

CHAPITRE 2-3

LES TESTS D'HYPOTHÈSE

Plan

2-3.1 Introduction aux tests d'hypothèse	2
2-3.2 Cas type : un test d'hypothèse simple sur une moyenne	9
2-3.3 Un peu de terminologie relative aux tests d'hypothèse	20
2-3.4 Tests unilatéraux (one-sided tests)	22
2-3.5 Test de probabilité critique, sans seuil de signification pré-défini (p-value test)	26
2-3.6 Intervalles de confiance et marges d'erreur (estimation de la moyenne)	30
2-3.7 Détermination de la taille requise d'un échantillon (estimation de la moyenne)	35
2-3.8 Autres tests fréquemment utilisés	39
Table des valeurs critiques du test de Student (test bilatéral)	43

CHAPITRE 2-3

LES TESTS D'HYPOTHÈSE

Réf. : Wonnacott et Wonnacott (1992, chap. 9) présentent les tests d'hypothèse *après* avoir présenté l'estimation par intervalle (intervalles de confiance). Il n'est donc pas possible d'établir un parallèle parfait entre ce manuel et les présentes notes de cours.

2-3.1 Introduction aux tests d'hypothèse

Il y a plusieurs éléments qui entrent en jeu dans la formulation d'un test d'hypothèse :

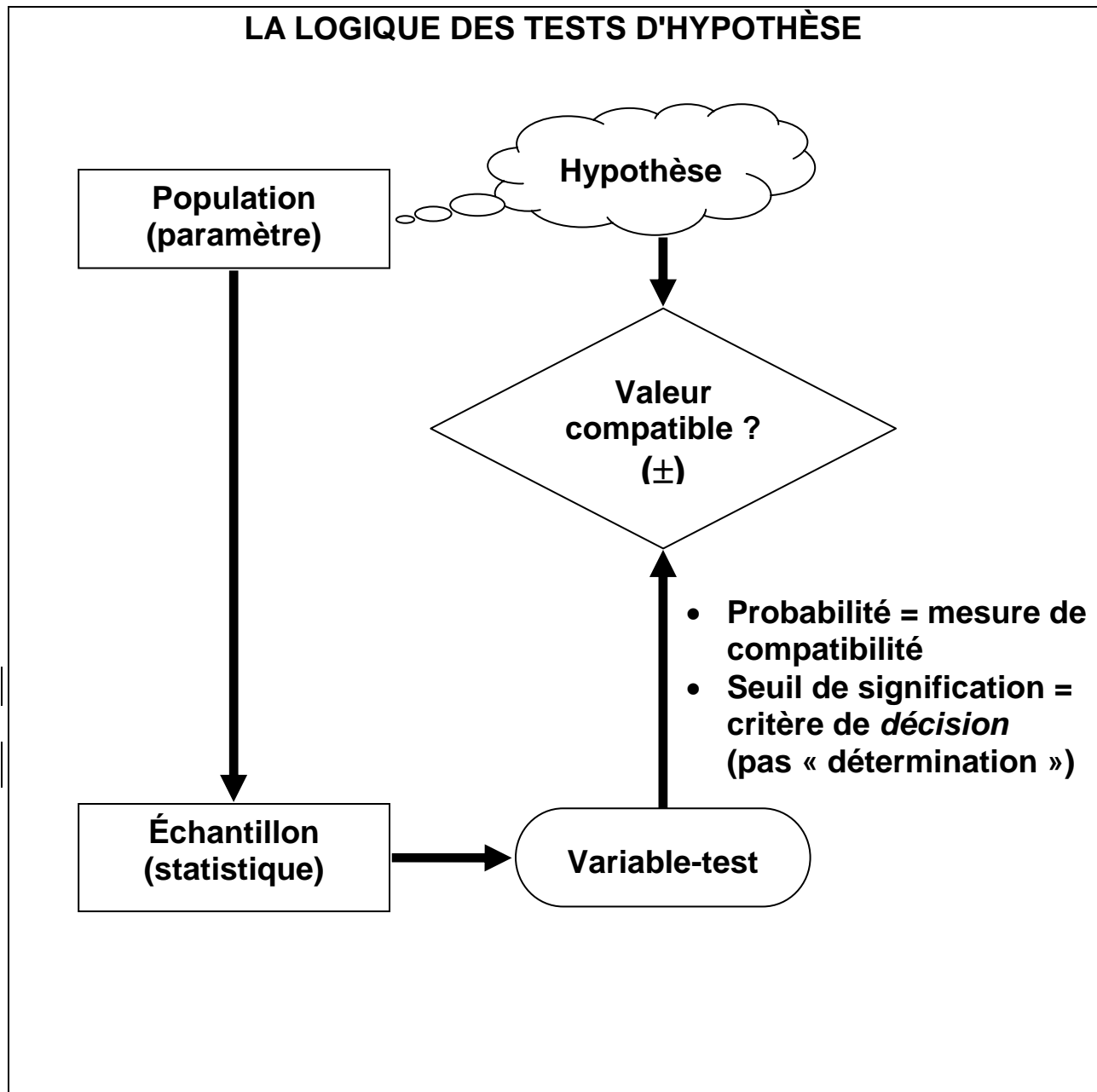
- la **population** (plus exactement, les paramètres de la population), qui est la réalité inconnue à propos de laquelle on veut tester...
- l'**hypothèse**, qui est un énoncé qui porte sur cette population (plus exactement, sur un ou plusieurs paramètres de cette population) et dont on ne sait pas si elle est vraie ou fausse.
- l'**échantillon**, c'est-à-dire l'ensemble des observations tirées de la population, à partir desquelles on cherchera à décider ¹ si l'on considère l'hypothèse comme vraie ou fausse.
- la **variable-test**, qui est une statistique de l'échantillon que l'on utilisera pour décider si l'on considère l'hypothèse comme vraie ou fausse.
- la **probabilité** qui, dans le contexte d'une logique probabiliste, est une mesure inverse ² le degré d'incompatibilité de la valeur observée de la variable-test avec l'hypothèse.
- le **seuil de signification**, qui est le seuil de probabilité critique, en-deçà duquel on décide que les observations (tel que résumées dans la variable-test) seront jugées suffisamment improbables pour être incompatibles avec l'hypothèse.

Ces éléments et les relations qui les unissent sont représentés au schéma 2a.

¹ Noter que l'on dit bien « décider », et non pas « déterminer ». Pour « déterminer », il faudrait arriver à une certitude. On pourra cependant, le cas échéant, « décider » de rejeter une hypothèse tout en sachant que, ce faisant, il n'est pas impossible que l'on commette une erreur.

² Attention à la double négation ! Plus la probabilité est *grande*, plus la valeur observée est *compatible* avec l'hypothèse ; donc, plus la probabilité est *petite*, plus la valeur observée tend à être *incompatible* avec l'hypothèse. La probabilité est bien une mesure *inverse* de l'*incompatibilité*.

Schéma 2a



Si l'on tentait de formaliser l'exemple des dromadaires australiens, on dirait que

- la population étudiée est la faune australienne sauvage ;
- l'hypothèse à tester est que le nombre de dromadaires dans cette faune est nul ;
- l'échantillon est constitué de l'ensemble des animaux observés jusqu'au moment du test ;
- la statistique utilisée est le nombre de dromadaires observés.

Mais laissons de côté cet exemple, qu'il serait incongru de mener beaucoup plus loin. Car ce qui fait défaut dans cet exemple, c'est la possibilité de mesurer le degré d'incompatibilité entre la valeur observée de la variable-test et l'hypothèse. Concrètement, dans l'affaire des dromadaires, il est impossible de construire un énoncé comme : « J'ai repéré X dromadaires jusqu'à maintenant. Je suppose que le dromadaire ne fait pas partie de la faune sauvage australienne (c'est-à-dire que je suppose que ceux que j'ai vus s'étaient échappés d'un zoo ou d'un cirque). Si ma supposition est vraie, les chances d'observer X dromadaires évadés sont de Y sur un million »... Dans un test d'hypothèse, il faut pouvoir quantifier ce Y .

Dans la formulation d'un test d'hypothèse, le noeud du problème est le choix d'une variable-test. D'ailleurs, plusieurs des variables-test fréquemment utilisées portent le nom de leur inventeur (Student ³, Fisher, Durbin-Watson,...). Une variable-test doit posséder certaines propriétés indispensables :

1. La valeur d'une variable-test doit dépendre à la fois des données de l'échantillon et de l'hypothèse à tester.

La variable-test constitue en effet le point de jonction entre le modèle (l'hypothèse à tester) et les observations. Elle est en quelque sorte une mesure de distance ou de dissimilarité entre les observations et les prédictions du modèle ou de l'hypothèse ⁴. La variable-test ne serait d'aucune utilité si elle n'incorporait pas l'information contenue dans l'échantillon, c'est-à-dire si sa valeur était indépendante des observations. Elle ne serait pas davantage utilisable si sa valeur était la même, quelle que soit l'hypothèse particulière que l'on veuille tester : la variable-test doit permettre de faire la différence entre les hypothèses qu'on décide de rejeter et celles qu'on décide de ne pas rejeter.

³ Student (l'Étudiant) est un pseudonyme que s'était donné le mathématicien W. S. Gossett (1876-1937) : <http://www-gap.dcs.st-and.ac.uk/~history/Mathematicians/Gosset.html>.

⁴ Cela est particulièrement évident du Khi-deux de Pearson qui sert au test d'indépendance dans les tableaux de contingence; voir 4-1.

2. La valeur d'une variable-test doit pouvoir être calculée, c'est-à-dire qu'elle ne doit pas dépendre de valeurs inconnues, seulement des observations et des données de l'hypothèse.
3. La variable-test doit avoir une distribution d'échantillonnage connue. C'est ce qui permet de mesurer le degré d'incompatibilité entre l'hypothèse et les observations.

Pour le moment, réfléchissons un peu à la nature aléatoire de la variable-test. Car pour avoir une distribution de probabilité (sa distribution d'échantillonnage), il faut que la variable-test soit une variable aléatoire. Par contre, une fois l'échantillon tiré, les valeurs observées ne sont plus aléatoires, elles sont fixes (de la même façon que le nombre de dromadaires repérés, une fois qu'on les a comptés). La contradiction n'est qu'apparente si l'on se rappelle la distinction entre une variable aléatoire et les valeurs qu'elle peut prendre. En effet, l'échantillon tiré n'est que l'un des échantillons possibles. À chacun d'eux correspond une valeur de la variable-test (d'autres voyageurs, ou le même voyageur à un autre moment, n'auraient pas nécessairement vu le même nombre de dromadaires). *Avant* que l'échantillon ne soit tiré, il y avait donc une multitude (dans certains cas une infinité) de valeurs possibles de la variable-test. *Ex ante*, lorsque l'on se place en imagination avant le tirage de l'échantillon, la variable-test est donc bel et bien une variable aléatoire à laquelle est associée une distribution de probabilité (sa distribution d'échantillonnage).

Ce concept est moins exotique qu'il n'y paraît à première vue. Les linguistes aiment citer cet exemple de deux titres de journaux :

HOMME MORDU PAR UN CHIEN

et

CHIEN MORDU PAR UN HOMME

Il s'agit des mêmes mots dans les deux cas ; seul leur ordre a été légèrement modifié. Qu'est-ce qui fait que le second titre est digne de la première page du *Journal de Montréal* mais pas le premier ? C'est que le second relate un événement surprenant, surprenant parce que sa probabilité *ex ante* était faible.

Dans la même veine, on considère qu'une personne qui a gagné à la loterie a été chanceuse, parce que, *ex ante*, la probabilité que ce soit elle était faible.

En somme, la distinction entre la valeur observée et sa distribution *ex ante* est analogue à la distinction entre ce qui s'est effectivement produit et ce que l'on attendait. Poétiquement,

l'événement qui se réalise n'abolit pas le souvenir de ce à quoi l'on s'attendait ; au contraire, la surprise naît du choc entre les deux.

Idée-clé No 16 :

La valeur de la variable test peut être calculée. Elle est donc fixe, une fois l'échantillon tiré. Mais la variable elle-même n'en est pas moins une variable aléatoire : avant que l'échantillon ne soit tiré, il y avait donc une multitude (dans certains cas une infinité) de valeurs possibles de la variable-test.

Pour réaliser un test d'hypothèse, il faut donc pouvoir mesurer la surprise, c'est-à-dire déterminer, en supposant que l'hypothèse à tester soit vraie, quelle était, avant de l'observer, la probabilité d'observer ce que l'on a observé (et qui est résumé dans la variable-test). Pour pouvoir déterminer cette probabilité, il faut avoir défini un modèle d'échantillonnage, c'est-à-dire un modèle du lien entre la population et l'échantillon. Il s'ensuit que le choix d'une variable-test et la spécification du modèle d'échantillonnage vont de pair.

Idée-clé No 17 :

Le choix d'une variable-test et la spécification du modèle d'échantillonnage vont de pair.

Le modèle d'échantillonnage comprend le plus souvent deux éléments :

- une hypothèse quant à la forme générale des lois de probabilité qui régissent le phénomène étudié dans la population ;
- la spécification de la méthode d'échantillonnage.

Une fois déterminée la probabilité de ce qu'on a observé lorsqu'on suppose que l'hypothèse est vraie, il ne reste plus qu'à décider si le résultat conduit à rejeter l'hypothèse ou non. Pour ce faire, on compare cette probabilité avec le seuil de probabilité critique préalablement choisi, en-deçà duquel on juge que les observations sont suffisamment improbables pour être incompatibles avec l'hypothèse. Ce seuil critique s'appelle le *seuil de signification*, parce que c'est le niveau de probabilité en-deçà duquel on décide de considérer le désaccord entre les observations et l'hypothèse comme *statistiquement significatif*.

Pour bien fixer les idées, l'argument logique du test d'hypothèse classique est formalisé dans l'encadré qui suit. Dans cet énoncé, les termes en italiques entre accolades sont les « variables » de l'argument. On applique l'argument à un cas particulier en remplaçant ces variables par les données pertinentes au cas particulier. L'argument est donc présenté un peu comme une formule mathématique, dont on calcule le résultat en remplaçant les variables par

leur valeur. Ou encore, comme un formulaire à remplir. Un tableau subséquent donnera la « valeur » qu'il faut donner à chaque « variable » pour appliquer l'argument au test d'une hypothèse simple sur une moyenne.

ARGUMENT DU TEST D'HYPOTHÈSE CLASSIQUE

Modèle d'échantillonnage, hypothèse et implications (syllogisme)

Si {modèle d'échantillonnage},
alors {variable}⁵ a la distribution {distribution d'échantillonnage}.

Or,

si {hypothèse},
alors {variable} est égale à la statistique {variable-test}.

Donc,

si {modèle d'échantillonnage} **et si** {hypothèse},
alors {variable-test} a la distribution {distribution d'échantillonnage}.

Règle de décision

On rejettera {hypothèse} si la valeur observée de {variable-test} appartient à un ensemble de valeurs extrêmes dont la probabilité est inférieure ou égale à {seuil de signification}⁶.

Étant donné

- la distribution {distribution d'échantillonnage}
- l'orientation du test (bilatéral, ou unilatéral, à gauche ou à droite, dépendamment de l'hypothèse complémentaire H_A),

l'ensemble de valeurs extrêmes ayant une probabilité égale à {seuil de signification} est défini par {zone de rejet}.

Décision

Or, la valeur observée de {variable-test} {fait/ne fait pas} partie de l'ensemble de valeurs extrêmes⁷ défini par {zone de rejet}.

Donc, la règle de décision choisie conduit à {rejeter/ne pas rejeter} {hypothèse}

⁵ Cette variable n'est ni une statistique, ni un paramètre. Elle n'est pas une statistique parce que sa valeur dépend des paramètres, mais elle n'est pas un paramètre parce que sa valeur dépend aussi d'une statistique.

⁶ Il serait plus simple de raisonner en termes de la probabilité de la valeur observée. Mais, s'agissant d'une variable aléatoire continue, la probabilité d'une valeur en particulier est infiniment petite. C'est pourquoi on raisonne en termes d'un ensemble de valeurs extrêmes défini par une ou deux valeurs critiques (selon que l'on fasse un test unilatéral ou bilatéral).

⁷ Cette formulation un peu alambiquée est exigée par le fait que la variable-test est la plupart du temps une variable continue, de sorte que la probabilité d'une valeur particulière est infiniment petite.

2-3.2 Cas type : un test d'hypothèse simple sur une moyenne

Maintenant que nous avons exposé la logique fondamentale des tests d'hypothèse, voyons quelles sont les étapes à suivre pour appliquer cette logique. Nous le ferons en référence à un cas type : un test d'hypothèse simple sur une moyenne.

Exemple : nous voulons étudier le temps passé par les habitants de l'Île de Montréal à écouter la radio. L'indicateur retenu (la variable x) pourrait être le nombre de minutes pendant lesquelles un individu a écouté la radio le mercredi 23 septembre 1998. La moyenne inconnue μ_x pourrait être le nombre moyen de minutes d'écoute radiophonique des habitants de l'Île de Montréal ce jour-là⁸. Quant à l'échantillon nous supposerons qu'il comporte 25 observations et que la durée moyenne d'écoute de l'échantillon est égale à 110 minutes, avec un écart type $s_x = 20$. L'hypothèse à tester pourrait être qu'en moyenne, les Montréalais ont écouté la radio ce jour-là pendant cent minutes⁹ :

$$H_0 : \mu_x = 100$$

Plus généralement, nous voulons étudier, dans une population donnée, une caractéristique représentée par une variable x . Nous nous intéressons à μ_x , la moyenne de x dans la population. Cette moyenne, μ_x , est inconnue. Mais nous disposons d'un échantillon tiré de la population et nous avons pu calculer m_x , la moyenne de x dans l'échantillon. Il s'agit de tester l'hypothèse H_0 que la « vraie » valeur de la moyenne est égale à une valeur particulière donnée, que nous dénoterons ici par la lettre grecque gamma : γ .

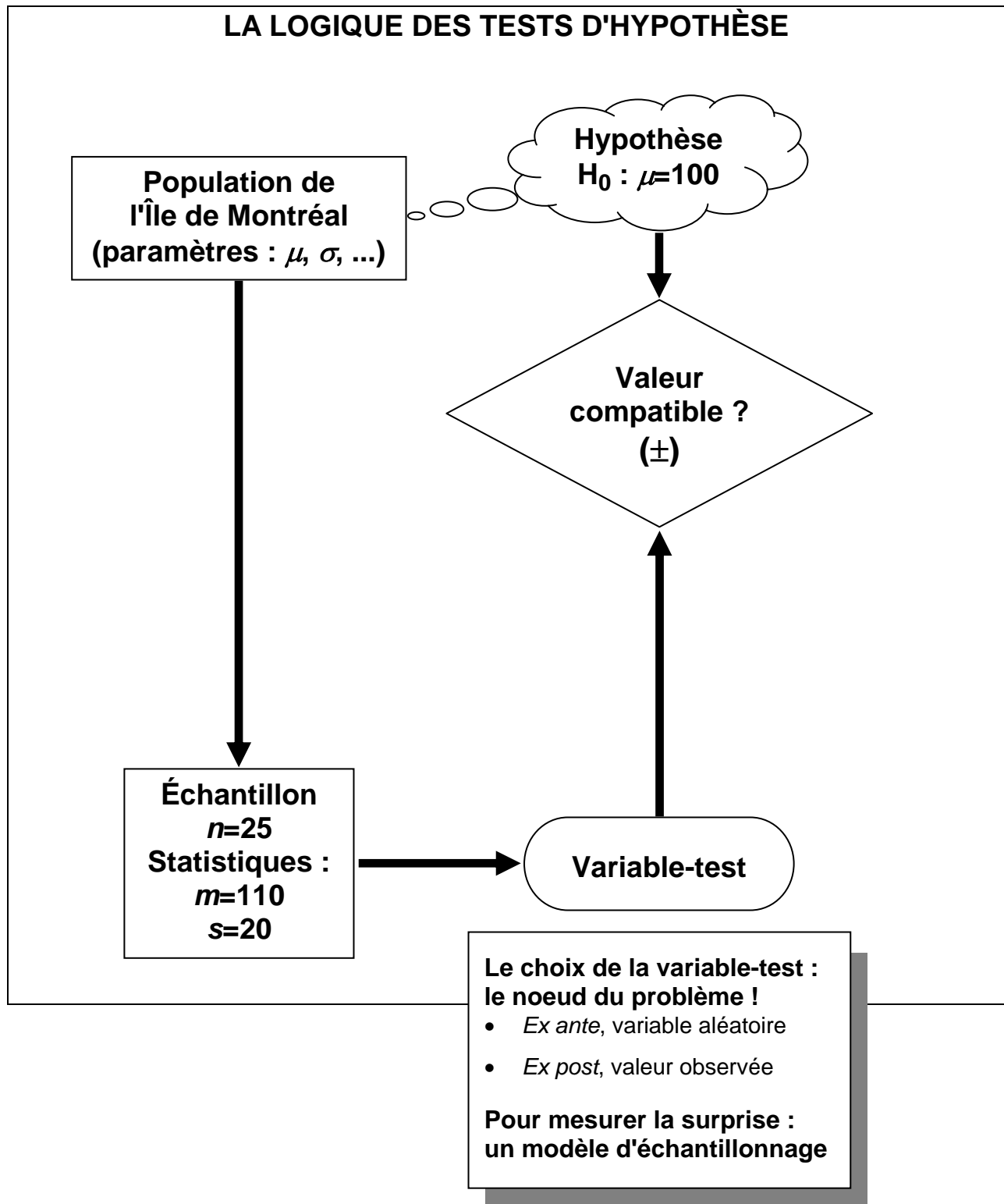
$$H_0 : \mu_x = \gamma$$

Pour compléter la présentation, le schéma 2b, en annexe, reprend le schéma 2a en y insérant les données de notre exemple de test d'hypothèse simple sur une moyenne.

⁸ Noter que la moyenne recherchée pourrait être, par exemple, le nombre moyen de minutes d'écoute radio pour un mercredi quelconque de la période du premier septembre au 30 octobre 1998. Cette autre moyenne réfère à une population différente. Mais si les observations ont toutes été recueillies le mercredi 23 septembre, l'échantillon risque de ne pas être représentatif de cette population plus large, à moins de croire que les comportements sont similaires pour tous les mercredis de la période ciblée (or il a fait mauvais le mercredi 23 septembre, mais pas le 16...).

⁹ Évidemment, rien n'empêche de tester l'hypothèse que $\mu_x = m_x = 110$. Mais cette hypothèse particulière n'est qu'une parmi l'infinité de possibilités.

Schéma 2b



Les étapes à suivre pour tester l'hypothèse sont les suivantes :

1. choisir une variable-test ;
2. vérifier que le modèle d'échantillonnage associé à cette variable-test est acceptable ;
3. calculer la valeur de la variable-test ;
4. choisir un seuil de signification ;
5. repérer les valeurs critiques de la variable-test (zone de rejet) ;
6. comparer la valeur de la variable-test avec les valeurs critiques et prendre la décision de rejeter ou non l'hypothèse.

Voyons maintenant de façon plus détaillée en quoi consiste chacune de ces étapes dans notre exemple.

Première étape : choix de la variable-test

Pour les utilisateurs que nous sommes, il n'est pas question d'inventer de toutes pièces une variable-test. Il s'agit plutôt de choisir dans la panoplie que la statistique met à notre disposition. En l'occurrence, nous allons appliquer le test de Student, qui utilise la variable-test suivante :

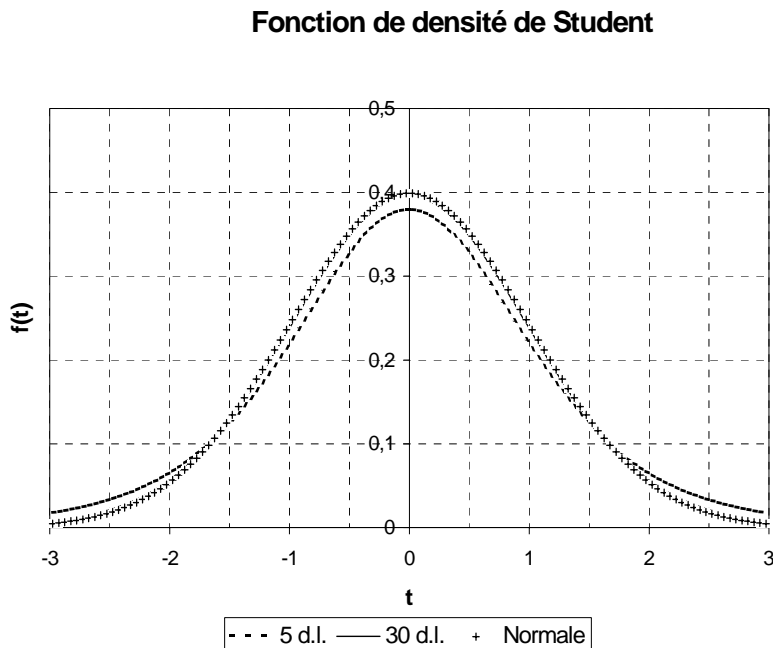
$$t_{n-1} = \frac{m_x - \gamma}{\left(\frac{s_x}{\sqrt{n}} \right)}$$

Le choix de cette variable-test se justifie parce que, *sous certaines conditions* (nous y reviendrons), la variable $\frac{m_x - \mu_x}{\left(\frac{s_x}{\sqrt{n}} \right)}$ a une distribution connue, que l'on désigne sous le nom de

distribution de Student. La distribution de Student ressemble à la normale, mais sa forme change un peu avec la valeur de n , tel qu'illustré à la figure 3 : on dit que cette variable a une distribution de Student avec $n-1$ degrés de liberté ¹⁰.

¹⁰ Cependant, pour des valeurs de n supérieures à 30, la distribution de Student est suffisamment proche de la normale pour que l'on considère souvent que la variable a une distribution approximativement normale.

FIGURE 3 – DISTRIBUTION DE STUDENT



La valeur de la variable-test est donc calculée simplement en substituant la valeur γ à μ_x dans la formule précédente. Et on peut affirmer que *si l'hypothèse H_0 est vraie*, alors $\gamma = \mu_x$ et la variable-test a une distribution de Student avec $n-1$ degrés de liberté :

$$\text{si } H_0 \text{ est vraie, alors } t_{n-1} = \frac{m_x - \gamma}{\left(\frac{s_x}{\sqrt{n}}\right)} = \frac{m_x - \mu_x}{\left(\frac{s_x}{\sqrt{n}}\right)}$$

Il faut noter que, dans la formule de calcul de la variable-test t_{n-1} , l'écart type utilisé est bien celui de l'échantillon

$$s_x^2 = \frac{1}{n-1} \sum_i (x_i - m_x)^2$$

et non celui de la population

$$\sigma_x^2 = \frac{1}{n} \sum_i (x_i - m_x)^2$$

On peut vérifier que cette variable-test a les qualités requises. D'abord, sa valeur dépend à la fois des données de l'échantillon (m_x , s_x , et n) et de l'hypothèse à tester (γ). Ensuite, cette valeur n'est pas une inconnue : on peut la calculer. Enfin, cette variable-test a une distribution

d'échantillonnage dont la forme générale est connue (distribution de Student), et dont la forme particulière dépend de ce sur quoi porte l'hypothèse à tester (la moyenne μ_x).

Deuxième étape : le modèle d'échantillonnage est-il acceptable ?

Nous avons choisi le test de Student, parce que, *sous certaines conditions*, la variable $\frac{m_x - \mu_x}{\left(\frac{s_x}{\sqrt{n}} \right)}$

a une distribution de Student. Quelles sont ces conditions ?

Les conditions suivantes sont suffisantes ¹¹

- Dans la population, la variable x a une distribution (approximativement ¹²) normale, avec une moyenne μ_x et un écart type σ_x inconnus.
- La population est de très grande taille et il en a été tiré un échantillon aléatoire simple de taille n .

Ces conditions constituent un modèle d'échantillonnage, qui spécifie la forme générale de la distribution de probabilité de x dans la population et le mode d'échantillonnage. Pour ce qui est de la distribution de probabilité de x , ce peut être un fait avéré ou une hypothèse, selon le contexte. Pour ce qui est du mode d'échantillonnage, c'est nous qui en avons décidé au moment de la constitution de l'échantillon : dans un échantillon aléatoire simple, chaque individu avait une chance égale de faire partie de l'échantillon.

Il revient au chercheur de décider si les conditions qui constituent le modèle d'échantillonnage sont acceptables. Le test ne porte pas sur le modèle d'échantillonnage, qui ne sera donc plus remis en question dans le cadre de ce test ¹³. Le test porte uniquement sur l'hypothèse H_0 .

Troisième étape : calcul de la valeur de la variable-test

Une fois choisie la variable-test, il suffit de calculer sa valeur en remplaçant les symboles par leur valeur numérique.

¹¹ Ces conditions sont suffisantes, mais non nécessaires. Si ces conditions sont réalisées, et si H_0 est vraie, alors la variable-test t_{n-1} aura une distribution de Student. Mais il existe d'autres jeux de conditions sous lesquelles la variable-test t_{n-1} aura aussi une distribution de Student.

¹² La variable x ne pourrait pas avoir une distribution *exactement* normale, parce qu'elle ne peut pas prendre de valeurs négatives, alors qu'une variable normale le peut.

¹³ Il est vrai qu'il existe des tests « de niveau supérieur », pour ainsi dire, qui portent sur certains aspects du modèle d'échantillonnage. Mais ces tests reposent eux-mêmes sur des modèles aléatoires plus généraux qui, à ce niveau, ne sont pas remis en question. On peut imaginer un test du modèle d'échantillonnage du test du modèle d'échantillonnage... Mais peu importe la « hauteur » du niveau auquel on s'élève, il y aura toujours au-dessus un modèle d'échantillonnage qui n'est pas remis en question.

Dans notre exemple de temps d'écoute radiophonique, la taille de l'échantillon est de $n=25$, la moyenne de x dans l'échantillon, $m_x=110$ et l'écart type, $s_x=20$; l'hypothèse à tester est

$$H_0 : \mu_x = 100$$

$$\text{Alors, } t_{24} = \frac{m_x - \gamma}{\left(\frac{s_x}{\sqrt{n}} \right)} = \frac{110 - 100}{\left(\frac{20}{\sqrt{5}} \right)} = 2,5$$

Quatrième étape : choix d'un seuil de signification

Il faut maintenant choisir un seuil de probabilité critique en-deçà duquel on jugera que les observations sont suffisamment improbables pour être incompatibles avec l'hypothèse. Les valeurs les plus fréquemment utilisées en sciences sociales sont 1 %, 5 % et 10 %. Aux fins de notre exemple, nous prendrons 5 %.

Cinquième étape : repérage des valeurs critiques de la variable-test (zone de rejet)

Consultons ensuite une table statistique (voir à la fin de ce chapitre, la « Table des valeurs critiques du test de Student »). Cette table nous apprend qu'avec $n-1=24$ degrés de liberté, il y a une probabilité de 0,05 (c'est-à-dire 5 %) que

$$t_{24} < -2,064 \text{ ou que } t_{24} > 2,064$$

De façon plus générale, la table statistique ¹⁴ du t de Student nous donne les valeurs critiques $\theta_{n-1}(\alpha)$ pour lesquelles, avec $n-1$ degrés de liberté, il y a une probabilité de α que

$$t_{n-1} < -\theta_{n-1}(\alpha) \text{ ou } t_{n-1} > +\theta_{n-1}(\alpha)$$

NOTE :

Certains auteurs utilisent la notation $t_{\alpha, n-1}$ pour désigner les valeurs critiques de la distribution de Student ; dans cette notation, α est le seuil de signification (ici, 0,05) et $n-1$ est le nombre de degrés de liberté (ici, 24). Pour éviter la confusion entre les valeurs critiques et la variable-test elle-même, nous éviterons cette notation et nous désignerons les valeurs critiques par $\theta_{n-1}(\alpha)$: ici par exemple,

$$\theta_{24}(0,05) = 2,064$$

¹⁴ Attention ! La table du t de Student donnée en annexe donne les valeurs du t pour un test bilatéral, comme celui que nous développons ici. Il faut cependant savoir qu'il existe des tables qui donnent les valeurs du t pour un test unilatéral (voir ci-après). Pour utiliser une table unilatérale pour faire un test bilatéral, il faut *multiplier* les probabilités par deux; nous verrons que pour utiliser une table bilatérale pour faire un test unilatéral, il faut *diviser* les probabilités par deux.

Sixième étape : comparaison de la valeur de la variable-test avec les valeurs critiques et décision

Nous avons maintenant en mains tous les éléments nécessaires pour conclure. Nous avons calculé que la variable-test $t_{n-1} = 2,5$. Ce que la table statistique nous apprend, c'est que, si H_0 est vraie, cette valeur est assez improbable : la probabilité d'observer une valeur aussi éloignée de zéro est de moins de 5 %. Puisque nous avons choisi 5 % comme seuil de signification, nous décidons de rejeter H_0 : nous concluons que la moyenne de x dans la population n'est pas égale à 100, parce que nous jugeons que nos observations sont probablement incompatibles avec cette hypothèse.

De façon plus générale, on rejettera l'hypothèse, à un seuil de signification de α , si

$$t_{n-1} < -\theta_{n-1}(\alpha) \text{ ou } t_{n-1} > +\theta_{n-1}(\alpha)$$

Évidemment, si la valeur de la variable-test n'avait pas dépassé les valeurs critiques (ce qui aurait pu arriver avec un autre échantillon), nous n'aurions pas rejeté l'hypothèse.

En résumé, nous avons suivi les étapes suivantes :

1. Nous avons choisi une variable-test ayant les propriétés requises, le t de Student.
2. Nous avons examiné les conditions sous lesquelles le test de Student est applicable (le modèle d'échantillonnage) et nous avons décidé qu'elles étaient acceptables.
3. Nous avons calculé la valeur de cette variable-test ($t_{24} = 2,5$).
4. Nous avons choisi un seuil de signification ($\alpha = 5 \%$).
5. Nous avons repéré les valeurs critiques dans la table statistique : si l'hypothèse est vraie, il y a une probabilité α que la variable-test tombe à l'extérieur de l'intervalle défini par les valeurs critiques $-\theta_{n-1}(\alpha)$ et $+\theta_{n-1}(\alpha)$ (ici, la probabilité que t_{24} soit inférieure à $-2,064$ ou supérieure à $+2,064$ est de 5 %).
6. Nous avons comparé la valeur de la variable-test ($t_{24} = 2,5$) avec les valeurs critiques données par la table. Dans notre exemple, nous avons constaté que, si l'hypothèse était vraie, les observations, telles que résumées dans la variable-test, auraient été improbables (probabilité inférieure à 5 %). Et puisque cette probabilité était inférieure au seuil de signification choisi, nous avons rejeté l'hypothèse.

Le tableau qui suit donne la « valeur » qu'il faut attribuer à chaque « variable », dans l'argument du test d'hypothèse classique, pour appliquer l'argument au test d'une hypothèse simple sur une moyenne.

APPLICATION DE L'ARGUMENT AU TEST D'UNE HYPOTHÈSE SIMPLE SUR UNE MOYENNE

Formulation générale	Exemple : $n = 25 ; m_x = 110 ; s_x = 20 ; \alpha = 0,05$
{hypothèse}	
$H_0 : \mu_x = \gamma$	$H_0 : \mu_x = 100$
{modèle d'échantillonnage}	
<ul style="list-style-type: none"> • Dans la population, la variable x a une distribution (approximativement) normale, avec une moyenne μ_x et un écart-type σ_x inconnus. • La population est de très grande taille et il en a été tiré un échantillon aléatoire simple de taille... 	
n	25
{variable}	
$\frac{m_x - \mu_x}{\left(\frac{s_x}{\sqrt{n}}\right)}$	$\frac{110 - \mu_x}{\left(\frac{20}{\sqrt{25}}\right)}$
{distribution d'échantillonnage}	
Distribution de Student avec...	
$n-1$ degrés de liberté	24 degrés de liberté
{variable-test}	
$t_{n-1} = \frac{m_x - \gamma}{\left(\frac{s_x}{\sqrt{n}}\right)}$	$t_{24} = \frac{110 - 100}{\left(\frac{20}{\sqrt{25}}\right)} = 2,5$
{seuil de signification}	
α	0,05
Orientation du test et {zone de rejet}	
test bilatéral ($H_A : \mu_x \neq \gamma$) :	test bilatéral ($H_A : \mu_x \neq 100$) :
$t_{n-1} < -\theta_{n-1}(\alpha)$ ou $t_{n-1} > +\theta_{n-1}(\alpha)$	$t_{24} < -2,064$ ou $t_{24} > 2,064$
OU	
test unilatéral à droite ($H_A : \mu_x > \gamma$) :	test unilatéral à droite ($H_A : \mu_x > 100$) :
$t_{n-1} > +\theta_{n-1}(2\alpha)$	$t_{24} > 1,711$
OU	
test unilatéral à gauche ($H_A : \mu_x < \gamma$) :	test unilatéral à gauche ($H_A : \mu_x < 100$) :
$t_{n-1} < -\theta_{n-1}(2\alpha)$	$t_{24} < -1,711$

Revenons sur le choix du seuil de signification. Que serait-il arrivé si nous avions choisi un critère différent, 1 % par exemple ? La table en annexe nous apprend que

$$t_{24}(0,01) = 2,797$$

c'est-à-dire qu'avec $n-1=24$ degrés de liberté, la probabilité que

$$t_{24} < -2,797 \text{ ou que } t_{24} > 2,797$$

est de 0,01 (c'est-à-dire 1 %). Ainsi donc, si nous avons choisi un seuil de signification de 1 %, la valeur de la variable-test (2,5) se serait trouvée à l'intérieur des bornes fixées par les valeurs critiques $-2,797$ et $+2,797$: avec ce critère plus exigeant, nous ne pourrions pas rejeter l'hypothèse. Toutefois, cela ne veut pas dire pour autant que nous accepterions l'hypothèse !

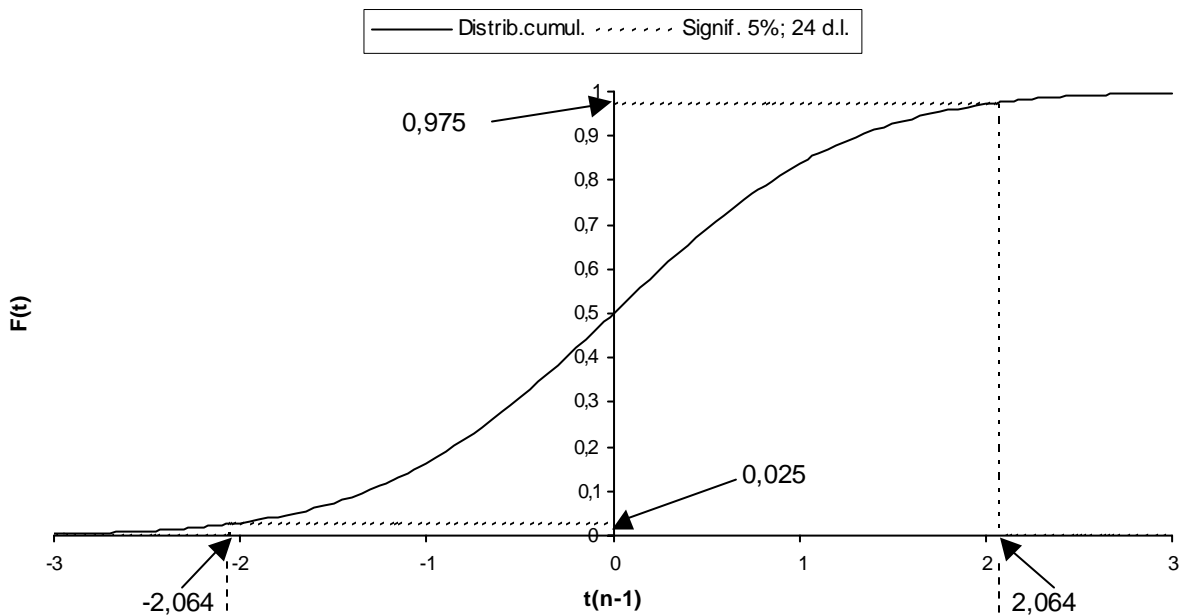
En termes généraux, plus le seuil de signification choisi est petit, plus la valeur critique est grande. Si l'on compare les décisions qui seraient prises à deux seuils de signification différents, il est évident qu'il y a des hypothèses qui correspondent à des valeurs de la variable-test qui seront au-delà de la valeur critique pour le seuil de signification le plus élevé, mais pas au-delà de la valeur critique pour le seuil de signification le plus exigeant (le plus petit). Telles hypothèses seraient rejetées au seuil de signification plus élevé (moins exigeant), mais pas au seuil de signification plus faible (plus exigeant).

La relation entre le choix du seuil de signification, les valeurs critiques et la zone de rejet est illustrée à la figure 4.

En guise de synthèse, le schéma 2c reprend la structure du schéma 2a, en y intégrant l'ensemble des concepts que nous venons d'expliquer et qui interviennent dans un test d'hypothèse.

FIGURE 4 – TEST DE STUDENT BILATÉRAL

Test t de Student : distribution cumulative



Test t de Student : fonction de densité

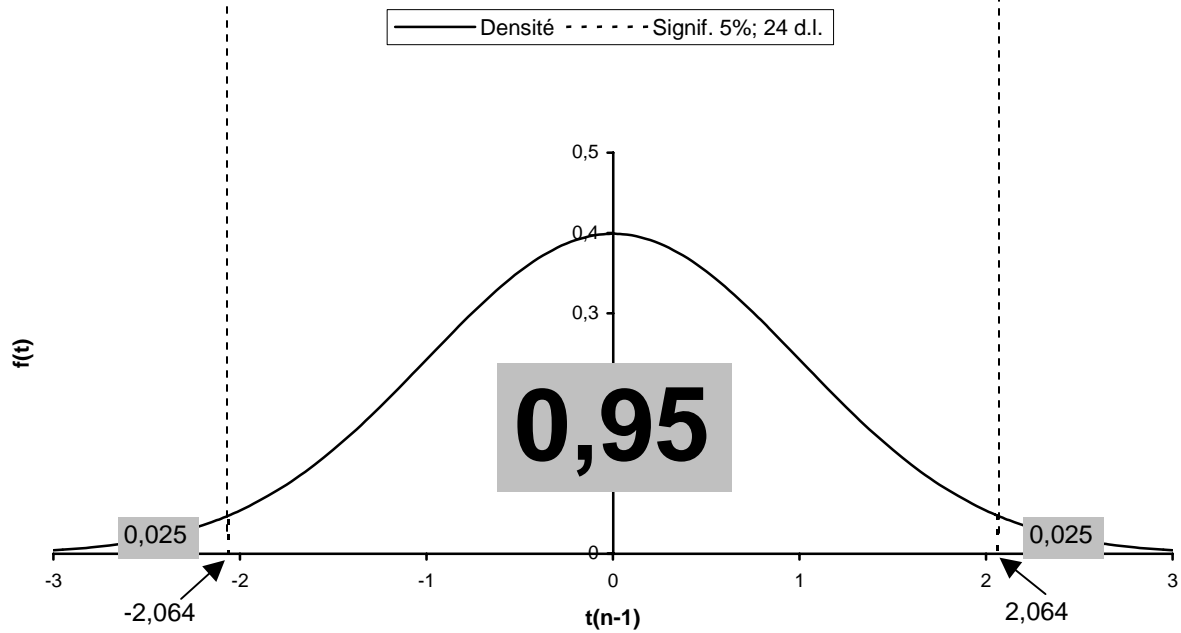
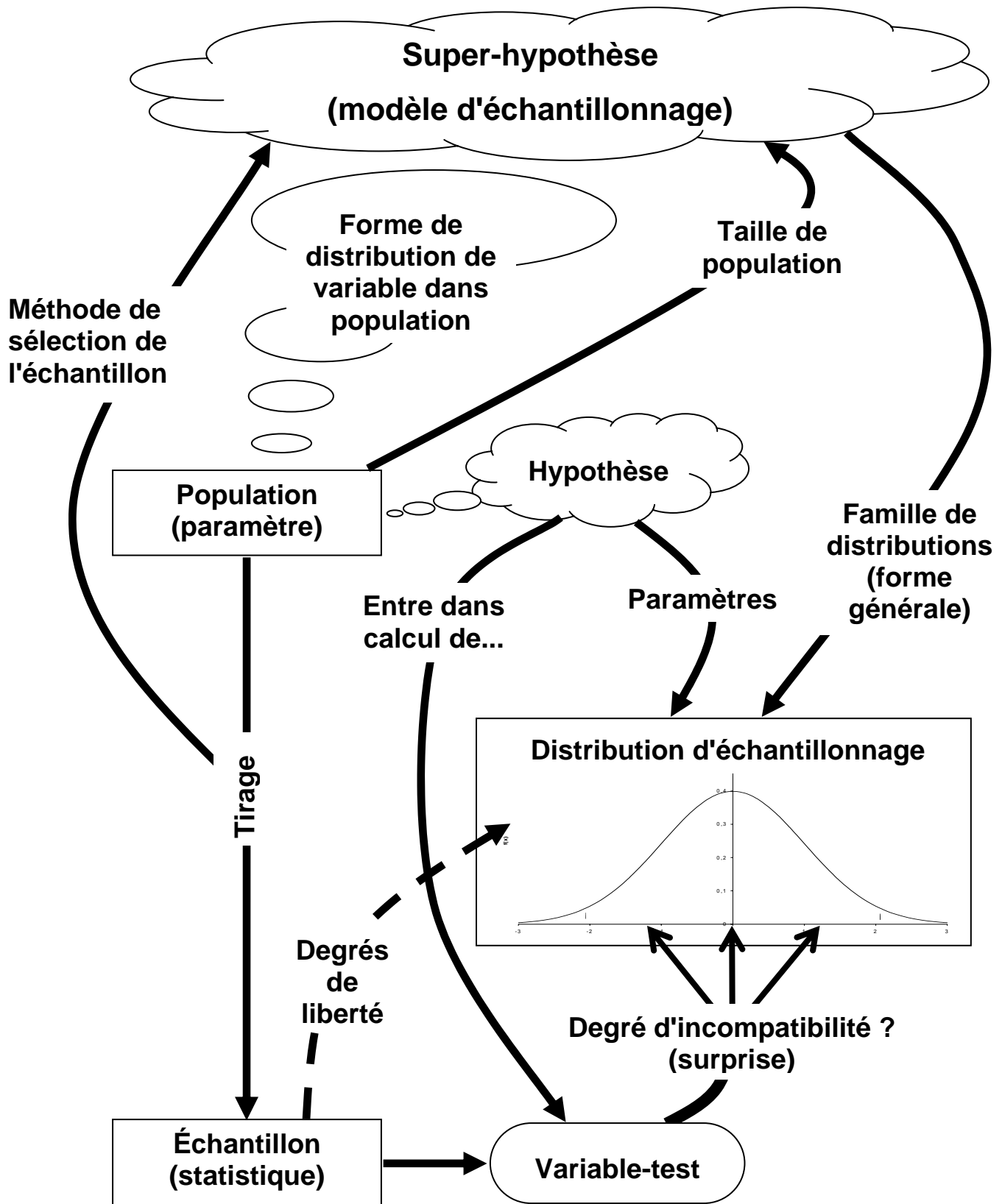


SCHÉMA 2C : LA LOGIQUE DES TESTS D'HYPOTHÈSE



2-3.3 Un peu de terminologie relative aux tests d'hypothèse

Réf. : Freund (1962), chap. 11.

Hypothèse simple, hypothèse composée ; hypothèse nulle, hypothèse complémentaire

La démarche qui vient d'être exposée est celle de tous les tests d'une hypothèse simple. Une hypothèse *simple* est une hypothèse qui spécifie complètement la distribution de la variable-test ; en pratique, une hypothèse simple attribue à un paramètre une valeur unique. Une hypothèse qui porte sur un éventail de valeurs possibles est une hypothèse *composée*. Par exemple

- hypothèse simple : $\mu_X = 0$
- hypothèse composée : $\mu_X > 0$

Dans le test d'une hypothèse simple, l'hypothèse testée est souvent appelée l'*hypothèse nulle*¹⁵ et notée H_0 . Lorsqu'un test conduit au rejet de l'hypothèse nulle, il s'ensuit logiquement que l'on accepte l'hypothèse *complémentaire* (*alternate hypothesis*), souvent notée H_A . L'hypothèse complémentaire d'une hypothèse simple est généralement une hypothèse composée. Par exemple,

$$H_0 : \mu_X = 0$$

$$H_A : \mu_X \neq 0$$

Seuil de signification, zone de rejet et erreurs de type I et II

Réf. : Wonnacott et Wonnacott (1992, p. 344-345, 349-350 et 354-357)

Une fois qu'on a choisi un seuil de signification, l'ensemble des valeurs de la variable-test dont la probabilité de réalisation est en-dessous du seuil de signification s'appelle la *zone de rejet* (ou *région critique*, ou *zone critique*) du test (voir figure 3).

Un test statistique est basé sur un raisonnement probabiliste. Sa conclusion n'est donc pas certaine ; elle est seulement probable. Une décision prise à la lumière d'un test statistique comprend toujours un certain risque d'erreur.

¹⁵ Selon Knapp (1996), cette expression s'explique de multiples façons : (1) l'hypothèse testée est souvent que la valeur du paramètre est nulle ; (2) c'est l'hypothèse « neutre », selon laquelle il n'y a rien qui sorte de l'ordinaire ; (3) le chercheur espère souvent que les données « annulent » cette hypothèse (personnellement, je considère cette dernière explication comme un peu « tirée par les cheveux »).

Le tableau suivant résume les différents types erreurs qu'on peut commettre :

		Situation (inobservable)	
		H_0 est vraie	H_0 est fausse
Décision	Rejeter H_0	Erreur de type I	Bonne décision
	Ne pas rejeter H_0	Bonne décision	Erreur de type II

Pour chacune des deux situations possibles, les probabilités associées à ces possibilités sont :

		Situation (inobservable)	
		H_0 est vraie	H_0 est fausse
Décision	Rejeter H_0	Seuil de signification α	Puissance $(1 - \beta)$
	Ne pas rejeter H_0	$(1 - \alpha)$	β

Dans le jargon de la statistique, on appelle *erreur de type I* l'erreur que l'on commet quand on rejette l'hypothèse H_0 alors qu'en fait, elle est vraie. Lorsqu'on applique un test statistique à l'hypothèse H_0 , la probabilité de commettre une erreur de type I est la probabilité que la variable-test tombe dans la zone de rejet même si H_0 est vraie. Et quelle est cette probabilité ? C'est, par définition, le seuil de signification choisi ! *Si H_0 est vraie, la probabilité α de commettre une erreur de type I est le seuil de signification choisi pour le test.* Le choix du seuil de signification est donc le choix du niveau accepté du risque de commettre une erreur de type I si H_0 est réellement vraie.

Une *erreur de type II* consiste à accepter (plus exactement à ne pas rejeter) une hypothèse alors qu'elle est fausse. Dans l'éventualité où H_0 est fausse, la probabilité β de commettre une erreur de type II est difficile à évaluer, parce qu'il y a alors plusieurs distributions possibles (généralement une infinité) pour la variable-test. Lorsqu'on peut évaluer cette probabilité, alors la probabilité d'éviter une erreur de type II ($1-\beta$) s'appelle la *puissance* du test.

Idéalement, on souhaiterait un test pour lequel les probabilités des deux types d'erreurs sont faibles (α et β petits). Mais on peut voir intuitivement que, pour un test donné, plus α est petit, plus β est grand. En effet, lorsque α est petit, la zone de rejet est petite elle aussi, ce qui

augmente les chances de ne pas rejeter H_0 et, donc, augmente β . En somme, la décision que l'on prend sur la base d'un test d'hypothèse est un pari, dans lequel on fait un compromis entre deux risques d'erreurs : le risque d'erreur de type I et le risque d'erreur de type II. Un bon test statistique est donc un test qui, pour tout niveau donné de probabilité d'erreur de type I, a la plus petite probabilité possible d'erreur de type II ; en d'autres mots, le meilleur test est celui qui est le plus puissant à chaque seuil de signification.

Distributions asymptotiques

Réf. : Wonnacott et Wonnacott (1992, p. 224-226)

Le modèle d'échantillonnage ne permet pas toujours de spécifier complètement la distribution d'échantillonnage d'une variable-test. On peut souvent sortir de l'impasse en ayant recours à la *distribution asymptotique* de la variable-test. On peut montrer en effet que plusieurs distributions d'échantillonnage tendent à se rapprocher d'une distribution connue à mesure que la taille de l'échantillon augmente. Cette distribution connue s'appelle une distribution asymptotique. Lorsque l'échantillon est « suffisamment grand », la distribution asymptotique peut être prise comme approximation de la distribution d'échantillonnage exacte.

Par exemple, la distribution asymptotique d'une distribution de Student est la distribution normale (voir la figure 3). Dans ce cas particulier, il n'y a pas d'impasse et l'on pourrait spécifier complètement la distribution d'échantillonnage de la variable-test pour chaque valeur du nombre de degrés de liberté. Néanmoins, au-delà d'une certaine taille d'échantillon (et du nombre de degrés de liberté qui s'ensuit), on considère que la distribution de Student est tellement proche de la normale qu'il ne vaut pas la peine de référer à la distribution exacte. En pratique, l'échantillon est souvent considéré comme « suffisamment grand » lorsque la distribution de Student a plus de 30 degrés de liberté.

Dans un tableau de contingence, le test d'indépendance du Khi-deux de Pearson (voir le chapitre 4-1) est basé sur une distribution asymptotique, en l'occurrence la distribution du χ^2 .

2-3.4 Tests unilatéraux (one-sided tests)

L'exemple présenté en 2-3.2 était un test bilatéral (*two-sided test*). Dans ce test, on accorde la même importance aux déviations à la hausse ou à la baisse par rapport à l'hypothèse nulle

$$H_0 : \mu_x = \gamma$$

Dans certaines circonstances toutefois, on n'accorde d'importance qu'à l'une des deux possibilités. Par exemple, supposons qu'un acheteur veuille s'assurer qu'un produit respecte une norme de qualité moyenne. Disons que la qualité est mesurée par un indicateur x et que la norme à respecter est que la valeur moyenne μ_x de l'indicateur de qualité x , soit au moins égale à γ . Pour décider s'il acceptera le lot (la population), l'acheteur examine un échantillon tiré du lot, dont il calcule la qualité moyenne m_x . Il va de soi que l'acheteur ne sera pas déçu si la qualité moyenne du produit excède la norme. Dans ce cas, l'hypothèse complémentaire pertinente n'est pas

$$H_A : \mu_x \neq \gamma$$

mais plutôt

$$H_A : \mu_x > \gamma$$

En d'autres mots, pour l'acheteur, rejeter H_0 , cela signifie accepter le lot, c'est-à-dire accepter l'hypothèse que la qualité du lot excède la norme. Dans ces conditions, la zone de rejet se situe d'un seul côté du zéro, à droite. La logique est simple, pour peu que l'on s'arrête à y réfléchir : si m_x est suffisamment grande pour que l'on rejette l'hypothèse $H_0 : \mu_x = \gamma$, alors *a fortiori*, on rejettera toute hypothèse $H_0' : \mu_x = \gamma'$ pour toute valeur $\gamma' < \gamma$. Il est à noter que ce raisonnement amène l'acheteur à n'accepter le lot que si la qualité moyenne de l'échantillon dépasse la norme d'une marge suffisante ¹⁶. En d'autres mots, lorsque l'acheteur n'accepte pas le lot, ce n'est pas parce qu'il rejette l'hypothèse que le lot respecte la norme ; c'est plutôt parce qu'il est statistiquement incapable de rejeter l'hypothèse que le lot ne respecte pas la norme.

Il faut se rappeler par ailleurs que l'application d'un test unilatéral change la relation entre le seuil de signification et les valeurs critiques qui définissent la zone de rejet. Par exemple, si la zone de rejet utilisée est définie par

$$t_{24} > 2,064$$

au lieu de

$$t_{24} < -2,064 \text{ ou } t_{24} > 2,064$$

alors la probabilité de rejeter H_0 alors que cette hypothèse serait vraie n'est pas de 5 %, mais de 2,5 %. Le seuil de signification de ce test unilatéral serait donc de 2,5 %.

¹⁶ Pour un test unilatéral avec un seuil de signification égal à α , cette marge est la marge d'erreur bilatérale associée à un niveau de confiance de $(1-2\alpha)$. Voir 2-3.4.

Exemple ¹⁷ : Un acheteur doit décider s'il accepte un lot de 100 000 tubes cathodiques. La norme de qualité exigée est que la durée moyenne des tubes du lot soit d'au moins 1200 heures. Des tests sont effectués sur un échantillon de 100 tubes : la durée de vie moyenne des tubes de l'échantillon se révèle être 1265 heures, l'écart type étant de 300 heures. L'hypothèse à tester est

$$H_0 : \mu_x = 1200$$

La statistique t est donnée par

$$t_{n-1} = \frac{m_x - \gamma}{\left(\frac{s_x}{\sqrt{n}} \right)} = \frac{1265 - 1200}{\left(\frac{300}{\sqrt{100}} \right)} = 2,17$$

Si le seuil de signification retenu est de 0,005, avec 99 degrés de liberté, la valeur critique se situe entre 2,626 et 2,632 (voir la table des valeurs critiques du test de Student) ¹⁸. On ne peut pas rejeter H_0 et l'acheteur n'acceptera pas le lot.

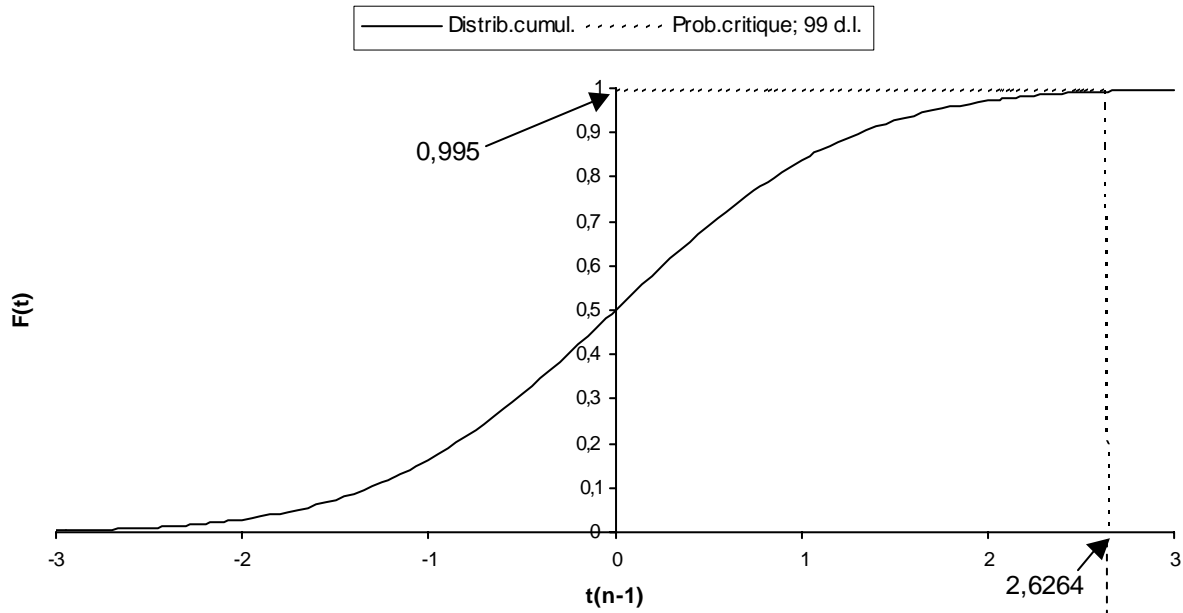
Dans l'exemple qui précède, la zone de rejet est à droite du zéro. Il y a évidemment des circonstances où la zone de rejet serait à gauche.

¹⁷ Cet exemple est inspiré de Wonnacott et Wonnacott (1991, p. 333-334).

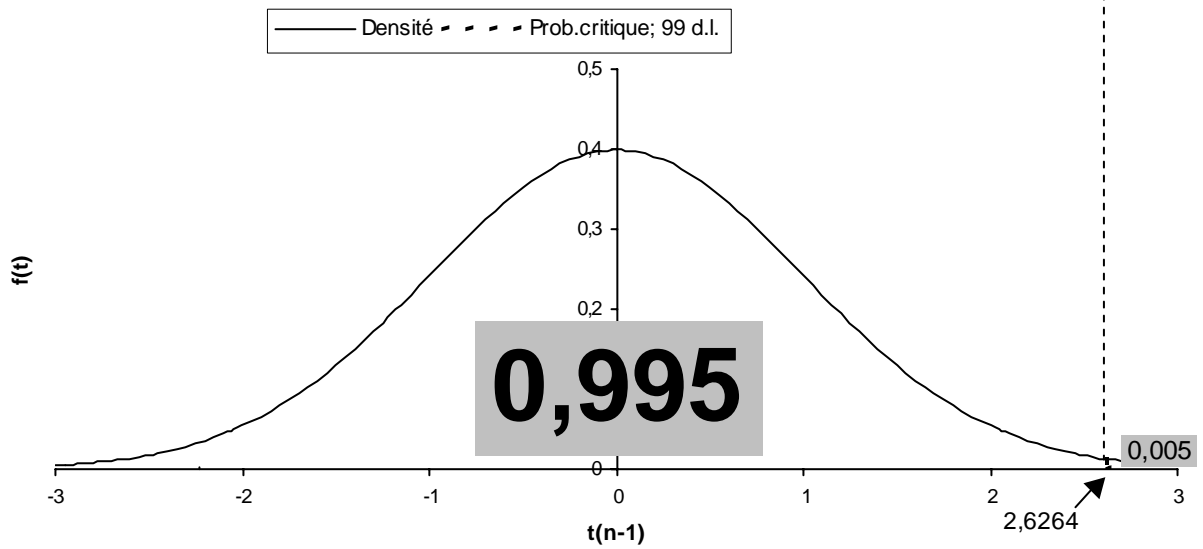
¹⁸ Bien entendu, on peut calculer la valeur critique exacte au moyen de la fonction `LOI.STUDENT.INVERSE` du logiciel Excel (`TINV` dans la version anglaise), comme on a calculé les valeurs de la table : la valeur exacte est 2,626402, obtenue au moyen de la commande `= LOI.STUDENT.INVERSE(0,005*2;99)`, où la multiplication par deux est rendue nécessaire par le fait que la fonction calcule la valeur critique pour un test bilatéral.

FIGURE 5A – TEST D'HYPOTHÈSE UNILATÉRAL

Test t de Student : distribution cumulative



Test t de Student : fonction de densité



L'argument logique du test d'hypothèse classique, formalisé dans l'encadré présenté à la fin de 2-3.1, s'applique évidemment à un test unilatéral. On peut l'appliquer à un test d'hypothèse unilatéral à droite en faisant, dans le tableau de 2-3.2 qui donne la « valeur » qu'il faut attribuer à chaque « variable », la substitution suivante :

Pour un test unilatéral à droite ($H_A : \mu_x > \gamma$) : avec 24 degrés de liberté et $\alpha = 0,05$,

$$\{\text{zone de rejet}\} = t_{n-1} > +t_{n-1}(2\alpha) = t_{24} > 1,711 \text{ au lieu de}$$

$$\{\text{zone de rejet}\} = t_{n-1} < -t_{n-1}(\alpha) \text{ ou } t_{n-1} > +t_{n-1}(\alpha) = t_{24} < -2,064 \text{ ou } t_{24} > 2,064$$

2-3.5 Test de probabilité critique, sans seuil de signification pré-défini (p-value test)

Réf. : Wonnacott et Wonnacott (1992, p. 333-337)

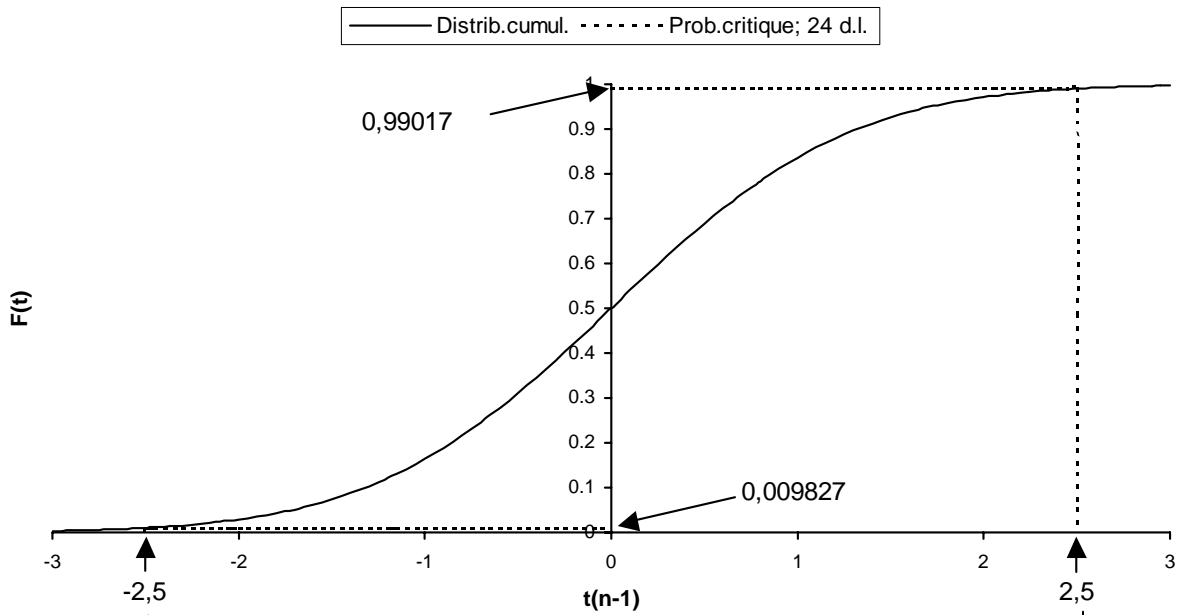
Les tests statistiques classiques se font en comparant la valeur calculée d'une variable-test avec les valeurs repères données dans les tables. Cependant, pour plusieurs variables-test couramment utilisées, les logiciels d'application statistiques donnent maintenant le seuil de signification pour lequel la valeur de la statistique serait exactement à la limite de la zone de rejet¹⁹. Ce seuil de signification s'appelle la *probabilité critique* (*p-value*). Dans la présentation de résultats, on donne donc de plus en plus la valeur de cette probabilité, plutôt que d'indiquer que l'hypothèse est ou n'est pas rejetée au seuil de signification de 1 %, 5 % ou 10 %. C'est là en quelque sorte une façon de rapporter les résultats avec un maximum de transparence, et qui laisse le lecteur libre de choisir le seuil de signification et de décider s'il est d'accord ou non avec le rejet de l'hypothèse.

Le test de probabilité critique peut être unilatéral ou bilatéral. Les deux cas sont illustrés aux figures 5B et 5C, ci-après ; son argument logique est représenté dans l'encadré qui suit, sur le même modèle que celui de 2-3.1.

¹⁹ Lorsqu'il s'agit du *t* de Student, on peut trouver cette valeur au moyen de la fonction `LOI.STUDENT` du logiciel Excel (`TDIST` en anglais) :
 $\alpha = \text{LOI.STUDENT}(t.d.d.; 1 \text{ pour uni-}, 2 \text{ pour bilatéral}).$

FIGURE 5B – TEST DE PROBABILITÉ CRITIQUE (BILATÉRAL)

Test t de Student : distribution cumulative



Test t de Student : fonction de densité

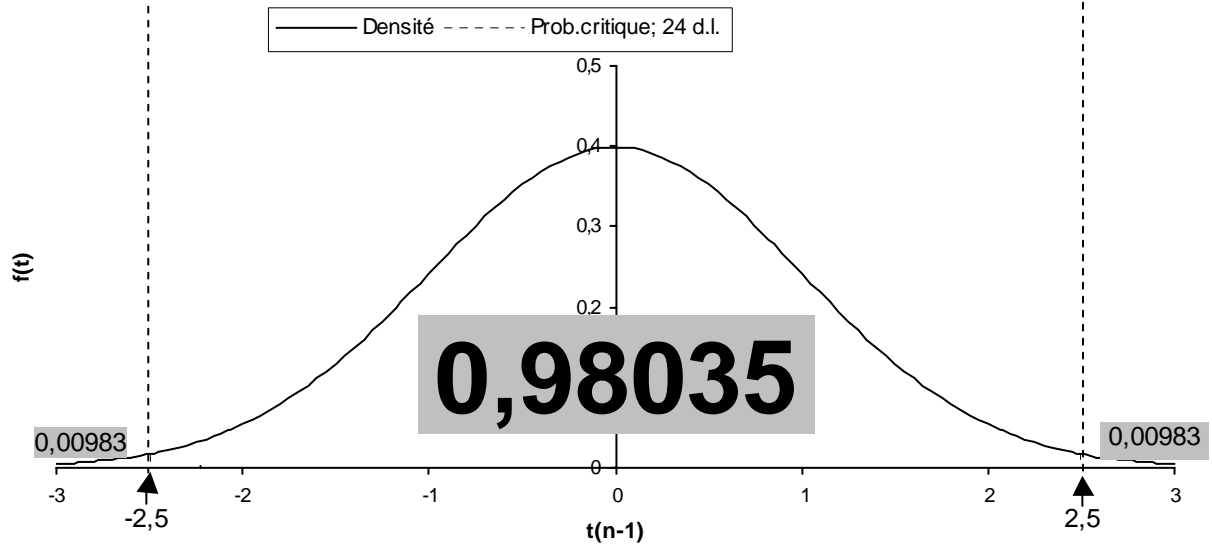
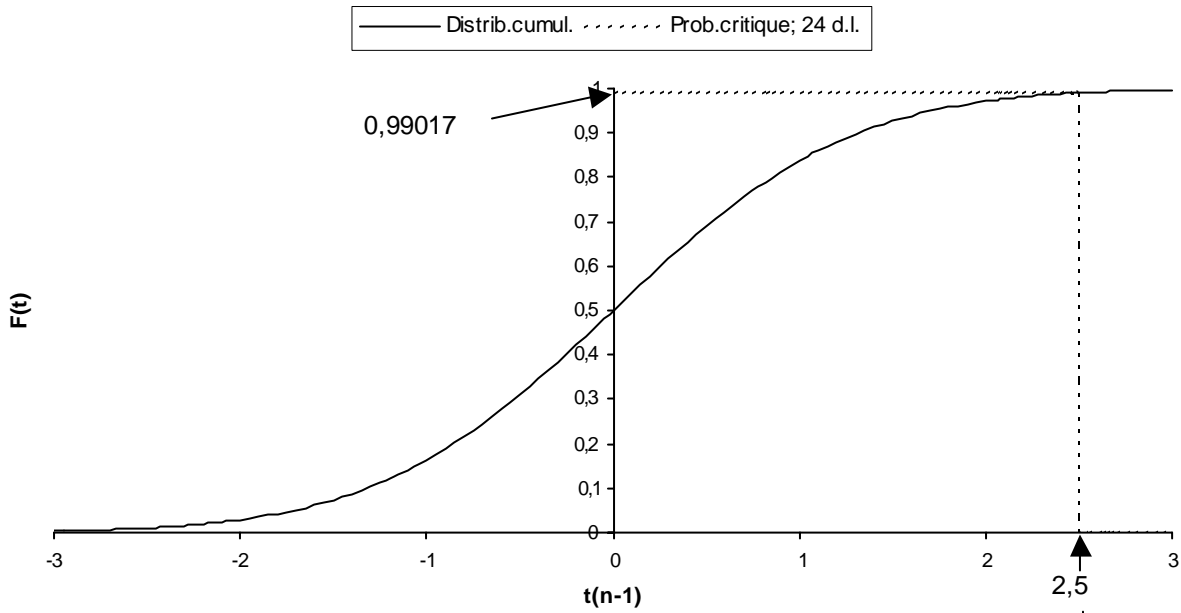
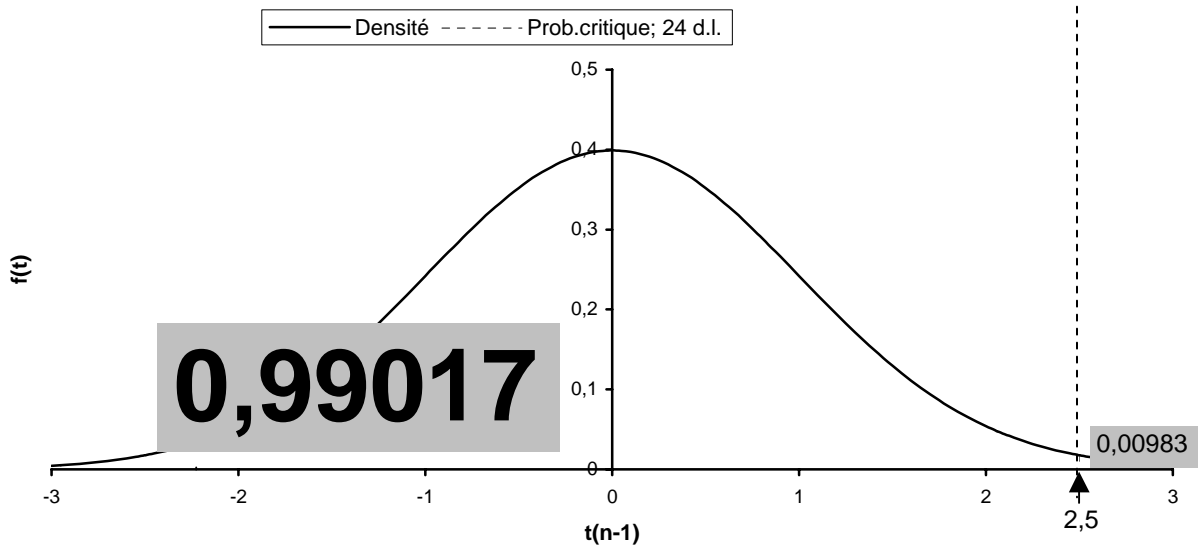


FIGURE 5C – TEST DE PROBABILITÉ CRITIQUE (UNILATÉRAL)

Test t de Student : distribution cumulative



Test t de Student : fonction de densité



ARGUMENT DU TEST DE PROBABILITÉ CRITIQUE (TEST D'HYPOTHÈSE SANS SEUIL DE SIGNIFICATION PRÉ-DÉFINI)

Modèle d'échantillonnage, hypothèse et implications (syllogisme)

Si {modèle d'échantillonnage},
alors {variable} a la distribution {distribution d'échantillonnage}.

Or,

si {hypothèse},
alors {variable} est égale à la statistique {variable-test}.

Donc,

si {modèle d'échantillonnage} **et si** {hypothèse},
alors {variable-test} a la distribution {distribution d'échantillonnage}.

Évaluation de la crédibilité de l'hypothèse

Étant donné

- la distribution {distribution d'échantillonnage}
- l'orientation du test (bilatéral, ou unilatéral, à gauche ou à droite),
l'ensemble de valeurs extrêmes dont la limite est donnée par la valeur observée de {variable-test} a une probabilité égale à {probabilité critique}.

Donc, si {hypothèse},

alors, la valeur observée de {variable-test} appartient à un ensemble de valeurs extrêmes dont la probabilité est égale à {probabilité critique}.

Conclusion

On décide que {probabilité critique} {est/n'est pas} suffisamment petite pour conclure que les observations sont probablement incompatibles avec {hypothèse}, et alors on {rejette/ne rejette pas} {hypothèse}.

2-3.6 Intervalles de confiance et marges d'erreur (estimation de la moyenne)

Réf. : Wonnacott et Wonnacott (1992, p. 286-296)

Dans l'exemple présenté en 2-3.2, nous avons rejeté l'hypothèse que $\mu_x = 100$, avec un seuil de signification de 5 %. Nous pourrions répéter le test avec d'autres hypothèses : $\mu_x = 101$, $\mu_x = 102$, ..., $\mu_x = 110$, etc. En testant toutes les valeurs possibles, nous pourrions faire un inventaire des hypothèses qui ne seraient *pas* rejetées (qui seraient « acceptables ») avec un seuil de signification de 5 %. L'ensemble des hypothèses qui ne seraient *pas* rejetées à un seuil de signification donné constitue un *intervalle de confiance*²⁰.

Idée-clé No18 :

L'ensemble des hypothèses qui ne seraient *pas* rejetées à un seuil de signification donné constitue un *intervalle de confiance*.

Il y a cependant une façon plus directe d'arriver au même résultat. Nous savons que, pour chaque hypothèse possible de la forme $\mu_x = \gamma$, on aura la variable-test

$$t_{n-1} = \frac{m_x - \gamma}{\left(\frac{s_x}{\sqrt{n}} \right)}$$

c'est-à-dire, pour notre exemple,

$$t_{24} = \frac{110 - \gamma}{\left(\frac{20}{\sqrt{5}} \right)}$$

On rejettera toutes les hypothèses pour lesquelles

$$t_{n-1} < -\theta_{n-1}(\alpha) \text{ ou } t_{n-1} > +\theta_{n-1}(\alpha)$$

c'est-à-dire, pour notre exemple,

$$t_{24} < -2,064 \text{ ou } t_{24} > +2,064$$

L'ensemble des hypothèses qui ne seraient ***pas rejetées*** à un seuil de signification de α est donc défini par

$$-\theta_{n-1}(\alpha) < t_{n-1} < +\theta_{n-1}(\alpha)$$

c'est-à-dire, pour notre exemple,

$$-2,064 < t_{24} < +2,064$$

En remplaçant t_{n-1} , on obtient

$$-\theta_{n-1}(\alpha) < \frac{m_x - \gamma}{\left(\frac{s_x}{\sqrt{n}}\right)} < +\theta_{n-1}(\alpha)$$

C'est-à-dire

$$-\theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}}\right) < (m_x - \gamma) < +\theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}}\right)$$

$$-m_x - \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}}\right) < -\gamma < -m_x + \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}}\right)$$

$$m_x + \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}}\right) > \gamma > m_x - \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}}\right)$$

$$m_x - \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}}\right) < \gamma < m_x + \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}}\right)$$

ou, pour notre exemple,

$$110 - 2,064 \left(\frac{20}{5}\right) < \gamma < 110 + 2,064 \left(\frac{20}{5}\right)$$

$$101,744 < \gamma < 118,256$$

Ainsi donc, pourvu que γ fasse partie de l'intervalle

$$\left[m_x - \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}}\right); m_x + \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}}\right) \right]$$

([101,744 ; 118,256] dans notre exemple), l'hypothèse $\mu_x = \gamma$ ne sera pas rejetée à un seuil de signification de α (5 %). Cela conduit naturellement à formuler l'hypothèse composée

$$C : \boxed{m_x - \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}}\right) < \mu_x < m_x + \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}}\right)}$$

c'est à dire, pour notre exemple, où $\theta_{n-1}(\alpha) = 2,064$,

$$110 - 2,064 (20/5) < \mu_x < 110 + 2,064 (20/5)$$

²⁰ La terminologie statistique traditionnelle distinguait l'estimation « ponctuelle » et l'estimation « par intervalle », cette dernière en référence aux intervalles de confiance.

Quelle est la probabilité que la condition C soit vraie ? En un certain sens, cette question n'a pas de sens, puisque les valeurs m_x , s_x et n sont données par l'échantillon, tandis que μ_x est inconnue, mais fixe : il n'y a rien d'aléatoire dans l'énoncé de la condition C ! Cependant, si nous nous plaçons en imagination *avant* que l'échantillon ne soit tiré ²¹, nous savons ceci : quelle que soit la valeur fixe mais inconnue de μ_x , il y a une probabilité de α (5 %, ou 0,05 dans notre exemple) que les valeurs de m_x et de s_x qui sortiront de l'échantillon violent la condition C. Autrement dit, *avant* que l'échantillon ne soit tiré, il y a une probabilité de $(1-\alpha)$ que la condition C soit respectée (95 % ou 0,95).

L'intervalle $\left[m_x - \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right); m_x + \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right) \right]$ s'appelle un *intervalle de confiance*,

dont le *niveau de confiance* est donné par

$$1 - \alpha = 1 - \text{seuil de signification}$$

(ici, $0,95 = 1 - 0,05$).

L'intervalle de confiance et le niveau de confiance sont indissociables. Parler d'un intervalle de confiance sans mentionner son niveau de confiance, c'est comme rapporter le résultat « partiel » d'un match sportif en annonçant le nombre de buts comptés par l'une des deux équipes, sans mentionner le nombre de buts comptés par l'autre...

Idée-clé No 19 :

L'intervalle de confiance et le niveau de confiance sont indissociables.

On définit pareillement la *marge d'erreur* : la marge d'erreur est d'ampleur de l'intervalle de part et d'autre de la moyenne observée qui définit l'intervalle de confiance. Ainsi, si on prend m_x comme valeur estimée de μ_x , on dira que la marge d'erreur est de $\pm \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right)$ avec un niveau de confiance de $1 - \alpha$ (dans notre exemple, $\pm 8,256$, avec un niveau de confiance de 95 % ou, comme on dit dans les comptes rendus journalistiques des sondages, « 19 fois sur 20 »). La marge d'erreur, comme l'intervalle de confiance, est dénuée de signification si elle n'est pas accompagnée de son niveau de confiance.

²¹ C'est-à-dire *avant* de connaître la valeur de m_x et celle de s_x .

Idée-clé No 20 :

La marge d'erreur est d'ampleur de l'intervalle de part et d'autre de la moyenne observée qui définit l'intervalle de confiance. La marge d'erreur, comme l'intervalle de confiance, est dénuée de signification si elle n'est pas accompagnée de son niveau de confiance.

Puisque μ_x est fixe, ce n'est pas une variable aléatoire et sa valeur n'est pas sujette à une distribution de probabilité. Il n'est donc pas rigoureusement exact d'affirmer que la valeur du paramètre est comprise dans l'intervalle de confiance avec une probabilité de 95 %. C'est d'ailleurs pourquoi la statistique utilise une formulation différente, en termes de « confiance » (probabilité subjective). Ce qui est exact en revanche, c'est que, *si l'on tirait des échantillons répétés de la même population*, l'écart entre la valeur estimée du paramètre et sa vraie valeur serait inférieur à la marge d'erreur²² dans 95 % des cas : c'est le sens du fameux « 19 fois sur 20 », qui signifie qu'en moyenne, sur 20 échantillons différents, il y en aurait 19 pour lesquels la marge d'erreur ne serait pas dépassée.

Il est à noter ici que le processus d'induction statistique nous a conduits à formuler un énoncé affirmatif, plutôt qu'un non-rejet. Mais cette affirmation, fort nuancée au demeurant, découle d'une logique de non-rejet : on affirme que, dans un ensemble donné d'hypothèses, il s'en trouve probablement une qui est vraie et on qualifie ce « probablement » au moyen d'une évaluation de la confiance qu'on accorde à l'affirmation.

L'encadré qui suit résume le développement qui conduit à définir un intervalle de confiance ou une marge d'erreur.

²² Marge d'erreur qui est cependant différente d'un échantillon à l'autre, puisque sa valeur dépend de l'écart type de l'échantillon s_x .

INTERVALLES DE CONFIANCE ET MARGES D'ERREUR

Formulation générale

$$t_{n-1} = \frac{m_x - \gamma}{\left(\frac{s_x}{\sqrt{n}} \right)}$$

Exemple :

$$n = 25 ; m_x = 110 ; s_x = 20 ; \alpha = 0,05$$

$$t_{24} = \frac{110 - \gamma}{\left(\frac{20}{\sqrt{25}} \right)}$$

Hypothèses rejetées à un seuil de signification de α (5 %)

$$t_{n-1} < -\theta_{n-1}(\alpha) \text{ ou } t_{n-1} > +\theta_{n-1}(\alpha)$$

$$t_{24} < -2,064 \text{ ou } t_{24} > 2,064$$

Hypothèses non rejetées à un seuil de signification de α (5 %)

$$-\theta_{n-1}(\alpha) < t_{n-1} < +\theta_{n-1}(\alpha)$$

$$-2,064 < t_{24} < +2,064$$

$$-\theta_{n-1}(\alpha) < \frac{m_x - \gamma}{\left(\frac{s_x}{\sqrt{n}} \right)} < +\theta_{n-1}(\alpha)$$

$$-\theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right) < (m_x - \gamma) < +\theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right)$$

$$-m_x - \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right) < -\gamma < -m_x + \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right)$$

$$m_x + \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right) > \gamma > m_x - \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right)$$

$$m_x - \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right) < \gamma < m_x + \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right)$$

$$110 - 2,064 \left(\frac{20}{5} \right) < \gamma < 110 + 2,064 \left(\frac{20}{5} \right)$$

$$110 - 8,256 < \gamma < 110 + 8,256$$

$$101,744 < \gamma < 118,256$$

Avant que l'échantillon ne soit tiré, quelle que soit la valeur de μ_x , il y a une probabilité de $(1-\alpha)$

(95 %) que soit respectée la condition C : $m_x - \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right) < \mu_x < m_x + \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right)$

L'intervalle de confiance...

$$\left[m_x - \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right) ; m_x + \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right) \right]$$

$$\left[110 - 2,064 \left(\frac{20}{5} \right) ; 110 + 2,064 \left(\frac{20}{5} \right) \right]$$

$$[101,744 ; 118,256]$$

... et son niveau de confiance

$$1 - \alpha = 1 - \text{seuil de signification}$$

$$0,95 = 1 - 0,05$$

Marge d'erreur avec un niveau de confiance de $(1-\alpha)$ (95 %)

$$\pm \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right)$$

$$\pm 8,256$$

Il y a deux conclusions à tirer de ce qui précède :

1. L'exemple de la moyenne montre clairement que, plus le niveau de confiance choisi est élevé, plus l'intervalle de confiance doit être large, et plus la marge d'erreur est grande : on gagne en confiance au détriment de la précision.
2. Cet exemple illustre aussi comment la précision des estimations dépend de la taille de l'échantillon. S'agissant d'estimer la moyenne, lorsque l'on augmente la taille de l'échantillon, la marge d'erreur diminue avec la *racine carrée* de la taille de l'échantillon ²³. Le gain de précision est moins que proportionnel à l'augmentation de la taille de l'échantillon : c'est, transposée au domaine de la statistique, la loi des rendements décroissants de l'économie...

Nous avons présenté les notions d'intervalle de confiance et de marge d'erreur dans le contexte de l'estimation de la moyenne d'une variable approximativement normale, à l'aide d'un échantillon aléatoire simple tiré d'une population de très grande taille. Bien sûr, ces notions se transposent à d'autres situations, où les conclusions que nous venons d'énoncer demeurent valides ²⁴.

2-3.7 Détermination de la taille requise d'un échantillon (estimation de la moyenne)

Comment déterminer la taille de l'échantillon en fonction du niveau de précision désiré ? Dans le contexte de l'estimation d'une moyenne, on voit que la valeur de la marge d'erreur dépend de l'écart type de l'échantillon, s_x , de sorte qu'aucune formule ne permet de connaître le degré de précision avant d'avoir tiré l'échantillon. Tout au plus peut-on déterminer la taille nécessaire pour que, dans le pire des cas, la marge d'erreur demeure acceptable. Et qu'entend-on par « le pire des cas » ? Le pire des cas est évidemment celui où, dans l'échantillon tiré, l'écart type est le plus grand. Voyons cela de plus près.

La marge d'erreur, avons-nous vu, est donnée par

$$\varepsilon = \pm t_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right)$$

²³ Il y a aussi un gain de précision qui résulte de l'accroissement du nombre de degrés de liberté associés au t de Student : dans la table, on peut voir que les valeurs critiques $t_{n-1}(\alpha)$ diminuent lorsque le nombre de degrés de liberté augmente. Mais à mesure que l'on approche de 30 degrés de liberté, les gains sont de plus en plus faibles.

²⁴ Si les principes demeurent valides dans d'autres situations, il importe de rappeler que la forme particulière des formules dépend du modèle d'échantillonnage que nous avons défini à la section 2-3.2.

Il s'agit de déterminer n , la taille de l'échantillon ; les tables statistiques nous donnent les valeurs de $\theta_{n-1}(\alpha)$; et s_x est inconnue tant que l'on a pas tiré l'échantillon. Voici les étapes à suivre :

1. Décider de la marge d'erreur acceptable ε .
2. Choisir le niveau de confiance désiré $(1 - \alpha)$.
3. Repérer dans la table les valeurs de $\theta_{n-1}(\alpha)$ pour les différentes tailles d'échantillon n .
4. Formuler à propos de s_x l'hypothèse du pire, c'est-à-dire la plus grande valeur de s_x que l'on puisse obtenir dans un échantillon.
5. Résoudre pour n l'équation suivante

$$\varepsilon = \pm \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right)$$

$$\sqrt{n} = \theta_{n-1}(\alpha) \left(\frac{s_x}{\varepsilon} \right)$$

$$n = \left[\theta_{n-1}(\alpha) \left(\frac{s_x}{\varepsilon} \right) \right]^2$$

Par exemple, supposons que la marge d'erreur acceptable soit de 10 ($\varepsilon = 10$), le niveau de confiance désiré de 90 % ($\alpha = 0,10$) et l'hypothèse du pire $s_x = 20$. On aurait alors

$$n = \left[\theta_{n-1}(\alpha) \left(\frac{s_x}{\varepsilon} \right) \right]^2 = \left[\theta_{n-1}(0,10) \left(\frac{20}{10} \right) \right]^2 = 4 \left[\theta_{n-1}(0,10) \right]^2$$

Puisque $\theta_{n-1}(\alpha)$ dépend de n , il s'agit d'une équation sous forme implicite. On peut la résoudre par approximations successives.

Solution par approximations successives

On peut amorcer le processus d'approximation en supposant un échantillon de très grande taille ($n \rightarrow \infty$). La table nous donne

$$\theta_{\infty}(0,10) = 1,645$$

Et alors

$$n_0 = \left[\theta_{n-1}(\alpha) \left(\frac{s_x}{\varepsilon} \right) \right]^2 = \left[\theta_{n-1}(0,10) \right]^2 = 4 \left[1,645 \right]^2 = 10,8$$

ce qui signifie que l'échantillon pourrait être plus petit que l'infini, tout en étant plus grand que 10,8. Par contre, avec $n = 11$ (11 est le premier nombre entier supérieur à 10,8),

$$\theta_{11-1}(0,10) = 1,812 \text{ et}$$

$$n_1 = 4[\theta_{10}(0,10)]^2 = 4[1,812]^2 = 13,1 > 11$$

ce qui signifie que l'échantillon doit être plus grand que 11. Avec $n = 13$ (13 est le premier nombre entier inférieur à 13,1),

$$\theta_{13-1}(0,10) = 1,782 \text{ et}$$

$$n_2 = 4[\theta_{12}(0,10)]^2 = 4[1,782]^2 = 12,7 < 13$$

Cela signifie que l'échantillon pourrait être plus petit. Mais avec $n = 12$,

$$\theta_{12-1}(0,10) = 1,796 \text{ et } 4[\theta_{11}(0,10)]^2 = 4 \times (1,796)^2 = 12,9 > 12$$

L'échantillon doit être plus grand.

Conclusion, puisque 12 est insuffisant, et que 13 est plus que suffisant pour obtenir la marge d'erreur désirée, il faut un échantillon de taille 13.

On peut vérifier le résultat en recalculant la marge d'erreur :

$$\varepsilon = \pm \theta_{n-1}(\alpha) \left(\frac{s_x}{\sqrt{n}} \right) = 1,782 \left(\frac{20}{\sqrt{13}} \right) = 9,885 < 10$$

Cas où la marge d'erreur acceptable est fixée en termes relatifs

En général, on s'intéresse davantage à la marge d'erreur en termes relatifs, c'est-à-dire en fraction de la moyenne estimée (fraction que l'on exprime le plus souvent comme un pourcentage) :

$$\frac{\varepsilon}{m_x} = \pm \theta_{n-1}(\alpha) \frac{\left(\frac{s_x}{m_x} \right)}{\sqrt{n}}$$

On voit que la méthode de détermination de la taille d'échantillon requise est essentiellement la même lorsqu'on veut fixer la marge d'erreur en pourcentage de la moyenne estimée. La seule différence est que l'hypothèse du pire est formulée par rapport au coefficient de variation s_x/m_x , plutôt que par rapport à l'écart type. L'avantage de cette approche est qu'elle permet la construction de tables d'usage général qui donnent la taille d'échantillon requise en fonction de la marge d'erreur relative acceptable et du coefficient de variation.

Cas où la moyenne recherchée est une proportion

Réf. : Wonnacott et Wonnacott (1992, p. 232-240 et 309-311)

Il n'est peut-être pas toujours facile de formuler l'hypothèse du pire à propos de l'écart type ou du coefficient de variation. Il y a cependant une classe de situations où on n'a pas besoin

d'hypothèse : cela arrive quand on veut estimer une proportion. Par exemple, on cherche à savoir quelle proportion d'une population est favorable à un certain projet d'aménagement. Un sondage est réalisé, et on définit une variable dichotomique représentant les réponses à la question sur le projet d'aménagement :

$x_i = 1$ si le répondant i est favorable

$x_i = 0$ si le répondant i n'est pas favorable

Dans ce cas,

$$m_x = \frac{\sum_i x_i}{n} = \frac{\text{Nombre de réponses favorables}}{\text{Nombre total de répondants}}$$

La moyenne m_x est donc la proportion de personnes favorables dans l'échantillon, qu'on représente habituellement par la lettre p (pour « proportion »), plutôt que par m_x . On veut estimer μ_x , la proportion de personnes favorables dans la population, avec une certaine marge d'erreur.

Pour déterminer la taille d'échantillon nécessaire, il faut spécifier ce que nous avons appelé l'hypothèse du pire. Or on peut montrer que, pour une variable dichotomique,

$$s_x^2 = p(1 - p)$$

où la valeur de p est forcément entre zéro et un. On peut aussi démontrer que, pour les valeurs de p comprises entre zéro et un, s_x atteint son maximum pour $p = 0,5$: on a alors $s_x^2 = 0,25$ et $s_x = 0,5$. Ainsi se trouve résolue la question de la spécification du plus grand écart type possible.

NOTE : Puisque la variable étudiée est une variable dichotomique, on ne peut prétendre qu'elle ait une distribution normale dans la population. Il s'ensuit que, strictement parlant, le test de Student et les procédures qui s'y rattachent ne s'appliquent pas au cas d'une proportion. Cependant, si l'on a une population de très grande taille et qu'on en tire un échantillon aléatoire simple, la statistique mathématique montre que le test de Student est approximativement valide, à la condition que μ_x ne soit pas trop éloignée de 0,5.

Il existe évidemment des tables statistiques qui donnent la taille d'échantillon nécessaire, en fonction de la marge d'erreur acceptable, pour différentes hypothèses sur la plus grande valeur possible de m_x (c'est-à-dire de p).

MISE EN GARDE : Il faut se garder de confondre l'erreur relative sur une moyenne qui n'est pas une proportion et l'erreur absolue sur une proportion. Par exemple, si j'estime qu'en moyenne, le 23 septembre 1998, les habitants de l'Île de Montréal ont écouté la radio pendant 120 minutes, avec une marge d'erreur de douze minutes (à un niveau de confiance de 95 %), j'ai une marge d'erreur relative de 10 % (12/120). Par ailleurs, si je dis que j'ai estimé à 80 % la proportion des Montréalais qui ont écouté la radio pendant au moins dix minutes le 23 septembre 1998, avec une marge d'erreur de plus ou moins 10 % (à un niveau de confiance de 95 %), il y a ambiguïté : sont-ce 10 % de 80 % ou 10 % tout court ? En d'autres mots, l'intervalle de confiance à 95 % s'étend-il de 72 % à 88 %, ou de 70 % à 90 % ? C'est généralement la seconde interprétation qui est la bonne, parce qu'il devient pas mal alambiqué de parler de pourcentage sur un pourcentage (on pourrait ajouter que si, effectivement, l'intervalle de confiance s'étendait de 72 % à 88 %, la maison de sondage aurait tout intérêt à proclamer que sa marge d'erreur est de 8 %, plutôt que de dire qu'elle est de 10 % de 80 %).

2-3.8 Autres tests fréquemment utilisés

Nous n'avons parlé jusqu'ici que du test de Student, et d'une seule utilisation, le test d'une hypothèse simple sur une moyenne. Il y a d'autres applications du test de Student. Par exemple, lorsqu'on compare deux échantillons, on peut se servir du test de Student pour tester l'hypothèse que les deux moyennes sont égales²⁵. Si l'on rejette l'hypothèse que les deux moyennes sont égales, on rejette par le fait même l'hypothèse que les deux échantillons proviennent de la même population.

Soit

$$H_0 : \mu_1 - \mu_2 = \delta$$

Dans le cas particulier où les deux échantillons sont de même taille n , on a

$$t_{2(n-1)} = \frac{m_1 - m_2 - \delta}{\left(\frac{s_1 + s_2}{\sqrt{n}} \right)}$$

En général, si les deux échantillons sont de taille n_1 et n_2 respectivement,

²⁵ Wonnacott et Wonnacott (1992, p. 299-307).

$$t_{2(n-1)} = \frac{m_1 - m_2 - \delta}{\left(\frac{\sqrt{\frac{(n_1 - 1)s_1 + (n_2 - 1)s_2}{n_1 + n_2 - 2}}}{\sqrt{\frac{n_1 + n_2}{n_1 n_2}}} \right)}$$

Un autre test fréquemment utilisé est le test du χ^2 (*chi-carré* ou *Khi-deux*). Il est notamment appliqué à l'hypothèse d'indépendance dans les tableaux de contingence (*Khi-deux* de Pearson : voir le chapitre 4-1). On peut aussi s'en servir pour tester une hypothèse simple sur une variance. En effet, dans un échantillon aléatoire simple tiré d'une population normale de grande taille, la variable

$$\frac{s^2}{\left(\frac{\sigma^2}{n-1} \right)}$$

a la distribution du χ^2 avec $n-1$ degrés de liberté²⁶. L'encadré qui suit donne la « valeur » qu'il faut attribuer à chaque « variable », dans l'argument du test d'hypothèse classique, pour appliquer l'argument au test d'une hypothèse simple sur un écart type (voir l'encadré inséré à la fin de 2-3.1).

Mentionnons aussi le test F de Fisher, dont il sera question à propos de l'analyse de régression. La distribution F de Fisher dépend de deux paramètres : le nombre de degrés de liberté du numérateur et le nombre de degrés de liberté du dénominateur (la signification de ces expressions s'éclaircira quelque peu dans le contexte des tests F sur les régressions). On peut s'en servir notamment pour tester un coefficient de corrélation simple. Dans l'hypothèse où le « vrai » coefficient de corrélation $\rho = 0$, la variable-test

$$\frac{r^2}{\left[\frac{(1-r^2)}{(n-2)} \right]} = (n-2) \frac{r^2}{1-r^2}$$

a la distribution du F de Fisher avec 1 degré de liberté au numérateur et $(n-2)$ degrés de liberté au dénominateur. Dans cette expression, r est le coefficient de corrélation de l'échantillon :

$$r = \frac{s_{xy}}{s_x s_y}.$$

²⁶ Wonnacott et Wonnacott (1992, chap. 17).

APPLICATION DE L'ARGUMENT AU TEST D'UNE HYPOTHÈSE SIMPLE SUR UN ÉCART-TYPE

Formulation générale	Exemple : $n = 20 ; m_x = 110 ; s_x = 65 ; \alpha = 0,05$
{hypothèse}	
$H_0 : \sigma_x = \gamma$	$H_0 : \sigma_x = 70$
{modèle d'échantillonnage}	
<ul style="list-style-type: none"> • Dans la population, la variable x a une distribution (approximativement) normale, avec une moyenne μ_x et un écart-type σ_x inconnus. • La population est de très grande taille et il en a été tiré un échantillon aléatoire simple de taille... 	
n	20
{variable}	
$\frac{s_x^2}{\left(\frac{\sigma_x^2}{n-1}\right)}$	$\frac{65^2}{\left(\frac{\sigma_x^2}{19}\right)}$
{distribution d'échantillonnage}	
Distribution du χ^2 avec...	
$n-1$ degrés de liberté	19 degrés de liberté
{variable-test}	
$\chi_{n-1}^2 = \frac{s_x^2}{\left(\frac{\gamma^2}{n-1}\right)}$	$\chi_{19}^2 = \frac{65^2}{\left(\frac{70^2}{(20-1)}\right)} = 16,38$
{seuil de signification}	
α	0,05
Orientation du test et {zone de rejet}	
test unilatéral à droite ($H_A : \sigma_x > \gamma$) :	test unilatéral à droite ($H_A : \sigma_x > 70$) :
$\chi_{n-1}^2 > \chi_{n-1}^2(\alpha)$	$\chi_{19}^2 > 30,144$
OU	
test unilatéral à gauche ($H_A : \sigma_x < \gamma$) :	test unilatéral à gauche ($H_A : \sigma_x < 70$) :
$\chi_{n-1}^2 < \chi_{n-1}^2(1 - \alpha)$	$\chi_{19}^2 < 10,117$
OU	
{seuil de signification}	
2α	0,10
test bilatéral asymétrique ($H_A : \sigma_x \neq \gamma$) :	test bilatéral asymétrique ($H_A : \sigma_x \neq 70$) :
$\chi_{n-1}^2 < \chi_{n-1}^2(1 - \alpha)$ ou $\chi_{n-1}^2 > \chi_{n-1}^2(\alpha)$	$\chi_{19}^2 < 10,117$ ou $\chi_{19}^2 > 30,144$

APPLICATION DE L'ARGUMENT AU TEST DU KHI-2 DE PEARSON D'INDÉPENDANCE DANS UN TABLEAU DE CONTINGENCE

Formulation générale	Exemple : 2 col.; 5 lignes; $\alpha = 0,05$²⁷
{hypothèse}	
H_0 : Indépendance entre les deux variables catégoriques d'un tableau à deux dimensions	H_0 : Indépendance entre sexe et motif de présence à Place Versailles
{modèle d'échantillonnage}	
<ul style="list-style-type: none"> Le tableau de contingence est le résultat d'un processus de Poisson ou d'une distribution multinomiale 	
{variable}	
$\chi^2 = \sum_i \sum_j \frac{(x_{ij} - x_{ij}^*)^2}{x_{ij}^*}$ où x_{ij}^* dépend de H_0	$\chi^2 = \sum_i \sum_j \frac{(x_{ij} - x_{ij}^*)^2}{x_{ij}^*}$ où x_{ij}^* dépend de H_0
{distribution d'échantillonnage}	
Distribution asymptotique du χ^2 avec...	
(C - 1) (L - 1) degrés de liberté (tableau de C colonnes et lignes)	4 degrés de liberté
{variable-test}	
$\chi^2 = \sum_i \sum_j \frac{(x_{ij} - x_{ij}^*)^2}{x_{ij}^*}$ avec $x_{ij}^* = \frac{x_{i\bullet} \cdot x_{\bullet j}}{x_{\bullet\bullet}}$	Khi-2 = 6,111436
{seuil de signification}	
α	0,05
Orientation du test et {zone de rejet}	
test unilatéral à droite (H_A : dépendance) : $\chi^2 > \chi^2_{n-1}(\alpha)$	test unilatéral à droite (H_A : dépendance) : $\chi^2 > 9,488$

²⁷ Voir 4-1 Acétates du cours du 21 janvier 2003.pdf, dernière page.

TABLE DES VALEURS CRITIQUES DU TEST DE STUDENT (TEST BILATÉRAL)

Degrés de liberté	Probabilité α			
	0,10	0,05	0,02	0,01
1	6,314	12,706	31,821	63,656
2	2,920	4,303	6,965	9,925
3	2,353	3,182	4,541	5,841
4	2,132	2,776	3,747	4,604
5	2,015	2,571	3,365	4,032
6	1,943	2,447	3,143	3,707
7	1,895	2,365	2,998	3,499
8	1,860	2,306	2,896	3,355
9	1,833	2,262	2,821	3,250
10	1,812	2,228	2,764	3,169
11	1,796	2,201	2,718	3,106
12	1,782	2,179	2,681	3,055
13	1,771	2,160	2,650	3,012
14	1,761	2,145	2,624	2,977
15	1,753	2,131	2,602	2,947
16	1,746	2,120	2,583	2,921
17	1,740	2,110	2,567	2,898
18	1,734	2,101	2,552	2,878
19	1,729	2,093	2,539	2,861
20	1,725	2,086	2,528	2,845
21	1,721	2,080	2,518	2,831
22	1,717	2,074	2,508	2,819
23	1,714	2,069	2,500	2,807
24	1,711	2,064	2,492	2,797
25	1,708	2,060	2,485	2,787
26	1,706	2,056	2,479	2,779
27	1,703	2,052	2,473	2,771
28	1,701	2,048	2,467	2,763
29	1,699	2,045	2,462	2,756
30	1,697	2,042	2,457	2,750
40	1,684	2,021	2,423	2,704
50	1,676	2,009	2,403	2,678
60	1,671	2,000	2,390	2,660
70	1,667	1,994	2,381	2,648
80	1,664	1,990	2,374	2,639
90	1,662	1,987	2,368	2,632
100	1,660	1,984	2,364	2,626
∞	1,645	1,960	2,326	2,576

Source : valeurs calculées à l'aide de la fonction `LOI.STUDENT.INVERSE` (TINV en anglais)
 du logiciel Excel : $t = \text{LOI.STUDENT.INVERSE}(\alpha; d.d.)$