

REVUE DE STATISTIQUE APPLIQUÉE

GUY D'HERBEMONT

Aspects de la théorie statistique des décisions

Revue de statistique appliquée, tome 11, n° 3 (1963), p. 41-82

http://www.numdam.org/item?id=RSA_1963__11_3_41_0

© Société française de statistique, 1963, 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/>

ASPECTS DE LA THÉORIE STATISTIQUE DES DÉCISIONS (1)

Guy D'HERBEMONT

Conseiller en Statistique à la Compagnie des Machines Bull

SOMMAIRE

- I - JEUX ET DECISIONS STATISTIQUES.
 - II - STRATEGIES PURES DU STATSTICIEN.
 - III - STRATEGIES ADMISSIBLES ET STRATEGIES MIXTES.
 - IV - STRATEGIES DE BAYES.
 - V - LES REGLES DE SCORE.
 - VI - DETERMINATION DES STRATEGIES DE BAYES.
 - VII - STRATEGIES MINIMAX.
 - VIII - CRITERES DE CHOIX.
 - IX - CONCLUSION.
- BIBLIOGRAPHIE.
- ANNEXE : Probabilités a priori et a posteriori.

(1) Etude élaborée pour le numéro 4 d'"Informations Scientifiques BULL" et éditée avec l'accord du Comité de Rédaction.

La théorie statistique des décisions a été créée en 1939 par Abraham Wald, professeur de statistique mathématique à l'Université de Columbia.

Les résultats obtenus par Wald à ce sujet sont restés assez longtemps des secrets militaires. Ils n'ont guère été publiés qu'après la fin de la deuxième guerre mondiale.

L'ouvrage fondamental de Wald "Statistical Decision Functions" est paru en 1950. A l'heure actuelle la plupart des statisticiens estiment que les méthodes utilisées en théorie des décisions doivent servir de base à tous les travaux fondamentaux de statistique mathématique.

Ces méthodes sont d'ailleurs basées sur le même principe que celui sur lequel ont insisté, les premiers, Neyman et Pearson, créateurs de la théorie des tests d'hypothèses qui porte leur nom. Ce principe peut être énoncé de la façon suivante : toute méthode statistique doit être rigoureusement évaluée par les conséquences que son application implique face aux diverses circonstances dans lesquelles on l'utilise.

Le texte qui suit ne donne qu'une image élémentaire et très incomplète de la théorie de Wald, mais il s'efforce, sans faire appel à des êtres mathématiques trop compliqués, d'évoquer certaines des notions fondamentales qui mériteraient d'être mieux connues et plus souvent appliquées.

I - JEUX ET DECISIONS STATISTIQUES -

Décisions, états, et tableau de perte.

Lorsqu'on a à faire face à une situation, plus ou moins bien connue, en prenant une décision, on peut souvent analyser le problème ainsi posé en distinguant diverses modalités que la situation en question est susceptible de revêtir. Nous dirons que ces modalités sont les divers "états" possibles. La situation réelle correspond donc à l'un de ces états, mais on ignore lequel.

Quant à la décision à prendre elle résulte généralement d'un choix entre plusieurs décisions éventuelles, dont certaines sont mieux adaptées que d'autres à tel ou tel "état".

Poussant l'analyse plus loin il arrive que l'on puisse chiffrer la perte ou le gain qui résulterait de l'application de chaque décision face à chaque état possible. Ceci détermine alors ce que l'on appelle le tableau de perte qui caractérise le problème (les gains sont considérés comme des pertes négatives).

Premier exemple de problème de décision.

Imaginons par exemple un fabricant dont la capacité de production est déterminée, mais qui puisse modifier la proportion de l'un de ses modèles (le modèle rouge) par rapport à un autre (le modèle jaune). Le choix de cette proportion p constitue la décision à prendre. La situation face à laquelle se trouve placé le fabricant est caractérisée par la proportion q des clients qui préfèrent le modèle rouge. Chaque proportion q constitue un "état" et le tableau de perte est une certaine fonction $f(p, q)$ qui peut être calculée en tenant compte des inventus et du coût qu'implique une modification des programmes de fabrication actuels.

Pour simplifier l'exposé nous limiterons à deux le nombre des "états" possibles :

- état R : $2/3$ de la clientèle préfère le produit rouge $q = 2/3$
- état J : $1/3$ de la clientèle préfère le produit rouge $q = 1/3$

Pour la même raison, nous limiterons à trois les décisions envisagées :

- la première décision (décision R) est bien adaptée à l'état rouge
- la deuxième décision (décision J) est bien adaptée à l'état jaune
- la troisième décision résulte d'un compromis que nous ne décrirons pas.

Les six valeurs correspondantes de la fonction $f(p, q)$ seront prises égales à celles déterminées par le tableau suivant :

Tableau des pertes du fabricant	Décisions du fabricant		
	1 décision R	2 décision J	3 compromis
1 - état R	0	50	60
2 - état J	100	0	60

Analogie avec la théorie des jeux à deux personnes.

Ce tableau de perte du fabricant en fonction des états et des décisions pourrait être considéré comme définissant, sous forme normale, un jeu à deux joueurs, le 1er joueur étant la clientèle et le second le fabricant.

C'est en effet un résultat important de la théorie des jeux, que tout jeu de paiement à somme nulle (perte d'un joueur - gain de l'autre) peut être mis, au moins théoriquement sous la forme d'un tableau de perte semblable au précédent.

Ainsi, par exemple, dans le jeu de Bridge, imaginons que l'équipe Nord-Sud (joueur 1) décrive à des remplaçants sans initiative le compor-

tement exact qu'ils doivent avoir face à toutes les circonstances qui peuvent se présenter. Dans ces conditions les remplaçants connaissant les annonces déjà faites et leurs cartes, savent exactement ce qu'ils doivent annoncer eux-mêmes ou ce qu'ils doivent jouer dans toutes les phases du déroulement du jeu.

Définition des stratégies pures d'un joueur.

Une telle définition de la manière de jouer dans tous les cas possibles sans incertitude ni doute porte le nom de stratégie pure. (Nous verrons plus loin que ces stratégies pures servent à définir encore d'autres stratégies, comportant cette fois une part d'incertitude : Les stratégies mixtes).

Ce sont toutes les stratégies pures du joueur 1 qui constitueraient les intitulés des lignes du tableau de perte - De même, les stratégies pures du joueur 2 (équipe Est-Ouest) intituleraient les colonnes.

Forme normale.

Ceci posé, si le joueur 1 choisit une stratégie s_1 parmi les lignes du tableau et le joueur 2 une stratégie s_2 parmi les colonnes, le déroulement du jeu est entièrement déterminé dès que les cartes sont distribuées.

A chaque donne, le couple s_1, s_2 fait donc correspondre une perte particulière.

Considérons alors les diverses donnes possibles ; celles-ci, compte tenu de leurs probabilités respectives, font fluctuer les pertes d'une certaine façon autour d'une perte moyenne relative à toutes les donnes. C'est cette perte moyenne⁽¹⁾ que l'on pourrait théoriquement mentionner dans la case s_1, s_2 du tableau de perte.

En ce qui concerne les pertes moyennes, le jeu de Bridge est alors équivalent au suivant (forme normale) :

Chaque joueur choisit, indépendamment, le premier une ligne et le second une colonne ; la case correspondante définit la perte du joueur 2 donc le gain du joueur 1.

Exemple de mise sous forme normale d'un jeu.

Pour montrer sur un exemple simple comment se définissent les stratégies et s'établit le tableau de perte, considérons le jeu de "bluff" basé sur la règle de jeu suivante : "Dans un jeu de 32 cartes bien mélangées, le joueur 1 tire une carte. C'est une basse carte (7 - 8 - 9 - 10) ou une haute carte avec probabilité 1/2 pour chaque cas.

(1) Il s'agit ici d'une moyenne au sens statistique du terme. C'est-à-dire qu'à chaque perte possible on attribue sa probabilité et que la perte moyenne est égale à la somme des produits des pertes possibles par la probabilité correspondante. Cette grandeur est aussi appelée "espérance mathématique" par opposition à la moyenne arithmétique usuelle qui porte le nom de "moyenne empirique". Cette dernière serait par exemple la moyenne arithmétique de pertes réelles résultant d'un nombre de parties déterminées. On montre que la moyenne empirique tend en probabilité vers l'espérance mathématique lorsque le nombre de parties est assez grand.

Si le joueur 1 tire une haute carte il doit annoncer 200 - S'il tire une basse carte il doit soit payer 100 au joueur 2, soit annoncer 200 (bluff).

Le joueur 2 peut, soit admettre sans contrôle l'annonce du joueur 1 et lui payer 100, soit demander à contrôler et il gagne alors 200 si le joueur 1 a bluffé, mais lui paye 200 dans le cas contraire".

Le joueur 1 a deux stratégies pures (il annonce nécessairement 200 si la carte est haute) :

- a - si la carte est basse : payer 100
- b - si la carte est basse : annoncer 200

Le joueur 2 a deux stratégies pures également (si le joueur 1 lui paye 100 il n'a rien à décider) :

- A - si le joueur 1 annonce 200 : payer 100
- B - si le joueur 1 annonce 200 : contrôler

On a donc 4 combinaisons de stratégies pures à étudier.

Calcul du tableau de perte.

Le calcul du tableau de perte s'effectue de la façon suivante pour chacune de ces quatre combinaisons :

aA - le joueur 1 ne bluffe pas et le joueur 2 ne contrôle pas. Alors:
. si la carte est haute (probabilité 1/2) le joueur 1 annonce 200 et le joueur 2 paye 100 la perte est donc 100 avec probabilité 1/2

. si la carte est basse (probabilité 1/2) le joueur 1 paye 100 le joueur 2 reçoit 100 la perte est donc - 100 avec probabilité 1/2

La perte moyenne de la combinaison a A est donc :

$$1/2 \times (100) + 1/2 \times (-100) = 0$$

aB - le joueur 1 ne bluffe pas et le joueur 2 contrôle. Alors :

. si la carte est haute (probabilité 1/2) le joueur 1 annonce 200 le joueur 2 contrôle ; comme il n'y a pas bluff, il paye 200 la perte est donc 200 avec probabilité 1/2

. si la carte est basse (probabilité 1/2) le joueur 1 paye 100 le joueur 2 reçoit 100 la perte est donc - 100 avec probabilité 1/2

La perte moyenne de la combinaison a B est donc :

$$1/2 \times (200) + 1/2 \times (-100) = 50$$

bA - le joueur 1 bluffe et le joueur 2 ne contrôle pas. Alors :

. si la carte est haute (probabilité 1/2) le joueur 1 annonce 200 et le joueur 2 ne contrôle pas et paye donc 100 la perte est donc 100 avec probabilité 1/2

. si la carte est basse (probabilité 1/2) le joueur 1 annonce 200 et le joueur 2 ne contrôle toujours pas il paye donc 100 la perte est 100 avec probabilité 1/2

La perte moyenne de la combinaison bA est donc :

$$1/2 \times (100) + 1/2 \times (100) = 100$$

On constate d'ailleurs que dans ce cas la perte de 100 est certaine.

bB - le joueur 1 bluffe et le joueur 2 contrôle. Alors :

. si la carte est haute (probabilité 1/2) le joueur 1 annonce 200 et le joueur 2 contrôle et perd 200
la perte est 200 avec probabilité 1/2

. si la carte est basse (probabilité 1/2) le joueur 1 annonce 200 et le joueur 2 contrôle et gagne 200 -
la perte est - 200 avec probabilité 1/2

La perte moyenne de la combinaison bB est donc :

$$1/2 \times (200) + 1/2 \times (-200) = 0$$

Résultats du calcul.

		Joueur 2	
		A	B
Joueur 1	a	0	50
	b	100	0

On voit donc que le tableau de perte de ce jeu de bluff est le suivant :

Ce tableau est identique, en ce qui concerne les deux premières colonnes au tableau posé initialement. On voit que des problèmes très différents peuvent conduire au même tableau de perte.

Il peut ne pas exister de stratégie pure judicieuse contre un joueur intelligent.

Supposons que le joueur 1 adopte la stratégie a à tous les coups - Le joueur 2 s'en apercevra et jouera A à tous les coups - Jouer "a" à tous les coups n'est donc pas judicieux pour le joueur 1. Supposons que le joueur 1 choisisse alors de jouer b à tous les coups ; bluffant à chaque fois, le joueur 2 s'en apercevra et jouera B à tous les coups.

Si "1" se borne à jouer à chaque coup une stratégie pure déterminée il peut seulement s'attendre à gagner 0 à tous les coups. Nous verrons qu'il y a mieux à faire pour ce joueur, en rendant son attitude imprévisible pour le joueur 2, grâce à une stratégie mixte qui lui fera jouer tantôt a tantôt b sans aucun systématisme qui puisse être exploité par "2". (Les mêmes remarques s'appliqueraient aussi au joueur 2).

Différences entre jeu à deux personnes et problème de décision face à une nature indifférente.

Revenant au problème initial, on peut dire que la clientèle a le choix entre deux stratégies pures et le fabricant entre trois stratégies pures. Toutefois il n'apparaît guère possible de considérer que les états R et J puissent être des stratégies conscientes des clients. Ceux-ci pourraient d'ailleurs ne subir ni gain ni perte du fait des décisions du fabricant.

Dans les problèmes de décision statistique ce type de joueur indifférent est habituellement désigné sous le nom de "la nature", entité abstraite qui joue les stratégies et reçoit les paiements.

L'autre joueur est "le statisticien". Il a la possibilité de sonder la nature, par échantillonnage et de prendre ses décisions d'après les résultats obtenus. Cette possibilité transforme le jeu initial en un jeu avec espionnage unilatéral au détriment du joueur 1. Autrement dit le joueur 2 peut obtenir par des paiements supplémentaires des renseignements sur la stratégie choisie par le joueur 1, après que celui-ci ait pris position.

Les stratégies avec échantillonnage.

Les stratégies pures du joueur 2 sont constituées alors par l'ensemble des manières (fonctions de décision) dont ce joueur peut attribuer aux résultats d'échantillonnage l'une ou l'autre des décisions finales qu'il doit prendre. Les pertes totales de ces nouvelles stratégies pures sont obtenues en ajoutant aux pertes du tableau celles qui résultent du coût de l'échantillonnage.

Ces nouvelles stratégies sont bien des stratégies pures car les incertitudes qui résultent de l'échantillonnage ne sont pas le fait du statisticien et celui-ci est supposé se baser sur des règles de décision déterministes telles que les résultats obtenus lors des sondages fixent à eux seuls sa conduite.

Exemple de transformation du problème de décision simple en un problème de décision statistique.

Dans notre problème l'échantillonnage consiste à faire une étude de marché, c'est-à-dire à interroger des clients éventuels sur leurs préférences. Si l'état est R il y a une probabilité $2/3$ pour qu'un client interrogé préfère le modèle rouge mais il y a une probabilité $1/3$ pour que le client préfère le modèle jaune. Si, au contraire, l'état est J ces probabilités deviennent $1/3$ en faveur de modèle rouge et $2/3$ en faveur du modèle jaune.

Tout se passe donc comme si l'état de la nature était déterminé par la couleur des boules contenues dans un sac :

- 2 boules rouges et une jaune si l'état est R
- 2 boules jaunes et une rouge si l'état est J,

car, à condition de tirer des boules à l'aveugle une à une en les remettant à chaque fois et en mélangeant entre chaque tirage, les probabilités de :

tirer R si l'état est R
tirer J si l'état est R
tirer R si l'état est J
tirer J si l'état est J

sont les mêmes que pour les réponses correspondantes du client.

La situation peut donc être schématisée de la façon suivante.

Le statisticien a, à portée de la main, le sac considéré. Il ignore son contenu exact mais peut se renseigner s'il le désire en extrayant des boules dans les conditions indiquées, pour les observer avant de prendre sa décision.

Le nombre de tirages n'est pas fixé a priori ; c'est une inconnue du problème, qui contribue à la description des stratégies pures du statisticien. Mais les études de marché dont ces tirages sont l'image, ne sont pas gratuites. Nous poserons donc que pour tirer une boule du sac il faut payer une certaine somme que nous prendrons comme unité monétaire.

Le problème de décision statistique du fabricant est maintenant complètement posé en termes mathématiques :

- on dispose d'un tableau de perte relatif aux décisions finales.
- on connaît les probabilités relatives aux résultats d'échantillonnage pour chaque état de la nature (2/3 et 1/3 ou 1/3 et 2/3)
- on connaît le coût de l'échantillonnage (ici 1 unité monétaire par tirage(1)).

Autres exemples de problèmes de décision statistique.

Avant de montrer comment se construit la solution d'un tel problème remarquons que, mis sous cette forme schématique, il a la même structure que l'un quelconque des suivants :

- Accepter (décision R) ou refuser (décision J) un lot de pièces qui peut être satisfaisant (état R) ou non (état J), connaissant les préjudices causés par la recette d'un mauvais lot ou le refus d'un lot satisfaisant et le coût du contrôle des pièces

- Etudier un phénomène physique dont on se demande s'il suit une certaine loi (état R) ou une autre loi (état J) connaissant le coût de conclusions erronées, celui des essais plus ou moins nombreux que l'on peut entreprendre pour en décider (décision R ou décision J) et l'ordre de grandeur des erreurs expérimentales (incertitudes de tirage).

- Mesurer une grandeur inconnue m (au lieu de deux états R et J on a une infinité d'états : m_1 , m_2 , et au lieu de trois décisions possibles il y en a aussi une infinité : décider $m = m_1$, $m = m_2$, ...) connaissant le coût d'une décision erronée (par exemple : coût proportionnel au carré de l'écart entre la grandeur vraie et la grandeur décidée) et la précision de l'appareil de mesure (chaque mesure peut, par exemple être considérée comme un tirage au sort dans une population de moyenne inconnue m).

(1) La théorie statistique des décisions étudie également des cas où le coût d'échantillonnage n'est pas proportionnel au nombre de tirages.

Rapports entre théorie des décisions et statistique classique:

D'une façon générale, la théorie statistique des décisions comporte comme cas particulier d'application possible, pratiquement toute la statistique classique : Théorie des tests d'hypothèse, Théorie de l'estimation et des intervalles de confiance, Théorie des plans d'expérimentation etc., à condition d'affecter à chaque cas un tableau de perte convenable.

Mais la statistique classique conserve cependant toute sa valeur ne serait-ce que parce que dans de nombreux problèmes l'introduction d'un tableau de perte est impossible ou même n'a aucun sens (peut-on chiffrer le coût d'une erreur de détermination sur la masse d'une particule élémentaire de la physique moderne ou le coût d'admettre une théorie erronée sur le champ magnétique terrestre ?).

Mentionner toutefois que l'on s'est efforcé de tourner cette difficulté en définissant une notion mathématique d'Utilité qui, tout au moins dans certains cas, peut remplacer la notion de perte monétaire.

II - STRATEGIES PURES DU STATISTICIEN -

Un exemple de stratégie du statisticien.

Revenant à notre exemple, pris sous la forme schématique que nous lui avons donnée, voici une stratégie pure que peut envisager le statisticien :

stratégie n° 4 : tirer une boule du sac
 si celle-ci est rouge décider R
 si celle-ci est jaune décider J

Calcul des pertes de cette stratégie.

Quelles sont les pertes qui résulteraient d'une telle stratégie ? Pour cela il faut examiner le cas de l'état R puis celui de l'état J. Dans chaque cas on ajoutera la perte due à la décision finale, telle qu'elle résulte du tableau de perte, au coût de l'échantillonnage.

Si l'état est R :

. avec une probabilité $2/3$ on tirera une boule rouge et on décidera R ce qui correspond à une perte : 0 puisque l'état est R

. avec une probabilité $1/3$ on tirera une boule jaune et on décidera donc J ce qui correspond à une perte : 50 puisque l'état est R

La perte moyenne qui résulte de la décision finale est donc :

$$2/3 \times 0 + 1/3 \times 50 = 16,6$$

Le coût de l'échantillonnage (1 seul tirage) est 1 si bien que la perte totale est :

$$16,6 + 1 = 17,6$$

Si l'état est J :

on verrait de la même manière que la perte due à la décision finale est :

$$1/3 \times 100 + 2/3 \times 0 = 33,3$$

à laquelle il faut ajouter le coût de l'échantillonnage qui est toujours 1

$$33,3 + 1 = 34,3$$

pour la perte totale.

Les deux pertes totales ainsi obtenues sont à porter dans une colonne 4 du tableau de perte.

1-état R	17,6
2-état J	34,3

Un exemple de stratégie séquentielle.

Une stratégie pure plus élaborée pourrait être basée sur la règle (ou fonction) de décision suivante :

stratégie n° 5 ; ne rien décider avant d'avoir tiré au moins deux boules ; si elles sont toutes rouges décider R si elles sont toutes deux jaunes décider J.

Restent les cas où l'on trouverait une boule de chaque couleur (R J ou J R) ; étant donné qu'on se trouve alors dans une situation d'ignorance aussi grande qu'au début il pourrait paraître logique d'opérer de même. On tirera donc à nouveau deux boules et on se ne décidera que si les deux boules sont de même couleur et ainsi de suite...

Une telle stratégie pour laquelle le nombre de tirages à effectuer n'est pas fixé a priori porte le nom de stratégie séquentielle.

C'est le plus souvent parmi les stratégies séquentielles que se trouvent les meilleures stratégies possibles.

Calcul des pertes de cette stratégies séquentielle.

Nous n'effectuerons le calcul des pertes de la stratégie n° 5 que dans le cas de l'état R (Le raisonnement à faire dans le cas de l'état J serait tout à fait analogue).

Pour cela nous déterminerons d'abord la perte due à la décision finale puis la perte due à l'échantillonnage, obtenant ainsi la perte totale qui est égale à la somme de ces deux nombres.

Probabilités de décider R ou J au premier coup et au deuxième coup.

Puisque l'état est R, la probabilité de :

$$\text{tirer R R est } 2/3 \times 2/3 = 4/9$$

$$\text{tirer J J est } 1/3 \times 1/3 = 1/9$$

$$\text{tirer R J ou J R } 2(2/3 \times 1/3) = 4/9$$

(On vérifie que tous les cas sont bien envisagés puisque le total de ces trois probabilités est égal à 1).

Donc, la probabilité de :

- décider R du 1er coup est 4/9
- décider J du 1er coup est 1/9
- de faire un second prélèvement est 4/9

Cette fraction $4/9$ des cas indécis se subdivise à son tour en 3 éventualités dont les probabilités sont :

- décider R au 2 ^è coup	:	$4/9 \times 4/9$
- décider J au 2 ^è coup	:	$4/9 \times 1/9$
- faire un 3 ^è prélèvement	:	$4/9 \times 4/9$
total	:	$4/9$

Probabilités de décider R ou J au n ième coup.

De proche en proche, de la même façon, on trouverait les probabilités de :

- décider R au n ième coup	:	$(4/9)^n$
- décider J au n ième coup	:	$(4/9)^{n-1} \times 1/9$
- être indécis au n ième coup	:	$(4/9)^n$
total (probabilité d'indécision au (n-1)ième coup		$(4/9)^{n-1}$

Probabilité de décider R ou J à un moment quelconque.

La probabilité de décider R à un moment quelconque est donc la somme de la progression géométrique :

$$4/9 + (4/9)^2 + \dots + (4/9)^n + \dots = 4/5 = 0,8$$

La probabilité de décider J à un moment quelconque est de même :

$$1/9 (1 + (4/9) + \dots + (4/9)^{n-1} + \dots) = 0,2$$

Quant à la probabilité de ne rien décider c'est la limite de $(4/9)^n$ pour n tendant vers l'infini, soit zéro. Pratiquement on se décidera donc en un nombre fini de coups.

Ainsi, la probabilité de n'avoir pas encore pris de décision au 10ième coup par exemple n'est que $(4/9)^{10} \approx 0,0003$, si bien qu'un tel évènement ne se produit que 3 fois sur 10.000 en moyenne !

Pertes correspondantes.

La perte due à la décision finale, l'état étant R, est donc :

0, avec la probabilité 0,8
et 50, avec la probabilité 0,2 soit, en moyenne :

$$0 \times 0,8 + 50 \times 0,2 = 10$$

Pertes dues à l'échantillonnage.

A ceci il faut ajouter le coût de l'échantillonnage. Or si la décision finale (R ou J) est prise au nième coup le nombre de prélèvements est 2 n avec une probabilité :

$$(4/9)^n + (4/9)^{n-1} \times 1/9 = 5/4 (2/3)^{2n}$$

(qui est la somme des probabilités des deux évènements exclusifs : décider R ou décider J ou n ième coup).

Le coût de l'échantillon est donc la somme de la série :

$$5/4 (2 \times (2/3)^2 + 4 \times (2/3)^4 + \dots + 2n (2/3)^{2n} + \dots) = 3,6$$

Pertes totales.

La perte totale est donc : $10 + 3,6 = 13,6$ dans le cas ou l'état est R.

On verrait de même que si l'état est J la perte totale de la stratégie n° 5 est : 23,6.

Introduction des stratégies du statisticien dans le tableau de perte.

En notant par $W_R (5)$ la perte totale de la stratégie 5, si l'état est R et par $W_J (5)$ la perte totale de cette même stratégie si l'état est J, on a :

$$W_R (5) = 13,6$$

$$W_J (5) = 23,6,$$

si bien que le tableau de perte peut être complété de la façon suivante en ce qui concerne les cinq stratégies envisagées (trois stratégies pures sans tirage et deux stratégies avec tirage) :

	Stratégies pures sans tirage (décision finale)			Stratégies pures avec tirage	
	1	2	3	4	5
1 Etat R	0	50	60	17,6	13,6
2 Etat J	100	0	60	34,3	23,6

On ne peut énumérer toutes les stratégies du statisticien.

On conçoit qu'en opérant de façon analogue on pourrait pour toute stratégie s du statisticien, déterminer les pertes $W_R (s)$ et $W_J (s)$ correspondantes.

Mais on imagine facilement que l'entreprise consistant à énumérer toutes les stratégies pures du statisticien avant de choisir entre elles est absolument sans espoir. On se trouve dans une situation analogue à celle du joueur de Bridge évoqué plus haut, sauf qu'au lieu d'être simplement très grand, le nombre des stratégies possibles est ici infini...

Il est donc nécessaire d'établir d'abord un critère de comparaison entre les stratégies, qui permette d'écarter d'emblée celles qui ne sont pas à considérer comme acceptables.

III - STRATEGIES ADMISSIBLES - STRATEGIES MIXTES -

Stratégies admissibles.

On dit qu'une stratégie est admissible s'il n'existe pas de stratégie qui la domine, c'est-à-dire qu'il n'existe pas parmi toutes les stratégies, de stratégie dont les pertes soient inférieures pour tous les états de la nature.

Ainsi la stratégie n° 3 n'est pas admissible car il existe la stratégie 2 pour laquelle $0 < 60$ et $50 < 60$. En jouant (3) on serait sûr de perdre plus qu'avec (2).

De même la stratégie (4) n'est pas admissible, elle est dominée par la stratégie (5).

Mais dans certains jeux il n'y a pas de stratégie admissible. Ainsi, dans l'exemple étudié, il n'y aurait pas de stratégie admissible avec tirage, si ceux-ci étaient gratuits ; à toute stratégie on pourrait en opposer une meilleure comportant des prélèvements plus nombreux.

Classes complètes.

On dit qu'une classe de stratégies est complète si, à toute stratégie située hors de la classe on peut opposer au moins une stratégie de la classe qui la domine.

Il y a toujours au moins une classe complète : celle de toutes les stratégies. Toute classe complète contient nécessairement les stratégies admissibles si elles existent.

Nous allons maintenant définir la notion de stratégie mixte en utilisant l'exemple du jeu de bluff cité dans le premier paragraphe.

Stratégies mixtes et pertes correspondantes.

Nous avons vu que si le joueur 1 s'en tient à l'une de ses stratégies pures, par exemple en bluffant en permanence et en espérant gagner 100, cette attitude risque assez tôt d'être décelée par le joueur 2 qui lui opposera B de sorte que le joueur 1, qui peut cependant très légitimement espérer un gain, ne gagnera très vite plus rien.

		Joueur 2	
		ne contrôle pas A	contrôle B
joueur 1	a = pas de bluff	0	50
	b = bluff	100	0

Mais si le joueur 1 ne craint pas de faire appel aux probabilités il peut alors imaginer d'adopter tantôt sa stratégie a, tantôt sa stratégie b et cela tout à fait au hasard de façon à déjouer toute tentative du joueur 2 pour deviner ses coups.

Il peut le faire, par exemple, en tirant à pile ou face à l'insu de son adversaire, entre ses deux stratégies pures (il blufferait alors "au hasard" mais "en moyenne" une fois sur deux).

Supposons maintenant que le joueur 2 oppose à cela sa stratégie A ; alors le joueur 1 gagne :

0 avec probabilité 1/2
100 avec probabilité 1/2,

soit, en moyenne : $0 \times 0,5 + 100 \times 0,5 = 50$

Si, au contraire, le joueur 2 lui oppose sa stratégie B, on voit de même que le tirage à pile ou face permet au joueur 1 de gagner en moyenne :

$$50 \times 0,5 + 0 \times 0,5 = 25$$

Si bien qu'en faisant intervenir sciemment le hasard le joueur 1 est sûr de gagner en moyenne au moins 25, quelle que soit la conduite de B.

Le joueur 1 peut même s'assurer davantage si au lieu de tirer à pile ou face il tire à l'aide d'un dé à 6 faces et ne bluffe pas si le dé montre 3, 4, 5 ou 6 (probabilité 2/3). En effet, que le joueur 2 adopte alors sa stratégie A ou sa stratégie B, le joueur gagne en moyenne :

$$0 \times 2/3 + 100 \times 1/3 = 50 \times 2/3 + 0 \times 1 = 33,33 \dots$$

De telles stratégies, obtenues par tirage au sort entre deux ou plus de deux stratégies pures, avec affectation de probabilités à chacune d'elles, sont désignées sous le nom de stratégies mixtes⁽¹⁾.

Stratégies mixtes dans les problèmes de décision.

Dans les problèmes de décision il n'y a souvent aucune raison pour que le statisticien s'interdise d'envisager de telles stratégies mixtes, dont on a vu le grand parti que l'on pouvait tirer.

Considérons par exemple la stratégie n° 6, mixte, obtenue en donnant une probabilité de 1/5 à la stratégie 2 et de 4/5 à la stratégie 5. Les pertes de cette nouvelle stratégie sont donc :

$$W_R(6) = 1/5 \times 50 + 4/5 \times 13,6 = 20,9$$

$$W_J(6) = 1/5 \times 0 + 4/5 \times 23,6 = 18,8$$

Nous verrons que de telles stratégies mixtes du statisticien jouent un rôle fondamental pour obtenir des classes complètes de stratégies.

Représentation géométrique remplaçant le tableau de perte.

L'ensemble des stratégies mixtes que l'on peut engendrer à l'aide des stratégies pures n'est pas commodément représentable par un tableau

(1) On dit aussi "stratégies combinées".

de perte. C'est pourquoi chaque fois que les états de la nature sont en nombre fini il est souhaitable de représenter le jeu par un diagramme géométrique.

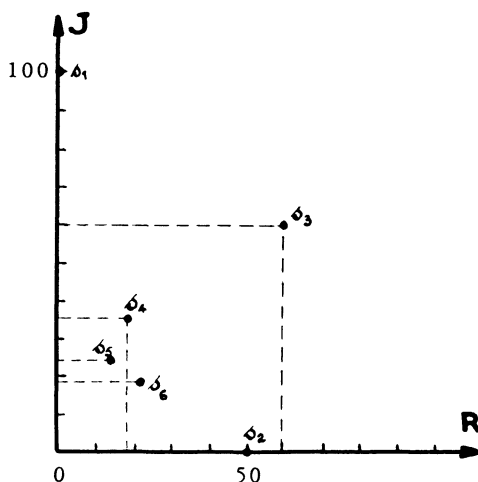
S'il y a N états de la nature, on fait correspondre à chacun d'eux un axe dans un système de coordonnées rectangulaires. Toute stratégie du joueur 2 est alors représentable par un point dont la coordonnée sur l'axe i est la perte W_i correspondant à l'état i de la nature.

Ici $N = 2$. Soient donc deux axes rectangulaires R et J .

La stratégie s_1 a pour coordonnées :

$$W_R = 0 \text{ sur l'axe } R$$

$$W_J = 100 \text{ sur l'axe } J$$



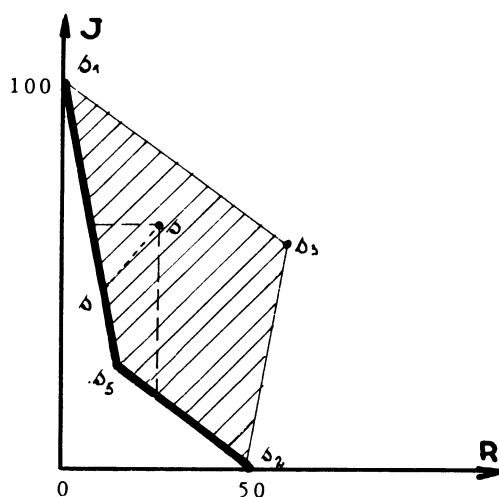
Sur la figure ci-dessus on a représenté les 6 stratégies envisagées jusqu'à présent.

On remarquera que la stratégie mixte s_6 est représentée par le centre de gravité des deux points s_2 et s_3 affectés des masses respectives $1/5$ et $4/5$.

Toutes les stratégies mixtes obtenues par mélange de deux stratégies pures sont figurées par le segment défini par les deux points représentatifs de ces deux stratégies pures.

D'une façon plus générale l'ensemble des stratégies mixtes que l'on peut obtenir à partir d'un ensemble de stratégies pures est représentable par la surface (frontière comprise) du polygone de sustentation engendré par les points figuratifs des stratégies pures.

Les stratégies mixtes que l'on peut obtenir à l'aide des stratégies 1 à 6 sont représentées sur la figure ci-après.



On voit immédiatement sur la figure que toute stratégie située hors de la partie de frontière s_1, s_5, s_2 , marquée en traits gras n'est pas admissible et qu'à chacune d'entre elles, telle que s , on peut en opposer une qui la domine et qui soit située sur la ligne considérée (une stratégie s' domine une stratégie s si le segment allant de s' vers s est dirigé dans le 1er quadrant des axes, limites comprises).

La ligne s_1, s_5, s_2 , ne constitue pas à elle seule une classe complète, mais il suffirait d'y ajouter les stratégies (encore inconnues) situées du côté de l'origine par rapport à elle pour en constituer une.

Stratégies mixtes de la nature.

Le joueur 1 peut également utiliser des stratégies mixtes (il pleut p jours sur q , les lots présentés au contrôle sont mauvais 3 fois sur cent etc.).

Dans notre exemple la nature peut présenter tantôt l'état R tantôt l'état J avec des probabilités respectives P_R et P_J : $P_R + P_J = 1$. Peut-être n'est-il pas réaliste de faire une telle hypothèse sur l'évolution des goûts, des couleurs. Peu importe, tant que nous ne choisissons pas arbitrairement une valeur de P_R - En effet on peut toujours considérer P_R comme un paramètre permettant de représenter notamment l'état R ($P_R = 1$ certitude de l'état R) ou l'état J ($P_R = 0$ certitude de l'état J) aussi bien que des cas intermédiaires.

On peut représenter graphiquement une telle stratégie mixte de la nature par une droite D passant par 0 et de pente :

$$r = \frac{P_J}{P_R}$$

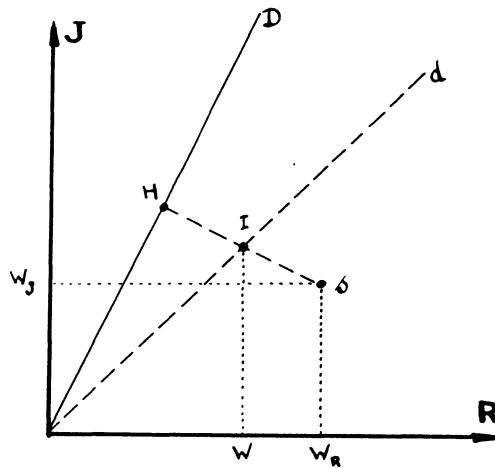
Cette droite peut venir se confondre avec l'un ou l'autre des axes, représentant ainsi l'Etat R ou l'Etat J purs.

Le jeu est alors transformé de la façon suivante. Indépendamment la nature choisit une droite D de pente r et le statisticien un point s parmi l'ensemble figuratif de ses stratégies.

La perte qui résulte de la combinaison s, D : $W_0(s)$ peut être obtenue en abaissant la perpendiculaire de s sur D, et en prenant son intersection avec la diagonale des axes (point I) de coordonnées W ; on a alors :

$$W_0(s) = W$$

Cette règle s'applique évidemment, en particulier, si D est confondu avec l'un des axes.



En effet, si s a pour coordonnées W_r (perte contre l'état R) et W_j (perte contre l'état J), la perte moyenne contre D s'obtient en pondérant ces deux pertes par leur probabilité :

$$W_0 = P_r W_r + P_j W_j$$

Cette expression est le produit scalaire du vecteur \vec{os} par le vecteur \vec{p} de coordonnées P_r, P_j porté par D. Or :

$$\begin{aligned} W_0 &= \vec{os} \cdot \vec{p} = \vec{OH} \cdot \vec{p} = \vec{OI} \cdot \vec{p} = \\ &= W P_r + W P_j = W (P_r + P_j) = W \end{aligned}$$

Cette construction met en évidence que toutes les stratégies s qui seraient situées sur une même droite sH, ont même perte contre D. Cette perte est d'autant plus faible que le point H est situé plus bas sur la droite D.

IV - STRATEGIES DE BAYES -

Thomas Bayes.

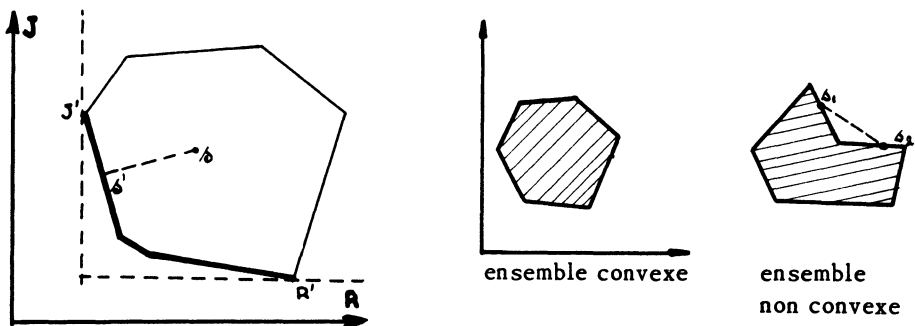
Thomas Bayes est un ecclésiastique britannique du 18ème siècle dont l'oeuvre posthume : "An Essay toward solving a problem in the doctrine of chance" parue en 1763 introduisit pour la première fois les notions de probabilité a priori et de probabilité a posteriori.

A cette époque évidemment il ne pouvait guère être question de théorie des jeux, mais nous verrons plus loin pourquoi le nom de Bayes est particulièrement justifié pour désigner toute une classe de stratégies jouant ici un rôle central.

L'ensemble convexe des stratégies.

Supposons le problème résolu et considérons l'ensemble des stratégies du joueur 2 (y compris les stratégies mixtes). Les points représentatifs de ces stratégies forment un ensemble convexe, c'est-à-dire un ensemble qui, s'il contient deux points s_1 et s_2 , contient tous les points du segment s_1, s_2 . Tout point de ce segment peut en effet être obtenu par une stratégie mixte déduite de s_1 et s_2 comme il a été constaté plus haut.

Nous admettrons ici que cet ensemble convexe contient sa frontière. Considérons alors la partie de cette frontière qui est constituée par ceux de ses points qui sont à la fois les plus bas et les plus à gauche (partie $J'R'$ en trait gras). On constate alors qu'à toute stratégie s située en dehors de $J'R'$ on peut opposer une stratégie de $J'R'$ qui la domine (par exemple s').



Par ailleurs il n'existe pas de stratégie qui domine une stratégie de la ligne $J'R'$.

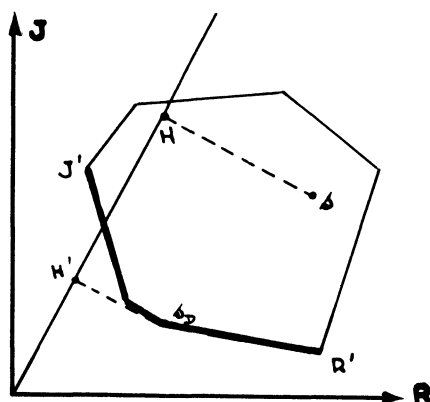
La ligne en trait gras représente donc une classe complète de stratégies admissibles.

Les stratégies de Bayes.

Soit maintenant une droite D représentative d'une stratégie mixte de la nature et cherchons, parmi les stratégies s celles qui rendent mi-

nimum la perte contre D. Il faut pour cela que le point H soit le plus bas possible comme nous l'avons remarqué plus haut.

On voit sur la figure que l'on obtient nécessairement au moins une solution S_0 si l'ensemble convexe contient sa frontière.



Les stratégies qui, comme s_0 , rendent minimum la perte contre D sont appelées stratégies de Bayes contre D.

La connaissance des stratégies de Bayes détermine les stratégies admissibles.

On voit que, faisant varier la droite D on peut obtenir toutes les stratégies admissibles (ligne J'R') et réciproquement. (Des cas spéciaux peuvent toutefois se produire si D est parallèle à un axe et que le contour comporte un segment parallèle à l'autre axe).

Le problème de la recherche des stratégies admissibles se ramène donc à celui de la recherche des stratégies de Bayes contre les diverses probabilités a priori attribuables aux états de la nature.

Dans certains problèmes on peut effectivement savoir que la nature agit selon un schéma probabiliste et appliquer la stratégie de Bayes correspondante. Dans d'autres problèmes on n'utilise la notion de probabilité a priori que pour décrire l'ensemble des stratégies admissibles.

Comment déterminer les stratégies de Bayes.

Revenons maintenant à l'exemple proposé et montrons sous quelle forme se présentent les stratégies de Bayes. Le raisonnement que nous utiliserons pourra s'adapter à la plupart des jeux statistiques comportant deux états de la nature, deux décisions finales et un coût d'échantillonnage proportionnel au nombre de tirages.

Nous constaterons d'abord qu'il n'est pas simple d'exprimer quelle est la perte contre D d'une stratégie quelconque (nous n'avons d'ailleurs pas décrit ici comment serait constituée une telle stratégie), pour ensuite trouver "celle" qui conduit au minimum, alors qu'il peut en exister plusieurs.

Pour tourner cette difficulté on procèdera par récurrence : on examinera s'il y a intérêt du point de vue de la perte contre D, à faire un premier tirage, puis dans l'affirmative s'il y a intérêt à en effectuer un second etc.. On obtiendra ainsi, une description détaillée, coup par coup, des stratégies de Bayes. Les solutions multiples se trouveront ipso facto éliminées si l'on convient d'adopter de préférence une stratégie basée sur une décision immédiate si celle-ci a même perte contre D qu'une stratégie équivalente basée sur un tirage supplémentaire.

Les stratégies s_1 et s_2 , sans tirage sont admissibles.

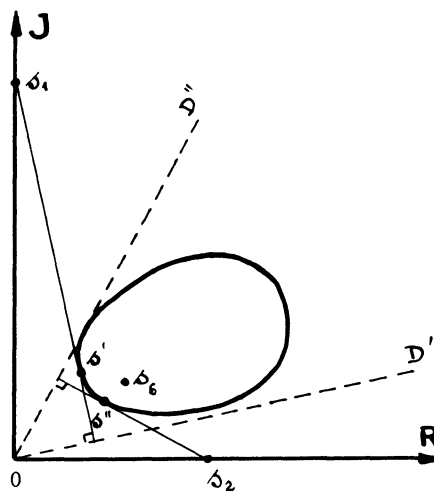
Il importe en premier lieu de remarquer que les stratégies s_1 et s_2 du statisticien (se décider sans tirer) sont nécessairement admissibles. A ce propos il paraît souhaitable d'attirer l'attention sur le fait que lorsqu'un terme (ici le terme admissible) a pris un sens mathématique précis il pourrait être désastreux de l'interpréter au sens courant : quel service de contrôle jugerait comme admissible de prendre ses décisions sans aucun examen des lots qui lui sont présentés ?

Pourtant tout tirage d'une boule augmente toutes les cases du tableau de perte d'une unité quelle que soit la suite donnée à ce tirage. Il en résulte que la perte contre R aussi bien que la perte contre J ne peuvent être inférieures à 1. Or aucune des stratégies telles que $W_R \geq 1$ et $W_J \geq 1$ ne peut dominer s_1 ni s_2 (les stratégies qui peuvent dominer s sont en effet dans le quadrant inférieur gauche par rapport à s). s_1 et s_2 sont donc admissibles au sens mathématique.

Faut-il effectuer un premier tirage.

Considérons alors l'ensemble convexe des stratégies "avec tirage". Cet ensemble contient s_0 qui est du côté de l'origine par rapport à la droite $s_1 s_2$.

Il existe donc deux droites $s_1 s'$ et $s_2 s''$ s'appuyant en s' et s'' sur le contour des stratégies avec tirage.



Menons par O les perpendiculaires D' et D'' à ces deux droites - D' a pour pente r' et D'' : r''.

Soit une droite D de pente $r > r''$; on voit que la stratégie de Bayes est s_2 car le pied de la perpendiculaire menée de s_2 sur D est le plus proche de 0.

De même, pour une droite D de pente $r < r'$; la stratégie de Bayes est s_1 .

Finalement la figure montre que ce n'est que lorsque :

$$r' < r < r''$$

qu'il faut effectuer un tirage au moins.

Dans les cas limites, $r = r'$ par exemple, on voit que l'on peut indifféremment utiliser s_1 , s' ou l'une des stratégies mixtes du segment $s_1 s'$, la perte contre D' étant la même.

Il en résulte que la règle à suivre est la suivante :

si $r \leq r'$ ou $r \geq r''$: décider sans tirage en faveur de s_1 ou s_2 selon le cas

si $r' < r < r''$: tirer au moins une boule

Reste à déterminer ce qu'il faut faire ensuite, dans ce dernier cas.

Faut-il ayant tiré une fois tirer une seconde fois.

Tirons donc une boule. Sa couleur x (x = rouge ou x = jaune) donne une information sur l'état de la nature.

Désignons par $f_R(x)$ la probabilité de tirer x sachant que l'état est R et par $f_J(x)$ la probabilité de tirer x sachant que l'état est J.

On a donc :

$$\begin{array}{ll} f_R(\text{rouge}) = 2/3 & f_R(\text{jaune}) = 1/3 \\ f_J(\text{rouge}) = 1/3 & f_J(\text{jaune}) = 2/3 \end{array}$$

La probabilité a priori de l'état R : P_R , est modifiée par la connaissance de x et devient :

$$P_{R/x}$$

probabilité a posteriori de l'état R, qui s'énonce "probabilité de l'état R sachant x'' ".

De même P_J devient $P_{J/x}$.

Or on démontre⁽¹⁾ en calcul des probabilités que :

$$P_{R/x} = \frac{P_R f_R(x)}{P_R f_R(x) + P_J f_J(x)}$$

(1) Les notions de probabilité a priori et a posteriori ainsi que l'établissement de la formule utilisée sont donnés en annexe.

$$P_{J/x} = \frac{P_J f_J(x)}{P_R f_R(x) + P_J f_J(x)}$$

si bien que, connaissant x , le rapport des probabilités de l'état jaune et de l'état rouge est devenu :

$$r_1 = \frac{P_{J/x}}{P_{R/x}} = \frac{P_J}{P_R} \cdot \frac{f_J(x)}{f_R(x)}$$

soit :

$$r_1 = r y_1$$

en posant $y = \frac{f_J(x)}{f_R(x)}$, nombre bien déterminé dès que x est connu.

On se trouve alors dans la même situation qu'avant tirage sauf que le rapport "a priori" r se trouve remplacé par le rapport "a posteriori" r_1 .

Il en résulte que la stratégie de Bayes se poursuit en appliquant la règle énoncée plus haut en y remplaçant r par r_1 :

si $r_1 \leq r'$ ou $r_1 \geq r''$: prendre la décision correspondante

si $r' < r_1 < r''$: faire un deuxième tirage

Ayant procédé à n tirages sans décision, que faire après le $(n + 1)$ ième tirage.

De proche en proche, on verrait qu'après n tirages on est encore dans l'indécision si le rapport r_n des probabilités de l'état J et de l'état R connaissant les n résultats : x_1, x_2, \dots, x_n , est tel que :

$$r' < r_n < r''$$

Un nouveau tirage permet d'observer x_{n+1} et de calculer :

$$y_{n+1} = \frac{f_J(x_{n+1})}{f_R(x_{n+1})} \text{ et } r_{n+1} = r_n y_{n+1},$$

nouveau rapport des probabilités a posteriori, connaissant les résultats déjà donnés et, en outre, x_{n+1} - On est encore dans la même situation qu'au début mais r est devenu r_{n+1} ; la règle à suivre est donc :

Si $r_{n+1} \leq r'$ décider R

si $r_{n+1} \geq r''$ décider J

si $r' < r_{n+1} < r''$ ne pas décider et continuer le processus.

Deux inconnues suffisent à déterminer toutes les stratégies de Bayes.

On voit donc que toutes les stratégies de Bayes peuvent être explicitement énoncées dès que l'on connaît r' et r'' . Si bien que le problème ne comporte que ces deux inconnues.

Dans le cas général, celles-ci dépendent du tableau de perte, du coût des essais et de la forme des fonctions $f_R(x)$ et $f_J(x)$ (le nombre de valeurs que x est susceptible de prendre n'est pas limité à deux).

Par ailleurs il existe des cas où l'ensemble convexe des stratégies avec tirage est situé en haut et à droite par rapport à la droite $s_1 s_2$. Aucun tirage n'est alors admissible. Quel que soit r il faut décider sans tirer !

V - LES REGLES DE SCORE -

Règles de score pour les stratégies de Bayes.

Du fait que pour obtenir r_{n+1} on doit multiplier r_n par y_{n+1} , il résulte que :

$$r_n = r y_1 y_2 \dots y_n$$

C'est l'expression explicite du rapport des probabilités a posteriori, connaissant x_1, x_2, x_n .

La double inégalité qui conduit à effectuer un nouveau tirage peut donc s'écrire :

$$r'/r < y_1 y_2 \dots y_n < r''/r.$$

En prenant le logarithme des deux membres et en posant :

$$-b = \log r'/r, \quad a = \log r''/r, \quad z_i = \log y_i.$$

On obtient la règle de décision équivalente :

$$-b < \sum_1^n z_i < a$$

Il s'agit là d'une règle dite de score, bien connue en contrôle statistique. Ce "score" est le cumul des valeurs z_i obtenues au cours de prélèvements successifs ; tant qu'il est compris entre les deux limites fixées $-b$ et a , on continue à prélever sinon on prend la décision conforme à la limite franchie.

Cas de l'exemple étudié.

Dans le cas de l'exemple proposé initialement :

$$\text{si } x = \text{rouge} \quad z = \log \frac{f_J(\text{rouge})}{f_R(\text{rouge})} = \log \frac{1/3}{2/3} = -\log 2$$

$$\text{si } x = \text{jaune} \quad z = \log \frac{f_J(\text{jaune})}{f_R(\text{jaune})} = \log \frac{2/3}{1/3} = +\log 2$$

si bien que les stratégies pures de Bayes sont de la forme (après division par $\log 2$) :

- partir de zéro
- compter - 1 pour une boule rouge
- compter + 1 pour une boule jaune

- décider R si le score devient $\leq \frac{-b}{\log 2}$
- décider J si le score devient $\geq \frac{a}{\log 2}$
- sinon continuer à prélever.

Les règles de score en contrôle statistique.

r' et r'' étant déterminés par le tableau de perte les stratégies pures de Bayes ne dépendent que d'un seul paramètre r ; on a entre a et b la relation :

$$a + b = \log \frac{r''}{r'}$$

Par contre, un contrôle statistique usuel (ou en théorie des tests d'hypothèses) on aboutit bien à des règles de score de même type mais a et b dépendent de deux paramètres :

α : risque de première espèce (ici : probabilité de décider J alors que l'état est R)

β : risque de première espèce (ici : probabilité de décider R alors que l'état est J).

On voit donc que si l'on se donne a et b de façon quelconque (ce qui revient à se donner α et β) on obtient une stratégie pure du statisticien qui n'est pas nécessairement admissible pour le tableau de perte considéré. Mais on montre cependant que ces stratégies constituent un optimum, considérées à un autre point de vue : du fait que la variable z attachée à x est déterminée par le logarithme de quotient des probabilités a priori de x , le prélèvement moyen, avant de parvenir à une décision, est minimum pour les risques α et β fixés.

Si l'on a choisi une stratégie de risques α et β donnés, il n'est pas difficile de calculer les pertes résultant des décisions finales. Ainsi, si l'on se fixe une stratégie comportant les risques $\alpha = \beta = 0,10$:

si l'état est R on décide R avec la probabilité $1 - \alpha = 0,9$ et J avec la probabilité $\alpha = 0,10$; la perte moyenne due à la décision finale est donc dans le cas de notre problème :

$$0,9 \times 0 + 0,1 \times 50 = 5$$

. si l'état est J on trouve de la même façon une perte moyenne :

$$0,10 \times 100 + 0,9 \times 0 = 10$$

Mais le calcul de la perte par échantillonnage est moins simple.

Prélèvement moyen et risques d'une règle de score.

Soit donc une variable susceptible de prendre des valeurs z avec une probabilité $g(z)$. Notons par :

\bar{n} ($-b, a$) le nombre moyen des valeurs à observer avant d'atteindre ou dépasser l'une des limites,

$\pi(-b, a)$ la probabilité d'atteindre ou dépasser la limite supérieure ($1 - \pi$ est donc la probabilité d'atteindre ou de dépasser la limite inférieure).

La relation exacte qui détermine \bar{n} et π en fonction de $g(z)$ et des limites n'est pas connue sous forme explicite dans le cas général.

Des valeurs approximatives peuvent être obtenues de la façon suivante.

On cherche la racine différente de zéro de l'équation en t :

$$1 = \sum_z e^{tz} g(z) \quad \text{et on pose } E = \sum_z z g(z)$$

On a alors si $-b < 0 < a$:

$$\pi(-b, a) \quad \# \quad \frac{e^{-bt} - 1}{e^{(a+b)t} - 1} \quad (1)$$

$$\bar{n}(-b, a) \quad \# \quad \frac{a \pi(-b, a) + (1 - \pi) b}{E} \quad (2)$$

Ces formules sont rigoureuses lorsque le score ne peut sortir des limites qu'en passant exactement par celles-ci.

C'est le cas des règles de Bayes relatives à notre problème initial car en prenant $z = \frac{1}{\log 2} \log \frac{f_J(x)}{f_R(x)}$, le score qui est alors somme de valeurs égales à ± 1 ne peut être qu'un nombre entier et rien n'est changé à la règle de score si l'on prend pour a et b le plus petit entier égal ou supérieur.

Si le passage de la limite a coûte W_a et le passage de la limite $-b$: W_b , en prenant toujours comme unité le coût du tirage d'une valeur z , la règle de score coûte en moyenne :

$$\rho(-b, a) = \bar{n}(-b, a) + \pi(-b, a) W_a + (1 - \pi) W_b$$

Cette perte dépend de $g(z)$.

Pertes d'une règle de score.

Si z peut suivre l'une ou l'autre de deux lois de probabilité $g_1(z)$ ou $g_2(z)$ selon l'état de nature on a deux pertes moyennes correspondantes:

$$\rho_1(-b, a) \text{ et } \rho_2(-b, a)$$

Dans notre problème :

$$\begin{aligned} g_1(-1) &= 2/3 & g_1(+1) &= 1/3 \\ g_2(-1) &= 1/3 & g_2(+1) &= 2/3 \end{aligned}$$

Si bien que t_1 et t_2 , sont donnés par :

$$1 = e^{-t} g_i(-1) + e^{+t} g_i(+1)$$

qui a pour racines non nulles $t_1 = \ln 2$ ou $t_2 = -\ln 2$

ρ_1 et ρ_2 se trouvent donc donnés par les formules (1) et (2).

Limites du score en contrôle statistique.

Remarquons enfin que lorsque, comme il est d'usage, on prend comme variable de score :

$$z = \log \frac{f_2(x)}{f_1(x)}$$

on a alors $t_1 = \ln 10$ et $t_2 = -\ln 10$, et les deux égalités (1) résolues en a et b donnent les formules classiques :

$$a = \log \frac{1 - \beta}{\alpha}$$

$$- b = \log \frac{\beta}{1 - \alpha}$$

tandis que les formules 2 donnent le prélèvement moyen de la règle de score.

VI - DETERMINATION DES STRATEGIES DE BAYES -

Mode de calcul des deux inconnues r' et r'' .

Nous avons vu que la connaissance des deux inconnues r' et r'' suffisait à déterminer toutes les stratégies de Bayes. On obtiendra ces inconnues en résolvant les équations exprimant que pour $r = r'$ et $r = r''$, la perte totale est la même pour une stratégie (règle de score) avec tirage que pour une décision immédiate. Les limites de score correspondant à ces deux cas ne font intervenir que r' et r'' :

$$\text{contre } r = r' \quad - b = \frac{\log r'/r'}{\log 2} = 0 \quad a = \frac{\log r''/r'}{\log 2} = a_0$$

$$\text{contre } r = r'' \quad - b = \frac{\log r'/r''}{\log 2} = -a_0 \quad a = \frac{\log r''/r''}{\log 2} = 0$$

On trouve finalement, dans le cas de notre exemple :

$$a_0 = 7 \quad (\text{valeur arrondie à l'entier supérieur, du fait que le score ne prend que des valeurs entières})$$

$$r' = 0,075$$

$$r'' = 4,866$$

Il en résulte que l'ensemble des stratégies pures de Bayes contre r sont les suivantes :

- si $r \leq 0,075$ décider R sans tirage
- si $r \geq 4,866$ décider J sans tirage
- si $0,075 < r < 4,866$ utiliser la règle de score énoncée plus haut avec les limites :

$$- b = - \left[\frac{\log \frac{r}{0,075}}{\log 2} \right] \quad a = \left[\frac{\log \frac{4,866}{r}}{\log 2} \right]$$

où le crochet signifie qu'il faut prendre le plus petit entier égal ou supérieur.

Exemple de stratégie de Bayes.

Ainsi, si l'on sait (par une étude préalable de l'évolution des goûts de la clientèle) que l'état R et l'état J ont a priori une probabilité de 0,5, les limites du score seront (car alors $r = 1$) :

$$-b = - \left[\frac{\log \frac{1}{0,075}}{\log 2} \right] = -4 \qquad a = \left[\frac{\log 4,866}{\log 2} \right] = +3$$

La stratégie de Bayes correspondante consiste donc à effectuer des tirages en comptant +1 pour une boule jaune -1 pour une boule rouge jusqu'à ce que le total devienne égal à +3 ou à -4. Si l'on atteint +3 on décide J si l'on atteint -4, on décide R.

Description de toutes les stratégies pures de Bayes pour le problème étudié.

En donnant à r toutes les valeurs :

$$0,075 < r < 4,866$$

un calcul analogue montrerait qu'il existe 13 stratégies pures de Bayes avec tirage correspondant aux treize couples de limites (-b, a) suivantes :

valeur de b	-1	-1	-2	-2	-3	-3	-4	-4	-5	-5	-6	-6	-7
valeur de a	+7	+6	+6	+5	+5	+4	+4	+3	+3	+2	+2	+1	+1

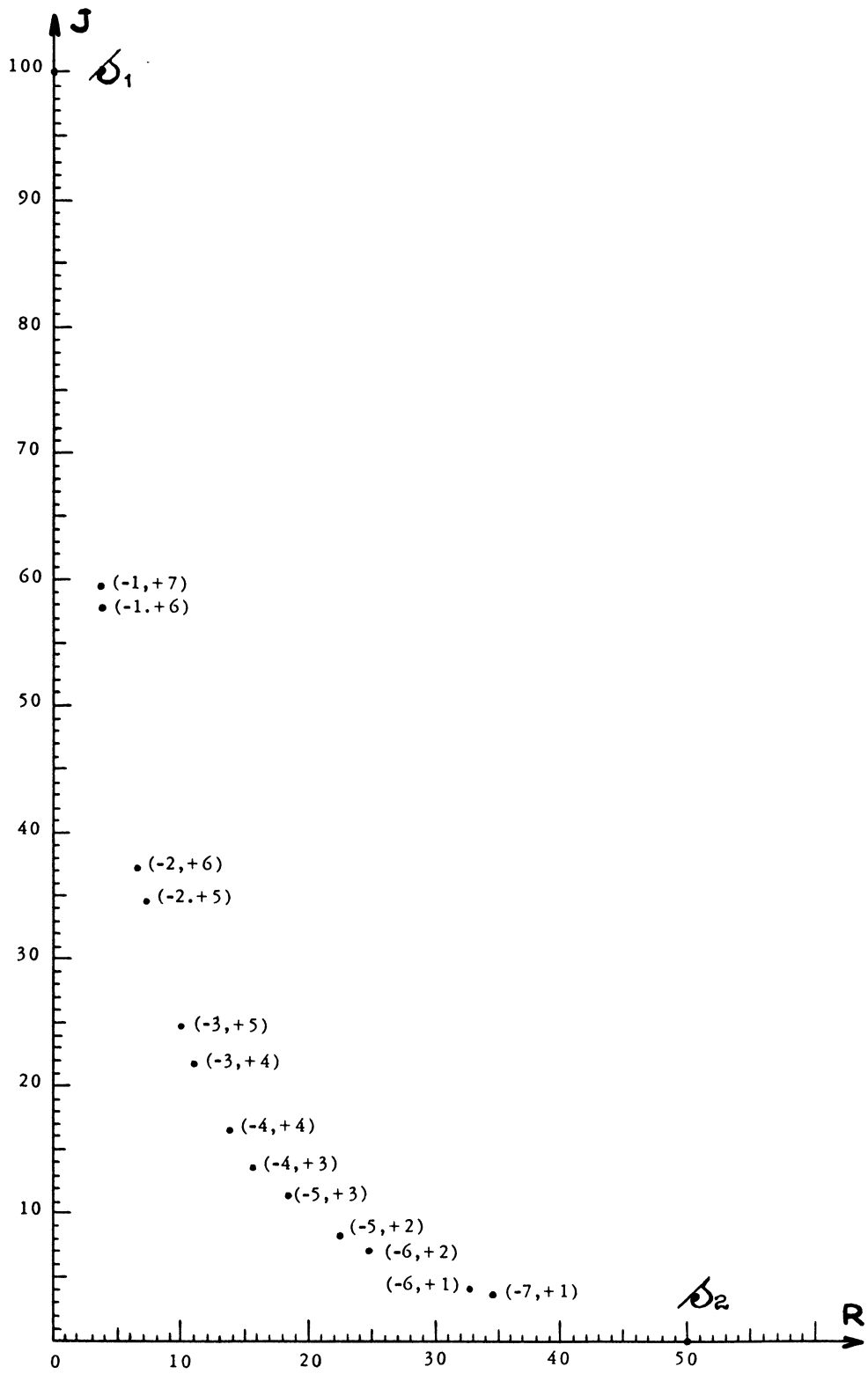
citées dans l'ordre croissant des valeurs de r, c'est-à-dire dans l'ordre décroissant des probabilités a priori P_R de l'état R. La stratégie s_1 (décision R sans tirage) se placerait la première à gauche de cette liste car c'est la stratégie de Bayes contre $r = 0$; la stratégie s_2 (décision J sans tirage) se placerait à l'autre extrémité car c'est la stratégie de Bayes contre $r = +\infty$ ou $P_R = 0$.

Il existe 14 types de stratégies mixtes de Bayes qui peuvent être obtenues en affectant des probabilités à deux stratégies pures successives de la liste des 15 stratégies ainsi énumérées.

Si l'on calcule les pertes totales W_R et W_J , de chacune des 15 stratégies pures on obtient le diagramme géométrique qui correspond au problème posé. (A côté de chaque point représentatif d'une stratégie pure on a indiqué entre parenthèses les limites (-b, a) du score correspondant). On n'a pas figuré pour plus de clarté les segments rectilignes représentatifs des stratégies mixtes.

lère méthode générale de résolution.

D'une façon générale lorsqu'il y a deux états de nature et deux décisions finales, le coût d'un tirage étant constant, on prendra celui-ci comme unité monétaire. De plus r' et r'' n'étant pas modifiés par une translation de l'ensemble des points représentatifs des stratégies on peut toujours amener le tableau de perte à la forme :



O	W_{12}
W_{21}	O

avec $W_{12} > 0$, $W_{21} > 0$ en ajoutant ou retranchant un nombre convenable à chaque ligne (on suppose évidemment que l'une des décisions ne domine pas l'autre car dans le cas contraire il n'y a plus de problème).

La connaissance des deux lois de probabilités $f_1(x)$ et $f_2(x)$ correspondant aux états de nature 1 ou 2 permet de déterminer les lois $g_1(z)$ et $g_2(z)$ de :

$$z = \frac{f_2(x)}{f_1(x)}$$

Si l'on peut déterminer $\rho_1(-b, a)$ et $\rho_2(-b, a)$, r' et r'' sont solutions des deux équations :

$$\begin{aligned} W_{12} &= 1 + r' + W_{12} \sum_{z \geq 0} g_1(z) + r' W_{21} \sum_{z < -b_0} g_2(z) + \\ &+ \sum_{-b_0 < z < 0} \rho_1(-b_0 + z, -z) g_1(z) + r' \rho_2(-b_0 + z, -z) g_2(z) \\ r'' W_{21} &= 1 + r'' + W_{12} \sum_{z \geq b_0} g_1(z) + r'' W_{21} \sum_{z \leq 0} g_2(z) + \\ &+ \sum_{0 < z < b_0} \rho_1(-z, b_0 - z) g_1(z) + r'' \rho_2(-z, b_0 - z) g_2(z) \end{aligned}$$

ou $b_0 = \log r'' - \log r'$

Réciproquement si l'on se donne r' et r'' on peut en déduire le tableau de perte correspondant. Les deux équations sont alors linéaires en W_{12} et W_{21} .

2ème méthode générale de résolution.

On peut aussi déterminer r' et r'' par approximations successives. Pour cela notons par p la probabilité a priori de l'état 1 et par $\rho(p)$ la perte inconnue de la stratégie de Bayes contre p .

La méthode consiste à déterminer une suite croissante de limites inférieures $L''(p)$ et une suite décroissante de limites supérieures $L'(p)$:

$$L''_n < \rho(p) < L'_n,$$

qui tendent vers $\rho(p)$ et déterminent donc la perte de Bayes par approximations successives.

Pour cela on utilise les formules :

$$L_n(p) = \min \left\{ \psi(p), 1 + \sum_x (p f_1(x) + (1-p) f_2(x)) L_{n-1} \left(\frac{p f_1(x)}{p f_1(x) + (1-p) f_2(x)} \right) \right\}$$

avec :

$$\psi(p) = \min \{ p W_{12} \quad (1-p) W_{21} \}$$

L'_n est obtenu en prenant comme valeur initiale $L'_0(p) = \psi(p)$

L''_n en prenant $L''_0(p) = 0$.

A chaque fonction $L_n(p)$ on peut faire correspondre deux valeurs de p : p_1 et $p_2 < p_1$ qui sont les deux solutions de l'équation :

$$\psi(p) = 1 + \sum_x (p f_1(x) + (1-p) f_2(x)) L_n \left(\frac{p f_1(x)}{p f_1(x) + (1-p) f_2(x)} \right)$$

alors $(1 - p_1)/p_1$ tend vers r'

$(1 - p_2)/p_2$ tend vers r''

VII - STRATEGIES MINIMAX -

Quelle stratégie choisir parmi les stratégies admissibles ?

Dans le cas particulier correspondant aux hypothèses faites on a trouvé une classe complète de stratégies admissibles.

Il serait absurde de choisir une stratégie hors de cette classe alors que, dans la classe on peut trouver une stratégie qui la dominerait. Mais, dans la classe de stratégies admissibles aucune n'en domine une autre.

Dans ces conditions, quel critère faut-il adopter pour choisir entre les stratégies ainsi trouvées ?

Certes, si l'on peut attribuer aux états de la nature une probabilité à priori (avec $\frac{P_j}{P_r} = r$) il convient de choisir la stratégie de Bayes contre r (il peut d'ailleurs en exister plusieurs équivalentes) qui minimise la perte correspondante.

Mais connaître la droite D de pente r revient à connaître la stratégie mixte de l'adversaire et bien que, dans les jeux contre un adversaire indifférent, ce puisse être le cas il est nécessaire de pouvoir faire face à des circonstances où r est inconnu.

Dans ce cas, en théorie des jeux, on choisira une stratégie minimax au sens que nous allons maintenant définir. Mais en théorie statistique des décisions une telle réponse est douteuse parce que la nature n'a généralement aucune raison a priori pour présenter le même comportement qu'un joueur intelligent dont les intérêts seraient opposés à ceux du statisticien.

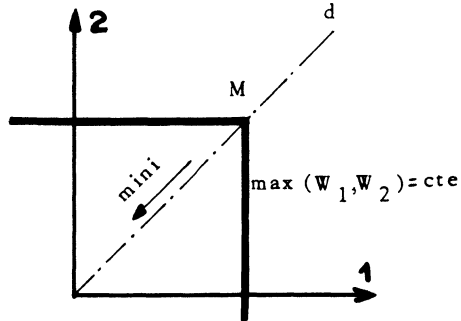
Définition des stratégies minimax.

Une stratégie minimax, minimise le maximum de la perte du joueur considéré.

Si l'on connaît les stratégies admissibles du statisticien et que celles-ci forment une classe complète on peut facilement déterminer graphiquement les minimax.

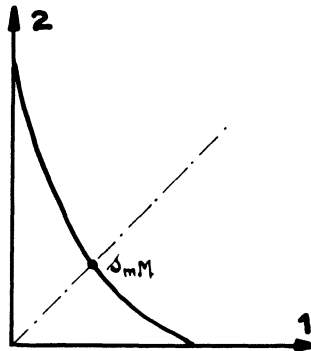
Soit une stratégie s du joueur 2 dont les pertes sont W_1 contre l'état 1 et W_2 contre l'état 2. La plus grande perte contre une stratégie quelconque (pure ou mixte) de la nature est :

$$W_M = \max (W_1, W_2)$$



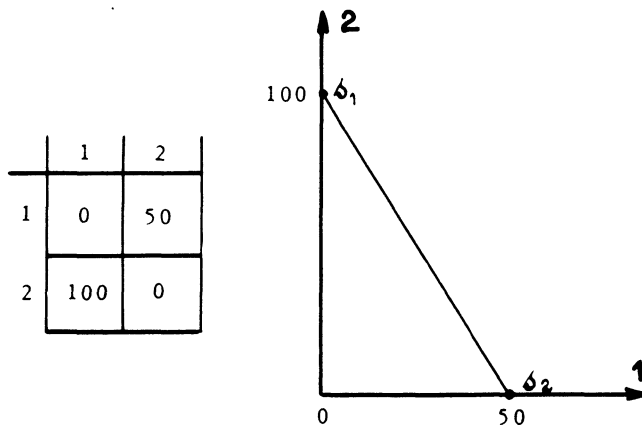
On voit sur la figure que les lignes $\max (W_1, W_2) = cte$ sont formées de deux segments à angle droit se coupant sur la diagonale d en un point M . Les stratégies minimax sont donc telles que ce point M est le plus bas possible.

Dans le cas de figure ci-dessous, le point représentatif de la stratégie minimax est s_{mm} intersection de la diagonale d avec la ligne des stratégies admissibles (ce point peut représenter plusieurs stratégies équivalentes et, dans certains cas de figure il peut y avoir des stratégies minimax situées hors de la diagonale).



Instabilité des stratégies pures dans un jeu à deux personnes.

Supposons maintenant que l'on ait à considérer un jeu ordinaire entre joueurs intelligents - Pour fixer les idées prenons le jeu de bluff déjà décrit et dont le tableau de perte et la représentation graphique sont les suivants :



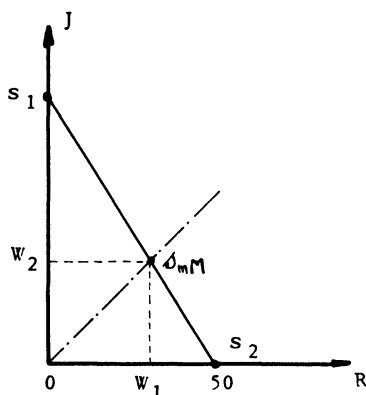
Si l'on exclut les stratégies mixtes, la stratégie minimax du joueur 2 est sa deuxième stratégie car $50 < 100$; les deux stratégies du joueur 1 sont minimax pour lui (perte max du joueur 1 : - 0 dans les deux cas).

Si le joueur 1 pense que le joueur 2 adopte sa stratégie maximax il a intérêt à lui opposer plutôt sa stratégie 1. Mais le joueur 2 peut imaginer que le joueur 1 raisonne ainsi et alors il n'a pas intérêt à jouer sa stratégie minimax mais sa stratégie 1 qui lui assure dans ces conditions une perte nulle. Mais le joueur 1 imaginant cela à son tour aurait intérêt à choisir alors sa stratégie 2 ; d'où pour le joueur 2 sa stratégie 2 ... etc..

Ce cercle vicieux disparaît si l'on introduit les stratégies mixtes et la conduite à tenir par chacun des joueurs peut être déterminée.

Stratégie minimax du joueur 2.

En effet supposons que le joueur 2 joue sa stratégie minimax ; c'est alors la stratégie mixte s_{mM} intersection de s_1 s_2 avec la diagonale des axes, et de coordonnées $W_1 = W_2 = W$. C'est un mélange de s_1 et s_2 dans le rapport :



$$\frac{50 - W}{W} = \frac{W}{100 - W} = \frac{50}{100}$$

soit $1/3$ pour s_1 et $2/3$ pour s_2 avec $W = 33,3$.

On voit que quelle que soit la stratégie du joueur 1 le joueur 2 peut assurer une perte de $W = 33,3$.

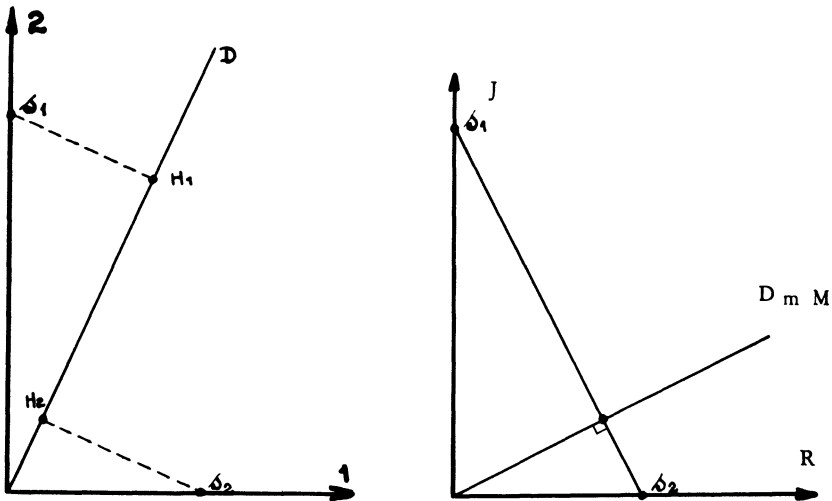
Stratégie minimax du joueur 1.

Soit maintenant une stratégie D du joueur 1. Menons par s_1 et s_2 les perpendiculaires à D : $s_1 H_1$ et $s_2 H_2$.

La stratégie minimax de ce joueur (maximin de son gain) est celle qui confond H_1 avec H_2 . C'est donc la perpendiculaire menée par O à la ligne $s_1 s_2$, D_{mm} .

Alors, que le joueur 2 joue s_1 s_2 ou n'importe quelle stratégie mixte la perte est la même et est égale à

$$W = 33,3.$$



Les stratégies minimax fixent un point d'équilibre pour les jeux à deux personnes.

En jouant leur stratégie minimax les deux joueurs peuvent quelle que soit la stratégie de l'autre assurer une perte $W = 33,3$ - Ce nombre est appelé la valeur du jeu, et les remarques suivantes justifient cette dénomination.

Supposons que l'un des joueurs s'écarte de sa stratégie minimax où nous supposons aussi que les deux joueurs sont initialement placés. Ceci ne change rien, la perte est toujours $W = 33,3$. Le joueur qui s'écarte de sa stratégie minimax n'a rien à y gagner. Par contre il risque, ce faisant, que l'autre joueur s'en apercevant lui inflige la stratégie de Bayes correspondante contre laquelle il ne peut que perdre davantage. On voit que pour que le jeu soit équitable il faut que la "valeur du jeu" soit nulle.

On retrouve ainsi sur un cas particulier l'énoncé central de la théorie des jeux finis à deux personnes et à somme nulle. Résoudre le jeu c'est trouver les stratégies minimax de chacun des joueurs et on utilise généralement pour trouver les solutions les méthodes de programmation linéaire.

Dans les problèmes de décision statistique on peut certes aussi s'assurer la perte minimax. Mais il n'y a généralement pas à craindre que si l'on s'en écarte pour appliquer une stratégie de Bayes, la nature s'en aperçoive...! Il peut donc y avoir mieux à faire si l'on dispose d'information a priori sur la fréquence avec laquelle la nature prend ses états.

Stratégie minimax pour l'exemple étudié.

La figure de la page 25 permet de trouver la stratégie minimax dans le cas étudié. L'intersection avec la diagonale montre qu'il s'agit d'un mélange des stratégies pures de limites :

$$(-4, 3) \text{ et } (-4, 4)$$

dans les proportions 0,586... et 0,413... respectivement.

La perpendiculaire menée par 0 à ce segment donne la perte r_{nm} de la stratégie minimax de la nature : $r_{nm} = \frac{0,378}{0,621}$.

C'est la stratégie de la nature la plus défavorable pour le statisticien.

La valeur du jeu (perte minimax du statisticien) est 14,642, si bien qu'en adoptant sa stratégie minimax le statisticien peut s'assurer en moyenne cette perte quelle que soit la stratégie de la nature.

VIII - CRITERES DE CHOIX -

Pour choisir une stratégie parmi les stratégies admissibles, nous avons défini jusqu'à présent deux critères :

- le critère de Bayes, choix d'une stratégie de Bayes contre la probabilité à priori de la nature.

- le critère du minimax des pertes : choix d'une stratégie minimax.

Objections au critère de Bayes.

Le critère de Bayes semble s'imposer lorsque la probabilité à priori est connue. Certaines objections lui ont cependant été faites⁽¹⁾.

Lorsque la probabilité a priori est inconnue il a été proposé de prendre arbitrairement une même probabilité pour les divers états de la nature ($P_R = P_J = 1/2$ s'il y a deux états). Mais une telle façon de faire, toute subjective et sans support expérimental réel paraît généralement à exclure (savoir qu'à pile ou face une pièce est équilibrée n'est certainement pas équivalent à ignorer si elle l'est). En tous cas, au point de vue statistique les conséquences d'une telle convention devraient alors être évaluées avec précision dans le cadre du problème considéré. Le choix d'une stratégie de Bayes contre une équirépartition des probabilités a priori, non justifiée par une étude préalable, est désignée parfois sous le nom de critère de Laplace.

Objections au critère du minimax.

Le critère du minimax des pertes est applicable dans tous les cas, mais des objections fondamentales peuvent lui être faites. Considérons, par exemple, le tableau de perte II qui est obtenu en retranchant 50 à la première ligne du tableau I et en ajoutant 50 à la deuxième ligne de ce même tableau.

	décision R	décision J
état R	0	50
état J	100	0

Tableau I

	décision R	décision J
état R	- 50	0
état J	150	50

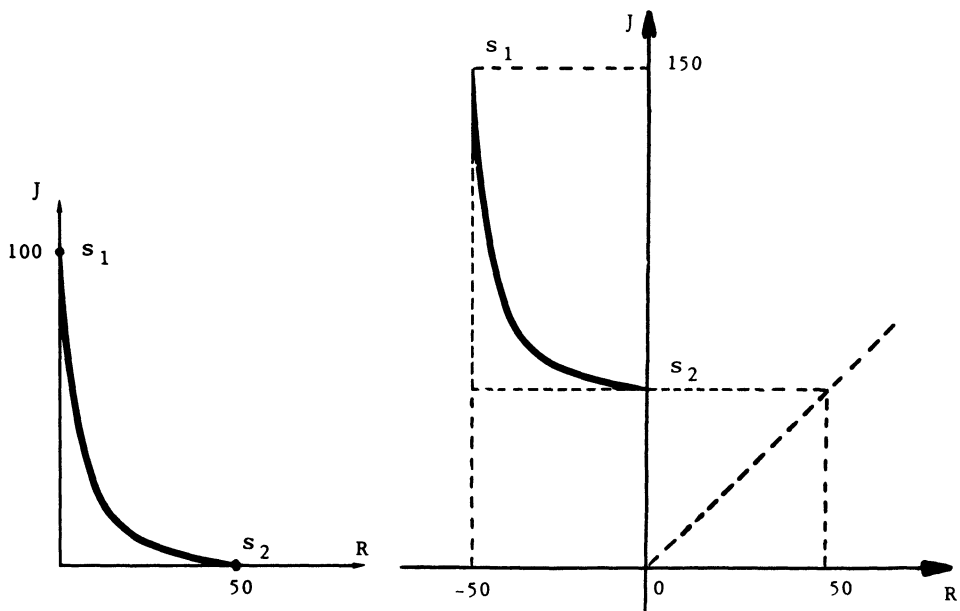
Tableau II

l'ensemble figuratif des stratégies, relatif au tableau II, se déduit simplement de celui relatif au tableau I par une translation de tous les points s . Les composantes de cette translation sont - 50 selon l'axe R et + 50 selon l'axe J. La figure 1 donne le contour des stratégies admissibles tel qu'il résulte du tableau I et on voit sur la figure II que c'est la stratégie s_2 (décision J sans tirage) qui est la stratégie minimax du tableau II.

Ainsi, bien que l'état J conduise maintenant à des pertes plus élevées et que l'état R soit devenu très favorable, on renoncerait à tout échantillonnage pour se décider comme si l'on était certain que l'état J devait se produire ! On aboutirait encore à la même conclusion si le chiffre qui figure dans la case état R - décision R était - 1.000.000 au lieu de - 50 : alors, sous prétexte de ne pas risquer de perdre 150 on refuserait de faire quelques prélèvements pour voir s'il n'y a pas des chances de gagner 1.000.000.!

Il semble que la seule justification que l'on puisse imaginer à une conduite aussi déraisonnable serait que la nature risque de se rendre compte que sa stratégie "état J" domine considérablement sa stratégie "état R" et n'a aucune raison, si elle est douée de la moindre intelligence, de choisir l'état R.

(1) cf réf 2 - section 4. 3 -



On conçoit que dans de nombreux problèmes cette attitude anthropomorphique est sans fondement, et que la théorie des jeux usuelles ne s'applique pas.

Critère du minimax des regrets.

Au lieu de considérer les pertes elles-mêmes on porte son attention le plus souvent en théorie des décisions, sur les regrets. Ainsi, si la nature prenait l'état J, le tableau II indiquerait qu'en tout état de cause, on perdrait au moins 50. Face à l'état J, contre lequel on ne peut rien, la décision R conduirait à regretter : $150 - 50 = 100$ et la décision J, à regretter $50 - 50 = 0$. De même, si par une heureuse coincidence, la nature prenait l'état R, on n'aurait rien à regretter d'une décision R mais une décision J donnerait un regret de $0 - (-50) = 50$.

C'est pourquoi le tableau I est dit être le tableau des regrets du tableau II. Pour ces deux tableaux considérés comme tableaux de pertes les stratégies de Bayes sont les mêmes, ainsi que les stratégies admissibles. Par contre les stratégies minimax, comme nous l'avons vu, diffèrent.

Certains auteurs ont suggéré de choisir comme critère le minimax des regrets ; c'est-à-dire qu'au lieu de prendre la stratégie minimax du tableau II on prend celle du tableau I.

Il en résulte que les deux pertes de la stratégie minimax des regrets sont, pour le tableau II⁽¹⁾ :

(1) Rappelons qu'il a été déterminé plus haut que la valeur du jeu, avec tirage, pour le tableau I, est 14,642.

$$W_R = 14,642 - 50 = -35,358 \quad (\text{gain})$$

$$W_J = 14,642 + 50 = 64,642 \quad (\text{perte})$$

Objections au critère du minimax des regrets.

Mais, là encore, des objections fondamentales peuvent être faites à l'emploi systématique du critère du minimax des regrets. On montre en effet, que contrairement au critère de Bayes et au critère minimax des pertes, ce critère ne satisfait pas à la condition suivante qui semblerait cependant devoir être respectée par tout critère raisonnable :

- si le critère conduit, face à deux stratégies s_1 et s_2 , à choisir s_1 la prise en considération d'une troisième stratégie s_3 , ne doit pas conduire à choisir s_2 .

Pour montrer le non respect de ce principe, considérons le tableau de perte suivant, relatif à 4 stratégies s_1 à s_4 :

	s_1	s_2	s_3	s_4
état R	0	50	80	200
état J	100	50	30	0

C'est également un tableau de regrets (car chaque ligne comporte un zéro et des pertes positives) et la stratégie de regrets minimax est donc s_2 .

Supposons que pour une raison quelconque s_1 ne puisse être envisagée alors le choix doit s'opérer entre s_2 , s_3 et s_4 . Les regrets correspondants sont donnés par le tableau :

	s_2	s_3	s_4
état R	0	30	150
état J	50	30	0

Cette fois c'est s_3 qui est à retenir et non plus s_2 . La relation de préférence entre s_2 et s_3 s'inverse donc par l'intervention de la stratégie s_1 dont l'application est hors de question.

Critères basés sur les risques.

Evoquons enfin un dernier type de critère de nature différente de celle des précédents.

Un usage constant et qui a fait ses preuves en statistique classique est de choisir la règle de décision en se basant sur les risques de première et de seconde espèce. Ainsi la stratégie s_1 (décider R sans tirer), qui comporte un risque de première espèce $\alpha = 0$ mais un risque de seconde espèce $\beta = 1$ ne serait pas prise en considération. La stratégie s_2 serait écartée pour la même raison ($\alpha = 1$, $\beta = 0$).

Il semble que, même si l'on ne juge pas convenable de prendre comme critère celui qui consisterait à minimiser, parmi l'ensemble des stratégies admissibles, le plus grand des deux risques (minimax $\{\alpha, \beta\}$),

il serait conforme à des usages établis de faire intervenir ces risques dans les critères de choix. Ceci correspondrait au fait que le statisticien a non seulement le souci de ne pas faire de dépenses inconsidérées mais aussi celui de ne pas se tromper trop souvent. L'objection faite au critère de minimax des pertes n'est elle pas précisément de ne pas être capable de prévoir le comportement d'une nature indifférente ?

IX - CONCLUSION -

Nous avons montré que lorsqu'un problème de décision, comportant la possibilité d'effectuer des investigations, a pu être complètement posé, diverses conduites raisonnables se trouvent déterminées et d'autres conduites sont reconnues à proscrire.

Si l'on parvient à décrire les divers états possibles de la nature, les décisions finales que l'on peut prendre et les pertes possibles qui en résulteraient, le problème devient un jeu dans lequel les stratégies du statisticien peuvent être évaluées du point de vue des pertes totales auxquelles elles conduisent (investigations comprises).

La solution de ce jeu n'est pas toujours aussi simple que nous nous sommes efforcé de le montrer, mais de toutes façons un problème convenablement posé n'est il pas déjà, plus qu'à demi, résolu ?

Puissent en tous cas ces quelques pages suggérer à chacun que, dans son domaine, un effort profitable peut vraisemblablement être fait en vue d'un meilleur traitement de certaines questions techniques ou économiques qui sont encore actuellement abordées par des méthodes anciennes moins satisfaisantes.

BIBLIOGRAPHIE

- A. WALD - Statistical Decision Functions - (J. Wiley 1950) ouvrage fondamental d'accès difficile.
- D. BLACKWELL et M.A. GIRSHICK - Theory of Games and Statistical Decisions (J. Wiley 1950) ouvrage fondamental d'accès difficile.
- H. CHERNOFF and L. E. MOSES - Elementary Decision Theory (J. Wiley 1959). Les auteurs ont fait un effort considérable pour rendre plus facilement accessible une partie de la théorie, et y sont parvenus.

NOTE ANNEXE

PROBABILITES A PRIORI ET A POSTERIORI

Probabilité des hypothèses en tenant compte des observations.

Considérons une épreuve aléatoire et son ensemble fondamental E des "cas possibles". Soient deux événements relatifs à cette épreuve : A et B. Ces deux événements, constitués de "cas possibles", sont des parties de E, au sens de la théorie des ensembles (on supposera $A \neq B$).

Supposons que A soit observable (ou son contraire A') et que B ne le soit pas (ni son contraire B'). La réalisation de B n'est alors qu'hypothétique, ce qui justifie le langage suivant :

A et son contraire A' sont les deux observations possibles
 B et son contraire B' sont les deux hypothèses possibles.

Le problème posé ici est alors le suivant : on observe x ($x = A$ ou $x = A'$), que peut-on dire de B ?

Si l'hypothèse B a *a priori* la probabilité $Pr(B)$, x nous apprend que, dans l'ensemble E seuls les cas vérifiant x sont maintenant possibles. La probabilité de B s'en trouve donc généralement modifiée et devient la probabilité *a posteriori* ou probabilité de B sachant x qui se note

$$Pr(B/x)$$

Nous allons montrer comment se calcule cette nouvelle probabilité.

L'épreuve aléatoire dans le cas du problème étudié.

Pour éviter les abstractions, nous raisonnerons sur l'exemple étudié dans le texte, en nous plaçant d'abord dans le cas du tirage d'une seule boule.

L'épreuve aléatoire à considérer ici est celle qui conduit à l'une des 4 possibilités exclusives :

- l'état est R et on tire x = rouge
- l'état est R et on tire x = jaune
- l'état est J et on tire x = rouge
- l'état est J et on tire x = jaune

Soient a, b, c, d les probabilités correspondantes (proportions théoriques) que l'on peut faire figurer dans les cases du tableau suivant

	x = rouge	x = jaune
Etat R	a	b
Etat J	c	d

On a : $a + b + c + d = 1$ puisque tous les cas sont envisagés et que ceux-ci sont exclusifs les uns des autres.

R et J (= R') sont les hypothèses sur lesquelles on désire être informé par l'observation de x.

Probabilité a priori des hypothèses.

Les cas favorables a priori à R sont en proportion a + b, si bien qu'avec nos notations antérieures :

$$P_R = a + b \quad (1)$$

de même :

$$P_J = c + d \quad (2)$$

Probabilité a posteriori des hypothèses.

Supposons maintenant connu le résultat x de tirage. On sait alors que l'on est dans l'un des cas relatifs à la colonne x, si bien que la probabilité a posteriori de l'état R est :

$$\frac{\text{proportion des cas favorables à R dans la colonne x}}{\text{proportion des cas de la colonne x}}$$

C'est cette quantité que nous avons notée par P .

Ainsi, si x = rouge, ceci correspond à une proportion a + c des cas parmi lesquels une proportion a est favorable à R ; on a donc :

$$P_{R/\text{rouge}} = \frac{a}{a + c} \quad (3)$$

On verrait de même :

$$P_{R/\text{jaune}} = \frac{b}{b + d} \quad (4)$$

$$P_{J/\text{rouge}} = \frac{c}{a + c} \quad (5)$$

$$P_{J/\text{jaune}} = \frac{d}{b + d} \quad (6)$$

Probabilité conditionnelle des observations.

Imaginons maintenant que l'on sache que l'état est R, la proportion des "cas possibles" n'est plus que a + b et celle des cas "favorable", dans ces conditions, à un tirage rouge est a. Il en résulte que la probabilité de tirer une boule rouge si l'état est R est :

$$\frac{\text{proportion des cas favorables à } x = \text{rouge si état R}}{\text{proportion des cas possibles si état R}} = \frac{a}{a + b}$$

On a donc avec les notations du texte :

$$f_{R/(\text{rouge})} = \frac{a}{a + b} \quad (7)$$

et de même :

$$f_{R(\text{jaune})} = \frac{b}{a + b} \quad (8)$$

$$f_{J(\text{rouge})} = \frac{c}{c + d} \quad (9)$$

$$f_{J(\text{jaune})} = \frac{d}{c + d} \quad (10)$$

Formule de Bayes pour une seule observation.

Il en résulte que si P_R (et $P_J = 1 - P_R$), $f_R(x)$ et $f_J(x)$ sont connus, les égalités (1) à (10) déterminent toutes les autres quantités évoquées.

Notamment, (1) et (7) donnent :

$$a = P_R f_R(\text{rouge})$$

et, de même :

$$b = P_R f_R(\text{jaune})$$

$$c = P_J f_J(\text{rouge})$$

$$d = P_J f_J(\text{jaune})$$

Alors, d'après (3) :

$$P_{R/\text{rouge}} = \frac{P_R f_R(\text{rouge})}{P_R f_R(\text{rouge}) + P_J f_J(\text{rouge})} \quad (11)$$

l'égalité (4) conduirait à une égalité analogue à (11), mais où $x = \text{rouge}$ serait remplacé par $x = \text{jaune}$. On peut donc écrire en laissant x indéterminé :

$$P_{R/x} = \frac{P_R f_R(x)}{P_R f_R(x) + P_J f_J(x)} \quad (12)$$

L'expression de P_J/x peut être obtenue de la même façon ou en remarquant simplement que :

$$P_{R/x} + P_{J/x} = 1$$

Cas de plusieurs observations.

Si, plus généralement on suppose que le tirage est répété n fois on peut calculer de même :

$$P_{R/x_1, x_2 \dots x_n}, (\text{probabilité de l'état } R, \text{ connaissant } x_1 \dots x_n)$$

en remplaçant le tableau à $2^2 = 4$ cases que nous venons d'envisager, par un tableau à 2^{n+1} cases. On trouverait alors :

$$P_{R/x_1, x_2 \dots x_n} = \frac{P_R f_R(x_1) f_R(x_2) \dots f_R(x_n)}{P_R f_R(x_1) f_R(x_2) \dots f_R(x_n) + P_J f_J(x_1) f_J(x_2) \dots f_J(x_n)} \quad (13)$$

Méthode de calcul progressif.

Il est important de remarquer que la méthode de calcul progressif que nous avons utilisée par ailleurs conduit exactement au même résultat. On a obtenu en effet :

$$r_n = r y_1 y_2 \dots y_n \quad r = \frac{P_J}{P_R} \quad y_i = \frac{f_J(x_i)}{f_R(x_i)}$$

formules où r_n est le rapport des probabilités *a posteriori* connaissant $x_1 x_2 \dots x_n$:

$$r_n = \frac{P_J / x_1 x_2 \dots x_n}{P_R / x_1 x_2 \dots x_n} = \frac{1 - P_R / x_1 x_2 \dots x_n}{P_R / x_1 x_2 \dots x_n}$$

d'où :

$$P_R / x_1 x_2 \dots x_n = \frac{1}{1 + r}$$

formule qui n'est autre que (13).

Cas de l'indépendance stochastique.

Remarque : si l'on avait $f_R(x) = f_J(x)$, c'est-à-dire si x était aussi probable dans l'hypothèse R que dans l'hypothèse J, on aurait identiquement:

$$P_{R/x} = P_R$$

et les observations n'apporteraient évidemment aucune information sur les hypothèses. On dirait alors qu'il y a indépendance stochastique, ou en probabilité, entre les observations et les hypothèses.