Principe des tests d'hypothèses

d'après Statistiques en seconde et … un peu après

Université de Bourgogne IREM, Mai 2000

<u>Exemple</u> : Je participais à un jeu de Pile ou Face. Je pariai pour le côté Pile. Mon adversaire Arnak obtint 5 fois Face en 6 coups. Doutant de l'honnêteté de mon adversaire, j'interrompis le jeu quelques instants. Je m'étais rendu compte alors, qu'il y avait 7 chances sur 64 pour qu'une pièces montre par hasard au moins 5 fois sur 6 sa face (p = 0,109). Cette probabilité était certes faible, mais pas assez pour accuser Arnak de tricheur. Je repris donc le jeu en estimant que mon seuil de doutes n'était pas atteint. En 18 coups, Arnak montra 14 faces. Or la probabilité d'obtenir, par hasard, au moins 14 faces en 18 coups est $C^{14}{}_{18}\left(\frac{1}{2}\right)^{18} + \dots + C^{18}{}_{18}\left(\frac{1}{2}\right)^{18} \approx 0,015$. Donc j'avais 985 chances sur 1000 de ne pas me tromper en refusant d'attribuer au hasard seul ma perte au jeu. Je décidai alors d'interrompre définitivement le jeu contre Arnak le prétendu tricheur.

Dans cet exemple élémentaire apparaissent les principales étapes d'un test statistique d'hypothèse.

<u>Formulation des hypothèses</u> : Problème Arnak est-il tricheur ou honnête ?

H_0 : Arnak est honnête, la probabilité pour que la pièce tombe sur face à chaque essai est $p = \frac{1}{2}$ et seul le hasard gouverne le jeu.

H_1 : Arnak est tricheur, il ne gagne pas au hasard, la probabilité de Face est plus grande que celle du Pile

<u>Choix du modèle statistique</u> : Choisir une statistique adaptée au problème posée. Dans cet exemple, on désigne par S_n la v.a. égale au nombre de faces obtenues pour des échantillons de n lancers. La loi de S_n est une binomiale $B(n,p)$ avec $p = \frac{1}{2}$: $P(S_n = k) = C_n^k\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{n-k} = C_n^k\left(\frac{1}{2}\right)^n = \frac{C_n^k}{2^n}$

Donc l'hypothèse H_0 est testable et H_1 ne l'est pas.

<u>Choisir un seuil de signification</u> :

Décisions / Etats	Accepter H_0	Accepter H_1
H_0 vraie	Bon : 1 - α	Erreur : α
H_1 vraie	Erreur : β	Bon : 1 - β

α = P(Accepter H_1/H_0 est vraie) est l'erreur de première espèce. Quel risque maximal dois-je accepter de courir s'il m'arrive d'accuser à tort mon adversaire ?

Suivant la nature du problème et la gravité qu'entraînerait une mauvaise décision, on prend $\alpha = 0,01$ (test très significatif) ou $0,05 = \alpha$ (test assez significatif).

En innocentant mon adversaire dans la première partie, je risque de me tromper et commettre l'erreur (Accepter H_0/H_1 est vraie).

L'erreur de deuxième espèce est donc $\beta = P$(Accepter H_0/H_1 est vraie)

La puissance du test est par définition, $\eta = 1 - \beta$

<u>Définir la région critique</u> :

C'est la plus grande région K_α de probabilité inférieure au seuil α et où H_0 ne peut être acceptée. Dans notre exemple, on rejette l'hypothèse H_0 si le nombre de Faces observé est trop grand au niveau α.

Dans le cas de la partie 1 :

Si $n = 6$ et S_n suit $B(6,\frac{1}{2})$, on refuse H_0 si S_n est plus grand qu'une valeur seuil k_α :

$K_\alpha = \{k \geqslant k_\alpha\}$ et $P(K_\alpha) \leqslant \alpha = 0{,}05$

$P(S_n \geqslant 6) = \dfrac{1}{2^6} = 0{,}015$ et $P(S_n \geqslant 5) = \dfrac{7}{2^6} = 0{,}109$ donc $K_\alpha = \{6\}$

(cf visualisation sur les histogrammes)

Dans le cas de la seconde partie $n = 18$, $S_n = B(18, \dfrac{1}{2})$

$P(S_n \geqslant 18) = 3{,}8.10^{-6}$, $P(S_n \geqslant 17) = \dfrac{19}{2^{18}} = 7{,}2.10^{-5}$, …etc $P(S_n \geqslant 13) = 0{,}048$,, $P(S_n \geqslant 12) = 0{,}1189$

Par conséquent, $k_\alpha = 13$ et $K_\alpha = \{13,14,15,16,17,18\}$

Dans la partie 1, $5 \notin K_\alpha$ donc j'accepte H_0

Dans la partie 2, $14 \in K_\alpha$ donc j'accepte H_1

Puissance du test :

La puissance du test est la probabilité $\eta = 1 - \beta = P(\text{Accepter } H_1 / H_1 \text{ vraie})$

Dans l'exemple étudié, $\eta(p) = \sum_{k \in K_\alpha} C_n^k p^k (1 - p)^{n - k}$ pour une valeur quelconque de p telle que $p > \dfrac{1}{2}$

Pour $n = 6$

$K_{0,05} = \{6\}$, $\eta(p) = p^6$ d'où $\eta(0{,}6) \approx 0{,}0467$, $\eta(0{,}8) \approx 0{,}2621$, $\eta(0{,}9) \approx 0{,}5314$

Dans ce cas, le test n'est pas puissant, la puissance $\eta(0{,}8) \approx 0{,}2621$signifie que : même si Arnak triche de façon à gagner 8 fois sur 10, on aura seulement une probabilité de 0,2621 de l'accuser de tricheur quand on le juge sur une partie de 6 coups.

Pour $n = 18$

$K_{0,05} = \{13,14,15,16,17,18\}$ et $\eta(p) = \sum_{k=13}^{18} C_{18}^k p^k (1 - p)^{18 - k}$

On trouve, par exemple, $\eta(0{,}8) \approx 0{,}8671$