

# Seuil de décision et limite de détection : estimation, interprétation et optimisation 1<sup>re</sup> partie : les principes de base

A. VIVIER<sup>1</sup>, R. FOTTORINO<sup>2</sup>, B. ROUSSE<sup>3</sup>

(Manuscrit reçu le 7 janvier 2010, accepté le 2 avril 2010)

## RÉSUMÉ

La mise en œuvre des concepts de seuil de décision et de limite de détection pose souvent des problèmes aux métrologues des laboratoires d'analyse de biologie médicale chargés de la surveillance des salariés exposés aux rayonnements ionisants. Les problèmes mis en avant sont souvent relatifs au choix à faire entre diverses formules, dont les résultats numériques présentent rarement des différences significatives. Les raisons plus profondes des difficultés rencontrées sont en fait d'une part la signification et l'interprétation de ces tests statistiques, et d'autre part les valeurs utilisées dans ces formules parmi lesquelles l'estimation du bruit de fond reste la grandeur essentielle et la plus délicate à définir. L'objectif de ce premier article est de revenir dans un premier temps sur les fondements des concepts de seuil de décision et de limite de détection, en distinguant d'une part la méthode de détermination du seuil de décision basé sur un risque d'erreur *a priori* de première espèce, d'autre part l'interprétation du résultat après mesure de l'échantillon. Nous exposerons ensuite quelques exemples d'optimisation de ces seuils et limites en spectrométrie gamma, où nous verrons que tout repose sur la recherche de la meilleure estimation possible du bruit de fond qui nécessite bien plus une maîtrise de la mesure que des outils de statistique. Un second article, mettant en œuvre les principes exposés ici, sera plus particulièrement consacré à la démarche d'optimisation des seuils de décision en spectrométrie alpha pour les dosages d'uranium dans les urines.

## ABSTRACT

Decision threshold and detection limit: estimation, interpretation and optimization. 1st part: basic principles.

The use of decision threshold and detection limit concepts often poses many problems for metrologists in biological analysis laboratories in charge of the management of nuclear plant workers. These problems are frequently related to the understanding of a normalized formula or to the possible choice between several formulas whose numerical results are exceptionally significantly different. The true problems are in fact firstly the significance and the interpretation of these statistical tests, and secondly the numerical values used in these tests. Among those, the background estimate remains the essential and the most delicate value to define. The aim of this first article is to go back to the initial bases of the concepts of the decision threshold and detection limit. In the first step, a distinction will be made between the methods of decision threshold determination based upon definition of *a priori* first species error risks (*i.e.* before sample measurement) and the *a posteriori* interpretation of the sample

<sup>1</sup> CEA, Institut National des Sciences et Techniques nucléaires, Centre de Saclay, 91191 Gif-sur-Yvette Cedex, France.

<sup>2</sup> CEA, Laboratoires d'analyse bio-médicale LABM Centre de Cadarache, 13115 Saint-Paul Lez Durance, France.

<sup>3</sup> Société Automatismes et Mesures, Parc d'activité de Vausbesnard, 91410 Dourdan, France.

measurement result. In the second step, some examples of decision threshold and detection limit optimization in gamma spectrometry will be described. These examples will show that the essential parameter is the optimization of the estimation of background values, which requires the control of measurement rather than the management of statistical tools.

**Keywords:** Decision threshold / detection limits / spectrometry / measurement optimization

---

## 1. Introduction

L'amélioration des seuils de décision (SD) des analyses radiotoxicologiques, notamment en spectrométrie alpha, est une préoccupation des Laboratoires d'analyses de biologie médicale (LABM). Cette amélioration passe par la formation indispensable des opérateurs au concept de SD et peut être obtenue en optimisant les logiciels de calcul notamment en ce qui concerne la spectrométrie alpha des isotopes de l'uranium dans les urines. Ce sujet important de métrologie va être traité en deux parties complémentaires. L'objectif de la première partie est de revenir sur la problématique de base du seuil de décision et sur toutes les questions que l'on peut se poser à son sujet :

- Quand et comment apparaît la nécessité d'un seuil de décision ?
- Comment se détermine t-il ?
- Comment s'interprète t-il ?
- Quel est ensuite le rôle de la limite de détection par rapport au seuil de décision ?
- Quel sont les paramètres permettant d'améliorer et d'optimiser le seuil de décision ?

Ce premier article essaie de répondre le plus clairement possible à ces questions que se posent tous les laboratoires, questions auxquelles les normes actuelles, qui ne proposent souvent que des formules sans aucunes justifications, ne répondent que très peu. La seconde partie sera consacrée aux méthodes particulières d'optimisation des SD mises en œuvre au niveau des LABM dans le cas particulier de la recherche des isotopes de l'uranium par spectrométrie alpha.

## 2. Problématique des comptages à bas niveau

Soit un système de mesure constitué d'un détecteur associé à une simple échelle de comptage utilisé pour quantifier l'activité contenue dans divers échantillons. La procédure de comptage est la suivante : pour chaque échantillon deux mesures de durée égale sont réalisées : une mesure de bruit de fond d'une part et une mesure échantillon d'autre part. L'optimisation du partage des durées de comptages sera évoqué ultérieurement.

Notations :

- $BdF$  : la valeur du comptage de bruit de fond ;
- $Brut$  : la valeur du comptage brut ;
- $Net$  : la valeur du comptage net ;
- $T$  : la durée de comptage ;
- $R$  : le rendement de comptage défini par étalonnage suivant la relation

$$R = \frac{Net}{A_{\text{étalon}} T}, \text{ avec } A_{\text{étalon}} \text{ l'activité d'une source étalon.}$$

Pour des durées de comptages égaux nous avons :

$$Net = Brut - BdF. \quad (1.1)$$

À partir de la valeur du comptage net on calcule l'activité mesurée par :

$$A_{\text{mesurée}} = \frac{Net}{RT}. \quad (1.2)$$

Comme pour toute mesure il faut estimer ici l'incertitude sur l'activité moyenne vraie  $\tilde{A}$  estimée par l'activité mesurée  $A_{\text{mesurée}}$  obtenue. Cette incertitude est une incertitude composée, à partir de la relation (1.1), des incertitudes associées à chacun des comptages brut et bruit de fond. Ces incertitudes sont des incertitudes de type B (mesures uniques) et s'estiment grâce à l'hypothèse d'un comportement de type poissonien des comptages nucléaires bruts. On peut ainsi écrire dans cette hypothèse  $u_{BdF} = \sqrt{BdF}$  et  $u_{Brut} = \sqrt{Brut}$ . Ces incertitudes portent non pas sur les valeurs de mesures en soi, qui sont « certaines », mais sur la valeur des comptages moyens « vrais » estimés à partir de ces valeurs de comptage. Nous noterons ces valeurs moyennes vraies de comptage par  $\mu_{BdF}$  et  $\mu_{Brut}$ . Le comptage net, en tant que différence de deux variables de Poisson, ne possède plus le caractère poissonien mais l'on peut écrire, en appliquant à la relation (1.1) les règles élémentaires de compositions des incertitudes :

$$u_{Net} = \sqrt{u_{BdF}^2 + u_{Brut}^2} \Rightarrow u_{Net} = \sqrt{Brut + BdF}. \quad (1.3)$$

Cette incertitude porte là encore sur la valeur de la moyenne vraie des comptages nets définie par  $\mu_{Net} = \mu_{Brut} - \mu_{BdF}$ . Si l'on néglige l'incertitude sur le rendement<sup>4</sup>, alors l'incertitude finale sur l'activité mesurée est égale :

$$u_A = \frac{u_{Net}}{RT}. \quad (1.4)$$

<sup>4</sup> L'incertitude sur le rendement n'impacte pas le seuil de décision et n'impacte la limite de détection qu'à partir d'incertitude relative élevée, supérieure à 30 % (cf. § 4.4).

TABLEAU I

Exemple de valeurs numériques et incertitudes associées élargies à  $k = 2$ .  
 Numerical values exemple and associated expanded uncertainty with  $k = 2$ .

	Comptage (coups)	Incertitude élargie (coups)
BdF	46	$\pm 14$
Brut	64	$\pm 16$
Net	18	$\pm 21$
Activité mesurée	147 Bq	$\pm 171$ Bq

### Application numérique

Valeurs de comptage obtenues suivantes et incertitudes élargies associées avec un rendement de comptage  $R = 1,23 \%$ , un durée de comptage de 10 s et un facteur d'élargissement  $k = 2$  : voir tableau I.

L'incertitude relative sur l'activité élargie à un niveau de confiance de 95 % est égale ici à  $U\%_A = U_A/A_{mes} \approx 116 \%$ . Il est possible d'interpréter ce résultat en disant que l'activité moyenne vraie  $\tilde{A}$  est comprise dans l'intervalle de confiance  $[-24 \text{ Bq} ; 318 \text{ Bq}]$ . Ce résultat surprenant<sup>5</sup> pose entre autre le problème réel suivant : la valeur «  $\tilde{A} = 0$  » étant contenue dans cet intervalle, cela signifie qu'il est possible que l'échantillon ne contienne aucune radioactivité ! Ainsi une telle incertitude, plus grande que la valeur mesurée elle-même, pose naturellement une première question à laquelle il faut tenter de répondre avant de vouloir quantifier la radioactivité de l'échantillon : l'échantillon est-il radioactif ou non ? C'est l'objectif du seuil de décision associé à un risque alpha, ou risque de première espèce, de nous aider à répondre à cette question.

## 3. Seuil de décision dans le cadre des tests d'hypothèse

### 3.1. Construction du seuil de décision dans une approche a priori

La méthode des tests d'hypothèse élaborée dans les années 1930 par Neyman et Pearson est abondamment utilisée dans tous les domaines où des valeurs statistiques apparaissent en même temps qu'un doute sur leur significativité (Saporta, 1990 ; Kauffmann, 1994). Appliquée à notre problème, la méthode des tests d'hypothèse consiste à adopter la démarche suivante :

- (1) on se place volontairement dans l'hypothèse où l'on effectuerait, dans les mêmes conditions expérimentales que précédemment, la mesure d'un échantillon réellement non radioactif. C'est en cela que cette démarche, qui est une démarche

<sup>5</sup> Surprenant car une activité moyenne vraie ne peut être négative.

inductive de la cause – *échantillon non radioactif* – vers les effets – *valeurs expérimentales* – peut être appelée « *a priori* ». Cette hypothèse est appelée indistinctement *hypothèse préférentielle* ou *hypothèse nulle* et notée  $H_0$ . Dans le cas présent cette hypothèse s'écrit : «  $H_0 : \mu_{Net} = 0 (\pm 0)$  ». En se plaçant dans cette hypothèse  $H_0$ , le problème lié à l'incertitude n'existe plus :  $\mu_{Net}$  étant parfaitement déterminé, il n'y plus d'incertitude sur sa valeur. Pour autant subsiste le problème, non pas de l'incertitude, mais de la dispersion inévitable et aléatoire des valeurs de mesures ;

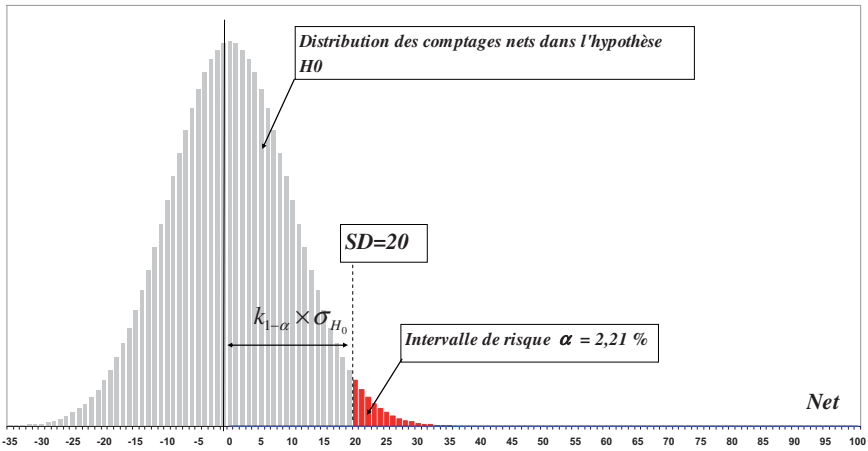
- (2) on évalue, à partir des informations connues, la distribution des valeurs des comptages nets obtenues que l'on pourrait obtenir sous cette hypothèse. La mesure d'un comptage net sous l'hypothèse  $H_0$  peut être modélisée comme la différence de deux réalisations indépendantes de la variable « bruit de fond » :  $Net_{[H_0]} = Bdf(2) - Bdf(1)$ , puisque sous  $H_0$  nous avons  $Brut_{[H_0]} = Bdf(2)$ . Cette nouvelle variable correspond à un bruit de fond net  $Bdf_{[Net]}$  dont les paramètres statistiques principaux, moyenne vraie et écart-type vrai, sont :
- moyenne vraie  $\mu_{H_0}$  : par construction, la variable  $Bdf_{[Net]}$  est centrée sur  $\mu_{H_0} = 0$  (d'où en fait le terme hypothèse « nulle ») ;
  - écart-type  $\sigma_{H_0}$  : on peut montrer que l'écart-type de la distribution  $H_0$  peut être estimé par  $\hat{\sigma}_{H_0} = \sqrt{2(Bdf_1 + 1)}$  (cf. Vivier et Aupiais, 2007).

C'est à partir de cette distribution  $H_0$ , dont on peut montrer par ailleurs qu'elle est de forme gaussienne (Vivier et Aupiais, 2007), évaluée à partir d'une mesure unique<sup>6</sup> de bruit de fond, qu'il devient possible de déterminer un seuil de décision. On peut noter ici que la construction de la distribution  $H_0$  peut dans la plupart des cas être déterminée expérimentalement par des mesures répétées du bruit de fond (mesure de  $\sigma_{H_0}$ ). Une telle méthode peut s'avérer plus simple mais surtout plus efficace que toute construction analytique, notamment si des facteurs de variabilité autres que poissonien interviennent.

*Remarque : le terme « +1 » sous la racine apparaît dans une détermination rigoureuse de la distribution  $H_0$  par une méthode bayésienne (Vivier et Aupiais, 2007). Ce terme est pour l'essentiel lié à l'incertitude résiduelle sur la valeur de l'écart-type de comptage de bruit de fond estimée à partir d'une mesure unique. Ce terme évite de plus d'obtenir un seuil de décision égal à zéro lorsque la mesure unique de bruit de fond est nulle, ce qui est courant en spectrométrie alpha.*

- (3) On construit un intervalle supérieur, appelé *intervalle de risque*  $\alpha$ , à partir d'une valeur particulière SD sur l'axe des comptages net respectant la condition  $P(Net > SD | H_0) = \alpha \%$  avec  $\alpha$  relativement petit, en général 2,5 % ou 5 %.

<sup>6</sup> Il est tout à fait possible d'effectuer plusieurs comptages de bruit de fond et de les prendre en compte dans la détermination du seuil de décision. Mais c'est là un autre sujet.



**Figure 1 – Exemple graphique : pour BdF = 46, SD = 20 pour un risque  $\alpha = 2,21$  %.**  
*Graphic example: for BdF = 46, SD = 20 for an  $\alpha$  risk equal to 2.21 %.*

La condition  $P(\text{Net} > SD \mid H_0) = \alpha \%$  se lit : « la probabilité que le comptage net soit supérieur à SD, sachant que l'échantillon est réellement non radioactif (hypothèse  $H_0$ ), est égale à  $\alpha \%$  ».

La valeur SD se détermine donc en unité d'écart-type par :

$$P(\text{Net} > SD = k_{1-\alpha} \sigma_{H_0} \mid H_0) = \alpha \% \Rightarrow SD = k_{1-\alpha} \sigma_{H_0} = k_{1-\alpha} \sqrt{2(BdF_1 + 1)}. \quad (3.1)$$

Le coefficient  $k_{1-\alpha}$  correspond au fractile de risque  $\alpha$  de la loi normale centrée réduite.

Il est important de noter que le seuil de décision n'est fonction que de la mesure de bruit de fond et non de la mesure échantillon (voir Fig. 1).

La définition  $P(\text{Net} > SD \mid H_0) = \alpha \%$  du risque  $\alpha$  sur laquelle s'appuie la construction du seuil de décision est une définition *a priori*, sous-entendu avant mesure échantillon, que l'on peut expliciter suivant deux interprétations équivalentes :

- interprétation statistique : « si l'on mesure  $n$  échantillons non radioactifs,  $\alpha \%$  d'entre eux seulement génèreront une valeur nette supérieure à SD » ;
- interprétation probabiliste : « on a a priori  $\alpha \%$  de chance d'obtenir une valeur supérieure à SD si l'on mesure un échantillon réellement non radioactif ».

Dans les deux cas cette approche permet de quantifier, dans une démarche *a priori*, le risque d'erreur  $\alpha$ , ou risque de première espèce, qui représente le risque<sup>7</sup> de déclarer « radioactif » un échantillon réellement non radioactif.

### 3.2. Mise en œuvre du seuil de décision et interprétation a posteriori

Nous venons de voir que la construction du seuil de décision s'appuie dans le cadre de l'hypothèse  $H_0$  sur une estimation du bruit de fond et sur un risque  $\alpha$  a priori. Lorsque le seuil de décision est déterminé, il reste alors à mettre en œuvre ce test après la mesure échantillon et à interpréter le résultat de ce test. Pour une mesure échantillon donnant une valeur de comptage net  $Net$  deux cas sont possibles.

#### 3.2.1. Premier cas possible : $Net > SD$

On décide alors de rejeter l'hypothèse  $H_0$  en s'appuyant sur le faible risque d'erreur *a priori* d'obtenir cette inégalité si l'échantillon était non radioactif. Rejeter l'hypothèse  $H_0$  revient à considérer la mesure comme significative d'une radioactivité avérée. L'estimation de la valeur vraie en termes de comptage ou activité, est alors tout à fait classique :

$$\begin{cases} \hat{\mu} = Net \pm k\sqrt{Brut + BdF} \\ \hat{A} = \frac{Net}{RT} \pm k\frac{\sqrt{Brut + BdF}}{RT} \end{cases}$$

#### Exemple numérique

En reprenant l'exemple avec  $BdF = 46$ , supposons que la mesure d'un échantillon donne la valeur  $Brut = 71$  soit  $Net = 25$ . Nous sommes dans le cas où  $Net = 25 > SD = 20$ . Nous pouvons considérer cette mesure comme significative et exprimer le résultat : « Comptage net significatif d'une radioactivité avérée avec  $\hat{\mu}_{Net} = 25 \pm 22$  pour un niveau de confiance de 95 % ».

- Pour autant cette démarche basée sur un raisonnement *a priori* ne permet pas :
- de donner la probabilité que l'échantillon soit non-radioactif lorsque  $Net > SD$  ;
  - de différencier en termes de risque d'erreur de décision un résultat légèrement supérieur au seuil de décision d'un résultat très supérieur au seuil de décision.

<sup>7</sup> En toute rigueur un risque se définit comme la probabilité d'occurrence de l'évènement redouté multiplié par les conséquences, sur une échelle donnée, de l'évènement s'il survient. Dans le domaine de la mesure le risque est assimilé à la seule probabilité de prendre la mauvaise décision.

Une approche bayésienne, c'est-à-dire *a posteriori*, ou si l'on préfère encore après la mesure échantillon, permet toutefois d'inverser les propositions et de donner une interprétation intéressante du résultat du test (Vivier et Aupiais, 2007) : « si  $Net > SD$  alors la probabilité que l'échantillon soit non-radioactif est inférieure à  $\alpha\%$   $\Leftrightarrow P(H_0 | Net > SD) \leq \alpha\%$  ».

De façon équivalente on peut dire aussi : « si  $Net > SD$  alors la probabilité que l'échantillon soit radioactif est supérieure à  $1 - \alpha\%$   $\Leftrightarrow P(\overline{H_0} | Net > SD) > 1 - \alpha\%$  »,  $\overline{H_0}$  étant l'hypothèse complémentaire de  $H_0$ , soit l'évènement « échantillon radioactif », toutes valeurs possibles confondues.

Dans la mesure où l'on prend en général des valeurs de risque  $\alpha$  faibles, entre 2,5 % et 5 %, alors l'expression précédente permet d'affirmer, lorsque  $Net > SD$ , que la probabilité que l'échantillon soit radioactif est supérieure à 95 %. Devant cette quasi-certitude on peut alors considérer que la valeur de mesure est significative et que la radioactivité de l'échantillon est avérée. Cette interprétation *a posteriori* permet ainsi à un opérateur de prendre cette décision – *échantillon radioactif* – avec plus d'assurance. Il est de plus possible, toujours dans une approche bayésienne, de déterminer ces probabilités en fonction de la valeur exacte du comptage net échantillon. Bien que dépassant le cadre de cet article et l'utilisation habituelle du seuil de décision, quelques constats apportent un éclairage intéressant : tout d'abord on peut montrer ce que l'intuition élémentaire permet de penser : plus la valeur nette est grande devant le seuil de décision et plus la probabilité que l'échantillon soit réellement radioactif est grande, et ainsi plus le risque de déclarer à tort « radioactif » un échantillon non-radioactif diminue.

Un exemple numérique qu'il est intéressant d'avoir à l'esprit est le suivant : en reprenant les valeurs numériques  $\{BdF = 46 ; \alpha = 2,2\%\} \Rightarrow SD = 20$ .

- Pour l'échantillon précédent avec  $Net = 25$ , valeur légèrement supérieure au seuil de décision, la probabilité d'être non radioactif est de 1,1 %.
- Pour un autre échantillon avec  $Net = 43$ , valeur de l'ordre de grandeur de la limite de détection (*voir chapitre suivant*), la probabilité d'être non radioactif n'est plus que de l'ordre de 0,1 %.

### 3.2.2. *Second cas possible : $Net < SD$* <sup>8</sup>

Il n'est plus possible d'écarter ici l'hypothèse  $H_0$  que l'échantillon soit non-radioactif, mais on ne peut pas pour autant rejeter l'hypothèse d'un échantillon

<sup>8</sup> Le cas rarissime  $Net = SD$  est à exclure car indécidable. Il faut continuer dans ce cas là la mesure, et le résultat finira par être supérieur ou inférieur au seuil de décision.



radioactif. Tout ce que l'on peut affirmer est que l'effet, en termes de comptage, d'une éventuelle radioactivité n'est pas détecté. Il faut bien comprendre ici qu'il n'est pas possible d'affirmer de façon absolue que la cause (*la radioactivité*) est absente parce que l'effet (*un comptage net significatif*) n'a pas été détecté. Cet effet peut être présent, mais masqué dans les fluctuations du bruit de fond. Le cas  $Net < SD$  est ainsi plus délicat à interpréter que le cas inverse. Si, par symétrie par rapport au cas précédent  $Net > SD$ , le résultat  $Net < SD$  était utilisé pour déclarer que l'échantillon est « selon toute vraisemblance » non-radioactif on prendrait alors le risque de commettre une erreur. En effet un échantillon de radioactivité donnée non nulle est susceptible, en raison de la fluctuation des mesures brutes et de bruit de fond, de générer malencontreusement un comptage net inférieur au seuil de décision. Cette situation conduirait à décider à tort qu'un échantillon est non radioactif alors qu'il est réellement radioactif. Ce risque d'erreur est appelé erreur de seconde espèce ou risque  $\beta$ . C'est ce risque d'erreur qui génère le concept de limite de détection.

## 4. Concept de limite de détection

### 4.1. Construction de la limite de détection

En fonction de ce qui a été dit précédemment la limite de détection est une grandeur secondaire que l'on n'utilise en toute rigueur que lorsque  $Net < SD$ . La détermination *a priori* et l'utilisation de la limite de détection s'appuie sur la démarche suivante :

- (1) on se place dans l'hypothèse alternative  $H_1$  où l'on postule que l'on mesure un échantillon radioactif dont on connaît parfaitement (*i.e. sans incertitude*) la valeur de la moyenne vraie  $\mu_{Net}$  non nulle des comptages nets qu'il génère. On note LD cette valeur. En résumé l'hypothèse  $H_1$  s'écrit : «  $\mu_{Net} = LD$  ». Cette valeur de comptage net moyen vrai sera ultérieurement reliée à l'activité moyenne vraie de l'échantillon par la relation  $\tilde{A} = \frac{\mu_{Net}}{RT}$  ;
- (2) on construit ensuite la distribution des valeurs nettes mesurées possibles d'un tel échantillon correspondant à l'hypothèse  $H_1$  » ;
- (3) on cherche la valeur de LD qui permette d'obtenir un risque  $\beta$  déterminé *a priori* avec  $P[Net < SD \mid H_1(LD)] = \beta\%$  (voir Fig. 2).

Pour une valeur de LD déterminée supérieure au seuil de décision on constate qu'une fraction  $\beta$  des comptages nets possibles, en queue de distribution à gauche, est inférieure au seuil de décision. De tels comptages amèneraient à conclure à tort que l'échantillon n'est pas radioactif, situation qui correspond bien au risque d'erreur de seconde espèce.

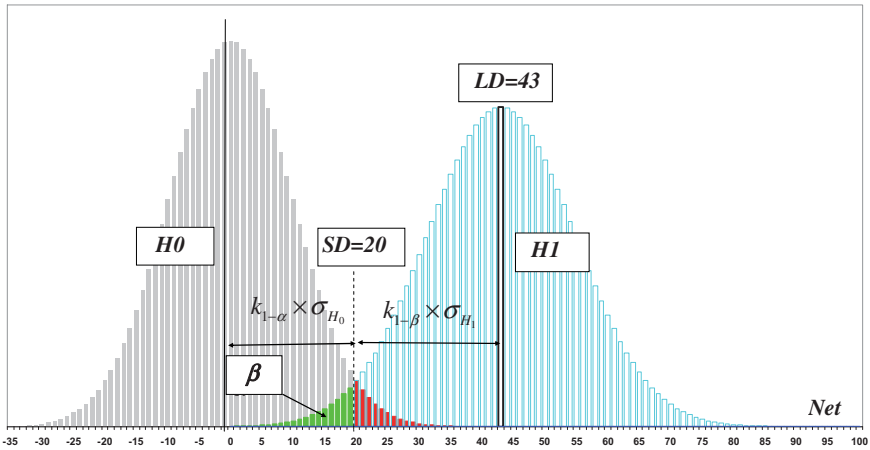


Figure 2 – Distribution hypothèses préférentielle  $H_0$  et alternative  $H_1$ .  
 $H_0$  preferential hypothesis and  $H_1$  alternative hypothesis distributions.

Pour des risques  $\alpha$  et  $\beta$  identiques (*risques symétriques cf. graphe*) la limite de détection est égale à :

$$LD_{\alpha = \beta} = k_{1-\beta}^2 + 2 SD^9. \quad (4.1)$$

#### 4.2. Utilisation de la limite de détection dans le cadre des tests d'hypothèse et interprétation a posteriori

Rappelons que la limite de détection n'est en toute rigueur utile que dans le cas  $Net < SD$ . La définition  $P[Net < SD | H_1 (LD)] = \beta\%$  de l'intervalle de risque  $\beta$  s'interprète de la façon suivante dans une démarche *a priori* :

- interprétation statistique : « si l'on mesure  $n$  échantillons radioactifs de moyenne vraie  $\mu_{net} = LD$ , alors  $\beta\%$  d'entre eux seulement génèreront une valeur nette inférieure à  $SD$  et seront à tort déclarés effets non détectés » ;
- interprétation probabiliste équivalente : « on a  $\beta\%$  de chance d'obtenir une valeur inférieure à  $SD$  si l'on mesure un échantillon radioactif de moyenne vraie  $\tilde{N}_{net} = LD$  ».

Nous avons vu que dans le cas  $Net > SD$  nous pouvions, au risque d'erreur  $\alpha$  près, rejeter l'hypothèse  $H_0$ , alors que dans le cas  $Net < SD$  il était assez difficile de se positionner par rapport à cette même hypothèse. Cependant on peut par raison de symétrie, lorsque  $Net < SD$ , rejeter selon toute vraisemblance, c'est-à-dire au

<sup>9</sup> Pour des valeurs de comptages suffisamment élevées cette formule converge vers la formule bien connue  $LD = 2 SD$ .

risque  $\beta$  près, l'hypothèse  $H_1$ . En effet comme pour le seuil de décision on peut interpréter, lorsque  $Net < SD$ , la limite de détection dans une démarche *a posteriori* (i.e. bayésienne) par :

- « si  $Net < SD$  alors la probabilité que la valeur vraie échantillon soit supérieure à  $LD$  est inférieure à  $\beta\%$  »

$$\Leftrightarrow P(\mu_{Net} \geq LD \mid Net < SD) \leq \beta\% \quad (4.2)$$

ou encore de façon équivalente :

- « si  $Net < SD$  alors la probabilité que la valeur vraie échantillon soit inférieure à  $LD$  est supérieure à  $1 - \beta\%$  »

$$\Leftrightarrow P(\mu_{Net} < LD \mid Net < SD) > 1 - \beta\% \quad (4.3)$$

Ces expressions montrent simplement que si un échantillon est mesuré inférieur au seuil de décision, alors sa valeur vraie  $\mu_{Net}$  a peu de chance de dépasser la limite de détection. Cette dernière doit ainsi être comprise non pas comme une limite basse de détection, au sens où l'on ne pourrait pas détecter de valeurs plus petites, mais comme une *limite basse de détection quasi-certaine*, au sens où, à partir de cette valeur et au-delà, la probabilité de détection est très élevée et à l'inverse la probabilité de non-détection très faible. Dans cette acceptation il n'y a plus de contradiction apparente au fait de détecter et de déclarer des valeurs inférieures à la limite de détection mais supérieure au seuil de décision. Ainsi, le résultat, lorsque  $Net < SD$ , doit s'exprimer de la façon suivante :

- aspect qualitatif : « effet non détecté » ou « radioactivité non avérée » ;
- aspect quantitatif :  $\mu_{Net} < LD$  avec un niveau de confiance égal à  $\beta\%$ .

Par symétrie par rapport au cas précédent, il est assez intuitif de comprendre que plus la valeur mesurée nette est petite devant le seuil de décision, plus la probabilité que la valeur vraie dépasse la valeur  $LD$  est faible. Citons quelques exemples calculées en mettant en œuvre une méthode bayésienne, en reprenant la valeur de référence  $BdF = 46$  soit  $SD = 20$  :

- pour un échantillon mesuré à  $Net = 17$ , la probabilité  $P(\mu_{Net} < LD)$  est égale à 99,3 % ;
- pour un échantillon mesuré à  $Net = 5$ , la probabilité  $P(\mu_{Net} < LD)$  est égale à 99,993 %.

### **Remarque importante**

La déclaration  $\mu_{Net} < LD$  correspond à la définition d'un intervalle de confiance, c'est-à-dire d'une incertitude, sur la valeur vraie. De fait l'intervalle  $[0 ; LD]$  représente l'intervalle de confiance au niveau  $1 - \beta\%$  de la valeur vraie  $\mu_{Net}$ . Cette remarque est importante dans la mesure où une certaine confusion peut apparaître

dans l'esprit des utilisateurs concernant la déclaration « < LD » lorsque le résultat expérimental est trouvé inférieur au seuil. Cette situation amène certains utilisateurs à remplacer dans ce cas là la déclaration « < LD » par « < SD ». À titre d'exemple, en reprenant l'exemple numérique précédent avec  $BdF = 46$ , soit  $SD \approx 20$  et  $LD \approx 43$  pour des risques  $\alpha$  et  $\beta$  de l'ordre de 2,5 % considérons :

- un premier échantillon générant un comptage  $Brut = 73$  soit  $Net = 27 > SD$ . On peut ainsi déclarer que la radioactivité dans l'échantillon est avérée et donner le résultat et son incertitude associée en termes de comptage net, soit  $27 \pm 22$  avec un niveau de confiance de 95 %.
- un second échantillon générant un comptage  $Brut = 61$  soit  $Net = 15 < SD$ . Dans ce cas là on déclare que la radioactivité n'est pas détectée en donnant sur un plan quantitatif « < 43 (LD) ».

Cette démarche peut déranger en donnant le sentiment de déclarer dans le second cas, alors que le comptage net est plus petit et ce malgré la présence du signe « inférieur à », une grandeur plus élevée (43) que dans le premier cas (27). Cette situation n'est pourtant en rien paradoxale si l'on compare ce qui est comparable. En effet l'expression du résultat pour le second cas, à savoir « < 43 », ne porte pas sur la valeur expérimentale<sup>10</sup>, mais sur la valeur vraie et doit se lire «  $\mu_{Net} < 43$  ». Cette expression doit être comparée par rapport au premier cas non pas à la valeur centrale (27) de l'intervalle de confiance, mais à sa borne supérieure. Ainsi dans le premier cas le résultat peut s'énoncer, pour ce qui est de la limite supérieure, par «  $\mu_{Net} < 49 (27 + 22)$  ». On voit ainsi que l'énoncé des résultats est cohérent avec les résultats expérimentaux.

### ***4.3. Utilisation de la limite de détection en tant que critère de performance ou de tolérance***

L'interprétation des résultats dans le cadre strict des tests d'hypothèse ne permet pas, lorsque le résultat est inférieur au seuil de décision, d'affirmer avec certitude que l'échantillon est non-radioactif. Ainsi pour le moment on ne peut répondre de façon stricte à la question initiale de savoir si l'échantillon est radioactif ou non. Cet état d'incertitude peut dans certains cas (gestion des effluents et des déchets, analyses biomédicales) être inacceptable et ingérable. Il faut alors renverser le problème et déterminer, en fonction de critères scientifiques et/ou réglementaires une valeur de tolérance, ou critère de performance, telle que si la radioactivité est déclarée inférieure à cette valeur alors on pourra considérer cette radioactivité comme négligeable et en tous les cas acceptable vis-à-vis des effets redoutés. Ainsi, pour un décideur, ce cas sera analogue, sur le plan de la gestion, au cas d'un

<sup>10</sup> Effectivement si  $Net < SD$  alors la relation  $Net < LD$  est vrai, mais triviale et sans intérêt.

échantillon strictement non radioactif. Cette valeur de tolérance doit alors être imposée a minima comme limite de détection du protocole de mesure, ce qui permet d'en déduire le seuil de décision que l'on doit obtenir. Cette démarche impose alors à rebours les conditions de mesures minimales, taux de comptage de bruit de fond, temps, rendement, etc., que l'on doit respecter dans le protocole de mesure pour assurer à termes cette limite de détection.

Ainsi en imposant une limite de détection égale à une telle valeur de tolérance, un échantillon pour lequel la valeur de mesure sera inférieure au seuil de décision pourra être considéré, ou tout au moins géré, comme un échantillon non-radioactif puisqu'en effet sa valeur vraie est inférieure, avec un niveau de confiance acceptable, à la valeur de tolérance imposée. À l'inverse nous avons vu que si la radioactivité vraie est supérieure à cette limite de détection imposée, alors cette radioactivité sera détectée de façon quasi-certaine, ce qui là encore ne peut que rassurer le décideur sur la pertinence des mesures.

#### ***4.4. Seuil de décision et limite de détection en terme d'activité***

Une décision prise en fonction du test «  $NET > SD ?$  » n'est pas modifiée si l'on convertit les valeurs nettes en termes d'activité. En effet :

$$Net > SD \Leftrightarrow \frac{Net}{RT} > \frac{SD}{RT}$$

avec  $\frac{Net}{RT} = A_{mes}$  (activité mesurée) et  $\frac{SD}{RT} = SD_{[Bq]}$  (seuil de décision en termes d'activité).

On peut noter qu'une éventuelle incertitude sur la valeur du rendement, même élevée, est sans effet sur le seuil de décision. En effet, le seuil ne dépend que du bruit de fond, qui lui ne dépend pas, pour un détecteur donné et sauf cas exceptionnels, du rendement de détection (*géométrie source et distance source-détecteur*). Comme pour le seuil de décision on peut tout à fait convertir la limite de détection en termes de comptage net en limite de détection en termes d'activité par :  $LD_{[Bq]} = LD_{[Net]}/RT$ .

Par contre une incertitude relative élevée (> 20 %) sur la valeur du rendement est susceptible de modifier la valeur de cette limite. Ce constat apparaît dans la formule donnant la limite de détection en termes d'activité compte tenu d'une

incertitude relative  $u\%_R$  sur le rendement :  $LD = \frac{k_{1-\beta}^2 + 2SD}{1 - u\%_R^2 \times k_{1-\beta}^2}$  (cf. Norme

ISO 11929, 2009).

Pour une incertitude relative de 30 % sur le rendement, la limite de détection est ainsi multipliée par un facteur 1,32 pour un risque  $\alpha$  de 5 % et par un facteur 1,56 pour un risque  $\alpha$  de 2,5 %.

## 5. Amélioration et optimisation du seuil de décision

Nous avons vu que la valeur du seuil de décision ne dépendait que de la valeur du bruit de fond. Une amélioration de la méthode d'estimation de ce bruit de fond permettra donc son optimisation. Nous allons détailler dans ce chapitre quelques axes d'amélioration du seuil de décision. Nous commencerons par traiter le cas des simples comptages où nous mettrons en évidence le principe général puis nous en verrons une application différente dans son aspect mais équivalente sur le fond en spectrométrie gamma.

### 5.1. Cas des comptages, influence des rapports de durées de comptage

Dans le chapitre précédent nous avons considéré le cas de durées de mesures allouées au bruit de fond et à l'échantillon identiques. Cependant rien ne s'oppose à ce que le temps  $T_0$  alloué à la mesure de bruit de fond soit différente du temps  $T_S$  alloué à la mesure échantillon. On notera  $BdF(T_0)$  la valeur du comptage unique durant  $T_0$  et  $n$  le rapport  $\frac{T_0}{T_S}$ . Dans ces conditions le comptage moyen durant  $T_S$

peut être estimé par :  $\overline{BdF}(T_S) = \frac{BdF(T_0)}{T_0} T_S = \frac{BdF(T_0)}{n}$ .

On peut alors montrer que dans ces conditions le seuil de décision est alors (Vivier et Aupiais, 2007) :

$$SD = k_{1-\alpha} \sqrt{\left(1 + \frac{1}{n}\right) \left(\overline{BdF}(T_S) + \frac{1}{n}\right)}. \quad (5.1)$$

#### Cas limites

Pour  $n = 1$  on retrouve bien évidemment l'expression  $SD = k_{1-\alpha} \sqrt{2(BdF(T_S) + 1)}$ .

Pour  $n$  suffisamment grand le seuil de décision tend vers la valeur :  $SD_{n \gg 1} \approx k_{1-\alpha} \sqrt{BdF(T_S)}$ , soit globalement un seuil de décision divisée par un facteur  $\sqrt{2} = 1,41$  par rapport au seuil de décision calculé avec un temps de mesure du bruit de fond égal à  $T_S$ .

**Exemple numérique**

Si l'on reprend l'exemple précédent avec  $T_S = T_0$ ,  $BdF_1 = 46$  on obtenait  $SD = 20$  pour un risque  $\alpha$  de 2,2 %. Si l'on accorde une durée de comptage bruit de fond supérieure à la durée de comptage échantillon, par exemple  $T_0 = 5T_S$ , avec

$$\overline{BdF}(T_S) \approx 46^{11} \text{ le seuil de décision devient } SD \approx 2 \times \sqrt{\left(1 + \frac{1}{5}\right)} \times 46 \approx 15.$$

**Remarque**

La formule (5.1) est valable lorsque  $n$  est inférieur à 1, dans le cas où la durée de comptage bruit de fond est inférieur à la durée de comptage échantillon. On constate alors que le seuil de décision est dégradé par rapport au cas  $n = 1$ . Ainsi pour une durée de comptage de bruit de fond égal au quart du temps alloué à la mesure échantillon, le seuil de décision est égal à  $SD \approx 2 \times \sqrt{(1 + 4)} \times 46 \approx 30$ .

Cet exemple illustre parfaitement le principe que l'on retrouve dans toutes les techniques de mesures : si le bruit de fond est mieux estimé, alors le seuil de décision est plus petit.

**5.2. Optimisation du rapport des durées de comptages**

Le problème que l'on peut se poser ici est celui de l'optimisation du temps  $T_0$  accordé à la mesure du bruit de fond et du temps  $T_S$  accordé à la mesure des échantillons, sous la contrainte d'un temps disponible total  $T_{total}$  déterminé. Considérons le cas général où l'on est amené à mesurer  $k$  échantillons,  $k$  étant un nombre prédéterminé : voir figure 3.

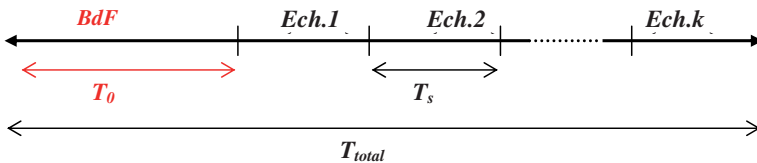


Figure 3 – Schéma de partage des durées de mesures.  
Sharing measurement duration scheme.

<sup>11</sup> Le comptage moyen de bruit de fond ramené à une durée  $T_S$  reste en moyenne indépendant du temps de mesure du bruit de fond  $T_0$ .

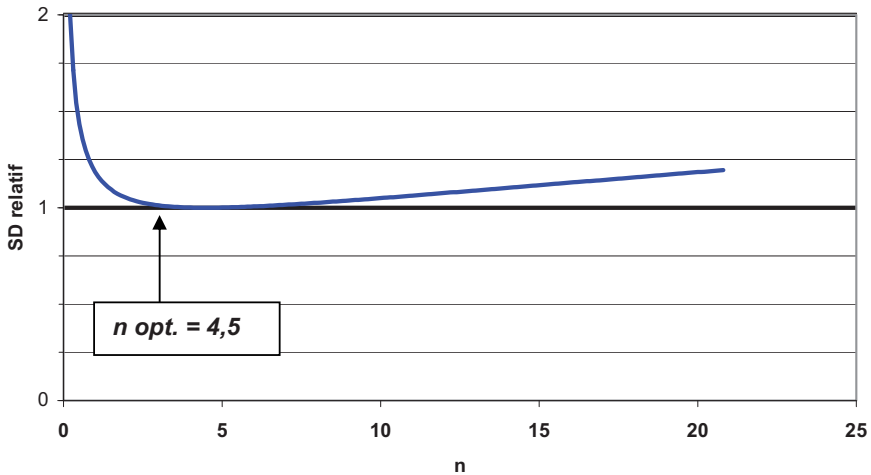


Figure 4 – *Optimisation des durées de comptages sous contrainte de temps total (k = 20).*  
*Optimization of time counting under total time constraint (k = 20).*

On montre alors que l’expression du seuil de décision pour chaque échantillon s’écrit :

$$SD_{[Bq]} = \frac{2k_{1-\alpha}\sqrt{r_{BdF}}}{R\sqrt{T_{total}}} \sqrt{\left(1 + \frac{1}{n}\right)(n+k)} \quad (5.2)$$

avec  $r_{BdF}$  le taux de comptage moyen de bruit de fond. On peut montrer que cette fonction possède une valeur qui minimise le seuil de décision. Cette valeur correspond à la valeur qui annule la dérivée :  $\frac{dSD(n)}{dn} = 0 \Leftrightarrow n = \sqrt{k}$ . On obtient alors un partage optimal des temps avec :

$$\begin{cases} T_{S\ opt} = \frac{1}{\sqrt{k} + k} T_{total} \\ T_{0\ opt} = \frac{\sqrt{k}}{\sqrt{k} + k} T_{total} \end{cases} \quad (5.3)$$

**Applications numériques (voir Fig. 4)**

1<sup>er</sup> cas :  $k = 1 \Rightarrow n = 1 \Rightarrow T_S = T_0 = \frac{T_{total}}{2}$  ;

2<sup>e</sup> cas :  $k = 20 \Rightarrow n = \sqrt{20} = 4,5 \Rightarrow \begin{cases} T_{S\ opt} = 4,1 \% \times T_{total} \\ T_{0\ opt} = 18,3 \% \times T_{total} \end{cases}$  .



La formule (5.2) montre l'influence de la durée de comptage dans le cas général : le seuil de décision en termes d'activité diminue comme l'inverse de la racine carrée du temps de comptage. Pour gagner un facteur 2 il faut ainsi compter 4 fois plus longtemps, ce qui peut être rédhibitoire dans certain cas, par exemple en analyse biomédicale avec des temps de comptages de l'ordre de 3 jours. On comprend alors l'intérêt d'optimiser le seuil de décision « toutes choses égales par ailleurs » (*dure de comptage, rendement, taux de bruit de fond...*) par de telles méthodes.

### 5.3. Cas de la spectrométrie gamma, influence du marquage des régions d'intérêt

#### 5.3.1. Optimisation du marquage de la région d'intérêt (ROI)

Nous considérerons ici le cas simple d'un pic sans interférence et d'un bruit de fond uniforme dans la région du pic considéré. La largeur de la ROI sera prise égale à  $b$  (en nombre de canaux). La présence du pic, largement supérieure au seuil de décision, n'est ici que pour visualiser la résolution (LTMH<sup>12</sup>) et n'est pas en toute rigueur indispensable. Le seuil de décision en termes de surface nette correspondant à ce schéma simplifié est proportionnel à la quantité  $\sqrt{BdF_{trap}}$ , avec  $BdF_{trap}$  le bruit de fond estimé sous pic par la méthode du trapèze. Or  $BdF_{trap} \approx bB$  ; où  $B$  est la valeur moyenne du bruit de fond par canal à gauche et à droite du pic. On constate donc que  $SD_{[surface\ nette]} \propto \sqrt{bB}$ . Une trop grande valeur de  $b$  aura donc tendance à augmenter le bruit de fond estimé et donc le seuil de décision (cf. Fig. 5).

À l'inverse une trop grande diminution de  $b$  aura tendance, en dessous d'une certaine valeur, à faire là encore augmenter le seuil de décision, à la condition que celui-ci soit exprimé en termes d'activité. Cet effet est lié pour l'essentiel à la diminution du rendement, comme on peut le voir sur un marquage différent du spectre précédent (cf. Fig. 6).

Le rendement  $R = Net/A_{\text{étalon}} I_{\gamma} T$  diminue et le seuil de décision en termes d'activité  $SD_{[Bq]} = SD_{[NET]}/R I_{\gamma} T$  augmente à nouveau lorsque  $b$  diminue. Il existe donc un optimum pour la largeur de la ROI résultant de ces deux effets inverses. Le calcul montre que lorsque le pic est gaussien la valeur optimale pour le marquage de la ROI est tel que  $b = 2,5$  LTMH, (cf. Figs. 7 et 8). Dans la pratique, et compte-tenu de la discrétisation des spectres et des effets de traîne, il est conseillé de prendre  $b \approx 3$  LTMH.

<sup>12</sup> La LTMH, ou largeur totale à mi-hauteur, est encore appelé « résolution ». Cette valeur, fonction croissante de l'énergie, est en général connue après étalonnage en résolution du spectromètre.

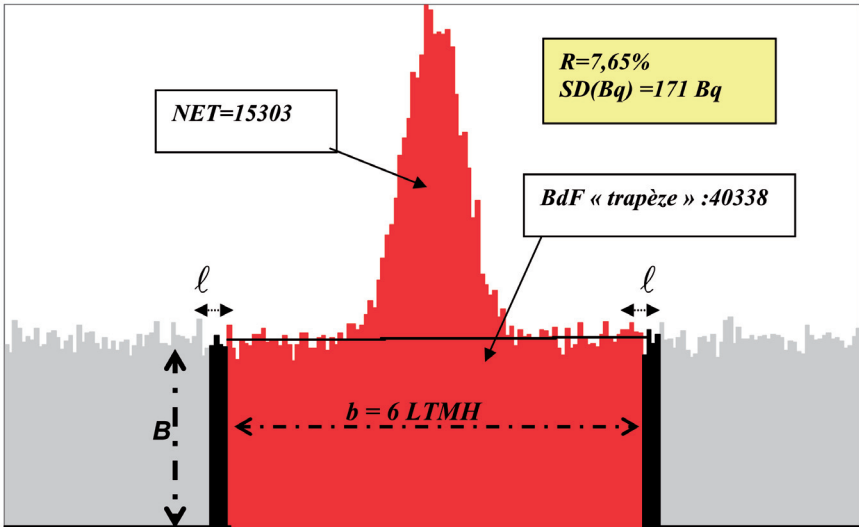


Figure 5 – Région d'intérêt non optimale – trop large.  
 Not optimal area of interest – too large.

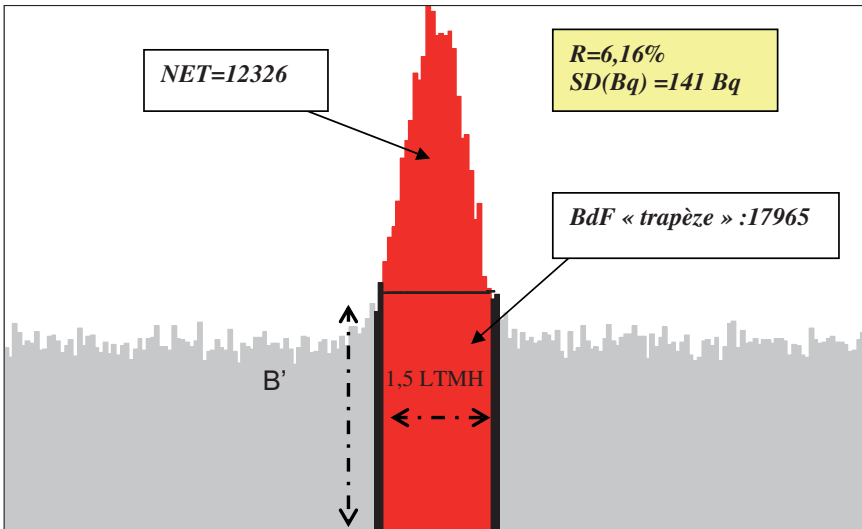


Figure 6 – Région d'intérêt non optimale – trop étroite.  
 Not optimale area of interest – too narrow.

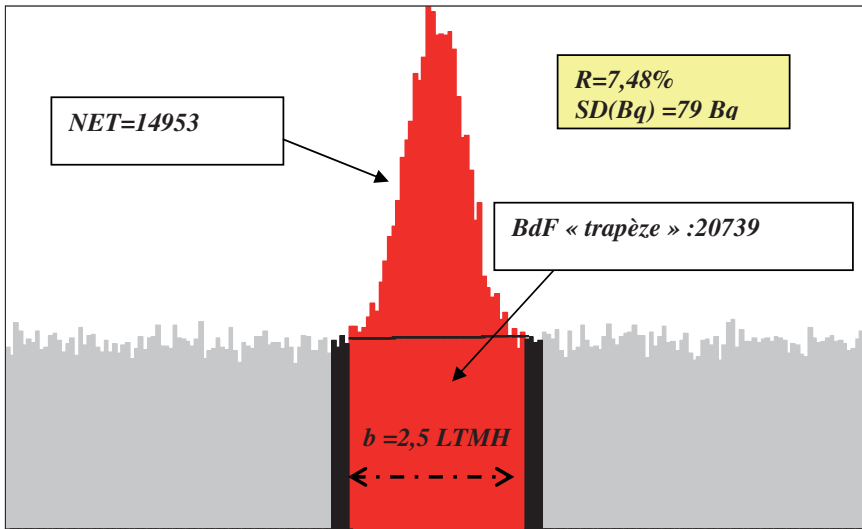


Figure 7 – Région d'intérêt optimale (2,5 LTMH).  
Optimal area of interest (2.5 FWHM).

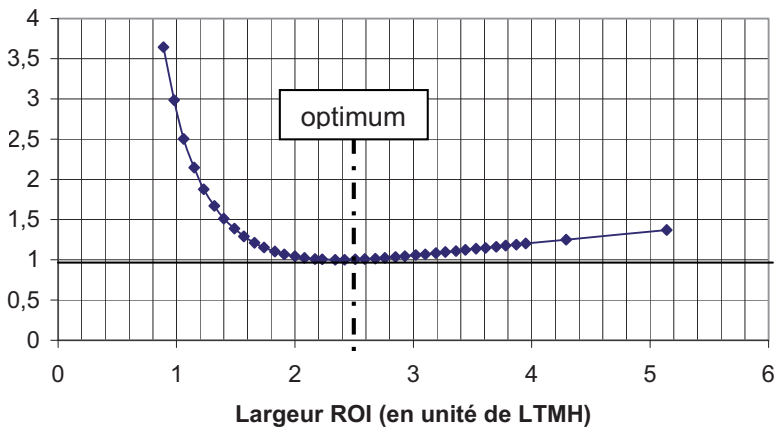


Figure 8 – Optimisation largeur de la région d'intérêt vs. LTMH.  
Area of interest width optimisation vs. FWHM.

### 5.3.2. Influence du marquage bruit de fond

Dans le problème du seuil de décision évalué directement sur le spectre échantillon, il n'est pas possible de jouer sur des durées de comptages différents<sup>13</sup> dans la mesure où le choix est fait d'estimer le bruit de fond d'une région d'intérêt (ROI) à partir du spectre échantillon<sup>14</sup>. Il n'y a donc dans ce cas qu'une seule mesure effective et implicitement des temps égaux. Cependant on peut retrouver une formule analogue à l'expression (5.1) en s'intéressant aux marquages utilisés pour estimer le bruit de fond.

Si l'on prend le cas de la spectrométrie gamma, où les pics sont relativement symétriques, on sait que l'on peut utiliser la méthode du trapèze pour estimer le bruit de fond « sous pics ». Le trapèze s'appuie à gauche et à droite du pic sur des valeurs de bruits de fond calculées sur un certain nombre de canaux. Ainsi cette estimation sera d'autant plus fiable que le nombre de canaux utilisés sera grand.

On retrouve ainsi en spectrométrie gamma la formule (5.1) avec ici  $n = 2\ell/b$  où  $\ell$  est égal au nombre de canaux à gauche et à droite utilisés pour construire le trapèze et  $b$  le nombre de canaux utilisés pour définir la ROI. On peut montrer en effet par un calcul de variance que le rapport  $n = 2\ell/b$ , rapport du nombre de canaux utilisés pour estimer le bruit de fond à gauche et à droite de la ROI au nombre de canaux de la ROI intervient dans le calcul du seuil de décision suivant la formule :

$$SD_{[Bq]} = k_{1-\alpha} \sqrt{\left(1 + \frac{1}{n}\right) B d F_{[trapèze]}}. \quad (5.4)$$

Le cas « classique »  $n = 1$  correspond au cas où la largeur totale utilisée pour estimer le bruit de fond est égale à la largeur de la ROI. Le cas où l'on ne prendrait par exemple qu'un seul canal à gauche et à droite pour estimer le bruit de fond, avec une largeur de ROI égale à 10 canaux soit  $n = 0,2$ , dégraderait fortement le seuil de décision avec  $SD_{n=0,2} = k_{1-\alpha} \sqrt{6 B d F(T_S)}$ . À l'inverse et lorsque l'allure du spectre le permet, en prenant une largeur  $2\ell$  plus grande que la largeur  $b$  de ROI telle que  $n \approx 10$ , alors pour le même spectre on obtiendrait  $SD_{n=10} \approx k_{1-\alpha} \sqrt{B d F(T_S)}$ . Le rapport entre la méthode avec  $n = 10$  et  $n = 0,2$  est alors égal à  $\frac{SD(n=0,2)}{SD(n=10)} \approx \sqrt{6} \approx 2,4$ .

<sup>13</sup> Ne pas confondre avec l'augmentation du durée de comptage qui lui aussi permet de baisser le seuil de décision en terme d'activité.

<sup>14</sup> Méthode appelée parfois « seuil de décision *a posteriori* », à ne pas confondre avec l'utilisation *a posteriori* du seuil décrite au chapitre 3.3.

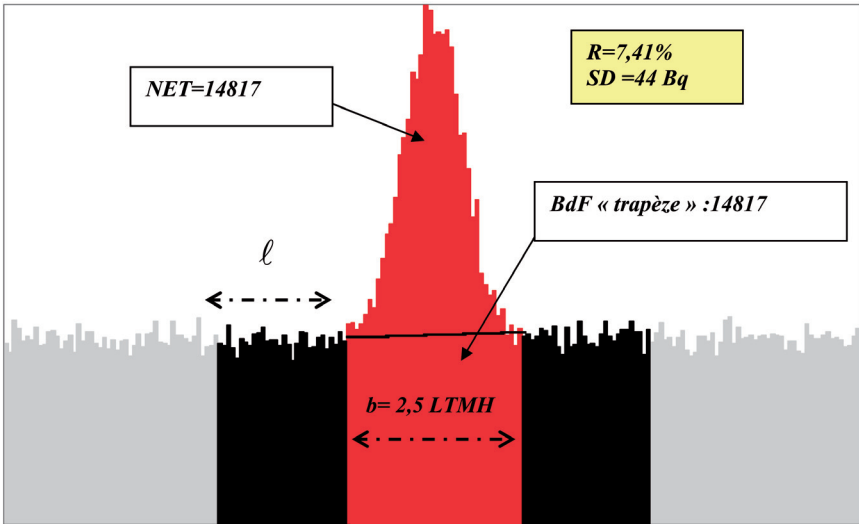
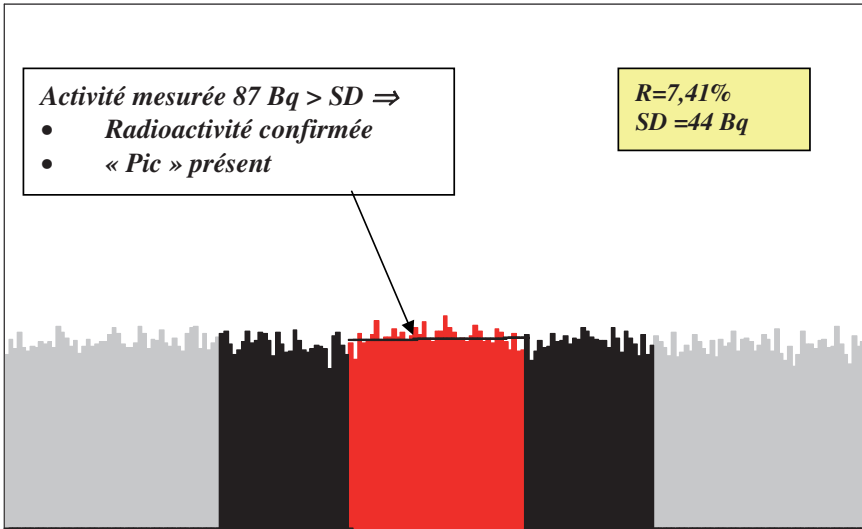


Figure 9 – Région d'intérêt optimale et meilleure estimation du bruit de fond ( $n = 1,5$ ).  
*Optimal area of interest and best estimate of background ( $n = 1.5$ ).*

Ces exemples montrent donc bien qu'il est possible de faire baisser le seuil de décision en améliorant la qualité d'estimation du bruit de fond sur un même spectre. C'est ce type de méthode qui a été mis en application à divers niveaux dans le sujet développé en seconde partie de cet article portant sur la spectrométrie alpha.

La mise en œuvre d'une part de l'optimisation du marquage ROI (§ 5.3.1.) et d'autre part du marquage le plus large possible à gauche et à droite du pic pour l'estimation du bruit de fond sous pic permet de réduire considérablement le seuil de décision en terme d'activité. Pour les figures précédentes nous avons  $n = 0,1$ . En prenant  $n = 1,5$  on obtient une diminution significative du seuil de décision : voir figure 9.

L'optimisation des marquages permet ainsi de diminuer dans cet exemple, pour un même spectre, le seuil de décision d'un facteur 3,8 entre le marquage figure 5 et le marquage figure 9, ce qui est considérable. Ces conditions sont ainsi optimales et représentent un moyen de recherche de pic très sensible, ainsi que le montre la figure 10 suivante. On notera que cette validation ne repose pas sur une recherche de forme, pour laquelle les algorithmes classiques seraient ici en défaut.



**Figure 10** – *Mesure échantillon positive avec marquage optimisé.*  
*Positive sample measurement with optimized marking.*

## 6. Conclusion

La nécessité de déterminer et d'utiliser un seuil de décision apparaît lorsque les fluctuations de bruit de fond rendent le résultat trop incertain et font ainsi apparaître un doute sur la présence même de radioactivité dans l'échantillon mesuré. Nous avons vu que ce seuil de décision permet, comme son nom l'indique, de prendre une décision dans un sens ou dans l'autre sans état d'âme. Mais il faut garder à l'esprit que ces décisions restent entachées d'un risque d'erreur. L'intérêt des tests d'hypothèse est de pouvoir quantifier de façon objective ces risques d'erreur, et ainsi décider en toute connaissance de cause. Nous avons vu que le seuil en terme de comptage net n'est fonction que de la valeur de bruit de fond et de la qualité de la méthode d'estimation de ce bruit de fond. Il importe alors de s'assurer que cette estimation est correcte, à savoir non biaisée et de variance minimale. Relativement simple dans le cas des comptages, ce type d'estimation peut devenir assez délicat comme par exemple en spectrométrie alpha à bas niveau. Ainsi la problématique du seuil de décision relève plus d'une compétence d'expérimentateur que d'une compétence de statisticien.

## RÉFÉRENCES

- Kauffmann P. (1994) *Statistique, information, estimation, tests*. Dunod, Paris.
- Norme ISO 11929 (2009) Determination of the characteristic limits (decision threshold, detection limit and limit of the confidence interval) for measurements of ionizing radiations.
- Saporta G. (1990) *Probabilités analyse de données et statistique*. Technip, Paris.
- Vivier A., Aupiais J. (2007) Optimization of the decision threshold for single radioactive counting, *Radiochim. Acta* **95**, 477-492.