B I P M

F-92310 SEVRES

Rapport BIPM-80/7

Statistiques de comptage

Jörg W. Müller

L'élaboration ou le perfectionnement de nouvelles méthodes de mesure est d'actualité permanente et l'on donnera d'abord quelques indications concernant la mesure de la période d'un radionucléide en tirant profit de la perturbation que la décroissance exerce sur la répartition des comptages. D'autre part, divers problèmes liés à la présentation d'incertitudes expérimentales ont aussi retenu notre attention, et il en sera question dans la seconde partie. Quelques aspects pourraient jouer un rôle dans la réunion d'experts organisée par le BIPM pour l'automne prochain, où l'on essaiera de trouver une solution nouvelle au vieux problème de l'expression de l'incertitude d'un résultat de mesure.

Perturbation des moments de comptage par la décroissance et le temps mort

Il est maintenant bien connu et confirmé par l'expérience, que la simple loi de Poisson est à modifier si l'on veut décrire la statistique d'impulsions provenant d'une source qui décroît de manière non négligeable pendant le temps d'observation. On a établi précédemment (Rapport BIPM-79/11) la loi modifiée pour la probabilité ${}_{\lambda}P(k)$ d'observer k impulsions pendant un intervalle de durée t_o pris parmi un grand nombre d'intervalles égaux qui se succèdent sur une période totale T, pour un radionucléide de vie moyenne $1/\lambda$.

Une estimation fiable de λ à partir de telles mesures demande que l'on tienne compte de plusieurs complications d'origine expérimentale, comme l'existence d'un mouvement propre (qui n'est pas nécessairement constant) et en particulier l'effet d'un temps mort pour les taux de comptage élevés. Leur influence sur la répartition des comptages est trop compliquée pour être incorporée de façon exacte dans les formules ; il est cependant possible d'en tenir compte par des calculs numériques faits à l'ordinateur en utilisant un procédé qui s'approche d'une simulation. Pour le cas fréquent de perturbations faibles par le temps mort, on peut se servir d'une méthode approchée qui est beaucoup plus rapide.

D'autre part, on a montré que l'utilisation des deux premiers moments permet une détermination simple et sensible de la vie moyenne de la source, en particulier dans le cas d'une déformation importante de la distribution initiale de Poisson. Il arrive que les moments en question peuvent être calculés de façon rigoureuse en tenant compte de tous les effets expérimentaux.

Si le taux de comptage moyen dû à la source et au mouvement propre (taux p) suit la loi

$$r_t = \rho_o \cdot e^{-\lambda t} + p,$$

on a après introduction d'un temps mort τ, pour $\lambda\tau \ll 1$,

$$\rho(t) = \frac{r_t}{1 + r_t \tau} \quad \text{ou} \quad \rho(t) = r_t \cdot \exp(-r_t \tau),$$

suivant que le temps mort est du type non cumulatif ou cumulatif. Pour $t = 0$, le taux de comptage non perturbé est ρ_o.

L'évaluation des expressions de l'espérance mathématique (E) et de la variance (V) du nombre k d'impulsions enregistrées pendant t_o a été faite, pour le cas général et les deux types de temps mort, dans les notes BIPM WPN-212 et 213. Pour simplifier l'écriture, nous nous contentons des formules valables en l'absence de mouvement propre ($p = 0$). Avec les notations

$$\theta = \lambda T, \quad b = \rho_o \tau, \quad B = b\ \mathrm{e}^{-\theta}, \quad c = 1 + b \text{ et } C = 1 + B,$$

on arrive aux expressions suivantes pour les valeurs moyennes et les variances

- pour un temps mort non cumulatif :

$$E(k) = \frac{\mu_o}{b\theta} (\ln c - \ln C),$$

$$_\lambda V(k) = \frac{\mu_o}{b} \left[{}_\lambda E(k) + \frac{\mu_o}{b\theta} \left(\frac{1}{c} - \frac{1}{C}\right) - \frac{1}{2\theta} \left(\frac{1}{c^2} - \frac{1}{C^2}\right) \right] - {}_\lambda E^2(k) \ ;$$

- pour un temps mort cumulatif :

$$_\lambda E(k) = \frac{\mu_o}{b\theta} (\mathrm{e}^{-B} - \mathrm{e}^{-b}),$$

$$_\lambda V(k) = {}_\lambda E(k) - {}_\lambda E^2(k) + \frac{\mu_o(\mu_o - 2b)}{4b^2\theta} \left[(1 + 2B) \cdot \mathrm{e}^{-2B} - (1 + 2b) \cdot \mathrm{e}^{-2b} \right],$$

toujours avec θ et $\tau \neq 0$.

En pratique, la détermination de la vie moyenne, et donc de la période $T_{1/2} = (1/\lambda) \cdot \ln 2$, peut se faire comme suit. L'expérience donne des valeurs numériques pour l'espérance E_{exp} et pour la variance V_{exp}. La valeur de τ et son type sont connus par des mesures indépendantes. A l'aide d'un programme d'ordinateur on détermine maintenant le taux initial ρ_o qui, pour une période admise (donc θ connu), donne pour l'espérance théorique E_{th} (en utilisant la formule appropriée) une valeur qui coïncide avec le résultat E_{exp}. La valeur correspondante pour la variance V_{th}, obtenue avec la formule théorique, s'écarte plus ou moins du résultat expérimental V_{exp}. Par un procédé itératif, la vie moyenne est alors ajustée jusqu'à ce que les valeurs expérimentales et théoriques coïncident pour E et pour V.

Pour une distribution empirique de comptage, tout le calcul nécessaire se fait automatiquement à l'aide d'un programme d'ordinateur qui donne en même temps une estimation de l'incertitude qu'il convient d'attribuer à la valeur ainsi trouvée pour la période d'un radionucléide.

Une publication sur l'ensemble des problèmes concernant la statistique de comptage d'une source décroissante est en préparation.

Détermination d'une distribution a priori

Dans les discussions sur la manière à adopter pour indiquer les incertitudes de résultats expérimentaux, il y a des problèmes controversés qui reviennent plus ou moins périodiquement ; après des échanges de vue généralement longs mais peu concluants, ils disparaissent de nouveau, tout en laissant une impression peu encourageante. La notion d'incertitude systématique, et en particulier le problème de la distribution de probabilité qu'il convient de lui associer, en fait partie. C'est pourquoi il nous semble utile d'indiquer qu'un progrès réel a été fait ces dernières années pour un problème restreint, certes, mais fondamental. Il s'agit de savoir comment choisir la densité de probabilité pour un paramètre sur lequel on ne sait rien.

Le problème qui consiste à décrire notre ignorance apparaît sous une forme particulièrement nette dans une classe de problèmes qui est liée au nom de Bayes. Il s'agit de combiner l'information provenant d'un nouveau résultat de mesure avec nos connaissances antérieures, pour en déduire de façon logique la distribution qui correspond à l'ensemble du savoir, ancien et nouveau. En particulier, il y a le cas de l'ignorance totale, et la détermination de la densité correspondante a été pendant longtemps considérée comme un problème sans solution. Dans cette situation, quelques-uns ont fait appel au "postulat" de Bayes (dont il n'est pas responsable) qui énonce que l'ignorance complète équivaut à l'attribution de probabilités égales pour tous les cas possibles. A vrai dire, cette règle est trop simpliste et elle conduit vite à des contradictions. Néanmoins, elle a aussi été utilisée dans le domaine des incertitudes systématiques où l'on ne dispose que de limites choisies de manière à ce qu'elles encadrent, avec une probabilité proche de l'unité, la valeur "vraie" dont l'emplacement à l'intérieur du domaine admis est inconnu, ce qui se traduit alors par une densité de forme rectangulaire. Or, ce raisonnement n'a jamais rencontré un accord unanime.

La notion d'ignorance totale a généralement été jugée trop vague pour être d'utilité pratique. Dans une étude tout à fait remarquable (bien que largement ignorée) le physicien Jaynes (1) montre, à l'aide d'exemples concrets, que ce n'est pas le cas. Une étude critique de ce qu'on entend par absence complète de connaissance et qui tire profit de toutes les symétries cachées établit que la densité de probabilité correspondante peut être bien définie. Prenons comme illustration le cas de la loi binomiale qui est donnée par

$$P(k,n|\Theta) = \binom{n}{k} \Theta^k (1 - \Theta)^{n - k}, \quad \text{avec} \quad 0 \leqslant k \leqslant n,$$

où l'on considère la probabilité Θ comme la grandeur recherchée. Désignons sa densité *a priori* par $f(\Theta)$.

---

(1) JAYNES (E.T.), Prior probabilities, *IEEE Trans. Syst. Sci. and Cyb.*, SSC-4, 1968, pp. 227-241.

L'application du théorème de Bayes pour la densité *a posteriori* $g(\Theta)$, qui tient aussi compte d'une expérience ayant donné comme résultat k "succès" pour n "possibilités", nous amène à

$$g(\Theta) = \frac{P(k,n|\Theta)\cdot f(\Theta)}{\int_0^1 P(k,n|\Theta)\cdot f(\Theta)\, d\Theta} .$$

Le résultat surprenant de Jaynes consiste à prouver que la fonction $f(\Theta)$ qui correspond à notre ignorance totale de Θ est donnée sans ambiguïté par la forme

$$f(\Theta) = \frac{\text{const.}}{\Theta(1-\Theta)} , \text{ pour } 0 < \Theta < 1,$$

qui est très différente d'un rectangle.

Il est vrai qu'il s'agit d'une densité "impropre", puisque non normalisable, mais cela n'a pas de conséquences fâcheuses. L'évaluation de $g(\Theta)$ conduit maintenant à une fonction bêta du type

$$g(\Theta) = A \cdot \Theta^{k-1} (1 - \Theta)^{n-k-1},$$

où $A = \binom{n-1}{k-1} (n-k)$ est un facteur de normalisation.

En utilisant les moments connus de la distribution bêta, on a pour l'espérance (E) et la variance (V) du paramètre recherché

$$E(\Theta) = \frac{k}{n} \quad \text{et} \quad V(\Theta) = \frac{k\ (n-k)}{n^2 (n+1)} .$$

Ces résultats correspondent bien à ce qu'on pourrait attendre. Mentionnons, d'autre part, que l'hypothèse d'une densité rectangulaire pour Θ (entre 0 et 1) nous amène aux expressions

$$E(\Theta) = \frac{k+1}{n+2} \quad \text{et} \quad V(\Theta) = \frac{(k+1)\ (n-k+1)}{(n+2)^2\ (n+3)} ,$$

qui étaient toujours difficiles à interpréter.
Pour plus de détails techniques et un autre exemple qui concerne la loi de Poisson, voir le Rapport BIPM-80/6.

On peut déduire de ces résultats que l'identification de l'ignorance avec une densité de probabilité constante pour le domaine en question est un procédé mal justifié et dangereux. Son emploi dans le contexte des erreurs dites systématiques est donc à éviter.

Les réponses de 21 laboratoires nationaux que le BIPM a reçues à son "Questionnaire sur les incertitudes" ont fait l'objet d'une brève analyse (Rapport BIPM-80/3). Quoiqu'il n'en découle pas pour le moment l'impression d'un rapprochement des opinions, parfois assez divergentes, cet exercice n'était certainement pas inutile. L'étude des réponses permettra à chaque participant d'être mieux informé sur l'ensemble des problèmes soulevés et, le cas échéant, de réfléchir à nouveau sur le bien-fondé de son attitude propre.

Pour une approche nouvelle qui, cependant, ne représente pas nécessairement l'opinion du BIPM, voir aussi notre récente petite étude "Les incertitudes de mesures".

Autres travaux

Parmi les autres travaux statistiques, on peut mentionner une étude assez détaillée (Rapport BIPM-80/4) qui traite des déformations apportées à la loi de Poisson si celle-ci est perturbée par un temps mort cumulatif et pour le cas d'un compteur libre. On trouve des expressions générales pour les moments en utilisant des nombres de Stirling de deuxième espèce.

Dans l'évaluation numérique des moments empiriques pour une longue série de mesures, l'inclusion d'un nouveau résultat nécessite en général de refaire tous les calculs. Pour éviter cela, on a établi des expressions qui permettent de voir directement l'influence d'une nouvelle mesure sur les trois premiers moments (BIPM WPN-214). Dans une autre note, on traite le problème analogue pour le cas où deux échantillons, à moments connus, sont réunis en un seul (BIPM WPN-215).

(Octobre 1980)