aleatoria de la población cuya función de densidad es

$$f(x | \theta) = \frac{\theta^2}{\theta + 1}(x + 1) \exp(-\theta x), \quad 0 < x < \infty, \quad 0 < \theta < \infty.$$

Obtenga una estadística suficiente minimal y completa. Además, obtenga el estimador por momentos y el estimador máximo verosímil de θ .

16. Considere el caso de una distribución $N(\theta, \theta^2)$ de la cual se observa una muestra aleatoria de tamaño n , siendo θ el parámetro desconocido.

(a) Muestre que se está frente a un caso de una familia exponencial en el que la dimensión de la estadística suficiente minimal, es 2; siendo la dimensión original igual a 1.

(b) Demuestra que T no es completa. (*Hint: Observe las distribuciones de las componentes de T y note por ejemplo que $\mathbb{E}(\bar{X}^2) = \frac{\theta(n+1)}{n}$.*)

17. Sea X_1, \dots, X_n una muestra aleatoria con distribución Poisson de parámetro λ , donde se desea estimar a $\tau(\lambda) = (1 + \lambda)e^{-\lambda}$.

(a) ¿Cuál es el estimador máximo verosímil de λ ?

(b) Obtén un estimador insesgado para $\tau(\lambda)$.

(c) Encuentra el UMVUE de $\tau(\lambda)$.

18. Sea X_1, \dots, X_n una muestra aleatoria con densidad $Be(\theta, 1)$.
- Encuentre el estimador máximo verosímil para θ y encuentre su varianza. ¿Qué puedes observar de esto?
 - Encuentra el estimador máximo verosímil para $\tau(\theta) = \theta^{-1}$. ¿Es insesgado? ¿Alcanza la cota inferior de Cramér - Rao?
 - ¿Es \bar{X} un estimador insesgado para $\theta/(1 + \theta)$? ¿Alcanza la cota inferior de Cramér - Rao?
19. Sea X_1, \dots, X_n una muestra aleatoria de una distribución Poisson de parámetro θ . Considerando la estadística

$$T = \frac{1}{2}(X_1 + X_2),$$

realice lo siguiente.

- Obtén la $\mathbb{E}(T)$. ¿Qué puedes decir del estimador?
 - Calcula la varianza de T .
 - Encuentra una estadística U que sea suficiente y completa para θ . Con ella mejora el estimador T y obtén su varianza. ¿Qué puedes concluir?
20. Sea X_1, \dots, X_n una muestra aleatoria de una distribución Poisson de parámetro θ . Encuentre el UMVUE para $\tau(\theta) = \theta e^{-\theta}$.