

# REVUE DE STATISTIQUE APPLIQUÉE

GÉRARD BRUNSCHWIG

## Les méthodes bayésiennes en statistique

*Revue de statistique appliquée*, tome 29, n° 1 (1981), p. 57-68

[http://www.numdam.org/item?id=RSA\\_1981\\_\\_29\\_1\\_57\\_0](http://www.numdam.org/item?id=RSA_1981__29_1_57_0)

© Société française de statistique, 1981, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

# LES METHODES BAYESIENNES EN STATISTIQUE

Gérard BRUNSCHWIG

Ingénieur en chef des Ponts et Chaussées  
Chargé de mission auprès du directeur du Laboratoire central

## INTRODUCTION

Les méthodes statistiques dites *classiques* reposent essentiellement sur la collecte et le traitement de résultats expérimentaux, dans des buts variés parmi lesquels les plus couramment rencontrés sont :

- l'estimation, soit ponctuelle, soit par intervalle de confiance (par exemple : étant donné  $n$  résultats de mesure d'un caractère, estimer la valeur moyenne du caractère) ;
- les tests d'hypothèse (par exemple : comparer la valeur moyenne d'un caractère à une valeur fixée) ;
- les plans d'expérience (par exemple : étude expérimentale de l'effet d'additifs sur l'adhésivité, en faisant varier les doses et les associations).

Les méthodes *bayésiennes* associent à ces résultats expérimentaux des distributions de probabilités *a priori*, sur les valeurs des paramètres par exemple, distributions étayées par le niveau des connaissances préalables sur le sujet, avant toute expérimentation.

L'information "objective" apportée par l'expérimentation intervient pour modifier les distributions *a priori*, et les remplacer par des distributions *a posteriori*, qui concernent des probabilités conditionnelles (et plus précisément, conditionnées par l'expérimentation).

Dans l'étude qui suit, on se propose de situer, de façon élémentaire, les techniques bayésiennes, en supposant connu le contexte théorique fondamental, qui se réduit d'ailleurs au théorème de Bayes et aux diverses façons de l'écrire.

On partira d'une étude de cas, en montrant comment, dans un problème de décision, les "règles bayésiennes" de décision permettent de réaliser un pourcentage global d'erreur (c'est-à-dire de décisions mal fondées) au plus égal à celui des règles classiques, et généralement inférieur.

---

Article repris du bulletin de liaison du Laboratoire Central des Ponts et Chaussées, publié avec l'aimable autorisation du Directeur du L.C.P.C..

Nous nous sommes inspirés d'un exemple figurant dans l'excellent ouvrage "Statistics by example", de F. Mosteller, W. Kruksal et al. Nous sommes également redevables d'observations très pertinentes de la part de MM. Vessereau et Bonitzer, que nous tenons à remercier.

Dans la seconde partie, on examine les applications possibles des méthodes bayésiennes, notamment en matière de contrôle de qualité, en insistant non seulement sur les avantages qu'elles peuvent apporter, mais aussi sur les limitations qui leur sont inhérentes, et les critiques qu'on peut leur faire.

## I. ETUDE D'UN CAS DE DECISION PAR MAXIMUM DE VRAISEMBLANCE OU PAR REGLE BAYESIENNE

### RESUME DU CAS

On veut établir la carte géologique d'une région inconnue et peu accessible au moyen de relevés obtenus à partir d'un survol aérien, et en utilisant une méthode physique qui permet de distinguer le basalte du granite. Malheureusement, la méthode n'est pas sûre à 100 %.

On l'a expérimentée sur une région connue, dont la carte existe, et on a ainsi pu établir les risques d'erreur de la méthode.

A partir de ces informations, on peut arriver à clarifier un peu les choses pour la région inconnue, en utilisant une des deux méthodes suivantes qui diffèrent quelque peu :

- La méthode du maximum de vraisemblance,
- La méthode de Bayes.

C'est cette dernière qui, dans tous les cas, conduit au taux minimal d'erreur.

### EXPOSE DETAILLE

Il s'agit en fait de construire un modèle probabiliste permettant d'établir une carte aussi précise que possible. On va voir comment le construire et le modifier à la lumière des résultats obtenus ; et comment on peut comparer différents modèles quant à la probabilité qu'ils donnent des résultats conformes à la réalité.

On peut distinguer *en laboratoire* les roches granitiques des roches basaltiques (on les appellera par la suite granite et basalte, pour simplifier) par diverses méthodes expérimentales. Une de ces méthodes consiste à analyser le spectre infrarouge de la lumière solaire réfléchi par la surface de la roche : les deux spectres permettent de distinguer la nature de la roche, spécialement à trois longueurs d'onde particulières. Si  $R_1$ ,  $R_2$  et  $R_3$  représentent les intensités spectrales pour ces trois longueurs d'onde, les granites se caractérisent par  $R_1 < R_2 < R_3$  et les basaltes, par  $R_3 < R_1 < R_2$ .

L'intérêt de cette méthode est de pouvoir être appliquée hors laboratoire, à partir d'un avion. Cependant, les conditions atmosphériques peuvent entraîner des erreurs.

A titre d'exemple, des vols au-dessus de deux régions *connues* pour être, l'une, uniquement granitique et, l'autre, uniquement basaltique, ont donné les pourcentages d'observations indiqués table I.

TABLE I

Type de roche	Classement des intensités			Total
	$R_1 < R_2 < R_3$	$R_1 < R_3 < R_2$	$R_3 < R_1 < R_2$	
Granite	60 %	25 %	15 %	100 %
Basalte	10 %	20 %	70 %	100 %

On va utiliser cette table pour identifier les roches de régions inconnues. La procédure suivie est la suivante. Un vol est effectué au-dessus d'une région semblable (du point de vue climat, notamment) aux précédentes, mais de géologie inconnue. On a relevé des spectres pour 1 000 points équidistants de 100 m, dans le but d'identifier les alternances granite-basalte. Les pourcentages de la table I nous incitent à adopter la règle n° 1 de classification, résumée table II.

TABLE II

Type de spectre en un point	Appellation de la roche
$R_1 < R_2 < R_3$	Granite
$R_1 < R_3 < R_2$	Granite
$R_3 < R_1 < R_2$	Basalte

Nous sommes conscients que cette règle conduira à mal classer certains points. Les granites ayant donné  $R_3 < R_1 < R_2$  seront incorrectement appelés basaltes. Cela affecte (table I) 15 % de tous les spectres pris sur granite.

Une bonne mesure de la qualité de la méthode consisterait dans la proportion P des 1000 spectres mal classés. Mais on ne peut la calculer que si on connaît les proportions *réelles* des granites et des basaltes. Appelons-les respectivement  $p_G$  et  $p_B = 1 - p_G$ . On a alors :

$$P = 0,15 p_G + 0,30 p_B \quad (0,30 = 0,10 + 0,20, \text{ table I}).$$

Ce qui s'écrit aussi :

$$P = 0,15 p_G + 0,30(1 - p_G) = 0,30 - 0,15 p_G .$$

Si on passe de la notation "proportionnelle" à la notation "probabiliste", on peut écrire :  $P = \text{Prob}(\text{erreur}) = \text{Prob}(\text{erreur/granite}) \times \text{Prob}(\text{granite}) + \text{Prob}(\text{erreur/basalte}) \times \text{Prob}(\text{basalte})$ .

Ce qui est une équation classique des probabilités conditionnelles.

A titre d'exemple, si  $p_G = 1/3$ ,  $P = 0,25$ , c'est-à-dire 250 des 1 000 points seraient classés de façon incorrecte. Si  $p_G = 2/3$ ,  $P = 0,20$ . Cela montre bien que la performance de la méthode est fonction de  $p_G$ .

On peut imaginer sur le même principe d'autres règles, ainsi, par exemple, les règles n° 2 et 3 qui correspondent aux tables III et IV.

TABLE III

Type de spectre en un point	Appellation de la roche
$R_1 < R_2 < R_3$ $R_1 < R_3 < R_2$ $R_3 < R_1 < R_2$	Granite Basalte Granite

TABLE IV

Type de spectre en un point	Appellation de la roche
$R_1 < R_2 < R_3$ $R_1 < R_3 < R_2$ $R_3 < R_1 < R_2$	Granite Basalte Basalte

Il est facile de voir que la règle n° 2 donnera *toujours* quel que soit  $p_G$  un taux d'erreur supérieur à la règle n° 1, car  $(0,80 - 0,55 p_G)$  est toujours supérieur à  $(0,30 - 0,15 p_G)$ , puisque  $p_G \leq 1$ .

En revanche, la règle n° 3 sera meilleure que la règle n° 1 pour certaines valeurs de  $p_G$  et moins bonne pour d'autres. On laisse au lecteur le plaisir de découvrir que la règle n° 3 est meilleure pour

$$0 < p_G < \frac{4}{9}.$$

On peut noter que les trois règles citées figurent parmi les huit règles qu'on pourrait imaginer *a priori*, et dont certaines seraient d'ailleurs absurdes.

Il est maintenant important de comprendre le principe de la règle n° 1. Cette règle consiste à nommer la roche inconnue du nom de celle *qui rend le plus probable le spectre observé*, c'est-à-dire en comparant les chiffres de la colonne (table I) correspondant au spectre observé, et en choisissant la roche donnant le plus grand des deux chiffres, qui sont des probabilités conditionnelles du type Prob. (Spectre observé/Roche réelle), le signe / signifiant "sachant que". Cette méthode s'appelle celle du *maximum de vraisemblance*.

Mais il peut arriver, comme on l'a vu, que d'autres règles que celle-ci conduisent, dans certains cas, à de plus faibles probabilités globales d'erreur.

## REGLE BAYESIENNE

Elle part d'un principe différent et consiste à affecter à *un spectre donné la roche la plus probable*.

Pour bien comprendre le lien et les différences entre les deux règles, prenons comme exemple un spectre donnant  $R_1 < R_3 < R_2$ . Le maximum de vraisemblance conduit à appeler ce spectre granite, parce que :

$$\text{Prob}(R_1 < R_3 < R_2/\text{granite}) (= 0,25) > \text{Prob}(R_1 < R_2 < R_3/\text{basalte}) (= 0,20).$$

Mais la roche la plus probable n'est le granite que si :

$$\text{Prob (granite/R}_1 < R_3 < R_2) > \text{Prob (basalte/R}_1 < R_3 < R_2).$$

Or le théorème de Bayes, qui s'écrit dans le cas présent :

$$\boxed{\text{Prob (granite/R}_1 < R_3 < R_2) = \frac{\text{Prob (R}_1 < R_3 < R_2/\text{granite}) \cdot p_G}{\text{Prob (R}_1 < R_3 < R_2)}} \quad (1)$$

permet de voir que le granite est le plus probable si et seulement si :

$$\text{Prob (R}_1 < R_3 < R_2/\text{granite}) \cdot p_G > \text{Prob (R}_1 < R_3 < R_2/\text{basalte}) (1 - p_G)$$

soit, en utilisant la table I,  $p_G > \frac{4}{9}$ .

Les deux règles diffèrent dans leur formule parce que la seconde pondère par  $p_G$  et  $p_B = 1 - p_G$  les probabilités que la première compare.

Leurs résultats ne coïncident que si  $p_G > \frac{4}{9}$  ; autrement, la règle de Bayes conduit à appeler basalte ce que la règle du maximum conduit à appeler granite.

L'ennui est évidemment qu'on ignore  $p_G$ . On examinera ce point ultérieurement.

Si on examine tous les cas possibles de règles bayésiennes, on verra aisément qu'il n'y en a que quatre qui sont indiquées table V ; on y a mentionné les plages de  $p_G$  qui conduisent, dans chaque cas, à *la règle la plus efficace de toutes les règles possibles*.

On remarquera (avec un peu d'étonnement) que les règles de Bayes n° 3 et n° 4 font comme si on n'avait obtenu aucun résultat. En fait, pour les valeurs faibles de  $p_G$ , ou de  $p_B$ , il vaut mieux ignorer les résultats qu'en tenir compte, du point de vue de la probabilité d'erreur globale.

TABLE V

Type du spectre observé	Appellation de la roche			
	Bayes n° 1	Bayes n° 2	Bayes n° 3	Bayes n° 4
$R_1 < R_2 < R_3$	Granite	Granite	Basalte	Granite
$R_1 < R_3 < R_2$	Granite	Basalte	Basalte	Granite
$R_3 < R_1 < R_2$	Basalte	Basalte	Basalte	Granite
Plages	$\frac{4}{9} < p_G < \frac{14}{17}$	$\frac{1}{7} < p_G < \frac{4}{9}$	$p_G < \frac{1}{7}$	$p_G > \frac{14}{17}$

On remarquera aussi qu'il n'y a aucune différence quant aux résultats entre les deux types de règles (règle de Bayes n° 1 et règle du maximum de vraisemblance) si :

$$\frac{4}{9} < p_G < \frac{14}{17}.$$

## ESTIMATION DE $p_G$

On va utiliser la table I pour estimer  $p_G$  en utilisant, par exemple, la méthode simple qui suit.

Appelons respectivement  $\pi_1$ ,  $\pi_2$  et  $\pi_3$  les proportions globales parmi les spectres observés, de ceux qui donnent  $R_1 < R_2 < R_3$ ,  $R_1 < R_3 < R_2$  et  $R_3 < R_1 < R_2$ . On a évidemment  $\pi_1 + \pi_2 + \pi_3 = 1$ .

*Avant l'observation*, on peut considérer les  $\pi_i$  comme des variables aléatoires, dont les espérances sont égales à ce que donne la théorie, à savoir :

$$\left. \begin{aligned} E(\pi_1) &= 0,60 p_G + 0,10 p_B = 0,50 p_G + 0,10 \\ E(\pi_2) &= 0,25 p_G + 0,20 p_B = 0,05 p_G + 0,20 \\ E(\pi_3) &= 0,15 p_G + 0,70 p_B = -0,55 p_G + 0,70 \end{aligned} \right\} \quad (2)$$

*Après l'observation*, on peut considérer les proportions observées comme des *estimations ponctuelles* de  $E(\pi_1)$ ,  $E(\pi_2)$  et  $E(\pi_3)$ , les substituer dans les équations (2), et en tirer par résolution une estimation de  $p_G$ .

En fait, il y a même deux équations de trop... , mais la seconde comporte une faible pondération de  $p_G(0,05)$ , conduisant à une mauvaise précision de l'estimation.

Il ne reste plus que la première et la troisième équation pour estimer  $p_G$  à partir des proportions observées  $\pi_1$  et  $\pi_3$ .

$$\left. \begin{aligned} 0,50 p_G + 0,10 &= \pi_1 \\ -0,55 p_G + 0,70 &= \pi_3 \end{aligned} \right\}$$

soit encore :

$$\left. \begin{aligned} p_G &= \frac{\pi_1 - 0,10}{0,50} \\ p_G &= \frac{0,70 - \pi_3}{0,55} \end{aligned} \right\} \quad (3)$$

Il y a deux conditions à satisfaire :

$$\begin{aligned} \pi_1 + \pi_3 &< 1 \\ 0 &< p_G < 1 \end{aligned}$$

ce qui s'écrit :

$$\begin{aligned} 0,10 &< \pi_1 < 0,60 \\ 0,15 &< \pi_3 < 0,70 \end{aligned}$$

Représentons la situation dans un système d'axes  $(\pi_1, \pi_3)$ . Le point  $(\pi_1, \pi_3)$  doit être intérieur au pentagone figuré en trait épais sur la figure 1.

Rappelons-nous, d'autre part, que les quatre règles bayésiennes de la table V sont recommandées chacune pour une plage de valeur de  $p_G$ .

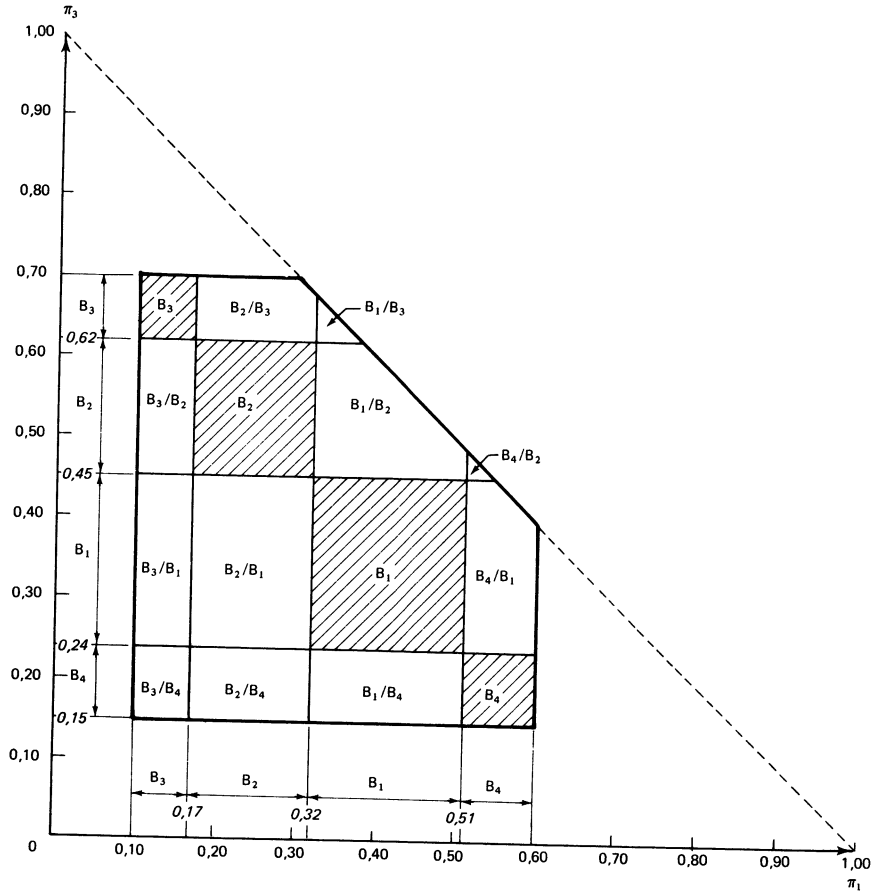


Figure 1. — Représentation dans le plan  $(\pi_1, \pi_3)$ .

Les conditions d'inégalité sur  $p_G$  se traduisent, grâce aux équations (3), par des conditions d'inégalités sur  $\pi_1$  et  $\pi_3$ , de la manière suivante :

$$\left. \begin{aligned}
 \text{R\`egle } B_3 \quad p_G < \frac{1}{7} &\rightarrow \left\{ \begin{array}{l} \pi_1 < 0,17 \\ \pi_3 > 0,62 \end{array} \right. \\
 \text{R\`egle } B_2 \quad \frac{1}{7} < p_G < \frac{4}{9} &\rightarrow \left\{ \begin{array}{l} 0,17 < \pi_1 < 0,32 \\ 0,45 < \pi_3 < 0,62 \end{array} \right. \\
 \text{R\`egle } B_1 \quad \frac{4}{9} < p_G < \frac{14}{17} &\rightarrow \left\{ \begin{array}{l} 0,32 < \pi_1 < 0,51 \\ 0,24 < \pi_3 < 0,45 \end{array} \right. \\
 \text{R\`egle } B_4 \quad p_G < \frac{14}{17} &\rightarrow \left\{ \begin{array}{l} \pi_1 > 0,51 \\ \pi_3 < 0,24 \end{array} \right.
 \end{aligned} \right\} \quad (4)$$



Les zones ainsi définies pour  $\pi_1$  et  $\pi_3$  découpent le pentagone en quinze aires. On peut remarquer qu'il n'y a que quatre aires où on peut choisir indifféremment l'une ou l'autre des équations (3) pour estimer  $p_G$  : on tombera sur la même règle de Bayes ; dans les onze autres, en revanche, l'une des équations conduit à une règle, l'autre à une règle différente (par exemple  $\pi_1 = 0,55$  et  $\pi_3 = 0,35$  ; la première conduit à la règle  $B_4$ , la seconde à la règle  $B_1$ ). C'est peut-être là un caractère critiquable des méthodes bayésiennes.

On peut d'autre part se poser la question suivante : peut-on utiliser les résultats de l'observation pour réestimer les chiffres de la table I ? En fait, il faut répondre par la négative : les résultats se résument à trois nombres, liés de plus par une relation (leur somme est égale à 1). Il y a trop de nombres à estimer à partir de ces trois résultats liés.

## II. APPLICATIONS POSSIBLES DU THEOREME DE BAYES, NOTAMMENT EN MATIERE DE TECHNIQUES DE CONTROLE STATISTIQUE

La plupart des applications possibles<sup>(1)</sup> du théorème de Bayes, ou encore, les *méthodes bayésiennes*, consistent à préciser, éventuellement par étapes successives, la probabilité d'un événement A dont on connaît la probabilité *a priori*  $P(A)$  ; ces précisions successives proviennent de l'*ajout d'informations supplémentaires*, p. ex. B.

En effet,  $P(A/B)$  peut être considérée à son tour comme probabilité "*a priori*", et on peut l'affiner par une nouvelle information supplémentaire concernant un événement C.

En fait, on peut aller jusqu'à se demander jusqu'à quel point il peut exister une probabilité "*a priori*". Toute probabilité d'un événement est en réalité conditionnée par "ce que je sais", "l'état actuel des connaissances", etc.

On peut rattacher ce processus de combinaison des "degrés de croyance"<sup>(2)</sup> préalables avec l'expérience pour affiner ces degrés de croyance, aux processus d'éducation, d'apprentissage, ou d'acquisition de connaissances.

Pour simplifier, on peut appeler "méthodes bayésiennes" celles qui réunissent les conditions suivantes :

- 1) la probabilité traduit un "degré de croyance",
- 2) il existe des distributions *a priori* des paramètres,
- 3) le théorème de Bayes fournit des distributions *a posteriori*,
- 4) ces distributions *a posteriori* sont utilisées :
  - pour l'estimation
  - pour des tests d'hypothèse.

Le grand, et peut-être le seul problème que posent ces méthodes réside dans le caractère 2. Il est souvent difficile d'approcher les distributions *a priori*,

.....  
(1) Un terrain très propice aux méthodes bayésiennes est constitué par les techniques d'analyse statistique de la décision.

(2) C'est une des définitions de la probabilité données par l'AFNOR.

et on se contente d'une approche intuitive, "à vue de nez". Lorsque l'expérimentation qui apporte une information supplémentaire n'est fondée que sur un faible effectif d'échantillon, la distribution *a priori* retenue imprimera sa marque de façon très forte, et la qualité du résultat dépendra essentiellement de la qualité de cette distribution. En revanche, si l'effectif est grand, la distribution *a priori* aura beaucoup moins d'importance.

## 1. EXEMPLE : TESTS DE FIABILITE D'UN MATERIEL DE SERIE

a) Supposons pour simplifier qu'il n'y ait que deux niveaux de fiabilité : le niveau requis (fiabilité 0,90) et un mauvais niveau (fiabilité 0,50).

On estime, à partir de l'expérience passée, et du "jugement de l'ingénieur", les probabilités *a priori* respectives de ces niveaux. Puis, on effectue N essais qui donnent E échecs.

Le théorème de Bayes permet de calculer la probabilité que le niveau requis est atteint dans le lot d'où sont extraits les N éléments testés.

*Exemple numérique* : Au départ, on suppose qu'il y a une probabilité de 0,50 d'avoir le niveau requis, et donc une probabilité de 0,50 aussi d'avoir le mauvais niveau.

On fait un essai sur 10 éléments. Aucun échec. Que deviennent les probabilités *a priori* si on les combine avec l'expérimentation ?

Par le théorème de Bayes : 0,01 pour le bas niveau et 0,99 pour le niveau requis.

Si on avait utilisé l'approche classique, la limite inférieure de l'intervalle de confiance de la fiabilité, au niveau 0,90, ne serait que de 0,80. Cela inciterait à faire des essais supplémentaires, probablement sans nécessité.

b) On peut voir par l'exemple suivant que Laplace n'avait pas entièrement tort de dire que "la théorie des probabilités n'est que le bon sens réduit au calcul" (bien que...).

Les méthodes analytiques, les essais de laboratoire, l'expérience de l'ingénieur permettent à celui-ci de dire : "Cet équipement fonctionnera parfaitement 19 fois sur 20 essais."

Supposons qu'on désire *démontrer* une fiabilité à 90 %, au niveau de confiance 90 %. Disons que cela revient à accepter 2 échecs au plus sur 50 essais. L'épreuve consistera donc à faire un essai sur 30 équipements et à n'accepter au plus qu'1 échec.

Jugement de l'ingénieur	1 échec sur 20 essais
Essai de démonstration	$\leq$ 1 échec sur 30 essais
Exigence	$\leq$ 2 échecs sur 50 essais

Tout repose sur la ... fiabilité de l'affirmation de l'ingénieur.

## 2. LES ARGUMENTS DES BAYESIENS

Les bayésiens inconditionnels ont le point de vue suivant.

Leurs méthodes leur paraissent meilleures parce que :

- Elles donnent les mêmes résultats que les méthodes classiques quand celles-ci s'appliquent ; en plus :
  - elles mettent en lumière les hypothèses sous-jacentes (1)
  - elles sont intuitivement plus acceptables (2).
- Elles donnent de meilleures réponses que les tests classiques, tels qu'on les utilise actuellement (3).
- Elles permettent de traiter des problèmes non accessibles aux méthodes classiques (4).

On va essayer de voir, dans ce qui suit, ce que l'on peut retenir de l'enthousiasme des néo-bayésiens, et ce qu'il convient de considérer avec quelque scepticisme.

#### a) *Commençons par une critique*

Les auteurs américains, statisticiens de profession et employés dans l'industrie, semblent avoir eu beaucoup de mal à faire comprendre à leurs patrons ce qu'est l'estimation par intervalle de confiance, basé sur un échantillon d'effectif  $n$ .

La bonne façon de présenter classiquement l'intervalle de confiance est de dire qu'un paramètre certain, mais de valeur inconnue, est *encadré par deux limites*, avec une probabilité (grande)  $1 - \alpha$  ( $\alpha$  petit). Ces limites sont des valeurs prises par des variables aléatoires sur l'épreuve que constituent la prise d'échantillon et les mesures adéquates qui s'ensuivent. Le "boss" américain n'arrive pas à comprendre cet encadrement par un intervalle qui change à chaque échantillon et trouve que le statisticien lui raconte des balivernes. Qu'à cela ne tienne. Notre néo-bayésien va faire un tour de passe-passe, qu'on va démystifier sur un exemple.

Appelons  $m$  le paramètre certain, mais inconnu, du collègue classique. Supposons que ce soit la moyenne de la loi de  $X$ , et qu'on connaisse  $\sigma$ , l'écart-type de cette loi ; le néo-bayésien considère *comme une variable aléatoire* la moyenne  $m$ .

Soit  $\bar{x}$  la moyenne d'échantillon (estimation ponctuelle de  $m$ , selon la méthode classique) ; en supposant  $n$  assez grand, la loi de probabilité  $P(m/\bar{x}, \sigma^2)$  est une loi normale de moyenne  $\bar{x}$  et de variance  $\sigma^2/n$ .

On écrira donc, tout simplement :

$$\text{Prob}\left(\bar{x} - \frac{2\sigma}{\sqrt{n}} < m < \bar{x} + \frac{2\sigma}{\sqrt{n}}\right) = 0,95$$

Les limites de l'intervalle de confiance sont *exactement les mêmes* qu'avec la méthode classique, mais on dira au "boss" que la probabilité que  $m$  soit entre  $\bar{x} \pm \frac{2\sigma}{\sqrt{n}}$  est de 0,95.

-----

(1) Objection possible : les problèmes classiques d'estimation ou de tests considèrent des paramètres inconnus, mais certains et non des variables aléatoires. On n'a donc nul besoin d'hypothèses sous-jacentes sur la "distribution" de ces paramètres.

(2) Objection possible : que veut dire exactement "intuitivement" ?

(3) Objection possible : que veut dire meilleures : mieux vérifiées, plus économiques ?

(4) Cela est sans doute exact.

Il aura l'impression d'une affirmation solide, et le tour sera joué.

Je crois qu'on n'a *aucun intérêt* à présenter les choses ainsi (ce que font d'ailleurs certains statisticiens classiques, par abus de langage).

b) L'exemple suivant fera toucher du doigt – on l'espère – le talon d'Achille des méthodes bayésiennes.

Supposons qu'on ait "une certaine idée" de la proportion de défectueux d'une fabrication de pièces. Cette idée est exprimée, par exemple, par le contremaître (pas le directeur, qui aura certainement tendance à dire 0%), sous la forme suivante : 75 chances sur 100 que cela soit inférieur à 0,10.

Le bayésien de service au bureau de contrôle part de cela et établit le petit graphique de la figure 2.

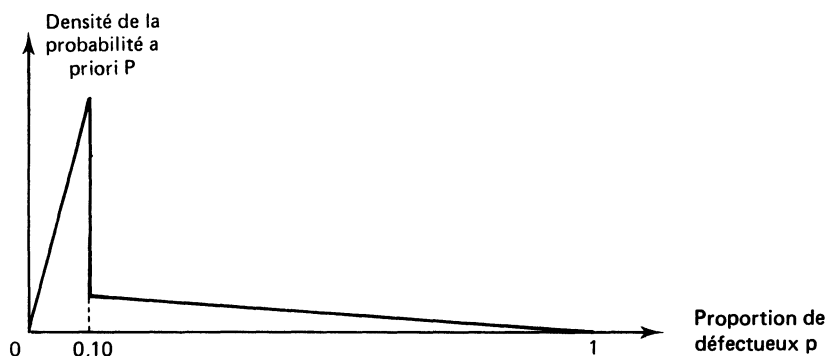


Figure 2

Les deux ordonnées correspondant à la discontinuité de la loi de densité pour  $p = 0,10$  sont calculées de façon que le triangle de gauche ait une surface de 0,75 et celui de droite, une surface de 0,25.

A partir de là, on peut fabriquer avec le théorème de Bayes, des abaques donnant par exemple, l'effectif  $n$  d'échantillon devant présenter au plus 1 défectueux pour "démontrer" qu'on a bien au plus 10% de défectueux dans le lot, avec un niveau de confiance  $1 - \alpha$  fixé, suivant les diverses valeurs  $P$  que va annoncer le contremaître (ici  $P = 0,75$ ). La famille de courbes (des morceaux de lignes droites) est indiquée en valeurs de  $p$  (ici 0,10); ces courbes sont toutes descendantes en  $P$  croissant, et d'autant plus fortement inclinées que  $p$  est plus petit. Les valeurs d'effectif sont inférieures à celles que donnent les méthodes classiques, pour les mêmes valeurs de  $p$ .

Les bayésiens considèrent donc leurs méthodes comme économiques.

Mais il convient d'attirer fortement l'attention sur deux points :

- "fiabilité" de l'affirmation du contremaître ?
- représentativité de la courbe de densité bi-triangulaire ? <sup>(1)</sup>.

-----  
(1) Il faut souligner le fait que ce diagramme bi-triangulaire est extrêmement simpliste. On aurait intérêt à choisir plutôt une loi Bêta, dont les paramètres seraient cohérents avec l'affirmation du contremaître.

Quand on ne sait rien, rétorquent les bayésiens, on peut toujours adopter la forme de densité (fig. 3), qu'on appellera pudiquement "diffuse".

*Remarque.* — On sait déjà quelque chose : la proportion de défectueux ne saurait dépasser 1. Quand il s'agit d'un paramètre dont on sait uniquement qu'il est positif, le rectangle est moins facile à tracer...

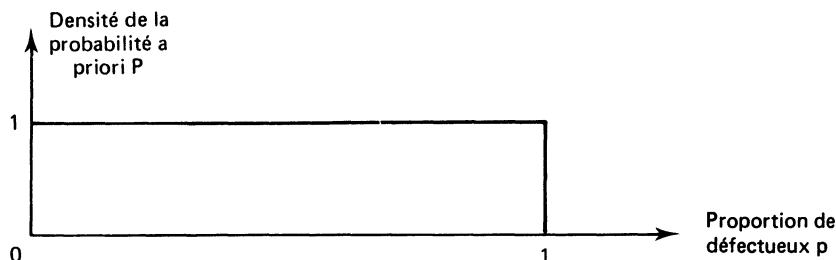


Figure 3

**Moralité.** *Les méthodes bayésiennes ne valent que ce que vaut l'information a priori. Si celle-ci est bonne, l'application est possible, et conduit effectivement à des décisions plus économiques. Si elle est mauvaise, ou vague, la prudence est de règle.*

C'est sur cette consigne que nous concluons.

#### QUELQUES REFERENCES BIBLIOGRAPHIQUES

- BENJAMIN J.R., CORNELL C.A. — *Probability, Statistics and Decision for Civil Engineers*, Mc Graw Hill Book Company, New York, 1970.
- LINDLEY D.V. — *Introduction to Probability and Statistics from a bayesian viewpoint*, Cambridge University Press, 1965.
- MOSTELLER F., KRUSKAL W.H. *et al.* — *Statistics by example*, Addison-Wesley Publ. Company, Reading (Mass.), 1973.
- WALLISER B. — *Ouvrage de méthode : I. Instruments d'analyse : I. Schéma d'un processus de décision. II. Théorie de la décision*, Documentation française, 1976, 1977.