Código de Honor: Como miembro de la comunidad de la Pontificia Universidad Católica de Chile, me comprometo a respetar los principios y normativas que la rigen. Asimismo, me comprometo a actuar con rectitud y honestidad en esta evaluación. Adicionalmente declaro estar en condiciones de salud adecuadas para rendir esta evaluación y que me presento a ésta bajo mi responsabilidad. En caso de sentirme mal o tener alguna complicación, deberé informarlo inmediatamente al ayudante o profesor en sala.

Pontificia Universidad Católica de Chile
Facultad de Economía y Administración

Primer Semestre 2021

Curso	**:**	**Inferencia Estadística**
Sigla	**:**	**EAA1520**
Profesores	**:**	**Valeria Leiva (Sec 1), Cristian Vásquez (Sec 2), Rafael Águila (Sec 4-5-6)**

Pauta Control 1

Problema 1

El gobierno de Chile está llevando a cabo el plan de vacunación COVID-19 "Yo me vacuno" por medio del Ministerio de Salud, de manera gradual y prograsiva. Este plan es voluntario y gratuito, y lo que busca es proteger a la población chilena contra la infección provacada por este virus. Las vacunas utilizadas en el plan se administran en dos dosis, con una diferencia de 28 días. Según las cifras oficiales, el avance en el plan de vacunación a la fecha 31 de marzo es el siguiente: Se han administrado 6,809,761 de primeras dosis de vacuna y 3,678,082 de segundas dosis de vacuna.

Suponga que el número diario de personas que se les adminstra la primera dosis Y es independiente e identicamente distribuído y, tiene la siguiente función de probabilidad:

$$\Pr(Y = y) = \frac{\left(e^{\theta}\right)^{y} e^{-e^{\theta}}}{y!}, \quad y = 0, 1, 2, \ldots.$$

donde $\theta \in \mathrm{R}$ es el parámetro desconocido, $\mathrm{E}(Y) = e^{\theta}$ y $\mathrm{Var}(Y) = e^{\theta}$. Suponga que se disponde una muestra aleatoria $\{Y_1, Y_2, \ldots, Y_n\}$ i.i.d. de administración de primeras dosis. Con los antecedentes anteriores realice lo siguiente:

(a) **[1.5 Ptos]** Determine el estimados de momento del parámetro desconocido θ.

(b) **[1.5 Ptos]** Muestre que el estimador de máxima verosimilitud del parámetro θ está dado por

$$\widehat{\theta} = \ln(\overline{Y}).$$

(c) **[0.5 Ptos]** En base a sus respuestas anteriores, ¿Son iguales el estimador de momento con el estimador de máxima verosimilitud?.

(d) **[1.5 Ptos]** Se dispone de la siguiente muestra observada desde el 15 de marzo hasta el 26 de marzo (**días hábiles considerando cifras oficiales**) de aplicación de la primera dosis

$$\{157684, 166332, 151032, 148588, 139677, 146323, 150071, 153761, 157276, 165806\}$$

Reporte el valor del estimador de máxima verosimilitud para θ. Aproxime su resultado al segundo decimal.

(e) **[1.0 Ptos]** Suponga que se desea analizar la siguiente función del parámetro desconocido $g(\theta) = e^{\theta}$ y se propone utilizar $g(\widehat{\theta}) = e^{\widehat{\theta}}$, donde $\widehat{\theta}$ corresponde al estimador de máxima verosimilitud. ¿Cuál es la distribución aproximada de $g(\widehat{\theta})$ cuando el tamaño muestral es grande?.

Solución

(a) Note que el número de parámetros desconocidos es $m = 1$ **[0.3 Ptos]**, por lo tanto, se debe igualar el primer momento poblacional $\mathrm{E}(Y)$ con el primer momento muestral $\overline{Y}$ y se despeja el parámetro desconocido:

$$\begin{aligned} \mathrm{E}(Y) &= \overline{Y}, \\ e^{\theta} &= \overline{Y} \ \textbf{[0.7 Ptos]}. \end{aligned}$$

De esta forma, aplicando logaritmo natural a ambos lados de la igualdad, el estimador de momento está dado por:

$$\widehat{\theta} = \ln(\overline{Y}) \ \textbf{[0.5 Ptos]}.$$

(b) Primero se debe determinar la función de verosimilitud:

$$\begin{aligned} L(\theta) &= \Pr(Y_1, Y_2, \ldots, Y_n), \\ &\overset{i.i.d.}{=} \prod_{i=1}^{n} \Pr(Y_i), \\ &= \prod_{i=1}^{n} \frac{\left(e^{\theta}\right)^{Y_i} e^{-e^{\theta}}}{Y_i!}, \\ &= \frac{\left(e^{\theta}\right)^{\sum_{i=1}^{n} Y_i} \cdot e^{-n\, e^{\theta}}}{\prod_{i=1}^{n} Y_i!} \ \textbf{[0.5 Ptos]}. \end{aligned}$$

Luego se puede determinar la función de log-verosimilitud aplicando el logaritmo natural a la función de verosimilitud,

$$\begin{aligned} \ell(\theta) &= \ln(L(\theta)), \\ &= \ln\left[\frac{\left(e^{\theta}\right)^{\sum_{i=1}^{n} Y_i} \cdot e^{-n\, e^{\theta}}}{\prod_{i=1}^{n} Y_i!}\right], \\ &= \ln(e^{\theta}) \sum_{i=1}^{n} Y_i - n e^{\theta} \ln(e) - \ln\left(\prod_{i=1}^{n} Y_i!\right) \ \textbf{[0.3 Ptos]}, \\ &= \theta \sum_{i=1}^{n} Y_i - n e^{\theta} - \ln\left(\prod_{i=1}^{n} Y_i!\right) \ \textbf{[0.2 Ptos]} \end{aligned}$$

Finalmente, hay que buscar el argumento θ que maximiza la función $\ell(\theta)$. Para este tipo de función, se puede derivar con respecto a θ, igualar a cero y despejar θ,

$$\begin{aligned} \frac{d}{d\theta} \ell(\theta) &= \frac{d}{d\theta}\left(\theta \sum_{i=1}^{n} Y_i - n e^{\theta} - \ln\left(\prod_{i=1}^{n} Y_i!\right)\right), \\ &= \sum_{i=1}^{n} Y_i - n e^{\theta} \ \textbf{[0.3 Ptos]}. \end{aligned}$$

Igualando a cero se tiene $\sum_{i=1}^{n} Y_i - ne^{\theta} = 0$, así $ne^{\theta} = \sum_{i=1}^{n} Y_i$, lo que implica que

$$\widehat{\theta} = \ln(\overline{Y}) \text{ [0.2 Ptos]}.$$

(c) Note que ambos estimadores son iguales. **[0.5 Ptos]**

(d) Ahora se debe calcular el estimador con la muestra observada

$$\{157684, 166332, 151032, 148588, 139677, 146323, 150071, 153761, 157276, 165806\}$$

Utilizando los datos $\overline{y} = 153655$ **[0.5 Ptos]**. Entonces la estimación puntual de θ a partir de esta muestra es

$$\begin{aligned}\widehat{\theta} &= \ln(153655),\\ &= 11.94 \text{ [1.0 Ptos]}.\end{aligned}$$

(e) Para analizar $g(\theta) = e^{\theta}$ se propone utilizar $g(\widehat{\theta}) = e^{\widehat{\theta}}$, donde el EMV de θ está dado por $\widehat{\theta} = \ln(\overline{Y})$. Reemplazando, puede notar que

$$g(\widehat{\theta}) = e^{\ln(\overline{Y})} = \overline{Y} \text{ [0.3 Ptos]}.$$

Así, $g(\widehat{\theta})$ corresponde al promedio muestral, por lo tanto, se puede utilizar el Teorema del Límite Central (TLC) para tener distribución aproximada.

A partir del enunciado, se tiene que $\mathrm{E}(Y) = e^{\theta}$ y $\mathrm{Var}(Y) = e^{\theta}$ **[0.2 Ptos]**, entonces, por el TLC

$$\sqrt{n}\,\frac{\overline{Y} - e^{\theta}}{\sqrt{e^{\theta}}} \overset{aprox}{\sim} \text{Normal}\,(0,\, 1) \text{ [0.5 Ptos]}$$

De esta forma, la distribución aproximada de $g(\widehat{\theta}) = \overline{Y}$ está dada por:

$$\overline{Y} \overset{aprox}{\sim} \text{Normal}\left(e^{\theta},\, \frac{e^{\theta}}{n}\right) \text{ o [0.5 Ptos] por esta respuesta}$$