

TP2

Colin Coërchon et Mathis Guillory

2023-02-17

Contents

Estimation du Maximum de vraisemblance et L'intervalle de confiance	2
Vraisemblance: La loi Bernoulli	2
Vraisemblance pour plusieurs paramètres: La Loi Gamma	6
Normalité asymptotique de l'EMV et l'intervalle de confiance	18

Estimation du Maximum de vraisemblance et L'intervalle de confiance

Vraisemblance: La loi Bernoulli

Question 1

```
n <- 10
x <- rbinom(n,1,0.6)
x

## [1] 1 1 1 1 1 0 0 1 1 1

moyenneEmpirique <- length(x[x==1])/n

moyenneEmpirique

## [1] 0.8
```

Question 2

Soit $X_1, \dots, X_n$ échantillon de n variables aléatoires i.i.d de même lot $B(1, p)$. On a alors:

$$\begin{aligned} L(p; x) &= f(p; x_1, \dots, x_n) \\ &= \prod_{i=1}^n f(p; x_i) && \text{(par indépendance des } (x_i)) \\ &= \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} && \text{(car } \forall x \in \{0, 1\}, \mathbb{P}(X = x) = p^x (1-p)^{1-x}) \\ &= p^{\sum_{i=1}^n x_i} (1-p)^{n - \sum_{i=1}^n x_i} \end{aligned}$$

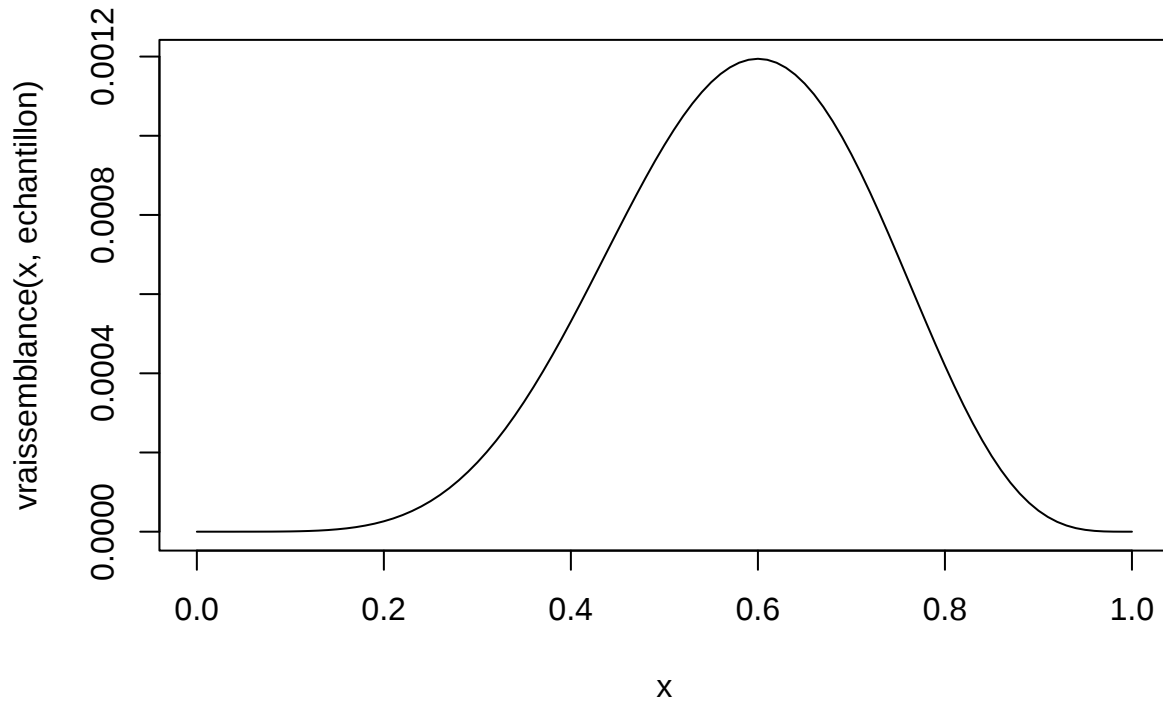
```
vraisemblance <- function (p,x) {
  somme <- sum(x)
  L_bern <- p^(somme) * (1-p)^(n-somme)
  L_bern
}
```

Question 3

```
n <- 10
L_bern <- vraisemblance(0.6, x)
L_bern
```

```
## [1] 0.002687386
```

```
echantillon <- rbinom(n,1,0.6)
curve(vraisemblance(x,echantillon))
```



Plus n est grand plus le maximum de vraisemblance est atteint en $p = 0.6$

Question 4

```
fn <- function (x,e) {
  - vraisemblance(x,e)
}

opti <- optim(0.2, fn, e = echantillon, method = "L-BFGS-B", lower = 0, upper = 1)
p_empirique <- opti$par

opti
```

```
## $par
## [1] 0.5999997
##
## $value
## [1] -0.001194394
##
## $counts
```

```
## function gradient
##      14      14
##
## $convergence
## [1] 0
##
## $message
## [1] "CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH"
```

Question 5

```
tabP <- c()

for(n in seq(10,2000,10)) {
  echant <- rbinom(n, 1, 0.6)
  op <- optim(0.2, fn, e = echant, method = "L-BFGS-B", lower = 0, upper = 1)$par
  tabP <- c(tabP, op)
}
tabP
```

```
## [1] 0.5000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [8] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [15] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [22] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [29] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [36] 0.2000000 0.2000000 0.6304512 0.6108399 0.6656712 0.6179549 0.5732144
## [43] 0.6075730 0.5929506 0.6133527 0.5762366 0.6149121 0.5763121 0.6167514
## [50] 0.6084924 0.6124100 0.5747334 0.6210245 0.5520251 0.6187477 0.5733991
## [57] 0.6214179 0.6333060 0.6074956 0.6246175 0.5674964 0.6170402 0.6083034
## [64] 0.6140674 0.6466872 0.6108292 0.5699798 0.6208689 0.5638962 0.2000000
## [71] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [78] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [85] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [92] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [99] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [106] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [113] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [120] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [127] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [134] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [141] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [148] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [155] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [162] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [169] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [176] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [183] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [190] 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000 0.2000000
## [197] 0.2000000 0.2000000 0.2000000 0.2000000
```

Pour combattre l'instabilité numérique due aux multiplications de probabilités, il peut être alors intéressant de travailler avec la log-vraisemblance. En effet, les multiplications deviennent alors des sommes.

Question 6

On prend comme estimateur: $\hat{p}_n = \frac{1}{n} \sum_{i=1}^n x_i$ qui approche p .

En effet, $\mathbb{E}_p[\hat{p}_n] = p$, et par définition des (X_i) , $V_p(\hat{p}_n) = \frac{p(1-p)}{n}$.

Donc, d'après l'inégalité de Bienaymé-Tchebychev:

$$\mathbb{P}(|\hat{p}_n - p| > \varepsilon) \leq \frac{V_p(\hat{p}_n)}{\varepsilon^2} \leq \frac{1}{4n\varepsilon^2} \quad (\text{car } \forall p \in]0; 1[, p(1-p) \leq \frac{1}{4})$$

Si nous prenons ε tel que $\frac{1}{4n\varepsilon^2} = \alpha$, $\varepsilon = 2\frac{1}{\sqrt{n\alpha}}$, nous obtenons:

$$\begin{aligned} \mathbb{P}(|\hat{p}_n - p| > \frac{1}{2\sqrt{n\alpha}}) &\leq \alpha \\ \implies \mathbb{P}(|\hat{p}_n - p| \leq \frac{1}{2\sqrt{n\alpha}}) &\geq 1 - \alpha \end{aligned}$$

En prenant $\alpha = 10\%$, et en notant: $I_n = \left[\hat{p}_n \pm \frac{1}{2\sqrt{n\alpha}} \right]$,

$$\boxed{\mathbb{P}(p \in I_n) \geq 0.90}$$

On sait que: $\forall i \in \llbracket 1; n \rrbracket$, $X_i \sim B(1, p)$, donc $X_i \in \{0; 1\}$

Donc, $\forall i \in \llbracket 1; n \rrbracket$, $\mathbb{P}(0 \leq X_i \leq 1) = 1$.

Donc d'après l'inégalité de Hoeffding (car $\mathbb{E}[\hat{p}_n] = p$), on a:

$$\mathbb{P}(|\hat{p}_n - p| \geq \varepsilon) \leq 2 \exp(-2n\varepsilon^2)$$

Si on prend: $\alpha = 2 \exp(-2n\varepsilon^2)$, $\varepsilon = \sqrt{\frac{1}{2n} \ln \left(\frac{2}{\alpha} \right)}$, on en déduit de la même manière que,

En posant: $I'_n = \left[\hat{p}_n \pm \sqrt{\frac{1}{2n} \ln \left(\frac{2}{\alpha} \right)} \right]$ avec $\alpha = 10\%$, on a:

$$\boxed{\mathbb{P}(p \in I'_n) \geq 0.90}$$

```
I <- c(p_empirique - (1/(2*sqrt(n*0.1))), p_empirique + (1/(2*sqrt(n*0.1))))
Isecond <- c(p_empirique - sqrt((1/(2*n))*log(2/0.1)), p_empirique + sqrt((1/(2*n))*log(2/0.1)))
I
```

```
## [1] 0.5646444 0.6353551
```

```
Isecond
```

```
## [1] 0.5726331 0.6273664
```

L'intervalle de confiance obtenu grâce à l'inégalité de Hoeffding est plus précis que celui obtenu par l'inégalité de Bienaymé-Tchebychev.

Notre paramètre $p = 0.6$ appartient à ces deux intervalles de confiances de niveau 0.90.

Vraisemblance pour plusieurs paramètres: La Loi Gamma

Question 7

On rappelle que la densité de la loi Gamma s'écrit:

$$\forall x > 0, f(x; \alpha, \beta) = x^{\alpha-1} \frac{\beta^\alpha e^{-\beta x}}{\Gamma(\alpha)}$$

```
n <- 25
theta <- c(2.5, 1.5)

xGamma <- rgamma(n, shape = theta[1], rate = theta[2])
xGamma

## [1] 0.50114043 1.33909692 1.97431135 0.89251218 0.06311232 0.39602974
## [7] 5.10905511 4.04559348 1.27267247 1.58569752 0.08223482 0.55947837
## [13] 1.46822364 2.18678496 0.82099155 3.33793727 0.88082295 2.13990056
## [19] 0.56485270 1.79767551 2.42586556 2.48371821 1.06255119 1.62484834
## [25] 0.72885869

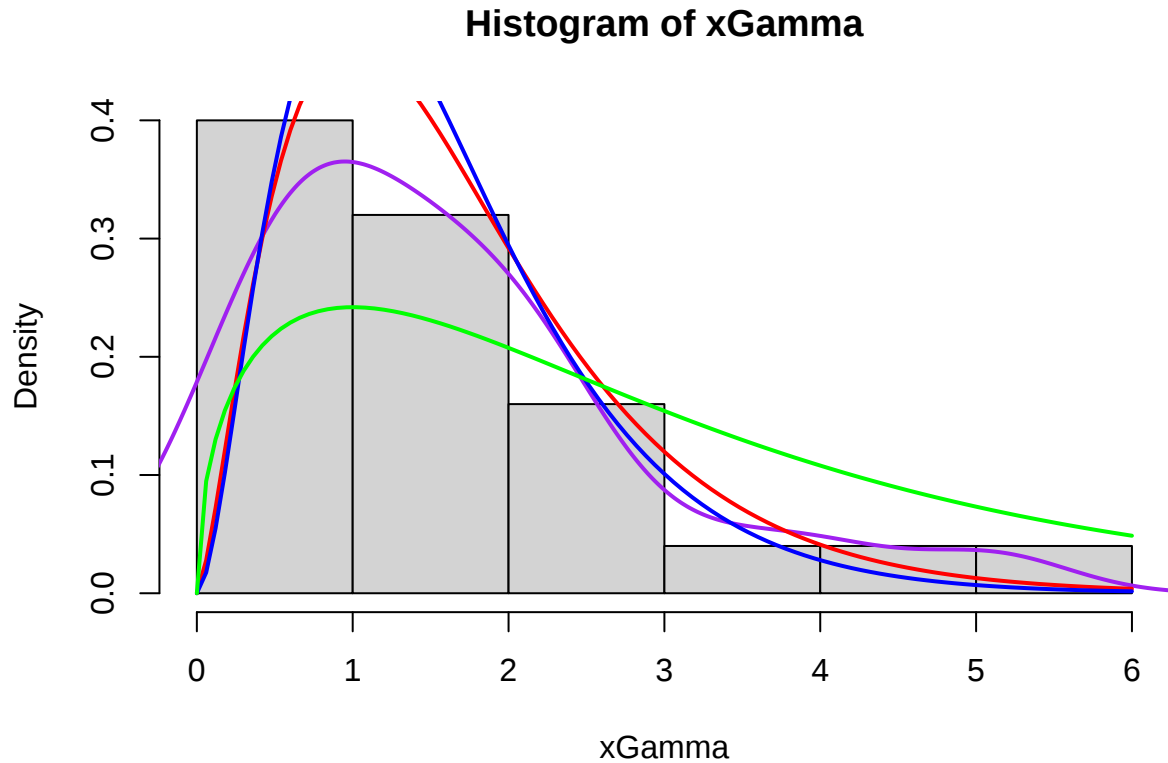
hist(xGamma, probability = TRUE)

lines(density(xGamma), col = "purple", lwd = 2)

curve(dgamma(x, shape = theta[1], rate = theta[2]), add = T, col = "red", lwd = 2)

curve(dgamma(x, shape = theta[1] + 0.3, rate = theta[2] + 0.3), add = T, col = "blue", lwd = 2)

curve(dgamma(x, shape = theta[1] - 1, rate = theta[2] - 1), add = T, col = "green", lwd = 2)
```



La courbe violette correspondant à la courbe “densité” de l’histogramme. La courbe rouge correspond au tracé théorique de la loi Gamma avec $(\alpha = 2.5, \beta = 1.5)$. La courbe bleue correspond au tracé théorique de la loi Gamma avec $(\alpha = 2.8, \beta = 1.8)$. La courbe verte correspond au tracé théorique de la loi Gamma avec $(\alpha = 1.5, \beta = 0.5)$.

Modifier α ne change pas grand chose (cf. rouge et bleu). Par contre changer β modifie beaucoup la loi gamma (cf. rouge et vert). On peut aussi remarquer que l’échantillon est de trop petite taille, $n=10$ ne donne évidemment pas une très bonne approximation de la loi Gamma théorique.

Question 8

$X_1, \dots, X_n$ est un échantillon i.i.d à n variables indépendantes de loi $\Gamma(\alpha, \beta)$.

On a alors:

$$\begin{aligned} L(x_1, \dots, x_n; \alpha, \beta) &= \prod_{i=1}^n x_i^{\alpha-1} \frac{\beta^\alpha e^{-\beta x_i}}{\Gamma(\alpha)} \\ &= \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} \prod_{i=1}^n x_i^{\alpha-1} e^{-\beta x_i} \end{aligned}$$

Donc:

$$\begin{aligned}\mathcal{L}(x_1, \dots, x_n; \alpha, \beta) &= \ln(L(x_1, \dots, x_n; \alpha, \beta)) \\ &= n\alpha \ln(\beta) - n \ln(\Gamma(\alpha)) + (\alpha - 1) \sum_{i=1}^n \ln(x_i) - \beta \sum_{i=1}^n x_i\end{aligned}$$

```
n <- 25
theta <- c(2.5, 1.5)
xGamma <- rgamma(n, shape = theta[1], rate = theta[2])

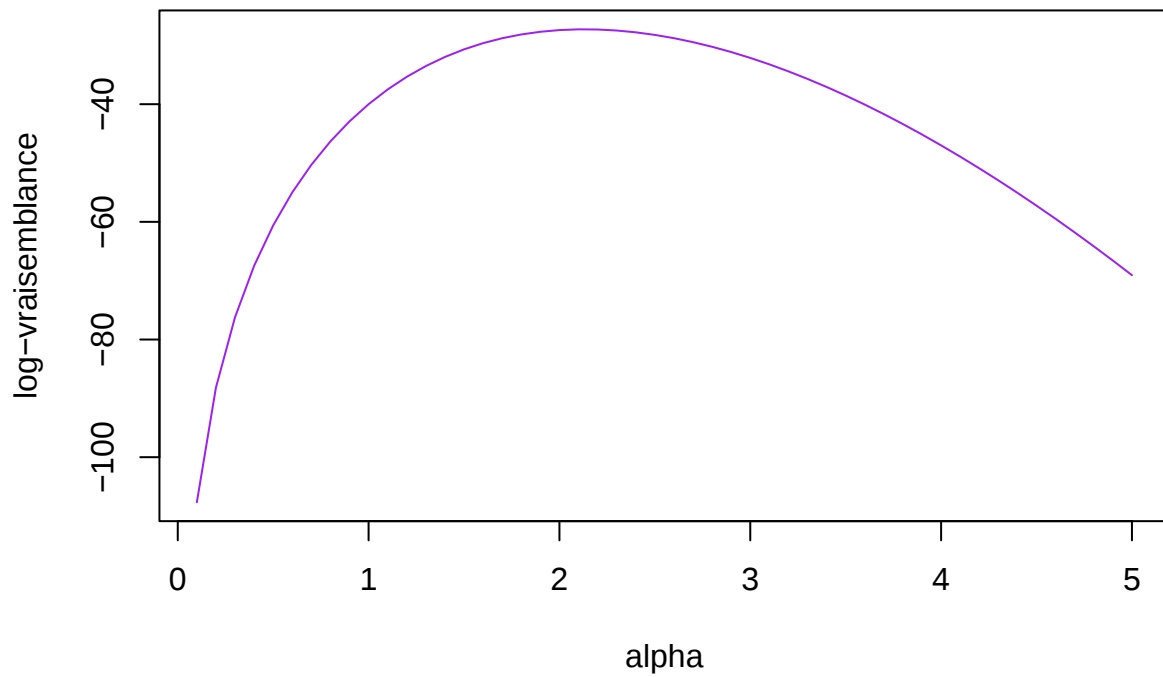
logL_gamma <- function (theta, x) {
  n <- length(x)
  alpha <- theta[1]
  beta <- theta[2]
  return(n * alpha * log(beta) - n * log(gamma(alpha)) + (alpha - 1) * (sum(log(x))) - beta * sum(x))
}

logL_gamma_alpha <- function(alpha, x) {
  beta <- theta[2]
  theta1 <- c(alpha, beta)
  logL_gamma(theta1, x)
}

logL_gamma_beta <- function (beta, x) {
  alpha <- theta[1]
  theta2 <- c(alpha, beta)
  logL_gamma(theta2, x)
}

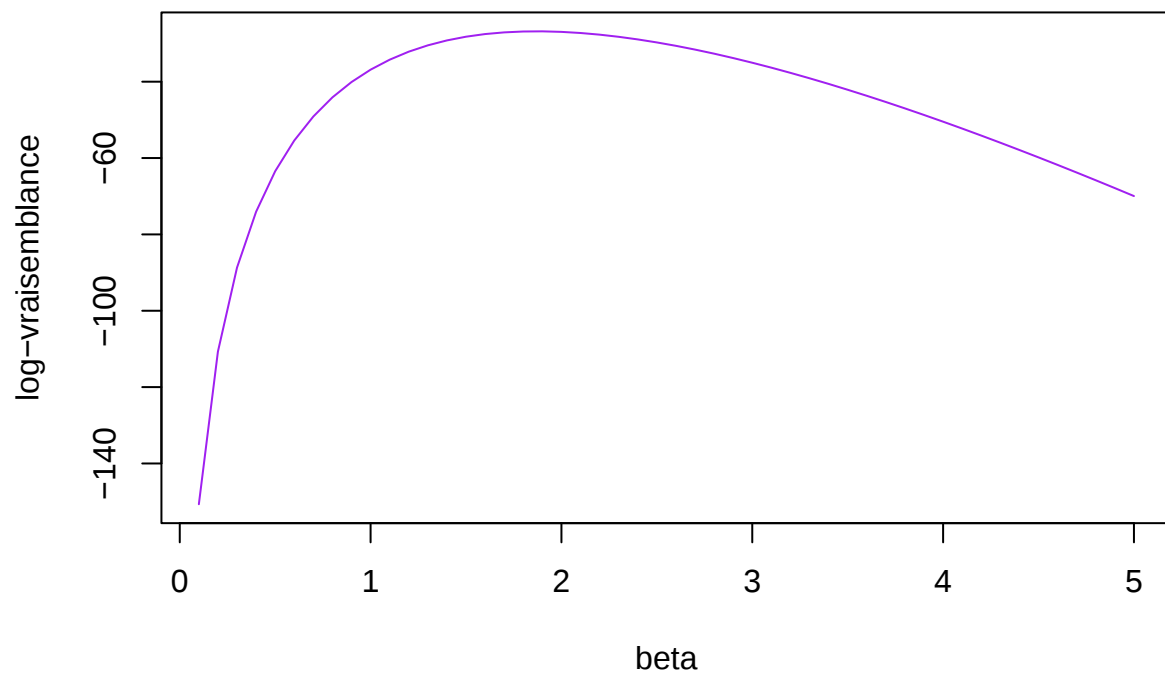
# Estimation de la log-vraisemblance en faisant varier alpha à partir de 0.1
alpha_seq <- seq(0.1, 5, by = 0.1)
logL_alpha <- rep(0, length(alpha_seq))

for (i in 1:length(alpha_seq)) {
  theta_i <- c(alpha_seq[i], theta[2])
  logL_alpha[i] <- logL_gamma(theta_i, xGamma)
}
plot(alpha_seq, logL_alpha, type = "l", xlab = "alpha", ylab = "log-vraisemblance", col = "purple")
```



```
# Estimation de la log-vraisemblance en faisant varier beta à partir de 0.1
beta_seq <- seq(0.1, 5, by = 0.1)
logL_beta <- rep(0, length(beta_seq))

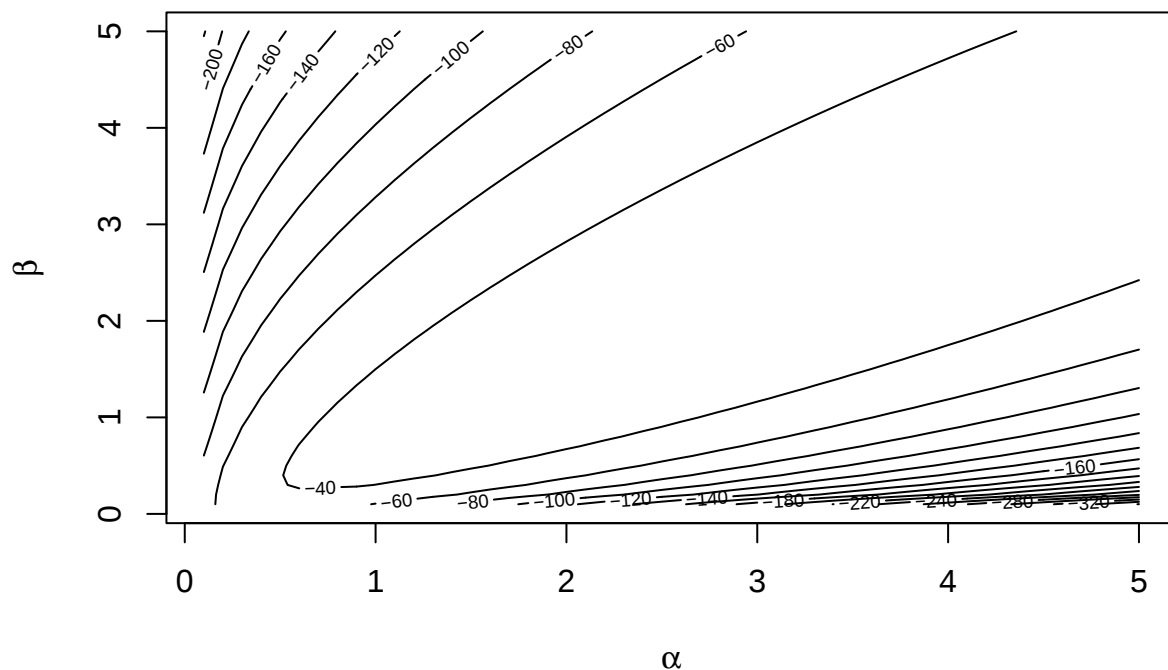
for (i in 1:length(beta_seq)) {
  theta_i <- c(theta[1], beta_seq[i])
  logL_beta[i] <- logL_gamma(theta_i, xGamma)
}
plot(beta_seq, logL_beta, type = "l", xlab = "beta", ylab = "log-vraisemblance", col = "purple")
```



Question 9

```
i <- seq(0.1, 5, by = 0.1)
z <- matrix(0, length(i), length(i))
for (j in 1:length(i)) {
  for (k in 1:length(i)) {
    z[j,k] <- logL_gamma(c(i[j], i[k]), xGamma)
  }
}
contour(i, i, z, nlevels=20,
        xlab = expression(alpha), ylab = expression(beta),
        main = "Contours de la log-vraisemblance")
```

Contours de la log-vraisemblance



Question 10

Le vecteur score $S(X; \alpha, \beta)$ d'une variable aléatoire $X \sim \Gamma(\alpha, \beta)$ s'écrit :

$$S(X; \alpha, \beta) = \nabla \mathcal{L}(x; \alpha, \beta) = \begin{pmatrix} \frac{\partial \mathcal{L}}{\partial \alpha}(x; \alpha, \beta) \\ \frac{\partial \mathcal{L}}{\partial \beta}(x; \alpha, \beta) \end{pmatrix}$$

Or, on déduit de ce qu'on a vu précédemment que :

$$\mathcal{L}(x; \alpha, \beta) = \alpha \ln(\beta) - \ln(\Gamma(\alpha)) + (\alpha - 1) \ln(x) - \beta x$$

Donc :

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \alpha}(x; \alpha, \beta) &= \ln(\beta) - \psi(\alpha) + \ln(x) \\ \frac{\partial \mathcal{L}}{\partial \beta}(x; \alpha, \beta) &= \frac{\alpha}{\beta} - x \end{aligned}$$

Avec $\psi : \alpha \mapsto \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$ la fonction *digamma*. On en déduit alors le vecteur score :

$$S(X; \alpha, \beta) = \begin{pmatrix} \ln(\beta) - \psi(\alpha) + \ln(x) \\ \frac{\alpha}{\beta} - x \end{pmatrix}$$

Or, l'EMV répond à l'équation : $\nabla \mathcal{L}(x_1, \dots, x_n; \hat{\alpha}, \hat{\beta}) = 0$. C'est-à-dire que :

$$\begin{aligned} S_n(X_1, \dots, X_n; \hat{\alpha}, \hat{\beta}) &= 0 \\ \Rightarrow \frac{1}{n} \sum_{i=1}^n S_1(X_i; \hat{\alpha}, \hat{\beta}) &= 0 \\ \Rightarrow \begin{pmatrix} \ln(\hat{\beta}) - \psi(\hat{\alpha}) + \frac{1}{n} \sum_{i=1}^n \ln(x_i) \\ \hat{\alpha} - \frac{1}{n} \sum_{i=1}^n x_i \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

Voici pour le résultat de la fonction optim pour notre estimateur de taille n = 25:

```
opt <- optim(par = c(1,1), fn = logL_gamma, x = xGamma, method = "L-BFGS-B", control=list(fnscale=-1))$
opt
```

```
## [1] 2.760227 2.064497
```

A noter que pour un échantillon n = 1000, le résultat se rapproche bien plus de $\theta_0 = (2.5, 1.5)$.

```
opt <- optim(par = c(1,1), fn = logL_gamma, x = rgamma(1000, 2.5, 1.5), method = "L-BFGS-B", control=list(fnscale=-1))$
opt
```

```
## [1] 2.483674 1.544899
```

Question 11

```
estimations <- matrix(nrow = 100, ncol = 2)

for (i in 1:100) {
  opt <- optim(par = c(1,1), fn = logL_gamma, x = rgamma(25,2.5,1.5), method = "L-BFGS-B", control=list(fnscale=-1))$
  estimations[i,] <- opt
}
```

```
# Calculer les intervalles de confiance empiriques à 95% pour alpha et beta
conf_int_alpha <- quantile(estimations[, 1], c(0.025, 0.975))
conf_int_beta <- quantile(estimations[, 2], c(0.025, 0.975))
```

```
# Afficher les résultats
```

```
cat("Intervalle de confiance empirique à 95% pour alpha:
    [", conf_int_alpha[1], ", ", conf_int_alpha[2], "]\n")
```

```
## Intervalle de confiance empirique à 95% pour alpha:
```

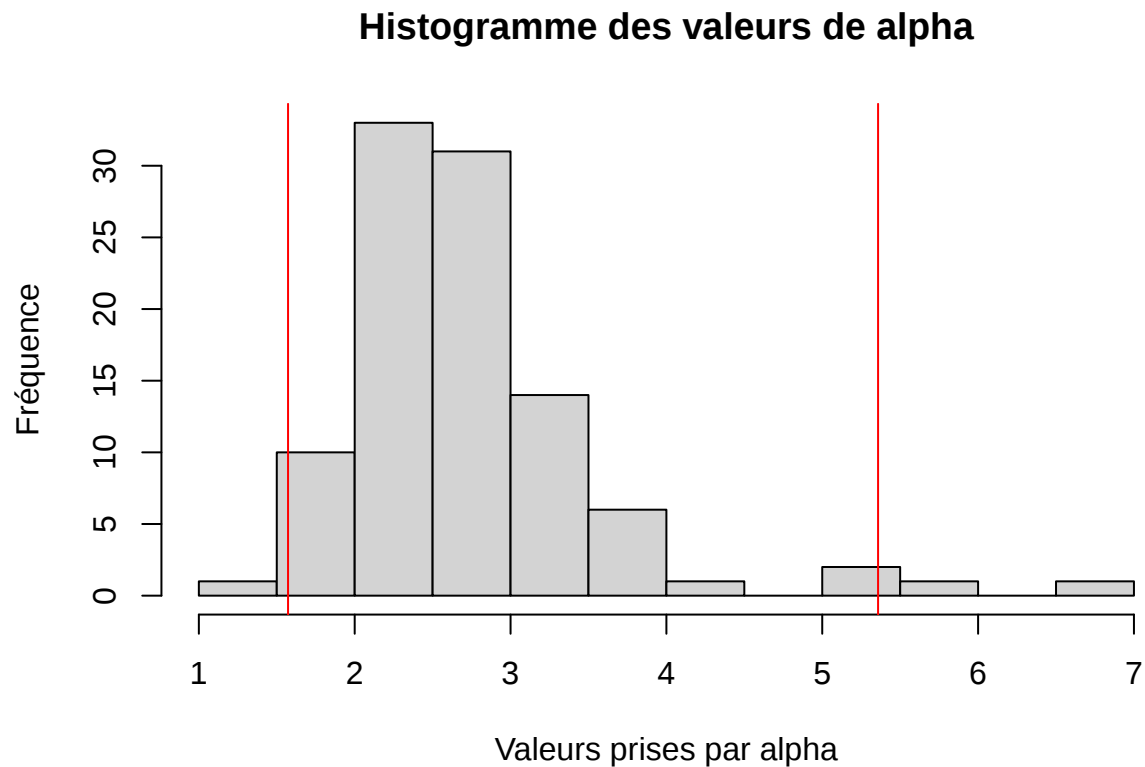
```
## [ 1.572458 , 5.358214 ]
```

```
cat("Intervalle de confiance empirique à 95% pour beta:
    [", conf_int_beta[1], ", ", conf_int_beta[2], "]\n")
```

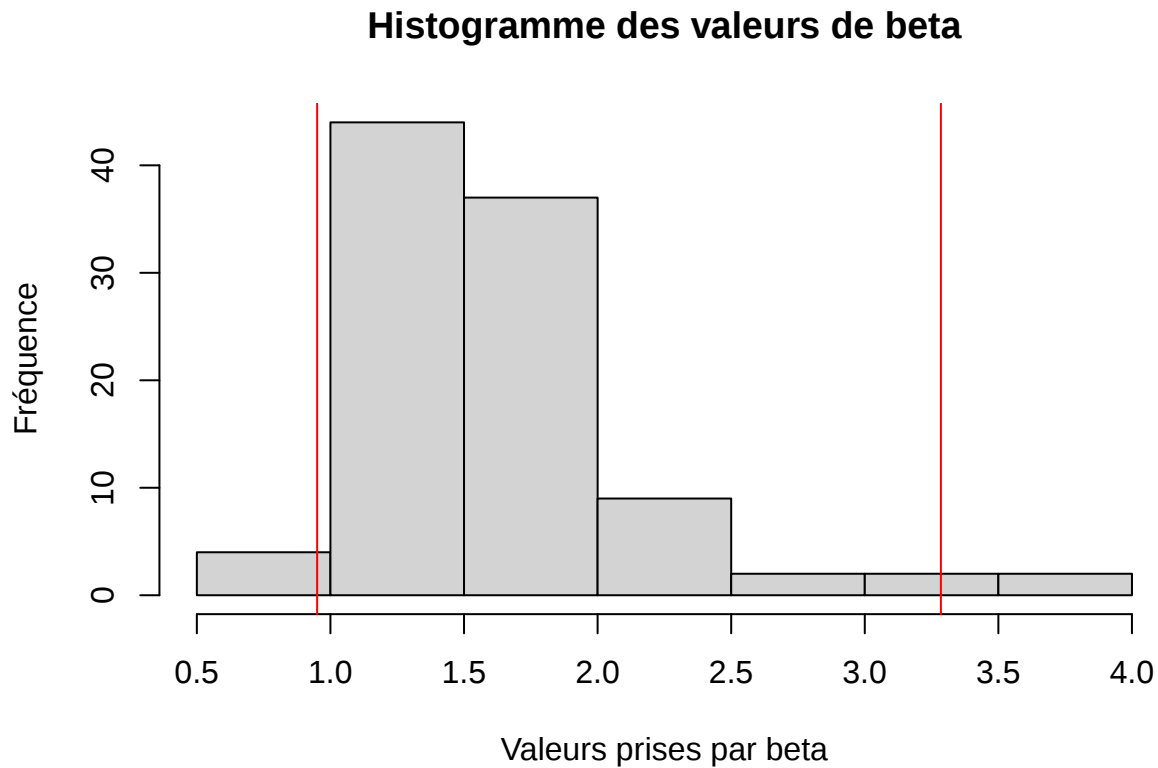
```
## Intervalle de confiance empirique à 95% pour beta:
```

```
## [ 0.9502639 , 3.284821 ]
```

```
hist(estimations[,1], main = "Histogramme des valeurs de alpha",
     xlab = "Valeurs prises par alpha", ylab = "Fréquence")
abline(v = conf_int_alpha[1], col = "red")
abline(v = conf_int_alpha[2], col = "red")
```



```
hist(estimations[,2], main = "Histogramme des valeurs de beta",
     xlab = "Valeurs prises par beta", ylab = "Fréquence")
abline(v = conf_int_beta[1], col = "red")
abline(v = conf_int_beta[2], col = "red")
```



On remarque que α^* et β^* appartiennent à leur intervalle de confiance respectif.

Question 12

Pour des échantillons de taille $n = 15$:

```
estimations <- matrix(nrow = 100, ncol = 2)

for (i in 1:100) {
  opt <- optim(par = c(1,1), fn = logL_gamma, x = rgamma(15,2.5,1.5), method = "L-BFGS-B", control=list(
    estimations[i,] <- opt
  })
}

# Calculer les intervalles de confiance empiriques à 95% pour alpha et beta
conf_int_alpha <- quantile(estimations[, 1], c(0.025, 0.975))
conf_int_beta <- quantile(estimations[, 2], c(0.025, 0.975))

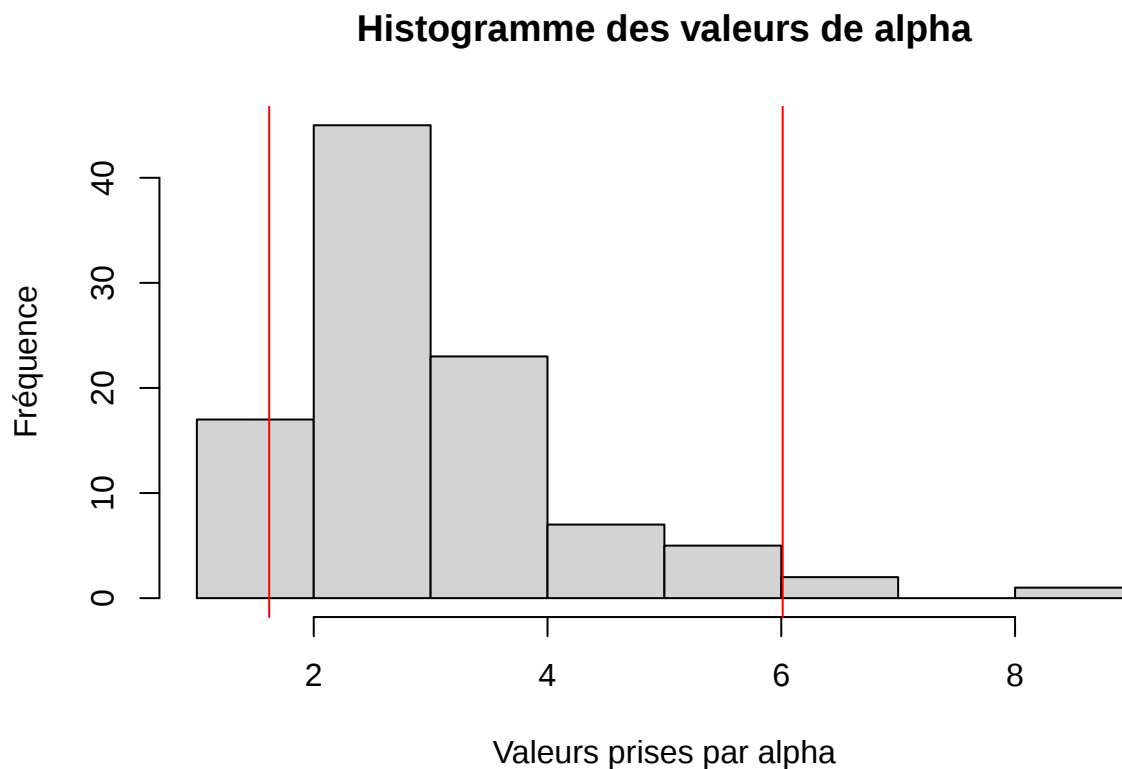
# Afficher les résultats
cat("Intervalle de confiance empirique à 95% pour alpha:
    [", conf_int_alpha[1], ", ", conf_int_alpha[2], "]\n")

## Intervalle de confiance empirique à 95% pour alpha:
##      [ 1.619031 ,  6.012115 ]
```

```
cat("Intervalle de confiance empirique à 95% pour beta:\n", conf_int_beta[1], ", ", conf_int_beta[2], "\n")
```

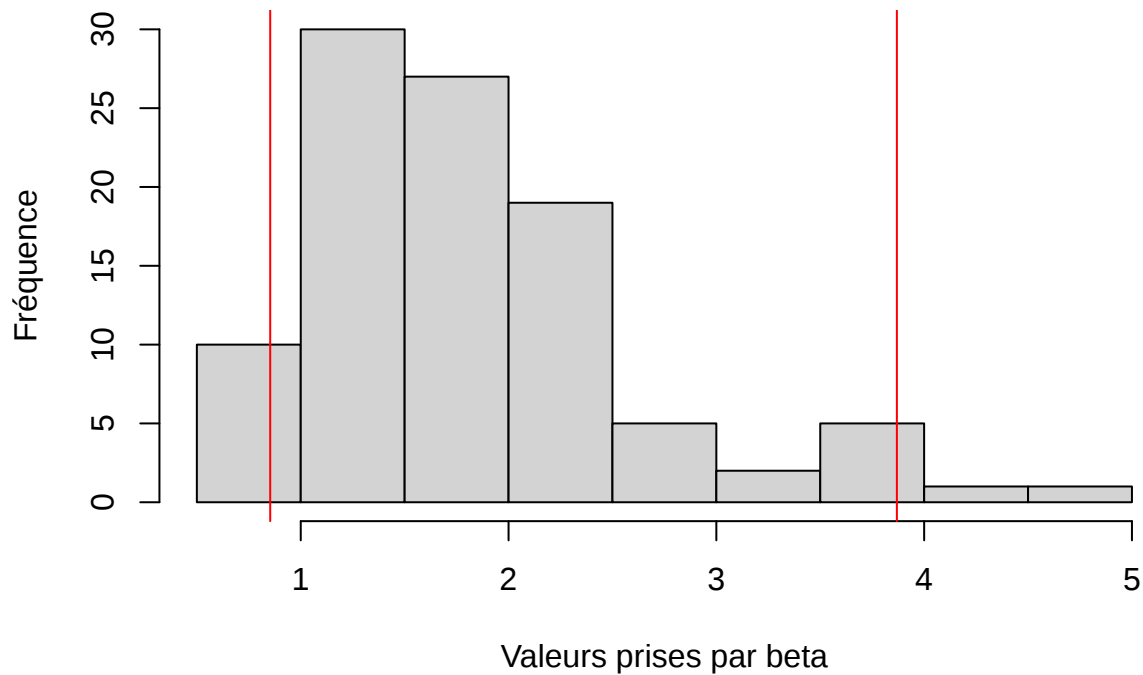
```
## Intervalle de confiance empirique à 95% pour beta:\n##      [ 0.8530253 ,  3.868777 ]
```

```
hist(estimations[,1], main = "Histogramme des valeurs de alpha",\n      xlab = "Valeurs prises par alpha", ylab = "Fréquence")\nabline(v = conf_int_alpha[1], col = "red")\nabline(v = conf_int_alpha[2], col = "red")
```



```
hist(estimations[,2], main = "Histogramme des valeurs de beta",\n      xlab = "Valeurs prises par beta", ylab = "Fréquence")\nabline(v = conf_int_beta[1], col = "red")\nabline(v = conf_int_beta[2], col = "red")
```

Histogramme des valeurs de beta



Pour des échantillons de taille $n = 100$:

```
estimations <- matrix(nrow = 100, ncol = 2)

for (i in 1:100) {
  opt <- optim(par = c(1,1), fn = logL_gamma, x = rgamma(100,2.5,1.5), method = "L-BFGS-B", control=list())
  estimations[i,] <- opt
}
```

```
# Calculer les intervalles de confiance empiriques à 95% pour alpha et beta
conf_int_alpha <- quantile(estimations[, 1], c(0.025, 0.975))
conf_int_beta <- quantile(estimations[, 2], c(0.025, 0.975))
```

```
# Afficher les résultats
```

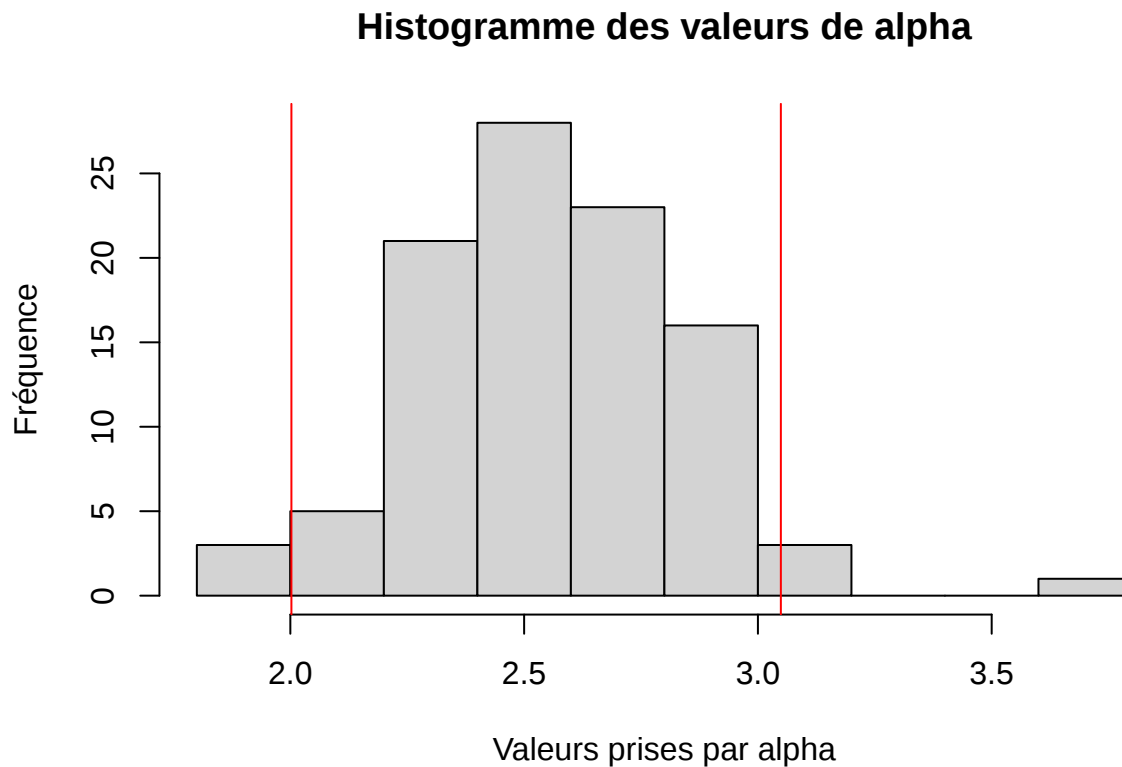
```
cat("Intervalle de confiance empirique à 95% pour alpha:
    [", conf_int_alpha[1], ", ", conf_int_alpha[2], "]\n")
```

```
## Intervalle de confiance empirique à 95% pour alpha:
##      [ 2.002279 ,  3.049049 ]
```

```
cat("Intervalle de confiance empirique à 95% pour beta:
    [", conf_int_beta[1], ", ", conf_int_beta[2], "]\n")
```

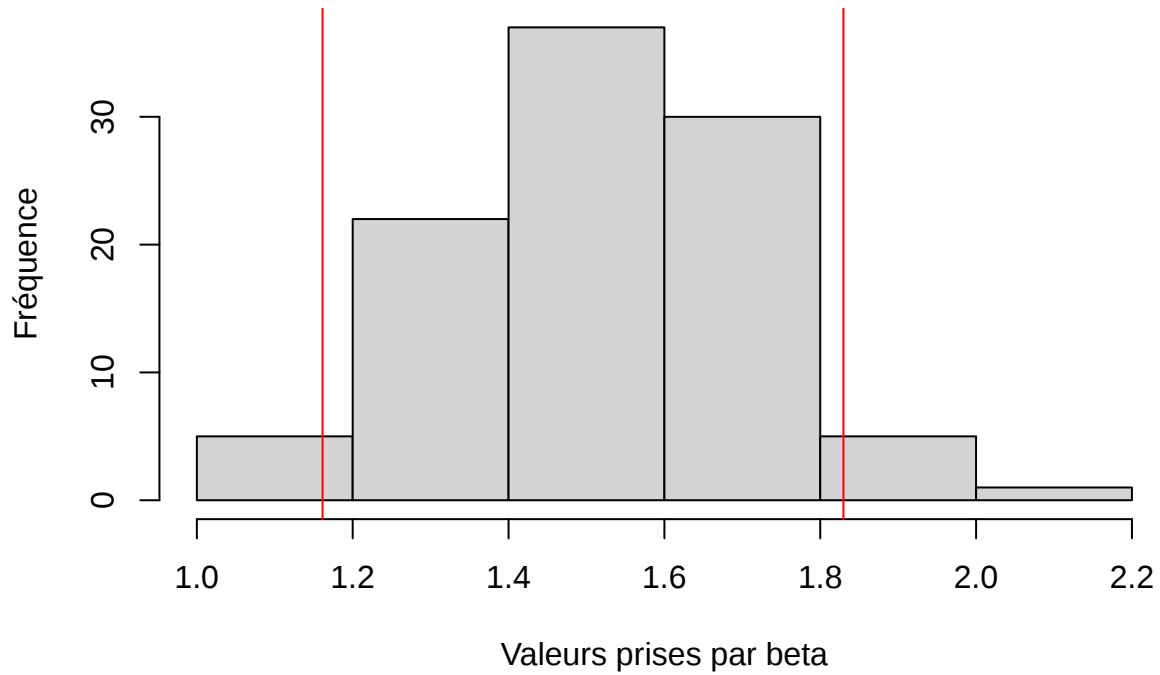
```
## Intervalle de confiance empirique à 95% pour beta:
##      [ 1.161315 ,  1.829742 ]
```

```
hist(estimations[,1], main = "Histogramme des valeurs de alpha",
     xlab = "Valeurs prises par alpha", ylab = "Fréquence")
abline(v = conf_int_alpha[1], col = "red")
abline(v = conf_int_alpha[2], col = "red")
```



```
hist(estimations[,2], main = "Histogramme des valeurs de beta",
     xlab = "Valeurs prises par beta", ylab = "Fréquence")
abline(v = conf_int_beta[1], col = "red")
abline(v = conf_int_beta[2], col = "red")
```

Histogramme des valeurs de beta



Effectivement, plus n est grand, plus l'intervalle de confiance à 95% se rétrécit vers la valeur du vrai paramètre $\alpha_0 = 2.5$ et $\beta_0 = 1.5$.

Normalité asymptotique de l'EMV et l'intervalle de confiance

Question 13

Par définition de l'information de Fisher, on a :

$$\begin{aligned} I(\theta) &= \mathbb{E}_{\theta} [\nabla_{\theta} \mathcal{L}(X, \theta)^t \nabla_{\theta} \mathcal{L}(X, \theta)] && \text{(par définition)} \\ &= -\mathbb{E}_{\theta} [\nabla_{\theta}^2 \mathcal{L}(X, \theta)] && \text{(propriété)} \end{aligned}$$

Pour revenir à l'exemple de la loi Gamma, avec $X \sim \Gamma(\alpha, \beta)$, on a :

$$\nabla_{\theta} \mathcal{L}(X, \theta) = \begin{pmatrix} \ln(\hat{\beta}) - \psi(\hat{\alpha}) + \ln(x) \\ \frac{\hat{\alpha}}{\hat{\beta}} - x \end{pmatrix}$$

Donc :

$$\begin{aligned}\nabla_{\theta}^2 \mathcal{L}(X, \theta) &= H_{\theta}(\mathcal{L}) = \begin{pmatrix} \frac{\partial^2 \mathcal{L}}{\partial \alpha^2}(X, \theta) & \frac{\partial^2 \mathcal{L}}{\partial \alpha \partial \beta}(X, \theta) \\ \frac{\partial^2 \mathcal{L}}{\partial \beta \partial \alpha}(X, \theta) & \frac{\partial^2 \mathcal{L}}{\partial \beta^2}(X, \theta) \end{pmatrix} \\ &= \begin{pmatrix} -\psi'(\alpha) & \frac{1}{\beta} \\ \frac{1}{\beta} & -\frac{\alpha}{\beta^2} \end{pmatrix}\end{aligned}$$

Et donc,

$$I(\theta) = \begin{pmatrix} \psi'(\alpha) & -\frac{1}{\beta} \\ -\frac{1}{\beta} & \frac{\alpha}{\beta^2} \end{pmatrix}$$

Et, en utilisant le fait que : $\hat{\theta} \approx \mathcal{N}(\theta, I_n(\theta)^{-1})$, ou écrit plus précisément :

$$\sqrt{n}(\hat{\theta} - \theta^*) \xrightarrow{\mathcal{L}} \mathcal{N}(0, I(\theta^*)^{-1}) \quad (1)$$

Donc, à l'aide de cette approximation asymptotique de la distribution, on a : $\text{Var}_{\theta}(\hat{\theta}) \approx I_n(\theta)^{-1}$. Or :

$$\begin{aligned}I_n(\theta) &= nI(\theta) = n \begin{pmatrix} \psi'(\alpha) & -\frac{1}{\beta} \\ -\frac{1}{\beta} & \frac{\alpha}{\beta^2} \end{pmatrix} \\ \Rightarrow I_n(\theta)^{-1} &= \frac{1}{n} I(\theta)^{-1} = \frac{1}{n \det(I(\theta))} \begin{pmatrix} \frac{\alpha}{\beta^2} & \frac{1}{\beta} \\ \frac{1}{\beta} & \psi'(\alpha) \end{pmatrix}\end{aligned}$$

Et on a :

$$\det(I(\theta)) = \psi'(\alpha) \times \frac{\alpha}{\beta^2} - \frac{1}{\beta^2} = \frac{1}{\beta^2} (\psi'(\alpha) - 1)$$

Et donc on peut alors estimer la covariance asymptotique :

$$\text{Var}_{\theta}(\hat{\theta}) \approx I_n(\theta)^{-1} = \frac{1}{n(\psi'(\alpha) - 1)} \begin{pmatrix} \alpha & \beta \\ \beta & \beta^2 \psi'(\alpha) \end{pmatrix}$$

À partir de là, on peut alors construire un intervalle de confiance asymptotique de niveau 0,95.

En effet, en reprenant la convergence en Loi (1), on peut écrire que :

$$\sqrt{n} \frac{\hat{\theta}^{MV} - \theta^*}{I^{-1/2}(\theta^*)} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$$

Et, comme nous nous sommes placés dans un modèle statistique **régulier**, le théorème de **consistance** de l'EMV nous assure que :

$$\hat{\theta}^{MV} \xrightarrow{\mathbb{P}} \theta^*$$

Et, comme la fonction $\theta \mapsto I(\theta)$ est continue, on aura aussi d'après le théorème de continuité :

$$I(\hat{\theta}^{MV}) \xrightarrow{\mathbb{P}} I(\theta^*)$$

Et donc, d'après le **lemme de Slutsky**, on en déduit le résultat suivant :

$$\sqrt{n} \frac{\hat{\theta}^{MV} - \theta^*}{I^{-1/2}(\hat{\theta}^{MV})} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$$

Et donc :

$$\mathbb{P} \left(u_{\frac{\alpha}{2}} \leq \sqrt{n} \frac{\hat{\theta}^{MV} - \theta^*}{I^{-1/2}(\hat{\theta}^{MV})} \leq u_{1-\frac{\alpha}{2}} \right) \xrightarrow{n \rightarrow +\infty} 1 - \alpha$$

Et donc, comme on sait que pour une estimation de niveau 0.95, on a : $u_{1-\frac{\alpha}{2}} \approx 1.96$. Donc, notre intervalle de confiance s'écrit :

$$\text{IC} = \left[\hat{\theta}^{MV} \pm 1.96 \frac{I^{-1/2}(\hat{\theta}^{MV})}{\sqrt{n}} \right]$$

Avec :

$$I^{-1/2}(\hat{\theta}^{MV}) = \frac{1}{\sqrt{\psi'(\alpha) - 1}} A \quad \text{où } A \text{ est telle que : } A^2 = I^{-1}(\hat{\theta}^{MV})$$

Note : Cela est possible de définir la matrice A comme **la racine carrée de la matrice** $I^{-1}(\hat{\theta}^{MV})$ car $I^{-1}(\hat{\theta}^{MV})$ est symétrique définie positive (i.e $I^{-1}(\hat{\theta}^{MV}) \in \mathcal{S}_2^{++}(\mathbb{R})$).

Je suppose aussi qu'il y a donc une petite erreur dans le cours correspondant de statistiques à la page sur la "région de confiance pour l'EMV" car la racine carrée semble avoir été perdue en chemin dans la démonstration de l'intervalle de confiance...

Note bis : J'ai de plus trouvé un ancien cours de l'ENSIIE où la racine carrée est bien présente : [cours Cordier](#) (page 43)

Question 14

Question 15