

<p style="text-align: center;">TD4</p> <p style="text-align: center;">Statistique Mathématique</p> <p style="text-align: center;">2022/2023</p>	<p style="text-align: center;">Université de Sousse</p> <p style="text-align: center;">  IHEC Institut des Hautes Études Commerciales de Sousse </p>	<p>Niveau : M1</p> <p>Finance & Actuariat</p> <p>Enseignant :</p> <p>Mohamed Essaied Hamrita</p> <p>mhamrita@gmail.com</p> <p>https://github.com/Hamrita</p>
--	--	--

Exercice 1

Soit X_1, X_2, \dots, X_n une séquence de variables aléatoires indépendantes issue d'une loi Bernoulli de paramètre θ

$$X_i = \begin{cases} 1 & \text{avec une probabilité } \theta \\ 0 & \text{avec une probabilité } 1 - \theta \end{cases}$$

Un échantillon aléatoire de taille $n = 15$ a donné : $X = (1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 0, 1)$.

- 1) Déterminer, $\hat{\theta}$, l'estimateur de θ par la méthode du maximum de vraisemblance.
- 2) Préciser le score, $s(\theta|x_i)$, la hessienne, $H(\theta|x_i)$ et $I_n(\theta)$ la quantité d'information de Fisher. En déduire $\mathbb{E}(\hat{\theta})$ et $\mathbb{V}(\hat{\theta})$. Conclure.
- 3) Effectuer le test $H_0 : \theta = 0.6$ en appliquant :
 - a) le test du rapport de vraisemblance;
 - b) le test de Wald;
 - c) le test de multiplicateur de Lagrange
- 4) Calculer la puissance du test LR pour $\theta_1 = 0.1, 0.2, 0.3, 0.6, 0.8$ et représenter la courbe de la puissance du test en fonction de θ_1 .
- 5) Reprendre les questions précédentes sous le logiciel R.

Solution 1

1) La densité de probabilité de la loi de Bernoulli de paramètre θ est donnée par :

$$f(x_i, \theta) = \theta^{x_i} (1 - \theta)^{1-x_i}, \quad x_i = \{0, 1\}$$

D'où le logarithme de f est $\ln(f(x_i, \theta)) = x_i \ln(\theta) + (1 - x_i) \ln(1 - \theta)$, donc

$$\ell(\theta|x_i) = \sum \ln(f(x_i, \theta)) = \ln(\theta) \sum x_i + \ln(1 - \theta) \sum (1 - x_i) = \ln(\theta) \sum x_i + \ln(1 - \theta)(n - \sum x_i).$$

$$\frac{\partial \ell}{\partial \theta} = 0 \iff \frac{\sum x_i}{\theta} - \frac{n - \sum x_i}{1 - \theta} = 0 \iff \hat{\theta} = \bar{X} = 0.4.$$

2) Le score est $s(\theta|x_i) = \frac{\partial \ell}{\partial \theta} = \frac{\sum x_i}{\theta} - \frac{n - \sum x_i}{1 - \theta} = \frac{\sum (x_i - \theta)}{\theta(1 - \theta)}$ et la hessienne est donnée par :

$$H(\theta|x_i) = \frac{\partial^2 \ell}{\partial \theta^2} = -\frac{\sum x_i}{\theta^2} + \frac{n - \sum x_i}{(1 - \theta)^2}.$$

$$I_n(\theta) = -\mathbb{E}(H(\theta|x_i)) = \mathbb{V}(s(\theta|x_i)) = \frac{n}{\theta(1 - \theta)}.$$

L'espérance mathématique de $\hat{\theta}$ est : $\mathbb{E}(\hat{\theta}) = \mathbb{E}(\bar{X}) = \theta$.

L'estimateur $\hat{\theta}$ est un estimateur sans biais.

La variance de $\hat{\theta}$ est égale à $\mathbb{V}(\bar{X}) = \frac{\theta(1 - \theta)}{n}$. On a $\lim_{n \rightarrow \infty} \mathbb{V}(\hat{\theta}) = 0$, donc $\hat{\theta}$ est convergent, de plus $\mathbb{V}(\hat{\theta}) = I_n^{-1}(\theta)$ donc $\hat{\theta}$ est un estimateur efficace.

3) a) Test LR : La statistique LR est :

$$LR = -2(\ell(\theta_0|x_i) - \ell(\hat{\theta})) \sim \chi^2(1)$$

$$LR = -2[(\ln(0.6) \times 6 + \ln(0.4) \times (15 - 6)) - (\ln(0.4) \times 6 + \ln(0.6) \times (15 - 6))]$$

$$= 2.4327$$

$LR < \chi_{0.95}^2(1)$, donc, au seuil $\alpha = 5\%$, on ne peut pas rejeter l'hypothèse nulle.

b) Test de Wald : La statistique de Wald est :

$$W = g(\hat{\theta})'(J(\theta)I^{-1}(\hat{\theta})J(\theta)')^{-1}g(\hat{\theta}); \quad g(\theta) = \theta - 0.6$$

$$W = (-0.2) \left(1 \times \frac{0.4 \times 0.6}{15} \times 1 \right)^{-1} \times (-0.2) = 2.5.$$

c) Test LM :

$$LM = s'(\theta|x_i) I^{-1}(\theta|x_i) s'(\theta|x_i)$$

$$LM = \left(\frac{\sum (x_i - \theta_0)}{\theta_0(1 - \theta_0)} \right)^2 \frac{\theta_0(1 - \theta_0)}{n} = \left(\frac{6 - 15 \times 0.6}{0.6 \times 0.4} \right)^2 \frac{0.6 \times 0.4}{15} = 2.5.$$

4) On rappelle que la puissance d'un test est égale à la probabilité de rejeter l'hypothèse nulle sachant que l'hypothèse alternative est vraie, i.e : $\pi(\theta) = Pr(D_1|H1) = Pr(W|\theta = \theta_1)$. Or, $W = \{\chi^2(1) < LR\}$, d'où

$$\pi(\beta) = Pr(\chi^2(1) < LR|\theta = \theta_1) = 1 - Pr(\chi^2(1) \geq LR|\theta = \theta_1), \quad LR = -2(\ell(\theta_1|x_i) - \ell(\hat{\theta}|x_i))$$

θ_1	0.1	0.2	0.3	0.6	0.8
LR	9.337	3.139	0.677	2.432	11.457
$\pi(\theta)$	0.9977	0.9235	0.5895	0.8811	0.999

```
x=c(1,0,0,0,1,1,0,0,0,1,0,0,1,0,1)
logL=function(theta,x){
  sum(x*log(theta)+(1-x)*log(1-theta) )
}
theta.hat=optim(fn=logL,par=0.1,control=list(fnscale=-1), hessian=TRUE, x=x)
```

La valeur de l'estimateur

```
theta.hat$par
```

```
[1] 0.4
```

La valeur de la vraisemblance en logarithme

```
theta.hat$value
```

```
[1] -10.09518
```

La puissance du test

```
vs=Vectorize(logL,"theta")
```

```
t1=c(0.1,0.2,0.3,0.6,0.8); (lr=-2*(vs(t1,x)-logL(theta.hat$par,x))) #
```

```
[1] 9.3371604 3.1394889 0.6774726 2.4327906 11.4572550
```

```
(pi_t=pchisq(lr,1))
```

```
[1] 0.9977545 0.9235822 0.5895416 0.8811788 0.9992878
```

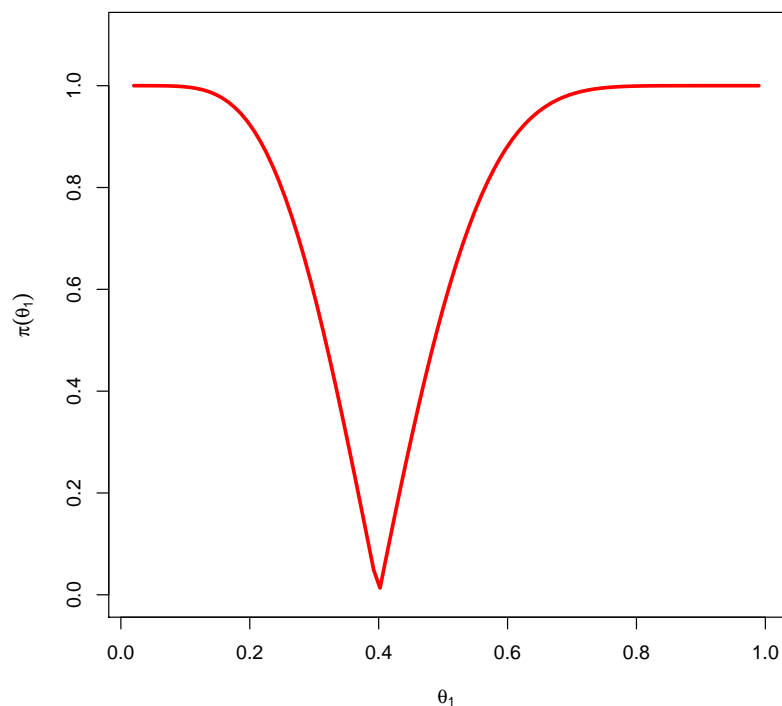
Graphique de la puissance du test

```
tt=seq(0.02,0.99,len=100); lrt=-2*(vs(tt,x)-logL(theta.hat$par,x))
```

```
pit=pchisq(lrt,1)
```

```
plot(tt,pit,type="l", lwd=3, col=2,xlab=expression(theta[1]),
```

```
ylab=expression(pi(theta[1])), ylim=c(0,1.1))
```



LR test

```
l10=logL(0.6,x); l11=logL(0.4,x)
(LR=-2*(l10-l11)); qchisq(0.95,1) # (quantile de chi-deux)

[1] 2.432791
[1] 3.841459
```

Wald test

```
g=function(x) x-0.6
vv=c(solve(-theta.hat$hessian)) # $
W=g(0.4)*vv^{-1}*g(0.4)
W

[1] 2.500024
```

LM test

```
ss=function(x,theta) sum(x-theta)/(theta*(1-theta))
LM=ss(x,0.6)^2*0.6*0.4/15
LM

[1] 2.5
```

Exercice 2

Soit le modèle linéaire suivant : $y_i = \beta x_i + \varepsilon_i$ où $\varepsilon \stackrel{i.i.d}{\sim} N(0, \sigma^2)$. Un échantillon aléatoire de taille $n = 20$ a donné les observations suivantes :

```
x  0  5  0  2  0  0 -2  0 -1  3 -1  2  1 -1  2 -1  3  1  4  2
y -2 13 -4  4 -4 -3 -6 -2 -5  7 -5  5  2 -7  4 -7  6 -1 12  5
```

On souhaite estimer $\theta = (\beta, \sigma^2)$ par la méthode du maximum de vraisemblance.

1) Écrire la densité de la variable $y_i|x_i$. En déduire sa vraisemblance en logarithme.

- 2) Déterminer le vecteur score $s(\theta|y_i)$, la matrice hessienne $H(\theta|y_i)$ et la quantité d'information de Fisher $I_n(\theta|y_i)$.
- 3) Déterminer l'estimateur du vecteur θ par la méthode du maximum de vraisemblance. Étudier ses qualités.
- 4) Tester $H_0 : \beta = 3$ et $\sigma^2 = 9$ par le trigoly (LR, Wald et LM).
- 4) Reprendre les questions avec le logiciel R.

Solution 2

1) On a $y_i = \beta x_i + \varepsilon_i$, donc $\mathbb{E}(y_i) = \beta x_i$ et $\mathbb{V}(y_i) = \sigma^2$. Ainsi, la densité de la variable $y_i|x_i$ est donnée par :

$$f(y_i|x_i, \theta) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(y_i - \beta x_i)^2\right)$$

$$\ln(f(y_i|x_i, \theta)) = -\frac{1}{2} \ln(2\pi\sigma^2) - \frac{1}{2}(y_i - \beta x_i)^2$$

$$\ell(y_i|x_i, \theta) = \sum_{i=1}^n \ln(f(y_i|x_i, \theta)) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \beta x_i)^2$$

2) Le vecteur score est donné comme suit :

$$s(\theta|y_i, x_i) = \left(\frac{\partial \ell}{\partial \beta}, \frac{\partial \ell}{\partial \sigma^2} \right)' = \left(\frac{1}{\sigma^2} \sum_{i=1}^n (x_i(y_i - \beta x_i)), -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (y_i - \beta x_i)^2 \right)'$$

et la matrice hessienne :

$$H(\theta|y_i, x_i) = \begin{pmatrix} -\frac{1}{\sigma^2} \sum_{i=1}^n x_i^2 & -\frac{1}{\sigma^2} \sum_{i=1}^n (x_i(y_i - \beta x_i)) \\ -\frac{1}{\sigma^2} \sum_{i=1}^n (x_i(y_i - \beta x_i)) & \frac{n}{2\sigma^4} - \frac{1}{\sigma^6} \sum_{i=1}^n (y_i - \beta x_i)^2 \end{pmatrix}$$

La quantité d'information de Fisher est définie par

$$I(\boldsymbol{\theta}|y_i, x_i) = -\mathbb{E}(H(\boldsymbol{\theta}|y_i, x_i)) = \mathbb{V}(s(\boldsymbol{\theta}|y_i, x_i))$$

3) La condition nécessaire est traduite par $s(\boldsymbol{\theta}|y_i, x_i) = (0, 0)$.

$$s_1 = 0 \iff \hat{\beta} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} = 2.647 \text{ et } s_2 = 0 \iff \hat{\sigma}^2 = \frac{\sum_{i=1}^n (y_i - \hat{\beta} x_i)^2}{n} = 6.120.$$

$$\mathbb{E}(\hat{\beta}) = \frac{\sum_{i=1}^n x_i \mathbb{E}(y_i)}{\sum_{i=1}^n x_i^2} = \frac{\beta \sum_{i=1}^n x_i^2}{\sum_{i=1}^n x_i^2} = \beta, \text{ donc } \hat{\beta} \text{ est sans biais.}$$

$$\mathbb{E}(\hat{\sigma}^2) = \frac{\sum_{i=1}^n \mathbb{E}(y_i - \mathbb{E}(y_i))^2}{n} = \frac{\sum_{i=1}^n \mathbb{V}(y_i)}{n} = \sigma^2, \text{ donc } \hat{\sigma}^2 \text{ est sans biais.}$$

$$I(\boldsymbol{\theta}|y_i, x_i) = -\mathbb{E}(H(\boldsymbol{\theta}|y_i, x_i)) = \begin{pmatrix} \frac{1}{\sigma^2} \sum_{i=1}^n x_i^2 & 0 \\ 0 & \frac{n}{2\sigma^4} \end{pmatrix}$$

En effet;

$$\mathbb{E}\left(\frac{1}{\sigma^2} \sum_{i=1}^n x_i^2\right) = \frac{1}{\sigma^2} \sum_{i=1}^n x_i^2 \text{ et } \mathbb{E}\left(\frac{1}{\sigma^2} \sum_{i=1}^n (x_i(y_i - \beta x_i))\right) = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i \mathbb{E}(y_i - \beta x_i)) = 0$$

$$\mathbb{E}\left(-\frac{n}{2\sigma^4} + \frac{1}{\sigma^6} \sum_{i=1}^n (y_i - \beta x_i)^2\right) = -\frac{n}{2\sigma^4} + \frac{1}{\sigma^6} \sum_{i=1}^n \mathbb{E}(y_i - \beta x_i)^2 = -\frac{n}{2\sigma^4} + \frac{n\sigma^2}{\sigma^6} = \frac{n}{2\sigma^4}$$

Donc la variance du vecteur $\hat{\boldsymbol{\theta}}$ est estimée par :

$$\hat{\mathbb{V}}(\hat{\boldsymbol{\theta}}) = \hat{I}^{-1}(\hat{\boldsymbol{\theta}}) = \begin{pmatrix} \frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2} & 0 \\ 0 & \frac{2\hat{\sigma}^4}{n} \end{pmatrix} = \begin{pmatrix} 0.0720 & 0 \\ 0 & 3.749 \end{pmatrix}$$

4)

```
x=c(0,5,0,2,0,0,-2,0,-1,3,-1,2,1,-1,2,-1,3,1,4,2)
y=c(-2,13,-4,4,-4,-3,-6,-2,-5,7,-5,5,2,-7,4,-7,6,-1,12,5)
vs=function(theta,x,y){ n=length(y)
-n/2*log(2*pi*theta[2])-1/(2*theta[2])*sum((y-theta[1]*x)^2)
}
theta=optim(fn=vs, par=c(0,1),control=list(fnscale=-1),x=x,y=y,hessian=T)
theta$par # estimateurs

[1] 2.647277 6.123271

theta$value # vraisemblance

[1] -46.49535

round(theta$hessian,3) # hessienne

      [,1] [,2]
[1,] -13.881 0.000
[2,]  0.000 -0.266

round(solve(-theta$hessian),3) # la matrice des variances

      [,1] [,2]
[1,] 0.072 0.000
[2,] 0.000 3.753
```

Test LR

```
l11=theta$value ; l10=vs(c(3,9),x,y); LR=-2*(l10-l11); LR

[1] 2.489103

qchisq(0.95,2)

[1] 5.991465
```


Test de Wald

$$g(\theta) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{\sigma}^2 \end{pmatrix} - \begin{pmatrix} 3 \\ 9 \end{pmatrix} \text{ et le jacobien est égale à 1.}$$

```
g=diag(1,2)%*%theta$par-c(3,9)
W=t(g)%*%(-theta$hessian)%*%g
c(W)

[1] 3.931256

# LM test
s.theta=c(1/(9)*sum(x*(y-3*x)),
          -length(x)/(2*9)+1/(2*9^2)*sum((y-3*x)^2))
LM=t(s.theta)%*%solve(-theta$hessian)%*%s.theta; c(LM)

[1] 1.116559
```

Exercice 3

Soit le processus auto-régressif d'ordre 1 défini par :

$$Y_t = \phi Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \stackrel{i.i.d}{\sim} N(0, \sigma^2)$$

On désire estimer le vecteur $\theta = (\phi, \sigma^2)$ par la méthode du maximum de vraisemblance.

Un échantillon de taille $n = 20$ a donné :

```
[1] 1.07 0.53 -0.17 -0.10 -1.67 1.07 -1.59 -0.06 -0.06 2.15 -0.74 0.95
[13] -0.96 -0.74 0.71 0.10 -1.71 0.02 0.63 0.57
```

- 1) Déterminer la densité de la variable aléatoire $Y_t|Y_{t-1}$. En déduire la vraisemblance associée.

- 2) Déterminer le vecteur score, $s(\boldsymbol{\theta}|Y_{t-1})$, la matrice hessienne, $H(\boldsymbol{\theta}|Y_{t-1})$ et la quantité d'information de Fisher, $I_n(\boldsymbol{\theta})$.
- 3) Déterminer l'estimateur de $\boldsymbol{\theta}$ par la méthode du maximum de vraisemblance. Étudier les qualités de l'estimateur (biais, convergence et efficacité).
- 4) Calculer la valeur des critères d'information AIC et BIC. On rappelle :

$$AIC = 2k - 2\log(\ell(\boldsymbol{\theta})); \quad BIC = k\ln(n) - 2\log(\ell(\boldsymbol{\theta}))$$

avec k est le nombre de paramètres estimés et n est nombre d'observations utilisé dans l'estimation.

- 5) En utilisant les tests LR, Wald et LM, tester, au seuil $\alpha = 5\%$, $H_0 : \phi = -0.5$ et $\sigma^2 = 1$.

Solution 3

- 1) Soit $Y_i = (Y_2, Y_3, \dots, Y_n)$ et $X_i = (Y_1, Y_2, \dots, Y_{n-1})$, d'où $Y_i = \phi X_i + \varepsilon_i$ pour $i = 1, 2, \dots, n-1$.

D'après l'exercice précédent, on a :

$$f(y_i|x_i) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(y_i - \phi x_i)^2\right)$$

La vraisemblance est donnée par :

$$\text{Log}L = \ell(y_i|x_i, \boldsymbol{\theta}) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum (y_i - \phi x_i)^2$$

```
Y=c(1.07,0.53,-0.17,-0.1,-1.67,1.07,-1.59,-0.06,-0.06,2.15,-0.74,0.95,-0.96,-0.74,
0.71,0.10,-1.71,0.02,0.63,0.57)
y=Y[-1]; x=Y[-20]
ll=function(theta,x,y){n=length(y)
-n/2*log(2*pi*theta[2])-1/(2*theta[2])*sum((y-theta[1]*x)^2)
}
```

```

estim=optim(fn=ll, par=c(0,0.1), control=list(fnscale=-1), x=x, y=y,hessian=T)
estim$par    $$ l'estimateur

[1] -0.2923829  0.8866790

estim$value    $$ valeur de log-vraisemblance

[1] -25.81531

(H=round(estim$hessian,3))    $$ la hessienne

      [,1]      [,2]
[1,] -21.779   -0.007
[2,]  -0.007  -12.079

round(solve(-H),3)    # la matrice des variances du theta

      [,1]      [,2]
[1,] 0.046  0.000
[2,] 0.000  0.083

(aic=2*2-2*estim$value)    $$ AIC (k=2: nbre de paramètres estimés)

[1] 55.63063

n=length(y)    # nbre d'observations
(bic=2*log(n)-2*estim$value)    $$ BIC

[1] 57.5195

```

5) Tests :

```

# LR test

l10=ll(c(-0.5,1),x,y); l11=estim$value    $$
(LR=-2*(l10-l11));    qchisq(0.95,2)

[1] 0.9672126
[1] 5.991465

```

```

# Wald test

g=diag(1,2)%*%estim$par-c(-0.5,1)
W=t(g)%*%(-estim$hessian)%*%g;  c(W)

[1] 1.093565

# LM test

ss=c(sum(x*(y+0.5*x)),
      -length(x)/2+1/2*sum((y+0.5*x)^2))
LM=t(ss)%*%solve(-estim$hessian)%*%ss;  c(LM)

[1] 0.7764187

```