Inferencia Estadística: Tarea 3

Estimación puntual

Fecha de entrega: 20 de octubre

1. (1 punto) Sea $X_1, \ldots, X_n$ 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) $f(x;\theta) = \theta x^{-2},\ 0 < \theta < x < \infty.$

   Es importante notar que el estimador puede no existir.

2. (1.5 puntos) 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^2.

   (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?

3. (1.5 puntos) 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?

4. (1 punto) Considera que la colección de variables aleatorias $X_1, \ldots, X_n$ son condicionalmente independientes dado un parámetro λ de tal forma que

$$f(x_1, \ldots, x_n \mid \lambda) \quad = \quad \prod_{i=1}^{n} f(x_i \mid \lambda).$$

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

(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.

5. (1 punto) Sean $\widehat{\theta_1}$ y $\widehat{\theta_2}$ dos estimadores insesgados de θ, con $\text{Var}(\widehat{\theta_1}) = \sigma_1^2$ y $\text{Var}(\widehat{\theta_2}) = \sigma_2^2$. Considere un tercer estimador dado por,

$$\widehat{\theta_3} \quad = \quad \alpha\widehat{\theta_1} + (1-\alpha)\widehat{\theta_2}, \quad 0 < \alpha < 1.$$

(a) ¿Es $\widehat{\theta_3}$ un estimador insesgado de θ?

(b) Si $\widehat{\theta_1}$ y $\widehat{\theta_2}$ son independientes. ¿Para qué valor de α se minimiza la varianza de $\widehat{\theta_3}$?

(c) Si $\widehat{\theta_1}$ y $\widehat{\theta_2}$ no son independientes y $\text{cov}(\widehat{\theta_1}, \widehat{\theta_2}) = c$. ¿Para qué valor de α se minimiza la varianza de $\widehat{\theta_3}$?

6. (1 punto) Sean $X_1, \ldots, X_n$ los tiempos de supervivencia de una muestra aleatoria de n individuos a los que se les administró un tratamiento médico. Asumiendo que siguen una distribución exponencial de media θ.

(a) Demuestre que $\widehat{\theta} = \bar{X}$ es un estimador insesgado de θ y obtén su varianza.

(b) ¿El estimador $\widehat{\theta}$ alcanza la cota inferior de Cramér - Rao?

(c) Sea $Y = \min\{X_1, \ldots, X_n\}$. 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?

7. (1 puntos) 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) Encuentra el estimador máximo verosímil de ν.

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

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

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

8. (1 punto) 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 $\widehat{\theta} = X_{(n)}$ el estimador máximo verosímil de θ, encuentre una constante c tal que $\tilde{\theta} = c\widehat{\theta}$ 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?

9. (1 punto) Sea $X_1, \dots, X_n$ una muestra aleatoria de una distribución Poisson de parámetro θ. Considerando el estimador para θ dado por:

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

   realice lo siguiente.

   (a) Obtén la $\mathbb{E}(T)$. ¿Qué puedes decir del estimador?

   (b) Calcula la varianza de T.

   (c) Encuentra una estadística S que sea suficiente y completa para θ. Con ella mejora el estimador T y obtén su varianza. ¿Qué puedes concluir?

Actividades de DataCamp

1. *Intermediate R*

2. *Introduction to Writing Functions in R*

3. *Exploratory Data Analysis in R*