

Baptiste Janvier ¹

A QUOI CORRESPOND LA SIGNIFICATIVITE STATISTIQUE

Résumé :

L'objectif est de comprendre ce que représente la significativité statistique d'un résultat ainsi que l'influence de la taille de l'échantillon sur celui-ci.

Qu'est que le p ?

La significativité statistique d'un résultat est une mesure estimée du degré pour lequel il est "vrai" (au sens de "représentatif de la population"). La valeur du niveau représente un indice décroissant de la fiabilité d'un résultat. Plus le niveau p est élevé, moins on peut croire que la relation observée entre les variables de l'échantillon est un bon indicateur de la relation entre les variables correspondantes dans la population. Le niveau p représente la probabilité d'erreur qui est impliquée dans l'acceptation du résultat observé comme valide (c'est-à-dire comme "représentatif de la population").

Par exemple, un niveau p de 0,05 indique qu'on a une probabilité de 5% que la relation entre les variables trouvées dans notre échantillon soit une "coïncidence" ou un "coup de chance". En d'autres termes, sous l'hypothèse que dans la population il n'y a pas de relation entre ces variables, si on répétait nos expériences les unes après les autres, on pourrait s'attendre à ce qu'approximativement une fois sur 20, on obtienne une relation entre les variables en question égale ou bien plus forte que dans les autres expériences. Dans beaucoup de domaines de la recherche, un niveau p de 0,05 est considéré selon l'usage comme une "limite" acceptable pour le niveau d'erreur.

¹ Baptiste Janvier, Etudiant thèse Université François Rabelais, département de psychologie, 3 rue des Tanneurs, 37041 Tours Cedex 01. b.janvier@laposte.net

** D'autres documents sur les analyses statistiques sont téléchargeables à l'adresse : <http://www.janvier.ca.tc> (puis page professionnelle et ensuite analyses statistiques).

Typiquement, dans beaucoup de domaines de la Science, les résultats avec $p < 0,05$ sont considérés comme étant statistiquement significatifs, mais souvenez-vous que ce niveau de significativité implique encore une probabilité d'erreur non négligeable de 5%. Les résultats qui sont significatifs au niveau $p < 0,01$ sont généralement considérés comme étant statistiquement significatifs et des niveaux $p < 0,005$ ou $p < 0,001$ sont souvent appelés "hautement" significatifs. Mais souvenez-vous que ces classifications ne représentent rien d'autre que des conventions arbitraires qui sont seulement basées sur une expérience générale de recherche.

Significativité et taille de l'échantillon ?

" Le résultat d'une épreuve de signification est également relatif au nombre d'observations sur lequel elle porte. En règle générale, pour une différence donnée et une dispersion donnée, un seuil sera d'autant plus facilement atteint que ce nombre sera grand. (...) Une différence donnée entre deux fréquences observées sur deux échantillons de taille N_1 et N_2 , sera significative d'autant plus facilement que la taille des échantillons sera grande, l'erreur type variant dans ce cas comme :

$$\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}$$

(Reuchlin, 1991, p.214).

Reuchlin M. (1991), *Précis de statistique*, col. Le Psychologue, P.U.F.

S'il y a très peu d'observations, alors il y a aussi peu de possibilités de combinaisons entre les valeurs des variables, et ainsi la probabilité d'obtenir par chance une combinaison de ces valeurs indiquant une forte relation est relativement élevée. La probabilité d'un écart aléatoire d'une taille donnée (écart par rapport à la moyenne de la population), décroît lorsque la taille de l'échantillon s'accroît.

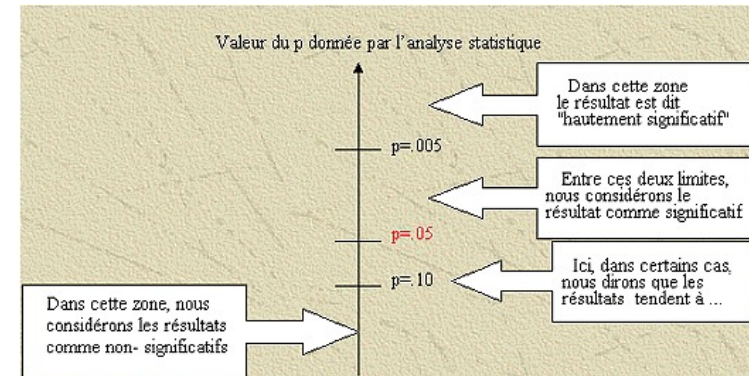
Supposons qu'on ait déjà calculé une mesure de la relation entre deux variables. La question suivante est "cette relation est-elle significative" ? Par exemple, est-ce que 40% de la variance expliquée entre les deux variables sont assez pour considérer la relation comme significative ? La réponse est : "cela dépend".

Explicitement, la significativité dépend essentiellement de la taille d'échantillon. Comme expliqué précédemment, dans des échantillons très grands, même des très faibles relations entre variables seront significatives, alors que dans des très petits échantillons, même des très fortes relations ne peuvent pas être considérées comme sérieuses (significatives). Ainsi, afin de déterminer le niveau de la significativité statistique, on a besoin d'une fonction qui représente la relation entre la "grandeur" et la "significativité" des relations entre deux variables, ceci dépendant de la taille de l'échantillon. La fonction dont on a besoin devrait nous dire exactement "quelle est la probabilité d'obtenir une relation d'une grandeur donnée à partir d'un échantillon d'une taille donnée, en supposant qu'il n'y ait pas de telle relation entre ces variables dans la population". En d'autres termes, cette fonction devrait nous donner le niveau de significativité (p) et devrait nous dire la probabilité d'erreur impliquée dans le rejet de l'idée que la relation en question n'existe pas dans la population. Cette hypothèse "alternative" (qu'il n'y a pas de relation dans la population) est habituellement appelée l'hypothèse nulle. Cela pourrait être idéal si la fonction de probabilité était linéaire, et par exemple, avait seulement différentes pentes pour des échantillons de tailles différentes. Malheureusement, la fonction est plus complexe, et n'est pas toujours exactement la même ; cependant, dans la plupart des cas on connaît sa forme et on peut l'utiliser pour déterminer les niveaux de significativité pour nos conclusions dans les échantillons d'une taille particulière. La plupart de ces fonctions sont liées à un type général de fonction qui est appelée normal.

Quel seuil choisir dans la description des résultats

A partir de la probabilité réelle, suite d'une analyse statistique, comment arriver simplement à considérer un résultat comme significatif ou non ?

En considérant que le logiciel statistique donne la valeur exacte de la probabilité associée au résultat, nous allons nous aider de ce schéma:



Dans un premier temps, il faut replacer la valeur de la probabilité donnée par l'analyse de l'analyse de variance sur ce schéma, et, ensuite, lire la description correspondante.

Imaginons que le résultat de l'analyse, effectuée sur deux groupes de 20 sujets soit :

dl Effet	MC Effet	dl Erreur	MC Erreur	F	niveau p
1	70,22500	38	3,803947	18,46109	.000116

Dans ce cas, la valeur .000116 se place alors dans la partie supérieure du schéma, ce qui revient à dire que ces deux groupes se différencient. Ces différences de niveaux sont hautement significatives [$F(1-38)=18,461$; $p<.005$]

L'explication de ce résultat est : si le résultat est "fiable" avec 0,116 % de chance de se tromper, alors nous pouvons affirmer avec une marge d'erreur de 0,5% (limite conventionnelle supérieure) que la relation (ici la différence) entre les variables ne soit pas une "coïncidence" ou un "coup de chance", mais bel et bien un effet statistique.

Si le résultat du niveau du p avait été .0008, nous aurions marqué dans le crochet; p<.001]

De même, si le résultat du niveau du p aurait été .06, nous l'aurions placé entre .05 et .10, ce qui, selon le contexte, peut nous amener à dire que ces deux groupes de sujets tendent à obtenir des résultats différents (et dans le crochet nous aurions indiqué F(... ; p<.10]. (il va de soit qu'il faut remplacer les ... par les valeurs des dl et du F).