EDITH MOURIER

Sur les probabilités de fausse alarme et de non-détection et le rapport signal/bruit

Publications des séminaires de mathématiques et informatique de Rennes, 1968-1969, fascicule 2
« Séminaire de probabilités, statistiques et analyse », , exp. nº 5, p. 1-13
<http://www.numdam.org/item?id=PSMIR_1968-1969___2_A5_0>

SUR LES PROBABILITES DE FAUSSE ALARME ET DE NON-DETECTION ET LE RAPPORT SIGNAL/BRUIT

par

Edith MOURIER

Professeur à la Faculte des Sciences de Paris

Résumé :

On étudie les principaux critères utilisés en théorie statistique de la décision, on rappelle les différentes définitions du rapport S/B et on compare le critère S/B aux critères probabilistes.

Summary :

The main tests used in statistical decision theory are considered, the different definitions of the signal to noise ratio are reminded, then the S/N criterion and probabilistic criteria are compared.

Séminaire sur la théorie de la Détection

SUR LES PROBABILITES DE FAUSSE ALARME ET DE NON-DETECTION ET LE RAPPORT SIGNAL SUR BRUIT

par

Edith MOURIER

Introduction.

L'usage s'est répandu d'utiliser la même expression : "critère de détection" dans des situations fort différentes, il en résulte une certaine confusion, de faux problèmes sont parfois soulevés et il semble utile de préciser à nouveau quelques notions fondamentales.

La théorie de la détection couvre un vaste ensemble de recherches qui ont en commun le but : au vu d'une observation, décider "au mieux" si pendant la durée de l'observation il y a eu seulement du bruit, hypothèse H_o, ou si un signal a été présent, hypothèse H_1. Dire "au mieux" sous-entend évidemment : relativement à un critère préalablement choisi.

Il y a deux problèmes distincts : éventuellement celui du choix d'un dispositif, disons de pré-traitement, partiellement imposé, qui, recevant à l'entrée une observation x_e, donne à la sortie une observation x, et toujours celui du choix d'une règle de décision, c'est-à-dire du choix d'une fonction de x, $\phi(x)$, (s'il n'y a pas de pré-traitement $x \equiv x_e$), et d'un partage des valeurs de cette fonction en deux domaines disjoints et complémentaires, D_0 et D_1, avec la règle : si $\phi(x) \in D_0$ accepter H_0, si $\phi(x) \in D_1$ accepter H_1.

Choix d'une règle de décision.

Le bruit à l'entrée, B_e, est un phénomène aléatoire, par conséquent, quelle que soit la nature du signal, quel que soit le pré-traitement, x est une réalisation d'un élément aléatoire X dont la loi de probabilité est μ_0 ou μ_1 selon qu'il a bruit seul ou présence d'un signal. μ_0 et μ_1 sont déterminés par la connaissance des propriétés statistiques du bruit et du mélange signal et bruit à l'entrée et du pré-traitement ; supposant ces deux lois connues, le choix d'une règle de décision est un problème de test de choix entre deux hypothèses statistiques, l'observation étant x.

Les deux erreurs que l'on peut commettre sont la fausse alarme (décider qu'il y a un signal alors qu'il y a seulement du bruit) et la non-détection (décider qu'il n'y a que du bruit alors qu'un signal est présent).

Puisqu'il s'agit d'un phénomène aléatoire la meilleure règle de décision - appelée "idéale" par Siegut - serait celle qui minimise la probabilité d'erreur, malheureusement cette probabilité

$$P = \pi_0 \alpha + \pi_1 \beta$$

dépend de la probabilité de fausse alarme α, de la probabilité de non-détection β, qui pour μ_0 et μ_1 donnés ne dépendent que de la règle de décision choisie, mais dépend aussi des probabilités a priori π_0 et $\pi_1 = 1-\pi_0$ d'absence et de présence du signal , probabilités qui, dans les applications, ne sont pas connues et par conséquent ce critère "idéal" est inutilisable.

Suivant Neyman et Pearson on remarque que des deux erreurs que l'on peut commettre, l'une a des conséquences plus graves que l'autre, on l'appelle erreur de première espèce - en théorie du signal, habituellement, c'est la fausse alarme - et on choisit le critère suivant : on se fixe la probabilité α que l'on accepte pour cette erreur - α est dit le niveau du test - et on décide que le "meilleur" test de choix entre μ_0 et μ_1, ou "test le plus puissant au niveau α" - souvent appelé test de Neyman - est celui qui, pour α fixé, minimise la probabilité β de l'erreur de seconde espèce, ici la probabilité de non-détection du signal alors qu'il est présent, donc maximiser la probabilité $1-\beta$ de détection.

On démontre [3] que quelle que soit la nature de l'élément aléatoire X, quelles que soient les lois de probabilités, supposées connues, μ_o et μ_1, il existe un meilleur test au niveau α, et que ce test est :

$$\text{(I)} \quad \begin{cases} \text{accepter } H_1 \text{ : signal présent si } L(x) > K(\alpha) & ^{(1)} \\ \text{accepter } H_o \text{ : bruit seul si } \quad L(x) < K(\alpha) \end{cases}$$

où $K(\alpha)$ est un seuil qui ne dépend que de α et où $L(x) = \dfrac{f_1(x)}{f_o(x)}$, $f_1(x)$ et $f_o(x)$ étant les densités de Radon-Nicodym de, respectivement, μ_1 et μ_o par rapport à une mesure μ par rapport à laquelle elles sont absolument continues ; on peut prendre, par exemple, $\mu = \mu_o + \mu_1$.

$L(x)$ est la généralisation du rapport de vraisemblance pour un élément aléatoire X de nature quelconque. Sous de très faibles hypothèses, on démontre [1], [4] que $L(x)$ est un résumé exhaustif , c'est-à-dire contient toute l'information utile, pour le choix entre μ_o et μ_1, ce qui explique la forme trouvée pour le test le plus puissant. Dans le cas particulier où X est une variable aléatoire à n dimensions, $x = (x_1,\dots,x_n)$ et $L(x)$ est le rapport de vraisemblance classique. Dans le cas où l'observation est continue sur un intervalle $[0,T]$, X est un processus aléatoire,

---

(1)
Si la probabilité de $L(X) = K$, quand il y a seulement du bruit est différente de zéro, quand $L(x) = K$ on décide H_o ou H_1 avec des probabilités γ et $1-\gamma$ respectivement, γ étant choisie de façon que la probabilité de fausse alarme soit exactement α. Pour ne pas alourdir les notations, nous négligerons ce cas dont l'intérêt est purement théorique.

généralement défini par sa loi temporelle, il s'agit alors de déterminer $L(x)$ à partir des lois temporelles relatives à H_o et H_1, ce calcul a été fait [1] dans le cas de la détection d'un signal non aléatoire connu dans un bruit gaussien additif, stationnaire ou non ; il n'a pas été fait dans le cas général et semble très difficile; on se ramène alors, par échantillonage ou en considérant les premiers termes d'un développement en série, de Karhunen-loève par exemple, à une variable à n dimensions et on calcule le rapport de vraisemblance correspondant $L_n(x)$; il faut alors étudier la convergence de L_n lorsque $n \to \infty$. Ce problème a été considéré par Grenander [2] et est actuellement étudié par M.M. Picinbono et Verrosi dans le cas d'un signal certain et d'un bruit additif qui est le produit d'un processus gaussien par une variable aléatoire.

Désignons par $\tilde{\beta}(\alpha)$ la probabilité de non-détection en utilisant le meilleur test au niveau α, c'est-à-dire le test (I), on a :

$$\alpha = P_r\left[L(X) > K(\alpha)/H_o\right] = \mu_o\{x : L(x) > K(\alpha)\}$$

$$\tilde{\beta}(\alpha) = P_r\left[L(X) < K(\alpha)/H_1\right] = \mu_1\{x : L(x) < K(\alpha)\}$$

Donc à deux valeurs α', α'' de la probabilité de fausse alarme $\alpha' < \alpha''$ correspondent :

$$K(\alpha') > K(\alpha'')$$

et $$\tilde{\beta}(\alpha') > \tilde{\beta}(\alpha'')$$

La probabilité de détection du signal lorsqu'il est présent, qui est égale à $1-\tilde{\beta}(\alpha)$, décroit donc en même temps que la probabilité de fausse alarme, résultat bien en accord avec l'intuition. $\tilde{\beta}(\alpha)$ caractérise la difficulté que l'on a à détecter le signal en utilisant le test de Neyman au niveau α; par conséquent, un pré-traitement T_1 est plus efficace qu'un pré-traitement T_2 si $\tilde{\beta}_1(\alpha) < \tilde{\beta}_2(\alpha)$ et donc le pré-traitement optimum, lorsqu'on appliquera à sa sortie le test de Neyman au niveau α est celui qui minimise $\tilde{\beta}(\alpha)$.

Pour appliquer le test de Neyman, outre $L(x)$, il faut calculer le seuil $K(\alpha)$. c désignant un nombre positif quelconque, soit

$$\gamma(c) = P_r \left|L(X) > c/H_o\right| = \mu_o\{x : L(x) > c\}$$

$K(\alpha)$ est le nombre, essentiellement unique, tel que :

$$\gamma(K(\alpha) \leq \alpha \leq \gamma(K(\alpha)-0)$$

La détermination du seuil $K(\alpha)$ nécessite donc la connaissance de la loi de la variable aléatoire $L(X)$ lorsqu'il y a seulement du bruit.

Souvent au lieu de choisir à priori la valeur α de la probabilité de fausse alarme et de construire le test (I), on choisit la valeur du seuil, soit c, et on adopte le test :

$$\text{(II)} \qquad \begin{cases} \text{accepter } H_1 \text{ : signal présent si } L(x) > c \\ \text{accepter } H_o \text{ : bruit seul} \quad \text{si } L(x) < c \end{cases}$$

Pour connaître complètement les qualités de ce test, il faut calculer la probabilité de fausse alarme, calcul qui nécessite la connaissance de la loi de $L(X)$ quand il y a bruit seul, puisque :

$$\alpha(c = P_r \left[L(X) > c/H_o\right]$$

et la probabilité de non-détection (quand le signal est présent) calcul qui nécessite la connaissance de la loi de L(X) lorsque le signal est présent, puisque :

$$\beta(c) = P_r \left[L(X) < c/H_1\right]$$

Evidemment, le test (II) est le test le plus puissant au niveau $\alpha(c)$.

Compte tenu de la signification intuitive du rapport de vraisemblance, on adopte souvent le test dit "du maximum de vraisemblance", correspondant au seuil c = 1. De tous les tests -et non seulement des tests de la forme (II)- c'est celui qui minimise la somme, $\alpha+\beta$, des probabilités de fausse alarme et de non-détection ; il coïncide avec le test "idéal" de Siegut dans le cas particulier où les probabilités a priori, Π_0 et Π_1, d'absence et de présence du signal sont égales à $\frac{1}{2}$.

Un test est choisi en supposant connues les lois μ_0 et μ_1. Puisqu'en fait la loi du bruit ne sera jamais parfaitement connue, il est nécessaire de savoir si le test possède une certaine "stabilité", plus précisément de savoir si en utilisant le test relatif à (μ_0,μ_1) alors que les vraies lois sont légèrement différentes, disons μ'_0,μ'_1, les probabilités d'erreur α', β' sont peu différentes des probabilités calculées α,β.

Dans le cas d'une observation x(t) sur un intervalle de temps $[0,T]$, pour détecter un signal connu S(t) dans un bruit additif gaussien, de carré sommable sur $[0,T]$, de covariance continue $\Gamma(t,\tau)$, Root [7] a montré qu'un test du type (II) est stable <u>si et seulement si</u> S(t) appertient au domaine de l'opérateur Γ^{-1}, opérateur inverse de l'opérateur intégral de noyau $\Gamma(t,\tau)$, et alors à une constante additive près

Log L(x) est de la forme :

$$\text{(III)} \qquad \int_0^T x(t)\, g(t)\, dt$$

g(t) étant définie par :

$$\text{(IV)} \qquad \int_0^T \Gamma(t,\tau)\, g(\tau)\, d\tau = S(t).$$

Rapport Signal sur Bruit.

Partant du fait bien évident que la facilité de détecter un signal dans un bruit dépend de leur importance relative on a défini un coefficient, appelé rapport signal sur bruit, en abrégé S/B, pour représenter cette importance relative. Malheureusement pour définir ce rapport il faut d'abord préciser la nature des signaux et des bruits en présence, et on doit choisir des définitions différentes selon les cas considérés.

Cette théorie suppose que le bruit à l'entrée est additif, indépendant du signal, centré de covariance stationnaire, et a une densité spectrale $\gamma_B(\tau)$; nous désignerons sa variance par σ_B^2.

Dans le cas d'un signal aléatoire et de même nature que le bruit, on définit le rapport S/B à l'entrée comme le rapport des puissances moyennes du signal et du bruit, c'est-à-dire :

$$\rho_e = [S/B]_{\text{entrée}} = \frac{\sigma_S^2}{\sigma_B^2}$$

A la sortie d'un filtre linéaire $\mathcal{F}$ le signal est $S' = \mathcal{F}(S)$ et le bruit $B' = \mathcal{F}(B)$. S' et B' sont de même nature que S et B, on définit donc :

$$\rho_s = [S/B]_{\text{sortie}} = \frac{\sigma^2_{S'}}{\sigma^2_{B'}}$$

Mais à la sortie d'un dispositif non linéaire , D, la situation est beaucoup plus compliquée. Il faut d'abord choisir ce que l'on appellera bruit de sortie et signal de sortie. Si $Y_B(+)$ et $Y_{B+S}(\tau)$ représentent la sortie de D , selon qu'il y a bruit seul ou signal présent à l'entrée, on prend les définitions, bien adaptées au système quadration, intégration, mais comportant un certain arbitraire :

$$\text{signal de sortie} = E\, Y_{B+S} - E\, Y_B$$

$$\text{bruit de sortie} = \left[E\, Y_B^2 - (E\, Y_B)^2\right]^{1/2}$$

$$\rho_s = [S/B]_{\text{sortie}} = \frac{E\, Y_{B+S} - E\, Y_B}{\left[E\, Y_s^2 - (E\, Y_B)^2\right]^{1/2}}$$

Dans le cas d'un signal certain S(t), de durée finie, ayant une transformée de Fourier $s(\tau)$ on a à considérer simultanément un signal d'énergie finie et un bruit de puissance finie donc d'énergie infinie; on définit alors le rapport S/B comme le rapport des puissances moyennes du signal et du bruit à un instant arbitraire t_c, donc

$$\rho = \frac{[S(t_o)]^2}{\sigma^2_B}$$

et à la sortie d'un filtre linéaire :

$$\rho_s = [S/B]_{sortie} = \frac{[S'(t_o)]^2}{\sigma^2_{B'}}$$

Le critère "signal sur bruit" est un critère de comparaison entre les dispositifs de traitement, du fait des différentes définitions du rapport signal sur bruit à la sortie, il n'a de signification bien précise que dans la comparaison de dispositifs analogues ; dans ce cas , par définition, on dit que de deux dispositifs le "meilleur" est celui qui, pour le même bruit et signal à l'entrée, donne le plus grand rapport signal sur bruit à la sortie. Lorsqu'on veut comparer des dispositifs qui ne sont pas analogues il est préférable de dire que le meilleur est celui qui pour le même $[S/B]_{sortie}$ a la plus faible $[S/B]_{entrée}$.

Dans le cas d'un signal certain, parmi tous les filtres linéaires, il en esiste un qui maximise le rapport $[S/B]_{sortie}$;il est défini, à un facteur constant près, par son gain :

(V) $$G(r) = \frac{s(r)^* \, e^{-2\Pi i r t_o}}{\gamma_B(\gamma)}$$

et on a à la sortie :

(VI) $$\rho_{max} = \int \frac{|s(\gamma)|^2}{\gamma_B(\gamma)} \, d\gamma$$

Ce filtre, appelé filtre adapté, se réduit dans le cas d'un bruit blanc, $\gamma_B(\gamma)$ = constante, au filtre de réponse percussionnelle :

(VII) $$R(t) = S(t_o - t).$$

On cherche souvent à opposer, ou à comparer, la théorie statistique de la décision, reposant sur le calcul des probabilités de détection et de fausse alarme, au critère signal sur bruit. Le problème de la détection n'est pas résolu par le choix du dispositif de traitement; il faut de plus préciser la règle de décision adoptée à la sortie du dispositif, en général c'est le choix d'un seuil, les probabilités de détection dépendent évidemment du choix de cette règle. Mais en supposant qu'à la sortie on adopte la règle qui pour une probabilité de fausse alarme donnée, α, maximise la probabilité de détection $1-\beta$, il est impossible de relier, de façon générale, cette probabilité au rapport signal sur bruit de sortie, donc de donner une signification probabiliste au critère signal sur bruit. Ceci s'explique par le fait que le rapport S/B ne dépend que des propriétés du 1° et du 2° ordre alors que les probabilités α et β dépendent des lois complètes.

Cependant, on peut comparer deux critères lorsque les lois sont gaussiennes, et alors complètement déterminées par les moments d'ordre un et deux.

Dans le problème de la détection d'un signal certain mélangé à du bruit gaussien centré, stationnaire le filtre optimum au sens du critère signal sur bruit, donne une sortie $y(t)$ qui à l'instant t_0 choisi pour donner une pointe de signal dans le bruit, est égale, à une constante près, $\frac{1}{2}\rho_{max}$, au logarithme du rapport de vraisemblance.

Ce résultat s'établit immédiatement dans le cas d'un bruit gaussien blanc, en effet la sortie du filtre adapté est alors :

$$y(t) = \int x(\Theta)\, S(t_o - t + \Theta)\, d\Theta$$

et il suffit alors de comparer :

$$y(t_o) = \int x(\Theta)\, S(\Theta)\, d\Theta$$

aux formules (III) et (IV).

Dans le cas d'un bruit gaussien stationnaire quelconque le résultat est moins simple à démontrer mais est encore valable.

De plus $Y(t_o) = \text{Log } L(X) + \frac{1}{2}\rho_{max}$ est une variable aléatoire gaussienne, de variance ρ_{max}, de moyenne 0 ou ρ_{max} selon qu'il y a bruit seul ou signal présent.

Lorsque le signal est de même nature que le bruit (gaussien, stationnaire, centré), le rapport de vraisemblance L(x) n'est plus une fonction linéaire de l'observation et on montre [6] que le système quadration, intégration précédé du filtrage optimal au sen critère S/B réalise une fonction monotone de L(x).

Or nous avons déjà remarqué que L(x) constitue un résumé exhaustif pour le choix entre deux lois, il en est de même de toute fonction monotone de L(x). Par conséquent dans les deux cas ci-dessus le système optimal au sens du critère S/B est aussi optimal au sens d'extraire de l'observation toute l'information utile.

BIBLIOGRAPHIE

[1] R. FORTET et E. MOURIER - Les fonctions aléatoires comme éléments aléatoires dans un espace de Banach. Journal de Math. Pures et Appliquées Tome 38. 1959.

[2] U. GRENANDER - Stochastic processes and statistical inference. Ark. Mat. 1 1950.

[3] E. L. LEHMANN - Testing statistical hypothses. Wiley Publications in statistics. 1959.

[4] E. MOURIER - Tests de choix entre diverses lois de probabilité. Trabajos de estadistica; Vol II Cuaduno III. 1951.

[5] E. MOURIER - Détection d'un signal non aléatoire connu en présence de bruit Laplacien de moyenne nulle, de covariance continue connue.
Rapport CNET NO 614. CME 1961.

[6] B. PICINBONO - Cours de 3ème cyle. Faculté des Sciences de PARIS.

[7] W. ROOT - Stability in signal detection problems. Proceedings of Symposia in Applied Mathematics. Vol XIV.