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. Dans le cas général, la variable aléatoire $T(X)$ est gaussienne, en tant que combinaison linéaire de gaussiennes. Son espérance est $\mathbb{E}(\bar{X}_A) - \mathbb{E}(\bar{X}_B) = \mu_A - \mu_B = \Delta$. Sa variance est (par indépendance) :

$$\mathbb{V}\mathrm{ar}(T(X)) = \mathbb{V}\mathrm{ar}(\bar{X}_A) + \mathbb{V}\mathrm{ar}(\bar{X}_B) = \left(\frac{1}{n_A} + \frac{1}{n_B}\right)\sigma^2 := \sigma_T^2.$$

Sous l'hypothèse nulle, on a donc $T(X) \sim \mathcal{N}\left(0, \left(\frac{1}{n_A} + \frac{1}{n_B}\right)\sigma^2\right) = \mathcal{N}(0, \sigma_T^2)$.

2. L'argument de la question 1) montre que sous $\tilde{H}_1$, avec $\Delta = \mu_A - \mu_B > 0$; $T(X) \sim \mathcal{N}(\Delta, \sigma_T^2)$ et sous H_0, $T(X) \sim \mathcal{N}(0, \sigma_T^2)$. L'observation étant $T(X)$, le rapport de vraisemblance s'écrit donc

$$\begin{aligned} Z = \frac{p_1(T(X)}{p_0(T(X))} &= \frac{\exp\left(\frac{-(T(X)-\Delta)^2}{2\sigma_T^2}\right)}{\exp\left(\frac{-T(X)^2}{2\sigma_T^2}\right)} \\ &= \exp\left(\frac{-(T(X) - \Delta)^2 + T(X)^2}{2\sigma_T^2}\right) \\ &= \exp\left(\frac{2\Delta T(X) - 2\Delta^2}{2\sigma_T^2}\right) \\ &= \Psi(\Delta, T(X)) \end{aligned}$$

où $t \mapsto \Psi(\Delta, t)$ est une fonction strictement croissante. Le test de Neyman-Pearson basé sur l'observation $T(X)$ est, d'après le cours, de type

$$\delta(T(X)) = \begin{cases} 1 & \text{si } Z > c \\ 0 & \text{sinon} \end{cases}$$

où c est tel que $\mathbb{P}_{\Delta=0}(Z > c) = \alpha = 0.05$. Or, $t \mapsto \Psi(\Delta, t)$ étant strictement croissante, pour tout $c \in \mathbb{R}$, il existe $C \in \mathbb{R}$ tel que

$$Z > c \iff \Psi(T(X), \Delta) > c \iff T(X) > C.$$

Le test de Neyman-Pearson revient donc à comparer $T(X)$ à un seuil C. On détermine C en se souvenant que le test doit satisfaire, sous H_0, $\mathbb{P}(\delta(T(X) = 1)) = \alpha$, c'est-à-dire $\mathbb{P}(T(X) > C) = \alpha$, avec $T(X) \sim \mathcal{N}(0, \sigma_T^2)$; ou encore $\mathbb{P}\Big(T(X)/\sqrt{\sigma_T^2} > C/\sigma_T^2\Big) = \alpha$ avec $T(X)/\sqrt{\sigma_T^2} \sim \mathcal{N}(0,1)$. Cette dernière condition est équivalente à $C/\sqrt{\sigma_T^2} = q_{\mathcal{N}}(1-\alpha)$ avec $q_{\mathcal{N}}(1-\alpha)$ le quantile d'ordre $1-\alpha$ de la loi normale, d'où

$$C = \sqrt{\sigma_T^2} q_{\mathcal{N}}(1-\alpha) = \sqrt{(\frac{1}{n_A} + \frac{1}{n_B})\sigma^2} \times q_{\mathcal{N}}(1-\alpha).$$

3. avec les données de l'énoncé,

$$C = \sqrt{\left(\frac{1}{4} + \frac{1}{3}\right)} \times 1.645 \ = \ \sqrt{\frac{7}{12}} \times 1.645 > \sqrt{1/2} \times 1.645 \simeq 1.645/1.414 > 1.$$

D'autre part,

$$T(x) = 13 - 12 = 1,$$

d'où $T(x) < C$, et $\delta(T(x)) = 0$. On ne rejette pas H_0 contre $\tilde{H}_1$.

4. On considère maintenant $H_0 : \Delta = 0$ contre $H_1 : \Delta > 0$. Le modèle pour l'observation $T(X)$ est $\mathcal{P} = \{P_\theta, \theta \geq 0\}$ avec $\theta = \Delta$, et $P_\theta = \mathcal{N}(\theta, \sigma_T^2)$ avec $\theta \geq 0$ inconnu, σ_T^2 connu. On peut écrire les deux hypothèses :

$$H_0 : \theta \in \Theta_0 = \{0\}$$

et

$$H_1 : \theta \in \Theta_1 =]0, \infty[.$$

la région d'acceptation construite à la question 2) ne dépend pas de la valeur de Δ et le risque de première espèce est

$$\sup_{\theta \in \Theta_0} \mathbb{P}_\theta(\delta(T(X)) = 1) = \mathbb{P}_{\theta=0}(\delta(T(X)) = 1) = \alpha$$

le test a donc un risque de première espèce égal à α : c'est un test de niveau α de H_0 contre H_1.

5. Il reste à montrer que δ est uniformément plus puissant, c'est-à-dire : pour tout autre test δ' de niveau α, pour tout $\theta \in \Theta_1$,

$$R(\theta, \delta') \geq R(\theta, \delta).$$

Soit δ' un autre test de niveau α. Pour tout $\theta \in \Theta_1$, le test δ construit au 2 est le test de Neyman-Pearson pour les hypothèses simples H_0, $\tilde{H}_1 = \{\Delta = \theta\}$. D'après le théorème de Neyman-Pearson, on a bien $R(\theta, \delta') \geq R(\theta, \delta)$, ce qu'il fallait démontrer.

Solution Exercice 3

1. Puisque l'on donne $s > u$, $g(\theta) = \mathbb{P}_\theta(X_1 > s) = 1 - F_\theta(s) = (u/s)^\theta$.
2. d'après le point précédent

$$g(\theta) \leq \rho_0 \iff \theta \log(u/s) \leq \log \rho_0 \iff \theta \geq \theta_0 = \frac{\log \rho_0}{\log(u/s)}$$

(puisque $\log(u/s) < 0$)
A.N : $\theta_0 = \log(1/1000)/\log(1/10) = 3/1 = 3$. $g(\theta) = 10^{-\theta}$.

3. d'après la question 1, en notant $p_1 = p_{\theta_1}^{\otimes n}$ et $p_0 = p_{\theta_0}^{\otimes n}$, le rapport de vraisemblance est

$$\begin{aligned}\Phi_{\theta_0,\theta_1}(X) &= \frac{p_1(X)}{p_0(X)} = \exp(\log p_1(X) - \log p_0(X)) \\ &= \exp\left(C_n(\theta_0, \theta_1) + (\theta_0 - \theta_1) \sum_{i=1}^n \log(X_i/u)\right)\end{aligned}$$

où $C_n(\theta_0, \theta_1)$ est une constante ne dépendant pas de X. Le test de Neyman-Pearson, d'après le cours, est de type $\delta(X) = \mathbb{1}_{\Phi_{\theta_0,\theta_1}(X) \geq c'}$. Or, $\theta_0 > \theta_1$, donc le rapport de vraisemblance est une fonction strictement croissante de $W = \sum \log(X_i/u)$. Il existe donc $c \in \mathbb{R}$ tel que $\Phi_{\theta_0,\theta_1}(X) \geq c' \iff W \geq c$, d'où le résultat.

4. D'après la question 3., sous $\tilde{H}_0$, $\theta_0 W \sim \mathcal{G}amma(n, 1)$. On a donc

$$\mathbb{P}_{\theta_0}(\theta_0 W > q_n(1-\alpha)) = \alpha.$$

i.e.

$$\mathbb{P}_{\theta_0}(W > \frac{q_n(1-\alpha)}{\theta_0}) = \alpha.$$

d'où $c = \frac{q_n(1-\alpha)}{\theta_0}$. Ce test est U.P.P. de niveau α d'après le cours.

5. Pour $t > \theta_0$, on a, sous $H_0(t)$, $tW \sim \mathcal{G}amma(n, 1)$, d'où

$$\mathbb{P}_t(W > c) = \mathbb{P}_t(tW > \frac{t}{\theta_0} q_n(1-\alpha)) < \mathbb{P}_t(tW > q_n(1-\alpha))$$

car $t/\theta_0 > 1$, et car la densité de la loi Gamma ne s'annulant pas sur $]0, \infty[$, $\mathbb{P}_t(tW \in [q_n(1-\alpha), \frac{t}{\theta_0} q_n(1-\alpha)]) > 0$.

6. D'après le point précédent, avec $\Theta_0 = [\theta_0, \infty[$, on a

$$\sup_{\theta \in \Theta_0} R(\theta, \delta) = \sup_{\theta \geq \theta} \mathbb{P}_\theta(W > c) = \mathbb{P}_{\theta_0}(W > c) = \alpha.$$

   Le niveau du test δ pour l'hypothèse composite H_0 est donc bien α. Soit δ' un autre test de niveau α pour H_0. en particulier, $R(\theta_0, \delta') \leq \alpha$ donc δ' est de niveau α pour $\tilde{H}_0$. Mais on sait que δ est U.P.P. pour $\tilde{H}_0$ donc $\beta(\theta_1, \delta') \leq \beta(\theta_1, \delta)$. Ceci étant vrai pour tout autre test δ' de niveau α pour H_0, le test δ et bien *U.P.P.* pour H_0 contre $\tilde{H}_1$.

7. Le raisonnement ci-dessus étant valide pour tout $\theta_1 < \theta_0$, on a bien, pour tout autre test δ', pour tout $\theta_1 \in \Theta_1$, $\beta(\theta_1, \delta') \leq \beta(\theta_1, \delta)$. Le test δ est donc U.P.P de niveau α pour tester H_0 contre H_1.