

---

# NORME INTERNATIONALE



# 3301

---

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION • МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ • ORGANISATION INTERNATIONALE DE NORMALISATION

---

## Interprétation statistique des données — Comparaison de deux moyennes dans le cas d'observations appariées

*Statistical interpretation of data — Comparison of two means in the case of paired observations*

Première édition — 1975-08-15

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

[ISO 3301:1975](https://standards.iteh.ai/catalog/standards/sist/4deaf99b-576c-4cfb-a97d-0851b0148c50/iso-3301-1975)

<https://standards.iteh.ai/catalog/standards/sist/4deaf99b-576c-4cfb-a97d-0851b0148c50/iso-3301-1975>

---

CDU 519.28

Réf. n° : ISO 3301-1975 (F)

Descripteurs : analyse statistique, moyenne arithmétique, comparaison.

Prix basé sur 6 pages

## AVANT-PROPOS

L'ISO (Organisation Internationale de Normalisation) est une fédération mondiale d'organismes nationaux de normalisation (Comités Membres ISO). L'élaboration de Normes Internationales est confiée aux Comités Techniques ISO. Chaque Comité Membre intéressé par une étude a le droit de faire partie du Comité Technique correspondant. Les organisations internationales, gouvernementales et non gouvernementales, en liaison avec l'ISO, participent également aux travaux.

Les Projets de Normes Internationales adoptés par les Comités Techniques sont soumis aux Comités Membres pour approbation, avant leur acceptation comme Normes Internationales par le Conseil de l'ISO.

La Norme Internationale ISO 3301 a été établie par le Comité Technique ISO/TC 69, *Application des méthodes statistiques*, et soumise aux Comités Membres en mars 1974.

Elle a été approuvée par les Comités Membres des pays suivants :

Afrique du Sud, Rép. d'	Hongrie	Roumanie
Allemagne	Inde	Royaume-Uni
Autriche	Israël	Suisse
Belgique	Italie	Tchécoslovaquie
Brésil	Pays-Bas	Turquie
Espagne	Pologne	U.R.S.S.
France	Portugal	Yougoslavie

Les Comités Membres des pays suivants ont désapprouvé le document pour des raisons techniques :

Suède  
U.S.A.

# Interprétation statistique des données — Comparaison de deux moyennes dans le cas d'observations appariées

## 0 INTRODUCTION

La méthode spécifiée dans la présente Norme Internationale, connue sous le nom de méthode des observations appariées, est un cas particulier de la méthode reproduite au tableau A' de l'ISO 2854, *Interprétation statistique des données — Techniques d'estimation et tests portant sur des moyennes et des variances*.<sup>1)</sup>

Ce cas particulier est mentionné dans la section deux de l'ISO 2854, immédiatement après l'exemple numérique du tableau A', et un exemple complet de la méthode de comparaison par paires a été donné dans l'annexe A de cette Norme Internationale. L'importance de la méthode et l'extension de son domaine d'application ont justifié qu'un document séparé lui soit consacré.

## 1 OBJET

La présente Norme Internationale spécifie une méthode de comparaison de la moyenne d'une population de différences entre des observations appariées à zéro ou à toute autre valeur donnée d'avance.

## 2 DÉFINITION

**observations appariées :** Deux observations  $x_i$  et  $y_i$  d'une certaine propriété ou caractéristique sont dites appariées si elles sont faites :

- soit sur le même élément  $i$  d'une population d'éléments, mais dans des conditions différentes (par exemple, comparaison des résultats de deux méthodes d'analyse pour un même produit),
- soit sur deux éléments distincts considérés comme aussi identiques que possible, sauf en ce qui concerne l'effet de causes différentes dont on veut tester l'action (par exemple, comparaison des rendements de deux parcelles voisines ensemencées avec deux variétés différentes).

On notera cependant que, dans ce second type d'essai, l'efficacité de la méthode dépend de la validité de l'hypothèse selon laquelle les deux individus d'une même

paire ne présentent pas de différence systématique autre que celle qui peut éventuellement résulter de l'effet des causes que l'on veut comparer.

## 3 DOMAINE D'APPLICATION

La méthode peut être utilisée pour faire ressortir une différence entre deux traitements. Dans ce cas, les observations  $x_i$  sont effectuées après le premier traitement et les observations  $y_i$  après le second. Les deux séries de résultats des observations ne sont pas indépendantes car tout résultat  $x_i$  de la première série (premier traitement) se trouve associé à un résultat  $y_i$  de la deuxième série (deuxième traitement). L'expression « traitement » doit être entendue dans le sens large. Les traitements comparés peuvent être, par exemple, deux méthodes d'essai, deux instruments de mesure ou deux laboratoires, afin de détecter une erreur systématique possible. Deux traitements effectués successivement sur le même matériau peuvent interagir, et la différence obtenue peut dépendre de l'ordre dans lequel ces traitements sont exécutés. Une étude expérimentale bien menée doit permettre d'éliminer ces effets. D'autre part, on peut aussi n'appliquer qu'un seul traitement et comparer son effet à l'absence de traitement; le but de la comparaison est alors d'établir la réalité de l'effet de ce traitement.

## 4 CONDITIONS D'APPLICATION

La méthode est applicable si les deux conditions suivantes sont remplies :

- la série de différences  $d_i = x_i - y_i$  peut être considérée comme une série d'individus aléatoires indépendants;
- la répartition des différences  $d_i = x_i - y_i$  entre les observations appariées est supposée normale ou approximativement normale.

Si la répartition de ces différences s'écarte de la loi normale, le procédé décrit reste applicable, à condition que l'effectif de l'échantillon soit suffisamment élevé; plus elle s'écartera de la normale, plus l'effectif nécessaire sera élevé. Toutefois, même dans les cas extrêmes, un échantillon de 100 sera suffisant pour la plupart des applications pratiques.

1) Actuellement au stade de projet.

## 5 PRÉSENTATION FORMELLE DES CALCULS

Problème étudié . . . . . Conditions expérimentales . . . . .	
<b>Données statistiques</b>  Effectif de l'échantillon : $n =$  Somme des valeurs observées : $\Sigma x_i =$ $\Sigma y_i =$  Somme des différences : $\Sigma d_i =$  Somme des carrés des différences : $\Sigma d_i^2 =$  Valeur donnée : $d_0 =$  Degrés de liberté : $\nu = n - 1 =$  Niveau de signification choisi : $\alpha =$	<b>Calculs</b>  $\bar{d} = \frac{1}{n} (\Sigma x_i - \Sigma y_i)$  $= \frac{1}{n} \Sigma d_i =$  $s_d^2 = \frac{1}{n-1} [\Sigma d_i^2 - \frac{1}{n} (\Sigma d_i)^2] =$  $\sigma_d^* = \sqrt{s_d^2} =$  $A_1 = [t_{1-\alpha}(\nu)/\sqrt{n}] \sigma_d^* =$  $A_2 = [t_{1-\alpha/2}(\nu)/\sqrt{n}] \sigma_d^* =$
<b>Résultats</b>  Cas bilatéral :  L'hypothèse de l'égalité de la moyenne de la population des différences à $d_0$ (hypothèse nulle) est rejetée si l'on a : $ \bar{d} - d_0  > A_2$  Cas unilatéral : a) L'hypothèse selon laquelle la moyenne de la population des différences est au moins égale à $d_0$ (hypothèse nulle) est rejetée si l'on a : $\bar{d} < d_0 - A_1$ b) L'hypothèse selon laquelle la moyenne de la population des différences est au plus égale à $d_0$ (hypothèse nulle) est rejetée si l'on a : $\bar{d} > d_0 + A_1$	

NOTE  $t_{1-\alpha}(\nu)$  représente le fractile d'ordre  $1-\alpha$  de la variable  $t$  de Student à  $\nu$  degrés de liberté. Les valeurs de  $t_{1-\alpha}(\nu)/\sqrt{n}$  sont données dans la table 1.

TABLE 1 — Valeurs du rapport  $t_{1-\alpha}(\nu)/\sqrt{n}$ , pour  $\nu = n - 1$ 

$\nu = n - 1$	Cas bilatéral		Cas unilatéral	
	$\frac{t_{0,975}}{\sqrt{n}}$	$\frac{t_{0,995}}{\sqrt{n}}$	$\frac{t_{0,95}}{\sqrt{n}}$	$\frac{t_{0,99}}{\sqrt{n}}$
1	8,985	45,013	4,465	22,501
2	2,434	5,730	1,686	4,021
3	1,591	2,920	1,177	2,270
4	1,242	2,059	0,953	1,676
5	1,049	1,646	0,823	1,374
6	0,925	1,401	0,734	1,188
7	0,836	1,237	0,670	1,060
8	0,769	1,118	0,620	0,966
9	0,715	1,028	0,580	0,892
10	0,672	0,956	0,546	0,833
11	0,635	0,897	0,518	0,785
12	0,604	0,847	0,494	0,744
13	0,577	0,805	0,473	0,708
14	0,554	0,769	0,455	0,678
15	0,533	0,737	0,438	0,651
16	0,514	0,708	0,423	0,626
17	0,497	0,683	0,410	0,605
18	0,482	0,660	0,398	0,586
19	0,468	0,640	0,387	0,568
20	0,455	0,621	0,376	0,552
21	0,443	0,604	0,367	0,537
22	0,432	0,588	0,358	0,523
23	0,422	0,573	0,350	0,510
24	0,413	0,559	0,342	0,498
25	0,404	0,547	0,335	0,487
26	0,396	0,535	0,328	0,477
27	0,388	0,524	0,322	0,467
28	0,380	0,513	0,316	0,458
29	0,373	0,503	0,310	0,449
30	0,367	0,494	0,305	0,441
40	0,316	0,422	0,263	0,378
50	0,281	0,375	0,235	0,337
60	0,256	0,341	0,214	0,306
70	0,237	0,314	0,198	0,283
80	0,221	0,293	0,185	0,264
90	0,208	0,276	0,174	0,248
100	0,197	0,261	0,165	0,235
200	0,139	0,183	0,117	0,165
500	0,088	0,116	0,074	0,104
$\infty$	0	0	0	0

**EXEMPLE :** Les données figurant dans la table ci-après ont été recueillies au cours d'une expérience réalisée en vue de déterminer si le degré moyen d'usure de vilebrequin causée par des paliers de métaux différents dans un moteur à combustion interne, varie avec le métal.

TABLE 2 – Usure du vilebrequin après une durée de fonctionnement donnée en 0,000 01 in

Vilebrequin $i$	Usure avec		Différence $d_i = x_i - y_i$
	cuivre-plomb $x_i$	métal blanc $y_i$	
1	3,5	1,5	2,0
2	2,0	1,3	0,7
3	4,7	4,5	0,2
4	2,8	2,5	0,3
5	6,5	4,5	2,0
6	2,2	1,7	0,5
7	2,5	1,8	0,7
8	5,8	3,3	2,5
9	4,2	2,3	1,9
Totaux	34,2	23,4	10,8

## Caractéristiques techniques

## Données statistiques

Effectif de l'échantillon :

$$n = 9$$

Somme des valeurs observées :

$$\Sigma x_i = 34,2 \quad \Sigma y_i = 23,4$$

Somme des différences :

$$\Sigma d_i = 10,8$$

Somme des carrés des différences :

$$\Sigma d_i^2 = 19,22$$

Valeur donnée :

$$d_0 = 0$$

Degrés de liberté :

$$\nu = 8$$

Niveau de signification choisi :

$$\alpha = 0,01$$

## Calculs

ISO 3301-1975

$$\bar{d} = \frac{1}{9} (34,2 - 23,4) = 1,2$$

$$s_d^2 = \frac{1}{8} \left( 19,22 - \frac{10,8^2}{9} \right) = 0,782 5$$

$$\sigma_d^* = \sqrt{0,782 5} = 0,884 6$$

$$t_{0,995}/\sqrt{9} = 1,118$$

$$A_2 = 1,118 \times 0,884 6 = 0,99$$

## Résultat

Comparaison de la moyenne de la population avec la valeur donnée 0 :

Cas bilatéral :

$$|\bar{d} - d_0| = 1,2 > 0,99$$

L'hypothèse de l'égalité du degré d'usure pour les deux métaux est rejetée au niveau de 1 %.

## 6 ERREURS DE SECONDE ESPÈCE

La probabilité de rejeter l'hypothèse nulle lorsqu'elle est vraie est au plus égale au niveau de signification  $\alpha$ . Rejeter une hypothèse nulle qui est vraie est appelé une erreur de première espèce, et le choix de  $\alpha$  limite, de ce fait, le risque d'une telle erreur.

D'autre part, il est possible de commettre une erreur de seconde espèce qui consiste à accepter l'hypothèse nulle lorsqu'elle est fausse. La probabilité  $1 - \beta$  de rejeter l'hypothèse nulle lorsqu'elle est fausse est appelée puissance du test; la probabilité d'une erreur de seconde espèce est, de ce fait,  $\beta$ .

Pour un échantillon donné  $n$  et une erreur de première espèce  $\alpha$ , ces probabilités sont liées à la moyenne vraie  $D$  de la population des différences observées  $d_i = x_i - y_i$  pour laquelle on peut envisager diverses hypothèses alternatives, mais elles dépendent aussi de la vraie valeur de l'écart-type  $\sigma_d$  inconnu de la population des différences, écart-type dont on ne possède qu'une estimation médiocre si  $n$  est petit.

Il en résulte qu'il est impossible d'imposer, de manière rigide, une limite supérieure à la probabilité d'erreur de seconde espèce.

Cependant, les graphiques qui suivent montrent la relation entre la puissance du test,  $1 - \beta$ , et la moyenne réelle de la population divisée par l'écart-type correspondant,  $D/\sigma_d$ , pour des tests unilatéraux de l'hypothèse  $H_0 : D \leq 0$ , et ce pour diverses valeurs de  $n$  et pour les niveaux de signification 0,05 et 0,01.

De ces graphiques, on tire les conclusions suivantes :

1) La puissance du test est uniquement déterminée par la moyenne vraie des différences, mesurée en prenant l'écart-type pour unité, par le niveau de signification et par l'effectif de l'échantillon.

2) La fonction de puissance est une fonction strictement croissante de la différence moyenne vraie.

Elle est également strictement croissante avec l'effectif de l'échantillon et le niveau de signification  $\alpha$ , à condition que  $D > 0$  et  $\alpha$  différent de 0 et de 1.

3) Avec un niveau de signification de 0,05 et un effectif de 50, une puissance d'au moins 0,95 est déjà atteinte lorsque la différence moyennée vraie dépasse la moitié de l'écart-type des différences. Pour  $n = 20$ , cette puissance est atteinte pour  $D/\sigma_d = 0,78$  ou plus.

iTeh STANDARD PREVIEW  
(standards.iteh.ai)

ISO 3301:1975

<https://standards.iteh.ai/catalog/standards/sist/4deaf99b-576c-4cfb-a97d-0851b0148c50/iso-3301-1975>

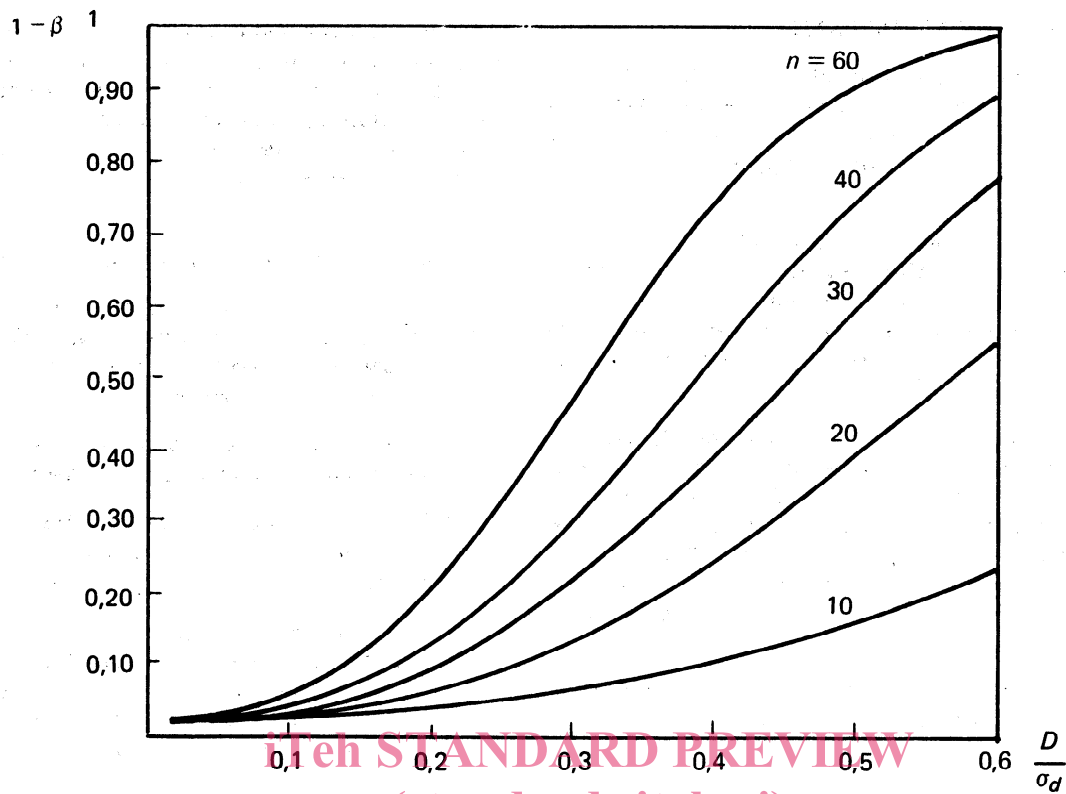


FIGURE 1 – Puissance du test de Student pour un échantillon (test unilatéral),  $\alpha = 0,01$

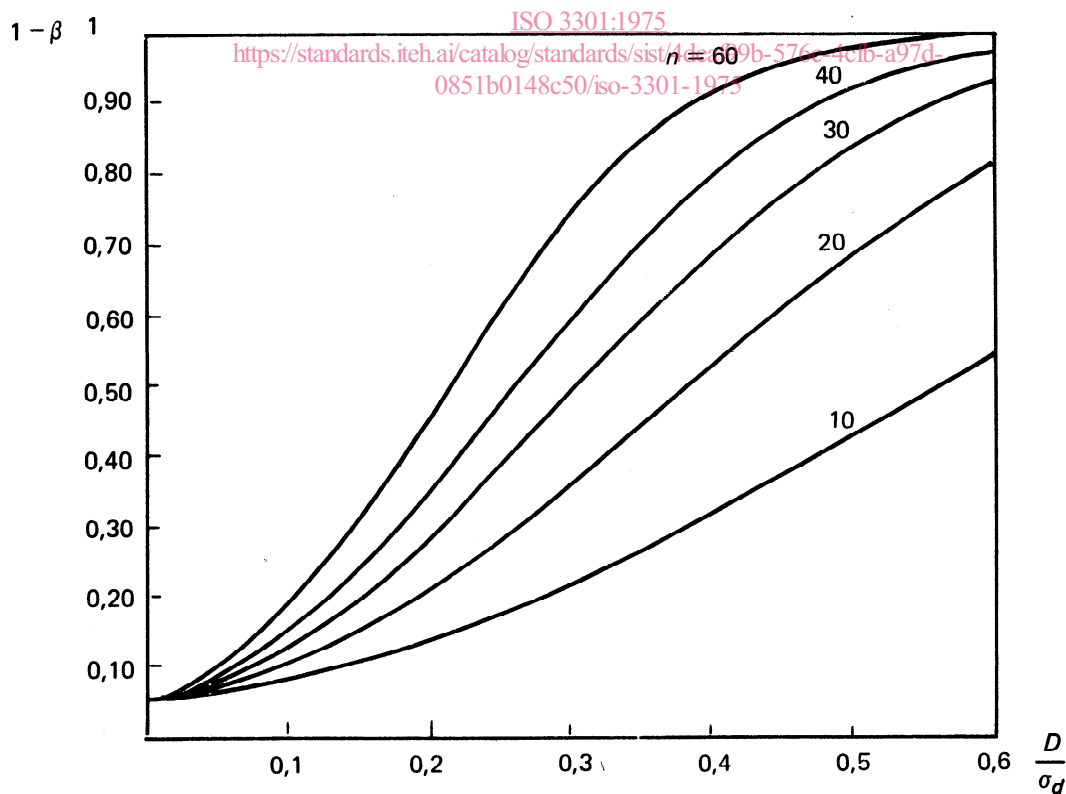


FIGURE 2 – Puissance du test de Student pour un échantillon (test unilatéral),  $\alpha = 0,05$

NOTE — Les graphiques sont basés sur l'ouvrage de D.B. OWEN, *Handbook of statistical tables*, Addison Wesley.