

Ejercicio 1, Práctico 2

Parte a

Calculemos el valor esperado de $X \sim \Gamma(\alpha, \beta)$ y su varianza. El valor esperado de una v.a. con pdf $f(x)$ se calcula como la integral

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx,$$

y dado que $f(x) = f(x|\alpha, \beta) = (\Gamma(\alpha)\beta^\alpha)^{-1}x^{\alpha-1}e^{-x/\beta}$, para $0 < x < \infty$ y $\alpha, \beta > 0$, tenemos que

$$E(X) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^{\infty} x^\alpha e^{-x/\beta} dx.$$

Con el c.v. $u = x/\beta$, la integral queda:

$$\int_0^{\infty} x^\alpha e^{-x/\beta} dx = \beta^{\alpha+1} \int_0^{\infty} u^\alpha e^{-u} du = \beta^{\alpha+1} \Gamma(\alpha + 1).$$

Sustituyendo en $E(x)$, y recordando que $\Gamma(z + 1) = z\Gamma(z)$ para todo z con $\Re(z) > -1$, concluimos que:

$$E(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \beta^{\alpha+1} \Gamma(\alpha + 1) = \alpha\beta.$$

Calculemos la varianza de X . Con $E(X^2)$ nos encontramos con la integral

$$\int_0^{\infty} x^{\alpha+1} e^{-x/\beta} dx = \beta^{\alpha+2} \int_0^{\infty} u^{\alpha+1} e^{-u} du = \beta^{\alpha+2} \Gamma(\alpha + 2).$$

Por tanto, $E(X^2) = (\beta^{\alpha+2} \Gamma(\alpha + 2)) / (\beta^\alpha \Gamma(\alpha)) = \beta^2(\alpha + 1)\alpha$, y

$$\text{Var}(X) = E(X^2) - E(X)^2 = (\alpha + 1)\alpha\beta^2 - \alpha^2\beta^2 = \alpha\beta^2.$$

Parte b

Recordemos la noción de función de verosimilitud.

Def. (función de verosimilitud). Sea $f(x|\theta)$ la función pdf conjunta o mdf de la muestra $X = (X_1, \dots, X_n)$. Entonces, dado que $X = x$ fue observado, la función de θ definida como

$$L(\theta|x) = f(x|\theta)$$

se llama función de verosimilitud.

La conexión entre la estimación puntual de un parámetro y la función de máxima verosimilitud, $L(\theta|x)$, está comprendida por la siguiente definición de estimador máximo verosímil (MLE).

Def. (MLE). Para cada dato muestral x , sea $\hat{\theta}$ el valor de un parámetro en el cual $L(\theta|x)$ alcanza su máximo como función de θ , dejando a x fijo. Un estimador de máxima verosimilitud (MLE) del parámetro θ basado en la muestra X es $\hat{\theta}(X)$.

En esta parte asumimos que el parámetro α es conocido y deseamos realizar la estimación del parámetro β utilizando el MLE asociado. La función de verosimilitud es

$$L(\beta|x) = \prod_{i=1}^n \frac{1}{\Gamma(\alpha)\beta^\alpha} x_i^{\alpha-1} e^{-x_i/\beta} = \frac{1}{\Gamma(\alpha)^n \beta^{n\alpha}} \left[\prod_{i=1}^n x_i \right]^{\alpha-1} e^{-\sum_i x_i/\beta}.$$

Es conveniente trabajar con la log-verosimilitud. Como el logaritmo es una función monótona, a los efectos de hallar el máximo de L podemos calcularlo utilizando la log-verosimilitud.

$$\log L(\beta|x) = -n \log \Gamma(\alpha) - n\alpha \log \beta + (\alpha - 1) \sum_i \log x_i - \beta^{-1} \sum_i x_i.$$

Por lo tanto,

$$\frac{\partial \log L}{\partial \beta} = -\frac{n\alpha}{\beta} + \beta^{-2} \sum_i x_i.$$

Haciendo esta derivada parcial igual a cero, y resolviendo para β , encontramos que

$$\hat{\beta} = \frac{\sum_i x_i}{n\alpha} = \frac{\bar{X}}{\alpha}.$$

Verifiquemos que efectivamente se trata de un máximo:

$$\frac{\partial^2 \log L}{\partial \beta^2} \Big|_{\beta=\hat{\beta}} = \left[\frac{n\alpha}{\beta^2} - \frac{2 \sum_i x_i}{\beta^3} \right] \Big|_{\beta=\hat{\beta}} = -\frac{(n\alpha)^3}{(\sum_i x_i)^2} < 0.$$

Dado que $\hat{\beta}$ es el único punto donde la derivada de L es cero y es un máximo local, entonces es un máximo global. Deducimos que $\hat{\beta}$ es el MLE.

Parte c

En esta parte, la función de verosimilitud es

$$L(\alpha, \beta|x) = \frac{1}{\Gamma(\alpha)^n \beta^{n\alpha}} \left[\prod_{i=1}^n x_i \right]^{\alpha-1} e^{-\sum_i x_i/\beta}.$$

Es decir la misma que en la parte (a), salvo que ahora tanto β como α se consideran variables.

El cálculo analítico de $\hat{\alpha}$ y de $\hat{\beta}$ no se puede hacer utilizando funciones elementales, ya que intervienen

derivadas de la función gamma. Estas derivadas se describen con la función especial [polygamma](#). Es decir, los valores analíticos de $\hat{\alpha}$ y de $\hat{\beta}$ nos quedarían expresados en términos de ciertas funciones especiales.

El enfoque alternativo sugerido en este ejercicio es realizar el cálculo de los MLE numéricamente, con la salvedad siguiente. Conocemos el máximo de L para cada valor fijo de α , e inyectamos el maximizador ($\beta = \beta(\alpha)$) en la fórmula de $L(\alpha, \beta)$. De esta forma hemos transformado un problema en principio bidimensional a un problema de optimización esencialmente unidimensional.

Entonces por la parte (b), tenemos una fórmula para $\hat{\beta}$ como función de $\hat{\alpha}$. Por tanto, podemos sustituir dicha fórmula en L :

$$L(\alpha, \beta = \bar{X}/\alpha|x) = \frac{1}{\Gamma(\alpha)^n (\bar{X}/\alpha)^{n\alpha}} \left[\prod_{i=1}^n x_i \right]^{\alpha-1} e^{-n\alpha}.$$

A continuación escribimos un programa en R para calcular el máximo de esta función dados los datos del problema.

El comando `c` de R se utiliza para concatenar argumentos:

```
1 # ingresar datos muestrales
2 x = c(22.0, 23.9, 20.9, 23.8, 25.0, 24.0, 21.7, 23.8, 22.8, 23.1, 23.1, 23.5,
```

Chequeo el primer elemento y la cantidad de datos en el vector `x`:

```
1 > x[1]
2 [1] 22
3
4 > n = length(x); n
5 [1] 14
```

Como `x` tiene números relativamente "grandes", conviene trabajar con la log-verosimilitud. Nótese que, por ejemplo

```
1 > gamma(20)
2 [1] 1.216451e+17
3
4 > log(gamma(20))
5 [1] 39.33988
6
7 > lgamma(20)
8 [1] 39.33988
```

Aquí `lgamma` es la función especial que corresponde al logaritmo natural del valor absoluto de la función gamma. Conviene trabajar con esta función ya que es más precisa que primero calcular `gamma` y luego tomar el logaritmo (ver ejemplos más abajo).

La log-verosimilitud es

$$\log L(\alpha, \beta = \bar{X}/\alpha|x) = -n \log \Gamma(\alpha) - n\alpha \log(\bar{X}/\alpha) + (\alpha - 1) \sum_i \log x_i - n\alpha.$$

La media muestral se calcula con la función `mean` :

```
1 > xmean = mean(x)
2 > xmean
3 [1] 23.11429
```

Primero ingresamos la función $\log L(\alpha, \beta = \bar{X}/\alpha|x)$:

```
1 logL <- function(alpha){
2   -n * lgamma(alpha) - n * alpha * log(xmean/alpha) + (alpha-1) * sum(log(x))
3 }
```

Evaluamos algunos valores:

```
1 > logL(10)
2 [1] -40.95251
3 > logL(100)
4 [1] -25.95473
5 > logL(500)
6 [1] -20.12502
```

A modo de comparación, si hubiéramos utilizado directamente la función L , no podríamos obtener estos valores, ya que escapamos del rango de valores admisibles de las funciones especiales que se utilizan en su definición:

```
1 > L <- function(alpha){
2   prod(x)^(alpha-1) * exp(-n*alpha) / gamma(alpha)^n / (sum(x)/(n*alpha))^(alp
3 }
4
5 > L(10)
6 [1] 1.638892e-18
7 > L(100)
8 [1] NaN
9 > L(500)
10 [1] NaN
11 Warning message:
12 In L(500) : value out of range in 'gammafn'
```

Este ejemplo demuestra la importancia de considerar el dominio de validez de los métodos numéricos que utilizamos, para obtener valores correctos.

Continuando con la función log-verosimilitud, queremos hallar el máximo respecto a α . Utilizaremos la función de R [optimize](#) para realizar dicho cálculo. En la documentación vemos que por defecto esta función

hace una minimización. Una alternativa es trabajar con la minimización de $-\log L$.

Por defecto, `optimize` utiliza un metodo de "line search" que necesita dos puntos para acotar el intervalo, tomemos `1` y `1000` con una tolerancia de una parte en un millon como criterio de terminacion:

```
1 > minus_logL <- function(alpha){
2   n * lgamma(alpha) + n * alpha * log(xmean/alpha) - (alpha-1) * sum(log(x)) +
3 }
4
5 > optimize(minus_logL, c(1, 1e3), tol=1e-6)
6 $minimum
7 [1] 514.3353
8
9 $objective
10 [1] 20.12224
11
12 > mean(x) / 514.3353
13 [1] 0.04494011
```

Concluimos que $\hat{\alpha} = 514$ y $\hat{\beta} = 0.0449$.

Nota. Inspeccionando la documentación de `optimize` vemos que tiene un argumento `maximum` que es por defecto falso (`FALSE`):

```
1 > optimize
2 function (f, interval, ..., lower = min(interval), upper = max(interval),
3   maximum = FALSE, tol = .Machine$double.eps^0.25)
4 {
5   if (maximum) {
6     val <- .External2(C_do_fmin, function(arg) -f(arg, ...),
7       lower, upper, tol)
8     list(maximum = val, objective = f(val, ...))
9   }
10  else {
11    val <- .External2(C_do_fmin, function(arg) f(arg, ...),
12      lower, upper, tol)
13    list(minimum = val, objective = f(val, ...))
14  }
15 }
16 <bytecode: 0x1123912a8>
17 <environment: namespace:stats>
```

En nuestro caso, como interesa el máximo, podemos trabajar directamente con `logL` utilizando el comando `optimize(logL, c(1, 1e3), maximum=TRUE, tol=1e-6)`.