Ejercicio de estimación de máxima verosimilitud

El tiempo de realización en minutos de una determinada tarea dentro de un proceso industrial es una variable aleatoria con función de densidad

$$f(x) = \frac{x}{\theta^2} e^{-x/\theta} \qquad \textbf{si } x > 0$$

donde $\theta > 0$.

a) **Calcular el estimador máximo-verosímil de θ para una muestra aleatoria simple de tamaño n.**

1. Escribir la verosimilitud:

$$\begin{aligned}
\mathcal{L}(\theta) &= \mathcal{L}(x_1, \ldots, x_n; \theta) \overset{ind.}{=} \prod_{i=1}^{n} f_{X_i}(x_i) \overset{i.d.}{=} \prod_{i=1}^{n} f_X(x_i) \\
&= \prod_{i=1}^{n} \left(\frac{x_i}{\theta^2} e^{-x_i/\theta} \right) = \left(\frac{1}{\theta^2} \right)^n \prod_{i=1}^{n} e^{-x_i/\theta} \prod_{i=1}^{n} x_i \\
&\overset{e^a \cdot e^b = e^{a+b}}{=} \left(\frac{1}{\theta^2} \right)^n \exp\left\{ -\frac{1}{\theta} \sum_{i=1}^{n} x_i \right\} \prod_{i=1}^{n} x_i; \qquad x_1, \ldots, x_n > 0
\end{aligned}$$

2. Escribir el logaritmo de la verosimilitud:

$$\begin{aligned}
\ell(\theta) &= \ln \mathcal{L}(\theta) = \ln \left(\left(\frac{1}{\theta^2} \right)^n \exp\left\{ -\frac{1}{\theta} \sum_{i=1}^{n} x_i \right\} \prod_{i=1}^{n} x_i \right) \\
&\overset{\ln(a \cdot b) = \ln a + \ln b}{=} \ln \left(\left(\frac{1}{\theta^2} \right)^n \right) + \ln \left(\exp\left\{ -\frac{1}{\theta} \sum_{i=1}^{n} x_i \right\} \right) + \ln \left(\prod_{i=1}^{n} x_i \right) \\
&= \ln\left(\theta^{-2n}\right) - \frac{1}{\theta} \sum_{i=1}^{n} x_i + \ln \left(\prod_{i=1}^{n} x_i \right) \\
&= -2n \ln \theta - \frac{1}{\theta} \sum_{i=1}^{n} x_i + \ln \left(\prod_{i=1}^{n} x_i \right); \qquad x_1, \ldots, x_n > 0
\end{aligned}$$

3. Obtener el θ tal que $\frac{\partial}{\partial \theta_j} \ell(\theta) = 0$.

$$\begin{aligned}
\frac{\partial}{\partial \theta} \ell(\theta) &= \frac{\partial}{\partial \theta} \left(-2n \ln \theta - \frac{1}{\theta} \sum_{i=1}^{n} x_i + \ln \left(\prod_{i=1}^{n} x_i \right) \right) \\
&= \frac{\partial}{\partial \theta} \left(-2n \ln \theta \right) - \frac{\partial}{\partial \theta} \left(\frac{1}{\theta} \sum_{i=1}^{n} x_i \right) + \frac{\partial}{\partial \theta} \left(\ln \left(\prod_{i=1}^{n} x_i \right) \right) \\
&= \frac{-2n}{\theta} - \left(-\frac{\sum_{i=1}^{n} x_i}{\theta^2} \right) + 0 \\
&= \frac{-2n}{\theta} + \frac{\sum_{i=1}^{n} x_i}{\theta^2}
\end{aligned}$$

Por lo tanto:

$$\begin{aligned}\frac{\partial}{\partial\theta}\ell(\theta)=0 \quad &\Leftrightarrow \quad \frac{-2n}{\theta}+\frac{\sum_{i=1}^n x_i}{\theta^2}=0 \Leftrightarrow \frac{\sum_{i=1}^n x_i}{\theta^2}=\frac{2n}{\theta}\\ &\overset{\times\theta}{\Leftrightarrow} \quad \frac{\sum_{i=1}^n x_i}{\theta}=2n \Leftrightarrow \theta=\frac{\sum_{i=1}^n x_i}{2n}\end{aligned}$$

El candidato a EMV es $\hat{\theta}_{MV}=\dfrac{\sum_{i=1}^n x_i}{2n}$.

4. Comprobar que realmente es un máximo, es decir, que $\frac{\partial^2}{\partial\theta^2}\ell(\theta)|_{\theta=\hat{\theta}_{MV}}<0$.

$$\begin{aligned}\frac{\partial^2}{\partial\theta^2}\ell(\theta) &= \frac{\partial}{\partial\theta}\left(\frac{-2n}{\theta}+\frac{\sum_{i=1}^n x_i}{\theta^2}\right)\\ &= (-2n)\frac{-1}{\theta^2}+\sum_{i=1}^n x_i\frac{-2}{\theta^3}\\ &= \frac{2n}{\theta^2}-\frac{2}{\theta^3}\sum_{i=1}^n x_i=\frac{2}{\theta^2}\left(n-\frac{1}{\theta}\sum_{i=1}^n x_i\right)\end{aligned}$$

Si ahora lo evaluamos en el candidato $\hat{\theta}_{MV}=\dfrac{\sum_{i=1}^n x_i}{2n}$, tenemos:

$$\begin{aligned}\frac{\partial^2}{\partial\theta^2}\ell(\theta)|_{\theta=\hat{\theta}_{MV}} &= \frac{2}{\hat{\theta}_{MV}^2}\left(n-\frac{1}{\hat{\theta}_{MV}}\sum_{i=1}^n x_i\right)=\frac{2}{\hat{\theta}_{MV}^2}\left(n-\frac{1}{\frac{\sum_{i=1}^n x_i}{2n}}\sum_{i=1}^n x_i\right)\\ &= \frac{2}{\hat{\theta}_{MV}^2}(n-2n)=\frac{2}{\hat{\theta}_{MV}^2}(-n)<0\end{aligned}$$

obtenemos que la segunda derivada de $\ell(\theta)$ es negativa en $\hat{\theta}_{MV}$ y por tanto es un máximo.

El estimador máximo verósimil de θ es $\hat{\theta}_{MV}=\dfrac{\overline{X}}{2}$. Para una muestra particular, la estimación máximo verosímil será $\hat{\theta}_{MV}=\dfrac{\overline{x}}{2}$.

b) Calcular el estimador máximo-verosímil de $E[X]$ para una muestra aleatoria simple de tamaño n.

Lo primero es calcular $E[X]$. Como X es una v.a. continua, sabemos que $E[X]=\int_{-\infty}^{\infty}x\,f(x)dx$. Pero como X sólo toma valores positivos:

$$E[X]=\int_0^\infty x\,f(x)\,dx = \int_0^\infty x\,\frac{x}{\theta^2}e^{-x/\theta}\,dx=\int_0^\infty \frac{x^2}{\theta^2}e^{-x/\theta}\,dx=-\frac{1}{\theta}\int_0^\infty x^2\left(\frac{-1}{\theta}\right)e^{-x/\theta}\,dx$$

$$\begin{aligned}
&= \left[\begin{array}{ll} \textit{integr. \quad por} & \textit{partes}^1 \\ u = x^2 & du = 2x\,dx \\ dv = \frac{-1}{\theta} e^{-x/\theta} dx & v = e^{-x/\theta} \end{array}\right] = -\frac{1}{\theta}\left(\left[x^2\, e^{-x/\theta}\right]_0^\infty - \int_0^\infty e^{-x/\theta} 2x\,dx\right) \\
&= -\frac{1}{\theta}\left[x^2\, e^{-x/\theta}\right]_0^\infty + \frac{1}{\theta}\int_0^\infty 2x\, e^{-x/\theta}\,dx = -\frac{1}{\theta}\left(\lim_{x\to\infty} x^2\, e^{-x/\theta} - \lim_{x\to 0} x^2\, e^{-x/\theta}\right) + \frac{1}{\theta}\int_0^\infty 2x\, e^{-x/\theta}\,dx \\
&=^2 -\frac{1}{\theta}(0-0) + \frac{1}{\theta}\int_0^\infty 2x\, e^{-x/\theta}\,dx = -2\int_0^\infty x\left(\frac{-1}{\theta}\right) e^{-x/\theta}\,dx = \left[\begin{array}{ll} \textit{integr. \quad por} & \textit{partes}^1 \\ u = x & du = dx \\ dv = \frac{-1}{\theta} e^{-x/\theta} dx & v = e^{-x/\theta} \end{array}\right] \\
&= -2\left(\left[x\, e^{-x/\theta}\right]_0^\infty - \int_0^\infty e^{-x/\theta} dx\right) =^2 -2\left(0 - \left[-\theta e^{-x/\theta}\right]_0^\infty\right) \\
&= -2\theta\left(\lim_{x\to\infty} e^{-x/\theta} - \lim_{x\to 0} e^{-x/\theta}\right) = -2\theta(0-1) = 2\theta
\end{aligned}$$

Tenemos que $E[X] = 2\theta$. Por el principio de invarianza del EMV, sabemos que si $\hat{\theta}_{MV}$ es el estimador máximo verosímil de θ, el estimador máximo verosímil de cualquier función $h(\theta)$ es $h(\hat{\theta}_{MV})$.
Por tanto, el EMV de $E[X]$ será:

$$\widehat{E[X]}_{MV} = 2\,\hat{\theta}_{MV} = 2\,\frac{\overline{X}}{2} = \overline{X}.$$

c) **Mediante un muestreo aleatorio simple se han recogido los siguientes 15 tiempos de realización de la tarea:**

5.56	2.23	0.58	1.37	0.21	1.98	2.44	2.71
10.12	4.69	3.47	1.73	3.51	1.19	0.97	

Obtener la estimación máximo-verosímil del tiempo medio de realización del proceso.

Para esta muestra se tiene que $\overline{x} = \dfrac{1}{15}\displaystyle\sum_{i=1}^{15} x_i = 2.8507$, por tanto, la estimación máximo verosímil para el tiempo medio será:

$$\widehat{E[X]}_{MV} = \overline{x} = 2.8507 \quad \text{minutos.}$$

[1]Integración por partes: $\int u\cdot dv = u\cdot v - \int v\cdot du$. Si la integral es definida, entonces $\int_a^b u\cdot dv = [u\cdot v]_a^b - \int_a^b v\cdot du$.

[2]La exponencial negativa decrece mucho más rápido hacia cero que lo que x^2 crece hacia infinito. En general, cuando tenemos una indeterminación en la que aparece una exponencial y cualquier función polinomial, siempre "gana" la exponencial. De forma rigurosa, tenemos:

$$\lim_{x\to\infty} x^2\, e^{-x/\theta} = \lim_{x\to\infty}\frac{x^2}{e^{x/\theta}} \overset{\text{L'Hôpital}}{=} \lim_{x\to\infty}\frac{(x^2)'}{(e^{x/\theta})'} = \lim_{x\to\infty}\frac{2x}{e^{x/\theta}/\theta} \overset{\text{L'Hôpital}}{=} \lim_{x\to\infty}\frac{2}{e^{x/\theta}/\theta^2} = 0$$

d) Para la muestra del apartado anterior, dar una aproximación de la varianza asintótica del estimador de máxima verosimilitud de θ.

Sabemos que el EMV es asintóticamente normal, y que su varianza asintótica es aproximadamente $-\frac{1}{\frac{\partial^2}{\partial\theta^2}\ell(\theta)|_{\theta=\hat{\theta}_{MV}}}$. En nuestro caso, en el apartado a) (paso 4) hemos obtenido que

$$\frac{\partial^2}{\partial\theta^2}\ell(\theta)|_{\theta=\hat{\theta}_{MV}} = \frac{-2n}{\hat{\theta}_{MV}^2}$$

por tanto, la aproximación de la varianza asintótica de $\hat{\theta}_{MV}$ para esta muestra particular será:

$$\begin{aligned} Var\left[\hat{\theta}_{MV}\right] &\overset{A}{\approx} -\frac{1}{\frac{\partial^2}{\partial\theta^2}\ell(\theta)|_{\theta=\hat{\theta}_{MV}}} = -\frac{1}{\frac{-2n}{\hat{\theta}_{MV}^2}} = \frac{\hat{\theta}_{MV}^2}{2n} \\ &= \frac{(\overline{x}/2)^2}{2n} \overset{c)}{=} \frac{2.8507^2}{8\cdot 15} = 0.0677 \end{aligned}$$