

# Fiche 1

## Pourquoi des tests statistiques

---

### */ Principe*

On réalise des tests statistiques dans de nombreuses disciplines scientifiques parce que l'on utilise des échantillons de population. Un échantillon est une partie de la population que l'on utilise parce qu'il est trop coûteux en temps, en énergie et en argent d'interroger tout le monde. Pour les sondages électoraux, on interroge souvent entre neuf cents et mille personnes ce qui prend déjà du temps. Imaginez alors le temps qui serait nécessaire pour interroger les quarante-huit millions d'électeurs. C'est pour cette raison que l'on interroge un nombre restreint de personnes. Cela revient moins cher, prend moins de temps et, si l'échantillon est prélevé avec sérieux, et bien il y a de grandes chances que l'on obtienne le même niveau d'information que celui que l'on obtiendrait en interrogeant tout le monde.

Toutefois, en matière de constitution d'échantillon, même avec la meilleure volonté du monde, il y aura toujours des petites différences dans des échantillons par rapport aux caractéristiques de la population et notamment si ceux-ci sont de petites tailles. Or il faut pouvoir estimer le poids de ces petites variations pour apprécier des différences qui, elles, sont expliquées par l'intervention d'un facteur. En effet,

si vous avez besoin de constituer deux échantillons et que vous appliquez une variable à l'un et une autre variable à l'autre. Vous comparez les deux groupes et vous trouvez une différence : la question est maintenant de savoir si cette différence est imputable à la différence d'action des variables que vous avez manipulées ou si cela provient de différences « naturelles » de vos échantillons.

Les tests statistiques sont là pour nous aider à évaluer le poids respectif de ces différences : avec les tests, on fonctionne comme s'il y avait une grande urne dans laquelle se trouvaient des boules dont les chiffres représentent des données quelconques. On imagine alors que la main droite extrait des boules qui représentent l'échantillon 1 tandis que la main gauche extrait des boules qui représentent les valeurs de l'échantillon 2. Si l'on compare les valeurs, il y aura toujours des différences mais les fréquences de ces différences se répartissent de manière non homogène. Ainsi, la probabilité d'obtenir une différence nulle (les moyennes sont identiques) est le cas le plus fréquent tandis qu'obtenir une différence non nulle est, théoriquement, moins fréquent et ce, d'autant plus que l'amplitude de la différence est grande.

*Exemple :* on fait passer à une population de huit sujets, une tâche d'apprentissage d'une liste de vingt mots concrets. Les scores obtenus sont les suivants :

Sujet	1	2	3	4	5	6	7	8
Score de rappel	6	8	7	10	9	11	12	5

Imaginons que nous prélevions, au hasard, deux sujets (S1 et S2) dans cette population et que nous fassions la différence entre leur score respectif. Celle-ci devrait obligatoirement correspondre à l'une des différences contenues dans l'un des deux tableaux ci-après :

S1 \ S2	5	6	7	8	9	10	11	12
5								
6	1							
7	2	1						
8	3	2	1					
9	4	3	2	1				
10	5	4	3	2	1			
11	6	5	4	3	2	1		
12	7	6	5	4	3	2	1	

S2 \ S1	5	6	7	8	9	10	11	12
5								
6	-1							
7	-2	-1						
8	-3	-2	-1					
9	-4	-3	-2	-1				
10	-5	-4	-3	-2	-1			
11	-6	-5	-4	-3	-2	-1		
12	-7	-6	-5	-4	-3	-2	-1	

La carte des effectifs des différences possibles nous donne alors, en ne tenant compte que des valeurs absolues des différences, les résultats suivants :

Amplitude de la différence	1	2	3	4	5	6	7	Total
Fréquences absolues (effectifs)	14	12	10	8	6	4	2	56
Fréquences relatives (en %)	25,00	21,43	17,86	14,29	10,71	7,14	3,57	100

Comme on le voit, plus l'amplitude de la différence est importante, plus la probabilité qu'elle a d'être obtenue est faible. C'est justement cette probabilité que les tests d'hypothèses tentent d'évaluer en posant la question suivante : quelle est la probabilité que j'aurais d'obtenir une telle différence par le hasard de l'échantillonnage ?

## // Le seuil de 5 % d'erreur en statistique

Les tests de comparaison de moyennes ou de fréquences vont répondre, avec précision, à cette question en évaluant la probabilité d'obtenir, de manière aléatoire, toute différence observée. Si cette probabilité est supérieure à un seuil fixé, par convention, à 5 %, alors on estimera que les différences sont imputables au hasard. On conclura que la différence est statistiquement significative et, par voie de conséquence, que la différence est imputable aux effets de la ou des variables indépendantes que l'on a pu manipuler ou qui ont été définies.

Le graphique ci-dessous résume bien comment on doit apprécier cette différence : imaginons que nous comparons deux groupes d'élèves dont l'un a fait l'objet d'une pédagogie A pour apprendre de l'arithmétique et l'autre à fait l'objet d'une pédagogie B. On leur fait passer un test et on constate des différences de performance dans les deux groupes. Pour apprécier l'effet du style pédagogique, on doit essayer d'évaluer la probabilité d'obtenir un tel effet par le hasard puis, en fonction de cette probabilité, on va décider si oui ou non on peut estimer que les différences entre les deux groupes proviennent bien du type de pédagogie.

