Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Primer Semestre 2016

Curso	**:**	**Probabilidad y Estadística**
Sigla	**:**	**EYP1113**
Pauta	**:**	**I3**
Profesores	**:**	**Ricardo Aravena C. y Ricardo Olea O.**

Problema 1

Durante el último tiempo, Chile ha estado sufriendo por la floración de algas nocivas (conocida como marea roja y que ha afectado a Chile más de un par de veces, muchas en forma inocua y otras con consecuencias catastróficas - de hecho, pudo haber sido la causa de la mortalidad masiva de Colonos en el siglo XVI en Magallanes, caso conocido como Puerto de Hambre). Suponga que a objeto de determinar el nivel de contaminación sobre algunos bivalvos (por ejemplo, almejas) un sector de Chiloé (por ejemplo, Dalcahue) usted toma una muestra de 120 bivalvos determinando que solo 24 de ellos muestran niveles altos. Más aún, un análisis especial realizado sobre estos 24 muestra que en promedio alcanzan un nivel de 0,33 mg con una desviación estándar de 0,08 mg de la toxina paralizante y que el nivel se comporta como una distribución normal.

(a) **[2.0 Ptos.]** Si la autoridad ha determinado que si menos del 25 % del marisco analizado presenta contaminación alta, el producto puede ser consumir con cuidado. En caso contrario, se debe declarar una veda absoluta. Determine el valor-p. Para un nivel de significancia del 10 %, ¿cuál debe ser la decisión de la autoridad?

(b) **[2.0 Ptos.]** ¿Existe evidencia que permita afirmar que, entre los mariscos contaminados, el nivel de la toxina paralizante es superior a 0,30 mg (situación que lo hace muy peligroso)? (Use $\alpha = 5\,\%$).

(c) **[2.0 Ptos.]** Si repite (b) pero con desviación estándar conocida e igual a 0,10 mg, ¿cambia la conclusión anterior?

Solución

(a) Se pide realizar la siguiente prueba de hipótesis

$$\mathrm{H}_0 : p = p_0 \quad \text{vs} \quad \mathrm{H}_a : p < p_0 \quad \textbf{[0.3 Ptos.]}$$

Bajo H_0 se tiene que

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0\,(1-p_0)}{n}}} \overset{\text{aprox}}{\sim} \text{Normal}(0, 1) \quad \textbf{[0.5 Ptos.]}$$

con $n = 120$ y $\hat{p} = \frac{24}{120}$. **[0.2 Ptos.]**

El valor-p para esta prueba está dado por

[0.4 Ptos.] $\text{valor-p} = \Phi(Z_0) = \Phi(-1{,}264911) = 1 - \Phi(1{,}264911) \approx 1 - \Phi(1{,}27) = 0{,}102 > 0{,}10 = \alpha$ **[0.3 Ptos.]**

Es decir, no existe suficiente evidencia para levantar la veda absoluta. **[0.3 Ptos.]**

(b) Se pide realizar la siguiente prueba de hipótesis

$$\mathrm{H}_0 : \mu = \mu_0 \quad \text{vs} \quad \mathrm{H}_a : \mu > \mu_0 \quad \textbf{[0.3 Ptos.]}$$

Bajo H_0 se tiene que

$$T_0 = \frac{\overline{X}_n - \mu_0}{S/\sqrt{n}} \sim \text{t-Student}(n-1) \quad \textbf{[0.5 Ptos.]}$$

con $n = 24$, $\overline{X}_n = 0{,}33$ y $S = 0{,}08$. **[0.4 Ptos.]**

Alternativa 1: El valor-p para esta prueba está dado por

$$\textbf{[0.3 Ptos.]} \quad \text{valor-p} = P(T > T_0) = P(T > 1{,}837117) \to 2{,}5\,\% < \text{valor-p} < 5\,\% = \alpha \quad \textbf{[0.3 Ptos.]}$$

Es decir, existe suficiente para apoyar que el nivel de la toxina paralizante es superior a 0,30 mg. **[0.2 Ptos.]**

Alternativa 2: El estadístico de prueba es

$$\textbf{[0.3 Ptos.]} \quad T_0 = 1{,}837117 > 1{,}714 = t_{1-0{,}05}(23) \quad \textbf{[0.3 Ptos.]}$$

Es decir, existe suficiente para apoyar que el nivel de la toxina paralizante es superior a 0,30 mg. **[0.2 Ptos.]**

(c) Se pide realizar la siguiente prueba de hipótesis

$$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_a : \mu > \mu_0 \quad \textbf{[0.3 Ptos.]}$$

Bajo H_0 se tiene que

$$Z_0 = \frac{\overline{X}_n - \mu_0}{\sigma/\sqrt{n}} \sim \text{Normal}(0,1) \quad \textbf{[0.5 Ptos.]}$$

con $n = 24$, $\overline{X}_n = 0{,}33$ y $\sigma = 0{,}10$. **[0.4 Ptos.]**

Alternativa 1: El valor-p para esta prueba está dado por

$$\textbf{[0.3 Ptos.]} \quad \text{valor-p} = 1 - \Phi(Z_0) = 1 - \Phi(1{,}469694) \approx 1 - \Phi(1{,}47) = 0{,}0708 > 0{,}05 = \alpha \quad \textbf{[0.3 Ptos.]}$$

Es decir nuestra conclusión en (b) cambia, ya que NO existe suficiente para apoyar que el nivel de la toxina paralizante es superior a 0,30 mg. **[0.2 Ptos.]**

Alternativa 2: El estadístico de prueba es

$$\textbf{[0.3 Ptos.]} \quad Z_0 = 1{,}469694 < 1{,}645 = k_{1-0{,}05} \quad \textbf{[0.3 Ptos.]}$$

Es decir nuestra conclusión en (b) cambia, ya que NO existe suficiente para apoyar que el nivel de la toxina paralizante es superior a 0,30 mg. **[0.2 Ptos.]**

+ 1 Punto Base

Problema 2

En la construcción en altura un factor predeterminante es el efecto del viento sobre la estructura, entre otros tales como rigidez, volcamiento, etc. La acción del viento, debido a la presión que provoca sobre la estructura plana, produce oscilaciones. Por lo cual es necesario tomar decisiones sobre la altura y tipo de estructura que puede realizarse en un sector dado (ej. Dunas de Montemar) afectado por vientos. Si se define la presión básica media Q como el producto entre k y la velocidad media al cuadrado, donde k es una constante conocida e igual a 1/16. Interesa determinar si hay evidencia estadística que permita determina que la presión básica media supera los 100 kg/m^2, situación que haría inviable la construcción del edificio de 40 mts. (norma NCH) de altura proyectada. Para tomar la decisión se dispone de una muestra de 60 velocidades de viento registradas en diferentes días, estos datos entregan la siguientes resultados: $\sum_{i=1}^n X_i = 1060$, $\sum_{i=1}^n \ln(X_i) = 174$. Si se pueden asumir que estas velocidades provienen de una distribución Log-Normal(λ, ζ) con ζ conocido e igual a $\sqrt{2}$. ¿Cuál es su decisión para un nivel de significancia del 2 %?

Solución

El estimador máximo verosímil de λ es

$$\textbf{[1.0 Ptos.]} \quad \hat{\lambda} = \frac{1}{n}\sum_{i=1}^n \ln(X_i) \sim \text{Normal}\left(\lambda,\, \sqrt{\frac{2}{n}}\right) \quad \textbf{[1.0 Ptos.]}$$

Del enunciado se tiene que

$$Q(\lambda) = \frac{1}{16} \cdot e^{2\lambda} \cdot e^2 \quad \textbf{[0.5 Ptos.]}$$

y por invarianza de los EMV

$$\widehat{Q(\lambda)} = Q(\hat{\lambda}) \overset{\text{aprox.}}{\sim} \text{Normal}\left(Q(\lambda),\, \sqrt{\frac{2\cdot[2\,Q(\lambda)]^2}{n}}\right) \quad \textbf{[0.5 Ptos.]}$$

Se pide

$$\text{H}_0 : Q(\lambda) = Q_0 \quad \text{vs} \quad \text{H}_a : Q(\lambda) > Q_0 \quad \textbf{[0.5 Ptos.]}$$

con $Q_0 = 100$ kg/m^2.

Bajo H_0 se tiene que

$$Z_0 = \frac{\hat{Q} - Q_0}{2\,Q_0\sqrt{\frac{2}{n}}} \overset{\text{aprox}}{\sim} \text{Normal}(0,1) \quad \textbf{[0.5 Ptos.]}$$

con

$$\widehat{Q(\lambda)} = \frac{e^2}{16}\cdot \exp\left[\frac{2}{n}\sum_{i=1}^n \ln(X_i)\right] = 152{,}5376 \quad \textbf{[0.5 Ptos.]}$$

Alternativa 1: El valor-p para esta prueba está dado por

$$\textbf{[0.5 Ptos.]} \quad \text{valor-p} = 1-\Phi(Z_0) = 1-\Phi(1{,}438801) \approx 1-\Phi(1{,}44) = 1-{,}9251 = 0{,}0749 > 0{,}02 = \alpha \quad \textbf{[0.5 Ptos.]}$$

Es decir, no existe suficiente evidencia que permita determina que la presión básica media supera los 100 kg/m^2 y por ende la construcción de un edificio de 40 mts. cumpliría con la norma NCH. **[0.5 Ptos.]**

Alternativa 2: El estadístico de prueba es

$$\textbf{[0.5 Ptos.]} \quad Z_0 = 1{,}438801 < 2{,}055 \approx k_{1-0{,}02} \quad \textbf{[0.5 Ptos.]}$$

Es decir, no existe suficiente evidencia que permita determina que la presión básica media supera los 100 kg/m^2 y por ende la construcción de un edificio de 40 mts. cumpliría con la norma NCH. **[0.5 Ptos.]**

+ 1 Punto Base

Problema 3

La encuesta nacional de salud 2009-2010 entrego muchos resultados relevantes en temas de salud. Por ejemplo, nos informa que los mayores de 15 años tienen un peso medio igual a 72,3 kilos y una mediana igual a 71,0 kilos, una talla media de 1,657 metros y una mediana de 1,630 metros. Suponga que el peso y talla se comportan como variables aleatorias independientes con distribución Log-Normal. Si el indice de masa corporal de una persona es igual al peso/talla^2, entregue una valor aproximado de 1er orden para la esperanza, varianza y coeficiente de variación del indice de masa corporal de los mayores de 15 años.

Solución

Definamos como X e Y, al peso y talla respectivamente.

Del enunciado

$$X \sim \text{Log-Normal}(\lambda_X, \zeta_X) \quad \text{e} \quad Y \sim \text{Log-Normal}(\lambda_Y, \zeta_Y)$$

con

$$\mu_X = 72{,}3, \quad \lambda_X = \ln(71{,}0), \quad \mu_Y = 1{,}657, \quad \lambda_Y = \ln(1{,}630) \quad \textbf{[0.8 Ptos.]}$$

y a partir del formulario

$$\sigma_X^2 = \mu_X^2\,(e^{\zeta_X^2} - 1) \quad \text{y} \quad \sigma_Y^2 = \mu_Y^2\,(e^{\zeta_Y^2} - 1) \quad \textbf{[0.8 Ptos.]}$$

con

$$\zeta_X^2 = 2\ln(\mu_X) - 2\lambda_X \quad \text{y} \quad \zeta_Y^2 = 2\ln(\mu_Y) - 2\lambda_Y \quad \textbf{[0.8 Ptos.]}$$

Sea

$$Z = \frac{X}{Y^2} = g(X, Y) \quad \textbf{[0.3 Ptos.]}$$

donde su aproximación de primer orden está dada por:

$$Z = g(X, Y) \approx \frac{\mu_X}{\mu_Y^2} + (X - \mu_X)\,\frac{1}{\mu_Y^2} + (Y - \mu_Y)\,\frac{(-2\,\mu_X)}{\mu_Y^3} \quad \textbf{[1.0 Ptos.]}$$

Aplicando operador esperanza

$$\textbf{[0.5 Ptos.]} \quad \mu_Z = \text{E}(Z) \approx \frac{\mu_X}{\mu_Y^2} = 26{,}33257 \quad \textbf{[0.5 Ptos.]}$$

y aplicando varianza (considerando la indepencia entre X e Y) tenemos

$$\textbf{[0.5 Ptos.]} \quad \sigma_Z^2 = \text{Var}(Z) \approx \frac{\sigma_X^2}{\mu_Y^4} + \frac{4\,\mu_X^2\,\sigma_Y^2}{\mu_Y^6} = 118{,}2725 \quad \textbf{[0.5 Ptos.]}$$

Finalmente

$$\delta_Z = \frac{\sigma_Z}{\mu_Z} \approx 0{,}4129987 \quad \textbf{[0.3 Ptos.]}$$

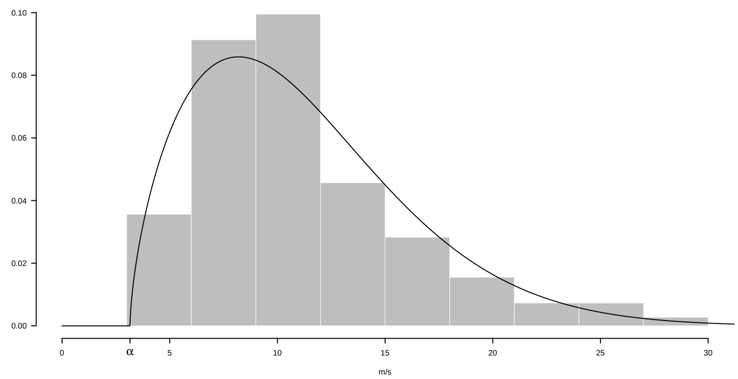
+ 1 Punto Base

Problema 4

Un parque eólico es una central eléctrica donde la producción de la energía eléctrica se consigue a partir de la fuerza del viento, mediante aerogeneradores que aprovechan las corrientes de aire. El viento es un efecto derivado del calentamiento desigual de la superficie de la Tierra por el Sol y el principal problema es la incertidumbre respecto a la disponibilidad de viento cuando se necesita. La siguiente figura, muestra el comportamiento de la máxima velocidad diaria durante el año 2010 en el sector donde actualmente se construye el Parque Eólico Piñón Blanco, IX Región, las cuales fueron ajustadas por un modelo Weibull trasladado definido como:

$$f(x) = \frac{\beta}{\eta}\left[\frac{(x-\alpha)}{\eta}\right]^{\beta-1}\exp\left\{-\left[\frac{(x-\alpha)}{\eta}\right]^{\beta}\right\}, \quad F(x) = 1 - \exp\left[-\left(\frac{x-\alpha}{\eta}\right)^{\beta}\right]$$

con $x \geq \alpha$, $\alpha > 0$, $\eta > 0$ y $\beta > 0$. Para $r > 0$ se tiene que $\mathrm{E}[(X-\alpha)^r] = \eta^r\,\Gamma(1 + r/\beta)$.



Suponga que β y η son parámetros conocidos, y para la estimación del parámetro α se tienen dos alternativas: $\widehat{\alpha}$ (estimador de momento) y $\widetilde{\alpha}$ (mínimo entre las mediciones).

(a) **[2.0 Ptos.]** Determine el estimador de momentos $\widehat{\alpha}$ de α y calcule el error cuadrático medio.

(b) **[3.0 Ptos.]** Calcule el error cuadrático medio de $\widetilde{\alpha}$.

(c) **[1.0 Ptos.]** ¿Cuál de los dos estimadores recomendaría? Justifique su respuesta.

Solución

(a) Del enunciado se tiene que

$$E(X) = \alpha + \eta\,\Gamma(1 + 1/\beta)$$

igualando a $\overline{X}_n$ se tiene que el estimador de momentos de α es

$$\hat{\alpha} = \overline{X}_n - \eta\,\Gamma(1 + 1/\beta) \quad \textbf{[1.0 Ptos.]}$$

Aplicando operador esperanza se tiene que

$$\mathrm{E}(\hat{\alpha}) = \mathrm{E}(\overline{X}_n) - \eta\,\Gamma(1+1/\beta) = \alpha + \eta\,\Gamma(1+1/\beta) - \eta\,\Gamma(1+1/\beta) = \alpha \quad \textbf{[0.5 Ptos.]}$$

es decir, es un estimador insesgado para α.

Luego su ECM, asumiendo independencia, está dada por

$$\mathrm{ECM}(\hat{\alpha}) = \mathrm{Var}(\hat{\alpha}) = \mathrm{Var}(\overline{X}_n) = \frac{\eta^2}{n}\left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right)\right] \quad \textbf{[0.5 Ptos.]}$$

(b) Tenemos que la distribución de $\widetilde{\alpha}$ está dada por

$$f_{\widetilde{\alpha}}(x) = \frac{\beta}{\eta\, n^{-1/\beta}} \left[\frac{(x-\alpha)}{\eta\, n^{-1/\beta}}\right]^{\beta-1} \exp\left\{-\left[\frac{(x-\alpha)}{\eta\, n^{-1/\beta}}\right]^{\beta}\right\} \quad \textbf{[1.0 Ptos.]}$$

que corresponde a una Weibull($\eta\, n^{-1/\beta}$, β) trasladada en α. **[0.5 Ptos.]**

Del enunciado se tiene que

$$\mathrm{E}(\widetilde{\alpha}) = \alpha + \frac{\eta}{n^{1/\beta}}\,\Gamma(1+1/\beta) \quad \textbf{[0.5 Ptos.]}$$

y

$$\mathrm{Var}(\widetilde{\alpha}) = \frac{\eta^2}{n^{2/\beta}}\left[\Gamma\left(1+\frac{2}{\beta}\right) - \Gamma^2\left(1+\frac{1}{\beta}\right)\right] \quad \textbf{[0.5 Ptos.]}$$

Luego

$$\mathrm{ECM}(\widetilde{\alpha}) = \frac{\eta^2}{n^{2/\beta}}\,\Gamma\left(1+\frac{2}{\beta}\right) \quad \textbf{[0.5 Ptos.]}$$

(c) Se recomienda el estimador más eficiente en base al que tiene menor ECM.

$$\frac{\mathrm{ECM}(\hat{\alpha})}{\mathrm{ECM}(\widetilde{\alpha})} = n^{2/\beta-1}\left[1 - \Gamma^2\left(1+\frac{1}{\beta}\right)\Big/\Gamma\left(1+\frac{2}{\beta}\right)\right] \quad \textbf{[0.4 Ptos.]}$$

- Para $\beta < 2$:

  $$\frac{\mathrm{ECM}(\hat{\alpha})}{\mathrm{ECM}(\widetilde{\alpha})} \to +\infty, \qquad \text{cuando } n \to +\infty$$

  $\widetilde{\alpha}$ es más eficiente **[0.3 Ptos.]**

- Para $\beta > 2$:

  $$\frac{\mathrm{ECM}(\hat{\alpha})}{\mathrm{ECM}(\widetilde{\alpha})} \to 0, \qquad \text{cuando } n \to +\infty$$

  $\hat{\alpha}$ es más eficiente **[0.3 Ptos.]**

+ 1 Punto Base

Formulario

Propiedades función $\Gamma(\cdot)$

$$(1)\quad \Gamma(k) = \int_0^\infty u^{k-1}\, e^{-u}\, du; \quad (2)\quad \Gamma(a+1) = a\,\Gamma(a);$$

$$(3)\quad \Gamma(n+1) = n!, \quad \text{si } n \in \mathbb{N}_0; \quad (4)\quad \Gamma(1/2) = \sqrt{\pi}$$

Propiedades función $B(\cdot,\,\cdot)$

$$(1)\quad B(q,\,r) = \int_0^1 x^{q-1}\,(1-x)^{r-1}\,dx; \quad (2)\quad B(q,\,r) = \frac{\Gamma(q)\,\Gamma(r)}{\Gamma(q+r)}$$

Igualdades

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k\, b^{n-k}, \qquad \sum_{k=x}^{\infty} \phi^k = \frac{\phi^x}{1-\phi} \quad \text{si } |\phi| < 1, \qquad \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \exp(\lambda)$$

Propiedad distribución Gamma

$$\text{Si } T \sim \text{Gamma}(k,\,\nu), \text{ con } k \in \mathbb{N} \longrightarrow F_T(t) = 1 - \sum_{x=0}^{k-1} \frac{(\nu\,t)^x\, e^{-\nu\,t}}{x!}$$

Transformación

Sea $Y = g(X)$ una función cualquiera, con k raíces:

$$f_Y(y) = \sum_{i=1}^{k} f_X\left(g_i^{-1}(y)\right) \cdot \left| \frac{d}{dy}\, g_i^{-1}(y) \right|$$

$$p_Y(y) = \sum_{i=1}^{k} p_X\left(g_i^{-1}(y)\right)$$

Sea $Z = g(X, Y)$ una función cualquiera:

$$p_Z(z) = \sum_{g(x,y)=z} p_{X,Y}(x, y)$$

Sea $Z = g(X, Y)$ una función invertible para X o Y fijo:

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_{X,Y}(g^{-1}, y) \left| \frac{\partial}{\partial z} g^{-1} \right| dy \\ &= \int_{-\infty}^{\infty} f_{X,Y}(x, g^{-1}) \left| \frac{\partial}{\partial z} g^{-1} \right| dx \end{aligned}$$

Esperanza y Varianza Condicional

$$\text{E}(Y) = \text{E}[\text{E}(Y \mid X)] \quad \text{y} \quad \text{Var}(Y) = \text{Var}[\text{E}(Y \mid X)] + \text{E}[\text{Var}(Y \mid X)]$$

Teorema del Límite Central

Sean $X_1, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas, entonces

$$Z_n = \frac{\displaystyle\sum_{i=1}^{n} X_i - n \cdot \mu}{\sqrt{n}\,\sigma} = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \longrightarrow Z \sim \text{Normal}(0, 1),$$

cuando $n \to \infty$, $\text{E}(X_i) = \mu$ y $\text{Var}(X_i) = \sigma^2$.

Propiedades Esperanza, Varianza y Covarianza

Sean $X_1,\, X_2,\, \ldots,\, X_n,\, Y_1,\, Y_2,\, \ldots,\, Y_m$ variables aleatorias y $a_0,\, a_1,\, \ldots,\, a_n,\, b_0,\, b_1,\, \ldots,\, b_m$ constantes conocidas.

- $\mathrm{E}\left(a_0 + \sum_{i=1}^{n} a_i \cdot X_i\right) = a_0 + \sum_{i=1}^{n} a_i \cdot \mathrm{E}(X_i)$.
- $\mathrm{Cov}\left(a_0 + \sum_{i=1}^{n} a_i \cdot X_i,\, b_0 + \sum_{j=1}^{m} b_j \cdot Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i \cdot b_j \cdot \mathrm{Cov}\,(X_i, Y_j)$.
- $\mathrm{Var}\left(a_0 + \sum_{i=1}^{n} a_i \cdot X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i \cdot a_j \cdot \mathrm{Cov}\,(X_i, X_j)$.
- Si $X_1, \ldots, X_n$ son variables aleatorias independientes, entonces

$$\mathrm{Var}\left(a_0 + \sum_{i=1}^{n} a_i \cdot X_i\right) = \sum_{i=1}^{n} a_i^2 \cdot \mathrm{Var}\,(X_i)$$

Mínimo y Máximo

Sean $X_1, \ldots, X_n$ variables aleatorias continuas independientes con idéntica distribución (f_X y F_X), entonces para:

$$Y_1 = \min\{X_1, \ldots, X_n\} \longrightarrow f_{Y_1} = n\,[1 - F_X(y)]^{n-1}\; f_X(y)$$
$$Y_n = \max\{X_1, \ldots, X_n\} \longrightarrow f_{Y_n} = n\,[F_X(y)]^{n-1}\; f_X(y)$$

Mientras que la distribución conjunta entre Y_1 e Y_n está dada por:

$$f_{Y_1,Y_n}(u,v) = n\,(n-1)\;[F_X(v) - F_X(u)]^{n-2}\; f_X(v)\, f_X(u), \qquad u \leq v$$

Aproximación de Momentos

Sean $X_1, \ldots, X_n$ variables aleatorias con valores esperados $\mu_{X_1}, \ldots, \mu_{X_n}$ y varianzas $\sigma^2_{X_1}, \ldots, \sigma^2_{X_n}$ e Y una función de ellas. La aproximación de primer orden está dada por

$$Y \approx g(\mu_{X_1}, \ldots, \mu_{X_n}) + \sum_{i=1}^{n} (X_i - \mu_{X_i})\,\frac{\partial}{\partial X_i} g(\mu_{X_1}, \ldots, \mu_{X_n})$$

Estimador Máximo Verosímil

Sea $X_1,\ldots,X_n$ una muestra aleatoria independiente e idénticamente distribuida con función de probabilidad p_X o de densidad f_X, determinada por un parámetro θ. Si $\hat{\theta}$ es el estimador máximo verosímil del parámetro θ, entonces:

- $\mathrm{E}(\hat{\theta}) \to \theta$, cuando $n \to \infty$.
- $\mathrm{Var}(\hat{\theta}) = \dfrac{1}{I_n(\theta)}$, con $I_n(\theta) = -\mathrm{E}\left[\dfrac{\partial^2}{\partial\,\theta^2}\,\ln L(\theta)\right]$.
- $\hat{\theta} \,\dot{\sim}\, \text{Normal}\left(\theta,\, \sqrt{\dfrac{1}{I_n(\theta)}}\right)$, cuando $n \to \infty$.
- El estimador máximo verosímil de $g(\theta)$ es $g(\hat{\theta})$, cuya varianza está dada por: $\mathrm{Var}[g(\hat{\theta})] = \dfrac{[g'(\theta)]^2}{I_n(\theta)}$.

Error Cuadrático Medio

El error cuadrático medio de un estimador $\hat{\theta}$ de θ se define como:

$$\mathrm{ECM}(\hat{\theta}) = \mathrm{E}\left\{\left[\hat{\theta} - \mathrm{E}(\hat{\theta})\right]^2\right\} = \mathrm{Var}(\hat{\theta}) + \text{Sesgo}^2$$

Distribuciones Muestrales

Sean $X_1,\ldots,X_n$ variables aleatorias independientes e idénticamente distribuidas Normal(μ, σ), entonces

$$\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim \text{Normal}(0,1), \qquad \frac{\overline{X}_n - \mu}{s/\sqrt{n}} \sim \text{t-student}(n-1), \qquad \frac{s^2\,(n-1)}{\sigma^2} \sim \chi^2_{n-1}$$

con $s^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$.

Distribución	Densidad de Probabilidad	Θ_X	Parámetros	Esperanza y Varianza
Binomial	$\binom{n}{x}\, p^x\,(1-p)^{n-x}$	$x = 0, \ldots, n$	$n,\ p$	$\mu_X = n\,p$ $\sigma_X^2 = n\,p\,(1-p)$ $M(t) = [p\,e^t + (1-p)]^n, \quad t \in \mathbb{R}$
Geométrica	$p\,(1-p)^{x-1}$	$x = 1, 2, \ldots$	p	$\mu_X = 1/p$ $\sigma_X^2 = (1-p)/p^2$ $M(t) = p\,e^t/[1-(1-p)\,e^t], \quad t < -\ln(1-p)$
Binomial-Negativa	$\binom{x-1}{r-1}\, p^r\,(1-p)^{x-r}$	$x = r, r+1, \ldots$	$r,\ p$	$\mu_X = r/p$ $\sigma_X^2 = r\,(1-p)/p^2$ $M(t) = \left\{p\,e^t/[1-(1-p)\,e^t]\right\}^r, \quad t < -\ln(1-p)$
Poisson	$\dfrac{(\nu\,t)^x\,e^{-\nu\,t}}{x!}$	$x = 0, 1, \ldots$	ν	$\mu_X = \nu\,t$ $\sigma_X^2 = \nu\,t$ $M(t) = \exp\left[\lambda\left(e^t - 1\right)\right], \quad t \in \mathbb{R}$
Exponencial	$\nu\,e^{-\nu\,x}$	$x \geq 0$	ν	$\mu_X = 1/\nu$ $\sigma_X^2 = 1/\nu^2$ $M(t) = \nu/(\nu - t), \quad t < \nu$
Gamma	$\dfrac{\nu^k}{\Gamma(k)}\, x^{k-1}\, e^{-\nu\,x}$	$x \geq 0$	$k,\ \nu$	$\mu_X = k/\nu$ $\sigma_X^2 = k/\nu^2$ $M(t) = [\nu/(\nu-t)]^k, \quad t < \nu$
Normal	$\dfrac{1}{\sqrt{2\,\pi}\,\sigma}\ \exp\left[-\dfrac{1}{2}\left(\dfrac{x-\mu}{\sigma}\right)^2\right]$	$-\infty < x < \infty$	$\mu,\ \sigma$	$\mu_X = \mu$ $\sigma_X^2 = \sigma^2$ $M(t) = \exp(\mu\,t + \sigma^2\,t^2/2), \quad t \in \mathbb{R}$
Log-Normal	$\dfrac{1}{\sqrt{2\,\pi}\,(\zeta\,x)}\ \exp\left[-\dfrac{1}{2}\left(\dfrac{\ln\,x - \lambda}{\zeta}\right)^2\right]$	$x \geq 0$	$\lambda,\ \zeta$	$\mu_X = \exp\left(\lambda + \dfrac{1}{2}\,\zeta^2\right)$ $\sigma_X^2 = \mu_X^2\left(e^{\zeta^2} - 1\right)$ $E(X^r) = e^{r\,\lambda}\,M_Z(r\,\zeta)$, con $Z \sim$Normal(0,1)
Uniforme	$\dfrac{1}{(b-a)}$	$a \leq x \leq b$	$a,\ b$	$\mu_X = (a+b)/2$ $\sigma_X^2 = (b-a)^2/12$ $M(t) = [e^{t\,b} - e^{t\,a}]/[t\,(b-a)], \quad t \in \mathbb{R}$
Beta	$\dfrac{1}{B(q,\ r)}\ \dfrac{(x-a)^{q-1}\,(b-x)^{r-1}}{(b-a)^{q+r-1}}$	$a \leq x \leq b$	$q,\ r$	$\mu_X = a + \frac{q}{q+r}\,(b-a)$ $\sigma_X^2 = \dfrac{q\,r\,(b-a)^2}{(q+r)^2\,(q+r+1)}$
Hipergeométrica	$\dfrac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$	$\max\{0, n+m-N\} \leq x \leq \min\{n, m\}$	$N,\ m,\ n$	$\mu_X = n\,\frac{m}{N}$ $\sigma_X^2 = \left(\frac{N-n}{N-1}\right)\,n\,\frac{m}{N}\,\left(1 - \frac{m}{N}\right)$

Tablas de Percentiles p

Distribución Normal Estándar k_p

k_p	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990

Distribución t-student $t_p(\nu)$

ν	$t_{0,90}$	$t_{0,95}$	$t_{0,975}$	$t_{0,99}$
1	3,078	6,314	12,706	31,821
2	1,886	2,920	4,303	6,965
3	1,638	2,353	3,182	4,541
4	1,533	2,132	2,776	3,747
5	1,476	2,015	2,571	3,365
6	1,440	1,943	2,447	3,143
7	1,415	1,895	2,365	2,998
8	1,397	1,860	2,306	2,896
9	1,383	1,833	2,262	2,821
10	1,372	1,812	2,228	2,764
11	1,363	1,796	2,201	2,718
12	1,356	1,782	2,179	2,681
13	1,350	1,771	2,160	2,650
14	1,345	1,761	2,145	2,624
15	1,341	1,753	2,131	2,602
16	1,337	1,746	2,120	2,583
17	1,333	1,740	2,110	2,567
18	1,330	1,734	2,101	2,552
19	1,328	1,729	2,093	2,539
20	1,325	1,725	2,086	2,528
21	1,323	1,721	2,080	2,518
22	1,321	1,717	2,074	2,508
23	1,319	1,714	2,069	2,500
24	1,318	1,711	2,064	2,492
25	1,316	1,708	2,060	2,485
26	1,315	1,706	2,056	2,479
27	1,314	1,703	2,052	2,473
28	1,313	1,701	2,048	2,467
29	1,311	1,699	2,045	2,462
30	1,310	1,697	2,042	2,457
∞	1,282	1,645	1,960	2,326

Distribución Chi-Cuadrado $c_p(\nu)$

ν	$c_{0,025}$	$c_{0,05}$	$c_{0,10}$	$c_{0,90}$	$c_{0,95}$	$c_{0,975}$	$c_{0,99}$	$c_{0,995}$
1	0,00	0,00	0,02	2,71	3,84	5,02	6,63	7,88
2	0,05	0,10	0,21	4,61	5,99	7,38	9,21	10,60
3	0,22	0,35	0,58	6,25	7,81	9,35	11,34	12,84
4	0,48	0,71	1,06	7,78	9,49	11,14	13,28	14,86
5	0,83	1,15	1,61	9,24	11,07	12,83	15,09	16,75
6	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55
7	1,69	2,17	2,83	12,02	14,07	16,01	18,48	20,28
8	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,95
9	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
10	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19
11	3,82	4,57	5,58	17,28	19,68	21,92	24,72	26,76
12	4,40	5,23	6,30	18,55	21,03	23,34	26,22	28,30
13	5,01	5,89	7,04	19,81	22,36	24,74	27,69	29,82
14	5,63	6,57	7,79	21,06	23,68	26,12	29,14	31,32
15	6,26	7,26	8,55	22,31	25,00	27,49	30,58	32,80
16	6,91	7,96	9,31	23,54	26,30	28,85	32,00	34,27
17	7,56	8,67	10,09	24,77	27,59	30,19	33,41	35,72
18	8,23	9,39	10,86	25,99	28,87	31,53	34,81	37,16
19	8,91	10,12	11,65	27,20	30,14	32,85	36,19	38,58
20	9,59	10,85	12,44	28,41	31,41	34,17	37,57	40,00
21	10,28	11,59	13,24	29,62	32,67	35,48	38,93	41,40
22	10,98	12,34	14,04	30,81	33,92	36,78	40,29	42,80
23	11,69	13,09	14,85	32,01	35,17	38,08	41,64	44,18
24	12,40	13,85	15,66	33,20	36,42	39,36	42,98	45,56
25	13,12	14,61	16,47	34,38	37,65	40,65	44,31	46,93

Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Primer Semestre 2016

Curso	**:**	**Probabilidad y Estadística**
Sigla	**:**	**EYP1113**
Pauta	**:**	**Bonus I3**
Profesores	**:**	**Ricardo Aravena C. y Ricardo Olea O.**

¿Son las siguientes afirmaciones verdaderas o falsas? Jutifique en ambos casos, no basta con poner V o F.

Observación: respuesta correcta +0.05 puntos en nota I3, respuesta incorrecta -0.05 puntos en nota I3.

1. Los estimadores de momentos y de máxima verosimilitud siempre coinciden.

   Justificación FALSO, por ejemplo en el caso Log-Normal

2. Una Poisson(100) puede aproximarse por una Normal($\mu = 100$, $\sigma = 10$).

   Justificación VERDADERO, porque una Poisson(100) es equivalente a sumar cien Poisson(1)

3. El sesgo de un estimador siempre será un valor positivo.

   Justificación FALSO, por que una transformación de un estimador puede seguir siendo estimador del mismo parámetro pero sesgado contrariamente

4. Si un estimador es consistente, entonces es un estimador insesgado.

   Justificación FALSO, no necesariamente ya que podría ser asintóticamente insesgado