Pauta auxiliar 9: Estimación

MA3403 - Probabilidades y Estadistica
Profesor: Fernando Lema
Auxiliar: José Cereceda - Antonio Lizama
14 de noviembre de 2013

Problemas.

P1. Sea $X_1, ..., X_n$ una m.a.s. de una v.a. X, con $\mathbb{E}(X) = \mu$, $Var(X) = \sigma^2$. Sea además $S^2 = \dfrac{\sum(X_i - \bar{X})^2}{n-1}$.

a) Mostrar que $\hat{\sigma}^2 = S^2$ es insesgado.

Solución:

$$\begin{aligned}
\sum_{i=1}^{n}(X_i - \bar{X})^2 &= \sum(X_i - \mu + \mu - \bar{X})^2 \\
&= \sum\left[(X_i-\mu)^2 + 2(\mu - \bar{X})(X_i - \mu) + (\mu - \bar{X})^2\right] \\
&= \sum(X_i - \mu)^2 + 2(\mu - \bar{X})\sum(X_i - \mu) + n(\mu - \bar{X})^2 \\
&= \sum(X_i - \mu)^2 - 2n(\mu - \bar{X})\frac{1}{n}\sum(\mu - X_i) + n(\mu - \bar{X})^2 \\
&= \sum(X_i - \mu)^2 - 2n(\mu - \bar{X})^2 + n(\mu - \bar{X})^2 \\
&= \sum(X_i - \mu)^2 - n(\mu - \bar{X})^2
\end{aligned}$$

Así,

$$\begin{aligned}
\mathbb{E}(S^2) &= \frac{1}{n-1}\left[\sum(\mathbb{E}((X_i-\mu)^2) - n\mathbb{E}((\mu - \bar{X})^2)\right] \\
&= \frac{1}{n-1}\left[\sum Var(X_i) - n\mathbb{E}\left(\left[\mu - \frac{1}{n}\sum X_i\right]^2\right)\right] \\
&= \frac{1}{n-1}\left[nVar(X) - n\mathbb{E}\left(\frac{1}{n^2}\left[n\mu - \sum X_i\right]^2\right)\right] \\
&= \frac{1}{n-1}\left[nVar(X) - \frac{1}{n}\mathbb{E}\left(\left[n\mu - \sum X_i\right]^2\right)\right] \qquad (1)
\end{aligned}$$

Definamos, $Z = \sum X_i$. Luego,

$$\mathbb{E}(Z) = \mathbb{E}\left(\sum X_i\right) = \sum \mathbb{E}(X_i) = n\mu$$

Por otro lado, por definición:

$$Var(Z) = \mathbb{E}((\mathbb{E}(Z) - Z)^2) = \mathbb{E}\left((n\mu - Z)^2\right) = \mathbb{E}\left((n\mu - \sum X_i)^2\right)$$

y además, por propiedades de la varianza:

$$Var(Z) = Var\left(\sum X_i\right) = nVar(X) = n\sigma^2$$

Es decir,

$$\mathbb{E}\left((n\mu - \sum X_i)^2\right) = n\sigma^2$$

Reemplazando en (1) se obtiene:

$$\mathbb{E}(S^2) = \frac{1}{n-1}\left[n\sigma^2 - \frac{1}{n}n\sigma^2\right] = \sigma^2$$

Es decir, S^2 es un estimador insesgado de la varianza.

b) Mostrar que si $X \sim N(\mu, \sigma^2)$, entonces

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum(X_i - \bar{X})^2}{\sigma^2} \sim \chi^2_{n-1}$$

Considere sólo el caso $n = 2$.

Solución:
Estudiemos S^2 para el caso $n = 2$.

$$\begin{aligned}
(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 &= \left[X_1 - \frac{1}{2}(X_1 + X_2)^2\right]^2 + \left[X_2 - \frac{1}{2}(X_1 + X_2)^2\right]^2 \\
&= \frac{1}{4}[2X_1 - X_1 - X_2]^2 + \frac{1}{4}[2X_2 - X_1 - X_2]^2 \\
&= \frac{1}{4}[(X_1 - X_2)^2 + (X_2 - X_1)^2] \\
&= \frac{(X_1 - X_2)^2}{2}
\end{aligned}$$

Ahora, como $X_i \sim N(\mu, \sigma^2)$, entonces $X_1 - X_2 \sim N(0, 2\sigma^2)$. Además, si $Y \sim N(0,1) \Rightarrow Y^2 \sim \chi^2_1$. Usando esto, obtenemos que:

$$\begin{aligned}
\left(\frac{X_1 - X_2}{\sqrt{2}\sigma}\right)^2 &\sim \chi^2_1 \\
\frac{(X_1 - X_2)^2}{2\sigma^2} &\sim \chi^2_1 \\
\frac{S^2}{\sigma^2} &\sim \chi^2_1
\end{aligned}$$

Así, tenemos una distibución tabulada para una función del estimador insesgado S^2 de la varianza.

P2. Suponga que T, el tiempo para fallar cierta componente, está distribuido exponencialmente, de parámetro α, y considere $\lambda = 1/\alpha$. Se prueban n componentes, anotando el momento en que falla cada una, $T_1, ..., T_n$. Se desea obtener una estimación insesgada del tiempo esperado en fallar (λ). Para esto considere:

(i) $\hat{\lambda}_1 = \bar{T}$.

(ii) $\hat{\lambda}_2 = n \cdot min\{T_1, ..., T_n\}$.

Estudie ambos e indique cuál escogería.

Solución:

(i) Sabemos que:

$$\mathbb{E}(\hat{\lambda}_1) = \mathbb{E}(\bar{T}) = \mathbb{E}(T) = \lambda \qquad Var(\hat{\lambda}_1) = Var(\bar{T}) = \frac{Var(T)}{n} = \frac{\lambda^2}{n}$$

(ii) Veamos primero que $K = minT_1, ..., T_n \sim exp\,(n\alpha)$

$$\begin{aligned} F_K(x) &= \mathbb{P}(min \leq x) \\ &= 1 - \mathbb{P}(min \geq x) \\ &= 1 - \mathbb{P}(T_1 \geq x, T_2 \geq x, ..., T_n \geq x) \\ &= 1 - \mathbb{P}(T_1 \geq x)\mathbb{P}(T_2 \geq x) \cdots \mathbb{P}(T_n \geq x) \\ &= 1 - (e^{-\alpha x})^n \end{aligned}$$

con esto se tiene que $K = min\{T_1, ...T_n\} \sim exp(n\alpha)$, luego:

$$\mathbb{E}(\hat{\lambda}_2) = \mathbb{E}(nK) = n\frac{1}{n\alpha} = \lambda \qquad\qquad Var(\hat{\lambda}_2) = Var(nK) = n^2\frac{1}{(n\alpha)^2} = \lambda^2$$

Recordemos que un medidor de un buen estimador es el error cuadrático medio. Además, en general se buscan estimadores insesgados. Como hemos visto, en este caso ambos son insesgados. Luego, para minimizar el E.C.M. debemos buscar el de menor varianza. Por esto, en terminos generales nos quedamos con el promedio.

Sin embargo, hay casos en que tomar el mínimo puede ser conveniente, a pesar de tener mayor varianza. Puede ser que la diferencia entre los tiempos de falla de la primera componente y la última sea muy grande, por lo que resulte costoso utilizar el promedio, ya que necesita conocer toda la muestra para obtenerlo.

P3. Sea $\hat{\theta}$ estimador de cierto parámetro θ. Muestre que si $\hat{\theta}$ es insesgado o asintóticamente insesgado, esto esto, $\lim_{n\to\infty} \mathbb{E}(\hat{\theta}) = \theta$, entonces basta que $\lim_{n\to\infty} Var(\hat{\theta}) = 0$ para que el estimador $\hat{\theta}$ sea consistente.

Solución:

$$\begin{aligned} \mathbb{P}(|\hat{\theta} - \theta| \geq \varepsilon) &\overbrace{\leq}^{\text{Des. Chebyshev}} \frac{\mathbb{E}((\hat{\theta} - \theta)^2)}{\varepsilon^2} \\ &= \frac{Var(\hat{\theta}) + (\mathbb{E}(\hat{\theta}) - \theta)^2}{\varepsilon^2} \end{aligned}$$

Tomando límite se concluye.

P4. Sea $X_1, ..., X_n$ m.a.s. de $X \sim N(\mu, \sigma^2)$. Considere $\hat{\mu} = \bar{X}$, con σ^2 conocido. Determine n t.q.

$$\mathbb{P}(|\bar{X} - \mu| < \varepsilon) = \gamma = 1 - \alpha$$

Solución: Sabemos que $Z = \frac{\bar{X}-\mu}{\sigma^2/\sqrt{n}} \sim \mathcal{N}(0, 1)$, utilizaremos esto en lo que sigue para encontrar el n requerido

$$\begin{aligned} \mathbb{P}(|\bar{X} - \mu| < \varepsilon) &= \mathbb{P}(\frac{|\bar{X} - \mu|}{\sigma^2/\sqrt{n}} < \varepsilon\frac{1}{\sigma^2/\sqrt{n}}) \\ &= 1 - \mathbb{P}(\left|\frac{\bar{X} - \mu}{\sigma^2/\sqrt{n}}\right| > \varepsilon\frac{1}{\sigma^2/\sqrt{n}}) \\ &= 1 - \mathbb{P}(\frac{\bar{X} - \mu}{\sigma^2/\sqrt{n}} > \varepsilon\frac{1}{\sigma^2/\sqrt{n}}) - \mathbb{P}(\frac{\bar{X} - \mu}{\sigma^2/\sqrt{n}} < -\varepsilon\frac{1}{\sigma^2/\sqrt{n}}) \\ &= 1 - 2\mathbb{P}(\frac{\bar{X} - \mu}{\sigma^2/\sqrt{n}} < -\varepsilon\frac{1}{\sigma^2/\sqrt{n}}) \\ 1 - \alpha &= 1 - 2\Phi\left(-\varepsilon\frac{1}{\sigma^2/\sqrt{n}}\right) \end{aligned}$$

esto implica que

$$\frac{\alpha}{2} = \Phi\left(-\varepsilon\frac{1}{\sigma^2/\sqrt{n}}\right)$$
$$\Phi^{-1}\left(\frac{\alpha}{2}\right) = -\varepsilon\frac{1}{\sigma^2/\sqrt{n}}$$
$$\left(\Phi^{-1}\left(\frac{\alpha}{2}\right)\frac{\sigma^2}{\varepsilon}\right)^2 = n$$

P5. Se sabe que cierta proporción fija p de detonantes es defectuosa. De una gran partida, se eligen n al azar y se prueban.

a) Determine la función de verosimilitud para X v.a. que indica si un detonante es o no defectuoso.
Solución: Definiendo la v.a. $X_i = 1$ si el $i-$ésimo detonante es defectuoso y 0 en otro caso para $i = 1, ..., n$
Por lo tanto se tiene que $(X_1, ..., X_n)$ es una m.a.s. de la v.a. definida por su función de probabilidad $\mathbb{P}(X = 0) = f(0; p) = 1 - p$ y $\mathbb{P}(X = 1) = f(1; p) = p$. Esto se puede escribir de la forma $f(x; p) = p^x (1-p)^{1-x}$ para $x = 0, 1$, luego

$$L(X_1, ..., X_n; p) = p^k(1-p)^{n-k}$$

con $k = \sum_{i=1}^n X_i = \#$ de detonantes defectuosos. Luego $\ln L(X_1, ..., X_n; p) = k \ln p + (n-k)\ln(1-p)$.
Por tanto

$$\frac{\partial \ln L}{\partial p} = \frac{k}{p} - \frac{n-k}{1-p}$$

Luego $\dfrac{\partial \ln L}{\partial p}\hat{p} = 0$, encontramos como solución $\hat{p} = \frac{k}{n} = \overline{X}$, que es un estimados insesgado del parámetro buscado.

b) Suponga que ahora desea calcular el parámetro β en una ley exponencial del tiempode falla de los detonantes, al probar n detonantes llamamos $T_1, ..., T_n$ los tiempos de falla, utilizando lo anterior encuentre un parámetro de β
Solución: Si llamamos esta vez por p a la probabilidad de que un detonante falle en un tiempo T_0, es decir

$$p = \mathbb{P}(T \leq T_0) = 1 - e^{-\beta T_0}$$

Por la parte anterior, se tiene que el estimador de p es $\hat{p} = k/n$, aplicando la propiedad de invarianza del estimador ML, es decir $\hat{g(\theta)} = g\left(\hat{\theta}\right)$, ($p$ es creciente con respecto a β), obtenemos el estimados ML de β

$$1 - \widehat{e^{-\beta T_0}} = 1 - e^{-\hat{\beta} T_0} = \hat{p} = k/n$$

Lo que implica que $\hat{\beta} = -\dfrac{1}{T_0}\ln\left(\dfrac{n-k}{n}\right)$, por lo que el estimador de $1/\beta = \mathbb{E}(exp(\beta))$

$$\left(\hat{\frac{1}{\beta}}\right) = \frac{-T_0}{\ln[(n-k)/n]}$$