Projet de statistiques

Maxime Peterlin - Gabriel Vermeulen

ENSEIRB-MATMECA, Bordeaux

28, avril 2014

Table des matières

Chapitre 1

Autour des variables aléatoires gaussiennes

1.1 Variables aléatoires gaussiennes réelles

1.1.1 Densité de probabilité d'une variable aléatoire gaussienne de de moyenne μ et de variance σ^2

La densité de probabilité d'une loi normale de moyenne 0 et de variance 1 est :

$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot exp(-\frac{x^2}{2})$$

Pour obtenir la densité de probabilité d'une variable aléatoire gaussienne de de moyenne μ et de variance σ^2, on effectue le changement de variable $y = \frac{x-\mu}{\sigma}$:

$$\int f(x)dx = \int \frac{1}{\sqrt{2\pi}} \cdot exp(-\frac{x^2}{2})dx = \int \frac{1}{\sqrt{2\pi}} \cdot exp(-\frac{(y-\mu)^2}{2\cdot\sigma^2})\frac{dy}{\sigma} = \int f(y)dy$$

Ainsi, on a la densité de probabilité $f(y)$ recherchée :

$$f(y) = \frac{1}{\sqrt{2\pi\sigma}} \cdot exp(-\frac{(y-\mu)^2}{2\cdot\sigma^2})$$

1.1.2 Histogramme de 100 réalisations de la loi $\mathcal{N}_{\mathbb{R}}(5,1)$

1.1.3 Histogramme de 1000 réalisations de la loi $\mathcal{N}_{\mathbb{R}}(5,1)$

1.2 Variables aléatoires gaussiennes vectorielles

1.2.1 Fonction caractéristique de la variable $t^T X$

On a

$$\Phi_{t^T X}(u) = \mathbb{E}[e^{iut^T X}] = \mathbb{E}[e^{iu \sum\limits_{k=1}^{n} t_k X_k}]$$

1.2.2 Montrons que les $X_1, \dots, X_n$ sont indépentantes

En se servant du lemme donné, si

$$\mathbb{E}[e^{i \sum\limits_{k=1}^{n} t_k X_k}] = \prod_{k=1}^{n} \mathbb{E}[e^{it_k X_k}]$$

alors les variables aléatoires $X_1, \dots, X_n$ seront indépendantes. Soit $X \sim \mathcal{N}_{\mathbb{R}^{\mathbb{N}}}(\mu, \sigma^2)$ et $Y \sim \mathcal{N}_{\mathbb{R}^{\mathbb{N}}}(O, I)$, on a la relation suivante entre X et Y :

$$X = m + \sigma Y$$

On en déduit que

$$\Phi_X(u) = \mathbb{E}[e^{iuX}] = \mathbb{E}[e^{iu(m+\sigma Y)}] = e^{ium}\mathbb{E}[e^{iu\sigma Y}] = e^{ium - \frac{\sigma^2 u^2}{2}}$$

Comme $X \sim \mathcal{N}_{\mathbb{R}^{\mathbb{N}}}(O, I)$, alors $t^T X \sim \mathcal{N}_{\mathbb{R}^{\mathbb{N}}}(O, \sum\limits_{k=1}^{n} t_k^2)$.

Ainsi $\Phi_{t^T X}(u) = e^{-\frac{u^2}{2} \sum\limits_{k=1}^{n} t_k^2}$ et $\Phi_{t^T X}(1) = e^{-\frac{1}{2} \sum\limits_{k=1}^{n} t_k^2}$.

De plus, $\Phi_{X_k}(t_k) = e^{-\frac{t_k^2}{2}} = \mathbb{E}[e^{it_k X_k}]$, donc $\prod\limits_{k=1}^{n} \Phi_{X_k}(t_k) = \Phi_{t^T X}(1) = \mathbb{E}[e^{i \sum\limits_{k=1}^{n} t_k X_k}]$.

On a bien montré que

$$\mathbb{E}[e^{i \sum\limits_{k=1}^{n} t_k X_k}] = \prod_{k=1}^{n} \mathbb{E}[e^{it_k X_k}]$$

Donc les variables aléatoires $X_1, \dots, X_n$ sont indépendantes.

1.2.3 Densité de probabilité de la loi $\mathcal{N}_{\mathbb{R}^n}(0, I)$

On a $\forall k$, $X_k \sim \mathcal{N}_{\mathbb{R}}(0,1)$ indépendantes, ainsi

$$f(X) = f(X_1, \dots X_n) = \prod_{k=1}^{n} f_k(X_k) = (\frac{1}{\sqrt{2\pi}})^n e^{-\frac{1}{2} \sum\limits_{k=1}^{n} x_k}$$

1.2.4 Densité de probabilité de la loi $\mathcal{N}_{\mathbb{R}^n}(\mu, \Gamma)$

Pour toutes fonctions ϕ bornée, on a

$$\mathbb{E}[h(Y)] =$$

1.2.5 Montrons que les composantes d'un vecteur gaussien sont indépendantes si et seulement si sa matrice de covariance est diagonale

Si les composantes d'un vecteur gaussien sont indépendantes, alors $\forall i \neq j$ on a $COV(X_i, X_j) = 0$, ainsi, la matrice de covariance a tous ses éléments non diagonaux qui sont nuls. Cette dernière est donc diagonale.
Réciproquement, si la matrice de covariance est diagonale, alors :

$$f_k(X_k) = \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{1}{2}\sum\limits_{k=1}^{n} \frac{(x_1-\mu_1)^2}{\sigma_k^2}}$$

et

$$f(X) = \frac{1}{\sqrt{2\pi}^n \prod\limits_{k=1}^{n} \sigma_k} e^{-\frac{1}{2}(x-\mu)^T \begin{bmatrix} \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n^2 \end{bmatrix}^{-1} (x-\mu)} = \frac{1}{\sqrt{2\pi}^n \prod\limits_{k=1}^{n} \sigma_k} e^{-\frac{1}{2}\sum\limits_{k=1}^{n} \frac{(x_k-\mu_k)}{\sigma_k^2}}$$

On a bien $\prod f_k(X_k) = f(X)$, les variables sont donc indépendantes.

1.2.6 Tracé de la densité de la loi $\mathcal{N}_{\mathbb{R}^2}(\mu, \Gamma)$

1.2.7 Histogramme de 1000 réalisations de la loi $\mathcal{N}_{\mathbb{R}^2}(\mu, \Gamma)$

1.2.8 Montrons que $X \sim \mathcal{N}_{\mathbb{R}^d}(0, I) \Rightarrow UX \sim \mathcal{N}_{\mathbb{R}^d}(0, I), \forall U \in \mathbb{O}_d(\mathbb{R})$

On a $UX = [a_1, \ldots, a_d]$, avec $\forall k, a_k = \sum\limits_{i=1}^{d} u_{ki} X_i$.
Calculons l'espérance de ce vecteur :

$$\mathbb{E}[UX] = \mathbb{E}\left[[\sum_{i=1}^{d} u_{1i} X_i, \ldots, \sum_{i=1}^{d} u_{di} X_i] \right] = \left[\sum_{i=1}^{d} u_{1i} \mathbb{E}[X_i], \ldots, \sum_{i=1}^{d} u_{di} \mathbb{E}[X_i] \right] = 0$$

car $\forall i, \mathbb{E}[X_i] = 0$.

Calculons à présent sa variance

$$Var[UX] = \mathbb{E}[(UX)(UX)^T] = \mathbb{E}[UXX^TU^T] = U\mathbb{E}[XX^T]U^T$$

Nous savons également que

$$Var[X] = \mathbb{E}[XX^T] - \mathbb{E}^2[X] = \mathbb{E}[XX^T] = I$$

Ainsi, $Var[X] = UU^T = I$, car $U \in \mathbb{O}_d(\mathbb{R})$.
Nous avons donc montré que $X \sim \mathcal{N}_{\mathbb{R}^d}(0, I) \Rightarrow UX \sim \mathcal{N}_{\mathbb{R}^d}(0, I), \forall U \in \mathbb{O}_d(\mathbb{R})$.

1.2.9 Histogramme de 1000 réalisations du vecteur UX

1.2.10 Démonstration et vérification expérimentale du caractère gaussien d'un vecteur donné

On a $X \sim \mathcal{N}_{\mathbb{R}}(0, 1)$.
On pose $\mathbf{X} = [X_1, X_2]^T$, avec $X_1 = X et X_2 = ZX$. Z est une variable aléatoire indépendante de X de loi uniforme sur l'intervalle $\{-1, 1\}$

On veut montrer que X_1 et X_2 suivent des lois gaussiennes.
Il est évident que c'est le cas pour X_1, montrons le pour X_2. La densité de probabilité que suit Z est

$$f_Z(x) = \frac{1}{2}\delta(x+1) + \frac{1}{2}\delta(x-1)$$

La densité de probabilité que suit X est

$$f_X(x) = \frac{1}{\sqrt{2\pi}} exp(-\frac{x^2}{2})$$

La densité de probabilité que suit ZX est la convolution de leur densité respective.

$$f_{ZX}(x) = (f_X * f_Z)(x) = [f * (\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)](x) = \frac{1}{2}f(x+1) + \frac{1}{2}f(x-1)$$

La fonction caractéristique d'une variable $X \sim \mathcal{N}_{\mathbb{R}}(0, 1)$ est

$$\Phi_X(u) = \mathbb{E}[e^{iuX}] = e^{-\frac{u^2}{2}}$$

Ainsi

$$\begin{aligned}
\Phi_{ZX}(u) = \mathbb{E}[e^{iuZX}] = \int_{\mathbb{R}} f_{ZX}(y)e^{ity}dy &= \int_{\mathbb{R}} (\frac{1}{2}f_X(x+1) + \frac{1}{2}f_X(x-1))e^{ity}dy \\
&= \int_{\mathbb{R}} \frac{1}{2}f(x+1)e^{ity}dy + \int_{\mathbb{R}} \frac{1}{2}f(x-1)e^{ity}dy \\
&= \frac{1}{2}e^{-\frac{u^2}{2}} + \frac{1}{2}e^{-\frac{u^2}{2}} = e^{-\frac{u^2}{2}}
\end{aligned}$$

car f_X suit une loi $\mathcal{N}_{\mathbb{R}}(0, 1)$.
Nous avons donc prouvé que $X_2 \sim \mathcal{N}_{\mathbb{R}}(0, 1)$.

A présent, nous allons voir si le vecteur $\mathbf{X}$ est gaussien ou non.
Si le vecteur $\mathbf{X}$ était gaussien, il aurait la fonction caractéristique suivante :

$$\Phi_{\mathbf{X}}(u) = e^{-\frac{u^2}{2}}$$

Calculons sa fonction caractéristique.

$$\Phi_{\mathbf{X}}(u) = \mathbb{E}(e^{i<u,X>}) = \int_{\mathbb{R}} \int_{\mathbb{R}} f_{\mathbf{X}}(x,z) e^{iu_1 x + iu_2 xz} dxdz$$

X et ZX n'étant pas indépendant, on ne peut retrouver la bonne fonction caractéristique.
Le vecteur $\mathbf{X}$ n'est donc pas gaussien.

1.3 Variables aléatoires gaussiennes complexes

1.3.1 Variance et densité de probabilité de $Z \sim \mathcal{N}_{\mathbb{C}}(\mu, \sigma^2)$

Calculons la variance de la variable aléatoire $Z = X + iY$.

$$\begin{aligned} V[Z] = \mathbb{E}\left[|Z - \mathbb{E}[Z]|^2\right] &= \mathbb{E}\left[|Z|^2 + |\mathbb{E}[Z]|^2 - 2\Re(Z\overline{\mathbb{E}[Z]})\right] \\ &= \mathbb{E}[|Z|^2] - |\mathbb{E}[Z]|^2 = \mathbb{E}[|X+iY|^2] - |\mathbb{E}[X+iY]|^2 \\ &= \mathbb{E}[X^2 + Y^2] - (\mathbb{E}^2[X] + \mathbb{E}^2[Y]) = Var[X] + Var[Y] = \sigma^2 \end{aligned}$$

Calculons à présent la densité de probabilité de Z.
On sait que X et Y sont indépendantes, ainsi

$$\begin{aligned} f_{XY}(X,Y) &= f_X(X) f_Y(Y) \\ &= \frac{1}{\sqrt{2\pi}\frac{\sigma}{\sqrt{2}}} exp(-\frac{1}{2}\frac{(x-\Re(\mu))^2}{\frac{\sigma^2}{2}}) \cdot \frac{1}{\sqrt{2\pi}\frac{\sigma}{\sqrt{2}}} exp(-\frac{1}{2}\frac{(y-\Im(\mu))^2}{\frac{\sigma^2}{2}}) \\ &= \frac{1}{\sigma^2 \pi} e^{-\frac{|z-\mu|^2}{\sigma^2}} \end{aligned}$$

1.3.2 Histogramme de 1000 réalisations de la loi $\mathcal{N}_{\mathbb{C}}(0,1)$

1.4 Théorème de la limite centrale

1.4.1 Loi de W_n lorsque $X_k \sim \mathcal{N}_{\mathbb{R}}(\mu, \sigma^2)$

On sait que $\forall k, X_k \sim \mathcal{N}_{\mathbb{R}}(0,1)$, de plus, on a

$$W_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \frac{X_k - \mu}{\sigma}$$

Nous cherchons la loi de W_n en sachant que $\forall k, X_k$ suit une loi normale.
Posons
$$Y_k = \frac{X_k - \mu}{\sqrt{n}\sigma}$$
Commençons par calculer l'espérance de Y_k
$$\mathbb{E}[Y_k] = \mathbb{E}\left[\frac{X_k - \mu}{\sqrt{n}\sigma}\right] = \frac{\mathbb{E}[X_k] - \mu}{\sqrt{n}\sigma} = 0$$
Calculons à présent la variance de Y_k
$$Var[Y_k] = Var\left[\frac{X_k - \mu}{\sqrt{n}\sigma}\right] = \frac{1}{n\sigma^2}Var[X_k - \mu] = \frac{1}{n\sigma^2}\sigma^2 = \frac{1}{n}$$
Nous avons donc $Y_k \sim \mathcal{N}_{\mathbb{R}}(0, \frac{1}{n})$ Ainsi
$$W_n = \sum_{k=1}^{n} Y_k \sim \mathcal{N}_{\mathbb{R}}(0, \sqrt{\sum_{k=1}^{n}\frac{1}{n}} = 1)$$

1.4.2 Histogramme de 1000 réalisations de la loi W_n lorsque $X_k \sim \mathcal{N}_{\mathbb{R}}(\mu, \sigma^2)$

1.4.3 Histogramme de 1000 réalisations de la loi W_n lorsque $X_k \sim \mathcal{U}([0, 1])$

Calcul de l'espérance et de la variance de $X_k, \forall k$
$$\mathbb{E}[X_k] = \int_0^1 x dx = \frac{1}{2}$$
$$Var[X_k] = \mathbb{E}[X_k^2] - \mathbb{E}^2[X_k] = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

1.4.4 Histogramme de 1000 réalisations de la loi W_n lorsque $X_k \sim \mathcal{B}(\frac{1}{2})$

Calcul de l'espérance et de la variance de $X_k, \forall k$
$$\mathbb{E}[X_k] = p = \frac{1}{2}$$
$$Var[X_k] = p(1 - p) = \frac{1}{2}$$

1.5 Loi du χ^2

1.5.1 Densité de probabilité d'une variable suivant une loi $\chi^2(n)$

Montrons par récurrence que la densité de probabilité de $Y \sim \chi^2(n)$ est

$$f_n(x) = \frac{x^{\frac{n}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}\mathbb{1}_{\mathbb{R}^+}(x)$$

Tout d'abord, montrons le pour $n = 1$. On a $Y = X^2$.
Pour toutes fonctions ϕ bornée, on a :

$$\begin{aligned}
\mathbb{E}[\phi(y)] = \mathbb{E}[\phi(x^2)] &= \int_{\mathbb{R}} \phi(x^2) f_X(x)dx \\
&= \int_{-\infty}^{0} \phi(x^2) f_X(x)dx + \int_{0}^{\infty} \phi(x^2) f_X(x)dx \\
&= \int_{0}^{\infty} \phi(x^2) f_X(-x)dx + \int_{0}^{\infty} \phi(x^2) f_X(x)dx
\end{aligned}$$

Posons le changement de variable $x = \sqrt{u}$, $dx = \frac{du}{2\sqrt{u}}$.

$$\begin{aligned}
\mathbb{E}[\phi(y)] &= \int_{0}^{\infty} \phi(u) f_X(-\sqrt{u})\frac{du}{2\sqrt{u}} + \int_{0}^{\infty} \phi(u) f_X(\sqrt{u})\frac{du}{2\sqrt{u}} \\
&= \int_{0}^{\infty} \phi(u)[f_X(-\sqrt{u}) + f_X(\sqrt{u})]\frac{du}{2\sqrt{u}}
\end{aligned}$$

Ainsi, pour $Y = X^2$, la densité de probabilité s'exprime de la façon suivante :

$$f_Y(y) = \frac{1}{2\sqrt{y}}[f_X(-\sqrt{y}) + f_X(\sqrt{y})]\mathbb{1}_{\mathbb{R}^+}(x)$$

On en déduit donc que

$$f_1(x) = \frac{x^{-\frac{1}{2}}}{2\sqrt{2\pi}}[e^{-\frac{x}{2}} + e^{-\frac{x}{2}}]\mathbb{1}_{\mathbb{R}^+}(x) = \frac{x^{-\frac{1}{2}}e^{-\frac{x}{2}}}{\sqrt{2\pi}}\mathbb{1}_{\mathbb{R}^+}(x)$$

Supposons que la propriété soit vraie au rang n et montrons le au rang $n + 1$:

$$\begin{aligned}
f_{n+1}(x) = f_n(x) * f_1(x) &= \int_{\mathbb{R}} f_n(t) f_1(x-t)dt \\
&= \frac{1}{2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{1}{2})} e^{-\frac{x}{2}}\mathbb{1}_{\mathbb{R}^+}(x) \int_{0}^{x} t^{\frac{n}{2}-1}(x-t)^{-\frac{1}{2}}dt
\end{aligned}$$

Posons le changement de variable $t = xu$, $dt = xdu$.

$$\begin{aligned}
f_{n+1}(x) &= \frac{1}{2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{1}{2})} e^{-\frac{x}{2}} \mathbb{1}_{\mathbb{R}^+}(x) \int_0^1 (xu)^{\frac{n}{2}-1}(x-xu)^{-\frac{1}{2}} du \\
&= \frac{1}{2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{1}{2})} e^{-\frac{x}{2}} \mathbb{1}_{\mathbb{R}^+}(x) \int_0^1 (xu)^{\frac{n}{2}-1}(x-xu)^{-\frac{1}{2}} xdu \\
&= \frac{1}{2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{1}{2})} e^{-\frac{x}{2}} x^{\frac{n+1}{2}-1} \beta(\frac{n}{2}, \frac{1}{2}) \mathbb{1}_{\mathbb{R}^+}(x) \\
&= \frac{x^{\frac{n+1}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{n+1}{2}}\Gamma(\frac{n+1}{2})} \mathbb{1}_{\mathbb{R}^+}(x)
\end{aligned}$$

1.5.2 Moyenne et variance de la loi $\chi^2(n)$

Calculons l'espérance de Y

$$\mathbb{E}[Y] = \mathbb{E}[\sum_{k=1}^{n} X_k^2] = \sum_{k=1}^{n} \mathbb{E}[X_k^2] = \sum_{k=1}^{n} Var[X_k] + \mathbb{E}^2[X_k] = n$$

Calculons à présent la variance de X^2

$$Var[X^2] = \mathbb{E}[X_k^4] - \mathbb{E}^2[X_k^2]$$

Déterminons l'expression de $\mathbb{E}[X_k^4]$

$$\begin{aligned}
\mathbb{E}[X_k^4] &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^4 e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^3 \cdot xe^{-\frac{x^2}{2}} dx \\
&= \frac{1}{\sqrt{2\pi}} [-x^3 e^{-\frac{x^2}{2}}]_{\mathbb{R}} + \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} 3x^2 e^{-\frac{x^2}{2}} dx \\
&= \frac{3}{\sqrt{2\pi}} \int_{\mathbb{R}} x \cdot xe^{-\frac{x^2}{2}} dx \\
&= \frac{3}{\sqrt{2\pi}} [-xe^{-\frac{x^2}{2}}]_{\mathbb{R}} + \frac{3}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{x^2}{2}} dx \\
&= 3 \cdot \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 3
\end{aligned}$$

Ainsi

$$\begin{aligned}
\mathbb{E}[X_k^4] &= 3 \\
Var[X_k^2] &= \mathbb{E}[X_k^4] - \mathbb{E}^2[X_k^2] = 3 - 1 = 2
\end{aligned}$$

On a donc $Var[Y^2] = \sum_{k=1}^{n} Var[X_k^2] = 2n$, car les variables sont indépendantes.

1.5.3 Tracé de la densité de probabilité de la loi $\chi^2(n)$

1.5.4 Loi du vecteur $U^T X$

On a

$$U^T X = \sum_{i=1}^{k} < u_i, X > u_i = \begin{bmatrix} u_1^T X \\ \vdots \\ u_k^T X \end{bmatrix}$$

Ainsi,

$$\forall i, u_i^T X \sim \mathcal{N}_{\mathbb{R}}(u_i^T \cdot \mathbf{0}, u_i^T \sigma^2 \mathbb{I} u_i) \Leftrightarrow \forall i, u_i^T X \sim \mathcal{N}_{\mathbb{R}}(0, \sigma^2)$$

Chapitre 2

Estimation paramétrique

2.1 Estimation des paramètres d'une loi gaussienne

2.1.1 Démonstration de l'expression des composantes de $\hat{\theta}_n$

On cherche $\hat{\theta}_n$ tel que

$$\hat{\theta}_n = \arg\max_{\theta \in \Theta} f_\theta(X_1, \ldots, X_n) = \{\theta / \forall \theta', f_{\theta'}(\mathbf{X}) \leq f_\theta(\mathbf{X})\}$$

On a

$$\begin{aligned} f_\theta(X_1, \ldots, X_n) &= \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} e^{-\frac{1}{2} \frac{\sum\limits_{k=1}^{n} (X_k - \mu)^2}{\sigma^2}} \end{aligned}$$

$$\frac{\partial f_\theta}{\partial \mu}|_{\theta = \hat{\theta}_n} = 0 = \sum_{k=1}^{n} (X_k - \hat{\mu}_n) \Leftrightarrow \hat{\mu}_n = \frac{1}{n} \sum_{k=1}^{n} X_k$$

Dans l'expression précédente de f_θ on pose $y = \hat{\sigma}_n^2$, ainsi

$$\frac{\partial f_\theta}{\partial y}|_{\theta=\hat{\theta}_n} = 0$$

$$= \frac{-\frac{n}{2}}{(2\pi)^{\frac{n}{2}} y^{\frac{n}{2}+1}} e^{-\frac{1}{2}\frac{\sum\limits_{k=1}^{n}(X_k-\hat{\mu}_n)^2}{\sigma^2}} + \frac{1}{(2\pi)^{\frac{n}{2}} y^{\frac{n}{2}}} \left(\frac{1}{2} \frac{\sum\limits_{k=1}^{n}(X_k-\hat{\mu}_n)^2}{y^2} \right) e^{-\frac{1}{2}\frac{\sum\limits_{k=1}^{n}(X_k-\hat{\mu}_n)^2}{\sigma^2}}$$

$$\Leftrightarrow \frac{n}{2y} = \frac{1}{2} \frac{\sum\limits_{k=1}^{n}(X_k-\hat{\mu}_n)^2}{y^2}$$

$$\Leftrightarrow \hat{\sigma}_n^2 = \frac{1}{n}\sum_{k=1}^{n}(X_k-\hat{\mu}_n)^2$$

2.1.2 Matrice d'information de Fisher

Commençons par calculer $log f_\theta$

$$log f_\theta(X_1,\ldots,X_n) = log\left[\frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)}\right]$$

$$= -\frac{1}{2} log\left[(2\pi)^n\sigma^{2n}\right] - \frac{1}{2}\frac{\sum\limits_{k=1}^{n}(X_k-\mu)^2}{\sigma^2}$$

$$= -\frac{n}{2} log(2\pi) - n log(\sigma) - \frac{1}{2}\frac{\sum\limits_{k=1}^{n}(X_k-\mu)^2}{\sigma^2}$$

$$\frac{\partial log f_\theta}{\partial\mu} = \frac{\sum\limits_{k=1}^{n} X_k-\mu}{\sigma^2}$$

$$\frac{\partial^2 log f_\theta}{\partial\mu^2} = -\frac{n}{\sigma^2}$$

$$\frac{\partial^2 log f_\theta}{\partial\mu\partial\sigma^2} = -\frac{\sum\limits_{k=1}^{n} X_k-\mu}{\sigma^4}$$

Posons $y = \sigma^2$

$$logf_\theta(X_1, \ldots, X_n) = -\frac{n}{2}log(2\pi) - nlog(\sqrt{y}) - \frac{1}{2}\frac{\sum\limits_{k=1}^{n}(X_k-\mu)^2}{y}$$

$$\frac{\partial logf_\theta}{\partial y} = -\frac{n}{2y} + \frac{1}{2}\frac{\sum\limits_{k=1}^{n}(X_k-\mu)^2}{y^2}$$

$$\frac{\partial^2 logf_\theta}{\partial y^2} = -\frac{n}{2^2 y} - \frac{2\sum\limits_{k=1}^{n}(X_k-\mu)^2}{2y^3}$$

On obtient alors la matrice de Fisher suivante :

$$\begin{bmatrix} -\mathbb{E}[\frac{\partial^2 logf_\theta}{\partial\mu^2}] & -\mathbb{E}[\frac{\partial^2 logf_\theta}{\partial\mu\partial\sigma^2}] \\ -\mathbb{E}[\frac{\partial^2 logf_\theta}{\partial\mu\partial\sigma^2}] & -\mathbb{E}[\frac{\partial^2 logf_\theta}{\partial\sigma^4}] \end{bmatrix} = \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{n}{2\sigma^4}] \end{bmatrix}$$

2.1.3 Borne de Cramer-Rao

On déduit directement de la matrice d'information de Fisher la borne de Cramer-Rao

$$I_\theta^{-1} = \begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2\sigma^4}{n}] \end{bmatrix}$$

2.1.4 Loi de l'estimateur $\hat{\mu}_n$

On a

$$\mathbb{E}[\hat{\mu}_n] = \frac{1}{n}\sum_{k=1}^{n}\mu = \mu$$

$$Var[\hat{\mu}_n] = \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n}$$

Ainsi $\hat{\mu}_n \sim \mathcal{N}_\mathbb{R}(\mu, \frac{\sigma^2}{n})$

2.1.5 Biais de l'estimateur $\hat{\mu}_n$

On a

$$\|\mu - \mathbb{E}[\hat{\mu}_n]\|^2 = \|\mu - \mu\|^2 = 0$$

2.1.6 Risque quadratique de l'estimateur $\hat{\mu}_n$

On a

$$\mathbb{E}[\|\mu - \hat{\mu}_n\|^2] = \int_\mathbb{R} f(t)(\mu - t)dt = \int_\mathbb{R}(\mu - t)^2\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\frac{(t-\mu)^2 n}{\sigma^2}}dt$$

On effectue le changement de variable $u = \frac{t-\mu}{\sqrt{2}\frac{\sigma}{\sqrt{n}}}$, $du = \frac{\sqrt{n}}{\sqrt{2}\sigma}dt$

$$\mathbb{E}[\|\mu - \hat{\mu}_n\|^2] = \frac{2\sigma^2}{n\sqrt{\pi}} \int_{\mathbb{R}} u^2 e^{-u^2} du$$

On effectue à présent le changement de variable $u = \sqrt{y}$, $du = \frac{dy}{2\sqrt{y}}$

$$\mathbb{E}[\|\mu - \hat{\mu}_n\|^2] = \frac{\sigma^2}{n\sqrt{\pi}} \int_{\mathbb{R}} y^{-\frac{1}{2}} e^{-y} dy = \frac{\sigma^2}{n\sqrt{\pi}}\Gamma(\frac{1}{2}) = \frac{\sigma^2}{n}$$

2.1.7 Tracé de l'évolution du risque quadratique de l'estimateur $\hat{\mu}_n$

2.1.8 Loi de l'estimateur $\hat{\sigma}_n^2$

On a

$$\hat{\sigma}_n^2 = \frac{1}{n}\|(\mathbb{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T)X\| = \frac{1}{n}\sum_{k=1}^{n}(X_k - \frac{1}{n}\sum_{l=1}^{n} X_l)^2 = \|MX\|^2$$

Avec

$$M = \frac{1}{\sqrt{n}} \begin{bmatrix} 1-\frac{1}{n} & 1 & \cdots & 1 \\ 1 & \ddots & & \vdots \\ \vdots & & \ddots & 1 \\ 1 & \cdots & 1 & 1-\frac{1}{n} \end{bmatrix}$$

On a $MX = [M_1, \ldots, M_n]^T$ et

$$\forall k, M_k = \frac{1}{\sqrt{n}}[(1-\frac{1}{n})X_k - \frac{1}{n}\sum_{\substack{l=1 \\ l\neq k}}^{n} X_l]$$

On en déduit l'espérance et la variance de $\hat{\sigma}_n^2$

$$\forall k, \mathbb{E}[M_k] = \frac{1}{\sqrt{n}}[(1-\frac{1}{n})\mu - \frac{1}{n}\sum_{\substack{l=1 \\ l\neq k}}^{n} \mu] = \frac{1}{\sqrt{n}}[\mu - \frac{1}{n}\mu - \mu + \frac{1}{n}\mu] = 0$$

$$\forall k, VarE[M_k] = \frac{1}{\sqrt{n}}[(1-\frac{1}{n})\sigma^2 - \frac{1}{n}\sum_{\substack{l=1 \\ l\neq k}}^{n} \sigma^2] = \frac{\sigma^2}{n}(1 - \frac{3}{n} + \frac{2}{n^2})$$

Comme $\sigma^2 = \|M\|^2 = \sum_{k=1}^{n} M_k^2$, alors

$\sigma^2 \sim \chi^2(n)$, avec $\forall k, X_k \sim \mathcal{N}_R(0, \frac{\sigma^2}{n}(1 - \frac{3}{n} + \frac{2}{n^2}))$

2.1.9 Biais de l'estimateur $\hat{\sigma}_n^2$

2.1.10 Risque quadratique de l'estimateur $\hat{\sigma}_n^2$

2.1.11 Tracé de l'évolution du risque quadratique de l'estimateur $\hat{\sigma}_n^2$

Chapitre 3

Introduction à la statistique descriptive

3.1 Principe de l'ACP

3.1.1 Montrons que le vecteur maximisant l'énergie de projection est donné par u_1

$\forall v \in \mathbb{C}^d$, tel que $\|v\| = 1$ on a

$$\mathbb{E}[(v^T X_k)^2] = \mathbb{E}[< v^T X_k, v^T X_k >] = \mathbb{E}[v^T X_k X_k^T v] = v^T \mathbb{E}[X_k X_k^T] v \quad = v^t R v$$

R étant diagonalisable, on peut l'écrire sous la forme $R = PDP^{-1}$.
$P = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$, de plus, comme P est une matrice orthonormée, on a $P^{-1} = P^T$.
Ainsi,

$$\mathbb{E}[(v^T X_k)^2] = v^T PDP^T v = \sum_{k=1}^{n} \lambda_k < v, u_k >^2$$

$$\mathbb{E}[(v^T X_k)^2] \leq \lambda_1 \sum_{k=1}^{n} < v, u_k >^2$$

$\forall v \in \mathbb{C}^d$, tel que $\|v\| = 1$, on a égalité dans l'inégalité large précédente, lorsque $v = u_1$.
Ainsi, c'est bien le vecteur u_1 qui maximise l'énergie de projection et cette dernière vaut $\mathbb{E}[(v^T X_k)^2] = \lambda_1 \|u_1\| = \lambda_1$.

3.1.2 Montrons que le vecteur orthogonal à u_1 maximisant l'énergie de projection est donné par u_2

On se base sur la relation trouvée précédemment

$$\mathbb{E}[(v^T X_k)^2] = \sum_{k=1}^{n} \lambda_k < v, u_k >^2$$

On cherche un vecteur $v \in \mathbb{C}^d$, avec $\|v\| = 1$ et $< v, u_1 >= 0$.
Ainsi,

$$\mathbb{E}[(v^T X_k)^2] = \sum_{k=2}^{n} \lambda_k < v, u_k >^2 \leq \lambda_2 \sum_{k=2}^{n} < v, u_k >^2$$

On a égalité lorsque $v = u_2$, donc le vecteur orthogonal à u_1 maximisant l'énergie de projection est bien donné par u_2 et l'énergie de projection vaut, avec le même raisonnement que précédemment, λ_2.

3.1.3 Montrons que le vecteur orthogonal à $u_1, \cdots, u_{l-1}$ maximisant l'énergie de projection est donné par u_l

On cherche un vecteur $v \in \mathbb{C}^d$, avec $\|v\| = 1$
et $< v, u_1 >= \ldots =< v, u_{l-1} >= 0$.
Ainsi,

$$\mathbb{E}[(v^T X_k)^2] = \sum_{k=l}^{n} \lambda_k < v, u_k >^2 \leq \lambda_l \sum_{k=l}^{n} < v, u_k >^2$$

On a égalité lorsque $v = u_l$, donc le vecteur orthogonal à $u_1, \cdots, u_{l-1}$ maximisant l'énergie de projection est bien donné par u_l et l'énergie de projection vaut, toujours avec le même raisonnement que précédemment, λ_l

3.2 ACP et classification d'échantillons gaussiens

3.3 ACP et compression d'images