1. (4 puntos) Sea X una población con distribución $\gamma(p, a)$ de la que se obtiene la muestra aleatoria simple $\vec{x}$ de tamaño 20 siguiente:

```
## c(1.3, 1.2, 3.5, 4, 2.8, 5.1, 1.5, 2.3, 3.8, 4.3, 3, 2, 0.9,
## 2.5, 1.7, 1.6, 1.1, 1.8, 1.4, 1)
```

Teniendo en cuenta que

$$f_X(x) = \frac{a^p}{\Gamma(p)} x^{p-1} e^{-ax} I(x > 0) \qquad E(X) = \frac{p}{a} \qquad \mathrm{Var}(X) = \frac{p}{a^2}$$

a) (1 punto) Halla un estadístico suficiente minimal para $\theta = (p, a)$.

b) (1 punto) Mediante el método de los momentos, halla estimadores para p y a y aplícalos a $\vec{x}$.

c) (1 punto) Halla la estimación máximo-verosímil de p y de a. Compara el resultado con el del método de los momentos.

d) (1 punto) Calcula mediante *bootstrap* un intervalo de confianza para $\mu = p/a$ a nivel $1 - \alpha = 0{,}95$.

Resolución

1. *a*) Halla un estadístico suficiente minimal para $\theta = (p, a)$.

Para $x > 0$, $f(x, p, a) = \frac{a^p}{\Gamma(p)} x^{p-1} e^{-ax} \implies \mathcal{L}(\vec{x}, p, a) = \frac{a^{np}}{\Gamma(p)^n} \left(\prod x_i\right)^{p-1} e^{-a \sum x_i} \implies$ $T := (\prod X_i, \sum X_i)$ es suficiente por el teorema de factorización, tomando $h(\vec{x}) = 1$ y $g\left(\vec{t}, \vec{\theta}\right) = g(t_1, t_2, p, a) = \frac{a^{np}}{\Gamma(p)^n} (t_1)^{p-1} e^{-at_2}$, $t_1(\vec{x}) = \prod x_i$, $t_2(\vec{x}) = \sum x_i$. [Con más rigor, se podría considerar x arbitrario y tomar $h(\vec{x}) = I(x_{(1)})$, etcétera.]

Es minimal porque $\frac{\mathcal{L}(\vec{x},p,a)}{\mathcal{L}(\vec{y},p,a)} = \left(\frac{\prod x_i}{\prod y_i}\right)^{p-1} e^{-a(\sum x_i - \sum y_i)}$ es independiente de (p, a) $\iff \frac{\prod x_i}{\prod y_i} = 1 \wedge \sum x_i - \sum y_i = 0 \iff (\prod x_i, \sum x_i) = (\prod y_i, \sum y_i) \iff T(\vec{x}) = T(\vec{y})$.

Otra forma de verlo sería considerando que la distribución γ pertenece a la familia exponencial 2-paramétrica, pues su soporte $(0, \infty)$ no depende de θ y

$$f(x, p, a) = \frac{a^p}{\Gamma(p)} \cdot e^{(p-1)\ln x - ax} = c(\theta) h(x) e^{Q_1(\theta)T_1(x) + Q_2(\theta)T_2(x)}$$

con $c(p, a) = \frac{a^p}{\Gamma(p)}$, $h(x) = 1$, $Q_1(p, a) = p - 1$, $T_1(x) = \ln x$, $Q_2(p, a) = -a$ y $T_2(x) = x$, o bien

$$f(x, \eta_1, \eta_2) = \frac{(-\eta_2)^p}{\Gamma(\eta_1 + 1)} \cdot e^{\eta_1 \ln x + \eta_2 x} = c(\eta) h(x) e^{\eta_1 T_1(x) + \eta_2 T_2(x)}$$

con la parametrización natural $\eta_1 = p - 1$, $\eta_2 = -a$, $\eta = (\eta_1, \eta_2)$, cuyo espacio paramétrico $\mathrm{H} = \{(\eta_1, \eta_2) \mid \eta_1 > -1, \eta_2 < 0\}$ contiene un abierto de $\mathbb{R}^2$ (por ejemplo, el abierto $(0, 1) \times (-1, 1)$ está contenido en H), luego $\left(\sum T_1(X_i), \sum T_2(X_i)\right) =$ $(\sum \ln X_i = \ln \prod X_i, \sum X_i)$ es suficiente minimal, luego también $(\prod X_i, \sum X_i)$ es suficiente minimal, pues inducen la misma partición (ln es inyectiva).

b) Mediante el método de los momentos, halla estimadores para p y a y aplícalos a $\vec{x}$.
$\mu = p/a$, $\sigma^2 = p/a^2 \implies p = \mu^2/\sigma^2$, $a = \mu/\sigma^2 \implies \hat{p}_{\text{MM}} = \bar{X}^2/\hat{S}^2$, $\hat{a}_{\text{MM}} = \bar{X}/\hat{S}^2$ o, con el estimador sesgado varianza muestral, $\hat{p}_{\text{MM}} = \bar{X}^2/S^2$, $\hat{a}_{\text{MM}} = \bar{X}/S^2$

```
mu <- mean(x); s2 <- var(x)
(pMM <- mu^2/s2)

## [1] 3.574151

(aMM <- mu/s2)

## [1] 1.527415

## con el estimador sesgado varianza muestral:
s2 <- mean(x^2) - mu^2
(pMM <- mu^2/s2)

## [1] 3.762265

(aMM <- mu/s2)

## [1] 1.607805
```

Con más rigor, se puede partir de los momentos absolutos:

$$\begin{aligned}
E(X) &= p/a \implies p = a \cdot E(X) \\
E(X^2) &= \text{Var}(X) + E^2(X) = \frac{p}{a^2} + \frac{p^2}{a^2} = \frac{p+p^2}{a^2} \\
&\implies E(X^2) = \frac{a \cdot E(X) + a^2 \cdot E^2(X)}{a^2} \\
&\implies a^2 \cdot E(X^2) = a \cdot E(X) + a^2 \cdot E^2(X) \\
&\implies a \cdot E(X^2) = E(X) + a \cdot E^2(X) \\
&\implies a = \frac{E(X)}{E(X^2) - E^2(X)} \implies \hat{a}_{\text{MM}} = \frac{\bar{X}}{\overline{X^2} - \bar{X}^2} = \frac{\bar{X}}{S^2} \\
&\implies p = \frac{E(X)}{E(X^2) - E^2(X)} E(X) \implies \hat{p}_{\text{MM}} = \frac{\bar{X}}{\overline{X^2} - \bar{X}^2} \bar{X} = \frac{\bar{X}^2}{S^2}
\end{aligned}$$

c) Halla la estimación máximo-verosímil de p y de a. Compara el resultado con el del método de los momentos.
$\mathcal{L}(\vec{x}, p, a) = \frac{a^{np}}{\Gamma(p)^n} \prod x_i^{p-1} e^{-a \sum x_i} \implies$
$l(a) := \ln \mathcal{L}(\vec{x}, p, a) = np \ln a - n \ln \Gamma(p) + (p-1) \sum \ln x_i - a \sum x_i \implies$
$0 = l'(a) = \frac{np}{a} - \sum x_i \implies \begin{cases} l''(a) = -\frac{np}{a^2} < 0 \text{ (máximo)} \\ \hat{a}_{\text{MV}} = \hat{p}_{\text{MV}}/\bar{X} \end{cases}$
$l(p) := \ln \mathcal{L}(\vec{x}, p, a) = np \ln a - n \ln \Gamma(p) + (p-1) \sum \ln x_i - a \sum x_i \implies$
$0 = l'(p) = n \ln a - n \frac{\Gamma'(p)}{\Gamma(p)} + \sum \ln x_i = n \ln \frac{p}{\bar{X}} - n\psi(p) + \sum \ln x_i =$
$n \ln p - \ln \bar{X} - n\psi(p) + \sum \ln x_i = n \ln p - n\psi(p) + \text{constante} \implies$

$\ln p - \psi(p) = \text{constante} \Longrightarrow$ existe a lo sumo una única p que lo satisfaga, ya que ln y ψ son estrictamente crecientes[1] en $(0, \infty)$

Hay que hallar p mediante métodos numéricos, bien como raíz de la ecuación anterior, donde ψ=`digamma`,

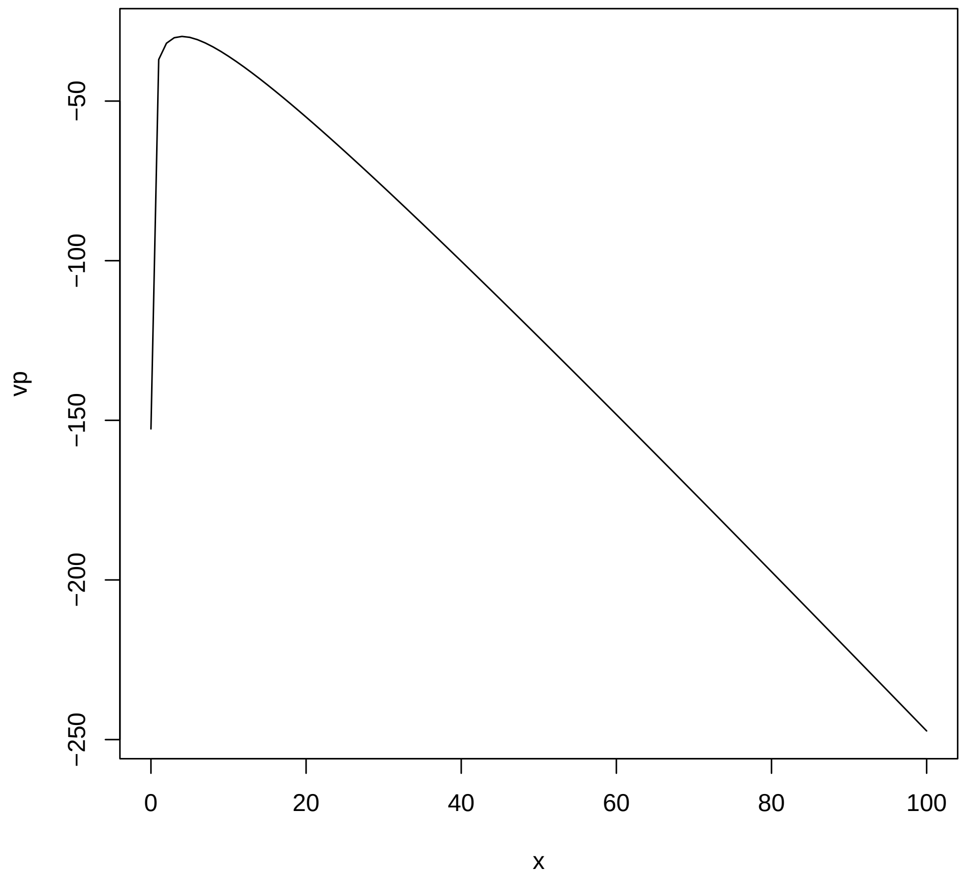
```
g <- function (p) n*log(p/mean(x)) - n*digamma(p) + sum(log(x))
uniroot (g, c (1e-100, 100))

## $root
## [1] 3.997045
##
## $f.root
## [1] -6.296073e-08
##
## $iter
## [1] 12
##
## $init.it
## [1] NA
##
## $estim.prec
## [1] 6.103516e-05
```

bien mediante la maximización de la verosimilitud perfilada en p:
$\ln \mathcal{L}(x, p) = np \ln(p/\bar{x}) - n \ln \Gamma(p) + (p-1) \sum \ln x_i - np$

```
vp <- function(p)
    n*p*log(p/mean(x))-n*log(gamma(p))+(p-1)*sum(log(x))-n*p
plot(vp,.001,100) # para comprobar comportamiento
```

[1] `https://en.wikipedia.org/wiki/Digamma_function`




```
sol <- optimize(vp, c(0,100), maximum=TRUE)
(pMV <- sol$maximum)

## [1] 3.997042

(aMV <- pMV/mean(x))

## [1] 1.708138

## otra forma: verosimilitud perfilada en "a"
vp <- function(a)
    n*a*mean(x)*log(a)-n*log(gamma(a*mean(x)))+
        (a*mean(x)-1)*sum(log(x))-n*a*mean(x)
sol <- optimize(vp, c(0,100), maximum=TRUE)
(aMV <- sol$maximum)
```

```
## [1] 1.708139
```

```r
(pMV <- aMV*mean(x))
```

```
## [1] 3.997044
```

```r
## otra forma: maximizar conjuntamente en (p,a)
optim(c(pMM,aMM),
      function(pa)
      {
          p <- pa[1]
          a <- pa[2]
          sum(dgamma(x,p,a,log=TRUE))
      },
      control=list(fnscale=-1))
```

```
## $par
## [1] 3.998465 1.709014
##
## $value
## [1] -29.75126
##
## $counts
## function gradient
##       41       NA
##
## $convergence
## [1] 0
##
## $message
## NULL
```

Se observa que las estimaciones obtenidas maximizando la verosimilitud son similares a las del método de los momentos. Pueden preferirse las máximo-verosímiles ya que el estimador de máxima verosimilitud suele tener mejores propiedades que el del método de los momentos.

d) Calcula mediante *bootstrap* un intervalo de confianza para $\mu = p/a$ a nivel $1 - \alpha = 0{,}95$.

```r
## bootstrap básico paramétrico
distri <- replicate(10000,
{
    mu <- mean(x); s2 <- var(x); p <- mu^2/s2; a <- mu/s2
    mean(rgamma(n,p,a)-mean(x))
})
mean(x)-quantile(distri,c(.975,.025),names=FALSE)
```

```
## [1] 1.772318 2.858683
```

```r
## bootstrap básico no paramétrico
distri <- replicate(10000, mean(sample(x,,TRUE)-mean(x)))
mean(x)-quantile(distri,c(.975,.025),names=FALSE)
```

```
## [1] 1.794875 2.840000
```

```r
## t-bootstrap: error típico de estima mu = raíz(sigma2/n)
distri <- replicate(10000,
                    mean((xb<-sample(x,,TRUE))-mean(x))/sqrt(var(xb)/n))
mean(x)-quantile(distri,c(.975,.025),names=FALSE)*sqrt(var(x)/n)
```

```
## [1] 1.834237 3.030428
```

```r
## bootstrap percentil
distri <- replicate(10000, mean(sample(x,,TRUE)))
quantile(distri,c(.025,.975),names=FALSE)
```

```
## [1] 1.830 2.905
```

El método paramétrico aprovecha el conocimiento de la distribución, por lo que sería en principio el mejor de los cuatro métodos. También el *t-bootstrap* no paramétrico, en parte, ya que aprovecha la expresión paramétrica del error típico. Podrían combinarse ambos:

```r
## t-bootstrap paramétrico
mu <- mean(x); s2 <- var(x); p <- mu^2/s2; a <- mu/s2
distri <- replicate(10000,
                    mean(((xb<-rgamma(n,p,a))-mean(x))/sqrt(var(xb)/n)))
mean(x)-quantile(distri,c(.975,.025),names=FALSE)*sqrt(var(x)/n)
```

```
## [1] 1.840667 3.045840
```