

Introduction

On utilise le test t pour échantillons pairés (ou échantillons appariés) quand les deux moyennes que l'on compare proviennent du même échantillon. Il s'agit du même groupe de participants que nous avons testés deux fois, il s'agit donc d'un plan à mesures répétées. Un bon exemple de situation dans laquelle on utilise ce test est un essai clinique. C'est-à-dire une étude dans laquelle on teste un traitement pour un trouble donné. Dans ce genre de situation, les chercheurs évaluent les symptômes avant le traitement, administrent le traitement aux participants, puis évaluent leurs symptômes une seconde fois. Finalement, ils regardent si les symptômes ont significativement diminué.

Pour calculer la statistique t avec deux échantillons appariés, vous aurez besoin de l'écart-type de vos scores de différences. En d'autres mots, vous aurez besoin des deux scores de chaque participant, vous calculerez la différence entre T2 et T1 (c'est le score de différence) et vous calculez l'écart-type de cette distribution.

Formule

$$t = \frac{\bar{D} - 0}{s_D / \sqrt{N}}$$

Moyenne des scores de différence

Statistique t

Écart-type des scores de différence

Moyenne des score de différence à laquelle on s'attend selon l'hypothèse nulle

Nombre de scores de différences

Mise en situation

Les chercheurs Kirk et McCoy se demandent si leur nouveau traitement pour la dépression qui consiste à envoyer les gens dans l'espace est optimal. Pour ce faire, ils recrutent 5 participants dépressifs, évaluent leur niveau de dépression, les envoient dans l'espace, puis évaluent leur niveau de dépression à leur retour.

H: La moyenne de nos différences est différente de 0

H₀: La moyenne de nos différences est de 0

Voici leurs résultats:

	Avant	Après	Différences (T2-T1)
1	17	5	-12
2	14	12	-2
3	34	11	-23
4	21	18	-3
5	26	13	-13
Moyenne	22.4	11.8	-10.6

Calculs

Étape 1 : calculer l'écart-type

*Si vous n'êtes pas familier avec l'écart-type, je vous conseille la capsule sur les mesures de tendance centrale ou de vous référer à votre manuel si vous en avez un.

$$s = \sqrt{\frac{\sum(X - \underline{X})^2}{N - 1}}$$

s

$$= \sqrt{\frac{(-12 - -10.6)^2 + (-2 - -10.6)^2 + (-23 - -10.6)^2 + (-3 - -10.6)^2 + (-13 - -10.6)^2}{5 - 1}}$$

$$= \sqrt{\frac{293.2}{4}} = \sqrt{73.3} = 8.56$$

Étape 2: calculer la statistique t

$$t = \frac{\underline{D} - 0}{\frac{S_D}{\sqrt{N}}}$$

$$t = \frac{-10.6 - 0}{\frac{8.56}{\sqrt{5}}} = \frac{-10.6}{\frac{8.56}{2.24}} = \frac{-10.6}{3.66} = 2.90$$

*La raison pour laquelle on soustrait 0 à la moyenne des scores de différences est que la moyenne à laquelle on s'attend est de 0. On le garde quand même dans la formule car dans

certain cas, il est possible que les chercheurs s'attendent à une autre moyenne et c'est cette moyenne à laquelle ils s'attendent qu'ils vont soustraire à la moyenne des scores de différences et non 0.

** $-10.6 \div 3.66 = -2.90$. Toutefois, les valeurs critiques dans la table t sont positives, il est de coutume de mettre sa statistique t au positif si celle-ci est négative pour la comparaison.

Étape 3: Le degré de liberté

Dans un test t pour échantillons appariés, vous calculez votre degré de liberté en faisant votre nombre de paires - 1

dl = Nombre de paires - 1

dl = 5 - 1

dl = 4

Étape 4: table t

Il s'agit de la même table que pour le test t pour échantillons indépendants. **Cependant, il s'agit d'une étude clinique, nous utilisons donc un seuil de rejet de 0.01** (rappelez-vous Kirk et McCoy envoient les gens dans l'espace pour les guérir de leur dépression). Votre valeur critique est donc de 4.604.

Si votre statistique t est **supérieure** à votre valeur critique, vous pouvez rejeter l'hypothèse nulle.

$2.90 < 4.604$

Dans notre cas, la statistique t n'est pas supérieure à notre valeur critique, on ne peut pas rejeter l'hypothèse nulle et donc conclure que le traitement de Kirk et McCoy est efficace.

TABLE C.2 Critical Values of the t Distribution for $\alpha = .05$ and $.01$, Two-Tailed Test

Degrees of Freedom	.05	.01
1	12.706	63.657
2	4.303	9.925
3	3.182	5.841
4	2.776	4.604
5	2.571	4.032
6	2.447	3.707
7	2.365	3.499
8	2.306	3.355
9	2.262	3.250
10	2.228	3.169
11	2.201	3.106
12	2.179	3.055
13	2.160	3.012
14	2.145	2.977
15	2.131	2.947
16	2.120	2.921
17	2.110	2.898
18	2.101	2.878
19	2.093	2.861
20	2.086	2.845
21	2.080	2.831
22	2.074	2.819
23	2.069	2.807
24	2.064	2.797
25	2.060	2.787
26	2.056	2.779
27	2.052	2.771
28	2.048	2.763
29	2.045	2.756
30	2.042	2.750
40	2.021	2.704
60	2.000	2.660
120	1.980	2.617
∞	1.960	2.576

Source: Adapted from Table 9 in *Biometrika Tables for Statisticians*, vol. 1, 3d ed., edited by E. S. Pearson and H. O. Hartley (New York: Cambridge University Press, 1958).

Signification des symboles

X : Valeur

N : Nombre d'observations total

n : Nombre d'observations d'un groupe

- n_1 : nombre d'observations du groupe 1 ;
- n_2 : nombre d'observations du groupe 2 ;
- etc.

K : Nombre de groupes

Σ : Somme

\bar{X} : Moyenne d'un échantillon

μ : Moyenne d'une population

s : Écart-type d'un échantillon

σ : Écart-type d'une population

s^2 : Variance d'un échantillon

σ^2 : Variance d'une population

χ^2 : Khi-carré

A : Fréquence attendue

O : Fréquence observé

L : Ligne

C : Colonne

t : Statistique t (ou score t dans le cas d'une corrélation/régression linéaire)

r : Coefficient de Pearson

Z : Score Z

b : Coefficient de régression

a : Ordonnée à l'origine

\hat{Y} : Valeur de la VD qu'on veut prédire l'aide la VI

F : Statistique F

Ouvrages de référence

- Field, A. (2017). *Discovering Statistics Using IBM SPSS Statistics: North American Edition* (5th ed.). Sage Edge
- Howell, D.C. (2008). *Méthodes statistiques en sciences humaines* (M. Rogier, V. Yzerbyt, & Y. Bestgen, Trans.). (6th ed.). De boeck. (Original work published 2008)
- Tabachnick, B. G., & Fidell, L. S. (2021). *Using Multivariate Statistics* (7th ed.). Pearson