Facultad de Ingeniería y Tecnologías
Departamento de Matemática
Probabilidad y Estadística

UCU

Segundo Informe de Probabilidad y Estadística

Cada grupo trabajará con la distribución 2 que fue usada en el primer informe.

Ejercicio 1. El objetivo de este ejercicio es generar muestras de la distribución 2 y construir intervalos de confianza aproximados para la media. También se quiere interpretar el significado de la confianza. Dadas $X_1, X_2, \ldots, X_{50}$ iid con distribución 2 consideramos $\overline{X}_{50}$ su promedio empírico.

a) Hallar la media y la varianza **teóricas** de $\overline{X}_{50}$.

b) Construir un vector v de largo 10^4 que contenga simulaciones independientes de $\overline{X}_{50}$. Construir un histograma a partir del vector v, y superponer la densidad de una normal $N(\mu, \sigma^2)$, donde μ y σ^2 son los valores calculados en la parte anterior. Justifique el gráfico obtenido.

c) Construir 10^4 intervalos de confianza para la media a partir de las muestras generadas en la parte anterior. Utilice una confianza de 95 % y asuma varianza desconocida.

d) De los 10^4 intervalos de confianza construidos en la parte anterior, calcular la proporción de ellos que no contiene el valor esperado teórico de X_1. ¿A qué debería tender dicha proporción?

Ejercicio 2. El objetivo de este ejercicio es estudiar el test de hipótesis bilateral para la media de una población, también se quiere estimar y calcular el p-valor del test a partir de una muestra. Nuevamente trabajaremos con una muestra $X_1, X_2, \ldots, X_{50}$ iid con distribución 2. Sea $\mu_0 = E(X_1)$. Planteamos el test de hipótesis bilateral para la media:

$$\begin{cases} H_0: & \mu = \mu_0 \\ H_1: & \mu \neq \mu_0 \end{cases}$$

donde consideramos $\alpha = 0,01$ y asumimos que $\sigma^2 = \mathrm{Var}(X_1)$ es conocida.

Recordemos que la región crítica del test es

$$RC = \left\{ |\overline{X}_n - \mu_0| > \frac{\sigma z_{\frac{\alpha}{2}}}{\sqrt{n}} \right\}$$

Además si $x_1, x_2, \ldots, x_n$ es una muestra iid con distribución 2 y $\overline{x}_n$ es su promedio entonces el p-valor del test para esta muestra es $\alpha^* = P_{H_0}(|\overline{X}_n - \mu_0| > |\overline{x}_n - \mu_0|)$.

a) Para el vector v hallado en la parte b) del Ejercicio 1, hallar la proporción de valores de $\overline{X}_{50}$ para los cuales se rechaza H_0. ¿A qué debería tender esta proporción?

b) Generar una nueva muestra $x_1, x_2, \ldots, x_{50}$ iid con distribución 2 y estimar el p-valor del test para esta muestra usando la proporción de valores de $\overline{X}_n$ en el vector v hallado en la parte b) del Ejercicio 1 que cumplen $|\overline{X}_n - \mu_0| > |\overline{x}_n - \mu_0|$.

c) Hallar aproximadamente el p-valor del test para la muestra de la parte anterior calculando la probabilidad $P_{H_0}(|\overline{X}_n - \mu_0| > |\overline{x}_n - \mu_0|)$. Comparar el resultado con la estimación hallada en la parte anterior.

Sobre el informe:

- **Fecha de entrega a confirmar.** Más adelante se creará un recurso en webasignatura para realizar la entrega online.
- El informe deberá estar en formato pdf, debe incluir los gráficos solicitados y los scripts utilizados. No es necesario entregar por separado los scripts utilizados.
- El informe deberá contener título, fecha, nombre y cédula de los integrantes del grupo.
- Se evaluará: prolijidad del informe, utilización correcta del idioma español, redacción, prolijidad del código presentado en los scripts, calidad de los gráficos, conclusiones.