Auxiliar 12: Pre-Control

MA3403 - Probabilidades y Estadistica
Profesor: Fernando Lema
Auxiliar: Martín Castillo & José Cereceda
13 de junio de 2014

P1. Sea $X \sim N(\mu, \sigma^2)$, con σ^2 conocido. Suponga que tiene una muestra $X_1, ..., X_n$.

a) Construya un intervalo de confianza para μ a nivel $(1-\alpha)$.
Analice qué sucede frente a variaciones en los parámetros n, σ, α.

Solución:

En virtud del teorema central del límite, se tiene que la distribución muestral de $Z = \frac{\overline{X}-\mu}{\sigma}\sqrt{n}$ es una normal $N(0,1)$, es decir que

$$\mathbb{P}(Z \leq z) = \Phi(z)$$

entonces tendremos que

$$\begin{aligned} 2\Phi(z) - 1 &= \mathbb{P}\left(-z \leq \frac{\overline{X}-\mu}{\sigma}\sqrt{n} \leq z\right) \\ &= \mathbb{P}\left(\frac{-z\sigma}{\sqrt{n}} - \overline{X} \leq -\mu \leq \frac{z\sigma}{\sqrt{n}} - \overline{X}\right) \\ &= \mathbb{P}\left(\overline{X} - \frac{z\sigma}{\sqrt{n}} \leq \mu \leq \frac{z\sigma}{\sqrt{n}} + \overline{X}\right) \end{aligned}$$

Luego esto nos permite decir que la probabilidad de que el intervalo aleatorio contenga a el parámetro real μ es $2\Phi(z) - 1$, luego como hemos escogido un z arbitrario, podemos imponer que este z sea tal que

$$2\Phi(z) - 1 = 1 - \alpha$$

es decir, escogemos nuestro z tal que $\Phi(z^*) = 1 - \alpha/2$, con esto decimos que el intervalo (aleatorio) $\left(\overline{X} - \frac{\sigma z^*}{\sqrt{n}}, \overline{X} + \frac{\sigma z^*}{\sqrt{n}}\right)$ es intervalo de confianza, con un *nivel* o *coeficiente de confianza* de $1-\alpha$.

Observemos que si aumentamos n, que se traduce en un aumento en la cantidad de muestras nuestro intervalo de confianza disminuye su longitud, es decir nuestro intervalo se vuelve más preciso.

Si consideramos nuestro σ conocido, y este aumenta, lo que se traduce en un aumento de la varianza de nuestra v.a. el intervalo aleatorio aumenta de tamaño y se vuelve menos preciso.

Si disminuimos nuestro parámetro α observemos que esto implica en la ecuación $\Phi(z^*) = 1-\alpha/2$ que nuestro z^* aumentará, todo esto se traduce que nuestro intervalo aleatorio tendrá una probabilida más cercana a uno de contener a μ, sin embargo nuestro intervalo se vuelve más impreciso con una longitud más larga.

b) Suponga que desea estudiar el valor de la media μ. Ud. sabe que $\overline{X} = 70$. ¿Qué nivel de confianza le da el intervalo (68,76), si $n = 25$ y $\sigma = 12$?

Solución:

Primero cabe notar que el intervalo de confianza que buscamos para μ se trata de un intervalo no simetrico, por lo que debemos replantear el análisis de la parte anterior, de la siguiente forma para Z una v.a. normal (0,1) (considerando $z_1, z_2 > 0$)

$$\begin{aligned}
\Phi(z_1) - (1 - \Phi(z_2)) &= \mathbb{P}(Z \leq z_1) - \mathbb{P}(Z \leq -z_2) \\
&= \mathbb{P}(-z_2 \leq Z \leq z_1) \\
&= \mathbb{P}\left(-z_2 \leq \frac{\overline{X} - \mu}{\sigma}\sqrt{n} \leq z_1\right) \\
&= \mathbb{P}\left(\frac{-z_2\sigma}{\sqrt{n}} - \overline{X} \leq -\mu \leq \frac{z_1\sigma}{\sqrt{n}} - \overline{X}\right) \\
&= \mathbb{P}\left(\underbrace{\overline{X} - \frac{z_1\sigma}{\sqrt{n}}}_{68} \leq \mu \leq \underbrace{\overline{X} + \frac{z_2\sigma}{\sqrt{n}}}_{76}\right)
\end{aligned}$$

luego ya que tenemos como datos, $\sigma, n, \overline{X}$, luego imponemos $\frac{z_1\sigma}{\sqrt{n}} = 2 \Rightarrow \{z_1 = ,83\}$ y $\frac{z_2\sigma}{\sqrt{n}} = 6 \Rightarrow \{z_2 = 2{,}5\}$, luego reemplazamos los valores de z_1, z_2 en la ecuación,

$$\begin{aligned}
\Phi(,83) + \Phi(2{,}5) - 1 &= 1 - \alpha \\
,797 + ,993 - 1 &= 1 - \alpha
\end{aligned}$$

luego alpha es $\alpha = ,21$, es decir el intervalo construido tiene un coeficiente de confianza $\alpha =, 21$

P2. El auto solar "solarium" se alista para competir en una carrera de 3600 (km). La distancia que puede recorrer en el dia es de $X = \sqrt{10R}$ donde R es la radiación solar en el dia y sigue $R \sim U(1000, 2000)$.

Usando TCL (Asuma que la radiación en dias distintos son independientes).

a) Calcule la probabilidad que el "Solarium" llegue a la meta dentro del plazo de 30 dias.

Solución:

Queremos calcular la $\mathbb{P}\left(\sum_{i=1}^{30} X_i \geq 3600\right)$. Para esto calcularemos $\mathbb{E}(X)$ y $Var(X)$

$$\begin{aligned}
\mathbb{E}(X) &= \int_{1000}^{2000} \sqrt{10R}\frac{1}{1000}dR \\
&= \frac{\sqrt{10}}{1000}\,\frac{2}{3}R^{3/2}\Big|_{1000}^{2000}
\end{aligned}$$

y para calcular la varianza

$$
\begin{aligned}
\mathbb{E}\left(X^2\right) &= \int_{1000}^{2000} 10R\frac{1}{1000}dR \\
&= \frac{1}{100}\frac{1}{2}R^2\Big|_{1000}^{2000}
\end{aligned}
$$

luego utilizando TCL, con $Z \sim (0,1)$

$$
\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{30} X_i \geq 3600\right) &= \mathbb{P}\left(\frac{\sum_{i=1}^{30} X_i - \mu}{\sigma^2/30} \geq \frac{3600-\mu}{\sigma^2/30}\right) \\
&= \mathbb{P}\left(Z \geq \frac{3600-\mu}{\sigma^2/30}\right) \\
&= 1 - \mathbb{P}\left(Z \leq \frac{3600-\mu}{\sigma^2/30}\right) \\
&= 1 - \Phi\left(\frac{3600-\mu}{\sigma^2/30}\right)
\end{aligned}
$$

b) ¿Cuantos días necesita para asegurarse de llegar a la meta con un 99 % de seguridad?

Solución:
De la misma forma que el item anterior, imponemos $\mathbb{P}\left(\sum_{i=1}^{n} X_i \geq 3600\right) = 0{,}99$

$$
\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{n} X_i \geq 3600\right) &= \mathbb{P}\left(\frac{\sum_{i=1}^{n} X_i - \mu}{\sigma^2/n} \geq \frac{3600-\mu}{\sigma/n}\right) \\
&= \mathbb{P}\left(Z \geq \frac{3600-\mu}{\sigma/\sqrt{n}}\right) \\
&= 1 - \mathbb{P}\left(Z \leq \frac{3600-\mu}{\sigma/\sqrt{n}}\right) \\
&= 1 - \Phi\left(\frac{3600-\mu}{\sigma/\sqrt{n}}\right)
\end{aligned}
$$

Luego, $0{,}01 = \Phi\left(\frac{3600-\mu}{\sigma/\sqrt{n}}\right)$, por lo que $n \geq \left(\Phi^{-1}(0{,}01)\frac{\sigma}{3600-\mu}\right)^2$.

P3. La velocidad de un grupo de partículas es una v.a. (X_1) normal (en mt/seg) de media 10 (mt/seg) y desviación estándar 3 (mt/seg).

a) Se toman dos partículas independientes. Calcule la probabilidad que sus velocidades difieran en menos de 1 (mt/seg).
Calcule la probabilidad que su velocidad promedio supere los 11 (mt/seg).

b) Otro grupo de partículas tiene velocidad (X_2) $N(14,4)$ y los dos grupos se mezclan. Dado que se mide la velocidad de una partícula se desea saber que grupo pertenece. Se propone determinar v que minimiza

$$\mathbb{P}(X_1 > v) + \mathbb{P}(X_2 < v) \qquad (1)$$

y decidir según si la velocidad es mayor o menor a v. ¿Qué interpretación tiene (1)? Analice gráficamente.

c) Un tercer grupo de partículas tiene velocidad $N(\mu, \sigma^2)$. Se toma una muestra de cinco de ellos y sus velocidades son 10, 12, 9, 8, 13.

i) Construya un I.C. para σ^2, con confianza 0.9, de tipo cota superior (ie, $\sigma^2 <$).

ii) Si $\sigma^2 = 3{,}5$, determine con qué confianza se construyó el intervalo de confianza para μ dado por $(10; 11{,}5)$.

P4. Sea $X \sim G(r, \alpha)$, y $X_1, ..., X_n$ m.a.s.

a) Muestre que el estimador de los momentos de α, para r conocido, es $\hat{\alpha} = \frac{r}{\overline{X}}$.

Solución:

Recordando que para una v.a. con distribución gamma, se tendrá que $\mathbb{E}(X) = \frac{r}{\alpha}$ y luego de acuerdo al metodo de los momentos, se estima $\mathbb{E}(X)$ por $\frac{\sum X_i}{n}$, luego tendremos que $\overline{X} = \widehat{\left(\frac{r}{\alpha}\right)}$ como suponemos r conocido, y de acuerdo a la invarianza del estimador, se concluye que $\overline{X} = \dfrac{r}{\hat{\alpha}}$ lo que implica que $\hat{\alpha} = \frac{r}{\overline{X}}$.

b) Muestre que el estimador de los momentos de $\lambda = 1/\alpha$, para r conocido, es $\hat{\lambda} = \frac{\bar{X}}{r}$. Encuentre además la distribución muestral de $\hat{\lambda}$ para n grande.

Solución:

En vitrud del TCL, tendremos que para n 'grande', $\overline{X}$ sigue una distribución normal de parámetros $N\left(\mathbb{E}(X), \mathbb{V}(X)\right) = N\left(\frac{r}{\alpha}, \frac{r}{n\alpha^2}\right)$ luego $\dfrac{\overline{X}}{r} \sim N\left(\frac{1}{\alpha}, \frac{1}{rn\alpha^2}\right)$

c) Obtenga el estimador de máxima verosimilitud $\hat{\alpha}$ para α.

Solución:

Recordando que la distribución de la función gamma es de la forma $f(x) = \dfrac{\alpha^r x^{r-1}}{\Gamma(r)\, e^{-\alpha x}}$ luego la función de verosimilitud muestral es

$$
\begin{aligned}
L(X_1, ..., X_2; \alpha, r) &= \Pi_{i=1}^{n} f(X_i) \\
&= \frac{\alpha^{nr} (\Pi X_i)^{r-1}}{\Gamma(r)^n} e^{\{-\alpha \sum_{i=1}^{n} X_i\}} \\
\log(L) &= \log\left(\frac{\alpha^{nr} (\Pi X_i)^{r-1}}{\Gamma(r)^n} e^{\{-\alpha \sum_{i=1}^{n} X_i\}}\right) \\
&= \log\left(\frac{\alpha^{nr}}{\Gamma(r)^n}\right) + \log\left((\Pi X_i)^{r-1}\right) + \log\left(e^{\{-\alpha \sum_{i=1}^{n} X_i\}}\right) \\
&= nr\log(\alpha) + (r-1)\sum_{i=1}^{n} \log(X_i) - \alpha \sum_{i=1}^{n} X_i - n\log(\Gamma(r))
\end{aligned}
$$

luego buscando los parámetros que máximizan, imponemos

$$
\frac{\partial \log L}{\partial r} = \frac{\partial \log L}{\partial \alpha} = 0
$$

P5. Volviendo a la pregunta 1, ¿Cómo encontraría un intervalo de confianza para μ, esta vez con σ desconocido?

Solución:

Sabemos que la v.a. $Z = \dfrac{\overline{X}-\mu}{\sigma}\sqrt{n}$ sigue una normal (0,1) y también sabemos que $V = \dfrac{(n-1)\,\hat{\sigma}^2}{\sigma^2} = \sum_{i=1}^{n}(X_i - \overline{X})^2/\sigma^2$ sigue una ${\chi^2}_{n-1}$, se puede mostrar que Z, V son v.a. independientes, veamos que si definimos $t = Z/\sqrt{V/(n-1)}$

$$\begin{aligned} Z/\sqrt{V/(n-1)} &= \frac{\overline{X}-\mu}{\sigma}\sqrt{n}\frac{1}{\sqrt{\frac{(n-1)\hat{\sigma}^2}{\sigma^2}/(n-1)}} \\ &= \frac{\overline{X}-\mu}{\sigma}\sqrt{n}\frac{1}{\sqrt{\frac{\hat{\sigma}^2}{\sigma^2}}} \\ &= \frac{\overline{X}-\mu}{\hat{\sigma}}\sqrt{n} \end{aligned}$$

es decir, para encontrar un intervalo de confianza para μ, podemos utilizar el estadistico $\frac{\overline{X}-\mu}{\hat{\sigma}}\sqrt{n}$ que sigue la distribución de una v.a. t-student, de $n-1$ grados de libertad, y se tiene que la función de densidad de una v.a. $t-$student es

$$h_{n-1}(t) = \frac{\Gamma[n/2]}{\Gamma\left((n-1)/2\right)\sqrt{\pi(n-1)}}\left(1+\frac{t^2}{n-1}\right)^{-n/2} \qquad -\infty < t < \infty$$