partie 5: LES TESTS D'HYPOTHESE

I - Généralités et définitions

Un test d'hypothèse consiste à définir **une règle de décision** concernant la validité d'une hypothèse portant sur la valeur d'une caractéristique (moyenne, variance, proportion…) d'une distribution dans une population dont on observe un échantillon aléatoire.

La procédure générale d'un test comprend les éléments suivants:

1. définir la variable aléatoire étudiée, sa loi, et les paramètres connus de la loi,
2. définir l'hypothèse nulle Ho et son alternative H1 ; un test statistique est par nature négatif,
3. fixer à priori une valeur α pour le risque de refuser Ho alors qu'elle serait vraie ; accepter Ho ne signifie pas que cette hypothèse est vraie, mais seulement que les observations disponibles n'ont pas permis de mettre en évidence qu'elle n'était pas vraie,
4. vérifier que la constitution de l'échantillon conduit à des observations aléatoires et indépendantes,
5. définir un critère statistique dont on connaît la loi quand Ho est vraie,
6. définir une région critique de rejet de H_0 telle que α soit la probabilité, si Ho est vraie, pour que le critère statistique appartienne à cette région,
7. énoncer la règle de décision correspondant à la valeur numérique prise par le critère statistique, à savoir rejeter Ho si ce critère est dans la région critique de rejet, ou accepter Ho si le critère est dans la région d'acceptation, région complémentaire de la précédente.

Rejeter Ho alors que celle-ci est vraie se fait avec un risque de 1ère espèce α.
Accepter Ho alors que celle-ci est fausse se fait avec un risque de 2ème espèce β.

Attention ! On accepte H_0 parce que l'on n'a pas pu mettre en évidence que H_1 soit vraie !

α **: risque du fabricant**, c'est le risque que prend le fabricant en décidant que sa fabrication n'est pas aux normes alors qu 'en fait elle l'est. Il va donc intervenir à tort sur son procédé de fabrication qu'il pense être mal réglé.
β **: risque du client**, c'est l'erreur que commet le fabricant en déclarant son procédé aux normes alors qu'en fait il ne l'est pas. Le client achètera donc un produit non conforme.

Lequel des 2 risques doit-on privilégier ? A vous de juger !.. Qu'en pensez-vous si on considère que α est le risque de déclarer coupable quelqu'un d'innocent et si β est le risque de déclarer innocent quelqu'un de coupable !?…

II - Choix entre deux paramètres

Les paramètres que nous étudierons sont la moyenne et la fréquence. Dans les 2 cas, le problème consistera à décider de quelle population est extrait un échantillon donné, le nombre de populations possibles étant réduit à 2.

II – 1 - Choix entre 2 moyennes

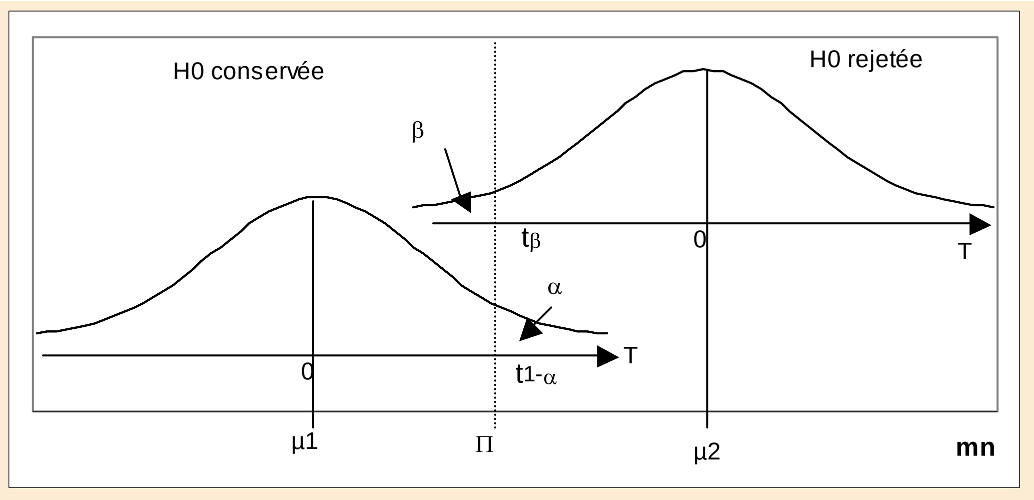
Soit un échantillon de taille n et de moyenne m. La population dont il est issu est-elle centrée sur $\mu = \mu_1$ (en général une norme ou un point référence de fabrication) ou sur $\mu = \mu_2$?

H_0 : $\mu_1 = \mu$ hypothèse rejetée avec un risque de **1ère espèce** α
H_1 : $\mu_2 = \mu$ hypothèse rejetée avec un risque de **2ème espèce** β

Deux cas peuvent alors se présenter :

1er cas : $\mu_1 < \mu_2$

$t_{1-\alpha} = \dfrac{\pi - \mu_1}{\sigma_1(m)}$ $\qquad$ $t_{\beta} = \dfrac{\pi - \mu_2}{\sigma_2(m)}$

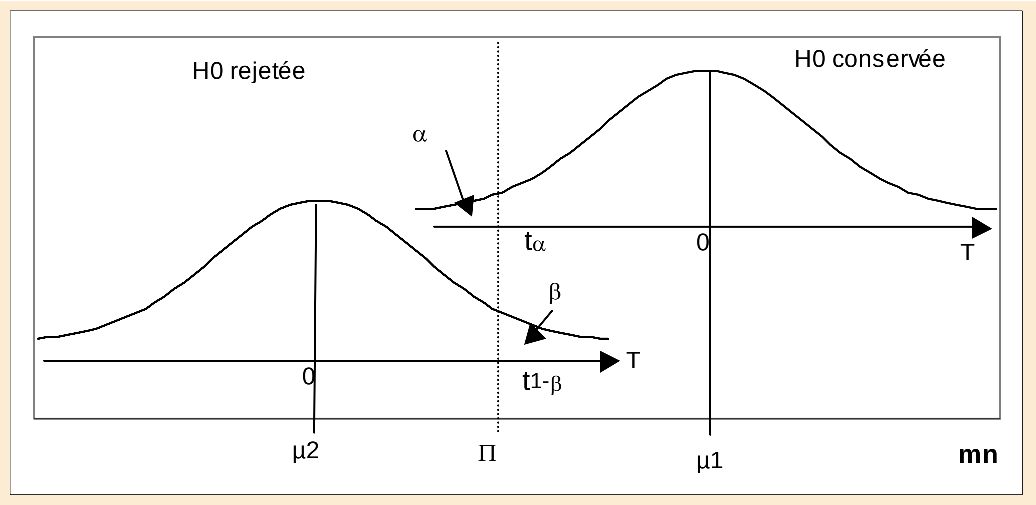


Π = valeur critique au dessus de laquelle on rejette H_0 avec α% de risque de le faire à tort.
Si la moyenne m d'un échantillon de taille n est supérieure à Π alors on est dans la zone de rejet de H_0 et dans celle d'acceptation de H_1. On conclura que l'on a pu mettre en évidence que l'échantillon est issu d'une population centrée sur μ_2 avec α% de risque de se tromper en disant cela.

Si la moyenne m d'un échantillon de taille n est inférieure à Π alors on est dans la zone de rejet de H_1 et dans celle d'acceptation de H_0. On conclura que l'on n'a pas pu mettre en évidence que l'échantillon est issu d'une population centrée sur μ_2 , on conserve l'idée que l'échantillon est issu d'une population centrée sur μ_1 avec β% de risque de se tromper en disant cela.

2ème cas : μ_1 > μ_2

$t_{\alpha} = \dfrac{\pi - \mu_1}{\sigma_1(m)}$ $\qquad$ $t_{1-\beta} = \dfrac{\pi - \mu_2}{\sigma_2(m)}$



Π = valeur critique au-dessous de laquelle on rejette H_0 avec α% de risque de le faire à tort.

Si la moyenne m d'un échantillon de taille n est supérieure à Π alors on est dans la zone de rejet de H_1 et dans celle d'acceptation de H_0. On conclura que l'on n'a pas pu mettre en évidence que l'échantillon est issu d'une population centrée sur μ_2 , on conserve l'idée que l'échantillon est issu d'une population centrée sur μ_1 avec β% de risque de se tromper en disant cela.
Si la moyenne m d'un échantillon de taille n est inférieure à Π alors on est dans la zone de rejet de H_0 et dans celle d'acceptation de H_1. On conclura que l'on a pu mettre en évidence que l'échantillon est issu d'une population centrée sur μ_2 avec α% de risque de se tromper en disant cela.

II – 2 - Choix entre 2 fréquences

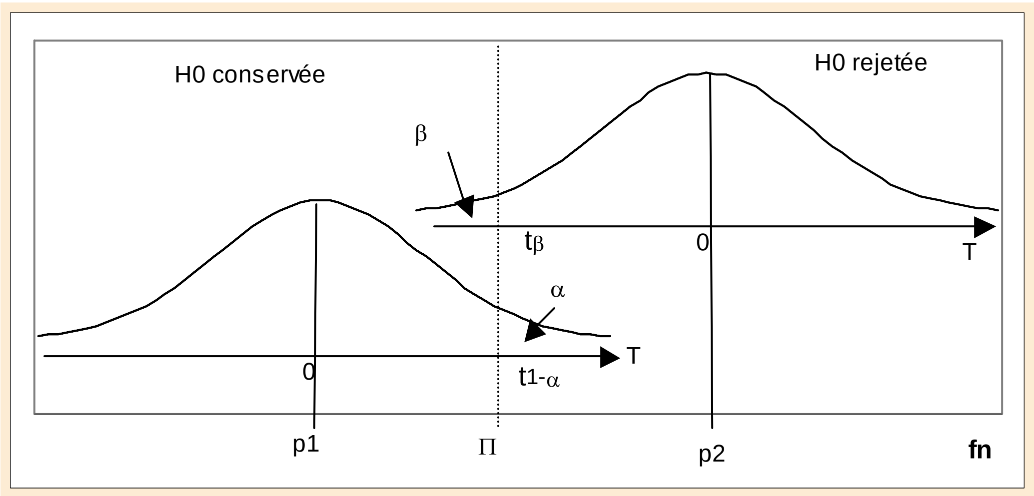
Soit un échantillon de taille n sur lequel on observe la fréquence f d'apparition d'un caractère. La population dont il est issu est-elle centrée sur $p = p_1$ (en général une proportion qui sert de référence à la fabrication) ou sur $p = p_2$?

H_0 : $p_1 = p$ hypothèse rejetée avec un risque de **1ère espèce** α
H_1 : $p_2 = p$ hypothèse rejetée avec un risque de **2ème espèce** β

Deux cas peuvent alors se présenter :

1er cas : $p_1 < p_2$

$$t_{1-\alpha} = \frac{\pi - p_1}{\sigma_1(f)} \qquad t_\beta = \frac{\pi - p_2}{\sigma_2(f)}$$

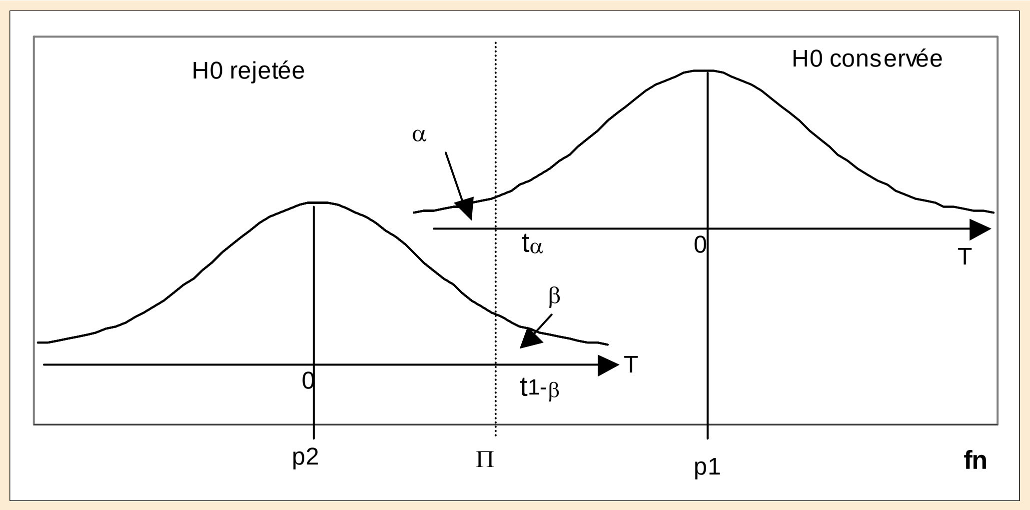


Π = valeur critique au dessus de laquelle on rejette H_0

Si la fréquence f dans un échantillon de taille n est supérieure à Π alors on est dans la zone de rejet de H_0 et dans celle d'acceptation de H_1. On conclura que l'on a pu mettre en évidence que l'échantillon est issu d'une population centrée sur p_2 avec α% de risque de se tromper en disant cela.
Si la fréquence f dans un échantillon de taille n est inférieure à Π alors on est dans la zone de rejet de H_1 et dans celle d'acceptation de H_0. On conclura que l'on n'a pas pu mettre en évidence que l'échantillon est issu d'une population centrée sur p_2 , on conserve l'idée que l'échantillon est issu d'une population centrée sur p_1 avec β% de risque de se tromper en disant cela.

$2^{ème}$ cas : $p_1 > p_2$

$$t_{\alpha} = \frac{\pi - p_1}{\sigma_1(f)} \qquad t_{1-\beta} = \frac{\pi - p_2}{\sigma_2(f)}$$



Π = valeur critique au dessous de laquelle on rejette H_0

Si la fréquence f dans un échantillon de taille n est inférieure à Π alors on est dans la zone de rejet de H_0 et dans celle d'acceptation de H_1. On conclura que l'on a pu mettre en évidence que l'échantillon est issu d'une population centrée sur p_2 avec α% de risque de se tromper en disant cela.
Si la fréquence f dans un échantillon de taille n est supérieure à Π alors on est dans la zone de rejet de H_1 et dans celle d'acceptation de H_0. On conclura que l'on n'a pas pu mettre en évidence que l'échantillon est issu d'une population centrée sur p_2 , on conserve l'idée que l'échantillon est issu d'une population centrée sur p_1 avec β% de risque de se tromper en disant cela.

III - Comparaison d'une moyenne à une valeur donnée

Soit un échantillon de taille n et de moyenne m. La population dont il est issu est-elle centrée sur μ = a ? On formule **l'hypothèse nulle** en disant que l'échantillon est bien issu d'une population centrée sur μ connue (μ = a), autrement dit m = μ , l'écart entre m et μ n'est dû qu'aux fluctuations d'échantillonnage de la moyenne ; m et μ ne sont pas statistiquement différentes.

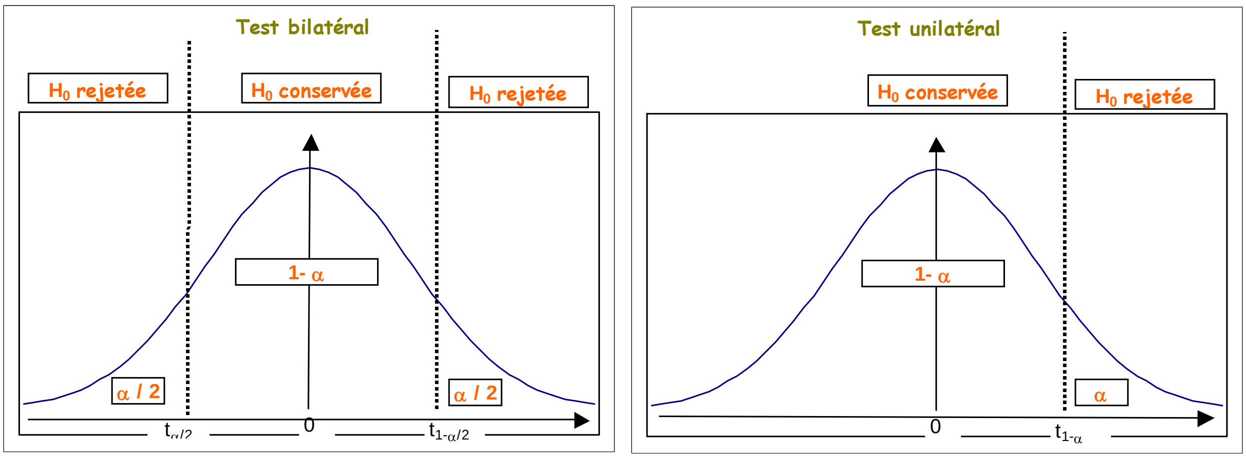
Au sens statistique une différence est non significative (noté NS) si on a un risque α supérieur à 5% de se tromper en rejetant H_0 l'absence de différence.

Au sens statistique une différence est significative (noté * ou S) si on a un risque α compris entre 1 et 5% de se tromper en rejetant H_0 l'absence de différence.

Au sens statistique une différence est très significative (noté ** ou TS) si on a un risque α compris entre 0.1 et 1% de se tromper en rejetant H_0 l'absence de différence.

Au sens statistique une différence est hautement significative (noté *** ou HS) si on a un risque α inférieur à 0.1% de se tromper en rejetant H_0 l'absence de différence.

Lorsque l'on cherche à mettre en évidence une différence on réalise un test bilatéral , si on cherche à prouver une supériorité (ou infériorité) on réalise un test unilatéral. Le choix de l'une ou l'autre des méthodes dépend de l'énoncé du problème et il se formalise dans l'hypothèse alternative H_1.



III - 1 - variance de la population connue

H_0 : μ = a
H_1 : $\mu \neq$ a (bilatéral) ou μ > a (ou μ < a) (unilatéral)

Variable aléatoire: m_n
Loi de la v.a: $L(m_n)=N\ (E(m);\sigma(m))$ si $L(X)=N\ (E(X);\sigma(x))$ <u>ou</u> si n > 30

Avec $$\sigma(m) = \frac{\sigma(x)}{\sqrt{n}}$$ (avec remise)

<u>ou</u> $$\sigma(m) = \frac{\sigma(x)}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$ (sans remise)

Critère statistique calculé : $t_{calc} = \frac{(m-\mu)}{\sigma(m)}$

Critère statistique théorique : $t_{1-\alpha/2}$ (bilatéral) ou $t_{1-\alpha}$ (unilatéral).

Conclusion : Si on rejette l'hypothèse nulle ($|t\ calc| > t$ théo) on a un risque égal à α% de se tromper.

III - 2 - variance de la population inconnue

H_0 : μ = a
H_1 : $\mu \neq$ a (bilatéral) ou μ > a (ou μ < a) (unilatéral)

Variable aléatoire: m_n
Loi de la v.a: $L(m_n)=S(\nu)\ (E(m);\sigma(m))$ si $L(X)=N\ (E(X);\hat{\sigma}(x))$ <u>et</u> si n < 30

Avec $$\sigma(m) = \frac{\hat{\sigma}(x)}{\sqrt{n}} = \frac{s(x)}{\sqrt{n-1}}$$ (avec remise) et **ν = n-1**

<u>ou</u> $$\sigma(m) = \frac{\hat{\sigma}(x)}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{s(x)}{\sqrt{n-1}}\sqrt{\frac{N-n}{N-1}}$$ (sans remise)

Critère statistique calculé : $t_{calc} = \frac{(m-\mu)}{\sigma(m)}$

Critère statistique théorique : $t_{1-\alpha/2}$ (ν)(bilatéral) ou $t_{1-\alpha}$ (ν)(unilatéral).

Si ddl >30 alors t est pris dans la table de la loi Normale.

Conclusion : Si on rejette l'hypothèse nulle ($|t\ calc| > t$ théo) on a un risque égal à α% de se tromper.

IV - Comparaison d'une proportion à une valeur donnée

Soit un échantillon de taille n sur lequel on observe la fréquence f d'apparition d'un caractère. La population dont il est issu est-elle centrée sur p = a ?

H_0 : p = a
H_1 : p $\neq$ a (bilatéral) ou p > a (ou p < a) (unilatéral)

Variable aléatoire: f_n
Si la loi de la v.a: $L(f_n) = N(E(f) ; \sigma(f))$

Avec $\sigma(f) = \sqrt{\frac{pq}{n}}$.(avec remise)

<u>Ou</u> $\sigma(f) = \sqrt{\frac{pq}{n}}\sqrt{\frac{N-n}{N-1}}$ (sans remise)

Critère statistique calculé : $t_{calc} = \frac{(f-p)}{\sigma(f)}$

Critère statistique théorique : t $_{1-\alpha/2}$ (bilatéral) ou t $_{1-\alpha}$ (unilatéral).

Conclusion : Si on rejette l'hypothèse nulle ($|t\ calc| > t$ théo) on a un risque égal à α% de se tromper.

V - Comparaison d'une variance à une valeur donnée

H_0 : $\sigma^2 = \sigma^2_0$
H_1 : $\sigma^2 \neq \sigma^2_0$ (bilatéral) ou $\sigma^2 > \sigma^2_0$ ($\sigma^2 < \sigma^2_0$) (unilatéral)

Variable aléatoire: SCE / σ^2_0
La variable aléatoire suit une loi du X^2 avec ν = n - 1

Critère statistique calculé : $X^2_{calc} = \frac{SCEx(\overline{x})}{\sigma^2_0}$

On accepte l'hypothèse nulle si :

$X^2_{\alpha/2}(\nu) < X^2_{calc} < X^2_{1-\alpha/2}(\nu)$ (bilatéral)
$X^2_{\alpha}(\nu) < X^2_{calc}$ (ou $X^2_{calc} < X^2_{1-\alpha}(\nu)$) (unilatéral)

Sinon on rejette l'hypothèse nulle avec un risque égal à α% de se tromper.

VI - Comparaison de 2 séries

VI – 1 - échantillons indépendants

(voir synthèse tableau qui suit)

Soient 2 échantillons de taille respectives n_1 et n_2 et de moyennes m_1 et m_2. La question est de savoir si les échantillons 1 et 2 sont bien issus de 2 populations respectivement centrées sur μ_1 et μ_2 avec $\mu_1 - \mu_2 = a$.

Remarque : si on veut savoir si les échantillons proviennent d'une même population cela sous entend que $\mu_1 - \mu_2 = 0$.

H_0 : $\mu_1 - \mu_2 = a$

H_1 : $\mu_1 - \mu_2 \neq a$ (bilatéral) ou $\mu_1 - \mu_2 > a$ (ou $\mu_1 - \mu_2 < a$) (unilatéral).

Variable aléatoire: Δm

Il faut considérer 3 cas pour connaître la loi suivie par la variable aléatoire :

. ***populations normales de variances connues.***

Loi de la v.a : $L(\Delta m) = N(\mu_1 - \mu_2 ; \sigma_d(m))$

$= N(0 ; \sigma_d(m))$ si $a = 0$

. ***lois et variances de populations inconnues; grands échantillons.***

Loi de la v.a : $L(\Delta m) = N(\mu_1 - \mu_2 ; \sigma_d(m))$

$= N(0 ; \sigma_d(m))$ si $a = 0$

. ***populations normales de variances inconnues; petits échantillons.***

Loi de la v.a : : $L(\Delta m) = S(\nu)\,(\mu_1 - \mu_2 ; \sigma_d(m))$ avec $\nu = n_1 + n_2 - 2$

$= S(\nu)\,(0 ; \sigma_d(m))$ si $a = 0$

Dans tous les cas:

Critère statistique calculé : $t_{calc} = \dfrac{(m_1 - m_2) - (\mu_1 - \mu_2)}{\sigma_d(m)}$

avec $\sigma_d(m) = \sqrt{\sigma_1^2(m) + \sigma_2^2(m)}$

Conclusion : Si on rejette l'hypothèse nulle ($|t_{calc}| > t_{théo}$) on a un risque égal à α% de se tromper.

Tableau de synthèse de comparaison de 2 moyennes d'échantillons indépendants, avec remise, pour a = 0

Populations:	μ_1 $\sigma^2{}_1(x)$	et	μ_2 $\sigma^2{}_2(x)$
Echantillons:	n_1 m_1 $s^2{}_1(x)$	et	n_2 m_2 $s^2{}_2(x)$

Hypothèses: H_0: $\mu_1 = \mu_2$ — Hypothèse nulle

H_1: $\mu_1 \neq \mu_2$ (test bilatéral) — Hypothèse alternative

ou H'_1: $\mu_1 < \mu_2$ (test unilatéral) ou $\mu_1 > \mu_2$

Populations normales de variances connues. Petits ou grands échantillons.

$$t\ calc = \frac{(m_1 - m_2)}{\sqrt{\frac{\sigma_1{}^2(x)}{n_1} + \frac{\sigma_2{}^2(x)}{n_2}}}$$

(avec remise)

Lois et variances inconnues dans les populations. Grands échantillons (n > 30).

$$t\ calc = \frac{(m_1 - m_2)}{\sqrt{\frac{\hat{\sigma}_1{}^2(x)}{n_1} + \frac{\hat{\sigma}_2{}^2(x)}{n_2}}} \quad \text{ou} \quad \frac{(m_1 - m_2)}{\sqrt{\frac{s_1{}^2(x)}{n_1 - 1} + \frac{s_2^2(x)}{n_2 - 1}}}$$

(avec remise)

t théorique = t $_{(1 - \alpha / 2)}$ dans la loi Normale (bilatéral)
t théorique = t $_{(1 - \alpha)}$ dans la loi Normale (unilatéral).

Populations normales de variances inconnues mais égales (à tester à l'aide d'un test F vu ultérieurement). Petits échantillons (n < 30).

$$t\ calc = \frac{(m_1 - m_2)}{\sqrt{\frac{\hat{\sigma}_1{}^2(x)}{n_1} + \frac{\hat{\sigma}_2{}^2(x)}{n_2}}}$$

avec $\hat{\sigma}^2(x) = \frac{n_1 s_1{}^2(x) + n_2 s_2{}^2(x)}{n_1 + n_2 - 2}$

(avec remise)

t théorique = t $_{(1 - \alpha / 2)}$ dans la loi Student (bilatéral)
ou t $_{(1 - \alpha)}$ (unilatéral) avec un ddl = $(n_1 + n_2 - 2)$
<u>remarque</u>: si ddl > 30 on utilise la loi normale

Dans tous les cas:
- si $|$ t calc $|$ < t théorique on accepte H_0.
- si $|$ t calc $|$ > t théorique on rejette H_0 avec α% de risque de le faire à tort.

VI – 2 - échantillons appariés

Des échantillons appariés sont des échantillons dont les éléments sont liés 2 à 2 par une relation limitant les risques d'incidence de facteurs susceptibles d'affecter la variable aléatoire à laquelle on s'intéresse, autre que celui dont on veut garder le contrôle.

unité	1	2	...	i	...	n
X	x_1	x_2	...	x_i	...	x_n
X'	x'_1	x'_2	...	x'_i	...	x'_n
d	$d_1 = x_1 - x'_1$	$d_2 = x_2 - x'_2$	...	$d_i = x_i - x'_i$	...	$d_n = x_n - x'_n$

On travaille donc sur la série des n différences $d_i = x_i - x'_i$ (on peut aussi faire la différence entre x'i et xi, ce qui est important c'est de la faire toujours dans le même sens , dans la série des di on aura des valeurs positives, négatives mais aussi nulles !! on les garde toutes !!!)

Dans l'échantillon on a n différences di de moyenne $\overline{d}$ et d'écart-type estimé $\hat{\sigma}$ (d). La question est de savoir si la moyenne des différences sur cette série correspond à une différence théorique égale à la valeur a.
Remarque : quand on veut vérifier l'existence d'un écart entre les 2 séries appariées on pose a = 0.

$H_0 : \delta = a$
$H_1 : \delta \neq a$ (bilatéral) ou $\delta < a$ (ou $\delta > a$) (unilatéral).

Variable aléatoire: $\overline{d}_n$
Loi de la v.a : $L(\overline{d}_n) = S(\nu)\ (\delta; \sigma(\overline{d}))$ avec ν =ddl = n-1
$= S(\nu)\ (0; \sigma(\overline{d}))$ si a = 0.

avec $\sigma(\overline{d}) = \dfrac{\hat{\sigma}_d}{\sqrt{n}}$ (avec remise)

ou $\sigma(\overline{d}) = \dfrac{\hat{\sigma}_d}{\sqrt{n}}\sqrt{\dfrac{N-n}{N-1}}$ (sans remise)

Critère statistique calculé : t calc = $\dfrac{\overline{d} - \delta}{\sigma(\overline{d})}$

Critère statistique théorique : t $_{1-\alpha/2}$ (ν) (bilatéral) ou t $_{1-\alpha}$ (ν) (unilatéral).

Conclusion : Si on rejette l'hypothèse nulle ($|t\ calc|$ > t théo) on a un risque égal à α% de se tromper.

VII - Comparaison de 2 proportions

Soit 2 populations comportant respectivement des proportions p_1 et p_2 inconnues d'unités possédant un caractère étudié ; Soit 2 échantillons d'effectif n_1 et n_2 supérieurs à 100, aléatoires et indépendants, extraits de chacune des populations. Les nombres d'individus possédant le caractère spécifié dans les échantillons sont notés k_1 et k_2.

H_0 : $p_1 = p_2$
H_1 : $p_1 \neq p_2$ (bilatéral) ou $p_1 <$ (ou $>$) p_2 (unilatéral).

Variable aléatoire: Δf
Loi de la v.a : $L(\Delta f) = N(0 ; \sigma_d(f))$

Avec $p_0 = \dfrac{k_1 + k_2}{n_1 + n_2}$ $\quad q_0 = 1 - p_0$

et $\sigma_d(f) = \sqrt{\sigma_1^2(f) + \sigma_2^2(f)} = \sqrt{\dfrac{p_0q_0}{n_1} + \dfrac{p_0q_0}{n_2}}$ (avec remise)

<u>ou</u> $\sigma_d(f) = \sqrt{\dfrac{p_0q_0}{n_1}\dfrac{N_1 - n_1}{N_1 - 1} + \dfrac{p_0q_0}{n_2}\dfrac{N_2 - n_2}{N_2 - 1}}$ (sans remise)

Critère statistique calculé : t calc = $\dfrac{(f_1 - f_2)}{\sigma d(f)}$

Critère statistique théorique : t $_{1-\alpha/2}$ (bilatéral) ou t $_{1-\alpha}$ (unilatéral).

Conclusion : Si on rejette l'hypothèse nulle ($|t \text{ calc}| > t$ théo) on a un risque égal à α% de se tromper.

VIII - Comparaison de 2 variances

Deux populations de variances σ_1^2 et σ_2^2 inconnues.
Deux échantillons aléatoires, tirés de façon indépendante dans chacune des populations d'où sont prélevées respectivement n_1 et n_2 unités indépendantes.
La distribution de la variable, dans chacune des populations, suit une loi Normale sinon l'effectif de chaque échantillon doit au moins être égal à 30.

H_0 : $\sigma_1^2 / \sigma_2^2 = 1$
H_1 : $\sigma_1^2 / \sigma_2^2 \neq 1$ (bilatéral) ou $\sigma_1^2 / \sigma_2^2 > 1$(unilatéral).

Variable aléatoire : $\hat{\sigma}_1^2 / \hat{\sigma}_2^2$ (on suppose que la numérotation des échantillons conduit à reporter au numérateur la plus forte des variances estimées).
Loi de la v.a : Loi F (ν_1 ; ν_2) avec $\nu_1 = n_1 - 1$ et $\nu_2 = n_2 - 1$

Critère statistique calculé : F calc = $\dfrac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$

Critère statistique théorique : F $_{1-\alpha/2}$ (bilatéral) ou F $_{1-\alpha}$ (unilatéral).

Conclusion : Si F calc < F théo on conserve H_0.

IX - Comparaison de plusieurs variances : test de Bartlett

Ce test est utilisé lorsque l'on doit vérifier l'homogénéité des variances intra groupes de plusieurs séries statistiques.

k populations de variances σ_1^2, σ_2^2,…σ_i^2,……σ_k^2 inconnues.
k échantillons aléatoires, tirés de façon indépendante dans chacune des population d'où sont prélevées respectivement n_1, n_2,…….n_k unités indépendantes.
La distribution de la variable, dans chacune des populations, suit une loi Normale et aucune des variances empiriques n'est nulle ni trop petite.

H_0 : les variances sont homogènes.
H_1 : au moins une variance est supérieure aux autres. (test unilatéral)

$$\chi^2_{calculé} = \left(\frac{2.3026}{C}\right) * \left[\nu * \log_{10}\hat{\sigma}^2 - \sum_{j=1}^{j=k}(\nu_j * \log_{10}\hat{\sigma}_j^2)\right]$$

$$C = 1 + \frac{1}{3(k-1)}\left[\sum_{j=1}^{j=k}\frac{1}{\nu_j} - \frac{1}{\nu}\right] \qquad \nu_j = n_j - 1 \qquad \nu = \sum_{j=1}^{j=k}\nu_j \qquad \hat{\sigma}_j^2 = \frac{SCE_{x_{ij}}(\bar{x}_j)}{n_j - 1}$$

$$\hat{\sigma}^2 = \frac{1}{\nu}\sum_{j=1}^{j=k}\nu_j\hat{\sigma}_j^2$$

Critère statistique théorique : $X^2_{1-\alpha}$ (k-1)

Conclusion : X^2 calc < X^2 théo on ne peut pas mettre en évidence qu'au moins une variance est supérieure aux autres, on garde l'hypothèse de l'homogénéité des variances.