

# LE TEST DE LA MAJORITÉ CONFIRMÉE ET SON UTILISATION DANS L'INFERENCE STATISTIQUE \*)

par

EDOUARD FRANCKX

Bruxelles

## I. *Le problème posé*

J'ai longtemps hésité à donner cette conférence. J'ai craint de déborder largement le cadre des thèmes proposés à Rättvik. Cependant, je pense que la question que je voudrais exposer se rattache plus ou moins à la question que nous avons discutée aujourd'hui et qui consiste à choisir pour un risque déterminé la fonction de distribution adéquate.

Ma première réflexion à ce sujet pourra paraître surprenante: je ne suis pas du tout convaincu que nous devons toujours et à tout prix essayer de trouver une représentation analytique du risque étudié. Remarquons que l'on ait ou non une telle possibilité, elle ne change rien au fait que l'on a traité dans le passé et que l'on continue à pratiquer l'assurance des risques divers sans avoir une représentation mathématique plus au moins bien approchée de la fonction de répartition du risque couru <sup>1)</sup>).

Il y a, par ailleurs, de bonnes raisons de craindre que l'on arrive jamais à une solution unique; en effet ce qui intéresse la Compagnie cédante c'est la partie de la fonction de répartition qui correspond aux petits et moyens sinistres: par contre pour le réassureur c'est

---

\*) Conférence du Colloque 1961 à Rättvik.

<sup>1)</sup> Nous avons de par notre formation mathématique, une satisfaction d'ordre intellectuel quand nous avons pu adjoindre une telle formule à un problème pratique. Mais ceci n'est qu'une déformation qui résulte d'une attitude très subjective, rien ne dit que nous avons raison lorsque nous voulons, a priori, imposer, à nous mêmes et aux autres, une telle solution au problème réel. Par contre, cela ne veut pas dire que nous ne devons pas étudier le problème du classement de nos lois, aussi il est particulièrement intéressant comme l'a fait notre collègue M. Delaporte de situer les lois de distributions des risques divers sur le diagramme de Karl Pearson. Mais il résulte également de nos discussions d'aujourd'hui, qu'il est bien plus délicat de choisir sur ce diagramme „le point représentatif” c'est à dire la loi de distribution particulière qui „théoriquement convient actuellement”.

la partie relative aux grands sinistres qui importe. Comme dès lors, en pratique, c'est le réassureur qui imposera la distribution correspondant à ses propres risques et qui sera définie à l'échelle internationale. Celle-ci sera tout naturellement différente de celle admise par la Compagnie cédante pour une production plus locale. Il y a là, en fait, des exigences pratiques qui conduisent à des solutions multiples.

Par ailleurs, il n'est pas nécessaire pour conclure des affaires, de connaître avec une très grande précision la fonction de répartition du risque; celle-ci pourrait être obtenue, en pratique, par juxtaposition de différentes branches déduites de distributions de Pearson de types différents par exemple (cela n'a pas d'importance au sens du calcul des probabilités du moment que la fonction de distribution soit croissante de 0 à 1). D'autre part, pour passer à un domaine familier, ce qui nous donne la sécurité en assurance vie, c'est pas tellement le fait que la table de mortalité est régie par la loi de Makeham (la représentation analytique de la fonction de distribution n'y fait rien), c'est avant tout parce que la mortalité qui résulte de la fonction de distribution de Makeham est plus forte que la mortalité résultant de la fonction de distribution réelle qui elle reste inconnue. En d'autres termes, ce qui importe en pratique c'est de faire un choix judicieux de façon à ce que les résultats enregistrés soient en général favorables à l'organisme assureur. Dès lors se pose le problème beaucoup plus large et beaucoup plus important de comparer „les résultats obtenus, en utilisant deux distributions différentes et à la suite de cette comparaison de prendre la décision de choisir telle distribution plutôt que telle autre”. Encore une telle décision nous semblera rationnelle si elle est prise sur la base du critère que nous aurons admis souvent très subjectivement”.

## II. *La méthode de comparaison*

Permettez-moi de sortir du cadre du sujet traité à Rättvik et faire appel aux données plus générales du monde réel.

Le problème du choix d'une distribution appartient à la classe générale de la „comparaison des événements dont les résultats sont aléatoires”. Cela revient à introduire une méthode — un opérateur au sens abstrait — qui „induit un ordre” au vu des résultats aléatoires des épreuves.

Prenons, comme premier exemple, deux joueurs d'échec Pierre et Paul. Les résultats des parties successives sont, en général, aléatoires sauf si l'un des joueurs est réellement de force supérieure. Si Pierre gagne chaque partie contre Paul, nous n'hésitons pas à décider que „Pierre est meilleur que Paul”. Mais lorsque les joueurs sont de force à peu près égale (c'est le cas lors d'un championnat mondial) les résultats deviennent de plus en plus aléatoires et la décision est plus longue à tomber. Mais cet exemple nous fournit déjà tous les éléments fondamentaux qui sont nécessaires pour mettre le problème en route. Nous le schématisons ci-après

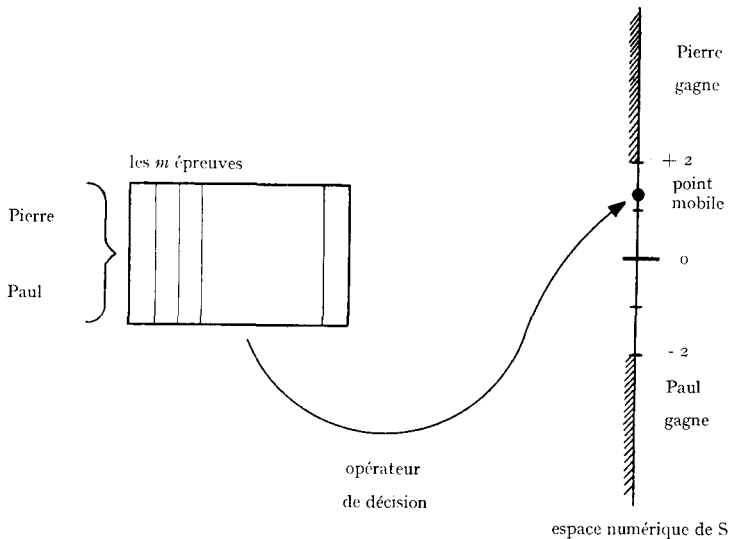


Fig. 1

Pierre mis en présence de Paul, joue une série d'épreuves.

La suite d'épreuves constitue un jury progressif, chaque épreuve nouvelle constitue un nouveau membre du jury, dont l'importance doit avoir le même „poids” qui celui des épreuves précédentes. Au sens de la théorie du jury, cela veut dire qu'au „stade atteint” les  $m$  épreuves, une permutation entre les membres du jury ne peut affecter la décision finale — donc la règle de décision doit être invariante par rapport aux permutations des épreuves. Toutes les règles de décision ne satisfont pas à un tel critère. Mais dans la vie courante, les jeux admettent de telles règles.



1) Nous obtenons ainsi:

Epreuves	$E_1$	$E_2$	$E_3$	$E_4$	$E_5$	$E_6$	$E_7$
Points acquis	+1	+1	-1	-1	-1	+1	+1
Total progressif	+1	+2	1	0	-1	0	+1

et la règle de décision admise ne permet pas de départager Pierre et Paul sur l'ensemble de ces sept épreuves considérées. Pierre n'est ni meilleur ni pire que Paul et nous conviendrons de les classer ex aequo. Cela revient à constater qu'ils sont équivalents au sens de la

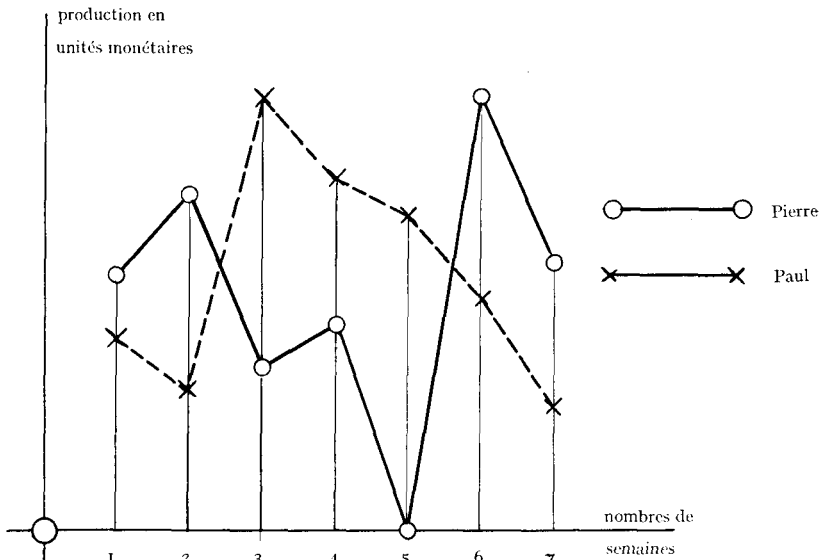


Fig. 2.  
 ——— trajet de Pierre  
 - - - - - trajet de Paul

règle de décision admise, ou bien, que cette règle n'est pas suffisamment forte pour les départager.

Considérons comme dernier exemple, le problème de la comparaison de deux lois de probabilités représentatives d'un même phénomène. Ces deux lois seront définies par leur fonction de distribution, la première est inconnue, elle est relative par exemple aux sinistres réellement subis, nous la désignerons par „Pierre”, la seconde est connue, c'est la loi analytique ou empirique qui est supposée représenter le phénomène, nous la désignons par „Paul”. Le pro-

blème consiste à classer Pierre et Paul. A cet effet, nous retournons au schéma de la fig. 2. En abscisse, nous portons la suite des sinistres déclarés, en ordonnée le coût du sinistre. Le trajet stochastique de Pierre est ainsi obtenu directement.

Pour obtenir celui de Paul, nous pouvons engendrer par la méthode de Monte Carlo une suite de montants de sinistres fictifs conformes à la loi désignée „Paul”. Dès lors, nous sommes revenus au cas de l'exemple précédent et le processus de décision suit.

Remarquons, d'ailleurs pour le cas de l'exemple 2, que si Pierre représente le producteur „moyen” nous aurions été amenés à engendrer de la même manière une séquence de Monte Carlo conforme au producteur moyen (séquence de Poisson par exemple), car il est difficile d'admettre que le producteur moyen est un individu dont les résultats hebdomadaires sont uniformes (cela revient à dire que le trajet stochastique de Pierre sera représenté par une horizontale). Nous concluons: quel que soit le problème, il nous sera toujours possible par le processus indiqué d'établir un jury d'épreuves. Ensuite ce jury doit faire choix d'une règle de décision. L'exemple 1 nous-en a indiqué une telle règle. Mais il est bien certain qu'on peut en imaginer bien d'autres qui répondent aux axiomes de symétrie énoncés.

### III. *Les règles de décision du jeu de tennis*

La pratique des jeux a conduit à admettre certaines règles de décision qu'il est intéressant de mettre en lumière.

De toute antiquité, l'homme quel que soit le jeu auquel il participe, n'a pu admettre qu'il soit battu sur le résultat d'une épreuve unique. Il ne considère le résultat acquis — c'est à dire que son adversaire est plus fort que lui si, lors d'un enchaînement de jeux successifs, le gain est plus souvent acquis par son adversaire.

L'enchaînement des jeux définit la règle „de décision” du jeu. A ce titre l'examen de la structure du jeu de tennis est plus particulièrement instructive.

L'examen révèle trois types différentes d'épreuves, auxquelles sont soumis Pierre et Paul à trois stades successifs.

- le jeu ordinaire
- le set, comportant une suite de jeux ordinaires
- la partie, comportant une suite de sets

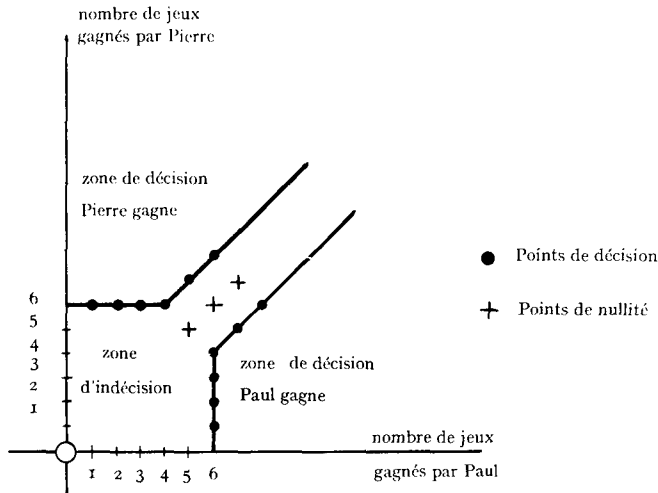


Fig. 3

Examinons la règle de décision d'un set. A cet effet, traçons le diagramme plan de la fig. 3. Supposons au départ une boule à l'origine 0

- elle se déplacera d'une unité à droite chaque fois que Paul gagne
- elle montera d'une unité chaque fois que Pierre gagne.

Dès lors on voit que Pierre gagne un set

- 1) s'il atteint l'horizontale 6 (il gagne par 6/0 6/1 6/2 6/3 6/4)
- 2) ensuite s'il atteint les scores de 7/5 8/6 etc . . . c'est-à-dire s'il gagne par deux points d'écart.

Remarquons qu'une telle règle de décision possède différentes propriétés

- 1) elle est séquentielle, au sens de Wald.
- 2) elle est parfaitement symétrique par rapport aux joueurs.
- 3) si à un stade quelconque du jeu, on intervertit l'ordre des épreuves, on ne modifie jamais le point final atteint par la boule.

Une telle règle possède les propriétés de double symétrie, admises comme axiomes, ce qui montre que l'on peut construire de multiples règles ayant ces mêmes propriétés de symétrie et satisfaisant aux axiomes.

En réalité aux trois stades on utilise trois règles différentes qui

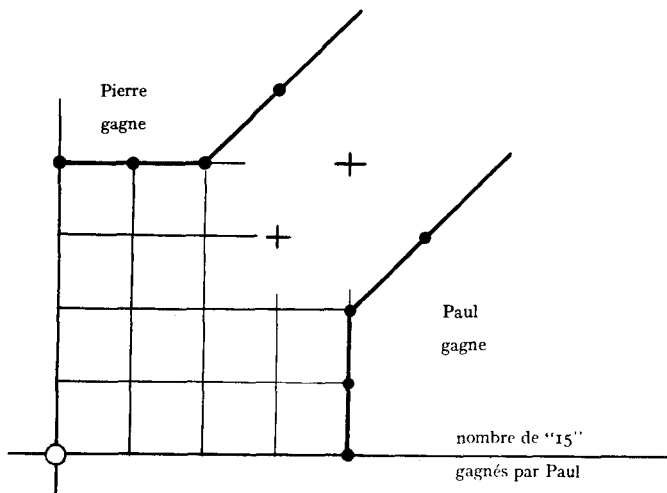


Fig. 4

sont toutes semblables, car elles peuvent être définies par trois diagrammes satisfaisant aux axiomes.

Mais les schémas des fig. 4 et fig. 5 sont différents puisque la règle de décision de la fig. 5 comporte un nombre maximum de sets

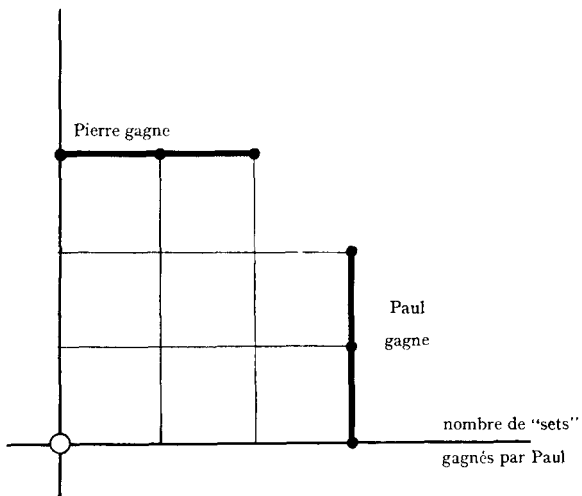


Fig. 5



qui est 5, alors que, théoriquement du moins, le nombre d'épreuves dans les autres cas peut être infini.

Ce que nous retenons également de cette étude, c'est le fait qu'il y a des épreuves à difficultés croissantes.

- 1) des épreuves d'ordre 1, qui sont les jeux ordinaires
- 2) des épreuves d'ordre 2, les sets
- 3) l'épreuve d'ordre 3, la partie.

Nous avons ainsi assemblé les matériaux nécessaires à notre test.

#### IV. *Le test de la majorité confirmée*

Nous savons, d'après la théorie de Wald que le test séquentiel le plus économique est obtenu par le schéma symétrique analogue aux tests des fig. 3 et 4, mais où les horizontales sont supprimées c'est à dire le test défini par la figure 6

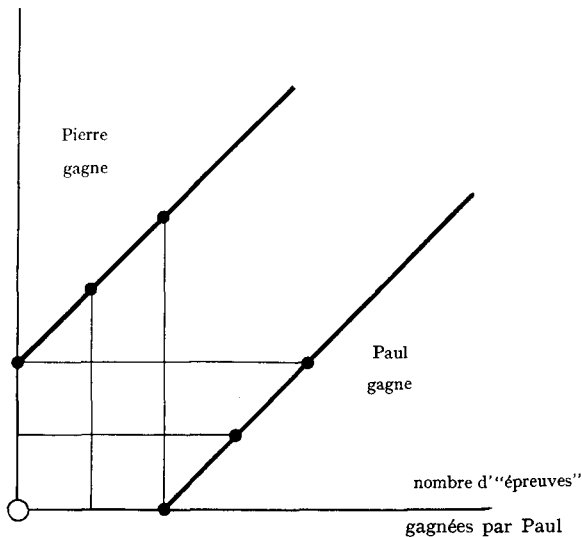


Fig. 6

c'est à dire

- 1) Pierre gagne s'il atteint les scores

$$2/0 \quad 3/1 \quad 4/2 \quad \dots \quad N + 2/N$$

- 2) Paul gagne s'il atteint les scores

$$0/2 \quad 1/3 \quad \dots \quad N/N + 2$$

3) le jeu continue pour tout score intermédiaire du type  $N + 1/N$ ,  $N/N$ ,  $N/N + 1$  (cas de nullité ou de non classement).

Or dans le cas  $N/N$  les joueurs sont „à égalité” des parties gagnées dans le cas  $N + 1/N$  ou  $N/N + 1$  l'un des joueurs atteint la majorité des épreuves gagnées.

Dans le cas  $N + 2/N$  ou  $N/N + 2$  l'un des joueurs *confirme qu'il a la majorité des nombres des épreuves gagnées*. C'est la raison pour laquelle nous appelons cette règle de décision: *le test de la majorité confirmée*.

Ce test possède les propriétés suivantes:

1<sup>0</sup>) les propriétés de symétrie par rapport aux deux joueurs et par rapport aux épreuves.

2<sup>0</sup>) en vertu de la manière dont il a été construit, il est impossible à tout autre test imaginable de comporter en moyenne un nombre plus petit d'épreuves tout en confirmant la majorité des épreuves gagnées. C'est par suite un test optimum au sens de Wald. Dès lors nous pouvons imaginer, comme au jeu de tennis une superposition d'épreuves d'ordre différent avec la convention:

1<sup>0</sup>) un joueur gagne une épreuve d'ordre  $N + 1$ , si au diagramme de la fig. 6, il gagne le test au nombre d'épreuves d'ordre  $N$

2<sup>0</sup>) les épreuves d'ordre 1 sont les épreuves ordinaires et la règle de décision du test d'ordre  $K$ .

Pierre est meilleur que Paul, s'il gagne une épreuve d'ordre  $K$ ;  $K$  étant un entier positif dont la valeur sera fixée conventionnellement.

#### V. Courbe opérationnelle d'un test d'ordre $K$ de la majorité confirmée et ses propriétés asymptotiques

Nous pouvons à ce stade abandonner la méthode de recherche opérationnelle qui nous a conduit au travers de la théorie du jury au test d'ordre  $K$ , pour aborder les problèmes plus connus de la théorie de la statistique mathématique. Désignons par  $p$ , la probabilité a priori pour que Pierre gagne une épreuve ordinaire, nous calculons la probabilité a priori  $P(p)$  pour qu'il gagne l'épreuve d'ordre  $K$  (avec  $P^1(p) = p$  pour convention).

*Théorème 1* Nous avons:

$$P_k(p) = \frac{1}{1 + \left(\frac{1-p}{p}\right)^{2^{k-1}}}$$

En effet la formule est exacte pour  $K = 1$

$$P^1(\phi) = \frac{1}{1 + \frac{1 - \phi}{\phi}} = \phi.$$

Supposons qu'elle soit valable au stade  $n$ . Si nous reprenons la figure 2, les probabilités d'atteindre les points de coordonnées

$$(2,0) \quad (1,1) \quad (0,2)$$

sont respectivement:

$$\phi_n = (P^n(\phi))^2 \quad q_n = 2P^n(\phi) [1 - P^n(\phi)] \quad r_n = [1 - P^n(\phi)]^2$$

avec

$$\phi_n + q_n + r_n = 1$$

Dès lors, puisqu'au premier point Pierre gagne; qu'au second point les „coups antérieurs sont nuls” (ce qui revient à dire qu'au point de vue du jeu (1, 1) est équivalent au point de départ (0, 0)), qu'au troisième point Pierre perd. Nous obtenons par décomposition

$$P^{n+1}(\phi) = \phi_n + q_n P^{n+1}(\phi)$$

d'où

$$P^{n+1}(\phi) = \frac{\phi_n}{\phi_n + r_n} = \frac{1}{1 + \frac{r_n}{\phi_n}}$$

De cette formule nous déduisons directement que

$$\frac{P^{n+1}(\phi)}{1 - P^{n+1}(\phi)} = \frac{\phi_n}{r_n} = \left[ \frac{P^n(\phi)}{1 - P^n(\phi)} \right]^2 = \left( \frac{\phi_{n-1}}{r_{n-1}} \right)^2$$

par suite

$$\frac{\phi_n}{\phi_n} = \left( \frac{\phi_1}{r_1} \right)^{2^{n-1}} = \left( \frac{1 - \phi}{\phi} \right)^{2^{n-1}}$$

Nous pouvons, suivant les méthodes classiques, définir la courbe opérationnelle du test d'ordre  $K$ .

Pour cela, fig. 7, nous considérons la fonction qui définit  $P_k(\phi)$  en fonction de la valeur de la probabilité a priori  $\phi$ . Le graphe de cette correspondance définit la courbe opérationnelle du test.

Nous avons obtenu de la Compagnie française des „Machines Bull” le calcul des courbes opérationnelles de ces tests. Elles sont annexés à cette étude sous forme de tableau. Nous tenons à exprimer notre gratitude à la maison Bull pour le concours qu'elle a bien voulu nous accorder dans cette étude.

*Propriétés des courbes opérationnelles*

A) Quel que soit leur ordre  $K$

$$\text{si} \quad p_1 + p_2 = 1 \quad P^K(p_1) + P^K(p_2) = 1$$

$$\text{en particulier} \quad P^K(1/2) = 1/2$$

$$\text{Car si} \quad 2^{K-1} = s$$

$$P^K(p_1) = \frac{p_1^s}{p_1^s + p_2^s} \quad P^K(p_2) = \frac{p_2^s}{p_1^s + p_2^s}$$

$$\text{donc} \quad P^K(p_1) + P^K(p_2) = 1$$

il en résulte que toutes les courbes sont symétriques par rapport au point central  $C$ , qui leur est commun.

B) Si nous comparons deux courbes opérationnelles successives on constate que

$$\left\{ \begin{array}{ll} 0 < p < 1/2 & P^{k+1}(p) < P^k(p) \\ 1/2 < p < 1 & P^{k+1}(p) > P^k(p) \end{array} \right.$$

La deuxième inégalité résulte de la première, par application de la propriété A.

$$\text{Si } p < 1/2 \quad \frac{1-p}{p} > 1 \quad \text{donc a fortiori} \quad \left( \frac{1-p}{p} \right)^{2^{k-1}} > 1$$

or

$$P^{k+1}(p) = \frac{1}{1 + \left( \frac{1-p}{p} \right)^{2^{k-1}} \cdot \left( \frac{1-p}{p} \right)^{2^{k-1}}} < \frac{1}{1 + \left( \frac{1-p}{p} \right)^{2^{k-1}}} = P^k(p).$$

*Test idéal — Comparaison du test idéal et du test d'ordre  $k$ ,  
Probabilité de l'erreur de classement*

A) Le test idéal n'est pas difficile à définir

En effet si la probabilité a priori pour Pierre gagne contre Paul  $p$  est inférieure à  $1/2$ , le test idéal doit classer Paul avant Pierre.

Si la probabilité a priori pour que Pierre gagne contre Paul est supérieure à  $1/2$ , le test idéal doit classer Pierre avant Paul.

Si cette probabilité vaut  $1/2$ , alors Pierre et Paul sont à égalité. Désignons par  $P^*(p)$  la courbe opérationnelle du test idéal.

Nous devons avoir

$$\begin{aligned}
 P^*(\phi) &= 0 && \text{pour } \phi < 1/2 \\
 P^*(\phi) &= 1/2 && \text{pour } \phi = 1/2 \\
 P^*(\phi) &= 1 && \text{pour } \phi > 1/2
 \end{aligned}$$

Elle est représentée à la figure 7.

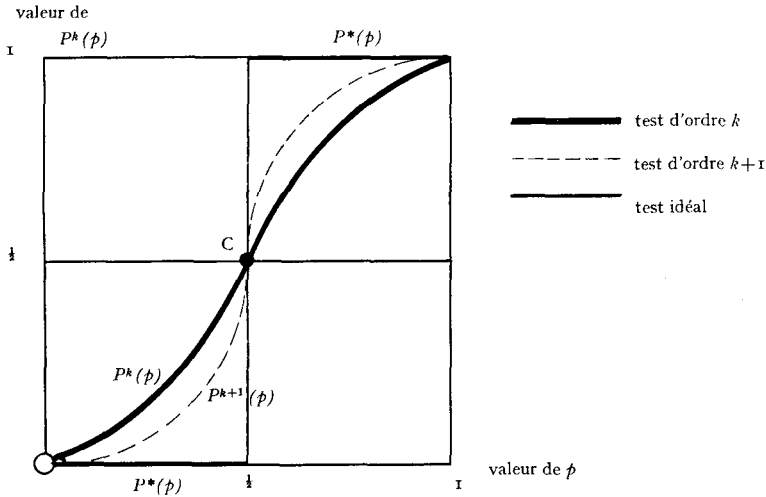


Fig. 7

B) Comparer le test idéal avec le test réel, revient à calculer les chances d'erreur en prenant une mauvaise décision après un test d'ordre  $k$ . Si  $\phi < 1/2$ , il y a une probabilité  $P^k(\phi)$  pour que Pierre gagne contre Paul, alors que nous admettons comme axiome dans le test idéal que Pierre doit perdre.

Donc la probabilité  $\phi_E(\phi)$  de l'erreur est telle que

$$\phi_E(\phi) = P^*(\phi) \text{ pour } \phi < 1/2$$

Si  $\phi > 1/2$ , il y a une probabilité  $1 - P^k(\phi)$  pour que Pierre perde contre Paul, alors que nous admettons que Pierre doit gagner.

La probabilité de commettre une erreur est par suite définie par la courbe symétrique de la figure 8.

Des propriétés des courbes opérationnelles de tests, il résulte que *l'on peut toujours faire un test tel que l'erreur de décision en dehors de*

$\varepsilon$ -voisinage de  $p = 1/2$  soit plus petit que toute limite de tolérance  $\eta$  arbitrairement petite.

C) En pratique cela revient à se fixer un seuil déterminé pour la limite de tolérance; ou ce qui revient au même à choisir conven-

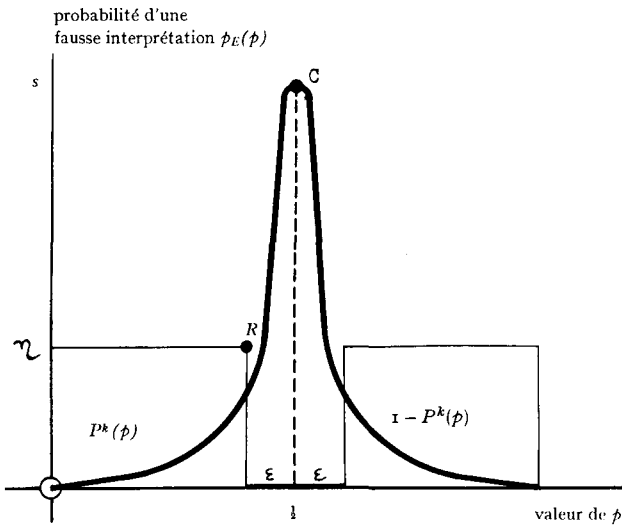


Fig. 8

tionnellement un point  $R$ , point de risque, en dessous duquel la courbe du test doit passer.

Par exemple, si  $\varepsilon = 0,05$  et  $k = 4$ , la probabilité de l'erreur en dehors du domaine  $0,45 \leq p \leq 0,55$  est limitée à 4 %

Si l'on veut diminuer la grandeur de l'intervalle central, il faut pousser le test à des ordres plus élevés.

En conclusion: *Les tests de la majorité confirmée, poussés suffisamment loin, permettent le classement de Pierre et Paul avec un risque d'erreur qui peut rendu arbitrairement petit, sauf dans le cas de la stricte égalité des joueurs. De toute façon le risque de fausse interprétation est mathématiquement évalué; c'est l'objet principal de l'inférence statistique.*

*Test de la majorité confirmée*

Epreuve ordinaire	Test d'ordre 2	Test d'ordre 3	Test d'ordre 4
0.01	0.00010		
0.02	0.00041		
0.03	0.00096		
0.04	0.00173		
0.05	0.00276	0.00001	
0.06	0.00406	0.00002	
0.07	0.00563	0.00003	
0.08	0.00750	0.00006	
0.09	0.00969	0.00010	
0.10	0.01220	0.00015	
0.11	0.01506	0.00023	
0.12	0.01826	0.00035	
0.13	0.02184	0.00050	
0.14	0.02582	0.00070	
0.15	0.03020	0.00100	
0.16	0.03501	0.00131	
0.17	0.04026	0.00176	
0.18	0.04597	0.00232	
0.19	0.05215	0.00302	0.00001
0.20	0.05882	0.00389	0.00002
0.21	0.06600	0.00496	0.00002
0.22	0.07369	0.00629	0.00004
0.23	0.08191	0.00790	0.00006
0.24	0.09068	0.00985	0.00010
0.25	0.10000	0.01220	0.00015

Epreuve ordinaire	Test d'ordre 2	Test d'ordre 3	Test d'ordre 4	Test d'ordre 5	Test d'ordre 6	Test d'ordre 7	Test d'ordre 8
0.26	0.10988	0.01501	0.00023				
0.27	0.12034	0.01837	0.00035				
0.28	0.13137	0.02236	0.00052				
0.29	0.14298	0.02708	0.00077				
0.30	0.15517	0.03263	0.00114				
0.31	0.16795	0.03914	0.00166				
0.32	0.18130	0.04675	0.00240	0.00001			
0.33	0.19523	0.05558	0.00345	0.00001			
0.34	0.20972	0.06579	0.00494	0.00002			
0.35	0.22477	0.07755	0.00702	0.00005			
0.36	0.24036	0.09100	0.00992	0.00010			
0.37	0.25646	0.10632	0.01396	0.00020			
0.38	0.27307	0.12366	0.01952	0.00040			
0.39	0.29016	0.14316	0.02720	0.00078			
0.40	0.30769	0.16495	0.03755	0.00152			
0.41	0.32565	0.18910	0.05158	0.00294	0.00001		
0.42	0.34399	0.21567	0.07029	0.00568	0.00003		
0.43	0.36269	0.24464	0.09493	0.01088	0.00012		
0.44	0.38170	0.27595	0.12683	0.02066	0.00044		
0.45	0.40099	0.30945	0.16723	0.03876	0.00162		
0.46	0.42051	0.34494	0.21708	0.07139	0.00588		
0.47	0.44022	0.38212	0.27665	0.12760	0.02095	0.00046	
0.48	0.46006	0.42063	0.34517	0.21743	0.07167	0.00592	0.00003
0.49	0.48001	0.46008	0.42067	0.34523	0.21752	0.07173	0.00594
0.50	0.50000	0.50000	0.50000	0.50000	0.50000	0.50000	0.50000