ENSAE

ECOLE NATIONALE
DE LA STATISTIQUE
ET DE L'ADMINISTRATION
ECONOMIQUE

Deuxième année
2002-2003

Théorie des tests
Enoncé des travaux dirigés n°1 à 1

Guillaume Lacôte
Bureau **E03**
✉ Guillaume.Lacote@ensae.fr
☞ http://ensae.no-ip.com/SE207

Dernière mise-à-jour: 20030318-09h21

Table des matières

1 Travaux Dirigés n°1

Exercice 1

Dans l'outillage de votre usine vous utilisez une grande quantité de pièces d'un certain modèle. Dans les conditions usuelles d'emploi, vous avez observé que la durée de vie de ces pièces est une variable aléatoire normale dont l'espérance mathématique est $m_0 = 120$ heures, et l'écart-type est $\sigma = 19,4$ heures.

Le représentant d'un fournisseur vous propose un nouveau modèle, en promettant un gain de performance en moyenne de 5%, pour une dispersion identique σ.

Vous décidez de tester le nouveau modèle sur un échantillon de $n = 64$ unités. On note $X_i (i = 1, ..., 64)$ la durée de vie des pièces testées.

☞ Q1 Soit μ l'espérance mathématique du nouveau modèle. Donner un estimateur sans biais de μ, et la variance de cet estimateur.

☞ Q2 Vous ne voulez pas changer de modèle si le nouveau n'est pas plus performant que l'ancien (hypothèse $\mathcal{H}$). Plus précisément vous voulez que la probabilité d'adopter à tort le nouveau modèle ne dépasse pas le seuil 0,05.

Quelle est alors la procédure de décision ?

Evaluez le risque que cette procédure vous fasse rejeter le nouveau modèle si l'annonce du représentant est exacte.

Les 64 pièces testées ont eu une durée de vie moyenne égale à 123,5 heures. Que concluez-vous ?

☞ Q3 *(facultatif)* Le représentant conteste cette procédure, prétextant qu'il vaut mieux partir de l'hypothèse $\mathcal{H}$', selon laquelle le gain de performance moyen est réellement de 5% (hypothèse $\mathcal{H}$'). Il souhaite que la probabilité de rejeter à tort le nouveau modèle ne dépasse pas le seuil 0,05.

Quelle est alors la procédure de décision ? Quel est le risque de l'acheteur ? Quel est le résultat de cette procédure au vu des observations faites. Commentez.

☞ Q4 *(facultatif)* Quelle procédure proposez-vous pour égaliser les risques de l'acheteur et du vendeur ? Quel est alors ce risque ?

Exercice 2

Les fonctions positives considérées sont à valeurs dans $\mathbb{R}^+ \cup \{+\infty\}$. Par convention, pour $a > 0$, $a/0 = +\infty$. Le rapport $0/0$ est indéterminé. Soit $x_1, \dots, x_n$ un échantillon d'observations de variables aléatoires $X_1, \dots, X_n$ i.i.d. uniformes sur l'intervalle $[0, \theta^\star]$, avec $\theta^\star > 0$. On note $x_{(1,n)} \leq \dots \leq x_{(n,n)}$ l'échantillon ordonné correspondant. Il est associé aux variables aléatoires ordonnées $X_{(1,n)} \leq \dots \leq X_{(n,n)}$.

☞ Q1 (a) Montrer que la fonction de répartition $F^{(n)}$ de la variable aléatoire $X_{(n,n)}$ vérifie $F^{(n)}(x) = 0$ pour $x < 0$, $F^{(n)}(x) = 1$ pour $x \geq \theta^\star$, et

$$F^{(n)}(x) = \left(\frac{x}{\theta^\star}\right)^n \quad \text{pour } 0 \leq x \leq \theta^\star.$$

(b) Montrer que la fonction de vraisemblance $\theta \to L_n(\theta)$ de l'échantillon vaut θ^{-n} si $x_{(n,n)} \leq \theta$ et 0 si $x_{(n,n)} > \theta$.

(c) Montrer que la statistique $X_{(n,n)}$ est exhaustive pour θ. Quel estimateur de θ proposer ?

(d) Soient $0 \leq \theta_0 < \theta_1$ deux valeurs distinctes du paramètre θ. On souhaite tester l'hypothèse $H_0 = [\theta = \theta_0]$ contre l'alternative $H_1 = [\theta = \theta_1]$. Quelle forme aurait le test le plus puissant de niveau $0 < \alpha < 1$ de H_0 contre H_1 ?

On cherche maintenant à déterminer un test déterministe sans biais de niveau α de $H_0 = [\theta \leq \theta_0]$ contre $H_1 = [\theta > \theta_0]$.

(a) i. Soient $0 < \theta_0 < \theta_1$ deux valeurs distinctes du paramètre θ. On souhaite d'abord construire un test déterministe de niveau α exactement de l'hypothèse simple $H_0 = [\theta = \theta_0]$ contre l'alternative simple $H_1 = [\theta = \theta_1]$.

ii. Déterminer $c_{\alpha,n}$ de sorte que le test de région de rejet $W_n = \{(x_1, \ldots, x_n) : x_{(n,n)} > c_{\alpha,n}\}$ soit de niveau α exactement. La valeur critique $c_{\alpha,n}$ dépend-elle de θ_1 ?

(b) b) Proposer une approximation simple de $c_{\alpha,n}$ pour α petit.

(c) Calculer la puissance du test ainsi construit. Quelle est la limite de cette puissance quand $n \to \infty$?

(d) Déterminer précisément la fonction

$$\theta \geq 0 \quad \longrightarrow \quad E_\theta \left[\mathbf{1}_{W_n}(x_1, \ldots, x_n) \right].$$

Montrer qu'elle est croissante. En déduire que le test est un test déterministe sans biais de niveau α de $H_0 = [\theta \leq \theta_0]$ contre $H_1 = [\theta > \theta_0]$.

(e) Comment construire un test déterministe sans biais de niveau α de $H_0 = [\theta \geq \theta_0]$ contre $H_1 = [\theta < \theta_0]$?

Exercice 3

☞ Q1 Variables suivant une loi de Poisson

Soient $X_1, \ldots, X_n$ et $Y_1, \ldots, Y_n$ deux échantillons de taille $n \geq 1$ de variables aléatoires (va) indépendantes et identiquement distribuées (iid), à valeurs dans l'ensemble $\mathbf{N}$ des entiers naturels. On suppose les va X_i et Y_j indépendantes. On suppose les va X_i issues d'une loi de Poisson de paramètre $\lambda_1 > 0$, et les va Y_j issues d'une loi de Poisson de paramètre $\lambda_2 > 0$. Les paramètres λ_1 et λ_2 sont inconnus. On souhaite étudier un test de H_0 $[\lambda_1 \leq \lambda_2]$ contre H_1 $[\lambda_1 > \lambda_2]$, puis un test de H_0 $[\lambda_1 = \lambda_2]$ contre H_1 $[\lambda_1 \neq \lambda_2]$. On note $T_1 = \sum_{i=1}^n X_i$ et $T_2 = \sum_{j=1}^n Y_j$.

(a) Exprimer $P\{X_1 = x_1, \ldots, X_n = x_n\}$ en fonction de $P\{T_1 = t_1\}$. Pourquoi cela établit-il que T_1 est une statistique exhaustive ?

(b) Exprimer $P\{T_1 = t_1, T_2 = t_2\}$ sous la forme

$$P\{T_1 = t_1, T_2 = t_2\} = Z_{\mathbf{T}}^{-1}(\theta_1, \theta_2) h_{\mathbf{T}}(t_1, t_2) \exp(\theta_1 t_1 + \theta_2 t_2),$$

en précisant ce que sont θ_1, θ_2, et en explicitant $Z_{\mathbf{T}}(\theta_1, \theta_2)$ et $h_{\mathbf{T}}(t_1, t_2)$.

☞ Q2 **Test bilatéral**

On veut tester H_0 $[\lambda_1 = \lambda_2]$ contre H_1 $[\lambda_1 \neq \lambda_2]$. On effectue le changement de paramètre $\tau_1 = \theta_1 - \theta_2$, $\tau_2 = \theta_2$, d'où

$$\exp(\theta_1 T_1 + \theta_2 T_2) = \exp(\tau_1 T_1 + \tau_2 (T_1 + T_2)),$$

et l'on note $U = T_1 + T_2$.

(a) Démontrer *de manière détaillée* que la loi conditionnelle de T_1 sachant que $U = u$, $u \geq 1$, est une loi binomiale $\mathrm{Bin}(u; p = \lambda_1/(\lambda_1 + \lambda_2))$. Exprimer cette loi conditionnelle à l'aide des paramètres θ_1 et θ_2, en faisant apparaître la différence $\tau_1 = \theta_1 - \theta_2$.

(b) Montrer *rigoureusement* que le test de H_0 $[\lambda_1 = \lambda_2]$ contre H_1 $[\lambda_1 \neq \lambda_2]$, de niveau de signification α, UPP parmi les tests α-semblables sur l'ensemble $\{\lambda_1 = \lambda_2\}$, admet une région de rejet de la forme

$$0 \leq T_1 \leq a(u) \quad \text{ou} \quad b(u) \leq T_1 \leq u$$

si $U = u$ avec $u \geq 1$, $a(u) \leq b(u)$ désignant des entiers, et écrire *avec précision* les deux équations permettant de déterminer les entiers $a(u)$ et $b(u)$, en faisant apparaître la symétrie par rapport à $u/2$.

(c) Si u est assez grand, comment pourrait-on déterminer $a(u)$ et $b(u)$ de manière approchée ? Donnez des *formules explicites* de ces approximations de $a(u)$ et $b(u)$, de telle sorte que l'on puisse utiliser ces formules pour implémenter effectivement ce test.