

Chapitre 2

Théorie de la Détection Statistique

1 Introduction

Dans notre vie de tous les jours, nous sommes toujours amenés à prendre des décisions. Étant donné certaines hypothèses, un critère de sélection est choisi, à partir duquel une décision est prise. En ingénierie, par exemple, nous sommes souvent confrontés à un problème de détection d'un signal radar pour savoir si une cible est présente ou absente. Auquel cas, les échos sont observés et une décision est prise quant à la présence ou pas d'une cible. Dans un système de communication digitale, une séquence de '0' et de '1' peut être transmise au moyen d'un support de communication quelconque. Au niveau du récepteur, les signaux reçus représentant les '0' et les '1', sont noyés dans un bruit additif. Le récepteur ne sait pas si le signal reçu représente un '0' ou un '1' mais doit prendre une décision quant à la présence d'un '0' ou d'un '1'. Le processus selon lequel le récepteur prend une décision relève de la théorie de la détection du signal.

Pour ce faire, un '1', est représenté par $r(t)=s_1(t) + n(t)$ et un '0', par $r(t)=s_0(t) + n(t)$ où $s_1(t) = \sin w_1 t$ et $s_0(t) = \sin w_0 t$, $0 \leq t \leq T$. Dans ce cas particulier, la seule erreur possible est associée à la présence du bruit additif. Si celui-ci n'était pas présent, nous parlerions d'un problème où l'entrée (récepteur) est complètement connue et les décisions ne seraient affectées d'aucune erreur. Ce problème est celui d'un signal connu et noyé dans du bruit. C'est le cas du détecteur le plus simple qui puisse exister (domaines d'applications: communication digitale synchrone, reconnaissance des formes, etc.)

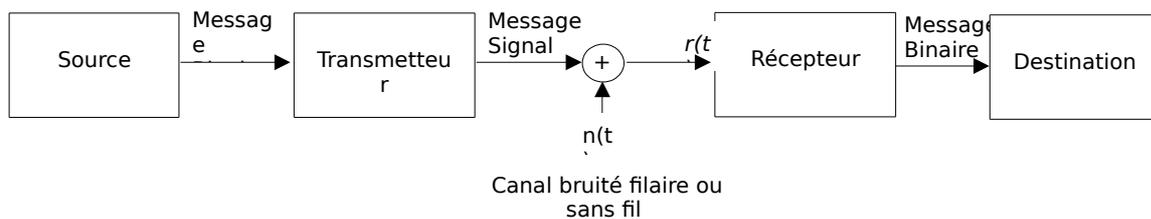


Figure 1 Système de communication digitale

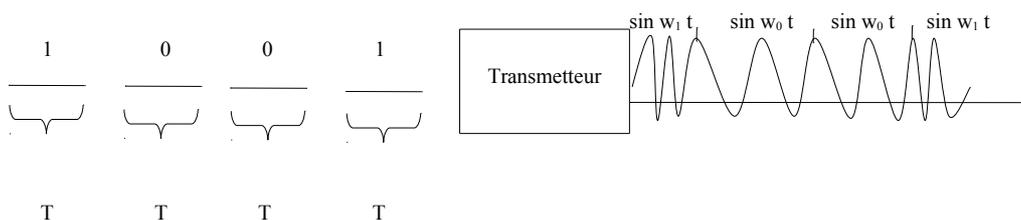


Figure 2 Séquences typiques dans un système de communication digitale

Par ailleurs, si $s_1(t) = \sin(w_1 t + \theta_1)$ et $s_0(t) = \sin(w_0 t + \theta_0)$, $0 \leq t \leq T$. θ_1 et θ_0 sont des constantes inconnues. Dans ce cas, même en l'absence du bruit, $r(t)$ n'est pas complètement connu. Si θ_1 et θ_0 varient lentement, alors on retombe sur le problème précédent.

Pour un signal radar, l'onde transmise est de la forme $s_t(t) = \sin w_c t$, $0 \leq t \leq T$. En présence d'une cible, l'onde reçue (écho) s'écrit $r(t) = V_r \sin [w_c(t-\tau) + \theta_r] + n(t)$, $\tau \leq t \leq \tau+T$ où V_r , θ_r et τ représentent l'amplitude, le déphasage et le temps d'un aller-retour (radar-cible et retour), respectivement. En l'absence d'une cible, $r(t) = n(t)$, $0 \leq t \leq \infty$. Dans ce cas, nous remarquons que même en l'absence du bruit, le signal présente trois quantités inconnues, à savoir V_r , θ_r et τ . Les deux cas précédents constituent le deuxième niveau de détection, connu sous le nom de *signal présentant des paramètres inconnus et noyé dans du bruit* (domaines d'applications: radar ou sonar à impulsions, détection de cibles, etc.).

Un troisième niveau de problème associé à la détection peut être représenté par l'exemple d'un sonar où le signal est modélisé par un signal aléatoire (signal acoustique) donné par $r(t) = s_r(t) + n(t)$ dans le cas de la présence d'une cible (s_r est une fonction aléatoire) et $r(t) = n(t)$ dans le cas contraire. En l'absence de bruit, nous remarquons que le signal est aléatoire (processus aléatoire). Dans ce cas, le troisième niveau de détection est dit *signal aléatoire et noyé dans du bruit* (domaines d'applications : sonar passif, astronomie radio, etc.).

Toutes les situations décrites précédemment peuvent être décrites par une source émettant deux sorties possibles et ce à n'importe quel instant. Comme cela est montré en Figure 3, les sorties se réfèrent à l'hypothèse nulle ou H_0 (absence de cible) et celle représentant l'hypothèse alternative ou H_1 (présence de cible). Chaque hypothèse correspond à une ou plusieurs observations qui sont représentées par des variables aléatoires. Compte tenu des valeurs observées, le récepteur décide en faveur de l'une des deux hypothèses (H_0 ou H_1). Supposons que le récepteur doit prendre une décision à la base d'une seule observation du signal reçu. L'intervalle des valeurs que prend la variable aléatoire Y constitue l'espace des observations Z . Ce dernier est partitionné en deux régions Z_0 et Z_1 , tel que si Y se trouve dans Z_0 , le récepteur décide en faveur de H_0 ; par contre, si Y se trouve dans Z_1 , le récepteur décide en faveur de H_1 (Figure 4).

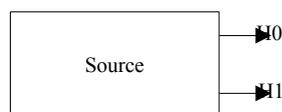


Figure 3 Source pour des hypothèses binaires

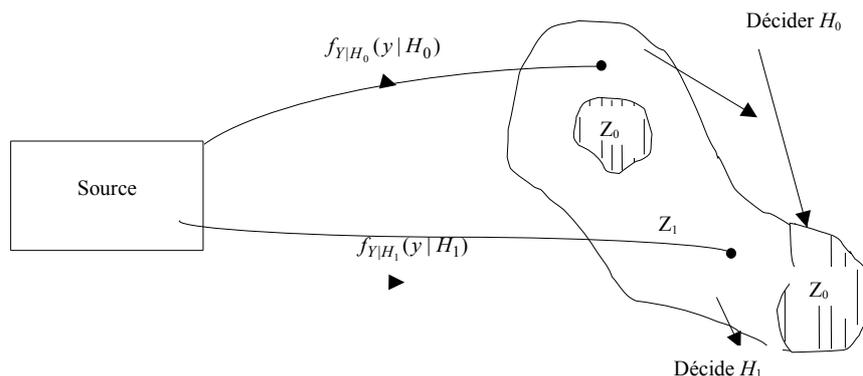


Figure 4 Régions de décision

L'espace d'observation est l'union de Z_0 et Z_1 . Autrement dit:

$$Z = Z_0 \cup Z_1 \quad (1)$$

Les fonctions de densités de probabilités de Y correspondant à chaque hypothèse sont $f_{Y|H_0}(y|H_0)$ et $f_{Y|H_1}(y|H_1)$, où y est une valeur particulière de la variable aléatoire Y .

Basé sur un critère de sélection, à ce problème de test d'hypothèses binaires correspond l'un des quatre cas possibles:

1. Décider H_0 quand H_0 est vraie.
2. Décider H_0 quand H_1 est vraie.
3. Décider H_1 quand H_0 est vraie.
4. Décider H_1 quand H_1 est vraie.

Dans les cas 1 et 4, le récepteur a pris une décision correcte. Cependant, dans les cas 2 et 3, le récepteur a pris une décision erronée.

Dans ce chapitre, nous développerons les principes de base du traitement des problèmes de décision. Les observations sont représentées par des variables aléatoires. L'extension de ces résultats peut être faite pour les processus stochastiques. Dans les sections suivantes, nous étudierons quelques critères utilisés dans la théorie de la décision et les conditions sous lesquelles ils sont utiles.

2. Critère de Bayes

L'utilisation du critère de Bayes suggère l'utilisation de deux hypothèses. Dans la première, nous supposons que les probabilités de sorties de la source sont connues. Ce sont les probabilités *a priori* $P(H_0)$ et $P(H_1)$. $P(H_0)$ est la probabilité de l'occurrence de l'hypothèse H_0 , alors que $P(H_1)$ est la probabilité de l'occurrence de l'hypothèse H_1 . En dénotant, respectivement, par P_0 et P_1 les probabilités *a priori* $P(H_0)$ and $P(H_1)$ et sachant que l'une ou l'autre hypothèse H_0 or H_1 est toujours vraie, nous pouvons écrire:

$$P_0 + P_1 = 1 \quad (2)$$

Dans la deuxième, nous supposons que nous connaissons les coûts associés à chaque décision. Si nous dénotons par D_i , $i=0,1$, où D_0 dénote "Décider H_0 " et D_1 dénote "Décider H_1 ," nous définissons C_{ij} , $i, j=0,1$, comme étant le coût associé à la décision D_i , sachant que l'hypothèse H_j est vraie. Autrement dit, $P(\text{d'encourir le coût } C_{ij}) = P(\text{décider } D_i, H_j \text{ vraie}), i, j=0,1$

Par conséquent, les coûts pour ce problème de test des hypothèses binaires sont C_{00} pour le cas (1), C_{01} pour le cas (2), C_{10} pour le cas (3), et C_{11} pour le cas (4). Le but du critère de Bayes est de déterminer la règle de décision tel que le coût moyen, $E[C]$ aussi appelé *risque*, \mathfrak{R} est minimum. $E[C]$ dénote l'opérateur "Valeur Moyenne". Nous supposons aussi que le coût d'une décision erronée est plus grand que celui d'une décision correcte; c'est-à-dire:

$$C_{01} > C_{11} \text{ et } C_{10} > C_{00} \quad (3)$$

Etant donné $P(D_i, H_j)$, la probabilité jointe de décider D_i , sachant que l'hypothèse H_j est vraie, le coût moyen s'écrit:

$$\mathfrak{R} = E[C] = C_{00}P(D_0, H_0) + C_{01}P(D_0, H_1) + C_{10}P(D_1, H_0) + C_{11}P(D_1, H_1) \quad (4)$$

A partir de la règle de Bayes, nous pouvons écrire:

$$P(D_i, H_j) = P(D_i | H_j)P(H_j) \quad i=0, 1 \text{ et } j=0, 1 \quad (5)$$

Les fonctions densité conditionnelles $P(D_i | H_j)$, $i, j=0,1$, en fonction des régions Z_0 et Z_1 sont alors:

$$P(D_0|H_0) \equiv P(\text{Décider } H_0|H_0 \text{ vraie}) = \int_{z_0} f_{y|H_0}(y|H_0) dy \quad (6)$$

$$P(D_0|H_1) \equiv P(\text{Décider } H_0|H_1 \text{ vraie}) = \int_{z_0} f_{y|H_1}(y|H_1) dy \quad (7)$$

$$P(D_1|H_0) \equiv P(\text{Décider } H_1|H_0 \text{ vraie}) = \int_{z_1} f_{y|H_0}(y|H_0) dy \quad (8)$$

et

$$P(D_1|H_1) \equiv P(\text{Décider } H_1|H_1 \text{ vraie}) = \int_{z_1} f_{y|H_1}(y|H_1) dy \quad (9)$$

En terminologie radar, les probabilités $P(D_0|H_0)$, $P(D_0|H_1)$, $P(D_1|H_0)$ et $P(D_1|H_1)$ représentent les probabilités nulle (probability of null) P_{Null} , de non détection (probability of miss) P_M , de fausse alarme (probability of false alarm) P_F , et de détection (probability of détection) P_D , respectivement. Nous remarquons aussi que:

$$P_M = 1 - P_D \quad (10)$$

et

$$P_{Null} = 1 - P_F \quad (11)$$

Par conséquent, la probabilité d'une décision correcte est:

$$\begin{aligned} P(\text{Décision correcte}) &= P(c) = P(D_0, H_0) + P(D_1, H_1) \\ &= P(D_0|H_0)P(H_0) + P(D_1|H_1)P(H_1) \\ &= (1 - P_F)P_0 + P_D P_1 \end{aligned} \quad (12)$$

et la probabilité d'une décision erronée ou probabilité d'erreur est:

$$\begin{aligned} P(\text{Décision erronée}) &= P(\varepsilon) = P(D_0, H_1) + P(D_1, H_0) \\ &= P(D_0|H_1)P(H_1) + P(D_1|H_0)P(H_0) \\ &= P_M P_1 + P_F P_0 \end{aligned} \quad (13)$$

Le coût moyen donné par (4) devient alors:

$$\mathfrak{R} = E[C] = C_{00}(1 - P_F)P_0 + C_{01}(1 - P_D)P_1 + C_{10}P_F P_0 + C_{11}P_D P_1 \quad (14)$$

En termes de régions de décision définies en (6)-(9), le coût moyen peut se mettre sous la forme:

$$\begin{aligned} \mathfrak{R} = & P_0 C_{00} \int_{Z_0} f_{Y|H_0}(y|H_0) dy + P_1 C_{01} \int_{Z_0} f_{Y|H_1}(y|H_1) dy \\ & + P_0 C_{10} \int_{Z_1} f_{Y|H_0}(y|H_0) dy + P_1 C_{11} \int_{Z_1} f_{Y|H_1}(y|H_1) dy \end{aligned} \quad (15)$$

En utilisant (1) et le fait que:

$$\int_Z f_{Y|H_0}(y|H_0) dy = \int_Z f_{Y|H_1}(y|H_1) dy = 1 \quad (16)$$

ceci implique que:

$$\int_{Z_1} f_{Y|H_j}(y|H_j) dy = 1 - \int_{Z_0} f_{Y|H_j}(y|H_j) dy, \quad j = 0, 1 \quad (17)$$

où $f_{Y|H_j}(y|H_j), j = 0, 1$, est la fonction densité de probabilité de Y correspondant à chaque hypothèse. En substituant (17) dans (15), nous obtenons:

$$\mathfrak{R} = P_0 C_{10} + P_1 C_{11} + \int_{Z_0} \{ [P_1(C_{01} - C_{11}) f_{Y|H_1}(y|H_1)] - [P_0(C_{10} - C_{00}) f_{Y|H_0}(y|H_0)] \} dy \quad (18)$$

Nous remarquons que la quantité $P_0 C_{10} + P_1 C_{11}$ est une constante positive. De (3), les termes $P_1(C_{01} - C_{11}) f_{Y|H_1}(y|H_1)$ et $P_0(C_{10} - C_{00}) f_{Y|H_0}(y|H_0)$, sont tous deux positifs.

Par conséquent, le risque est minimisé en sélectionnant la région Z_0 incluant uniquement les valeurs de Y pour lesquelles le second terme est plus grand que le premier terme. Donc, ceci donne alors un intégrande négatif. Plus précisément, nous attribuons à la région Z_0 les points pour lesquels

$$P_1(C_{01} - C_{11}) f_{Y|H_1}(y|H_1) < P_0(C_{10} - C_{00}) f_{Y|H_0}(y|H_0) \quad (19)$$

Toutes les valeurs de Y pour lesquelles le second terme est plus petit que le premier sont exclues de Z_0 et attribuées à Z_1 . Les valeurs pour lesquelles les deux termes sont égaux n'affectent pas le risque et peuvent être attribuées à l'une ou l'autre région. Par conséquent, nous disons que si:

$$P_1(C_{01} - C_{11}) f_{Y|H_1}(y|H_1) > P_0(C_{10} - C_{00}) f_{Y|H_0}(y|H_0) \quad (20)$$

alors, nous décidons H_1 . Sinon, nous décidons H_0 . Donc, la règle de décision qui résulte du critère de Bayes (Likelihood Ratio Test ou LRT) est:

$$\frac{f_{Y|H_1}(y|H_1)}{f_{Y|H_0}(y|H_0)} \underset{H_0}{\overset{H_1}{>}} \frac{P_0(C_{10}-C_{00})}{P_1(C_{01}-C_{11})} \quad (21)$$

Nous définissons le rapport de vraisemblance (Likelihood Ratio ou LR) comme:

$$\Lambda(y) = \frac{f_{Y|H_1}(y|H_1)}{f_{Y|H_0}(y|H_0)} \quad (22)$$

et le seuil (threshold) comme:

$$\eta = \frac{P_0(C_{10}-C_{00})}{P_1(C_{01}-C_{11})} \quad (23)$$

Si nous avons K observations où K désigne le nombre d'échantillons Y_1, Y_2, \dots, Y_K du signal reçu, le rapport de vraisemblance peut s'exprimer comme suit:

$$\Lambda(\mathbf{y}) = \frac{f_{\mathbf{Y}|H_1}(\mathbf{y}|H_1)}{f_{\mathbf{Y}|H_0}(\mathbf{y}|H_0)} \quad (24)$$

où \mathbf{Y} , est le vecteur reçu $\mathbf{Y}^T = [Y_1 \ Y_2 \ \dots \ Y_K]^T$.

La statistique de vraisemblance $\Lambda(\mathbf{Y})$ est une variable aléatoire car elle est fonction de la variable aléatoire \mathbf{Y} . De ce fait, le critère de Bayes qui minimise le coût moyen aboutit au test du rapport de vraisemblance suivant:

$$\Lambda(\mathbf{y}) \underset{H_0}{\overset{H_1}{>}} \eta \quad (25)$$

Une remarque importante est liée au fait que le test du rapport de vraisemblance est simplement effectué en considérant le vecteur reçu pour obtenir le rapport de vraisemblance pour qu'il puisse être comparé au seuil. Par conséquent, dans des situations réelles où les probabilités *a priori* et le coût peuvent changer, c'est le seuil qui change uniquement alors que le calcul du rapport de vraisemblance n'en est pas affecté.

Comme le logarithme naturel est une fonction monotone strictement croissante, une règle de décision équivalente à (25) est :

$$\begin{array}{c} H_1 \\ \ln \Lambda(\mathbf{y}) > \ln \eta \\ < \\ H_0 \end{array} \quad (26)$$

Si nous choisissons le coût d'une décision erronée égal à 1 et celui d'une décision correcte égal à 0, alors:

$$C_{01} = C_{10} = 1 \quad \text{et} \quad C_{00} = C_{11} = 0 \quad (27)$$

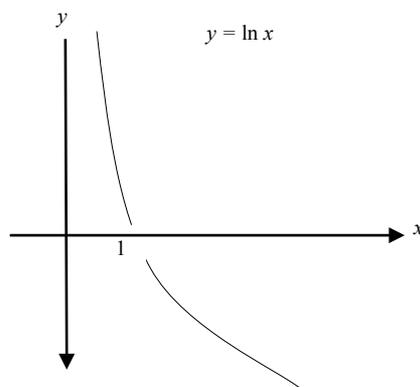


Figure 5 Fonction logarithme naturel

Le risque donné par (14) devient :

$$\mathfrak{R} = P_M P_1 + P_F P_0 = P(\varepsilon) \quad (28)$$

Autrement dit, dans ce cas, minimiser le coût moyen est équivalent à minimiser la probabilité d'erreur. De tels récepteurs sont appelés récepteurs à probabilité d'erreur minimale (minimum probability of error receivers)

Le seuil donné par (23) est réduit à:

$$\eta = \frac{P_0}{P_1} \quad (29)$$

Si, de plus, les probabilités *a priori* sont égales, η est égal à un et le logarithme du seuil devient nul.

Exemple 1 : Dans un système de communication digitale, nous considérons une source dont les sorties sont l'hypothèse H_1 correspondant à une tension constante de valeur m et l'hypothèse H_0 correspondant à une tension nulle. Le signal reçu est noyé dans un bruit additif N , blanc et Gaussien, de moyenne nulle et de variance σ^2 .

- (a) Obtenir le rapport de vraisemblance et déterminer la règle de décision ?
- (b) Calculer les probabilités de fausse alarme, P_F et de détection, P_D ?

Réponses :

(a) Le signal reçu sous chaque hypothèse est :

$$H_1 : Y = m + N$$

$$H_0 : Y = N$$

où le bruit N est Gaussien, de moyenne nulle et de variance σ^2 . Sous l'hypothèse H_0 ,

$$f_{Y|H_0}(y | H_0) = f_N(y) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{y^2}{2\sigma^2}\right)$$

Sous l'hypothèse H_1 , la moyenne de Y est $E[Y] = E[m + N] = m$, car $E[N] = 0$. La variance de Y est

$$\text{var}(Y) = \text{var}(m + N) = \text{Var}(m) + \text{Var}(N) = \sigma^2$$

Donc,

$$f_{Y|H_1}(y | H_1) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{1}{2} \frac{(y-m)^2}{\sigma^2}\right]$$

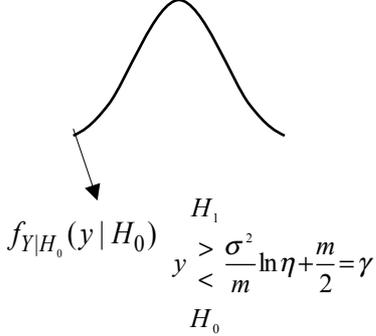
Le rapport de vraisemblance est alors:

$$\Lambda(y) = \frac{f_{Y|H_1}(y | H_1)}{f_{Y|H_0}(y | H_0)} = \exp\left(-\frac{m^2 - 2ym}{2\sigma^2}\right)$$

En prenant le logarithme naturel de chaque côté de l'équation ci-dessus, nous obtenons le test du rapport de vraisemblance suivant:

$$\ln \Lambda(y) = \frac{m}{\sigma^2} y - \frac{m^2}{2\sigma^2} \underset{H_0}{\overset{H_1}{>}} \ln \eta$$

ou, en réarrangeant les termes:



Ainsi, l'observation y reçue est comparée au seuil γ . Les régions de décisions sont montrées en Figure 6.

(b) Les probabilités de fausse alarme et de détection sont:

$$P_F = P(\text{Décider } H_1 | H_0 \text{ vraie}) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\gamma}^{\infty} e^{-\frac{y^2}{2\sigma^2}} dy = Q\left(\frac{\gamma}{\sigma}\right)$$

où

$$Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\infty} e^{-\frac{u^2}{2}} du$$

$$P_D = P(\text{Décider } H_1 | H_1 \text{ vraie}) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\gamma}^{\infty} e^{-\frac{(y-m)^2}{2\sigma^2}} dy = Q\left(\frac{\gamma-m}{\sigma}\right)$$

et

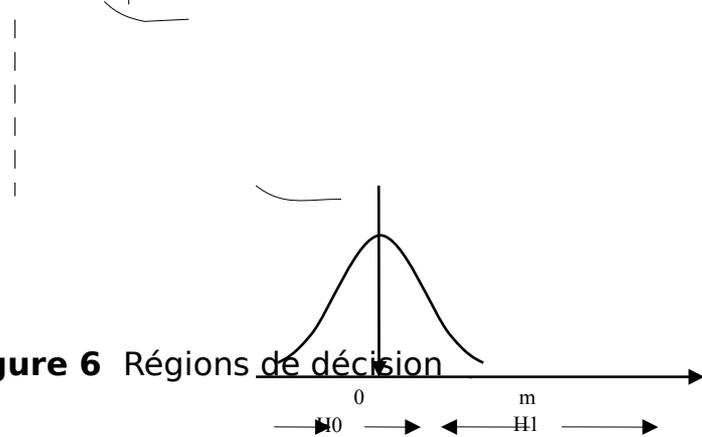


Figure 6 Régions de décision

Exemple 2 : En supposant que le récepteur de l'exemple 1 prenne K échantillons Y_1, Y_2, \dots, Y_K et que les échantillons du bruit soient des variables aléatoires Gaussiennes et indépendantes, chacune de moyenne nulle et de variance σ^2 , obtenir la règle de décision optimale ?

Réponse :

Le signal reçu sous chaque hypothèse est :

$$H_1 : Y_k = m + N_k, \quad k = 1, 2, \dots, K$$

$$H_0 : Y_k = N_k, \quad k = 1, 2, \dots, K$$

Sous l'hypothèse H_0 ,

$$f_{Y_k|H_0}(y_k | H_0) = f_{N_k}(y_k) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{y_k^2}{2\sigma^2}\right)$$

Sous l'hypothèse H_1 , le $k^{\text{ème}}$ échantillon reçu est une variable aléatoire Gaussienne de moyenne m et de variance σ^2 . Par conséquent,

$$f_{Y_k|H_1}(y_k | H_1) = f_{N_k}(y_k - m) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{(y_k - m)^2}{2\sigma^2}\right]$$

Il est clair qu'à partir de (22), nous avons besoin de $f_{Y|H_1}(y|H_1)$ et $f_{Y|H_0}(y|H_0)$. Comme les échantillons du bruit sont statistiquement indépendants, la densité de probabilité jointe des K échantillons est le produit de chacune des fonctions densités de probabilité. Ceci implique:

$$f_{Y|H_0}(y | H_0) = \prod_{k=1}^K \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{y_k^2}{2\sigma^2}}$$

$$\text{et } f_{Y|H_1}(y | H_1) = \prod_{k=1}^K \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y_k - m)^2}{2\sigma^2}}$$

où \prod désigne le produit. En utilisant le fait que $\prod_k e^{x_k} = e^{\sum_k x_k}$, le test de rapport de vraisemblance est :

$$\Lambda(y) = \exp\left[\sum_{k=1}^K \frac{y_k^2}{2\sigma^2} - \sum_{k=1}^K \frac{(y_k - m)^2}{2\sigma^2}\right] = \exp\left[\frac{m}{\sigma^2} \sum_{k=1}^K y_k - \frac{Km^2}{2\sigma^2}\right]$$

En prenant le logarithme naturel de chaque côté de l'équation ci-dessus, nous obtenons le test du rapport de vraisemblance devient:

$$\ln \Lambda(y) = \begin{matrix} H_1 \\ \frac{m}{\sigma^2} \sum_{k=1}^K y_k - \frac{Km^2}{2\sigma^2} > \ln \eta \\ H_0 \\ < \end{matrix}$$

ou, en réarrangeant les termes, nous obtenons :

$$T(y) = \sum_{k=1}^K y_k \begin{matrix} H_1 \\ > \frac{\sigma^2}{m} \ln \eta + \frac{Km}{2} = \gamma \\ H_0 \\ < \end{matrix}$$

Autrement dit, le récepteur effectue la somme des K échantillons et les compare au seuil γ .

Remarque: Une statistique est une variable aléatoire obtenue à partir d'une observation de données.

Du fait que la connaissance de \mathbf{Y} relève uniquement de $\sum_{k=1}^K y_k$ dans la prise

de décision et non des données individuelles, alors $T(\mathbf{Y}) = \sum_{k=1}^K Y_k$ est dite statistique suffisante. En d'autres termes, le récepteur n'a besoin que de

$T(\mathbf{y}) = \sum_{k=1}^K y_k$ pour prendre une décision sans pour autant se soucier des y_k , $k=1, 2, \dots, K$. Le rapport de vraisemblance dépend des observations y_k uniquement à travers la statistique $T(\mathbf{Y})$, mais pas de l'hypothèse H_j . Ceci veut dire que le test optimal peut être construit directement à partir de l'analyse de $T(\mathbf{y})$. Il n'est donc pas nécessaire de mémoriser toutes les données (valeurs) observées.

Exemple 3 : Considérer le cas où les fonctions densités conditionnelles sous chaque hypothèse sont données par :

$$f_{y|H_0}(y|H_0) = \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left(-\frac{y^2}{2\sigma_0^2}\right)$$

et $f_{y|H_1}(y|H_1) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{y^2}{2\sigma_1^2}\right)$

où $\sigma_1^2 > \sigma_0^2$. Déterminer la règle de décision.

Réponse :

L'application du test du rapport de vraisemblance (22), donne

$$\Lambda(y) = \frac{\frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{y^2}{2\sigma_1^2}\right) H_1}{\frac{1}{\sqrt{2\pi}\sigma_0} \exp\left(-\frac{y^2}{2\sigma_0^2}\right) H_0} > \eta$$

$$\frac{\sigma_0}{\sigma_1} \exp\left[\frac{y^2}{2} \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)\right] > \eta$$

ou

En prenant le logarithme naturel des deux côtés de l'inéquation, nous obtenons :

$$\ln \frac{\sigma_0}{\sigma_1} + \frac{y^2}{2} \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right) > \ln \eta$$

$$\begin{array}{l}
 H_1 \\
 y^2 > \frac{2\sigma_0^2\sigma_1^2}{\sigma_1^2 - \sigma_0^2} \ln \frac{\eta\sigma_1}{\sigma_0} = \gamma \\
 < \\
 H_0
 \end{array}$$

ou

Comme $T(Y) = Y^2$ est une statistique suffisante, le test peut s'écrire comme suit :

$$\begin{array}{l}
 H_1 \\
 T(y) > \gamma \\
 < \\
 H_0
 \end{array}$$

Exemple 4: Reprendre l'exemple précédent en considérant K observations indépendantes.

Réponse :

Comme les variables aléatoires Y_1, Y_2, \dots, Y_K sont indépendantes, la fonction densité de probabilité jointe est simplement le produit de chaque fonction densité individuelle:

$$f_{Y|H_0}(y|H_0) = \prod_{k=1}^K \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{y_k^2}{2\sigma_0^2}} \quad \text{et} \quad f_{Y|H_1}(y|H_1) = \prod_{k=1}^K \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{y_k^2}{2\sigma_1^2}}$$

En substituant dans (22) et en prenant le logarithme naturel, nous obtenons:

$$\begin{array}{l}
 H_1 \\
 \frac{1}{2} \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2} \right) \sum_{k=1}^K y_k^2 + K \ln \frac{\sigma_0}{\sigma_1} > \ln \eta \\
 < \\
 H_0
 \end{array}
 \quad \text{ou} \quad
 \begin{array}{l}
 H_1 \\
 \sum_{k=1}^K y_k^2 > \frac{2\sigma_0^2\sigma_1^2}{\sigma_1^2 - \sigma_0^2} \ln \eta \left(\frac{\sigma_1}{\sigma_0} \right)^K = \gamma \\
 < \\
 H_0
 \end{array}$$

Comme la statique $T(Y) = \sum_{k=1}^K Y_k^2$ est suffisante, le test peut s'écrire sous la forme:

$$\begin{array}{l}
 H_1 \\
 T(y) = \sum_{k=1}^K y_k^2 > \gamma \\
 < \\
 H_0
 \end{array}$$

Noter que si $\sigma_1^2 < \sigma_0^2$, alors la différence $\sigma_1^2 - \sigma_0^2$ est négative. Auquel cas, le sens de l'inégalité ci-dessus doit être inversé; c'est-à-dire :

$$\begin{array}{c}
 H_0 \\
 T(y) \begin{array}{l} > \\ < \end{array} \gamma \\
 H_1
 \end{array}$$

3. Critère du minimax

Dans plusieurs situations, il se peut que nous n'ayons pas assez d'information sur les probabilités *a priori*. Dans ce cas, le critère de Bayes ne peut être utilisé. Une autre approche consiste à choisir une valeur de P_1 , la probabilité *a priori* de l'hypothèse H_1 , pour laquelle le risque est maximal et puis minimiser cette fonction risque. Le principe de minimiser le maximum du coût moyen pour cette valeur de P_1 s'appelle le critère du minimax.

De (2), nous avons:

$$P_0 = 1 - P_1 \quad (30)$$

En substituant (30) dans (14), nous obtenons le risque en fonction de P_1 :

$$\mathfrak{R} = C_{00}(1 - P_F) + C_{10}P_F + P_1[(C_{11} - C_{00}) + (C_{01} - C_{11})P_M - (C_{10} - C_{00})P_F] \quad (31)$$

En supposant une valeur fixe de P_1 , le test de Bayes est donc:

$$\begin{array}{c}
 H_1 \\
 \Lambda(y) \begin{array}{l} > \\ < \end{array} \frac{(1 - P_1)(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} = \eta \\
 H_0
 \end{array} \quad (32)$$

De (32), nous remarquons que quand P_1 varie, le seuil η varie. Donc Z_0 et Z_1 varient et par conséquent P_F et P_M varient aussi. A cet effet, le risque lui aussi varie. Les trois courbes possibles du coût \mathfrak{R} sont montrées en Figure 7. Il y a deux cas extrêmes à investiguer; $P_1=0$ et $P_1=1$. Si $P_1=0$, alors le seuil est infini et la règle de décision est :

$$\begin{array}{c}
 H_1 \\
 \Lambda(y) \begin{array}{l} > \\ < \end{array} \infty \\
 H_0
 \end{array} \quad (33)$$

H_0 est toujours vraie. L'observation est Z_0 et les probabilités P_F et P_M sont telles que :

$$P_F = \int_{Z_1} f_{Y|H_0}(y|H_0)dy = 0 \quad (34)$$

et

$$P_M = \int_{Z_0} f_{Y|H_1}(y|H_1)dy = 1 \quad (35)$$

En substituant les valeurs de P_1, P_F , et P_M dans (31), nous obtenons le risque

$$\mathfrak{R} = C_{00} \quad (36)$$

Similairement, quand $P_1 = 1$, le seuil (32) est nul et la nouvelle règle de décision est :

$$\begin{array}{l} H_1 \\ \Lambda(y) > 0 \\ < 0 \\ H_0 \end{array} \quad (37)$$

Comme $\Lambda(y)$ est non négatif, nous décidons toujours H_1 . Donc, $P_F = 1$ et $P_M = 0$. Le risque résultant est:

$$\mathfrak{R} = C_{11} \quad (38)$$

A partir de (31), nous voyons que le risque \mathfrak{R} est linéaire en fonction de P_1 et le test de Bayes pour $P_1 = P_1^*$ donne le risque maximum $R_{\max} = \mathfrak{R}^*(P_1^*)$. Remarquons que la tangente au point R_{\max} est horizontale et que la courbe de Bayes doit être convexe. Donc, le coût moyen ne dépasse pas $\mathfrak{R}^*(P_1^*)$. En prenant la dérivée de \mathfrak{R} par rapport à P_1 et en l'égalant à zéro, nous obtenons l'équation minimax suivante:

$$(C_{11} - C_{00}) + (C_{01} - C_{11})P_M - (C_{10} - C_{00})P_F = 0 \quad (39)$$

Si le coût d'une décision correcte est nul ($C_{00} = C_{11} = 0$), alors l'équation minimax (39) pour $P_1 = P_1^*$ se réduit à:

$$C_{01}P_M = C_{10}P_F \quad (40)$$

Si, de plus, le coût d'une décision erronée est 'un' ($C_{01} = C_{10} = 1$), alors

$$P_F = P_M \quad (41)$$

Le coût minimax est:

$$\begin{aligned} \mathfrak{R} &= P_F(1-P_1) + P_1 P_M = P_0 P_F + P_1 P_M \\ &= P_M = P_F \end{aligned} \quad (42)$$

Il est clair (42) que le coût minimax est la probabilité d'erreur moyenne et c'est une constante indépendante de P_1 .

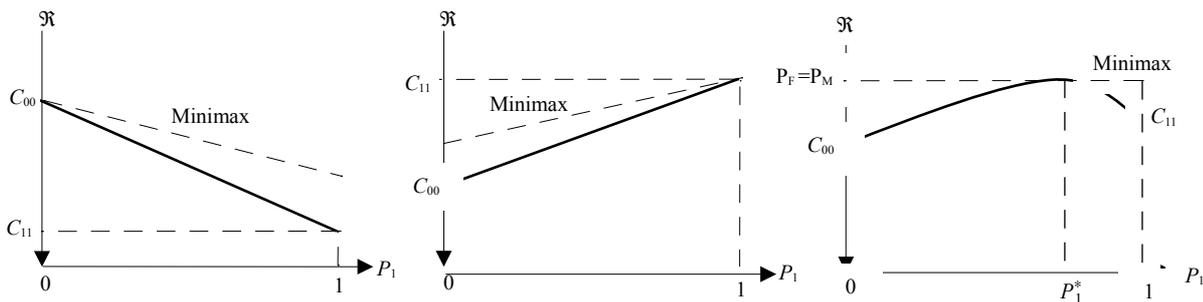


Figure 7 Les trois courbes possibles du risque en fonction de P_1

Exemple 5: Considérer le problème de l'exemple 1. Calculer la probabilité d'erreur minimum quand:

- $P_0 = P_1$.
- P_0 et P_1 sont inconnues.
- Comparer les résultats obtenus en (a) et (b). Que concluez-vous ?

Réponses :

(a) De l'exemple (1), nous avons trouvé que la règle de décision est:

$$\begin{aligned} &H_1 \\ y &> \frac{\sigma^2}{m} \ln \eta + \frac{m}{2} = \gamma \\ &< \\ &H_0 \end{aligned}$$

Sachant que $P_0 = P_1 = 1/2$, la probabilité d'erreur $P(\epsilon) = (1/2)(P_F + P_M)$, où

$$P_F = Q\left(\frac{\gamma}{\sigma}\right)$$

et

$$P_M = 1 - P_D = 1 - Q\left(\frac{\gamma - m}{\sigma}\right) = Q\left(\frac{m - \gamma}{\sigma}\right)$$

(b) Dans ce cas, le seuil optimum γ^* est obtenu quand $P_F = P_M$. Donc :

$$Q\left(\frac{\gamma^*}{\sigma}\right) = Q\left(\frac{m - \gamma^*}{\sigma}\right)$$

Le seuil γ^* est alors $\gamma^* = m/2$. Par conséquent, la probabilité d'erreur moyenne est:

$$P(\varepsilon) = P_0 P_F + P_1 P_M = (P_0 + P_1) P_M = Q\left(\frac{m}{2\sigma}\right)$$

(c) Afin de comparer les résultats de (a) et (b), nous devons normaliser l'écart type de l'observation de (a) à 1. Pour cela, nous posons $y' = y/\sigma$, et comme $\eta = 1$, la règle de décision devient :

$$\begin{array}{l} H_1 \\ y' > \frac{m}{2\sigma} = \gamma \\ H_0 \end{array}$$

Soit $\alpha = m/\sigma$. La règle de décision se réduit à:

$$\begin{array}{l} H_1 \\ y' > \frac{\alpha}{2} \\ H_0 \end{array}$$

Les probabilités P_F et P_D sont:

$$P_F = \int_{\alpha/2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y'^2}{2}} dy' = Q\left(\frac{\alpha}{2}\right)$$

et

$$P_D = \int_{\alpha/2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y' - \alpha)^2}{2}} dy' = Q\left(\frac{\alpha}{2} - \alpha\right) = Q\left(-\frac{\alpha}{2}\right)$$

Donc, $P_M = 1 - P_D = 1 - Q(-\alpha/2) = Q(\alpha/2)$. La probabilité d'erreur moyenne est $P(\varepsilon) = (1/2)[Q(\alpha/2) + Q(\alpha/2)] = Q(\alpha/2) = Q(m/2\sigma)$. Par conséquent, les résultats trouvés en (a) et (b) sont les mêmes.

4. Critère de Neyman-Pearson

Dans plusieurs autres applications telle que la détection radar, il est difficile d'attribuer des coûts réalistes et des probabilités *a priori* pour que nous puissions utiliser le critère de Bayes ou le critère du minimax dans des situations où il n'est pas possible de connaître les probabilités *a priori*. Dans de tels cas, nous utilisons les probabilités conditionnelles P_F et P_D . Le

test de Neyman-Pearson (N-P) fixe la P_F à une valeur α et maximise la P_D . Comme $P_M = 1 - P_D$; maximiser la P_D est équivalent à minimiser la P_M . Pour ce faire, nous formons la fonction objective J tel que :

$$J = P_M + \lambda(P_F - \alpha) \quad (43)$$

où $\lambda (\lambda \geq 0)$ désigne le *multiplicateur de Lagrange*. Etant donné l'espace d'observation Z , il existe plusieurs régions de décision Z_1 pour lesquelles $P_F = \alpha$. Le problème est de déterminer parmi ces régions celles qui garantissent une P_M minimale. Pour cela, écrivons J en fonction des régions de décision:

$$J = \int_{Z_0} f_{Y|H_1}(y|H_1)dy + \lambda \left[\int_{Z_1} f_{Y|H_0}(y|H_0)dy - \alpha \right] \quad (44)$$

En utilisant (17), (44) peut se mettre sous la forme

$$\begin{aligned} J &= \int_{Z_0} f_{Y|H_1}(y|H_1)dy + \lambda \left[1 - \int_{Z_0} f_{Y|H_0}(y|H_0)dy - \alpha \right] \\ &= \lambda(1 - \alpha) + \int_{Z_0} [f_{Y|H_1}(y|H_1) - \lambda f_{Y|H_0}(y|H_0)]dy \end{aligned} \quad (45)$$

Conséquemment, les valeurs qui minimisent J sont attribuées à la région de décision Z_0 , garantissant $f_{Y|H_1}(y|H_1) < \lambda f_{Y|H_0}(y|H_0)$. La règle de décision est alors donnée par:

$$\Lambda(y) = \begin{matrix} H_1 \\ \frac{f_{Y|H_1}(y|H_1)}{f_{Y|H_0}(y|H_0)} > \lambda \\ H_0 \end{matrix} \quad (46)$$

Le seuil η , dérivé à partir du critère de Bayes est équivalent à λ , le multiplicateur de Lagrange dans le test de Neyman-Pearson qui fixe P_F à α .

Si nous définissons $f_{\Lambda|H_0}(\lambda|H_0)$, la fonction densité de probabilité de Λ sachant que H_0 est vraie, alors $P_F = \alpha$ peut être réécrite comme suit :

$$P_F = \int_{Z_1} f_{Y|H_0}(y|H_0)dy = \int_{\lambda}^{\infty} f_{\Lambda(y)|H_0}[\lambda(y)|H_0]d\lambda \quad (47)$$

Un tel test est dit le plus puissant de niveau α (*most powerful test of level α*) si sa probabilité de rejection de l'hypothèse H_0 est α .

Exemple 6 : Considérer le problème des hypothèses binaires avec les probabilités conditionnelles suivantes :

$$f_{Y|H_0}(y|H_0) = \frac{1}{2(1-e^{-1})} e^{-|y|} \text{ pour } |y| \leq 1$$

et $f_{Y|H_1}(y|H_1) = \frac{1}{2} \text{rect}\left(\frac{1}{2}y\right)$

Les hypothèses H_0 et H_1 sont équiprobables.

- Trouver la règle de décision telle que la probabilité d'erreur est minimale.
- Calculer la probabilité d'erreur minimale.
- Trouver la règle de décision basée sur le critère de N-P telle que $P_F = 0.5$.
- Calculer la P_D pour le cas de la contrainte sur la P_F donnée en (c).

Réponses :

(a) Un récepteur à probabilité d'erreur minimale requiert que $C_{00} = C_{11} = 0$ et $C_{01} = C_{10} = 1$. Comme les probabilités *a priori* sont égales, le test du rapport de vraisemblance est donné par:

$$\Lambda(y) = \frac{f_{Y|H_1}(y|H_1)}{f_{Y|H_0}(y|H_0)} \underset{H_0}{>} \underset{H_1}{<} 1$$

C'est à dire, nous choisissons l'hypothèse pour laquelle $f_{Y|H_j}(y|H_j)$, $j=0,1$ est maximale. Les régions de décision sont représentées en Figure 8. Plus précisément, nous décidons H_1 pour $-1 \leq y \leq -0.459$ et $0.459 \leq y \leq 1$, et décidons H_0 pour $-0.459 < y < 0.459$.

(b) La probabilité d'erreur est $P(\varepsilon) = P_0 P_F + P_1 P_M$, où

$$P_F = P(\text{décider } H_1 | H_0 \text{ vraie}) = \frac{1}{2(1-e^{-1})} \left(\int_{-1}^{-0.459} e^y dy + \int_{0.459}^1 e^{-y} dy \right) = 0.418$$

et $P_M = P(\text{décider } H_0 | H_1 \text{ vraie}) = 2[(0.459)(1/2)] = 0.459$.

Donc, la probabilité d'erreur est $P(\varepsilon) = (1/2)(0.418 + 0.459) = 0.4385$.

(c) En utilisant le critère de Neyman-Pearson, nous obtenons :

$$\Lambda(y) = \begin{cases} \frac{1}{2} & H_1 \\ & > \lambda \\ \frac{1}{2(1-e^{-1})} e^{-|y|} & H_0 \\ & < \lambda \end{cases}$$

$$f_{Y|H_0}(y|H_0) = \frac{1}{2(1-e^{-1})} e^{-|y|} \text{ pour } |y| \leq 1$$

$$f_{Y|H_1}(y|H_1) = \frac{1}{2} \text{rect}\left(\frac{1}{2}y\right)$$

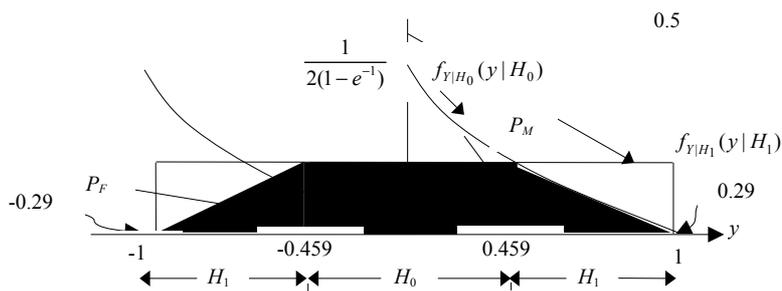


Figure 8 Régions de décision

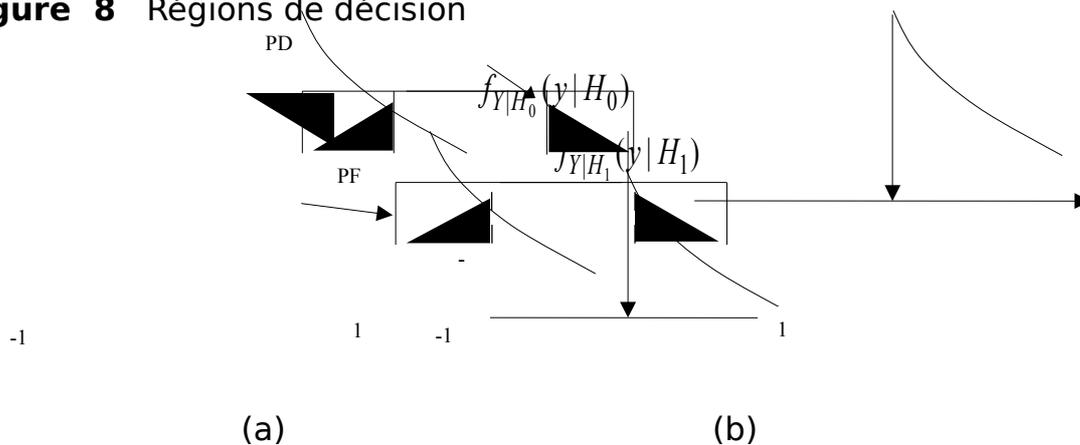


Figure 9 Régions de décision montrant (a) P_F , et (b) P_D .

Alors,

$$\frac{1-e^{-1}}{e^{-|y|}} > \lambda \quad \text{ou} \quad |y| > -\ln \frac{1-e^{-1}}{\lambda} = \gamma$$

$$\frac{1}{2} < \lambda \quad \text{ou} \quad |y| < -\ln \frac{1-e^{-1}}{\lambda} = \gamma$$

Comme P_F est tel que montrée en Figure 9(a), donc,

$$P_F = P(D_1|H_0) = \frac{1}{2(1-e^{-1})} \left(\int_{-1}^{-\gamma} e^y dy + \int_{\gamma}^1 e^{-y} dy \right) = 0.5$$

Ce qui donne un seuil $\gamma=0.38$.

(d) La P_D est tel que montrée en Figure 9(b), est $P_D = 2[(1-0.38)(1/2)] = 0.62$.

5. Détection séquentielle

Dans les sections précédentes, nous avons considéré la théorie des tests d'hypothèses pour un nombre fixe d'observations sur lequel le test est effectué. Dans plusieurs situations pratiques, les observations peuvent être prises de manière séquentielle de telle sorte à ce que le test soit effectué après chaque observation. A chaque fois qu'une observation est considérée, l'une des trois décisions peut être donc prise :

1. Décider H_1
2. Décider H_0
3. Information insuffisante pour décider en faveur de H_1 ou H_0 .

Si la décision (1) ou (2) est prise, le test s'arrête. Sinon, une observation additionnelle est considérée; et le test est effectué une nouvelle fois. Le processus doit continuer jusqu'à ce qu'une décision soit prise en faveur de H_1 ou H_0 . Notons que le nombre d'observations n'est pas fixe, mais il est une variable aléatoire.

Le test à effectuer après chaque observation calcule le rapport de vraisemblance et le compare à deux seuils, η_0 et η_1 . Un tel test qui considère les trois prises de décision possibles citées ci-dessus, après la $K^{\text{ème}}$ observation, est dit rapport de vraisemblance séquentiel.

Soit $Y_k, k=1,2,\dots,K$, le $K^{\text{ème}}$ échantillon du vecteur \mathbf{Y}_K est défini comme suit :

$$\mathbf{Y}_K^T = [Y_1 \ Y_2 \ \dots \ Y_K] \quad (48)$$

Le rapport de vraisemblance sur les K premières observations est:

$$\Lambda(\mathbf{y}_K) = \frac{f_{\mathbf{Y}_K|H_1}(\mathbf{y}_K | H_1)}{f_{\mathbf{Y}_K|H_0}(\mathbf{y}_K | H_0)} \quad (49)$$

Pour calculer le rapport de vraisemblance (49), nous avons besoin de la fonction de densité jointe de ces K observations. Pour simplifier, nous supposons que les observations sont identiquement distribuées. Le rapport de vraisemblance peut être écrit comme le produit de K rapports de vraisemblance.

$$\Lambda(\mathbf{y}_K) = \frac{f_{\mathbf{Y}_K|H_1}(\mathbf{y}_K | H_1)}{f_{\mathbf{Y}_K|H_0}(\mathbf{y}_K | H_0)} = \prod_{k=1}^K \frac{f_{Y_k|H_1}(Y_k | H_1)}{f_{Y_k|H_0}(Y_k | H_0)} \quad (50)$$

Le but est de déterminer η_0 et η_1 en fonction de la P_F et P_M . Nous posons alors :

$$P_F = \alpha \quad (51)$$

et

$$P_M = \beta \quad (52)$$

En effectuant le test suivant. Si

$$\Lambda(\mathbf{y}_K) \geq \eta_1 \quad (53)$$

nous décidons en faveur de H_1 . Si, par contre :

$$\Lambda(\mathbf{y}_K) \leq \eta_0 \quad (54)$$

nous décidons en faveur de H_0 . Sinon, si

$$\eta_0 < \Lambda(\mathbf{y}_K) < \eta_1 \quad (55)$$

nous prenons une observation additionnelle et effectuons un autre test. La P_D en termes d'intégrale sur l'espace des observations est :

$$P_D = P(\text{décider } H_1 | H_1 \text{ vraie}) = \int_{Z_1} f_{Y_K|H_1}(\mathbf{y}_K | H_1) d\mathbf{y}_K \quad (56)$$

En utilisant (49), la P_D peut s'écrire sous la forme :

$$P_D = \int_{Z_1} \Lambda(\mathbf{y}_K) f_{Y_K|H_0}(\mathbf{y}_K | H_0) d\mathbf{y}_K \quad (57)$$

La décision en faveur de H_1 veut dire que $\Lambda(\mathbf{y}_K) \geq \eta_1$. Donc, en substituant (53) dans (57), nous obtenons l'inégalité suivante :

$$P_D \geq \eta_1 \int_{Z_1} f_{Y_K|H_0}(\mathbf{y}_K | H_0) d\mathbf{y}_K \quad (58)$$

Notons que l'intégrale

$$\int_{Z_1} f_{Y_K|H_0}(\mathbf{y}_K | H_0) d\mathbf{y}_K = P_F = \alpha \quad (59)$$

et comme $P_D = 1 - P_M = 1 - \beta$, (58) devient

$$1 - \beta \geq \eta_1 \alpha \quad (60)$$

or, the seuil η_1 est

$$\eta_1 \leq \frac{1 - \beta}{\alpha} \quad (61)$$

Similairement, nous pouvons montrer que η_0 est :

$$\eta_0 \geq \frac{\beta}{1 - \alpha} \quad (62)$$

A ce niveau, quelques questions importantes méritent d'être posées et résolues. Pour ce faire, quelle la probabilité pour que la procédure ne se termine jamais ? Quelles sont quelques propriétés de la distribution de la variable aléatoire K ? En particulier, quelle est l'espérance mathématique de sa grandeur ? Pour répondre à ces questions, il est plus facile d'utiliser le logarithme de la fonction de vraisemblance. En prenant le logarithme naturel de (55), nous obtenons:

$$\ln \eta_0 < \ln \frac{f_{Y_1|H_1}(y_1 | H_1)}{f_{Y_1|H_0}(y_1 | H_0)} + \dots + \ln \frac{f_{Y_K|H_1}(y_K | H_1)}{f_{Y_K|H_0}(y_K | H_0)} < \ln \eta_1 \quad (63)$$

Soit le $k^{\text{ème}}$ terme, $k = 1, 2, \dots, K$, de la somme ci-dessus

$$L(y_k) = \ln \frac{f_{Y_k|H_1}(y_k | H_1)}{f_{Y_k|H_0}(y_k | H_0)} \quad (64)$$

alors, (63) devient:

$$\ln \eta_0 < L(y_1) + \dots + L(y_k) + \dots + L(y_K) < \ln \eta_1 \quad (65)$$

La somme peut être écrite de manière récursive comme suit:

$$L(\mathbf{y}_K) = L(\mathbf{y}_{K-1}) + L(y_K) \quad (66)$$

où

$$L(\mathbf{y}_{K-1}) = L(y_1) + L(y_2) + \dots + L(y_{K-1}) = \sum_{k=1}^{K-1} L(y_k) \quad (67)$$

Pour calculer $E[K]$, le nombre moyen des observations sous chaque hypothèse, nous supposons que le test se termine au bout de K

observations avec une probabilité égale à un. Cette supposition implique que $L(y_K)$ prend deux valeurs possibles, $\ln \eta_0$ et $\ln \eta_1$. Si l'hypothèse H_1 est vraie, une détection est déclarée quand $L(y_K) \geq \ln \eta_1$ avec une probabilité $P_D = 1 - P_M = 1 - \beta$. Un ratage arrive quand $L(y_K) \leq \ln \eta_0$ avec une probabilité $P_M = \beta$. Donc, l'espérance mathématique $L(y_K)$ sous l'hypothèse H_1 est:

$$E[L(y_K) | H_1] = \beta \ln \eta_0 + (1 - \beta) \ln \eta_1 \quad (68)$$

En suivant le même raisonnement, l'espérance mathématique de $L(y_K)$ sous l'hypothèse H_0 est :

$$E[L(y_K) | H_0] = \alpha \ln \eta_1 + (1 - \alpha) \ln \eta_0 \quad (69)$$

Soit B une variable aléatoire prenant les valeurs binaires zéro et un, tel que :

$$B_k = \begin{cases} 1, & \text{pas de décision prise jusqu'à } (k-1) \text{ échantillons} \\ 0, & \text{d'écision prise pour un échantillon antérieur} \end{cases} \quad (70)$$

En effet, B_k dépend des observations $Y_k, k=1, 2, \dots, K-1$, et non de Y_K . En réécrivant le logarithme du rapport de vraisemblance en fonction de la variable aléatoire B , nous obtenons :

$$L(y_K) = \sum_{k=1}^K L(y_K) = \sum_{k=1}^{\infty} B_k L(y_K) \quad (71)$$

Comme les observations sont IID, nous avons:

$$E[L(y_K) | H_j] = E[L(y) | H_j] \sum_{k=1}^{\infty} E[B_k], \quad j = 0, 1 \quad (72)$$

où

$$E[L(y) | H_j] = E[L(y_1) | H_j] = \dots = E[L(y_K) | H_j] \quad (73)$$

La somme de (72) est simplement:

$$\sum_{k=1}^{\infty} E[B_k] = \sum_{k=1}^{\infty} P(K \geq k) = \sum_{k=1}^{\infty} k P(K = k) = E[K] \quad (74)$$

En substituant (74) et (73) dans (68), nous obtenons :

$$E[L(y) | H_1] E[K | H_1] = \alpha \ln \eta_1 + (1 - \alpha) \ln \eta_0 \quad (75)$$

ou

$$E[K | H_1] = \frac{(1-\beta) \ln \eta_1 + \beta \ln \eta_0}{E[L(y) | H_1]} \quad (76)$$

Similairement, l'espérance mathématique de K sous l'hypothèse H_0 s'écrit:

$$E[K | H_0] = \frac{\alpha \ln \eta_1 + (1-\alpha) \ln \eta_0}{E[L(y) | H_0]} \quad (77)$$

Pour répondre à la question inhérente à la terminaison du processus avec une probabilité égale à un, il est facile de montrer que:

$$\lim_{k \rightarrow \infty} P(K \geq k) = 0 \quad (78)$$

De plus, il est possible de montrer que l'espérance mathématique du nombre d'observations K sous chaque hypothèse est minimale pour les valeurs spécifiées de P_F et P_M .

Exemple 7: Supposons que le récepteur de l'exemple 2 prend K observations séquentiellement. Soient la variance $\sigma^2 = 1$ et la moyenne $m = 1$.

- (a) Déterminer la règle de décision tel que $P_F = \alpha = 0.1 = P_M = \beta$.
- (b) Déterminer l'espérance mathématique de K sous chaque hypothèse.

Réponses :

(a) La définition de la règle de décision est donnée par (53), (54) et (55). Par conséquent, nous avons besoin de trouver le rapport de vraisemblance au $k^{\text{ème}}$ niveau et les seuils η_0 et η_1 . En substituant $\sigma^2 = 1$ et $m = 1$ dans le rapport de vraisemblance de l'exemple 2, nous obtenons le rapport de vraisemblance du $k^{\text{ème}}$ étage :

$$\Lambda(\mathbf{y}_K) = \exp\left(\sum_{k=1}^K y_k - \frac{K}{2}\right)$$

Le logarithme du rapport de vraisemblance est simplement:

$$L(\mathbf{y}_K) = \ln \Lambda(\mathbf{y}_K) = \sum_{k=1}^K y_k - \frac{K}{2}$$

De (61) and (62), les deux seuils sont:

$$\ln \eta_1 = 2.197 \text{ et } \ln \eta_0 = -2.197$$

Donc, la règle de décision en fonction du logarithme du rapport de vraisemblance est :

Si $L(y_K) \geq 2.197$, décider H_1 .

Si $L(y_K) \leq -2.197$, décider H_0 .

Si $-2.197 \leq L(y_K) \leq 2.197$, prendre l'observation additionnelle $K+1$ et effectuer un autre test.

(b) Les espérances mathématique de K sous les hypothèses H_1 et H_0 sont données par (76) et (77), respectivement. Nous remarquons que nous avons besoin d'obtenir $E[L(y)|H_1]$ et $E[L(y)|H_0]$. En supposant que les observations sont identiques, nous avons $E[L(y)|H_1] = 1 - (1/2) = 1/2$ et $E[L(y)|H_0] = 0 - (1/2) = -1/2$. En substituant pour les valeurs de $E[L(y)|H_1]$ et $E[L(y)|H_0]$ dans (76) et (77), nous obtenons $E[K|H_1] = 3.515$ et $E[K|H_0] = 3.515$. Autrement dit, nous avons besoin de quatre échantillons pour obtenir la performance correspondant à $P_F = P_M = 0.1$.