Tarea Examen

Estimación por intervalos

Fecha de entrega: 4 de noviembre

1. Utilizando el software de tu preferencia realiza lo siguiente.

   (a) Fijando una semilla inicial para que tus resultados sean reproducibles, genera 100 muestras de tamaño 100 de una distribución normal de parámetros $(\mu = 2, \sigma = 2)$. Guarda tus muestras en una matriz.

   (b) Calcula el estimador máximo verosímil para cada una de las 100 muestras.

   (c) Asumiendo σ conocido y $\alpha = 0.1, .05, .01$, construye los intervalos de confianza para la media (para cada una de las 100 muestras) y grafica tus resultados. ¿Cuántos contienen a μ?

   (d) Asumiendo σ desconocido y $\alpha = 0.1, .05, .01$, construye los intervalos de confianza para la media (para cada una de las 100 muestras) y grafica tus resultado. ¿Cuántos contienen a μ?

   (e) ¿Notas alguna diferencia significativa en los intervalos generados con σ conocida y desconocida? ¿Por qué? ¿Qué puedes concluir para los diferentes valores de α?

2. Utilizando el software de tu preferencia realiza lo siguiente.

   (a) Fijando una semilla inicial para que tus resultados sean reproducibles, genera una muestra de tamaño 50 de una distribución normal de parámetros $(\mu_x = -1, \sigma_x = 5)$ y otra muestra de tamaño 30 de una distribución normal con parámetros $(\mu_y = 1, \sigma_y = 5)$.

   (b) Calcula los estimadores máximo verosímiles para ambas muestras y dibuja ambas densidades en un mismo gráfico y comenta lo que observes sobre las medias poblaciones.

   (c) Asumiendo $\sigma_x = \sigma_y = \sigma$ conocida y $\alpha = 0.1, .05, .01$, construye los intervalos de confianza para la diferencia de las medias. ¿Puedes asumir que las medias son iguales?

   (d) Asumiendo $\sigma_x = \sigma_y = \sigma$ desconocida y $\alpha = 0.1, .05, .01$, construye los intervalos de confianza para la diferencia de las medias. ¿Puedes asumir que las medias son iguales?

3. Sea $X_1, \ldots, X_n$ una muestra de una distribución exponencial de parámetro θ, tal que $\mathbb{E}(X) = 1/\theta$.

   (a) Encuentre un intervalo de confianza al $100(1-\alpha)\%$ para la media de la población.

(b) Encuentre un intervalo de confianza al $100(1-\alpha)\%$ para la varianza de la población.

(c) Encuentre una cantidad pivotal basada únicamente en el mínimo y úsela para encontrar un estimador de intervalo para θ.

4. En una compañía farmacéutica se desea conocer si hay diferencia en el tratamiento de una enfermedad con dos dosis diferentes de un fármaco. Para esto, se tienen dos grupos de enfermos para los cuales se registró la dosis mínima (en escala logarítimica) que da los mejores resultados contra la enfermedad (ver, Tabla 1).

Tabla 1: Dosis de un fármaco para dos grupos.

Grupo 1	.0133	0.0265	0.0302	0.0444	0.0882	0.1461	0.1545	0.1585	0.2638	
Grupo 2	-0.1221	-0.1010	-0.0519	-0.0436	-0.0200	-0.0182	-0.0104	0.0879	0.1390	0.1945

(a) Construye un intervalo de confianza para el cociente de las varianzas a un nivel de confianza del 99%. ¿Puedes asumir que $\sigma_x^2 = \sigma_y^2$?

(b) Asumiendo que $\sigma_x^2 = \sigma_y^2 = .1$, construye un intervalo de confianza al 99% para la diferencia de las medias. ¿Hay posibilidad de asumir que son iguales?

(c) Asumiendo que $\sigma_x^2 = \sigma_y^2 = \sigma$ pero en este caso desconocida, construye un intervalo de confianza al 99% para la diferencia de las medias. ¿Hay posibilidad de asumir que son iguales?

5. Los siguientes datos representan el tiempo de vida útil de un artículo, medido en días: 29.1, 207.6, 81.8, 0.8, 76.1, 108.9, 48.4, 108.1, 52.2, 272.8, 150.5, 80.3, 97.4, 11.5, 46.2, 144.1, 62.5, 262.9, 247.6, 4.1. Este tiempo se asume que sigue una distribución exponencial con media θ.

(a) Encuentre un intervalo de confianza exacto al 95% para la media.

(b) Encuentre un intervalo de confianza aproximado al 95% usando el teorema central del límite.

(c) Encuentre un intervalo de confianza aproximado al 95% utilizando la distribución asintótica del estimador máximo verosímil.

(d) Comente en los resultados obtenidos y en las diferencias (si existen) entre los tres procedimientos.

6. Se lanza una moneda 500 veces y se obtienen 275 águilas y 225 soles. Obtenga un intervalo de confianza al 99% para la probabilidad de obtener un águila. ¿Qué puedes decir sobre la moneda?

7. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una población uniforme en el intervalo $(0, 1/\theta)$. Encuentre un intervalo de confianza para θ.

8. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución beta $(\theta, 1)$, y asuma que la distribución inicial de θ es gamma de parámetros (κ, β). Encuentra un intervalo de credibilidad al $100(1-\alpha)\%$.

9. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución normal (μ, τ) donde $\tau = 1/\sigma^2$ es la precisión. Esto es,

$$f(x|\mu, \tau) = \frac{\tau^{1/2}}{\sqrt{2\pi}} \exp\left(-\frac{\tau(x-\mu)^2}{2}\right).$$

Asuma que $\mu|\tau$ se distribuye normal $(\mu_0, \beta\tau)$ tal que $\beta\tau$ es la precisión, esto es,

$$f(\mu|\tau) = \frac{(\beta\tau)^{1/2}}{\sqrt{2\pi}} \exp\left(-\frac{\beta\tau(\mu-\mu_0)^2}{2}\right).$$

y τ se distribuye gamma de parámetros (κ, ρ). Por lo que la distribución inicial está dada por $f(\mu, \tau) = f(\mu|\tau)f(\tau)$, la cual se conoce como una distribución normal-gamma.

   (a) Encuentre la distribución posterior de (μ, τ) y muestre que es conjugada.

   (b) Obtenga un intervalo de credibilidad (de colas iguales) al 95% para cada parámetro. Para esto primero encuentre $f(\tau|\mathbf{X})$ y después realice un procedimiento similar al ejercicio 1 de la tarea 2 para derivar $f(\mu|\mathbf{X})$.

10. En este ejercicio se explorará brevemente la idea del bootstrap. Originalmente ideado por Efron en los años 70's, los métodos bootstrap son una clase de algoritmos de remuestreo con reemplazo (esto se puede lograr en R con la función sample) para hacer inferencia. Si $X_1, \ldots, X_n$ es la muestra original, la idea del método de Efron es obtener N muestras con reemplazo de tamaño n de $X_1, \ldots, X_n$ y con ellas realizar un proceso de inferencia. Utilizando los datos del ejercicio 5, realice lo siguiente.

   (a) Genere una matriz de $N = 1000$ renglones y $n = 20$ columnas, donde en cada renglón guardará un remuestreo con reemplazo de la muestra original.

   (b) Para cada una de las 1000 muestras, obtenga la media muestral. Dibuje un histograma y comente en los resultados obtenidos.

   (c) Obtenga la media de las medias calculadas en el inciso anterior. ¿Qué puede comentar de este estimador y del máximo verosímil?

   (d) Ordene de mayor a menor el vector de medias y construya el intervalo de confianza al 95% seleccionando los cuantiles empíricos. Esto es, el valor inferior estará dado por el cuantil .025 y el valor superior estará dado por el cuantil .975. ¿Cómo se compara con los obtenidos en el ejercicio 5?