

**TESTS STATISTIQUES : PRINCIPE
ET APPLICATIONS EN EPIDEMIOLOGIE**

**Olfa SAIDI^{1,2}, Dhafer MALLOUCHE^{2,3}, Nadia BEN
MANSOUR^{1,2}, Said HAJEM¹, Mohamed HSAIRI¹**

¹Institut National de Santé Publique

*²Laboratoire de Recherche en Epidémiologie et Prévention
des Maladies Cardiovasculaires*

³Ecole Supérieure de la Statistique et Analyse de l'Information

INTRODUCTION

En épidémiologie et en recherche clinique, l'analyse statistique des données consiste assez souvent à pratiquer des tests de comparaison. On cherche ainsi à comparer un paramètre à une valeur de référence ou plusieurs paramètres entre eux. Les paramètres peuvent être quantitatifs (moyenne, médiane...) ou qualitatifs (pourcentage).

A titre d'exemple, on peut citer les situations suivantes :

- La valeur moyenne de la pression artérielle systolique est-elle statistiquement différente selon que la personne soit obèse ou non ?
- La prévalence du tabac est-elle statistiquement différente selon le genre dans une population donnée ?
- Les pourcentages de guérison dans un essai thérapeutique évaluant des antihypertenseurs de 100 sujets hypertendus (un groupe de 50 sujets recevant le médicament A et un groupe de 50 sujets recevant le médicament B) sont-ils statistiquement différents ?

Dans ces situations, l'objet est de comparer des paramètres pour en sortir, le chercheur doit recourir aux tests statistiques.

Principe général des tests statistiques

Le principe général d'un test est de juger si la différence observée est due au hasard ou au contraire cette différence ne pourrait pas être due au hasard et elle est significative, mais indépendamment de la nature d'un test, son principe et son déroulement sont toujours les mêmes (figure n°1).

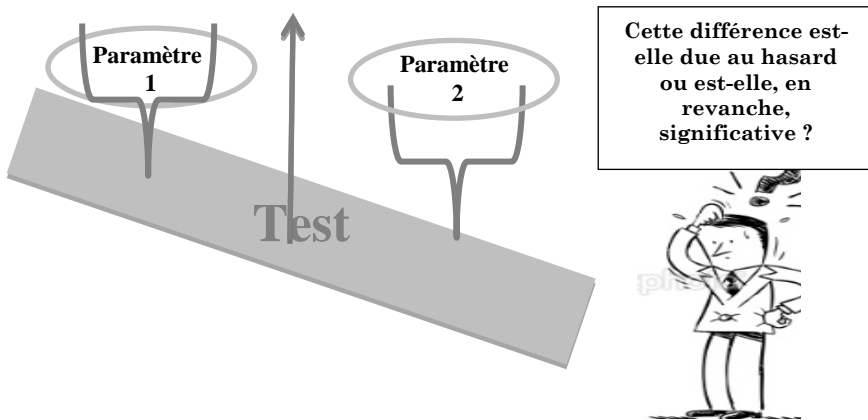


Figure 1 : Problématique du test statistique

Le principe des tests statistiques revient à formaliser un raisonnement s'appuyant sur les étapes suivantes :

- Formuler une hypothèse
- En déduire ce que devraient être les observations si cette hypothèse était vraie
- Vérifier si les observations sont conformes à ce que l'on attend sous l'hypothèse en question
- Conclure par un rejet ou un non rejet de l'hypothèse initiale

Formulation des hypothèses du test statistique

Un test statistique se base sur la formulation d'une hypothèse de départ dite hypothèse nulle (notée H_0), qui suppose l'absence de différence entre les deux paramètres à comparer, contre une autre hypothèse dite hypothèse alternative (notée H_1).

- L'hypothèse nulle notée H_0 est l'hypothèse que l'on désire contrôler : elle consiste à dire qu'il n'existe pas de différence statistiquement significative entre les paramètres comparés. Cette hypothèse est formulée dans le but d'être rejetée.
- L'hypothèse alternative notée H_1 est la négation de H_0 , elle permet de conclure que « H_0 est fausse». La décision de rejeter H_0 signifie que H_1 est réalisée ou H_1 est vraie.
- Un test statistique est utile quand il faut trancher entre ces deux hypothèses.

Si l'on reprend l'exemple de l'essai thérapeutique, pour déterminer s'il existe une différence d'efficacité entre les deux antihypertenseurs, on procède par un raisonnement par l'absurde qui consiste à supposer qu'il n'y a pas de différence entre P_A et P_B (ce qui suppose au préalable que H_0 est vraie). On calcule alors la probabilité d'observer les résultats obtenus sous cette hypothèse.

Si cette probabilité est inférieure à un seuil fixé au départ (en général 5% dit risque «de première espèce» ou alpha (α)-voir plus loin), on rejette l'hypothèse nulle et on accepte H_1 .

Test unilatéral ou bilatéral

La nature unilatérale ou bilatérale d'un test statistique dépend de la formulation de l'hypothèse alternative H_1 . Dans le cas de la comparaison de l'efficacité de deux traitements antihypertenseurs A et B :

- Le test est unilatéral si H_1 est formulée de la façon suivante : le traitement A est plus efficace que le traitement B dans le traitement de l'hypertension artérielle.
- Test bilatéral : On réalise un test bilatéral dans le cas où l'hypothèse alternative est formulée ainsi : l'efficacité du traitement A est différente de celle du traitement B dans le traitement de l'hypertension artérielle.

Initiation à la méthodologie statistique et épidémiologique

De façon globale, on utilise un test bilatéral quand on cherche à mettre en évidence une différence entre deux paramètres ou entre un paramètre et une valeur donnée sans se préoccuper du signe ou du sens de la différence (les 2 sens de la différence nous intéressent). En revanche, le test unilatéral est utilisé quand on cherche à savoir si un paramètre est supérieur (ou inférieur) à un autre ou à une valeur donnée (on ne s'intéresse qu'à la seule supériorité ou à la seule infériorité).

Choix du seuil de décision et risque α

Comme toute décision fondée sur les observations d'un échantillon et quel que soit le test utilisé, la conclusion d'un test statistique qui consiste à trancher entre les deux hypothèses nulle et alternative comporte un risque d'erreur.

Le risque α , encore appelé seuil de décision ou seuil de signification, est le risque de rejeter à tort l'hypothèse H_0 . Ce risque, fixé a priori avant de réaliser le test, fait partie de la règle de décision; il est aussi appelé risque de 1^{ère} espèce. Généralement le seuil de la signification statistique est fixé à 5%. Le choix d'une autre valeur plus restrictive est tout à fait possible.

Risque β et la puissance d'un test statistique

L'erreur de 2^{ème} espèce, notée β , est la probabilité d'accepter H_0 alors qu'elle est fautive. Il s'agit de la probabilité de ne pas conclure que deux groupes sont différents alors qu'ils le sont dans la réalité. Par exemple, c'est le risque de prouver qu'un médicament A n'est pas meilleur qu'un médicament B (ou un placebo) alors qu'il l'est en réalité. Il s'agit de la capacité d'éviter un faux négatif.

Quant à la puissance d'un test statistique, elle est donnée par la probabilité $1-\beta$. Ce concept mesure la capacité du test à conclure au rejet de H_0 quand celle-ci est effectivement fautive. Le test le plus puissant est celui qui fournit, pour une même valeur de α , l'erreur β la plus petite (ou encore la plus grande valeur de la puissance $1 - \beta$).

Degré de signification

Le degré de signification, noté «p», est la valeur qu'on compare au risque de 1^{ère} espèce α , fixé à l'avance souvent à 5%, et ce, pour trancher entre l'hypothèse nulle et l'hypothèse alternative.

Initiation à la méthodologie statistique et épidémiologique

Une valeur inférieure à 5% permet de conclure au rejet de H_0 en faveur de l'hypothèse alternative H_1 c'est-à-dire à l'existence d'une différence significative entre les paramètres à comparer. Dans ce cas de figure, le résultat du test est déclaré statistiquement significatif. Toutefois, une valeur supérieure à 5% ne permet pas de rejeter l'existence éventuelle d'une différence.

Application au calcul du nombre de sujets nécessaire

La puissance statistique consentie permet de calculer le nombre de sujets à inclure dans une étude. En général, on fixe la puissance désirée, le risque de première espèce et les paramètres associés aux groupes pour obtenir le nombre de sujets nécessaire dont le calcul doit impérativement figurer dans le protocole de l'étude.

CONCLUSION

Les tests statistiques sont d'usage très fréquent en épidémiologie et recherche clinique. Ils sont utiles en tant qu'outil d'évaluation des résultats des analyses comparatives. Ils permettent, en effet, de comparer une valeur observée à une valeur de référence ou bien deux valeurs observées. Ils concluent à l'existence ou non d'une différence significative entre les valeurs comparées. Ils ne permettent cependant pas d'établir un lien de causalité scientifique entre les paramètres étudiés. Tout lien statistique n'est donc pas synonyme de causalité.

Les tests statistiques s'articulent autour de la formulation de deux hypothèses (une hypothèse nulle notée H_0 et une hypothèse alternative notée H_1) dont une et une seule est vraie et permettent de trancher entre elles. L'hypothèse nulle est l'hypothèse qui stipule l'absence de différence significative entre les paramètres comparés. Quant à l'hypothèse alternative, elle correspond à l'existence d'une telle différence.

H_1 est vraie lorsque H_0 est rejetée. Le rejet de H_0 (quand elle est vraie) se fait avec un risque dit de première espèce et noté α . Celui-ci est fixé habituellement à 5%. La différence est dite statistiquement significative lorsque le degré de signification est inférieur à ce seuil. En revanche, le risque d'accepter H_0 alors qu'elle est fautive s'appelle le risque de deuxième espèce et est noté β . Le complément à 1 de β désigne la puissance du test.

BIBLIOGRAPHIE

1. *Bouyer J. Méthodes statistiques. Médecine- Biologie. Editions INSERM. Paris, 1996.*
2. *Ancelle T. Statistique Epidémiologie. 3^{ème} édition. Paris 2011.*
3. *Polycopié de biostatistiques. Faculté de Médecine de Montpellier-Nîmes. http://www.med.univ-montp1.fr/enseignement/cycle_2/Autres-Mod-Oblig/MB6/commun/polycop_biostat_tome_1_methodes_statistiques.pdf*
4. *R. MICHEL, L. OLLIVIER-GAY, A. SPIEGEL, J-P. BOUTIN. Les tests statistiques : intérêt, principe et interprétations. Médecine Tropicale. 2002.*
5. *Kleinbaum DG, Kupper LL, Morgenstern H. Epidemic Research: Principles and quantitative methods. New York 1982.*
6. *Rothman KJ. Modern Epidemiology. USA 1986.*
7. *Garel Bernard. Modélisation probabiliste et statistique. Editions CEPADUES.2002.*
8. *Engineering statistics handbook. Quantitative techniques. <http://www.itl.nist.gov/div898/handbook/eda/section3/eda35.htm>. Accessed September 2014.*
9. *Olivier Gaudoin. Principes et Méthodes Statistiques. INP Grenoble. 2014. <http://www-ljk.imag.fr/membres/Olivier.Gaudoin/PMS.pdf>.*
10. *Hela Ouaili-Mallek. Tests statistiques. ESSAIT Tunis 2013. https://www.academia.edu/6847/Documents_du_Cours_Tests_Parametriques;*