Université de Caen
Licence 3 MIASHS-Maths

Probabilités et Statistique
Année 2017-2018

TD - partie 3

Exercice 1. Soit X une variable aléatoire qui suit une loi normale $\mathcal{N}(\mu;\sigma^2)$. Sa densité est notée $f_X(x)$ et sa fonction de répartition est notée $F_X(x)$.

1. Soit la variable aléatoire Y définie par $Y = \frac{X-\mu}{\sigma}$
   (a) Quelle est la densité de Y ?
   (b) Déterminer la fonction de répartition de la variable $U = aY + b$, $a \in \mathbb{R}^\star$; et $b \in \mathbb{R}$, en fonction de la fonction de répartition de Y.
   (c) En déduire la loi de U.
2. Soit la variable aléatoire positive H telle que : $X = \ln H$
   (a) Déterminer la fonction de répartition de la variable H, notée $F_H(h)$, en fonction de $F_X(x)$.
   (b) En déduire la densité de H et reconnaître sa loi.
3. Soit la variable aléatoire suivante : $Z = \left(\frac{X-\mu)}{\sigma}\right)^2$
   (a) Déterminer la fonction de répartition de la variable Z, notée $F_Z(z)$, en fonction de $F_X(x)$.
   (b) En déduire la densité de Z et reconnaître sa loi.
   (c) Calculer l'espérance de Z.
   (d) On donne $\mathrm{Var}(Z) = 2$. En déduire $\mathbb{E}(X-\mu)^4$.

Solution.

Exercice 2. Soient X et Y deux v.a i.i.d suivant chacune la loi normale $\mathcal{N}(0,1)$ et soit $U = X + Y$ et $V = X - Y$.

1. Montrer que (U, V) est un couple Gaussien.
2. Montrer que les v.a U et V sont indépendantes.

Solution.

1. — *Première méthode.* Comme X et Y sont deux variables aléatoires réelles indépendantes et identiquement distribuées suivant une loi normale, le vecteur (X, Y) est gaussien. Pour tout $(a, b) \in \mathbb{R}^2$, on a
   $$aU + bV = (a + b)X + (a - b)Y.$$
   Comme (X, Y) est un vecteur gaussien, $aU + bV = (a + b)X + (a - b)Y$ suit une loi normale. Par conséquent, (U, V) est aussi un vecteur gaussien.

   — *Deuxième méthode.* On peut écrire $\mathbf{L} = (X + Y, X - Y)$ sous la forme
   $$\mathbf{L}^t = \mathbf{A}\mathbf{M}^t,$$
   où $\mathbf{M} = (X, Y)$ et
   $$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$
   Comme $\mathbf{M} = (X, Y)$ est un vecteur gaussien, $\mathbf{L} = (X+Y, X-Y)$ est aussi un vecteur gaussien.

2. Comme (U,V) est un vecteur gaussien, U et V sont indépendantes si, et seulement si, $\mathbb{C}(U,V)=0$. Or, en utilisant le fait que X et Y suivent la loi normale centrée réduite, on a

$$\begin{aligned}\mathbb{C}(U,V) &= \mathbb{E}(UV)-\mathbb{E}(U)\mathbb{E}(V)=\mathbb{E}(X^2-Y^2)-0\times 0\\ &= \mathbb{E}(X^2)-\mathbb{E}(Y^2)=1-1=0.\end{aligned}$$

Donc U et V sont indépendantes.

Exercice 3. Soit $V=(X,Y)^t$ un vecteur gaussien tel que $\mathbb{E}\left(X^2\right)=4$ et $\mathbb{E}\left(Y^2\right)=1$, et les v.a $2X+Y$ et $X-3Y$ sont indépendantes.

1. Déterminer la matrice de covariance de V.
2. Montrer que le vecteur $W=(X+Y,2X-Y)^t$ est gaussien.
3. Déterminer sa matrice de covariance.

Solution.

1. Comme (X,Y) un vecteur gaussien centré tel que $\mathbb{E}\left(X^2\right)=4$ et $\mathbb{E}\left(Y^2\right)=1$, on a

$$\mathbb{V}(X)=\mathbb{E}\left(X^2\right)-\left(\mathbb{E}\left(X\right)\right)^2=\mathbb{E}\left(X^2\right)-0^2=4$$

et

$$\mathbb{V}(Y)=\mathbb{E}\left(Y^2\right)-\left(\mathbb{E}\left(Y\right)\right)^2=\mathbb{E}\left(Y^2\right)-0^2=1.$$

De plus, comme $2X+Y$ et $X-3Y$ sont indépendantes, on a

$$\mathbb{C}(2X+Y,X-3Y)=0.$$

Or,

$$\begin{aligned}\mathbb{C}(2X+Y,X-3Y) &= \mathbb{C}(2X,X)+\mathbb{C}(2X,-3Y)+\mathbb{C}(Y,X)+\mathbb{C}(Y,-3Y)\\ &= 2\mathbb{C}(X,X)+(-2\times 3)\mathbb{C}(X,Y)+\mathbb{C}(Y,X)-3\mathbb{C}(Y,Y)\\ &= 2\mathbb{V}(X)-6\mathbb{C}(X,Y)+\mathbb{C}(X,Y)-3\mathbb{V}(Y)\\ &= 2\times 4-5\mathbb{C}(X,Y)-3\times 1\\ &= 5(1-\mathbb{C}(X,Y)).\end{aligned}$$

Donc $\mathbb{C}(2X+Y,X-3Y)=0$ entraîne

$$\mathbb{C}(X,Y)=1.$$

La matrice de covariance de (X,Y) est

$$\mathbf{V}=\begin{pmatrix}\mathbb{V}(X) & \mathbb{C}(X,Y)\\ \mathbb{C}(Y,X) & \mathbb{V}(Y)\end{pmatrix}=\begin{pmatrix}4 & 1\\ 1 & 1\end{pmatrix}.$$

2. — *Première méthode.* Pour tout $(a,b)\in\mathbb{R}^2$, on a

$$a(X+Y)+b(2X-Y)=(a+2b)X+(a-b)Y.$$

Comme (X,Y) est un vecteur gaussien, $a(X+Y)+b(2X-Y)=(a+2b)X+(a-b)Y$ suit une loi normale. Par conséquent, $(X+Y,2X-Y)$ est un vecteur gaussien. En utilisant le résultat de la question 1-, il vient

$$\mathbb{V}(X+Y)=\mathbb{V}(X)+\mathbb{V}(Y)+2\mathbb{C}(X,Y)=4+1+2\times 1=7.$$

De même

$$\mathbb{V}(2X-Y)=4\mathbb{V}(X)+\mathbb{V}(Y)-4\mathbb{C}(X,Y)=16+1-4\times 1=13.$$

On a également

$$\begin{aligned}
\mathbb{C}(X+Y,2X-Y) &= \mathbb{E}\left((X+Y)(2X-Y)\right)-\mathbb{E}\left(X+Y\right)\mathbb{E}\left(2X-Y\right)\\
&= \mathbb{E}(2X^2-XY+2XY-Y^2)\\
&= 2\mathbb{E}\left(X^2\right)+\mathbb{E}\left(XY\right)-\mathbb{E}\left(Y^2\right)\\
&= 8+1-1=8.
\end{aligned}$$

En combinant ces résultats, on obtient la matrice de covariance de $(X+Y,2X-Y)$:

$$\mathbf{W}=\begin{pmatrix}7 & 8\\ 8 & 13\end{pmatrix}.$$

— *Deuxième méthode.* On peut écrire $\mathbf{L}=(X+Y,2X-Y)$ sous la forme

$$\mathbf{L}^t=\mathbf{A}\mathbf{M}^t,$$

où $\mathbf{M}=(X,Y)$ et

$$\mathbf{A}=\begin{pmatrix}1 & 1\\ 2 & -1\end{pmatrix}.$$

Comme $\mathbf{M}=(X,Y)$ est un vecteur gaussien, $\mathbf{L}=(X+Y,2X-Y)$ est aussi un vecteur gaussien. La matrice de covariance de $\mathbf{L}$ est donnée par la formule

$$\mathbf{W}=\mathbf{A}\mathbf{V}\mathbf{A}^t,$$

où $\mathbf{V}$ désigne la matrice de covariance du vecteur $\mathbf{M}=(X,Y)$ obtenue au résultat de la question 1-. Il vient

$$\mathbf{W}=\begin{pmatrix}1 & 1\\ 2 & -1\end{pmatrix}\begin{pmatrix}4 & 1\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 2\\ 1 & -1\end{pmatrix}=\begin{pmatrix}7 & 8\\ 8 & 13\end{pmatrix}.$$

Exercice 4. Soient X_1 et X_2 deux v.a i.i.d suivant chacune la loi normale $\mathcal{N}(0,1)$. Soit $\mathbf{Y}=(Y_1,Y_2)^t$ le vecteur aléatoire tel que

$$\mathbf{Y}=\mathbf{A}\mathbf{X}+\mathbf{B},$$

où $\mathbf{A}=\begin{pmatrix}1 & 1\\ -1 & 2\end{pmatrix}$, $\mathbf{X}=(X_1,X_2)^t$ et $\mathbf{B}=(2,3)^t$.

1. Déterminer la loi de $\mathbf{Y}$.
2. Déterminer la loi de $\mathbf{Z}=(Y_1+Y_2+1,3Y_1-Y_2)^t$.
3. Déterminer la loi de Y_1+Y_2+1, et la loi de $3Y_1-Y_2$.

Solution.

1. Comme $\mathbf{X}=(X_1,X_2)$ est un vecteur gaussien et $\mathbf{Y}$ est de la forme $\mathbf{Y}^t=\mathbf{A}\mathbf{X}^t+\mathbf{M}^t$, alors $\mathbf{Y}$ est aussi un vecteur gaussien. Comme $\mathbf{X}=(X_1,X_2)$ est centré, on a

$$\mathbb{E}\left(\mathbf{Y}^t\right)=\mathbb{E}\left(\mathbf{A}\mathbf{X}^t+\mathbf{M}^t\right)=\mathbf{A}\mathbb{E}\left(\mathbf{X}^t\right)+\mathbf{M}^t=0+\mathbf{M}^t=\mathbf{M}^t.$$

La matrice de covariance de $\mathbf{X}=(X_1,X_2)$ est $\mathbf{I}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$. Par conséquent, la matrice de covariance de $\mathbf{Y}$ est

$$\mathbf{V}=\mathbf{A}\mathbf{I}\mathbf{A}^t=\begin{pmatrix}1 & 1\\ -1 & 2\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & -1\\ 1 & 2\end{pmatrix}=\begin{pmatrix}2 & 1\\ 1 & 5\end{pmatrix}.$$

On en déduit que $\mathbf{Y}=(Y_1,Y_2)$ est un vecteur gaussien de moyenne $\mathbf{M}=(2,3)$ et de matrice de covariance $\mathbf{V}=\begin{pmatrix}2 & 1\\ 1 & 5\end{pmatrix}$.

2. En posant $\mathbf{Z} = (Y_1 + Y_2 + 1, 3Y_1 - Y_2)$, on peut écrire

$$\mathbf{Z} = \mathbf{B}\mathbf{Y}^t + \mathbf{N}^t,$$

où

$$\mathbf{B} = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix},$$

$\mathbf{Y} = (Y_1, Y_2)$ et $\mathbf{N} = (1, 0)$. En utilisant le résultat de la question 1-, il vient

$$\begin{aligned} \mathbb{E}\left(\mathbf{Z}^t\right) &= \mathbb{E}\left(\mathbf{B}\mathbf{Y}^t + \mathbf{N}^t\right) = \mathbf{B}\mathbb{E}\left(\mathbf{Y}^t\right) + \mathbf{N}^t = \mathbf{B}\mathbf{M}^t + \mathbf{N}^t \\ &= \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}. \end{aligned}$$

La matrice de covariance de $\mathbf{Y} = (Y_1, Y_2)$ est $\mathbf{V} = \begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix}$. Par conséquent, la matrice de covariance de $\mathbf{Z}$ est

$$\mathbf{W} = \mathbf{B}\mathbf{V}\mathbf{B}^t = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 9 & 3 \\ 3 & 17 \end{pmatrix}.$$

On en déduit que $\mathbf{Z} = (Y_1 + Y_2 + 1, 3Y_1 - Y_2)$ est un vecteur gaussien de moyenne $\mathbf{N} = (6, 3)$ et de matrice de covariance $\mathbf{W} = \begin{pmatrix} 9 & 3 \\ 3 & 17 \end{pmatrix}$.

3. Le résultat de la question 2- nous assure que $Y_1 + Y_2 + 1$ suit la loi normale de moyenne 6 et de variance 9, et $3Y_1 - Y_2$ suit la loi normale de moyenne 3 et de variance 17.

$\star$

Exercice 5. Une densité de probabilité (ou loi de probabilité dans le cas discret) d'une v.a réelle X, notée $f(x;\theta)$ appartient à la *famille exponentielle* si $f(x;\theta)$ peut s'écrire sous la forme

$$f(x;\theta) = \exp[\eta(\theta)T(x) + a(\theta) + b(x)] \cdot \tag{1}$$

où les fonctions η, T, a et b sont à valeurs dans $\mathbb{R}$. La statistique $T(x)$ est appelée statistique exhaustive naturelle et le paramètre $\eta = \eta(\theta)$ est appelé paramètre naturel. Soit X une variable aléatoire de densité $f(x;\theta)$ appartenant à la famille exponentielle.
Déterminer une statistique exhaustive pour le paramètre θ pour un n-échantillon i.i.d $(X_1, \ldots, X_n)$ de v.a de densité $f(x;\theta)$.

Solution. La densité jointe pour les observations *i.i.d.* $\mathbf{x} = (x_1, \ldots, x_n)$ est donnée par :

$$\begin{aligned} f(\mathbf{x};\theta) &= \prod_{i=1}^{n} f(x_i;\theta) \\ &= \prod_{i=1}^{n} \exp[\eta(\theta)T(x) + a(\theta) + b(x)] \\ &= \exp\left[\sum_{i=1}^{n} \left(\eta(\theta)T(x_i) + a(\theta) + b(x_i)\right)\right] \\ &= \underbrace{\exp\left[\sum_{i=1}^{n} \eta(\theta)T(x_i) + n\,a(\theta)\right]}_{u(T(x_1,\ldots,x_n);\theta)} \underbrace{\exp\left[\sum_{i=1}^{n} b(x_i)\right]}_{v(x_1,\ldots,x_n)} \end{aligned}$$

A partir de ce résultat, et d'après le théorème de factorisation (critère de Fisher-Neyman), on a : $\sum_{i=1}^{n} T(x_i)$ est une statistique exhaustive pour θ.

Exercice 6. Soit X une v.a suivant une loi Poisson de paramètre $\lambda > 0$, i.e, $\forall x \in \mathbb{N}$:

$$p_X(x;\lambda) = \mathbb{P}(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}.$$

Étant donné un n-échantillon i.i.d $(X_1, \ldots, X_n)$ généré suivant une loi de Poisson de paramètre λ, $p_X(x;\lambda)$

1. Déterminer une statistique exhaustive pour λ.
2. Calculer l'estimateur du maximum de vraisemblance (EMV) de λ. En déduire son expression en fonction de la statistique exhaustive calculée dans la question précédente.

Solution. La loi jointe pour l'échantillon i.i.d $(x_1, \ldots, x_n)$ est donnée par :

$$\begin{aligned}
p(x_1, \ldots, x_n; \lambda) &= \prod_{i=1}^{n} p(x_i;\lambda) \\
&= \prod_{i=1}^{n} [e^{-\lambda}\frac{\lambda^{x_i}}{x_i!}] \\
&= \lambda^{\sum_{i=1}^n x_i} e^{-n\lambda} \frac{1}{\prod_{i=1}^n x_i!} \\
&= \underbrace{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}}_{u(T(x_1,\ldots,x_n);\theta)} \underbrace{\frac{1}{\prod_{i=1}^n x_i!}}_{v(x_1,\ldots,x_n)}
\end{aligned}$$

A partir de ce résultat, et d'après le théorème de factorisation (critère de Fisher-Neyman), on a : $\sum_{i=1}^n X_i$ est une statistique exhaustive pour λ. La vraisemblance s'écrit :

$$\begin{aligned}
L(\lambda; x_1, \ldots, x_n) &= p(x_1, \ldots, x_n; \lambda) = \prod_{i=1}^{n} p(x_i;\lambda) && (2) \\
&= \prod_{i=1}^{n} e^{-\lambda}\frac{\lambda^{x_i}}{x_i!} && (3) \\
&= e^{-\lambda n} \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} && (4)
\end{aligned}$$

La log-vraisemblance est donc donnée par

$$\begin{aligned}
\ln L(\lambda; x_1, \ldots, x_n) &= \ln e^{-\lambda n} \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} && (5) \\
&= \ln e^{-\lambda n} + \ln \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} && (6) \\
&= -\lambda n + \sum_{i=1}^{n} \ln \frac{\lambda^{x_i}}{x_i!} && (7) \\
&= -\lambda n + \ln\lambda \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \ln(x_i!) && (8) \\
& && (9)
\end{aligned}$$

La dérivée première s'annule quand :

$$\begin{aligned}
\frac{\partial \ln L(\lambda; x_1, \ldots, x_n)}{\partial\lambda} &= 0 && (10) \\
-n + \frac{\sum_{i=1}^n x_i}{\lambda} &= 0 && (11)
\end{aligned}$$

donc

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{n}$$

La dérivée seconde s'écrit :

$$\frac{\partial^2 \ln L(\lambda; x_1, \ldots, x_n)}{\partial \lambda^2} = -\frac{\sum_{i=1}^{n} x_i}{\lambda^2} \leq 0$$

car $x_i \in \mathbb{N}$. L'estimateur est donné par :

$$\Lambda_n = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X}$$

Remarque : on peut remarquer que l'estimateur s'exprime à l'aide de la statistique exhaustive pour λ : $\Lambda = \frac{T}{n}$

Exercice 7. Estimateurs et détermination de la Borne Inférieure de Cramér-Rao (CRLB) : Soit X une v.a. gaussienne univariée de densité :

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\,\pi}}\, \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

1. Calculer la borne inférieure de Cramér-Rao pour un estimateur sans biais de l'espérance μ (variance σ^2 connue)
2. Calculer l'estimateur du maximum de vraisemblance de l'espérance μ
3. Montrer qu'il est sans biais
4. En déduire qu'il est efficace
5. Calculer la la borne inférieure de Cramér-Rao pour un estimateur sans biais de la variance σ^2 (espérance μ connu)
6. La variance empirique notée S^2 est l'estimateur du maximum de vraisemblance de la variance σ^2. Son espérance est $\frac{n-1}{n}\sigma^2$ et sa variance est $\frac{2\sigma^4}{n-1}$.
   En déduire son efficacité.
   Comment peut-on faire pour trouver un estimateur efficace de la variance ?

Solution.

1. On a

$$\begin{aligned} \ln f(x; \mu) &= \ln \frac{1}{\sigma\sqrt{2\,\pi}}\, \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} && (12) \\ &= \ln\left[\frac{1}{\sigma\sqrt{2\,\pi}}\right] - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2. && (13) \end{aligned}$$

et

$$\begin{aligned} \frac{\partial \ln f(x; \mu)}{\partial \mu} &= \frac{1}{\sigma^2}\left(\frac{x-\mu}{\sigma}\right) && (14) \\ \left(\frac{\partial \ln f(x; \mu)}{\partial \mu}\right)^2 &= \frac{1}{\sigma^4}\left(\frac{x-\mu}{\sigma}\right)^2 && (15) \end{aligned}$$

Ainsi

$$\mathbb{E}\left[\left(\frac{\partial \ln f(x; \mu)}{\partial \mu}\right)^2\right] = \frac{1}{\sigma^4}\mathbb{E}\left[(x-\mu)^2\right] = \frac{1}{\sigma^4}\sigma^2 = \frac{1}{\sigma^2} \qquad (16)$$

La CRLB pour l'estimateur de μ est donc

$$\mathrm{CRLB} = \frac{1}{n\mathbb{E}\left[\left(\frac{\partial \ln f(x;\mu)}{\partial \mu}\right)^2\right]} = \frac{\sigma^2}{n} \qquad (17)$$

2. on a $\mathbb{E}[\bar{X}] = \mathbb{E}[\frac{1}{n}\sum_{i=1}^n X_i] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i] = \mu$. L'estimateur $\bar{X}$ est donc sans biais.
3. Ensuite, on a $\mathrm{Var}(\bar{X}) = \mathrm{Var}(\frac{1}{n}\sum_{i=1}^n X_i) = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}(X_i) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$
   Puisque la variance de $\bar{X}$ est $\frac{\sigma^2}{n}$, il est donc à variance minimale pour μ quand la v.a. X est a densité normale $\mathcal{N}(\mu, \sigma^2)$
4. Soit $\theta = \sigma^2$, on a

$$\begin{aligned} \ln f(x;\theta) &= \ln \frac{1}{\sqrt{\theta 2\,\pi}}\, \mathrm{e}^{-\frac{1}{2}\frac{(x-\mu)^2}{\theta}} && (18) \\ &= -\frac{1}{2}\ln 2\pi\theta - \frac{1}{2}\frac{(x-\mu)^2}{\theta}. && (19) \end{aligned}$$

et

$$\begin{aligned} \frac{\partial \ln f(x;\theta)}{\partial\theta} &= -\frac{1}{2\theta} + \frac{(x-\mu)^2}{2\theta^2} && (20) \\ \frac{\partial^2 \ln f(x;\theta)}{\partial^2\theta} &= \frac{1}{2\theta^2} - \frac{(x-\mu)^2}{\theta^3} && (21) \\ & && (22) \end{aligned}$$

Ainsi

$$\mathbb{E}\left[\frac{\partial^2 \ln f(x;\theta)}{\partial^2\theta}\right] = \frac{1}{2\theta^2} - \frac{\mathbb{E}\left[(x-\mu)^2\right]}{\theta^3} = \frac{1}{2\theta^2} - \frac{\theta}{\theta^3} = -\frac{1}{2\theta^2} = -\frac{1}{2\sigma^4} \quad (23)$$

La CRLB pour l'estimateur de σ^2 est donc donnée par [1]

$$\mathrm{CRLB} = -\frac{1}{n\mathbb{E}\left[\frac{\partial^2 \ln f(x;\mu)}{\partial^2\theta}\right]} = \frac{2\sigma^4}{n} \quad (24)$$

5. On a $\mathrm{Var}(S^2) = \frac{2\sigma^4}{n}$ donc l'efficacité de S^2 est donnée par :

$$e(S^2) = \frac{\mathrm{CRLB}}{\mathrm{Var}(S^2)} = \frac{2\sigma^4/n}{2\sigma^4/(n-1)} = \frac{n-1}{n}$$

S^2 n'est donc pas un estimateur efficace pour σ^2 (car $e(S^2) < 1$). Cependant, S^2 est *asymptotiquement* efficace car

$$\lim_{n\to+\infty} e(S^2) = 1$$

The likelehood to be maximized is given as the joint probability density function for sample $(x_1, \ldots, x_n)$ of n independent identically distributed normal random variables

$$f(x_1, \ldots, x_n; \mu, \sigma^2) = \prod_{i=1}^n f(x_i; \mu, \sigma^2) = \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\,\pi}}\, \mathrm{e}^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}. \quad (25)$$

Maximizing this likelihood is equivalent to maximzing the following log-likehood function

$$\begin{aligned} \mathcal{L}(\mu,\sigma^2) &= \log f(x_1, \ldots, x_n; \mu, \sigma^2) = \log \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\,\pi}}\, \mathrm{e}^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2} \\ &= \sum_{i=1}^n \log \frac{1}{\sigma\sqrt{2\,\pi}} + \sum_{i=1}^n \log \mathrm{e}^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2} \\ &= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2. && (26) \end{aligned}$$

1. ici on a utilisé la deuxième forme de la CRLB

Lets first start by maximzing (26) with respect to μ. We have to set its partial derivative w.r.t μ equat to zero. By taking into account the fact that in (26), only the quantity $-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$ depends on μ, we can therefore write

$$\frac{\partial \mathcal{L}(\mu,\sigma^2)}{\partial \mu} = \frac{\partial - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}{\partial \mu} = 0, \tag{27}$$

that is

$$-\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2\,(x_i-\mu)\,\frac{\partial(x_i-\mu)}{\partial\mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0. \tag{28}$$

Finally we have $\sum_{i=1}^{n}(x_i-\mu) = \sum_{i=1}^{n} x_i - n\mu = 0$ which results in the ML estimate $\hat{\mu}$ of μ

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

that is the sample mean!.
We can see that the ML estimator of μ given by

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

is unbiased. Indeed, we have

$$\mathbb{E}[\widehat{\mu}] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}n\mu = \mu$$

where we have used the fact that $\mathbb{E}[X_i] = \mu$ which is the true mean.
To estimate σ^2, we similarly differentiate the log-likelihood (26) with respect to σ and equate to zero. Since in (26) only the quantity $-\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$ depends on σ, we have

$$\frac{\partial \mathcal{L}(\mu,\sigma^2)}{\partial \sigma} = \frac{\partial\left(-\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)}{\partial \sigma} = 0, \tag{29}$$

that is

$$-\frac{n}{2\sigma^2} - \frac{-1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \tag{30}$$

whiche gives

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 = n. \tag{31}$$

Finally we therefore have

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 \tag{32}$$

and by replacing μ by its ML estimate we get the ML estimate $\hat{\sigma}^2$ for σ^2

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2 \tag{33}$$

which is the empirical variance!
However, in contrast to the ML estimator for the mean which is unbiased, the one for the variance, as

we can see it is biased.
To calculate the expectation of the ML estimator $\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \widehat{\mu}\right)^2$, by using the fact that

$$\left(\sum_{i=1}^{n} X_i\right)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j$$

we first write

$$\begin{aligned}
\widehat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \widehat{\mu}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n} X_j\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 + \left(\frac{1}{n}\sum_{j=1}^{n} X_j\right)^2 - \frac{2}{n}X_i\sum_{j=1} X_j\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 + \frac{1}{n^2}\sum_{j=1}^{n}\sum_{h=1}^{n} X_j X_h - \frac{2}{n}\sum_{j=1} X_i X_j\right),
\end{aligned} \tag{34}$$

the expectation of $\widehat{\sigma}^2$ is then given by

$$\mathbb{E}[\widehat{\sigma}^2] = \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{E}[X_i^2] + \frac{1}{n^2}\sum_{j=1}^{n}\sum_{h=1}^{n}\mathbb{E}[X_j X_h] - \frac{2}{n}\sum_{j=1}\mathbb{E}[X_i X_j]\right).$$

Since the variables are mutually independent, we have, for $i \neq j$, $\mathbb{E}[X_i X_j] = \mathbb{E}[X_i][X_j] = \mu^2$. When $i = j$, we have $\mathbb{E}[X_i X_i] = \mathbb{E}[X_i^2] = \sigma^2 + \mu^2$ from the variance formula. The previous equation is therefore wreritten as

$$\begin{aligned}
\mathbb{E}[\widehat{\sigma}^2] &= \frac{1}{n}\sum_{i=1}^{n}\left(\sigma^2 + \mu^2 + \frac{1}{n^2}\left(n\left((n-1)\,\mu^2 + \left(\sigma^2 + \mu^2\right)\right)\right) - \frac{2}{n}\left((n-1)\,\mu^2 + \left(\sigma^2 + \mu^2\right)\right)\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\sigma^2 + \mu^2 + \frac{1}{n^2}\left(n^2\mu^2 + n\sigma^2\right) - \frac{2}{n}\left(n\mu^2 + \sigma^2\right)\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\sigma^2 + \mu^2 + \mu^2 + \frac{1}{n}\sigma^2 - 2\mu^2 - \frac{2}{n}\sigma^2\right) \\
&= \frac{1}{n}(n\sigma^2 - \sigma^2) \\
&= \frac{n-1}{n}\sigma^2.
\end{aligned} \tag{35}$$

The ML estimator $\mathbb{E}[\widehat{\sigma}^2]$ for σ^2 is therefore unbiased since $\mathbb{E}[\widehat{\sigma}^2] \neq \sigma^2$. However it is asymptotically unbiased : $\lim_{n\to\infty}\mathbb{E}[\widehat{\sigma}^2] = \sigma^2$.
We can correct the bias by taking $\sigma^{\star 2} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \hat{\mu}\right)^2$ as an unbiased estimator of the variance, rather than the ML estimator $\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \hat{\mu}\right)^2$. It is indeed easy to check that $\mathbb{E}[\sigma^{\star 2}] = \sigma^2$.