Corrigé : Feuille de travaux dirigés 5

Solution Exercice 1

1. Le rapport de vraisemblance s'écrivant

$$\phi(Y) = \sqrt{\frac{det(\Sigma_0)}{det(\Sigma_1)}} \exp\left(-\frac{1}{2}Y^t(\Sigma_1^{-1} - \Sigma_0^{-1})Y\right),$$

le test de Neyman Pearson est

$$\mathbb{1}_{\{\phi(Y)\geq t_\alpha\}} = \mathbb{1}_{\left\{Y^t(\Sigma_1^{-1}-\Sigma_0^{-1})Y\leq 2\log(\sqrt{det(\Sigma_0)/det(\Sigma_1)}/t_\alpha)\right\}},$$

au niveau α (le seuil t_α étant choisi de façon tel que $\mathbb{P}_{Y\sim\mathcal{N}(0,\Sigma_0)}\{\phi(Y)\geq t_\alpha\} = \alpha$).

2. L'absence de signal correspondant au cas où $\sigma_X^2 = 0$, il s'agit de tester l'hypothèse nulle $H_0 : \sigma_X^2 = 0$ contre l'alternative $H_1 : \sigma_X > 0$. En désignant par I_n la matrice $n\times n$ unité, on se trouve dans le cadre de la question précédente avec $\Sigma_0 = \sigma_V^2 I_n$ et $\Sigma_1 = (\sigma_V^2 + \sigma_X^2)I_n$ et le test de Neyman-Pearson au niveau α sécrit

$$\mathbb{1}_{\{\sum_{i=1}^n Y_i^2\geq s_\alpha\}},$$

où le seuil s_α est tel que $\mathbb{P}_{Y\sim\mathcal{N}(0,\sigma_V^2)}\{\sum_{i=1}^n Y_i^2 \geq s_\alpha\} = \alpha$.

3. Sous H_0, la loi de Y_i/σ_V étant $\mathcal{N}(0,1)$, le seuil est donc $s_\alpha = \sigma_V^2 q_{1-\alpha}$ où $q_{1-\alpha}$ désigne le quantile de niveau $1-\alpha$ de la loi du chi-deux à n degrés de liberté.

Solution Exercice 2

1. Le rapport de vraisemblance relatif à l'hypothèse nulle $H_0 : \theta = \theta_0$ contre l'alternative $H_1 : \theta = \theta_1$, avec $\theta_1 > \theta_0$, est égal à :

$$\begin{aligned}\phi(X_1,\ \dots,\ X_n) &= \exp\left(\frac{1}{2}\sum_{i=1}^n (2X_i - (\theta_0+\theta_1))(\theta_1-\theta_0)\right)\\ &= \exp\left(n(\theta_1-\theta_0)\bar{X}_n - (n/2)(\theta_1^2-\theta_0^2)\right).\end{aligned}$$

Cette dernière expression est une fonction croissante de $\bar{X}_n$ car $\theta_1 > \theta_0$. Le test de Neyman-Pearson de niveau α peut donc sécrire :

$$\mathbb{1}_{\{\phi(X_1,\ \dots,\ X_n)\geq s_\alpha\}} = \mathbb{I}_{\{\bar{X}_n\geq t_{n,\alpha}\}}$$

avec $t_{n,\alpha} = \theta_0 + q_{1-\alpha}/\sqrt{n}$ où $q_{1-\alpha}$ désigne le quantile au niveau $1-\alpha$ de la loi $\mathcal{N}(0,1)$.

2. On remarque que $\theta \mapsto \mathbb{P}_\theta\{\bar{X}_{25} > t\}$ est croissante et on déduit ainsi de la question précdente que :
$$t_{25,5} = 1.645/\sqrt{25} = 0.329 \,.$$

3. Il s'agit de calculer
$$\begin{aligned}\mathbb{P}_{\log(1.5)}\{\bar{X}_{25} \leq 0.329\} &= \mathbb{P}\{Z \leq \sqrt{25}(0.329 - \log(1.5))\}\\ &\approx \mathbb{P}\{Z \leq 0.765\} = 0.778 \,.\end{aligned}$$

4. Si on pose $X'_i = -X_i$ pour tout $i \in \{1, \ldots, 25\}$ et $\theta' = -\theta$, les hypothèses se réécrivent $H_0 : \theta' \leq 0$ vs $H_1 : \theta' > 0$. Un test solution est donc $\mathbb{I}\{\bar{X}'_{25} > t_{25,5}\} = \mathbb{I}\{\bar{X}_{25} < -t_{25,5}\}$.

Solution Exercice 3

1. L'observation peut s'écrire de façon vectorielle $Y = \theta_1 A + V$ où $A = (x_1, \ldots, x_n)$ est un vecteur déterministe et $V = (V_1, \ldots, V_n)$ est un vecteur Gaussien (ses composantes sont des Gaussiennes i.i.d. $\mathcal{N}(0, \sigma^2)$) centré de matrice de covariance $\sigma^2 I_n$, I_n désignant la matrice $n \times n$ unité. La vraisemblance du modèle statistique (dominé par la mesure de Lebesgue) s'écrit donc :
$$p_\theta(Y) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (Y_i - x_i\theta_1)^2\right) .$$

2. On maximise la log-vraisemblance. Les equations de score s'écrivent :
$$\begin{aligned}0 &= \frac{\partial}{\partial \sigma^2}\log p_\theta(Y) = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^n (Y_i - \theta_1 x_i)^2,\\ 0 &= \frac{\partial}{\partial \theta_1}\log p_\theta(Y) = \frac{1}{\sigma^2}\sum_{i=1}^n x_i(Y_i - \theta_1 x_i).\end{aligned}$$

La solution (on vérifiera qu'il s'agit d'un maximum global en calculant la hessienne) est donc :
$$\begin{aligned}\widehat{\theta}_1 &= \frac{\sum_{i=1}^n Y_i x_i}{\sum_{i=1}^n x_i^2},\\ \widehat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^n (Y_i - \widehat{\theta}_1 x_i)^2.\end{aligned}$$

3. On vérifie que le maximum de $p((0,\sigma^2), Y)$ est atteint en $\widetilde{\sigma}^2 = (1/n)\sum_{i=1}^n Y_i^2$, ainsi le rapport de vraisemblance est
$$Z = \frac{p(\widehat{\theta}, Y)}{p((0, \widetilde{\sigma}^2), Y)} = \left(\frac{\widetilde{\sigma}^2}{\widehat{\sigma}^2}\right)^{n/2}$$
Le test basé sur le rapport de vraisemblance généralisée s'écrit
$$\delta(Y) = \mathbb{1}_{\frac{\widetilde{\sigma}^2}{\widehat{\sigma}^2} > c}$$
où c est une constante (qu'il n'est pas demandé de calculer), telle que $\mathbb{P}_{H_0}(Z > c) = \alpha$.