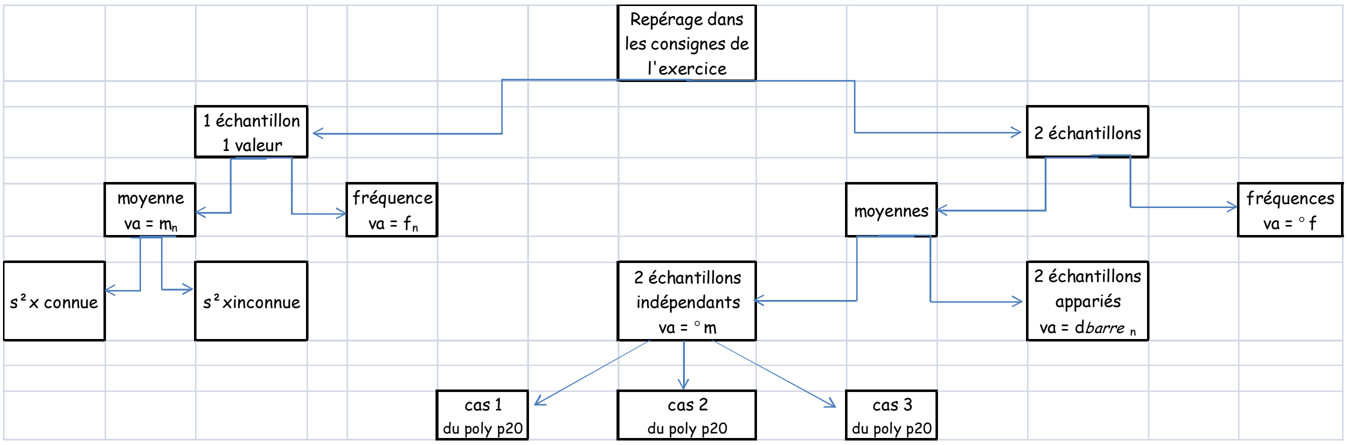
STATISTIQUE: 5EME PARTIE
CORRIGE D'EXERCICES

D'une manière générale quand une **v**ariable **a**léatoire de position notée **va** (moyenne m_n, fréquence f_n, différence de moyennes Δm, différence de fréquence Δf, moyenne des différences $\bar{d}_n$) est soumise à un test on calcule :

$$t_{calculé} = \frac{va - E(va)}{\sigma(va)}$$

Que l'on compare à un t $_{\text{théorique}}$ lu dans la table de la loi Normale ou dans la table de Student pour un risque α donné. Dans le cas où la va suit une loi Normale on peut, en plus, déterminer la p-value.

La méthode à suivre pour identifier le contexte de la question d'un exercice :



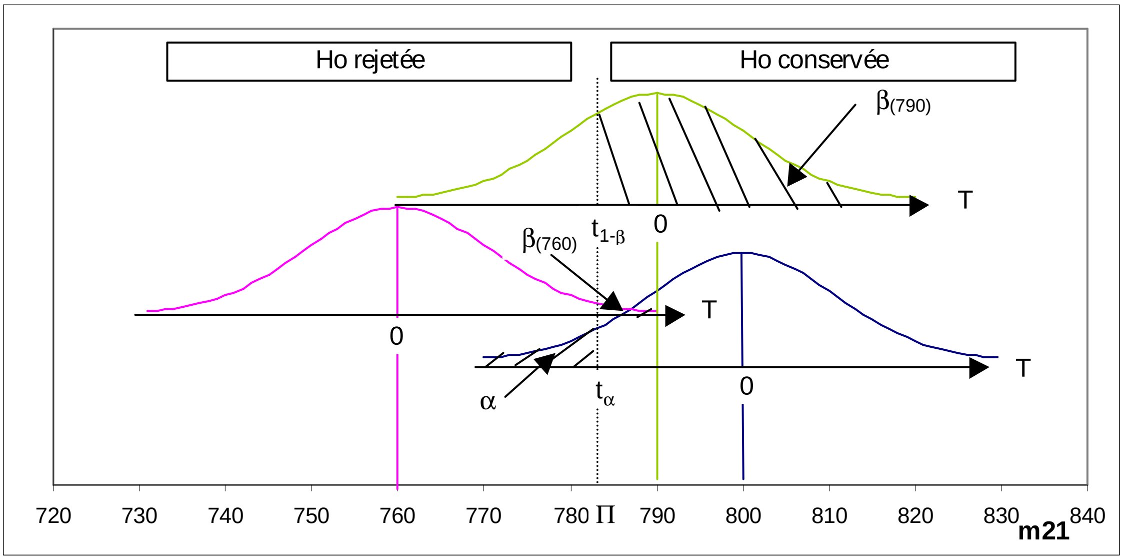
Exercice 1

Une entreprise fabrique des dispositifs électroniques dont la durée de vie suit une loi normale de moyenne 800 heures et d'écart-type 49 heures. Pour vérifier la qualité des dispositifs on prélève régulièrement un échantillon aléatoire de 21 dispositifs que l'on soumet à des essais de fiabilité. Le risque que prend le fabriquant de déclarer un lot non acceptable alors qu'il l'est est de 5%.
1) Calculer la région critique à ne pas dépasser pour considérer le lot acceptable.
2) Calculer le risque du client, à savoir accepter à tort un lot qui n'est pas aux normes si le processus de fabrication est centré sur une durée de vie moyenne de :
- 760 heures
- 790 heures
3) Le responsable du contrôle modifie son plan en prélevant 36 dispositifs de la production. Quelle règle de décision doit-il alors adopter?

X : durée de vie d'un dispositif en heures
$L(X) = N(\mu ; \sigma(x)) = N(800 ; 49)$
Echantillon n = 21
Variable aléatoire : m_{21}
Loi de m_{21} car L(X) = N et σ(x) connue.
$L(m_{21}) = N(E(m) ; \sigma(m))$ avec : $E(m) = \mu = 800$ et
N>>n donc avec remise, $\sigma^2(m) = 49^2 / 21$ et $\sigma(m) = 10.69$
Donc: $L(m_{21}) = N(800; 10.69)$

1°) α = risque de 1ère espèce, risque que prend le fabriquant de déclarer à tort que son lot est inacceptable ; c'est le risque de se tromper en rejetant Ho
Ho : lot acceptable $\mu = 800$
H1 : lot inacceptable $\mu < 800$
$\alpha = 0.05$ $t_\alpha = t_{0.05} = -1.645$
$t_\alpha = (\Pi - 800) / 10.69$ $\mathbf{\Pi = 782.4h}$

On considérera le lot acceptable si un échantillon de 21 dispositifs a une durée de vie moyenne supérieure à 782.4h.



2°)

Ho : $\mu = 800$
H1 : $\mu = 760$ lot inacceptable

Le risque de déclarer à tort que le lot est acceptable est le risque de 2ème espèce, risque du client noté β, c'est la probabilité de rejeter à tord l'hypothèse alternative H1.

$t_{1-\beta} = (782.4 - 760) / 10.69 = 2.097$ **$\beta = 0.0179$**

Le client à 1.79% de risque d'acheter un dispositif vendu comme pouvant durer en moyenne 800h alors qu'en réalité il durerait 760h en moyenne.

Ho : $\mu = 800$
H1 : $\mu = 790$ lot inacceptable

Le risque de déclarer à tort que le lot est acceptable est le risque de 2ème espèce, risque du client noté β, c'est la probabilité de rejeter à tord l'hypothèse alternative H1.

$t_{1-\beta} = (782.4 - 790) / 10.69 = -0.709$ **$\beta = 0.7611$**

Le client à 76.11% de risque d'acheter un dispositif vendu comme pouvant durer en moyenne 800h alors qu'en réalité il durerait 790h en moyenne.

3°)

X : durée de vie d'un dispositif en heures
$L(X) = N(\mu\ ; \sigma(x)) = N(800\ ; 49)$
Echantillon n = 36
Variable aléatoire : m_{36} durée moyenne de 36 dispositifs
Loi de m_{36} car $L(X) = N$ et $\sigma(x)$ connue mais plus simplement $n>30$.
$L(m_{36}) = N(E(m)\ ; \sigma(m))$ avec : $E(m) = \mu = 800$ et
$N>>n$ donc avec remise, $\sigma^2(m) = 49^2 / 36$ et $\sigma(m) = 8.167$
Donc: $L(m_{36}) = N(800; 8.167)$

α = risque de 1ère espèce, risque que prend le fabriquant de déclarer à tort que son lot est inacceptable ; c'est le risque de se tromper en rejetant Ho
Ho : lot acceptable $\mu = 800$
H1 : lot inacceptable $\mu < 800$

$\alpha = 0.05$ $t_{\alpha} = t_{0.05} = -1.645$

$t_{\alpha} = (\Pi - 800) / 8.167$ **$\Pi = 786.6h$**

On considérera le lot acceptable si un échantillon de 36 dispositifs a une durée de vie moyenne supérieure à 786.6h. En dessous de cette valeur le responsable du contrôle jugera son lot de fabrication non-conforme.

Exercice 2

Au cours de l'année écoulée, les ventes hebdomadaires du rayon parfumerie de " Casimouth " ont pu être approximées par une loi normale de moyenne 5000 et d'écart-type 600. Désirant développer ce rayon, " Casimouth " a procédé à un effort de promotion sur place (réaménagement du rayon, sélection des marques, démonstratrices…etc.) sans procéder à des campagnes publicitaires. Le but de la direction était d'obtenir un accroissement des ventes de 15% (avec l'idée de procéder à une action publicitaire si elle n'obtenait pas ce pourcentage d'augmentation). De plus elle a décidé de juger les résultats de son action en se basant sur les ventes des 9 premières semaines suivant la réorganisation du rayon. Sachant que la direction a fixé le risque de 1ère espèce à 2% (risque de ne pas procéder à l'action publicitaire alors qu'elle aurait été nécessaire) quelle aurait été votre conclusion si ce chiffre d'affaire enregistré sur ces 9 semaines avait été de 5400F ?

X: vente sur une semaine (hebdomadaire) $L(X) = N(5000; 600)$
Choix entre $\mu = 5000$ et $\mu = 5000 + 15/100*5000 = 5750$
$n = 9$ $\alpha = 0.02$ $m = 5400$

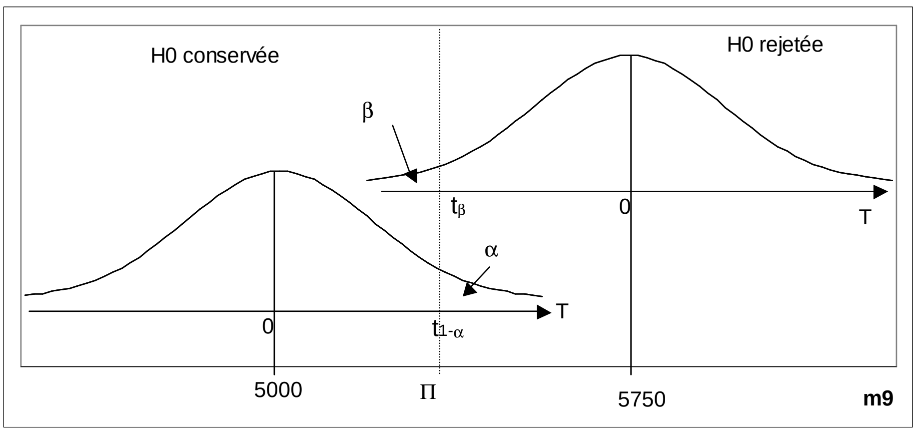
Variable aléatoire : m_9
Loi de m_9 :$L(m_9) = N(E(m)\,;\,\sigma(m))$:car $L(X) = N$ et $\sigma(x)$ connue.
Avec 1ère population: : $E(m) = \mu = 5000$ et
N>>n donc avec remise, $\sigma^2(m) = 600^2 / 9$ et $\sigma(m) = 200$
Donc: $L(m_9) = N(5000; 200)$
Avec 2ème population: : $E(m) = \mu = 5750$ et
N>>n donc avec remise, $\sigma^2(m) = 600^2 / 9$ et $\sigma(m) = 200$
Donc: $L(m_9) = N(5750; 200)$

α = risque de 1ère espèce, risque que prend le fabriquant de déclarer à tort que l'action publicitaire n'est pas nécessaire
Ho : vente insuffisante $\mu = 5000$
H1 : vente augmentée $\mu = 5750$ ($\mu > 5000$)

$\alpha = 0.02$ $t_{1-\alpha} = t_{0.98} = 2.054$ $t_{1-\alpha} = (\Pi - 5000) / 200$ $\mathbf{\Pi = 5410.74}$

Si les ventes sur les 9 semaines sont de 5400 alors on peut conclure que les ventes sont insuffisantes et qu'il faut engager une action publicitaire en plus de l'effort de promotion sur place. Mais on a β% de risque d'engager une action publicitaire alors qu'elle n'est pas nécessaire.
$t_\beta = (5410.74 - 5750) / 200 = -1.696$ ($t_\beta = -1.7$) $\boldsymbol{\beta = 0.0446}$

On a donc 4.46% de risque d'engager une action publicitaire inutile.



Exercice 3

Une culture soumise à des traitements curatif et préventif est habituellement contaminée par une maladie à raison d'un ratio de 10%. Pour des raisons de qualité de produit il est question de supprimer ces traitements si toutefois l'envahissement de la maladie n'atteint pas un taux moyen de 20%. L'étude porte sur un échantillon de 50 parcelles et le risque de première espèce (risque de juger à tort que le traitement est nécessaire) est fixé à 4%. En l'absence de tout traitement on a relevé un envahissement de 18% sur l'échantillon étudié.
1) Déterminer la région d'acceptation et celle de refus de la culture dans ces conditions
2) Calculer le risque de ne pas déceler à tort un envahissement de la maladie trop important.

Habituellement, avec traitement les cultures sont contaminées à raison d'un taux de 10% en moyenne. Ce taux est tolérable et doit rester tel quel. On veut juger des résultats que l'on aurait sans traitement en sachant qu'il serait intolérable d'atteindre un taux de 20%.

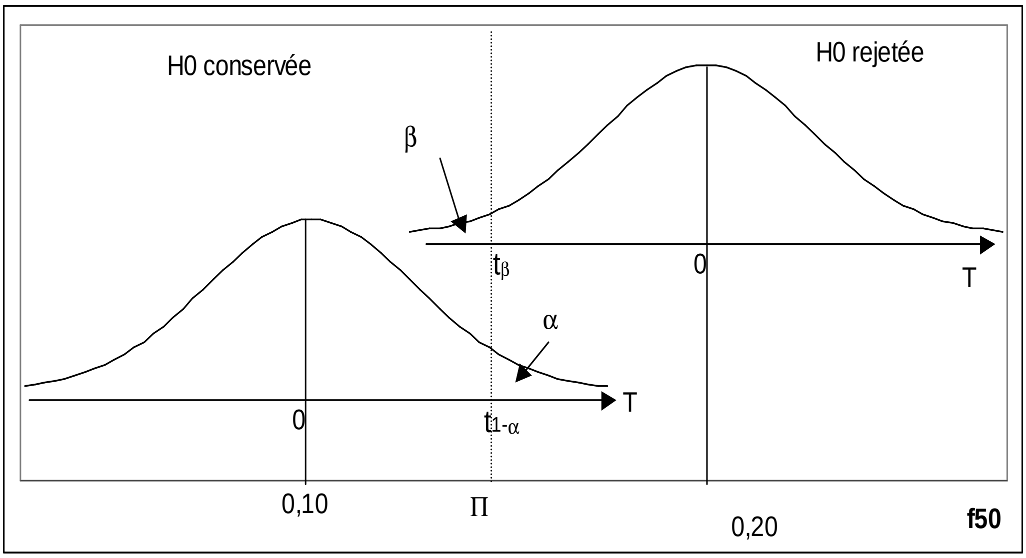
H0: $p = 0.10$ le traitement n'est pas nécessaire, si on rejette cette hypothèse on le fait avec un risque fixé à $\alpha = 0.04$ et dans ce cas on juge que le traitement est nécessaire avec 4% de risque de se tromper.
H1: $p = 0.20$ le traitement est nécessaire, la contamination est trop importante, si on rejette cette hypothèse on le fait avec un risque β et dans ce cas on juge que le traitement n'est pas nécessaire avec β% de risque de se tromper.

1°) $n = 50$ $p = 0.10$ $q = 0.9$ np et nq > 5
variable aléatoire: f_{50} taux de contamination de 50 parcelles en l'absence de traitement
$L(f_{50}) = N(E(f); \sigma(f))$ avec $E(f) = 0.10$ et $\sigma^2(f) = (0.1*0.9) / 50 = 0.0018$donc $\sigma(f) = 0.0424$

$L(f_{50}) = N(0.10; 0.0424)$
$\alpha = 0.04$ $t_{1-\alpha} = t_{0.96} = 1.751$ $t_{1-\alpha} = (\Pi - 0.10) / 0.0424$ **$\Pi = 0.1742$**

Si $f_{50} > \Pi$ alors on rejette Ho avec 4% de chance de se tromper, on décidera que le traitement est nécessaire
Si $f_{50} < \Pi$ alors on rejette H1 avec β % de chance de se tromper, on décidera que le traitement n'est pas nécessaire

Dans notre cas $f_{50} = 0.18 > 0.1742$ on décide que le traitement est nécessaire avec 4% de risque d'erreur.



2°) calculs de β

$L(f_{50}) = N(E(f); \sigma(f))$ avec $E(f) = 0.20$ et $\sigma^2(f) = (0.2*0.8) / 50 = 0.0032$donc $\sigma(f) = 0.0566$

$L(f_{50}) = N(0.20; 0.0566)$

$t_\beta = (0.1742 - 0.20) / 0.0566 = - 0.456$ $\beta = 1 - 0.6772 = 0.3228$ (32.28%)

Si on rejette H1 on le fait avec 32.28 % de chance de se tromper, dans ce cas on déciderait que le traitement n'est pas nécessaire.

Exercice 4

Une entreprise achète à un fournisseur 300 pièces d'une matière isolante dont la norme exige qu'elle ait une épaisseur de 6.34 mm. Cette entreprise veut vérifier si les normes sont respectées et sachant que la variance de la variable est de 0.25 elle prélève un échantillon de 20 pièces elle trouve une épaisseur moyenne de 6.50 mm .
1°) Que peut-on conclure sur le lot fourni au seuil de 5%?
2°) Même question si l'échantillon avait été de taille 40? Calculer le risque .

X : épaisseur d'une unité de matière isolante
$E(X) = \mu = 6.34$ $\sigma^2(x) = 0.25$ L(X) inconnue
1°) n = 20 m = 6.50
variable aléatoire m_{20}
$L(m_{20})$ = inconnue car L(X) inconnue et n < 30
On ne peut pas conclure sur ce lot.

2°) n = 40 m = 6.50
variable aléatoire m_{40}
$L(m_{40}) = N(E(m); \sigma(m))$ car n > 30
Paramètres de la loi: $E(m) = \mu = 6.34$
Taux de sondage : 100* 40/300 > 10% donc tirage sans remise
et $\sigma^2(m) = 0.25 / 40 *(300-40)/(300-1)$ d'où $\sigma(m) = 0.074$
$L(m_{40}) = N(6.34;0.074)$

H0: $\mu = 6.34$ (épaisseur aux normes)
H1: $\mu \neq 6.34$ (l'épaisseur n'est pas aux normes) test bilatéral

$t_{\text{calculé}} = (6.5 - 6.34) / 0.074 = 2.16$ *

1ère méthode: α donné = 0.05
$t_{\text{théorique}} = t_{1-\alpha/2} = t_{0.975} = 1.96$
$|t_{\text{calculé}}| > t_{\text{théorique}}$ On rejette Ho avec moins de 5% de chance de se tromper.

2ème méthode: α est à calculer
$\alpha_{\text{calculé}} = 2*P(T > 2.16) = 0.031$

Si on rejette Ho on a 3.1% de chance de se tromper. Ce risque est compris entre 1% et 5%, donc on conclue que l'on a pu mettre en évidence une différence significative (*) entre l'épaisseur de la matière isolante fabriquée et la norme.

Exercice 5

D'après des expériences antécédentes, il semble pour des individus de tranche d'âge comparable, que le temps moyen de réaction à un stimulus électrique soit de 76.26 millisecondes avec un écart-type de 2 millisecondes Parmi un groupe de 42 sujets de tranche d'âge supérieure on a trouvé un temps moyen de réponse à ce stimulus égal à 77 millisecondes . Peut-on conclure au seuil de 1% que ce groupe à un moins bon temps de réaction? Calculer le risque α.

X : temps de réaction d'un individu
$E(X) = \mu = 76.26$ $\sigma(x) = 2$ L(X) inconnue

$n = 42$ $m = 77$
variable aléatoire m_{42}
$L(m_{42}) = N(E(m); \sigma(m))$ car $n > 30$
Paramètres de la loi: $E(m) = \mu = 76.26$
N>>n donc tirage avec remise et $\sigma^2(m) = 4 / 42 = 0.0952$ d'où $\sigma(m) = 0.309$
$L(m_{42}) = N(76.26;0.309)$

H0: $\mu = 76.26$ (groupe comparable à la norme)
H1: $\mu > 76.26$ (moins bon temps de réaction) test unilatéral

$t_{\text{calculé}} = (77 - 76.26) / 0.309 = 2.4$ **

1ère méthode: α donné = 0.01
$t_{\text{théorique}} = t_{1-\alpha} = t_{0.99} = 2.326$
$|t_{\text{calculé}}| > t_{\text{théorique}}$ On rejette Ho avec moins de 1% de chance de se tromper.

2ème méthode: α est à calculer
$\alpha_{\text{calculé}} = P(T > 2.4) = 0.0082$

Si on rejette Ho on a 0.82% de chance de se tromper. Ce risque est compris entre 0.1% et 1%, donc on conclue que l'on a pu mettre en évidence que le groupe est très significativement (**) moins bon que la norme.

Exercice 6

Pour qu'un certain type de détecteur de fumée soit efficace il faut qu'il se déclenche en moyenne dans les 1.9 secondes qui suivent le début de l'incendie. Afin de vérifier une nouvelle marque un service de contrôle en teste 12 sur les 200 fabriqués, en les submergeant de fumée. Les temps de réponse sont les suivants :

1,4	1,2	1,0	2,2	2,0	1,2	1,6	1,7	1,8	1,5	1,9	2,0

Un autre service en contrôle 36 et obtient :
{ (1,4; 3); (1,2; 4); (2,0; 3); (1,7; 2); (2,2; 5); (1,9; 12); (1,8; 7) }

Pour chaque contrôle, pouvez-vous conclure que la nouvelle marque soit plus performante que ce que nécessite la norme au seuil de 2.5% ? Calculer le risque α si possible.

X : temps mit par un détecteur pour se déclencher

$E(X) = \mu = 1.9$ $\sigma(x)$ inconnu $L(X) = ?$ $N = 100$

1er cas:
Impossible, L(m) inconnue (L(X) inconnue et n<30)

2ème cas:
$n = 36$ $m = 1.8$ $\hat{\sigma}(x) = 0.2918$
variable aléatoire m_{36}
$L(m_{36}) = N(E(m); \sigma(m))$ car $n > 30$
Paramètres de la loi: $E(m) = \mu = 1.9$
Taux de sondage = 100* 36/200 = 18% (> 10%) donc tirage sans remise.
$\sigma^2(m) = \hat{\sigma}^2(x) * ((200 - 36) / (200-1)) / 36 = 0.00195$
$L(m_{36}) = N(1.9; 0.0441)$
Remarque : on peut noter $\sigma^2(m)$ ou $\hat{\sigma}^2(m)$, l'important est surtout de faire la différence entre $\sigma^2(x)$ (variance connue dans la population) et $\hat{\sigma}^2(x)$ (variance dans la population estimée = SCEx / (n-1) à partir des n données de l'échantillon)

H0: $\mu = 1.9$ (même performance)
H1: $\mu < 1.9$ (plus performant) test unilatéral

$t_{\text{calculé}} = (1.8 - 1.9) / 0.0441 = -2.27$ *

1ère méthode: α donné = 0.025
$t_{\text{théorique}} = t_{1-\alpha} = t_{0.975} = 1.96$
$|t_{\text{calculé}}| > t_{\text{théorique}}$ On rejette Ho avec moins de 2.5% de chance de se tromper.

2ème méthode: α est à calculer
$\alpha_{\text{calculé}} = P(T > 2.27) = 1-0.9884 = 0.012$

Si on rejette Ho on a 1.2% de chance de se tromper. Ce risque est compris entre 1% et 5%, donc on conclue que l'on a pu mettre en évidence que le nouveau type de détecteur est significativement (*) plus performant que l'ancien.

Exercice 7

Pour une tranche d'âge donnée la fréquence cardiaque des individus doit être en moyenne de 72 battements par minute. On supposera que la variable aléatoire suit une loi normale.
1) On a isolé un groupe de 12 individus et l'on souhaite savoir si ce groupe représente la population centrée sur 72. Les résultats suivants ont été obtenus :
88; 68; 72; 62; 80; 60; 72; 72; 72; 75; 72; 75
Que concluez-vous au seuil 4% ?
2) L'étude a porté sur un groupe de 53 individus où la fréquence cardiaque moyenne a été de 74 et la variance de 81. Pensez-vous que ce groupe représente une population centrée sur 72 au seuil de 4% ?

X : nombre de battements par minute d'un individu
$E(X) = \mu = 72$ $L(X) = N(72; \sigma(x))$

1°) $n = 12$ $m = 72.33$ $\hat{\sigma}(x) = 7.365$

variable aléatoire m_{12}
$L(m_{12}) = S(11)\,(E(m);\ \sigma(m))$ car $L(x) = N$ mais $\sigma(x)$ inconnue et $n < 30$
Paramètres de la loi: $E(m) = \mu = 72$

N>>n donc tirage avec remise et $\sigma^2(m) = \hat{\sigma}^2(x) / 12 = 4.520$ d'où $\sigma(m) = 2.126$

$L(m_{12}) = S\,(11)\,(72;\ 2.126)$

Remarque : on peut noter $\sigma^2(m)$ ou $\hat{\sigma}^2(m)$, l'important est surtout de faire la différence entre $\sigma^2(x)$ (variance connue dans la population) et $\hat{\sigma}^2(x)$ (variance dans la population estimée = SCEx / (n-1) à partir des n données de l'échantillon)

H0: $\mu = 72$ (groupe comparable à la norme)
H1: $\mu \neq 72$ (groupe différent de la norme) test bilatéral

$t_{\text{calculé}} = (72.32 - 72) / 2.126 = 0.16$ (NS)

Méthode: α donné = 0.04

$t_{\text{théorique}} = t_{1-\alpha/2}\,(\nu) = t_{0.98}\,(11) = 2.5$
$|t_{\text{calculé}}| < t_{\text{théorique}}$ si on rejette Ho on a un risque de se tromper > à 4%.
Remarque : $t_{0.975}\,(11) = 2.201$ la différence est non significative (NS) (α calculé = 0.8732)

On ne peut pas mettre en évidence que le groupe des 12 individus ne soit pas représentatif d'une population centrée sur 72.

2°) $n = 53$ $m = 74$ $s^2(x) = 81$
variable aléatoire m_{53}
$L(m_{53}) = N\,(E(m);\ \sigma(m))$ car $n > 30$
Paramètres de la loi: $E(m) = \mu = 72$
N>>n donc tirage avec remise et $\sigma^2(m) = s^2(x) / (n-1) = 81 / 52 = 1.558$ d'où $\sigma(m) = 1.248$
$L(m_{53}) = N\,(72;1.248)$

H0: $\mu = 72$ (groupe comparable à la norme)
H1: $\mu \neq 72$ (groupe différent de la norme) test bilatéral

$t_{\text{calculé}} = (74 - 72) / 1.248 = 1.6$ (NS)

Méthode: α donné = 0.04

$t_{\text{théorique}} = t_{1-\alpha/2} = t_{0.98} = 2.054$
$|t_{\text{calculé}}| < t_{\text{théorique}}$ si on rejette Ho on a un risque de se tromper > à 4%.
Remarque : $\alpha_{\text{calculé}} = 2{*}P(T > 1.6) = 0.1096$, la différence est non significative (NS)

On ne peut pas mettre en évidence que le groupe des 53 individus ne soit pas représentatif d'une population centrée sur 72.

Exercice 8

Une usine fabrique des pièces circulaires dont il garantit le diamètre moyen à 5 cm.
1) Un acheteur affirme que sur un lot de 37 pièces il a trouvé un diamètre moyen de 5,5 cm avec un écart-type de 1,25 cm et donc que le diamètre des pièces fournies n'est pas aux normes avancées par le fabricant mais lui est supérieur. Quel est votre avis?

2) Un autre acheteur, partant de l'hypothèse que cette dimension suit une loi normale a vérifié le diamètre moyen sur un échantillon de 25 pièces et a trouvé un diamètre moyen de 5,3 cm avec un écart-type de 0,6 cm. Pensez-vous que le lot fournit à cet autre acheteur ait un diamètre supérieur à 5 cm au seuil de 5% ? Calculer le risque α.

X : diamètre d'une pièce
$E(X) = \mu = 5$ $\sigma(x))$ inconnu L(X) inconnue

1°) $n = 37$ $m = 5.5$ $s(x) = 1.25$
variable aléatoire m_{37}
$L(m_{37}) = N(E(m); \sigma(m))$ car $n > 30$
Paramètres de la loi: $E(m) = \mu = 5$
N>>n donc tirage avec remise et $\sigma^2(m) = s^2(x) / (n-1) = 1.25^2 / 36$ d'où $\sigma(m) = 0.208$
$L(m_{37}) = N(5; 0.208)$

Remarque : on peut noter $\sigma^2(m)$ ou $\hat{\sigma}^2(m)$, l'important est surtout de faire la différence entre $\sigma^2(x)$ (variance connue dans la population) et $\hat{\sigma}^2(x)$ (variance dans la population estimée = SCEx / (n-1) à partir des n données de l'échantillon)

H0: $\mu = 5$ (diamètre à la norme)
H1: $\mu > 5$ (diamètre supérieur à la norme) test unilatéral

$t_{\text{calculé}} = (5.5 - 5) / 0.208 = 2.4$ **

Méthode: α est à calculer
$\alpha_{\text{calculé}} = P(T > 2.4) = 0.0082$

L'acheteur a probablement raison, on a 0.82% de chance de se tromper en décidant de renvoyer le lot car il n'est pas à la norme et ceci de façon très significative (**).

2°) $L(X) = N(5; \sigma(x))$ $n = 25$ $m = 5.3$ $s(x) = 0.6$
variable aléatoire m_{25}
$L(m_{25}) = S(24)(E(m); \sigma(m))$ car $L(X) = N$, $\sigma(x)$ inconnu et $n < 30$
Paramètres de la loi: $E(m) = \mu = 5$
N>>n donc tirage avec remise et $\sigma^2(m) = s^2(x) / (n-1) = 0.6^2 / 24$ d'où $\sigma(m) = 0.122$
$L(m_{25}) = N(5; 0.122)$

Remarque : on peut noter $\sigma^2(m)$ ou $\hat{\sigma}^2(m)$, l'important est surtout de faire la différence entre $\sigma^2(x)$ (variance connue dans la population) et $\hat{\sigma}^2(x)$ (variance dans la population estimée = SCEx / (n-1) à partir des n données de l'échantillon)

H0: $\mu = 5$ (diamètre à la norme)
H1: $\mu > 5$ (diamètre supérieur à la norme) test unilatéral

$t_{\text{calculé}} = (5.3 - 5) / 0.122 = 2.45$ *

1ère méthode: α donné = 0.05

$t_{\text{théorique}} = t_{1-\alpha}(\nu) = t_{0.95}(24) = 1.711$

$|t_{\text{calculé}}| > t_{\text{théorique}}$ On rejette Ho avec moins de 5% de chance de se tromper.

2ème méthode: α est à calculer

$\alpha_{\text{calculé}} = P(T > 2.45)$ à l'aide des tables dont on dispose on peut dire que le risque est compris entre 1% et 2.5% ($\alpha_{\text{calculé}} = 0.0159$ à l'aide d'Excel)

L'acheteur a probablement raison, on a entre 1 et 2.5% de chance de se tromper en décidant de renvoyer le lot car il n'est significativement (*) pas à la norme.

Exercice 9

Le directeur commercial d'un important quotidien affirme que plus de 80% des foyers du pays lisent ce quotidien. Un sondage est alors effectué auprès de 1000 foyers et indique que 840 foyers lisent ce quotidien. Est-ce que l'affirmation du directeur commercial est supportée par les résultats du sondage, au seuil de signification 0,06 ?

X: l'individu lit le quotidien (= 1) ou ne le lit pas (= 0)

$L(X) = B(n;p)$ avec $n = 1000$ et $p = 0.8$

Variable aléatoire: f_{1000} = proportion d'individus qui lit l quotidien dans un échantillon de 1000 personnes.

$L(: f_{1000}) = N(E(f); \sigma(f))$ car np et nq >5

Paramètres de la loi $E(f) = 0.8$

N>> n donc tirage avec remise et $\sigma^2(f) = p(1-p) / n = 0.8*0.2/1000$ d'où $\sigma(f) = 0.0126$

Donc $L(f_{1000}) = N(0.8; 0.0126)$

H0: $p = 0.8$ (80% lisent le quotidien)

H1: $p > 0.8$ (plus de 80% lisent le quotidien) test unilatéral

$n = 1000$ $k = 840$ $f = 840 / 1000 = 0.84$

$t_{\text{calculé}} = (0.84 - 0.8) / 0.0126 = 3.16$ ***

Méthode: α donné = 0.06

$t_{\text{théorique}} = t_{1-\alpha} = t_{0.94} = 1.555$

$|t_{\text{calculé}}| > t_{\text{théorique}}$ On rejette Ho avec 6% de chance de se tromper.

On a pu mettre en évidence que plus de 80% des foyers lisent le quotidien avec moins de 6% de chance de se tromper en disant cela. En réalité ce risque est de l'ordre de 0.07% ($P(t > 3.16)$ ce qui montre que le résultat est hautement significatif (***).

Exercice 10

Un laboratoire a mis au point un traitement contre une maladie et affirme que ce traitement a un taux de réussite de 79%. Dans un hôpital où cette maladie est traitée on administre ce traitement et on note que sur 200 malades traités, 150 sont guéris. Peut-on conclure que l'affirmation faite par le laboratoire est vraie au seuil de 5% ?

X: individu guéri (= 1) ou non (= 0)

$L(X) = B(n;p)$ avec $n = 200$ et $p = 0.79$
Variable aléatoire: f_{200} = proportion d'individus guéris dans un échantillon de 200 personnes.
$L(f_{200}) = N(E(f); \sigma(f))$ car np et nq >5
Paramètres de la loi $E(f) = 0.79$
N>>n donc tirage avec remise et $\sigma^2(f) = p(1-p) / n = 0.79*0.21/200$ d'où $\sigma(f) = 0.0288$
Donc $L(f_{200}) = N(0.79; 0.0288)$

H0: $p = 0.79$
H1: $p \neq 0.79$ test bilatéral

$n = 200$ $k = 150$ $f = 150 / 200 = 0.75$

$t_{\text{calculé}} = (0.75 - 0.79) / 0.0288 = -1.389$ NS

Méthode: α donné = 0.05
$t_{\text{théorique}} = t_{1-\alpha/2} = t_{0.975} = 1.96$
$|t_{\text{calculé}}| < t_{\text{théorique}}$ si on rejette Ho on a un risque de se tromper supérieur à 5%.
On conserve Ho, sinon on la rejette avec un risque d'erreur égal à 16.46% (2*P(T> 1.389).
On n'a pas pu mettre en évidence que l'affirmation faite par le laboratoire était vraie.

Exercice 11

Un échantillon aléatoire et indépendant de 10 individus est tiré d'une population où la variable étudiée (résultats à un test d'aptitude) suit une loi normale de variance égale à 100. Les résultats de ce test sont : 78; 60; 64; 82; 80; 66; 74; 61; 68; 57. Pensez-vous que la dispersion des valeurs est conforme à ce que l'on a habituellement dans la population, au risque de 5% ?

X: résultat à un test d'aptitude d'un individu
$L(X) = N(\mu; 10)$

$n = 10$ $SCE(x) = 720$
Ho: $\sigma^2(x) = 100$ (la dispersion est conforme)
H1 $\sigma^2(x) \neq 100$ (la dispersion n'est pas conforme) test bilatéral
$X^2_{\text{calculé}} = SCE(x) / \sigma^2(x) = 720 / 100 = 7.2$
$\alpha = 0.05$ $X^2_{0.025}(9) = 2.70$ $X^2_{0.975}(9) = 19$
$X^2_{0.025}(9) < X^2_{\text{calculé}} < X^2_{0.975}(9)$

On conserve Ho, on n'a pas pu mettre en évidence que la dispersion des valeurs n'était pas conforme à ce que l'on a habituellement.

Exercice 12

Une étude comparative du prix d'un même produits de 2 marques différentes distribuées dans différents magasins a donné les résultats suivants :
Marque A : moyenne égale à 138$ sur 25 magasins
Marque B : moyenne égale à 135$ sur 22 magasins
On sait que la variance des prix est commune et est égale à 100 et que la distribution des prix de ces produits suit une loi normale. La différence entre ces prix moyens est-elle significative au seuil de 1% ?

X: prix d'un produit $L(X) = N (E(X); 10)$

Variable aléatoire de l'étude $\Delta m = m_{25} - m_{22}$
Les distributions des variables aléatoires m_{25} et m_{22} suivent des lois normales car $L(X) = N$ et $\sigma^2(x)$ connue = 100 donc $L (\Delta m) = N (\mu_1 - \mu_2; \sigma_d(m))$
Paramètres de la loi: on pose $E(\Delta m) = \mu_1 - \mu_2$ (on la teste par rapport à 0)
On suppose que dans les 2 échantillons N>>n donc tirage avec remise, les 2 échantillons sont indépendants et donc :

$\sigma^2_d(m) = \sigma^2_1(m) + \sigma^2_2(m)$
$\sigma^2_d(m) = [\sigma^2(x) / n_1] + [\sigma^2 (x) / n_2] = 8.545$ et donc $\sigma_d(m) = 2.923$

$L(\Delta m) = N (0; 2.923)$
Ho: $\mu_1 - \mu_2 = 0$
H1: $\mu_1 - \mu_2 \neq 0$ test bilatéral
$m_1 = 138$ $m_2 = 135$
$t_{calculé} = (138 - 135) / 2.923 = 1.026$ NS

Méthode: α donné = 0.01

$t_{théorique} = t_{1-\alpha/2} = t_{0.995} = 2.576$
$|t_{calculé}| < t_{théorique}$ si on rejette Ho on a un risque de se tromper supérieur à 1%.
On conserve Ho, sinon on la rejette avec un risque d'erreur égal à 30.30% (2*P(T> 1.026).

On n'a pas pu mettre en évidence que la différence entre les prix était significative

Exercice 13

On veut comparer l'efficacité de 2 engrais sur le rendement de blé de parcelles en sachant que la variance liée au 1er engrais est de 100 et celle liée au 2ème engrais est 121. Les résultats suivants ont été obtenus :
Engrais 1 : moyenne égale à 616 sur 64 parcelles mesurées parmi les 700 traitées
Engrais 2 : moyenne égale à 620 sur 68 parcelles mesurées parmi les 800 traitées
Peut-on dire que l'on obtient un meilleur rendement avec le 2ème engrais au seuil 2,5% ?
Calculer le risque α.

X: rendement d'une parcelle   L(X) inconnue
$\sigma^2_1(x) = 100$ $\sigma^2_2(x) = 121$

Variable aléatoire de l'étude $\Delta m = m_{64} - m_{68}$
Les distributions des variables aléatoires m_{64} et m_{68} suivent des lois normales car n>30 donc $L (\Delta m) = N (\mu_1 - \mu_2; \sigma_d(m))$
Paramètres de la loi: on pose $E(\Delta m) = \mu_1 - \mu_2$ (on la teste par rapport à 0)
Dans les 2 échantillons n / N *100 < 10% donc tirages avec remise, de plus les 2 échantillons sont indépendants et donc:
$\sigma^2_d(m) = \sigma^2_1(m) + \sigma^2_2(m)$
$\sigma^2_d(m) = [\sigma^2_1 (x) / n_1] + [\sigma^2_2 (x) / n_2] = 3.342$ et donc $\sigma_d(m) = 1.83$

$L(\Delta m) = N(0; 1.83)$
Ho: $\mu_1 - \mu_2 = 0$
H1: $\mu_1 - \mu_2 < 0$ test unilatéral

$m_1 = 616$ $m_2 = 620$
$t_{\text{calculé}} = (616 - 620) / 1.83 = -2.19$ *

1ère méthode: α donné = 0.025
$t_{\text{théorique}} = t_{1-\alpha} = t_{0.975} = 1.96$
$|t_{\text{calculé}}| > t_{\text{théorique}}$ on rejette Ho avec un risque de se tromper inférieur à 2.5%.

2ème méthode: α à calculer
$\alpha_{\text{calculé}} = P(T > 2.19) = 0.0143$
On rejette Ho avec un risque d'erreur égal à 1.43%.

On a pu mettre en évidence que le deuxième engrais donnait un rendement significativement (*) meilleur.

Exercice 14

Un cours de programmation en langage basic est offert à 2 groupes d'individus d'un même programme. Les résultats suivants ont été obtenus en fin de session :

Groupe A :	$n_1 = 54$	$m_1 = 73$	$s_1 = 11$
Groupe B :	$n_2 = 58$	$m_2 = 77$	$s_2 = 11{,}5$

Peut- on conclure au seuil de signification 0,05 que le groupe B est supérieur au groupe A ?
Calculer le risque α.

X: nombre de points d'un individu L(X) inconnue
$\sigma^2_1(x)$ = inconnue $\sigma^2_2(x)$ = inconnue

Variable aléatoire de l'étude $\Delta m = m_{54} - m_{58}$
Les distributions des variables aléatoires m_{54} et m_{58} suivent des lois normales car n>30 donc $L(\Delta m) = N(\mu_1 - \mu_2; \sigma_d(m))$
Paramètres de la loi: on pose $E(\Delta m) = \mu_1 - \mu_2$ (on la teste par rapport à 0)
On suppose que dans les 2 échantillons N>>n donc tirage avec remise, les 2 échantillons sont indépendants et donc:
$\sigma^2_d(m) = \sigma^2_1(m) + \sigma^2_2(m)$
$\sigma^2_d(m) = [s^2_1(x) / (n_1-1)] + [s^2_2(x) / (n_2-1)] = 4.603$ et donc $\sigma_d(m) = 2.146$

$L(\Delta m) = N(0; 2.146)$
Ho: $\mu_1 - \mu_2 = 0$
H1: $\mu_1 - \mu_2 < 0$ test unilatéral

$m_1 = 73$ $m_2 = 77$
$t_{\text{calculé}} = (73 - 77) / 2.146 = -1.86$ *

1ère méthode: α donné = 0.05
$t_{\text{théorique}} = t_{1-\alpha} = t_{0.95} = 1.645$
$|t_{\text{calculé}}| > t_{\text{théorique}}$ on rejette Ho avec un risque de se tromper inférieur à 5%.

2ème méthode: α à calculer
$\alpha_{calculé}$ = P(T> 1.86) = 0.0314
On rejette Ho avec un risque d'erreur égal à 3.14%.

On a pu mettre en évidence que les résultats du deuxième groupe étaient significativement (*) meilleurs.

Exercice 15

On veut comparer les dépenses hebdomadaires moyennes liées à l'alimentation de familles de 4 personnes dans 2 régions de même caractéristiques sociologiques. On a obtenu :

Région A :	N1 = 300	$n_1 = 38$	$m_1 = 90\$$	$s_1 = 5.5\$$
Région B :	N2 = 350	$n_2 = 40$	$m_2 = 87.3\$$	$s_2 = 4.5\$$

L'hypothèse selon laquelle les dépenses hebdomadaires moyennes pour l'alimentation ne diffèrent pas de façon significative entre les familles de ces 2 régions est-elle vraisemblable au seuil de 5% ?
Calculer le risque α.

X: dépenses hebdomadaires d'une famille L(X) inconnue
$\sigma^2_1(x)$ = inconnue $\sigma^2_2(x)$ = inconnue

Variable aléatoire de l'étude $\Delta m = m_{38} - m_{40}$
Les distributions des variables aléatoires m_{38} et m_{40} suivent des lois normales car n>30 donc L (Δm) = N ($\mu_1 - \mu_2$; σ_d(m))
Paramètres de la loi: on pose $E(\Delta m) = \mu_1 - \mu_2$ (on la teste par rapport à 0)
Dans les 2 échantillons le tirage est sans remise (Taux < 10%), les 2 échantillons sont indépendants et donc:
$\sigma^2_d(m) = \sigma^2_1(m) + \sigma^2_2(m)$
$\sigma^2_d(m) = [K_1 * s^2_1 (x) / (n_1-1)] + [K_2 * s^2_2 (x) / (n_2-1)] =$
$[((300 - 38) / (300 - 1))* 5.5^2 / 37] + [((350 - 40) / (350 - 1))* 4.5^2 / 39] = 1.178$
et donc $\sigma_d(m) = 1.085$

L(Δm) = N (0; 1.085)
Ho: $\mu_1 - \mu_2 = 0$
H1: $\mu_1 - \mu_2 \neq 0$ test bilatéral

$m_1 = 90$ $m_2 = 87.3$

$t_{calculé}$ = (90-87.3) / 1.085 = 2.488 *

1ère méthode: α donné = 0.05
$t_{théorique} = t_{1-\alpha/2} = t_{0.975} = 1.96$
$|t_{calculé}| > t_{théorique}$ on rejette Ho avec un risque de se tromper inférieur à 5%.

2ème méthode: α à calculer
$\alpha_{calculé}$ = 2 * P(T> 2.49) = 0.0128
On rejette Ho avec un risque d'erreur égal à 1.28%.

On a pu mettre en évidence une différence significative (*) entre les dépenses des familles de 4 personnes entre les 2 régions.

Exercice 16

Un laboratoire a vérifié la résistance à l'éclatement (Kg/cm2) d'un réservoir à essence fabriqué par 2 entreprises. On sait que cette variable suit une loi normale. Les résultats suivants ont été obtenus :

A	218	220	222	220	216	224	221	224	219	216
B	218	216	217	218	219	220	219	223		

Peut-on conclure, au seuil de signification de 5% que les 2 types de réservoirs présentent une résistance moyenne à l'éclatement identique?

X: résistance à l'éclatement d'un réservoir L(X) = Normale

$\sigma^2_1(x)$ = inconnue $\sigma^2_2(x)$ = inconnue

Variable aléatoire de l'étude $\Delta m = m_{10} - m_8$

Les distributions des variables aléatoires m_{10} et m_8 suivent des lois de Student car L(X) =N, variances des populations inconnues et n<30 donc $L(\Delta m) = S(\nu)\ (\mu_1 - \mu_2; \sigma_d(m))$

Paramètres de la loi: on pose $E(\Delta m) = \mu_1 - \mu_2$ (on la teste par rapport à 0)

On suppose que dans les 2 échantillons N>>n donc tirage avec remise, les 2 échantillons sont indépendants et donc: $\sigma^2_d(m) = \sigma^2_1(m) + \sigma^2_2(m)$

Mais on est dans le cas de petits échantillons et si l'hypothèse d'invariance des variances est vérifiée on pourra calculer une estimation commune de la variance des populations $\hat{\sigma}^2(x)$

$\sigma^2_d(m) = [\hat{\sigma}^2(x) / n_1] + [\hat{\sigma}^2(x) / n_2]$

Vérification de l'invariance des variances:

Ho: $\sigma^2_1(x) / \sigma^2_2(x) = 1$

H1: $\sigma^2_1(x) / \sigma^2_2(x) \neq 1$ test bilatéral

$\hat{\sigma}^2{}_1(x) = 8.222$ $\nu_1 = 9$ $\hat{\sigma}^2{}_2(x) = 4.5$ $\nu_2 = 7$

F calculé = 8.222 / 4.5 = 1.83

$F_{\text{théorique}} = F_{1-\alpha/2}\ (9\ ,\ 7) = F_{0.975}\ (9\ ,\ 7)$ appartient à l'intervalle [4.76; 4.90]

$F_{\text{calculé}} < F_{\text{théorique}}$ on ne peut pas mettre en évidence l'existence d'une différence entre les variances des populations sinon il faudrait prendre un risque de se tromper supérieur à 5%.

$\hat{\sigma}^2(x) = (SCE_1(x) + SCE_2(x)) / (n_1 + n_2 - 2) = (73.998 + 31.5) / 16 = 6.59$

$\sigma^2_d(m) = [6.59/ 10] + [6.59/ 8] = 1.483$ et donc $\sigma_d(m) = 1.218$

$L(\Delta m) = S(16)\ (0; 1.218)$

Comparaison des moyennes:

Ho: $\mu_1 - \mu_2 = 0$

H1: $\mu_1 - \mu_2 \neq 0$ test bilatéral

$m_1 = 220$ $m_2 = 218.75$

$t_{\text{calculé}} = (220 - 218.75) / 1.218 = 1.026$

Méthode: α donné = 0.05

$t_{\text{théorique}} = t_{1-\alpha/2}\ (16) = t_{0.975}\ (16) = 2.12$

$|t_{\text{calculé}}| < t_{\text{théorique}}$ si on rejette Ho on le fait avec un risque de se tromper supérieur à 5%.

On n'a pas pu mettre en évidence une différence entre les résistances à l'éclatement des 2 types de réservoirs pour un niveau de confiance = 0.95.

Exercice 17

On a mesuré la consommation d'oxygène de 12 individus, avant et après un entraînement de 300 heures, en les soumettant à un effort physique intense pendant 30 minutes. Les résultats obtenus sont :

avant	2.4	2.5	2.6	2.7	2.6	2.5	2.4	2.6	2.5	2.5	2.6	2.7
après	2.6	2.6	2.8	2.9	2.7	2.8	2.7	2.9	2.6	2.5	2.8	2.9

Peut-on conclure à l'amélioration de la condition physique après entraînement au seuil de 2% ?
On admet que la consommation d'oxygène d'un individu (X) suit une loi Normale.

Dans cet exercice les échantillons ne sont pas indépendants il faudra appliquer la méthode des couples.

On détermine d'abord la série des 12 différences:
di: { (0;1), (0.1; 3), (0.2; 5), (0.3;3) }
Sur cette série on calcule: $\overline{d}$ = 0.1833 s_d = 0.08975 ou $\hat{\sigma}_d$ = 0.0937

La variable aléatoire est $\overline{d}_{12}$ et la loi L($\overline{d}_{12}$) = S(11) ((δ; σ($\overline{d}$)) car variance σ^2_d inconnue et n< 30
Les paramètres de la loi:E($\overline{d}$) = δ = 0 (ce que l'on posera dans H0)
On suppose N>>n donc tirage avec remise et $\sigma^2(\overline{d}) = \hat{\sigma}^2_d / n$ d'où σ($\overline{d}$) = 0.0271
Donc L($\overline{d}_{12}$) = S(11) ((0; 0.0271)

H0: δ = 0 (pas d'amélioration)
H1: δ > 0 (amélioration)

$t_{calculé}$ = 0.1833 / 0.0271 = 6.77 ***

Méthode: α donné = 0.02

$t_{théorique} = t_{1-\alpha}(11) = t_{0.98}(11)$ appartient à l'intervalle [2.201; 2.718]
$|t_{calculé}| > t_{théorique}$ on rejette Ho avec un risque de se tromper inférieur à 2%.
(même inférieur à 0.1% car $t_{0.999}(11) = 4.025$).

On a pu mettre en évidence une amélioration hautement significative de la condition physique après entraînement;

Exercice 18

On veut évaluer l'efficacité de 2 insecticides notés 1 et 2 . Avec le premier sur 250 insectes vaporisés 175 ont été éliminés alors qu'avec le deuxième sur 300 insectes vaporisés 80 ont pu survivre. Peut-on conclure, au seuil de 5%, que le premier insecticide est moins efficace que le deuxième? Calculer le risque α.

X: l'individu ne survit pas (= 1) ou survit (= 0) à l'insecticide.
2 échantillons: f_1 = 175 / 250 = 0.7 f_2 = 220 / 300 = 0.7333

Variable aléatoire de l'étude $\Delta f = f_{250} - f_{300}$
Les distributions des variables aléatoires f_{250} et f_{300} suivent des lois normales car n>100 et $0.1 < f < 0.9$ donc $L(\Delta f) = N(p_1 - p_2; \sigma_d(f))$
Paramètres de la loi: on pose $E(\Delta f) = p_1 - p_2 = 0$(on la teste par rapport à 0)
On suppose que dans les 2 échantillons N>>n donc tirage avec remise et donc:
$\sigma^2_d(f) = \hat{p} * \hat{q} (1 / n_1 + 1 / n_2)$ avec :

$\hat{p} = ((175 + 220) / (250 + 300) = 0.7182$

$\hat{q} = 0.2818$

$\sigma^2_d(f) = 1.48\ 10^{-3}$ et donc $\sigma_d(f) = 0.0385$
$L(\Delta f) = N(0; 0.0385)$

Ho: $p_1 - p_2 = 0$
H1: $p_1 < p_2$ test unilatéral

$t_{\text{calculé}} = (0.7-0.7333) / 0.0385 = -0.865$ NS

1ère méthode: α donné = 0.05
$t_{\text{théorique}} = t_{1-\alpha} = t_{0.95} = 1.645$
$|t_{\text{calculé}}| < t_{\text{théorique}}$ si on rejette Ho on le fait avec un risque de se tromper supérieur à 5%.

2ème méthode: α à calculer
$\alpha_{\text{calculé}} = P(T > 0.87) = 0.1922$
Le risque de se tromper en rejetant Ho est égal à 19.22%.

On n'a pas pu mettre en évidence que le premier insecticide était moins efficace que le deuxième.

Exercice 19

On veut comparer la variabilité de 2 procédés utilisés pour mesurer un caractère quantitatif qui suit une loi normale. Deux échantillons aléatoires indépendants donnent les résultats suivants :
Procédé A : n1 = 13 m1 = 1.5 $\Sigma xi^2 = SC1 = 49$
Procédé B : n2 = 21 m2 = 2 $\Sigma xi^2 = SC2 = 159$
Pouvez-vous conclure que la variabilité du procédé B est plus grande que celle du procédé A, au seuil de 5% ?

Ho: $\sigma^2_2(x) / \sigma^2_1(x) = 1$
H1: $\sigma^2_2(x) / \sigma^2_1(x) > 1$ test unilatéral

$SCE1(x) = 49 - 13 * 1.5^2 = 19.75$
$SCE2(x) = 159 - 21 * 2^2 = 75$

$\hat{\sigma}^2_1(x) = 1.6458$ $\nu_1 = 12$ $\hat{\sigma}^2_2(x) = 3.75$ $\nu_2 = 20$
en inversant l'ordre des variances:
$F_{\text{calculé}} = 3.75 / 1.65 = 2.27$ NS
$F_{\text{théorique}} = F_{1-\alpha}(20, 12) = F_{0.95}(20, 12) = 2.54$
$F_{\text{calculé}} < F_{\text{théorique}}$ on ne peut pas mettre en évidence la variabilité du procédé B est plus grande que celle du procédé A, sinon il faudrait prendre un risque de se tromper supérieur à 5%.

Exercice 20

On a mesuré une caractéristique quantitative , distribuée normalement, dans 3 échantillons aléatoires et indépendants et on a obtenu :

Ech 1	53	61	42	59	63	52						
Ech 2	83	46	73	85	73	63	66	54	60	84	62	59
Ech 3		48	94	76	73	65	21	48	62	94		

Les variances sont-elles comparables entre ces 3 échantillons, au seuil de 1% ?

Ho: $\sigma^2_1(x) = \sigma^2_2(x) = \sigma^2_3(x)$
H1: au moins une des variances est plus grande que les autres test unilatéral

Ech	$\hat{\sigma}_i^2$	ν_i	$\nu_i \log \hat{\sigma}_i^2$	$\nu_i\ \hat{\sigma}_i^2$
1	59.6	5	8.8762	298
2	154.97	11	24.0927	1704.67
3	548.53	8	21.9136	4388.22
total		24	54.88	6390.892

$C = 1 + 1/(3(3\text{-}1)(0.4159 – 0.0417) = 1.0624$
$\hat{\sigma}^2 = 6390.892 / 24 = 266.287$

$X^2_{\text{calculé}} = (2.3026/1.0624) * (24\log266.287 – 54.88) = 7.21$
$X^2_{\text{théorique}} = X^2\ 0.99\ (2) = 9.21$

X^2 calculé < X^2 théorique on conserve l'hypothèse d'invariance des variances, on ne peut pas mettre en évidence qu'au moins une des variances est supérieure aux autres à moins de prendre un risque supérieur à 1%.

Exercices supplémentaires

Exercice 21

Des personnes ont été choisies au hasard dans 2 régions afin d'évaluer leur préférence pour un nouveau produit A concurrentiel d'un produit B. Dans la première région, sur 200 personnes interrogées 108 ont préféré le produit A, alors que dans l'autre région sur 250 personnes testées 100 ont préféré le produit B. Pensez-vous qu'il y ait une différence significative entre ces régions au seuil de 5% ? Calculer le risque α.

X: l'individu préfère (= 1) ou ne préfère pas (= 0) le produit A.
2 échantillons: $f_1 = 108 / 200 = 0.54$ $f_2 = 150 / 250 = 0.60$

Variable aléatoire de l'étude $\Delta f = f_{200} – f_{250}$
Les distributions des variables aléatoires f_{200} et f_{250} suivent des lois normales car n>100 et $0.1 < f < 0.9$ donc $L(\Delta f) = N(p_1 - p_2; \sigma_d(f))$
Paramètres de la loi: on pose $E(\Delta f) = p_1 - p_2 = 0$(on la teste par rapport à 0)
On suppose que dans les 2 échantillons N>>n donc tirage avec remise et donc:
$\sigma^2_d(f) = \hat{p} * \hat{q}\,(1 / n_1 + 1 / n_2)$ avec :
$\hat{p} = ((108 + 150) / (200 + 250) = 0.5733$ et $\hat{q} = 0.4267$
$\sigma^2_d(f) = 2.20\ 10^{-3}$ et donc $\sigma_d(f) = 0.0469$
$L(\Delta f) = N(0; 0.0469)$

Ho: $p_1 = p_2$
H1: $p_1 \neq p_2$ test bilatéral

$t_{calculé}$ = (0.54-0.60) / 0.0469 = - 1.279 NS

1ère méthode: α donné = 0.05

$t_{théorique} = t_{1-\alpha/2} = t_{0.975} = 1.96$

$|t_{calculé}| < t_{théorique}$ si on rejette Ho on le fait avec un risque de se tromper supérieur à 5%.

2ème méthode: α à calculer

$\alpha_{calculé} = 2*P(T> 1.28) = 0.2005$

Le risque de se tromper en rejetant Ho est égal à 20.05%.

On n'a pas pu mettre en évidence une différence entre les réponses des individus des 2 régions.