UNIVERSIDAD NACIONAL DE INGENIERÍA
FACULTAD DE CIENCIAS
ESCUELA PROFESIONAL DE MATEMÁTICA CICLO 2019–II

SOLUCIONARIO DE LA PRÁCTICA CALIFICADA 2 - CM2H2

# 1 Parte teórica

Justifique el valor de verdad o falsedad de los siguientes afirmaciones.

1. Cuando tratamos con muestras pequeñas, hacemos uso del teorema del límite central. (1 pto.)

   **Solución 1.1** *(Falso) El teorema de límite central nos dice que cuando el tamaño de la muestra ($N$) es grande, la media de $\bar{Y}$ de una muestra aleatoria sigue una distribución normal centrada en la media de la población $\mu_Y$ y la desviación estándar es igual a la desviación estandar de la población $\sigma_Y$ dividido por la raíz cuadrada del tamaño de la muestra $N$.*

2. La $t$-student es una distribución simétrica alrededor de la media$=0$, con forma de campana, pero con las colas menos pesadas que la normal $N \sim (0, 1)$. (1 pto.)

   **Solución 1.2** *(Falso) La t de Student también es simétrica alrededor de la media=0, con forma de campana, pero con colas más pesadas, es decir, es más probable tener más observaciones más allá de 2 desvíos estándar respecto de la media si se la compara con la distribución Normal estándar.*

3. La $t$-student tiene un solo parámetro, llamado grados de libertad. (1 pto.)

   **Solución 1.3** *(Verdad) La t de Student tiene un solo parámetro, llamado grados de libertad, que determina cuan pesadas son las colas de la distribución.*

4. Si los estimadores $\hat{\theta}_1$ y $\hat{\theta}_2$ son insesgados, entonces podemos elegir cual de ellas es relativamente mas eficiente respecto a la otra haciendo uso del ratio de desviaciones estándar de los estimadores. (1 pto.)

   **Solución 1.4** *(Falso) Si ambos estimadores son insesgados, $\hat{\theta}_1$ es relativamente más eficiente que $\hat{\theta}_2$ si $V(\hat{\theta}_2) > V(\hat{\theta}_1)$. De hecho, se usa el ratio $V(\hat{\theta}_2)/V(\hat{\theta}_1)$ para definir la eficiencia relativa de dos estimadores insesgados.*

5. Un estimador consistente es aquel cuyo error de medida o sesgo se aproxima a cero cuando el tamaño de la muestra tiende a infinito. (1 pto.)

   **Solución 1.5** *(Verdad) Se dice que el estimador $\hat{\theta}_n$ es un estimador consistente de $\theta$ si, para cualquier número positivo $\epsilon > 0$,*

   $$\lim_{n\to\infty} P(|\underbrace{\hat{\theta}_n - \theta}_{sesgo}| \leq \epsilon) = 1.$$

6. Se dice que un estimador es suficiente cuando resume toda la información relevante contenida en la muestra, de forma que ningún otro estimador pueda proporcionar información adicional sobre el parámetro desconocido de la población. (1 pto.)

   **Solución 1.6** *(Verdad) Sea $Y_1$, $Y_2$,$\cdots$, $Y_n$ una muestra aleatoria de una distribución de probabilidad con parámetro desconocido $\theta$. Entonces se dice que el estadístico $U = g(Y_1, Y_2, \cdots, Y_n)$ es suficiente para $\theta$ si la distribución condicional de $Y_1$, $Y_2$,$\cdots$, $Y_n$ dada $U$, no depende de $\theta$. Es decir, una vez que se conozca $U$, ninguna otra función de $Y_1$, $Y_2$,$\cdots$, $Y_n$ proporcionará más información sobre el posible valor de $\theta$. En este sentido, $U$ contiene toda la información acerca de $\theta$. Por tanto, se dice que el estadístico $U$ es suficiente para $\theta$.*

7. El Teorema de Rao-Blackwell es una técnica para mejorar el estimador insesgado inicial $h(\theta)$. (1 pto.)

   **Solución 1.7** *(Verdad) Si $U$ es un estadístico suficiente para $h(\theta)$, se define $\widehat{h(\theta)}^* = E[\widehat{h(\theta)}|U]$. Entonces, para todo $h(\theta)$,*

   $$E[\widehat{h(\theta)}^*] = h(\theta) \quad y \quad Var[\widehat{h(\theta)}^*] \leq Var[\widehat{h(\theta)}]$$

   *Es decir, una vez que aplica el teorema de Rao-Blackwell se obtiene un mejor estimador insesgado para $h(\theta)$.*

## 2 Aplicaciones

1. Sea $Y_1, Y_2, \cdots, Y_n$ una muestra aleaoria de una distribución donde $P(Y_i = 1) = p$ y $P(Y_i = 0) = 1 - p$, con $p$ desconocida (es frecuente que tales variables aleatorias se denominen variables de Bernoulli). Use el criterio de factorización para hallar un estadístico suficiente que mejor resuma los datos. Proporcione un MVUE (Estimador insesgado de varianza mínima) para $p$. (5 ptos.)
   Sugerencia: Use la verosmilitud $L(y_1, y_2, \cdots, y_n|p)$

   **Solución 2.1** *Observe que la función de probabilidad anterior se puede escribir como*

$$P(Y_i = y_i) = p^{y_i}(1-p)^{1-y_i}, \qquad y_i = 0, 1$$

   *Por tanto, la verosimilitud $L(p)$ es*

$$\begin{aligned} L(y_1, y_2, \cdots, y_n|p) &= p(y_1, y_2, \cdots, y_n|p) \\ &= p^{y_1}(1-p)^{1-y_1} \times p^{y_2}(1-p)^{1-y_2} \times \cdots \times p^{y_n}(1-p)^{1-y_n} \\ &= \underbrace{p^{\sum y_i}(1-p)^{1-\sum y_i}}_{g(\sum y_i, p)} \times \underbrace{1}_{h(y_1, y_2 \cdots, y_n)} \end{aligned}$$

   *De acuerdo con el criterio de factorización, $U = \sum_{i=1}^n Y_i$ es suficiente para $p$. Este estadístico resume mejor la información acerca del parámetro $p$. Observe que $E(U) = np$, o bien, de la misma manera, $E(U/n) = p$. Así $U/n = Y$ es un estimador insesgado para $p$. Como este estimador es una función del estadístico suficiente $\sum_{i=1}^n Y_i$, el estimador $\hat{p} = \bar{Y}$ es el MVUE para $p$.*

2. Suponga que $Y_1, Y_2, \cdots, Y_n$ es una muestra aleatoria de una función de densidad de probabilidad de la familia exponencial(un parámetro) de modo que

$$f(y|\theta) = \begin{cases} a(\theta)b(y)e^{-[c(\theta)d(y)]}, & a \leq y \leq b; \\ 0, c.c. \end{cases}$$

   Donde $a$ y $b$ no dependen de $\theta$. Demuestre que $\displaystyle\sum_{i=1}^n d(Y_i)$ es suficiente para $\theta$. (5 ptos.)
   Suegerencia: Use la función de verosimilitud $L(\cdot)$ y aplique el criterio de factorización.

   **Solución 2.2** *La función de verosimilitud es*

$$L(\theta) = \prod_{i=1}^n f(y_i|\theta) = [a(\theta)]^n [\prod_{i=1}^n b(y_i)] e^{[-c(\theta)\sum_{i=1}^n d(y_i)]}$$

   *Así, $U = \sum_{i=1}^n d(Y_i)$ es suficiente para $\theta$ porque por el criterio de factorización $L(\theta)$ puede ser factorizado, donde $u = \sum_{i=1}^n d(y_i)$, $g(u, \theta) = [a(\theta)]^n e^{[-c(\theta)u]}$ y $h(y) = \prod_{i=1}^n b(y_i)$.*

3. Sea $Y_1, Y_2, \cdots, Y_n$ una muestra aleaoria de la función de densidad de probabilidad

$$f(y) = \begin{cases} \theta y^{\theta - 1}, & 0 < y < 1; \\ 0, c.c. \end{cases}$$

   donde $\theta > 0$. Demuestre que $\bar{Y}$ es un estimador consistente de $\frac{\theta}{\theta+1}$. (5 ptos.)

   **Solución 2.3** *Dado $f(y)$, tenemos que*

$$E[Y] = \frac{\theta}{\theta + 1} \quad y \quad Var[Y] = \frac{\theta}{(\theta+2)(\theta+1)^2}$$

   *Es fácil observar que $Y$ tiene distribución beta con parámetros $alpha = \theta$ y $\beta = 1$. Luego,*

$$E[\bar{Y}] = \frac{\theta}{\theta + 1} \quad y \quad Var[\bar{Y}] = \frac{\theta}{n(\theta+2)(\theta+1)^2}$$

   *Así,*

$$\lim_{n \to \infty} Var[\bar{Y}] = 0$$

   *Por lo tanto, $\bar{Y}$ es consistente.*

4. Sean $Y_1, Y_2, \cdots, Y_n$ que denotan una muestra aleatoria de una población con media $\mu$ y varianza $\sigma^2$. Considere los siguientes tres estimadores para $\mu$:

$$\hat{\mu}_1 = \frac{1}{2}(Y_1 + Y_2), \qquad \hat{\mu}_2 = \frac{1}{4}Y_1 + \frac{Y_2 + \cdots + Y_{n-1}}{2(n-2)} + \frac{1}{4}Y_n \qquad , \hat{\mu}_3 = \bar{Y}$$

- Demuestre que cada uno de los tres estimadores es insesgado. (2.5 ptos.)
- Encuentre la eficiencia de $\hat{\mu}_3$ con respecto a $\hat{\mu}_2$ y $\hat{\mu}_1$, respectivamente. (2.5 ptos.)

**Solución 2.4** *Sean $Y_1, Y_2, \cdots, Y_n$ que denotan una muestra aleatoria de una población con media $\mu$ y varianza $\sigma^2$.*

- *Veamos que los estimadores son insesgados*

$$\begin{aligned} E[\hat{\mu}_1] &= \frac{1}{2}\big(E[Y_1] + E[Y_2]\big) = \frac{1}{2}(\mu + \mu) = \mu \\ E[\hat{\mu}_2] &= \frac{\mu}{4} + \frac{(n-2)\mu}{2(n-2)} + \frac{\mu}{4} = \mu \\ E[\hat{\mu}_3] &= E[\bar{Y}] = \mu \end{aligned}$$

- *Las varianzas de los tres estimadores son*

$$\begin{aligned} Var[\hat{\mu}_1] &= \frac{1}{4}\big(\sigma^2 + \sigma^2\big) = \frac{\sigma^2}{2} \\ Var[\hat{\mu}_2] &= \frac{\sigma^2}{61} + \frac{(n-2)\sigma^2}{4(n-2)^2} + \frac{\sigma^2}{16} = \frac{\sigma^2}{8} + \frac{\sigma^2}{4(n-2)} \\ Var[\hat{\mu}_3] &= \frac{\sigma^2}{n} \end{aligned}$$

*Luego,*

$$eff(\hat{\mu}_3, \hat{\mu}_2) = \frac{n^2}{8(n-2)} \quad y \quad eff(\hat{\mu}_3, \hat{\mu}_1) = \frac{n}{2}$$

5. Un experimentador desea comprobar la variabilidad de mediciones obtenidas al usar equipo diseñado para medir el volumen de una fuente de audio. Tres mediciones independientes registradas por este equipo para la misma fuente de sonido fueron 4.1, 5.2 y 10.2. Estime $\sigma^2$ con coeficiente de confianza .90.
Sugerencia: $\chi^2_{0.05} = 5.991$ y $\chi^2_{0.95} = 0.103$. (3 ptos.)

**Solución 2.5** *Para los datos dados, $s^2 = 10,57$. Con $\frac{\alpha}{2} = 0,05$ y $(n-1) = 2$ grados de libertad.Entonces, el intervalo de confianza de 90% para $\sigma^2$ es*

$$\left(\frac{(n-1)s^2}{\chi^2_{0,05}}; \frac{(n-1)s^2}{\chi^2_{0,95}}\right) equiv \left(\frac{(2)(10,57)}{5,991}; \frac{(2)(10,57)}{0,103}\right) \equiv (3,53; 205,24)$$

*Observe que este intervalo para $\sigma^2$ es muy ancho, principalmente porque $n$ es muy pequeña.*

6. Un fabricante ha inventado una nueva pólvora que fue probada en ocho proyectiles. Las velocidades resultantes en la boca del cañón, en pies por segundo, fueron las siguientes:

| | | | |
|---|---|---|---|
| 3005 | 2925 | 2935 | 2965 |
| 2995 | 3005 | 2937 | 2905 |

Encuentre un intervalo de confianza de 95% para el verdadero promedio de velocidad $\mu$ para proyectiles de este tipo. Suponga que las velocidades en la boca del cañón están distribuidas normalmente en forma aproximada. (3 ptos.)
Sugerencia: $t_{0.25} = 2.365$

**Solución 2.6** *Si suponemos que las velocidades $Y_i$ están distribuidas normalmente, el intervalo de confianza para $\mu$ es*

$$\bar{Y} \pm t_{\frac{\alpha}{2}}\left(\frac{S}{\sqrt{n}}\right),$$

*donde $t_{\frac{\alpha}{2}}$ está determinado por un $n-1$ grados de libertad. Para los datos dados, $\bar{y} = 2959$ y $s = 39,1$. En este ejemplo tenemos $n-1 = 7$ grados de libertad. Entonces, obtenemos*

$$2959 \pm 3,365\left(\frac{39,1}{\sqrt{8}}\right) \equiv 2959 \pm 32,7$$

*como el intervalo de confianza observado para $\mu$.*

Los profesores.[1]
October 13, 2019

[1] Hecho en $\LaTeX$