DAVID SROUR

Problèmes de tests Ho composite contre h1 simple hypothèse la moins favorable et statistiques exhaustives en présence d'un bruit

Revue de statistique appliquée, tome 20, nº 1 (1972), p. 89-113
<http://www.numdam.org/item?id=RSA_1972__20_1_89_0>

PROBLÈMES DE TESTS H_o COMPOSITE CONTRE h_1 SIMPLE HYPOTHÈSE LA MOINS FAVORABLE ET STATISTIQUES EXHAUSTIVES EN PRÉSENCE D'UN BRUIT

David SROUR

Ingénieur en chef à CIM Développement

1 - INTRODUCTION

On se propose de traiter dans cet article des problèmes de tests paramétriques à hypothèse nulle composite, contre alternative simple :

$$(\mathrm{I}) \begin{cases} H_o : \theta \in \omega \\ h_1 : \theta = \theta_1 \quad (\theta_1 \in \Omega - \omega) \end{cases}$$

et plus précisément de la recherche d'un test le plus puissant pour le problème (I). L'espace des paramètres étant noté Ω et le sous-espace correspondant à l'hypothèse nulle ω, on cherchera occasionnellement à voir si le test le plus puissant défini pour (I), n'est pas uniformément le plus puissant (u. m. p) pour le problème plus général (II), à alternative composite :

$$(\mathrm{II}) \begin{cases} H_o : \theta \in \omega \\ H_1 : \theta \in \Omega - \omega \end{cases}$$

Les problèmes du type (I) se ramènent généralement au problème du test d'une hypothèse simple h_0 contre une hypothèse simple h_1, l'hypothèse nulle composite H_0 étant remplacée par l'hypothèse simple la "moins favorable", c'est-à-dire en quelque sorte "la plus difficile à tester contre h_1"

$$(1) \begin{cases} h_o : \text{hypothèse la moins favorable} \\ h_1 : \theta = \theta_1 \end{cases}$$

La notion d'hypothèse la moins favorable est rappelée dans le paragraphe 2 où il apparaît aussi que le test le plus puissant pour (1) (selon la méthode classique de Neyman et Pearson), est aussi le plus puissant pour le problème initial (I).

La définition ci-dessous de l'hypothèse h_o la moins favorable n'en rend pas la recherche facile. Bien que l'on ait établi certaines conditions d'existence Réf.[1], la recherche de l'hypothèse la moins favorable n'a pas été, du moins à notre connaissance, systématiquement entreprise. Cependant, la majorité des problèmes usuels, en particulier dans le cadre de lois normales, possèdent une hypothèse "la moins favorable" et peuvent par conséquent être traités selon la méthode exposée ci-dessus.

Pour les problèmes de tests à un paramètre du type (I), l'hypothèse h_o la moins favorable est le plus souvent immédiate à trouver, du moins dans le cadre d'ensembles ω "classiques". Divers exemples en seront donnés dans le paragraphe 3.

Le paragraphe 4 est consacré à l'étude des tests à deux paramètres, l'un d'entre eux (θ) faisant l'objet du test, l'autre (η) intervenant comme un paramètre de bruit. Les "équivalents" des problèmes (I) et (II) ci-dessus (respectivement (III) et (IV)) sont alors les suivants :

$$\text{(III)} \begin{cases} H_o & : \theta \in \omega, \ \eta \ \text{(non spécifié)} \\ h_1 & : \theta = \theta_1 \ ; \ \eta = \eta_1 \ (\theta_1 \in \Omega - \omega) \end{cases}$$

$$\text{(IV)} \begin{cases} H_o & : \theta \in \omega, \ \eta \ \text{(non spécifié)} \\ H_1 & : \theta \in \Omega - \omega, \ \eta = \eta_1 \end{cases}$$

mais il est aussi fort utile d'étudier le problème le plus général (V)

$$\text{(V)} \begin{cases} H_o & : \theta \in \omega, \ \eta \ \text{(non spécifié)} \\ H_1 & : \theta \in \Omega - \omega, \ \eta \ \text{(non spécifié)} \end{cases}$$

La recherche de l'hypothèse la moins favorable pour le problème (III) est plus délicate que pour le problème (I), mais devient cependant facile, lorsqu'il existe une "statistique exhaustive pour θ en présence du bruit η". Cette notion est précisément définie au cours du paragraphe 4, où l'on envisage par ailleurs plusieurs exemples.

Cet article se limitant pratiquement aux distributions normales, rappelons que U désignera la variable aléatoire normale centrée réduite. Sa densité de probabilité sera notée f(u), sa fonction de répartition F(u) et $u_{1-\alpha}$ désignera la valeur ayant la probabilité α d'être dépassée : $F(u_{1-\alpha}) = 1 - \alpha$.

De façon générale les hypothèses composites seront notées par H et les hypothèses simples par h. Pour une hypothèse nulle composite, H_o, h_o' désignera une hypothèse simple quelconque de H_o et h_o l'hypohtèse la moins favorable. Le passage d'un problème "composite-simple" (types I et III) au problème "simple-simple" correspondant à h_o "la moins favorable" se fera en remplaçant les chiffres romains par des chiffres arabes. Pour les problèmes du type (1), la région critique la plus puissante sera notée $W_\alpha(h_o)$ étant sous-entendu que toutes les régions critiques dont nous parlerons seront de niveau α. La puissance maximale sera noté indifféremment $P_{\theta_1}(h_o)$, ou $P_{\theta_1}(W_\alpha(h_o))$ et nous ne ferons par ailleurs aucune différence entre région critique $W_\alpha(h_o)$ et test $\Phi_{h_o}^{(\alpha)}$. L'ensemble des réels et celui des réels positifs seront désignés par R et R^+ respectivement et des renvois fréquents seront faits aux problèmes I, II, III, IV, V exposés dans ce paragraphe.

2 - DEFINITIONS ET RAPPELS

Hypothèse la moins favorable

Considérons le problème du type (I) exposé dans le § 1, où en l'absence de paramètre importun, l'hypothèse alternative est simple :

$$(I)\quad \begin{cases} H_o : \theta \in \omega \\ h_1 : \theta = \theta_1 \ (\theta_1 \in \Omega - \omega) \end{cases}$$

Afin de définir l'hypothèse h_o la moins favorable et pour les développements qui suivront, nous entendrons par hypothèse simple courante h'_o de H_o une distribution de probabilité λ' sur ω et non pas forcément une valeur θ' de ω. Pour deux hypothèses courantes h'_o et h''_o de H_o, l'hypothèse h_o est dite "moins favorable à tester contre h_1 au niveau α que l'hypothèse h''_o si :

$$(2.1)\qquad P_{\theta_1}(W_\alpha(h'_o)) < P_{\theta_1}(W_\alpha(h''_o))$$

c'est-à-dire si la puissance du test de niveau α le plus puissant pour le problème (1') $\begin{cases} h'_o \text{ dans } H_o \\ h_1 : \theta = \theta_1 \end{cases}$ est inférieure ou égale à la puissance du test de niveau le plus puissant pour le problème (1") $\begin{cases} h''_o \text{ dans } H_o. \\ h_1 : \theta = \theta_1. \end{cases}$

L'hypothèse simple h_o sera dite la moins favorable si l'on a :

$$(2.2)\qquad P_{\theta_1}(W_\alpha(h_o)) < P_{\theta_1}(W_\alpha(h'_o) \quad \forall \quad h'_o \in H_o$$

Plus précisément, si les observations utilisées pour le test sont celles d'une variable X de densité $g_\theta(x)$, le problème (1') par exemple, s'écrira comme suit :

$$(1')\quad \begin{cases} h'_o : \text{densité } \int_\omega g_\theta(x)\, d\lambda'(\theta) = g_{\lambda'}(x) \\ h_1 : \text{densité } g_{\theta_1}(x) \end{cases}$$

et l'on ne fera pas de différence formelle entre l'hypothèse h_o la moins favorable et la distribution λ la moins favorable.

Le passage de (I) à (1') revient à dire que l'on recherche, non pas une région critique $W(H_o)$ satisfaisant à

$$(2.3)\qquad P_\theta(W(H_o)) \leqslant \alpha \quad \forall\ \alpha \in \omega$$

mais $W(h'_o)$ satisfaisant à la condition moins restrictive :

$$(2.4)\qquad \int_\omega P_\theta(W(h'_o))\, d\lambda'(\theta) \leqslant \alpha$$

Cette procédure est justifiée par le théorème 1 ci-dessous dû à Lehmann et Stein.

Théorème 1 [Réf. 2, 3]

Considérons le problème (I) $\begin{cases} H_o : \theta \in \omega \\ h_1 : \theta = \theta_1 \ (\theta_1 \in \Omega - \omega) \end{cases}$

et le problème (1') $\begin{cases} h'_o : \text{densité } g_{\lambda'}(x) = \int_\omega g_\theta(x)\, d\lambda'(\theta) \\ h_1 : \theta = \theta_1 \end{cases}$,

et supposons que $g_\theta(x)$ soit une fonction mesurable de (x, θ) sur $\mathcal{X} \times \omega$ où $\mathcal{X}$ est l'espace des observations.

Soit $\widehat{W}_\alpha(h'_o)$ la région critique la plus puissante pour le problème (1') et supposons que l'on ait $P_\theta(W_\alpha(h'_o)) \leqslant \alpha, \quad \forall \theta \in \omega$:

- h'_o est alors effectivement l'hypothèse la moins favorable h_o
- $\widehat{W}_\alpha(h'_o)$ est la plus puissante pour le problème (I).

Pour la démonstration de ce théorème, nous renvoyons à [Réf. 2 p. 91-92] et à [Réf. 3], ainsi qu'à [Réf. 4 p. 81].

Du point de vue pratique l'étude du problème (I) se fera en passant par

l'étape (1) $\begin{cases} h_o : g_\lambda(x) = \int_\omega g_\theta(x)\, d\lambda(\theta) \\ h_1 : \theta = \theta_1 \end{cases}$

où h_o est l'hypothèse la moins favorable intuitivement pressentie. $W_\alpha(h_o)$ étant trouvée, on vérifiera la condition (2.5) $P_\theta(W_\alpha(h_o)) \leqslant \alpha \quad \forall\ \theta \in \omega$ (condition que nous noterons pour plus de commodité $P_{H_0}(W(h_o)) \leqslant \alpha$) et l'on en déduira que $W_\alpha(h_o)$ est la plus puissante pour le problème (I). Si de plus, $W_\alpha(h_o)$ est indépendante de θ_1, on aura ainsi trouvé un test u. m. p pour le problème (II) $\begin{cases} H_o : \theta \in \omega \\ H_1 : \theta \in \Omega - \omega. \end{cases}$

C'est seulement dans les cas où il semble évident à priori que h_o est concentrée en un point θ_o de ω, que l'on pourra étudier le problème (1') simplifié : (1') $\begin{cases} h'_o : \theta = \theta'_o \\ h_1 : \theta = \theta_1 \end{cases}$, en déduire $W_\alpha(h'_o)$ et $P_{\theta_1}(h'_o)$ et choisir le θ'_o qui minimise $P_{\theta_1}(h'_o)$ (cf. exemple 1 du paragraphe suivant).

En remarquant que les relations (2.5) et (2.6) :

(2.6) $$\int_\omega P_\theta(W_\alpha(h_o))\, d\lambda(\theta) = \alpha$$

(erreur de première espèce égale à α) ne sont compatibles que si la λ mesure de ω^* est unitaire :

(2.7) $$\omega^* = \{\theta \in \omega\ ;\ P_\theta(W_\alpha(h_o)) = \alpha\}$$

on peut écrire le théorème 1 sous une forme légèrement différente :

<u>Théorème 1'</u> [Réf. 3]

Considérons les problèmes (I) $\begin{cases} H_o : \theta \in \omega \\ h_1 : \theta = \theta_1 (\theta_1 \in \Omega - \omega) \end{cases}$

et (1) $\begin{cases} h_o : g_\lambda(x) = \int_\omega g_\theta(x)\, d\lambda(\theta) \\ h_1 : \theta = \theta_1 \ (\theta_1 \in \Omega - \omega) \end{cases}$

et soit $\omega^* \subset \omega$ tel que $\lambda(\omega^*) = 1$. Si $W_\alpha(h_o)$ constitue le test le plus puissant pour (1) et si $P_\theta(W_\alpha(h_o)) = \alpha \quad \forall \theta \in \omega^*$, alors $W_\alpha(h_o)$ est aussi le plus puissant pour (I).

Par la suite, nous essayerons presque toujours de nous placer dans les conditions du théorème 1, même dans le cas où ω^* est constitué par un ou deux points seulement de ω (exemples 1 et 2 respectivement). L'exemple 5 cependant sera traité selon le théorème 1'.

3 - EXEMPLES DE TESTS A UN PARAMETRE

L'exemple 1 ci-dessous se rapporte au cas où la distribution la moins favorable est concentrée en un point de ω.

a) Exemple 1

Considérons n observations indépendantes $x_1, x_2, \dots, x_n$ d'une variable x suivant la loi $\mathcal{N}(\theta, 1)$ et le problème de test suivant :

$$(\mathrm{I}) \quad \begin{cases} H_o : \theta \leqslant 0 \\ h_1 : \theta = \theta_1 \quad (\theta_1 > 0) \end{cases}$$

L'hypothèse la moins favorable h_o à tester contre h_1 est manifestement $h_o : \theta = 0$, valeur de ω_{H_o} la plus "proche" de h_1. En étudiant le problème (1) $\begin{cases} h_o : \theta = 0 \\ h_1 : \theta = \theta_1 \ (\theta_1 > 0) \end{cases}$, on trouve facilement le test le plus puissant correspondant :

$$(3a.1) \qquad W_\alpha(h_o) : \bar{x} > u_{1-\alpha}/\sqrt{n},$$

où $\bar{x}$ désigne la moyenne des x_i

$$(3a.2) \qquad P_{\theta_1}(h_o) = P_{\theta_1}(W_\alpha(h_o)) = 1 - F(u_{1-\alpha} - \theta_1\sqrt{n})$$

On vérifie par ailleurs que h_o est effectivement la moins favorable : pour θ quelconque dans ω_{H_o} en effet, on a :

$$P_\theta(W_\alpha(h_o)) = P(U > u_{1-\alpha} - \theta\sqrt{n}) \leqslant P(U > u_{1-\alpha}) = \alpha,$$

soit :

$$(3a.3) \qquad P_{H_o}(W_\alpha(h_o)) \leqslant \alpha \text{ (par convention de notation)}$$

Si l'on étudiait enfin le problème (1') $\begin{cases} h'_o : \theta = \theta'_o \ (\theta'_o \leqslant 0) \\ h_1 : \theta = \theta_1 \ (\theta_1 > 0) \end{cases}$ on trouverait :

$$(3a.4) \qquad P_{\theta_1}(h'_o) = 1 - F(u_{1-\alpha} + (\theta'_o - \theta_1)\sqrt{n}),$$

dont la valeur minimale s'obtient bien pour $\theta'_o = 0$ qui maximise la valeur de l'argument de F.

Le test le plus puissant pour (1) étant de niveau α pour H_o d'après (3a.3), il est par conséquent aussi le plus puissant pour (I). Ne dépendant d'ailleur pas de θ_1, il est aussi u. m. p pour le problème plus général

$$(\mathrm{II}) \quad \begin{cases} H_o : \theta \leqslant 0 \\ H_1 : \theta > 0 \end{cases}$$

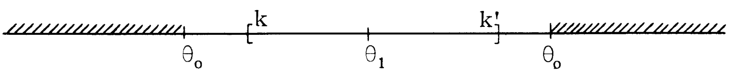
L'exemple 2 ci-dessous est consacré au cas où la distribution la moins favorable est concentrée en deux points de l'hypothèse composite H_o.

b) Exemple 2

Considérons le problème de test suivant et limitons-nous au cas d'une observation unique tirée d'une loi normale de variance unitaire $\mathcal{N}(\theta, 1)$:

$$(\mathrm{I}) \quad \begin{cases} H_o : \theta \leqslant \theta_o \text{ ou } \theta \geqslant \theta'_o \\ h_1 : \theta = \theta_1 = \dfrac{\theta_o + \theta'_o}{2} \end{cases}$$

(θ_1 point milieu de l'intervalle $]\theta_o\ ;\ \theta'_o[$)



Il semble manifeste a priori que l'hypothèse h_o la moins favorable sera :
$h_o \begin{cases} \theta = \theta_o \text{ avec la probabilité } P_o = 1/2 \\ \theta = \theta'_o \quad \text{"} \quad \text{"} \quad P'_o = 1/2 \end{cases}$ et il peut être facilement vu que le test le plus puissant pour le problème
$(1) \begin{cases} h_o : \theta = \theta_o \text{ avec la probabilité } 1/2 \text{ et } \theta = \theta'_o \text{ avec la probabilité } 1/2 \\ h_1 : \theta = \theta_1 = (\theta_o + \theta'_o)/2 \end{cases}$
conduit à la région critique $W(h_o)$:

(3b. 1) $\qquad g_{\theta_o.\theta'_o.\theta_1}(x) < K$ avec

(3b. 2) $$g_{\theta_o.\theta'_o.\theta_1}(x) = \exp\left[\frac{1}{2}(\theta_o - \theta_1)\,(2x - \theta_1 - \theta_o)\right] + \exp\left[\frac{1}{2}(\theta'_o - \theta_1)\,(2x - \theta_1 - \theta'_o)\right]$$

Or la relation vérifiée quelque soit x :

(3b. 3) $$g_{\theta_o,\theta'_o,\theta_1}(x) = g_{\theta_o,\theta'_o,\theta_1}(2\theta_1 - x)$$

montre que $g_{\theta_o\,\theta'_o,\,\theta_1}(x)$ est symétrique par rapport à la droite d'abscisse θ_1, de telle sorte que (3b. 1) est équivalente à :

$$W(h_o) \begin{cases} (3b.4)\ k < x < k' \\ \qquad \text{avec} \\ (3b.5)\ k + k' = 2\theta_1 = \theta_o + \theta'_o \end{cases}$$

(région critique symétrique par rapport à θ_1).

La condition de niveau α pour la région critique -que nous noterons $W_\alpha(h_o)$- s'écrit :

(3b. 6) $$P_{\theta_o}(W_\alpha(h_o)) + P_{\theta'_o}(W_\alpha(h_o)) = 2\alpha,$$

soit encore :

(3b. 6) $$F(k' - \theta_o) - F(k - \theta_o) + F(k' - \theta'_o) - F(k - \theta'_o) = 2\alpha$$

De plus, la symétrie du problème, conduit à l'écriture équivalente à (3b.7) (écrire $k + k' = \theta_o + \theta'_o$ et utiliser $F(-x) = 1 - F(x)$) :

(3b.8) $$F(k' - \theta_o) - F(k - \theta_o) = F(k' - \theta'_o) - F(k - \theta'_o) = \alpha$$

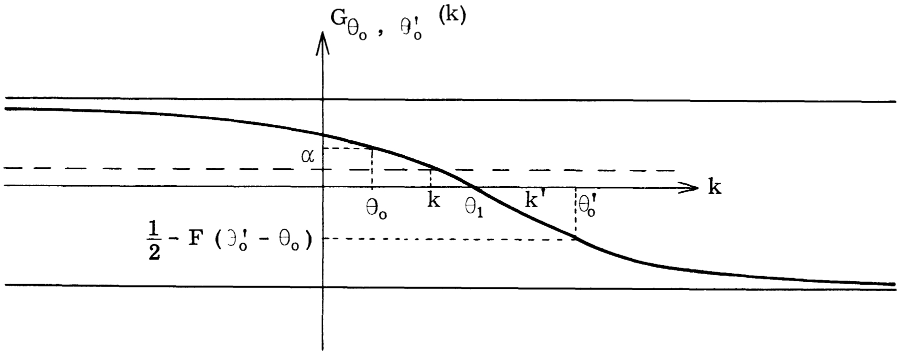
La détermination du seuil critique k se fait en écrivant :

(3b.9) $$F(k' - \theta_o) - F(k - \theta_o) = F(\theta'_o - k) - F(k - \theta_o) = \alpha$$

On peut en effet facilement voir que la fonction :

(3b.10) $$G_{\theta_o,\theta'_o}(k) = F(\theta'_o - k) - F(k - \theta_o)$$

décroît de la valeur 1 à -1, valeurs atteintes pour $k = -\infty$ et $+\infty$ respectivement. $G_{\theta_o,\theta'_o}(k)$ par ailleurs s'annule en son point d'inflexion $\theta_1 = (\theta_o + \theta'_o) / 2$ et le graphique ci-joint illustre la détermination de k



On notera que ce graphique a été construit dans le cas où

(3b.11) $$G_{\theta_o,\theta'_o}(\theta_o) = F(\theta'_o - \theta_o) - \frac{1}{2} = - G_{\theta_o,\theta'_o}(\theta'_o)$$

est supérieur à α ce qui, dans le cas pratique $\alpha = 5\ \%$ par exemple, conduit approximativement à :

(3b.12) $$\theta'_o - \theta_o > 0,125$$

Pour montrer que la distribution choisie est effectivement la moins favorable, prouvons que :

(3b.13) $$P_\theta(\widehat{W}_\alpha(h_o)) \leqslant \alpha \quad \forall\, \theta \leqslant \theta_o \quad \text{ou } \theta \geqslant \theta'_o$$

ou que :

(3b.14) $$F(k' - \theta) - F(k - \theta) \leqslant \alpha \quad \forall\, \theta \leqslant \theta_o \quad \text{ou} \quad \theta \geqslant \theta'_o$$

En considérant le cas $\theta \leqslant \theta_o$ par exemple et en tenant compte de (3b.8), il suffit de montrer que la fonction

(3b.15) $\quad H_k(\theta) = F(k' - \theta) - F(k - \theta)$ est croissante pour $\theta \leqslant \theta_o$.

Or cela est immédiat puisque la dérivée $H_k'(\theta)$ de $H_k(\theta)$:

(3b.16) $\quad H_k'(\theta) = - f(k' - \theta) + f(k - \theta)$ est positive pour tout $\theta \leqslant \theta_o$

Que $k - \theta$ soit positif ou négatif en effet, on a toujours

(3b.17) $$k' - \theta > |k - \theta| \text{ pour tout } \theta \leqslant \theta_o$$

Notons aussi la puissance du test le plus puissant relatif au problème (1) :

(3b.18) $$P_{\theta_1}(h_o) = P_{\theta_1}(W_\alpha(h_o)) = H_k(\theta_1)$$

puissance qui, de par la définition de l'hypothèse la moins favorable, est telle que :

(3b.19) $$P_{\theta_1}(h_o) = \underset{h_o'}{\operatorname{Min}} P_{\theta_1}(h_o')$$

où h_o' est une hypothèse simple "courante" quelconque de H_o.

D'après le théorème 1 rappelé dans le § 2, le test le plus puissant défini ci-dessus pour le problème (1) est aussi le plus puissant pour le problème initial (I). On note enfin d'après (3b.9) que le seuil critique k -et par conséquent k' de façon tout a fait comparable- est indépendant de θ_1 et l'on en déduit que le test défini ci-dessus est en fait le test u. m. p. pour le problème généralisé

$$\text{(II)} \begin{cases} H_o : \theta < \theta_o \quad \text{ou } \theta \geqslant \theta_o' \\ H_1 : \theta_o < \theta < \theta_o' \end{cases}$$

à hypothèses nulle et alternative composites.

Remarques

La considération pour le problème (I) de l'hypothèse $h_1 : \theta = \theta_1$ (point-milieu de $]\theta_o ; \theta_o'[$) évite une erreur commune, consistant dans le cadre du problème (I') par exemple : (I') $\begin{cases} H_o : \theta \leqslant \theta_o \quad \text{ou} \quad \theta \geqslant \theta_o' \\ h_1' : \theta = \theta_1' \qquad (\theta_1' < \theta_1) \end{cases}$ à choisir pour l'hypothèse la moins favorable : $h_o : \theta = \theta_o$ puisque θ_o est plus "proche" de θ_1'

que θ_o'. Le test le plus puissant pour (1') $\begin{cases} h_o : \theta = \theta_o \\ h_1' : \theta = \theta_1' \end{cases}$ est alors :

(3b.20) $$\widehat{W}_\alpha(h_o) : x > \theta_o + u_{1-\alpha}$$

(3b.21) $$P_{\theta_1'}(h_o) = P_{\theta_1'}(\widehat{W}_\alpha(h_o)) = 1 - F(\theta_o + u_{1-\alpha} - \theta_1')$$

Mais la condition (3b.13) ne s'avère alors pas vérifiée, puisque pour $\theta \geqslant \theta'_0$ $P_\theta(W_\alpha(h_o))$, qui d'après (3b.21) est égal à $1 - F(\theta_o + u_{1-\alpha} - \theta)$, dépasse le seuil α.

- L'existence d'un test u. m. p pour le problème II est vérifiée non seulement dans le cadre d'une loi normale à variance connue, mais aussi pour toute loi appartenant à une famille exponentielle à un paramètre (sans paramètre importun ou "bruit") [Réf. 2 p. 88-89]. On démontre aussi que la puissance de ce test possède un maximum en un point θ^* de l'intervalle $]\theta_o ; \theta'_o[$ et est strictement décroissante à gauche ou à droite de θ^*. Dans le cadre de notre exemple on montre facilement que θ^* n'est autre que θ_1 (en effet $H'_k(\theta)$ donné par 3b.16 s'annule au point θ_1) ce qui aurait pu être vu a priori en remarquant que la condition (3b.17) est faible, et qu'elle peut être remplacée par :

(3b.17') $$k' - \theta > |k - \theta| \text{ pour tout } \theta < \theta_1$$

- Pour une distribution autre que normale, la détermination des seuils critiques est moins simple. En particulier, pour une distribution discrète, il faut déterminer de plus les constantes de randomisation du test. Pour cela, on procédera généralement par itérations successives [Réf. 2 p. 90].

- Le problème II' $\begin{cases} H_o : \theta_o \leqslant \theta \leqslant \theta'_o \\ H_1 : \theta < \theta_o \text{ ou } \theta > \theta'_o \end{cases}$ ne possède pas de test u. m. p. On s'en rend facilement compte en étudiant le problème (2') à alternative simple

$$\text{II' (2')} \quad \begin{cases} H_o : \theta_o \leqslant \theta \leqslant \theta'_o \\ h_1 : \theta = \theta_1 \quad (\theta_1 < \theta_o \text{ ou } \theta_1 > \theta'_o) \end{cases}$$

et en remarquant que l'hypothèse la moins favorable h_o, dépend de la position de θ_1. On peut cependant trouver pour (II') un test u. m. p. u(u. m. p. dans la classe des tests sans biais, [Réf. 2 p. 126] et [Réf. 5 p. 215]) .

- L'exemple 5 du § 4 généralisera l'étude de l'exemple 2 dans le cas d'existence d'un paramètre importun, en l'occurence l'écart-type inconnu de la distribution.

4 - PROBLEMES DE TESTS A DEUX PARAMETRES

L'étude des tests à deux paramètres sera abordé à l'aide de trois exemples préliminaires.

a) Exemple 3

Soit n observations $x_1, \ldots, x_n$ indépendantes issues de $\mathcal{N}(\xi, \sigma)$. On se pose le problème du type III (H_o composite contre h_1 simple)

$$\text{(III)} \quad \begin{cases} H_o : \sigma \geqslant \sigma_o \quad \xi \in R \text{ (non spécifié)} \\ h_1 : \xi = \xi \quad ; \sigma = \sigma_1 \; (\sigma_1 < \sigma_o) \end{cases}$$

que l'on résoudra en passant par l'étape du problème (3)

$$(3)\quad \begin{cases} h_o : \text{la moins favorable} \\ h_1 : \xi = \xi_1 \; ; \; \sigma = \sigma_1 \quad (\sigma_1 < \sigma_o) \end{cases}$$

On peut raisonnablement penser que l'hypothèse h_o sera donnée par $h_o : \sigma = \sigma_o$; $\xi = \xi_1$. On se propose donc de conduire le test (3)

$$(3)\quad \begin{cases} h_o : \sigma = \sigma_o \; ; \; \xi = \xi_1 \\ h_1 : \sigma = \sigma_1 \; ; \; \xi = \xi_1 \; (\sigma_1 < \sigma_o) \end{cases}$$

et de vérifier a posteriori que h_o est effectivement la moins favorable pour (III). Pour le problème (3), le calcul du rapport de vraisemblance conduit à :

$$(4a.1) \qquad W(h_o) : \frac{(2\pi\sigma_o^2)^{-n/2} \exp\left[-\frac{1}{2\sigma_o^2}\sum_{i=1}^{n}(x_i - \xi_1)^2\right]}{(2\pi\sigma_1^2)^{-n/2} \exp\left[-\frac{1}{2\sigma_1^2}\sum_{i=1}^{n}(x_i - \xi_1)^2\right]} \leqslant k$$

soit de façon équivalente :

$$(4a.1') \qquad W(h_o) : \Sigma (x_i - \xi_1)^2 \leqslant k \quad \text{et}$$

$$(4a.2) \qquad W_\alpha(h_o) : \Sigma (x_i - \xi_1)^2 \leqslant \sigma_o^2 \, \chi^2_{n,\alpha}$$

ou $\chi^2_{n,\alpha}$ représente la valeur du χ^2 à n degrés de liberté, avec la probabilité $1 - \alpha$ d'être dépassée.

La puissance du test s'écrit :

$$(4a.3) \qquad P_{\xi_1, \sigma_1}(W_\alpha(h_o)) = \Pi_{(n)}\left(\frac{\sigma_o^2}{\sigma_1^2}\chi^2_{n,\alpha}\right)$$

où $\pi_{(n)}$ (t) représente la fonction de répartition du χ^2 à n degrés de liberté.

En démontrant la relation

$$(4a.4) \qquad P_{H_o}(W_\alpha(h_o)) \leqslant \alpha$$

soit :

$$(4a.4') \qquad P_{\xi,\sigma}(W_\alpha(h_o)) \leqslant \alpha \quad \forall \xi \in R, \quad \forall \sigma \geqslant \sigma_o \qquad \text{[Réf. 2, p. 96]}$$

on en déduit que la région critique $W_\alpha(h_o)$ donnée par (4a.2) est la plus puissante non seulement pour (3) mais aussi pour (III). De plus on a :

$$(4a.5) \qquad \Pi_{(n)}\left(\frac{\sigma_o^2}{\sigma_1^2}\chi^2_{n,\alpha}\right) = \min_{h'_o} P_{\xi_1,\sigma_1}(W_\alpha(h'_o))$$

où h'_o est une hypothèse simple courante de H_o.

Cependant, la région critique $\widehat{W}_\alpha(h_o)$ dépend de ξ_1, valeur de la moyenne dans l'hypothèse h_1. C'est donc une région u. m. p pour le problème du type

$$(IV) \quad \begin{cases} H_o : \sigma \geqslant \sigma_o \; ; \; \xi \in R \\ H_1 : \sigma < \sigma_o \; ; \; \xi = \xi_1 \end{cases}$$

mais non pour le problème le plus général

$$(V) \quad \begin{cases} H_o : \sigma \geqslant \sigma_o, \; \xi \\ H_1 : \sigma < \sigma_o, \; \xi \end{cases}$$

test "composite contre composite" sur l'écart-type, avec bruit ξ.

L'exemple 4 ci-dessous présentera un cas a priori semblable, mais où l'on verra que la recherche de l'hypothèse la moins favorable sera plus délicate. On pourra y trouver cependant un test u. m. p pour le problème du type (V) de l'introduction.

b) Exemple 4

Soit n observations $x_1, \ldots, x_n$ indépendantes issues de $\mathcal{N}(\xi, \sigma)$ et le problème du type (III)

$$(III) \quad \begin{cases} H_o : \sigma \leqslant \sigma_o, \; \xi \in R \\ h_1 : \sigma = \sigma_1, \; \xi = \xi_1 \quad (\sigma_1 > \sigma_o) \end{cases}$$

Assez paradoxalement, on peut montrer que l'hypothèse h_o du problème "simple contre simple" (3') $\begin{cases} h'_o : \sigma = \sigma_o \; ; \; \xi = \xi_1 \\ h_1 : \sigma = \sigma_1 \; ; \; \xi = \xi_1 \quad (\sigma_1 > \sigma_o) \end{cases}$ ne constitue pas effectivement l'hypothèse la moins favorable.

De la même façon que dans l'exemple 3 ci-dessus, on trouve en effet

$$(4b.1) \qquad \widehat{W}_\alpha(h'_o) \; : \; \Sigma (x_i - \xi_1)^2 \geqslant \sigma_o^2 \, \chi^2_{n, 1-\alpha}$$

et l'on se propose de montrer que la condition (2.5) du § 2 n'est pas vérifiée.

Plus exactement, on démontre que $\forall \xi \in R (\xi \neq \xi_1)$, $\exists \sigma \leqslant \sigma_o$ tel que

$$(4b.2) \qquad P_{\xi, \sigma}(\widehat{W}_\alpha(h'_o)) > \alpha .$$

Considérons pour cela ξ quelconque, non égal à ξ_1 :

$$(4b.3) \qquad P_{\xi,\sigma}(\widehat{W}_\alpha(h'_o)) = P_{\xi,\sigma}\left(\frac{\Sigma (x_i - \xi_1)^2}{\sigma^2} \geqslant \frac{\sigma_o^2}{\sigma^2} \chi^2_{n, 1-\alpha}\right)$$

$$= P\left(\chi^2_n(\tau^2) \geqslant \frac{\sigma_o^2}{\sigma^2} \chi^2_{n, 1-\alpha}\right)$$

où $\chi^2_n(\tau^2)$ est un χ^2 décentré de degré de liberté n et d'excentricité :

$$(4b.4) \qquad \tau^2 = n(\xi - \xi_1)^2 > 0$$

La relation (4b.2) sera vérifiée si l'on peut trouver $\sigma \leqslant \sigma_o$ tel que

$$\frac{\sigma_o^2}{\sigma^2} \chi^2_{n,1-\alpha} < \chi_{n,1-\sigma}(\tau^2) \tag{4b.5}$$

où $\chi^2_{n,1-\alpha}(\tau^2)$ représente la valeur du $\chi^2_n(\tau^2)$ à probabilité de dépassement égale à α.

Or il est toujours possible de trouver σ^2 vérifiant :

$$\frac{\sigma_o^2 \ \chi^2_{n,1-\alpha}}{\chi^2_{n,1-\alpha}(\tau^2)} < \sigma^2 \leqslant \sigma_o^2 \tag{4b.6}$$

puisque le $\chi^2_n(\tau^2)$ étant une variable aléatoire "stochastiquement plus grande" [Réf. 4 p. 160] que χ^2_n ($\forall$ n et $\forall \tau^2 > 0$), on a toujours

$$\chi^2_{n,1-\alpha} < \chi^2_{n,1-\alpha}(\tau^2), \quad \forall \alpha, \quad \forall n, \quad \forall \tau^2 > 0 \quad \text{[Réf. 6]} \tag{4b.7}$$

La puissance de la région critique (4b.1) s'écrit d'ailleurs :

$$P_{\xi_1,\sigma_1}(W_\alpha(h_o')) = 1 - \Pi_{(n)}\left(\frac{\sigma_o^2}{\sigma_1^2}\chi^2_{n,1-\alpha}\right) \tag{4b.8}$$

et l'on peut montrer qu'il existe au moins un couple (ξ^*, σ^*) $(\xi^* \neq \xi_1)$ tel que le problème

$$(3^*) \quad \begin{cases} h_o^* : \sigma = \sigma^* \ ; \ \xi = \xi^* \ (\xi^* \neq \xi_1) \\ h_1 : \sigma = \sigma_1 \ ; \ \xi = \xi_1 \end{cases}$$

conduite à :

$$P_{\xi_1,\sigma_1}(W_\alpha(h_o^*)) < 1 - \Pi_{(n)}\left(\frac{\sigma_o^2}{\sigma_1^2}\chi^2_{n,1-\alpha}\right) \tag{4b.9}$$

La démonstration n'est pas donnée dans cet article mais on notera cependant qu'il tel couple n'existe pas si $\xi^* = \xi_1$.

Revenons au problème du type (3) $\begin{cases} h_o \text{ la moins favorable} \\ h_1 : \xi = \xi_1 \ ; \ \sigma = \sigma_1 \ (\sigma_1 > \sigma_o) \end{cases}$

et recherchons h_o dans la classe des hypothèses nulles simples : $\sigma = \sigma_o$ et ξ distribué selon λ. En se limitant à la classe des tests de niveau α, fondés sur le couple exhaustif $(\overline{X}, \sum_i (X_i - \overline{X})^2)$ [Réf. 4], on voit que λ doit être cherchée de façon que la densité de $\overline{X}$ sous h_o, soit :

$$\int_R \left(\frac{n}{2\pi\sigma_o^2}\right)^{\frac{1}{2}} \exp\left[-\frac{n}{2\sigma_o^2}(\overline{x} - \xi)^2\right] d\lambda(\xi) \tag{4b.10}$$

apparaisse "aussi proche" que possible de la densité de $\overline{X}$ sous h_1, soit :

(4b. 11) $$\left(\frac{n}{2\pi\sigma_1^2}\right)^{\frac{1}{2}} \exp\left[-\frac{n}{2\sigma_1^2}(\bar{x}-\xi_1)^2\right]$$

Or la densité (4b.11) est celle de $\mathcal{N}(\xi_1, \sigma_1/\sqrt{n})$ alors que la densité (4b.10) est celle de la somme de deux variables indépendantes, l'une selon $\mathcal{N}(0, \sigma_o/\sqrt{n})$, l'autre de fonction de répartition λ. Il suffit donc de prendre ξ selon $\mathcal{N}\left(\xi_1, \sqrt{\frac{\sigma_1^2 - \sigma_o^2}{n}}\right)$ [Réf. 3, Réf. 4 p. 83] et d'effectuer le test

(3) $$\begin{cases} h_o : \sigma = \sigma_o, \ \xi \text{ selon } \mathcal{N}\left(\xi_1, \sqrt{\frac{\sigma_1^2 - \sigma_o^2}{n}}\right) \\ h_1 : \sigma = \sigma_1 \ (> \sigma_o), \ \xi = \xi_1 \end{cases}$$

sous réserve de vérifier a posteriori que h_o est effectivement la moins favorable.

Pour le problème (3), on trouve en se limitant aux tests basés sur $(\bar{X}, \Sigma(X_i - \bar{X})^2)$

(4b. 12) $$W(h_o) : \frac{\left[\Sigma(x_i - \bar{x})^2\right]^{\frac{n-3}{2}} \exp\left[-\frac{\Sigma(x_i - \bar{x})^2}{2\sigma_o^2}\right] \int_R \exp\left[-\frac{n}{2\sigma_o^2}(\bar{x} - \xi)^2\right] d\lambda(\xi)}{\left[\Sigma(x_i - \bar{x})^2\right]^{\frac{n-3}{2}} \exp\left[-\frac{\Sigma(x_i - \bar{x})^2}{2\sigma_1^2}\right] \exp\left[-\frac{n}{2\sigma_1^2}(\bar{x} - \xi_1)^2\right]} < K$$

soit après quelques calculs

(4b. 13) $$W(h_o) : \Sigma(x_i - \bar{x})^2 \geqslant k$$

et enfin :

(4b. 14) $$W_\alpha(h_o) : \Sigma(x_i - \bar{x})^2 \geqslant \sigma_o^2 \chi^2_{n-1, 1-\alpha}$$

région critique de niveaux α la plus puissante pour (3), de puissance :

(4b. 15) $$P_{\xi_1, \sigma_1}(W_\alpha(h_o)) = 1 - \Pi_{(n)}\left(\frac{\sigma_o^2}{\sigma_1^2}\chi^2_{n-1, 1-\alpha}\right)$$

Il faut maintenant vérifier que l'on a effectivement pour $W_\alpha(h_o)$ donné par (4b. 14)

(4b. 16) $$P_{\xi, \sigma}(W_\alpha(h_o)) \leqslant \alpha \quad \forall \xi \in R, \quad \forall \sigma \leqslant \sigma_o$$

Or on a (puisque $\sigma \leqslant \sigma_o$)

$$P_{\xi, \sigma}(W_\alpha(h_o)) = P\left(\chi^2_{(n-1)} \geqslant \frac{\sigma_o^2}{\sigma_1^2}\chi^2_{n-1, 1-\alpha}\right) \leqslant P(\chi^2_{(n-1)} \geqslant \chi^2_{n-1, 1-\alpha}) = \alpha$$

On pourra aussi noter, puisque h_o est effectivement la moins favorable que $P_{\xi_1, \sigma_1}(W_\alpha(h_o))$ donnée par (4 b. 15) est inféfieure à $P_{\xi_1, \sigma_1}(W_\alpha(h_o'))$ donnée par (4b. 8), soit :

$$(4b.17) \qquad 1 - \Pi_{(n-1)} \left(\frac{\sigma_o^2}{\sigma_1^2} \chi^2_{n-1, 1-\alpha}\right) < 1 - \Pi_{(n)} \left(\frac{\sigma_o^2}{\sigma_1^2} \chi^2_{n, 1-\alpha}\right)$$

La région critique définie par (4b. 14) est non seulement la plus puissante pour le problème III, mais aussi u. m. p pour le problème le plus général

$$(V) \quad \begin{cases} H_o : \sigma \leqslant \sigma_o, & \xi \in R \\ H_1 : \sigma > \sigma_o, & \xi \in R \end{cases}$$

puisqu'elle est indépendante des valeurs ξ_1 et σ_1 dans l'alternative h_1 de (III).

c) Exemple 5

Au cours de cet exemple, on reprend l'étude du deuxième exemple du § 3 en rajoutant un paramètre importun, en l'occurence l'écart-type de la distribution. Pour plus de généralité, nous prendrons n observations $x_1, \ldots, x_n$ indépendantes, de moyenne $\bar{x}$, issues d'une loi $\mathcal{N}(\theta, \sigma)$. Considérons donc le problème du type III :

$$(III) \quad \begin{cases} H_o : \theta \leqslant \theta_o \text{ ou } \theta \geqslant \theta'_o : \sigma \text{ inconnu} \\ h_1 : \theta = \theta_1 = (\theta_o + \theta'_o)/2 \ ; \ \sigma = \sigma_1 \end{cases}$$

qui sera conduit de façon très semblable au problème (I) de l'exemple 2. Nous en donnerons les différentes étapes sous forme de relations facilement compréhensibles.

Considérons le problème (3) $\begin{cases} h_o : \theta = \theta_o \text{ avec probabilité } 1/2 \text{ et } \theta'_o \\ \qquad \text{avec probabilité } 1/2 \ ; \sigma = \sigma_1 \\ h_1 : \theta = \theta_1 = \dfrac{\theta_o + \theta'_o}{2} \ ; \ \sigma = \sigma_1 \end{cases}$

et voyons si h_o est effectivement la moins favorable pour (III).

Pour le problème (3) :

$$(4c.1) \qquad W(h_o) : g_{\theta_o, \theta'_o, \theta_1, \sigma_1} (\bar{x}) < K$$

avec :

$$(4c.2) \qquad g_{\theta_o, \theta'_o, \theta_1, \sigma_1} = \exp\left[\frac{n}{2\sigma_1^2} (\theta_o - \theta_1)\ (2\bar{x} - \theta_1 - \theta_o)\right] + \exp\left[\frac{n}{2\sigma_1^2} (\theta'_o - \theta_1)\ (2\bar{x} - \theta_1 - \theta'_o)\right]$$

On a :

$$(4c.3) \qquad g_{\theta_o, \theta'_o, \theta_1, \sigma_1} (\bar{x}) = g_{\theta_o, \theta'_o, \theta_1, \sigma_1} (2\theta_1 - \bar{x}) \quad \forall \bar{x}$$

D'où :

$$(4c.4) \qquad W(h_o) : k < \bar{x} < k' \quad \text{avec } k + k' = 2\theta_1 = \theta_o + \theta'_o \quad \text{et}$$

(4c. 5) $W_\alpha(h_o)$: $k < \bar{x} < k'$ avec $k + k' = 2\theta_1$ et

$$F\left(\frac{k' - \theta_o}{\sigma_1/\sqrt{n}}\right) - F\left(\frac{k - \theta_o}{\sigma_1/\sqrt{n}}\right) + F\left(\frac{k' - \theta'_o}{\sigma_1/\sqrt{n}}\right) - F\left(\frac{k - \theta'_o}{\sigma_1/\sqrt{n}}\right) = 2\alpha$$

soit encore de façon équivalente (cf. 3b. 8)

(4c. 6) $W_\alpha(h_o)$: $k < \bar{x} < k'$ avec

$$F\left(\frac{k' - \theta_o}{\sigma_1/\sqrt{n}}\right) - F\left(\frac{k - \theta_o}{\sigma_1/\sqrt{n}}\right) = F\left(\frac{k' - \theta'_o}{\sigma_1/\sqrt{n}}\right) - F\left(\frac{k - \theta'_o}{\sigma_1/\sqrt{n}}\right) = \alpha$$

Considérons le domaine ω définissant H_o dans (III).

(4c. 7) $$\omega = (]-\infty, \theta_o] \cup [\theta', \infty[) \times R^+$$

où R^+ représente l'ensemble des réels positifs et X le produit cartésien classique. Le sous-ensemble ω^* de ω constitué par les deux couples (θ_o, σ_1) et (θ'_o, σ_1) possède bien entendu une λ mesure unitaire (λ désignant la distribution correspondant à h_o de (3)). D'après (4c. 6) et le théorème 1' du § 2, on en déduit par conséquent que $W_\alpha(h_o)$ est la plus puissante pour le problème (III).

Les seuils critiques k et k' sont indépendants de θ_1 mais pas de σ_1, de sorte que $W_\alpha(h_o)$ donnée par (4c. 6) constitue un test u. m. p pour le problème du type

$$\text{(IV)} \begin{cases} H_o : \theta \leqslant \theta_o \text{ ou } \theta \geqslant \theta'_o \ ; \ \sigma \in R^+ \\ H_1 : \theta_o < \theta < \theta'_o \ ; \ \sigma = \sigma_1 \end{cases}$$

mais non pour le problème plus général

$$\text{(V)} \begin{cases} H_o : \theta \leqslant \theta_o \text{ ou } \theta \geqslant \theta'_o \ ; \ \sigma \in R^+ \\ H_1 : \theta_o < \theta < \theta'_o : \ \sigma \in R^+ \end{cases}$$

Pour ce dernier problème et plus généralement pour toute distribution appartenant à une famille exponentielle à plusieurs paramètres, il existe cependant un test u. m. p. u. [Réf. 2, 5].

d) Statistiques exhaustives en présence d'un bruit : définition, utilité et applications

Dans ce paragraphe, on se propose d'introduire quelques notions permettant dans certains cas (assez rares en pratique cependant) de trouver un test u. m. p pour le problème V. Après avoir défini l'exhaustivité en présence d'un bruit, nous verrons dans le théorème 2, comment l'existence d'une telle statistique rend immédiate l'obtention d'une hypothèse h_o moins favorable pour le problème de type III de l'introduction. Les théorèmes 3 et 4 enfin donneront des conditions suffisantes d'existence d'un test u. m. p pour le problème du type (V).

Statistique exhaustive pour un paramètre θ en présence d'un bruit η

Soit une classe de distributions $P_{\theta\eta}$ à deux paramètres $\theta \in \Omega$, $\eta \in H$ et $\vec{x}$ un vecteur observations issu de $P_{\theta\eta}$. La statistique $T(\vec{x})$ est dite exhaustive pour θ en présence de η si :

(4d.1)
- sa distribution marginale est indépendante de η (probabilités transportées par T de la forme P_θ^T, $\theta \in \Omega$)
- et si la distribution de $\vec{x}$ conditionnée par T est indépendante de θ (forme $P_\eta(\vec{x}/T = t)$, $\eta \in H$)

Dans le cas où il n'existe pas de paramètre importun η, on remarquera que la définition ci-dessus se réduit à celle de l'exhaustivité classique pour θ.

Les conditions données par (4d.1) ne sont pas souvent satisfaites en pratique ; ainsi dans le cas d'une loi $\mathcal{N}(m, \sigma)$ remarquera-t-on :

- que ΣX_i n'est pas exhaustive pour m en présence de σ (première condition de 4d.1 non satisfaite)

- que $\Sigma(X_i - \bar{X})^2$ est pas exhaustive pour σ^2 en présence de m (première condition de 4d.1 non satisfaite)

On peut cependant parfois définir une transformation sur les paramètres du problème de façon à faire apparaître une statistique exhaustive pour l'un des paramètres transformés, en présence d'un bruit (cf. exemple 7 du § 4.e ci-dessous).

Théorème 2

Considérons le problème III $\begin{cases} H_0 : \theta \in \omega \quad \eta \in H \\ h_1 : \theta = \theta_1 \ (\theta \notin \omega) \ ; \eta = \eta_1 \end{cases}$ et supposons qu'il existe une statistique T exhaustive pour θ en présence de η.

L'hypothèse h_0 la moins favorable pour III est alors donnée par :

$$\text{(4d.2)} \qquad h_0 \begin{cases} \theta = \text{le moins favorable pour le problème (I)} \\ \eta = \eta_1 \end{cases} \begin{cases} H_0 : \theta \in \omega \\ h_1 : \theta = \theta_1 \\ \quad (\theta_1 \notin \omega) \end{cases}$$

Démonstration

Pour démontrer ce théorème, on se propose de conduire le test

$$\text{(3)} \quad \begin{cases} h_0 : \text{donnée par (4d.2)} \\ h_1 : \theta = \theta_1 \ (\theta_1 \notin \omega) \ ; \eta = \eta_1 \end{cases}$$

et de démontrer que la région critique optimale obtenue, soit $W_\alpha(h_0)$ est telle que :

(4d.3) $\quad P_{\theta\eta}(W_\alpha(h_0)) \leqslant \alpha \quad \forall\theta \in \omega, \quad \forall\eta$ (cf. théorème 1 et condition 2.5)

On peut écrire la région critique $W(h_0)$ pour le problème (3) sous la forme :

(4d. 4)
$$\frac{\int_{\omega} P_{\theta\eta_1}(\vec{x})\, d\lambda(\theta)}{P_{\theta_1\eta_1}(\vec{x})} < k$$

où λ définit l'hypothèse h_o la moins favorable pour (I). La relation (4d. 4) s'écrit encore, grâce à (4d. 1) :

(4d. 5)
$$\frac{\int_{\omega\times\tau} P_{\eta_1}(\vec{x}/t)\, P_{\theta}^{T}(t)\, d\lambda(\theta)}{\int_{\tau} P_{\eta_1}(\vec{x}/T = t)\, P_{\theta_1}^{T}(t)} < k$$

où τ est le transformé de l'espace $\mathscr{X}$ des observations, par la statistique T.

Pour la résolution des problèmes du type V et donc en particulier du type (3), le théorème de restriction à la classe des tests fondés sur une statistique T exhaustive pour θ en présence de η [Réf. 7], permet d'écrire :

(4d. 6)
$$W_{\alpha}(h_o) \ : \ T \in \mathscr{C}^*_{(\alpha)}$$

où $\mathscr{C}^*_{\alpha}$ est un sous-ensemble de $\mathscr{C}$ tel que :

(4d. 7)
$$\int_{\omega} P_{\theta}^{T}(\mathscr{C}^*_{(\alpha)})\, d\lambda(\theta) = \alpha$$

Il faut maintenant vérifier la condition (4d. 3), qui grâce à (4d. 1) s'écrit :

(4d. 8)
$$P_{\theta}^{T}(\mathscr{C}^*_{(\alpha)}) \leqslant \alpha \quad \forall\, \theta \in \omega$$

Or cette condition est une conséquence immédiate de (4d. 7) puisque λ est la distribution la moins favorable pour le problème (I) portant sur un paramètre unique θ.

Ce théorème montre que dans le cas favorable d'existence de T, statistique exhaustive pour θ en présence de η, l'hypothèse h_o la moins favorable pour (3) est obtenue en fixant le bruit à son niveau dans l'alternative et en choisissant indépendamment la moins favorable pour le problème à un paramètre :

$$\text{(I)} \begin{cases} H_o \ : \theta \in \omega \\ h_1 \ : \theta = \theta_1 \ (\theta_1 \notin \omega). \end{cases}$$

Il n'y a donc pas lieu comme dans l'exemple 4 du § 4b) de définir une distribution continue la moins favorable sur H. Il faut cependant remarquer que le théorème 2 donne une condition seulement suffisante pour qu'une telle situation se présente. Les exemples 3 et 5 montrent en effet qu'il n'est pas nécessaire pour cela qu'une statistique du type T ci-dessus existe.

Théorème 3

Considérons le problème (non mentionné dans l'introduction) que nous noterons :

$$(\text{III-V}) \quad \begin{cases} H_o : \theta \in \omega & \eta \in H \\ H_1 : \theta = \theta_1 & \eta \in H \quad (\theta_1 \notin \omega) \end{cases}$$

S'il existe T exhaustive pour θ en présence de η et si le problème

$$(\text{I}) \quad \begin{cases} H_o : \theta \in \omega \\ h_1 : \theta = \theta_1 \ (\theta_1 \notin \omega) \end{cases}$$

admet une hypothèse h_o la moins favorable, il existe un test u. m. p pour le problème III-V.

Démonstration

Considérons le problème (III) $\begin{cases} H_o : \theta \in \omega & \eta \in H \\ h_1 : \theta = \theta_1 & \eta = \eta_1 \quad (\theta_1 \notin \omega) \end{cases}$ obtenu en fixant dans III-V le bruit de H_1 au niveau η_1. D'après le théorème 2 ci-dessus, le problème (3) à h_o simple la moins favorable pour III, s'écrit :

$$(3) \quad \begin{cases} h_o : \theta = \text{le moins favorable pour I} \ ; \ \eta = \eta_1 \\ h_1 : \theta = \theta_1 \ (\theta_1 \notin \omega) \ ; \ \eta = \eta_1 \end{cases}$$

et l'on sait d'après les hypothèses qu'il existe une hypothèse la moins favorable pour I. Soit $W_\alpha^T(h_o)$ la région critique la plus puissante pour le problème simple contre simple :

$$(1) \quad \begin{cases} h_o : g_\lambda(x) = \int_\omega g_\theta(x) \, d\lambda(\theta) \\ h_1 : \theta = \theta_1 \quad (\theta_1 \notin \omega) \end{cases}$$

fondée sur la statistique T exhaustive pour θ. Nous montrerons successivement que $W_\alpha^T(h_o)$ est la plus puissante pour (3) et pour (III) et qu'elle est enfin u. m. p pour III-V.

. $W_\alpha^T(h_o)$ est la plus puissante au niveau α pour (3) ; en effet :

$$(4d.9) \qquad P_{\lambda\eta_1}(W_\alpha^T(h_o)) = P_\lambda^T(W_\alpha^T(h_o)) \leqslant \alpha$$

(d'après la première condition de 4d.1).

$$(4d.10) \qquad P_{\theta_1\eta_1}(W_\alpha^T(h_o)) = P_{\theta_1}^T(W_\alpha^T(h_o)) \geqslant P_{\theta_1,\eta_1}^X(W_\alpha^X(h_o)) \text{ où } W_\alpha^X(h_o)$$

est une région critique quelconque de niveau α pour (3). Cette dernière propriété (4d.10) résulte du théorème de restriction de la classe des tests à ceux fondés sur une statistique exhaustive [Réf. 4, p. 84-85].

. $W_\alpha^T(h_o)$ est la plus puissante pour III ; or cela est évident puisque h_o de (3) constitue la moins favorable pour (III) (cf. théorème 1) :

$$(4d.11) \qquad P_{\theta\eta}(W_\alpha^T(h_o)) = P_\theta^T(W_\alpha^T(h_o)) \leqslant \alpha \quad \forall \theta \in \omega$$

. Enfin $W_\alpha^T(h_o)$ étant indépendante de η_1, elle constitue ainsi une région u. m. p pour (III-V).

Remarquons que si $W_\alpha^T(h_o)$ est de plus indépendante de θ_1, on a ainsi défini un test u. m. p pour

$$(V) \quad \begin{cases} H_o : \theta \in \omega, \ \eta \\ H_1 : \theta \in \Omega - \omega. \ \eta \end{cases}$$

Dans ce cas $W_\alpha^T(h_o)$ n'est pas seulement la plus puissante pour (1) et (I), mais aussi u. m. p pour

$$(II) \quad \begin{cases} H_o : \theta \in \omega \\ H_1 : \theta \in \Omega - \omega. \end{cases}$$

D'où le théorème 4 ci-dessous.

Théorème 4

S'il existe un test u. m. p pour II et une statistique T exhaustive pour θ en présence de η, on a aussi un test u. m. p pour

$$(V) \quad \begin{cases} H_o : \theta \in \omega, \ \eta \\ H_1 : \theta \in \Omega - \omega, \ \eta \end{cases}$$

Il ne s'agit encore là aussi que de conditions suffisantes, puisque l'on a trouvé dans le cadre de l'exemple 4 un test u. m. p pour le problème du type

$$(V) \quad \begin{cases} H_o : \sigma \leqslant \sigma_o, \ \xi \\ H \ : \sigma > \sigma_o, \ \xi, \end{cases}$$

cela sans qu'il existe une statistique S exhaustive pour σ en présence de ξ.

Les exemples d'application des théorèmes 3 et 4 ne sont pas communs dans le cadre des distributions classiques. L'exemple 6 ci-dessous est emprunté aux problèmes de contrôle de qualité, le bruit η étant constitué par un couple de distributions conditionnelles inconnues.

Exemple 6

Soit X une caractéristique physique aléatoire d'une pièce fabriquée, considérée comme défectueuse si $X \leqslant v$ où v est connue. Soit p la probabilité que la pièce soit mauvaise :

> $p = P(X \leqslant v)$ et proposons-nous à l'aide de n observations indépendantes $x_1, x_2, \ldots, x_n$ de conduire le test :

$$(II) \quad \begin{cases} H_o : p \geqslant P_o \\ H_1 : p < P_o \end{cases},$$

cela sans aucune connaissance sur la distribution de X.

Plus exactement, le problème II est donc un problème du type V :

$$(V) \quad \begin{cases} H_o : p \geqslant P_o, & g \in G \\ H_1 : p < P_o, & g \in G \end{cases}$$

avec, pour simplifier, un bruit g qui est une densité continue de probabilité et G l'espace des densités correspondant.

Le bruit g peut être remplacé par un couple de bruits (g_+, g_-) représentant les densités liées par $X > v$ et $X \leqslant v$ respectivement et l'on aboutit donc au problème non paramétrique :

$$(V) \quad \begin{cases} H_o : p \geqslant P_o, & g_+, \ g_- \\ H_1 : p < P_o, & g_+, \ g_- \end{cases}$$

Soit M le nombre aléatoire d'observations x_i inférieures ou égales à v. On voit facilement que M est une statistique exhaustive pour p en présence de (g_+, g_-) :

(4d. 12)
- M suit la distribution binomiale de paramètres p et n : $\mathcal{B}(p, n)$
- la probabilité des observations $x_1, \ldots, x_n$, liée par $M = m$, ne dépend que de g_+ et g_-.

Par ailleurs, le bruit g étant fixé à la même "valeur" g_o dans V, on sait que le problème

$$(II) \quad \begin{cases} H_o : p \geqslant P_o \\ H_1 : p < P_o \end{cases}$$

densité de g_o connue, admet un test u.m.p (prendre $h_1 : p = p_1 < P_o$, traiter

$$(I) \quad \begin{cases} H_o : p \geqslant P_o \\ h_1 : p = p_1 < P_o \end{cases}$$

par la méthode de l'hypothèque la moins favorable $h_o : p = P_o$ et remarquer que $W_\alpha(h_o)$ pour :

$$(1) \quad \begin{cases} h_o : p = P_o \\ h_1 : p = p_1 < P_o \end{cases}$$

est indépendante de p_1). D'après le théorème 4, on a donc aussi un test u. m. p pour le problème V ci-dessus :

$$(4d.13) \qquad W_\alpha(H_o) \quad \begin{cases} M < C, \text{ avec :} \\ P_{p_o}(M < C) + \gamma \, P_{p_o}(M = C) = \alpha \end{cases}$$

relation permettant de trouver C et la constante γ de randomisation du test .

Pour l'aspect "test des signes" du problème V, exemple classique de la théorie non paramétrique, on se reportera à [Réf. 4 et 8].

Nous considérerons enfin l'exemple 7, fait de deux problèmes successifs de tests.

e) Exemple 7

Soit $x_1, x_2, \ldots, x_m$ m observations indépendantes issues de $\mathcal{N}(\xi, 1)$ et
$y_1, y_2, \ldots, y_n$ n " " " $\mathcal{N}(\eta, 1)$.

On suppose que les variables X et Y sont indépendantes et l'on se propose d'étudier le problème du type

$$\text{(II)} \quad \begin{cases} H_0 : \eta \leqslant \xi \\ H_1 : \eta > \xi \end{cases}$$

La procédure consiste à commencer par étudier le problème (I) à alternative simple :

$$\text{(I)} \quad \begin{cases} H_0 : \eta \leqslant \xi \\ h_1 : \eta = \eta_1 \,;\, \xi = \xi_1 \;\; (\xi_1 < \eta_1), \end{cases}$$

cela par la méthode de l'hypothèse h_0 la moins favorable.

Le problème (1) s'écrit alors :

$$\text{(1)} \quad \begin{cases} h_0 : \eta = \xi = \tau \\ h_1 : \eta = \eta_1 \;;\; \xi = \xi_1 (\xi_1 < \eta_1) \end{cases}$$

où τ est la valeur commune inconnue de η et ξ, telle que h_0 soit effectivement la moins favorable.

En se limitant à la classe des tests fondés sur le couple exhaustif $(\bar{x}, \bar{y})$, la méthode classique de Neyman et Pearson dans le cadre d'un test "simple contre simple" conduit à la région critique :

$$\text{(4e. 1)} \qquad g^X_{h_0}(\bar{x})\, g^Y_{h_0}(\bar{y}) \,/\, g^X_{h_1}(\bar{x})\, g^Y_{h_1}(\bar{y}) < K$$

où $g^x_{h_0}(\bar{x})$ par exemple représente la densité de probabilité de $\bar{x}$ dans l'hypothèse h . On trouve facilement :

$$\text{(4e. 2)} \qquad W(h_0) \;:\; m\bar{x}(\tau - \xi_1) + n\bar{y}(\tau - \eta_1) < k$$

La constante critique est déterminée par la condition de niveau α, ce qui donne :

$$\text{(4e. 3)} \qquad \begin{aligned} W_\alpha(h_0) \;:\; & m\bar{x}(\tau - \xi_1) + n\bar{y}(\tau - \eta_1) \\ & < u_{1-\alpha}\sqrt{m(\tau - \xi_1)^2 + n(\tau - \eta_1)^2} + \tau[m(\tau - \xi_1) + n(\tau - \eta_1)] \end{aligned}$$

$$\text{(4e. 4)} \qquad P_{\xi_1,\eta_1}(h_0) = P_{\xi_1,\eta_1}(W_\alpha(h_0)) = F\left(u_{1-\alpha} + \sqrt{m(\tau - \xi_1)^2 + n(\tau - \eta_1)^2}\right)$$

et τ, valeur commune de ξ et η dans l'hypothèse la moins favorable h_o, se calcule alors en écrivant :

(4e.5) $$\frac{\delta}{\delta\tau}\left(F_{\xi_1,\eta_1}(h_o)\right) = 0$$

On trouve facilement la solution

(4e.6) $$\tau = \frac{m\,\xi_1 + n\,\eta_1}{m + n}$$

Le problème (1) devient alors :

$$\begin{cases} h_o : \eta = \xi = \dfrac{m\xi_1 + n\eta_1}{m + n} \\ h_1 : \xi = \xi_1 \ ; \ \eta = \eta_1 \quad (\xi_1 < \eta_1) \end{cases}$$

Il conduit à la région critique :

(4e.7) $$W_\alpha(h_o) = \bar{y} - \bar{x} > u_{1-\alpha}\sqrt{\frac{1}{m} + \frac{1}{n}}$$

et à la puissance minimale

(4e.8) $$P_{\xi_1,\eta_1}(h_o) = 1 - F\left(u_{1-\alpha} - \frac{(\eta_1 - \xi_1)}{\sqrt{\frac{1}{m} + \frac{1}{n}}}\right)$$

On pourra bien entendu facilement vérifier que l'on a :

(4e.9) $$P_{\xi,\eta}(W_\alpha(h_o)) \leqslant \alpha \qquad \forall \xi \text{ et } \eta \text{ vérifiant } \eta \leqslant \xi$$

D'après le théorème 1, la région $W_\alpha(h_o)$ la plus puissante pour (1), l'est aussi pour (I) ; en outre, elle est indépendante de l'alternative simple h_1 et l'on a ainsi trouvé un test u. m. p pour le problème initial (II).

Le théorème 4 peut d'ailleurs aider à résoudre rapidement ce problème ; le test II s'écrit en effet :

$$\text{(II)} \quad \begin{cases} H_o : \theta \in \omega \\ H_1 : \theta \in \Omega - \omega \end{cases}$$

avec :

(4c.10) $$\begin{cases} \theta = (\xi, \eta) \\ \omega = \{(\xi, \eta) : \xi \geqslant \eta\} \\ \Omega - \omega = \{(\xi, \eta) \ ; \ \xi < \eta\} \end{cases}$$

et (II) peut se "transformer" en un problème du type V :

$$(\mathrm{V}) \quad \begin{cases} H_0 : \zeta \leqslant 0, & \xi \in R \\ H_1 : \zeta > 0, & \xi \in R \end{cases}$$

où :

(4e. 11) $$\zeta = \eta - \xi$$

La statistique $\overline{Y} - \overline{X}$ étant exhaustive pour ζ [Réf. 2 p. 171], elle est aussi exhaustive pour ζ en présence de ξ, puisque sa distribution

$$\mathfrak{N}\left(\zeta, \sqrt{\frac{1}{m} + \frac{1}{n}}\right)$$

ne dépend que de ζ. Par ailleurs, on sait qu'il existe un test u. m. p pour $\zeta \leqslant 0$ contre $\zeta > 0$ (cf. exemple 1). Ainsi d'après le théorème 4, existe-t-il aussi un test u. m. p pour (V).

On peut maintenant se poser la question de savoir ce qui se passe lorsqu'on rajoute au problème II ci-dessus un paramètre importun, en l'occurence, l'écart-type σ supposé le même pour X et Y ($\sigma_X = \sigma_Y = \sigma$). Soit donc le problème (V) :

$$(\mathrm{V}) \quad \begin{cases} H_0 : \eta \leqslant \xi, & \sigma_X = \sigma_Y = \sigma \in R^+ \\ H_1 : \eta > \xi, & \sigma_X = \sigma_Y = \sigma \in R^+ \end{cases}.$$

Le problème (III) de l'introduction correspondant à (V) s'écrit :

$$(\mathrm{III}) \quad \begin{cases} H_0 : \eta \leqslant \xi \ ; \sigma_X = \sigma_Y = \sigma \in R^+ \\ h_1 : \xi = \xi_1 \ ; \eta = \eta_1 \ ; \sigma_X = \sigma_Y = \sigma_1 \quad (\eta_1 > \xi_1) \end{cases}$$

problème que l'on se propose de résoudre en étudiant le test "simple contre simple"

$$(3) \quad \begin{cases} h_o : \xi = \eta = \dfrac{m\xi_1 + n\eta_1}{m + n} \ ; \sigma_X = \sigma_Y = \sigma_1 \\ h_1 : \xi = \xi_1 \ ; \eta = \eta_1 \ ; \sigma_X = \sigma_Y = \sigma_1 \quad (\eta_1 > \xi_1) \end{cases}$$

(cf. premier test de l'exemple).

Après quelques calculs élémentaires, le rapport de vraisemblance $\frac{L_o}{L_1}$ du problème (3) :

(4e. 12)
$$\frac{L_o}{L_1} = \frac{\exp\left\{-\dfrac{1}{2\sigma_1^2}\left[\sum\limits_{i=1}^{m}\left(x_i - \dfrac{m\xi_1 + n\eta_1}{m+n}\right)^2 + \sum\limits_{j=1}^{n}\left(y_j - \dfrac{m\xi_1 + n\eta_1}{m+n}\right)^2\right]\right\}}{\exp\left\{-\dfrac{1}{2\sigma_1^2}\left[\sum\limits_i (x_i - \xi_1)^2 + \sum\limits_j (y_j - \eta_1)^2\right]\right\}}$$

conduit à la région critique :

(4e.13) $$W_\alpha(h_o) : \bar{y} - \bar{x} > \sigma_1 u_{1-\alpha}\sqrt{\frac{1}{m} + \frac{1}{n}},$$

de puissance

(4e.14) $$P_{\xi_1,\eta_1,\sigma_1}(W_\alpha(h_o)) = 1 - F\left(u_{1-\alpha} - \frac{(\eta_1 - \xi_1)}{\sigma_1\sqrt{\frac{1}{m} + \frac{1}{n}}}\right) \quad (\eta_1 > \xi_1)$$

Le niveau de la région critique (4e.13) pour une hypothèse simple courante de l'hypothèse nulle H_o de III s'écrit :

(4e.15) $$P_{\xi,\eta,\sigma}(W_\alpha(h_o)) = 1 - F\left(\frac{\sigma_1}{\sigma} u_{1-\alpha} - \frac{(\eta - \xi)}{\sigma\sqrt{\frac{1}{m} + \frac{1}{n}}}\right) \quad (\eta \leqslant \xi)$$

et n'est inférieure ou égale à α que si :

(4e.16) $$\sigma \leqslant \sigma_1 - \frac{(\eta - \xi)}{u_{1-\alpha}\sqrt{\frac{1}{m} + \frac{1}{n}}}$$

à condition aussi que α soit inférieur à 0,5 ce qui est conforme aux risques de première espèce habituellement choisis. Le paramètre η étant inférieur ou égal à ξ dans l'hypothèse H_o de III, la condition (4e.16) est en particulier vérifiée si σ ne dépasse pas σ_1. Il apparaît ainsi que la région critique (4e.13), la plus puissante pour (3) ne l'est pas pour le problème III, mais pour le problème III'

$$(\text{III}') \begin{cases} H : \eta \leqslant \xi \ ; \sigma_x = \sigma_y = \sigma \in R^+ \quad (\sigma \leqslant \sigma_1) \\ h_1 : \xi = \xi_1 \ ; \eta = \eta_1 \ ; \sigma_x = \sigma_y = \sigma_1 \end{cases}$$

Par ailleurs, $W_\alpha(h_o)$ donnée par (4e.13) dépend de σ_1, de telle sorte qu'elle ne constitue ainsi qu'un test u. m. p pour le problème (IV') (cf. § 1), si $\alpha \leqslant 0,5$

$$(\text{IV}') \begin{cases} H_o : \eta \leqslant \xi \ ; \sigma_x = \sigma_y = \sigma \in R^+ \quad (\sigma \leqslant \sigma_1) \\ H_1 : \eta > \xi \ ; \sigma_x = \sigma_y = \sigma_1 \end{cases}$$

mais non pour le problème (V)

$$(\text{V}) \begin{cases} H_o : \eta \leqslant \xi \ ; \sigma_x = \sigma_y = \sigma \in R^+ \\ H_1 : \eta > \xi \ ; \sigma_x = \sigma_y = \sigma \in R^+ \end{cases}$$

Le problème du type V, dans le cas où les bruits σ_x et σ_y sont différents, n'est pas étudié dans le cadre de cet article.

REFERENCES

[1] LEHMANN E. L. - On the existence of least favorable distributions (Ann. Math. Stat. Vol 23, 1952, pp. 408-416).

[2] LEHMANN E. L. - Testing statistical hypotheses (Juillet 1966, John Wiley and Sons).

[3] LEHMANN E.L. and STEIN C. - Most powerful tests of composite hypotheses (Ann. Math. Stat. Vol 19, pp. 495-516).

[4] FRASER D.A.S. - Non parametric methods in Statistics (Juillet 1959, John Wiley and Sons)

[5] FERGUSON T.S. - Mathematical Statistics. A decision theoretic approach (1967, Academic Press).

[6] GALOT G. - Les échantillons bernoulliens de variables normales (cours polycopié à l'E.N.S.A.E.).

[7] FRASER D.A.S. - Sufficient statistics with nuisance parameters (Ann. Math. Stat. Vol 27, 1956, pp. 838-842).

[8] VAN der WAERDEN B.L. - Statistique mathématique (Dunod, 1967).