Ejercicio resuelto

Inferencia

4 de diciembre de 2023

Sea X la variable aleatoria que mide el tiempo, en meses, entre interrupciones consecutivas de suministro ocurridas en cierta instalación eléctrica. Se supone que X sigue una distribución con densidad

```
> f = function (x, mu) mu^(-1) * exp(-x/mu) * (x>0)
```

y que las interrupciones son independientes entre sí. Para estimar la distribución de X se tomó una muestra aleatoria simple cuyos resultados fueron los siguientes:

```
> X = c(3.35, 1.30, 1.23, 2.13, 7.70, 11.61, 8.02, 1.44, 3.26,
+  2.06, 6.06, 0.42, 5.58, 1.40, 6.26, 0.48, 1.46, 7.33, 7.55,
+  1.37, 5.41, 0.23, 3.96, 0.86, 0.85, 1.76, 1.3, 18.36, 2.76,
+  4.22)
```

1. Calcula el estimador máximo-verosímil de μ y el valor de la estimación para la muestra obtenida.

   densidad $f = e^{-x/\mu}/\mu \quad \Rightarrow \quad$ verosimilitud $L = e^{-\sum x_i/\mu}/\mu^n \quad \Rightarrow$
   $\ln L = -\sum x_i/\mu - n \ln \mu \quad \Rightarrow \quad \partial \ln L/\partial\mu = \sum x_i/\mu^2 - n/\mu \quad \Rightarrow$
   $\{ \quad 0 = \partial \ln L/\partial\mu \quad \Rightarrow \quad \hat{\mu}_{\text{MV}} = \sum x_i/n = \bar{x} \quad \}$

   o bien, sabiendo que X es exponencial$(1/\mu)$ y que $\hat{\lambda}_{\text{MV}} = 1/\bar{x}$, por la propiedad de invariancia funcional del EMV, $\hat{\mu}_{\text{MV}} = 1/\hat{\lambda}_{\text{MV}} = \bar{x}$

   ```
   > mean(X) # estimación
   ```

   ```
   [1] 3.990667
   ```

2. Halla la esperanza y la varianza del estimador máximo-verosímil de μ.

   X_i exponencial$(1/\mu)$, independientes $\quad\Rightarrow\quad \sum X_i$ gamma$(n, 1/\mu)$
   $\Rightarrow \quad \hat{\mu}_{\mathrm{MV}}$ gamma$(n, n/\mu) \quad\Rightarrow\quad$ esperanza $= n/(n/\mu) = \mu$,
   varianza $= n/(n/\mu)^2 = \mu^2/n$

   o bien, sin usar que X es exponencial,

$$E(\hat{\mu}_{\mathrm{MV}}) = E\left(\frac{\sum X_i}{n}\right) = \frac{\sum E(X_i)}{n} = \frac{\sum E(X)}{n} = \frac{n \cdot E(X)}{n} = E(X)$$

$$V(\hat{\mu}_{\mathrm{MV}}) = V\left(\frac{\sum X_i}{n^2}\right) = \frac{\sum V(X_i)}{n^2} = \frac{\sum V(X)}{n^2} = \frac{n \cdot V(X)}{n^2} = \frac{V(X)}{n}$$

$$\begin{aligned}
E(X) &= \int_0^\infty x \cdot \frac{1}{\mu} \cdot e^{-\frac{x}{\mu}} \cdot dx = \left\langle \begin{array}{cc} u = x/\mu & dv = e^{-x/\mu}dx \\ du = dx/\mu & v = -\mu e^{-x/\mu} \end{array} \right\rangle \\
&= \left[-x^2 e^{-x/\mu}\right]_{x=0}^\infty - \int_0^\infty -e^{-x/\mu} dx = 0 + \mu = \mu \\
E(X^2) &= \int_0^\infty x^2 \cdot \frac{1}{\mu} \cdot e^{-\frac{x}{\mu}} \cdot dx = \left\langle \begin{array}{cc} u = x^2/\mu & dv = e^{-x/\mu}dx \\ du = 2x\,dx/\mu & v = -\mu e^{-x/\mu} \end{array} \right\rangle \\
&= \left[-x^2 e^{-x/\mu}\right]_{x=0}^\infty - \int_0^\infty -2x e^{-x/\mu} dx = 0 + 2\int_0^\infty x e^{-x/\mu} dx \\
&= 2\mu \int_0^\infty x \cdot \frac{1}{\mu} \cdot e^{-x/\mu} dx = 2\mu \cdot E(X) = 2\mu^2 \\
V(X) &= E(X^2) - [E(X)]^2 = 2\mu^2 - \mu^2 = \mu^2
\end{aligned}$$

3. Obtén un intervalo de confianza con coeficiente $1 - \alpha = 0{,}95$ para la media y la mediana de X.

```
> (n = length(X))

[1] 30

> alfa = 1 - 0.95
> (a = qgamma(c(alfa/2, 1-alfa/2), n, n))

[1] 0.6746958 1.3882946
```

```
> (muMV = mean(X))

[1] 3.990667
```

media de $X = \mu$; $\hat{\mu}_{\mathrm{MV}}$ gamma$(n, n/\mu)$ $\Rightarrow$ $\hat{\mu}_{\mathrm{MV}}/\mu$ gamma(n, n)
$\Rightarrow$ $0{,}95 = \Pr(0{,}674695800714031 < \hat{\mu}_{\mathrm{MV}}/\mu < 1{,}38829458128622)$
$\Rightarrow$ intervalo de confianza para la media $= [\hat{\mu}_{\mathrm{MV}}/1{,}38829458128622;$ $\hat{\mu}_{\mathrm{MV}}/0{,}674695800714031] = [2{,}87451000706883;\ 5{,}91476434630739]$

mediana de $X = m$ $\Rightarrow$ $1/2 = F(m) = 1 - e^{-m/\mu}$ $\Rightarrow$ $\ln(1/2) = -\ln 2 = -m/\mu$ $\Rightarrow$ $\hat{m}_{\mathrm{MV}} = \ln 2 \cdot \hat{\mu}_{\mathrm{MV}}$ por equivariancia
$\Rightarrow$ intervalo de confianza para $m = \ln 2 \cdot$ intervalo de confianza para μ $= [1{,}99245850689111;\ 4{,}09980223031946]$

intervalo para μ mediante *bootstrap* :

```
>     B = 10000
>     ## bootstrap percentil
>     distri = replicate (B,
+     {
+         Xb = sample(X,,TRUE)
+         mean(Xb)
+     })
>     quantile (distri, c(alfa/2, 1-alfa/2))

    2.5%    97.5%
2.705333 5.491675

>     ## bootstrap básico
>     distri = replicate (B,
+     {
+         Xb = sample(X,,TRUE)
+         mean(Xb) - muMV
+     })
>     muMV - quantile (distri, c(1-alfa/2, alfa/2))

  97.5%    2.5%
2.49625 5.27935
```

intervalo exacto de longitud óptima:

```
> ## alfa == alfa1 + alfa2
> intervalo = function (alfa1)
+     muMV / qgamma(c(1-alfa+alfa1,alfa1), n, n)
> longitud = function (alfa1) diff (intervalo (alfa1))
> (o = optimize (longitud, c(0,alfa)))

$minimum
[1] 0.03802793

$objective
[1] 2.953237

> intervalo (o$minimum) # intervalo óptimo

[1] 2.739056 5.692293
```

4. Calcula la estimación máximo-verosímil de μ, suponiendo que los datos de la muestra estuvieran agrupados en las siguientes clases:

Intervalo	$[0;1)$	$[1;1{,}5)$	$[1{,}5;3)$	$[3;4{,}5)$	$[4{,}5;7{,}5)$	$\geq 7{,}5$
Frecuencia	5	7	4	4	5	5

```
> corte = c(0, 1, 1.5, 3, 4.5, 7.5, Inf)
> frecu = c(5, 7, 4, 4, 5, 5)
> L = function (mu) prod (diff (pexp (corte, 1/mu)) ^ frecu)
> lnL = function (mu)
+     sum (log (diff (pexp(corte, 1/mu))) * frecu)
> (muMVa <- optimize (L, c(0, 100), maximum=TRUE) $ maximum)

[1] 4.053007

> optimize (lnL, c(0, 100), maximum=TRUE) $ maximum

[1] 4.053015
```

salvo por el error de la optimización numérica, el máximo se alcanza en el mismo lugar, en torno a $\mu = 4{,}05301$