Scénario 10

UTILISATION D'UN TEST STATISTIQUE
Comparaison de deux échantillons

1 Présentation générale

1.1 Objectifs :

- Utiliser un test statistique imposé (comparaison de deux échantillons; test de rang; comparaison de moyennes ...)
- Connaître les conditions de validité du test employé.

1.2 Les étapes

1. Notions de probabilité; qu'est ce qu'une hypothèse?
2. Utilisation d'une statistique; l'hypothèse nulle (H0) et l'hypothèse alternative (H1) :
   Probabilité d'observer la valeur de cette statistique sous H0.
   Notion de valeur critique d'une statistique.
3. Etude et définition du risque α de 1ère espèce.
   Choix du risque α
4. Etude et définition du risque β de 2ème espèce.
   Notion de puissance d'un test $(1 - \beta)$.
5. Bilan des étapes précédentes : que fait on quand on utilise un test statistique?
   Nouvel exemple en utilisant un autre test, le t de Student.
   Utilisation de tables statistiques.
   Quelques conditions d'utilisation d'un test (pour le test de Wilcoxon et le test t).

1.3 Matériel

A) Pour le formateur :

- Un sac opaque
- 20 balles de ping-pong numérotées ainsi :
10 balles numérotées 0
2 balles numérotées 1
8 balles numérotées de 2 à 9 (attention à bien différencier 6 et 9)
- Fichier des agents INRA d'Antibes et de Nancy avec catégorie sexe et âge.
- Transparents (voir liste en annexe)
- Rétroprojecteur
- Un tableau
- Eventuellement; Stat ITCF, le "logiciel Belge" et S.

B) Pour les stagiaires :

- Fichier des agents d'Antibes et de Nancy.
- Feuilles d'exercices (voir liste en annexe)
- Tables de nombres au hasard

- Table pour le test Wilcoxon Signe et Rang
- Table t de Student
- Une calculette

1.4 Les Etapes

Etape 1

Travaux dirigés sous forme de dialogue. A partir d'une "expérience" de tirage de balles numérotées, amener les stagiaires à définir une hypothèse.

Etape 2

A partir de leur propre échantillon chaque stagiaire va calculer une statistique. Quelle décision prendre en fonction de la valeur de cette statistique, (accepter ou rejeter H0)? Réalisation simultanée de transparents par le formateur et de feuilles d'exercices par les stagiaires.

Etape 3

Formateur et stagiaires complètent les transparents précédents. Construction d'un graphique de la distribution d'une statistique sous H0.

Etape 4

Si possible à partir des exercices de chaque stagiaire, faire découvrir le risque β, si non utiliser la simulation de 1000 tirages de 2 échantillons sur les fichiers des agents d'Antibes et de Nancy (complément au scénario; annexe 4).

Etape 5

Bilan des étapes précédentes sous forme de tableau.
Le formateur réalise avec les stagiaires un résumé des différentes étapes nécessaires pour l'utilisation d'un test statistique à l'aide d'un nouveau test, le test de Student (test de t).
Utilisation d'une table statistique.
Le formateur donne les conditions d'utilisation pour le test de Wilcoxon et pour le test de Student.(Utiliser ici les transparents et le tableau pour une pédagogie plus "traditionnelle").

1.5 Forme de l'Evaluation du scénario

Faire utiliser sur un autre jeu de données le test de Wilcoxon.(voir étape 5).
Faire utiliser par les stagiaires le test de t (voir étape 5).

1.6 Les Prérequis

Notion de probabilité.
Paramètres calculés et paramètres de référence.
Estimateur et Intervalle de confiance d'un paramètre.
Paramètres de position.
Paramètres de dispersion.
Quelques notions sur la loi Normale et sa distribution (pour la partie facultative de l'étape 5).

1.7 Durée prévue des différentes étapes

Etape 1 : 1 h
Etape 2 : 1 h à 1 h 30 mn
Etape 3 : 1 h
Etape 4 : 1 h
Etape 5 : 1 h 30 mn

TOTAL : 5 h 30 mn à 6 h réparties en deux séances.

2 Les étapes

2.1 Etape 1 : Probabilité-Hypothèse

Le formateur présente aux stagiaires un sac contenant 20 balles de ping-pong numérotées de 0 à 9, (Aucune précision supplémentaire n'est apportée), et pose la question suivante :
- Quelle est la probabilité de tirer une balle numérotée zéro?
Rappel sur les probabilités : soit l'événement X on a

$$p(X) = \frac{\text{Cas favorables}}{\text{Cas Possibles}}$$

Le formateur propose aux stagiaires une hypothèse; qui peut être :

$$P(0) = \frac{2}{20} = \frac{1}{10}$$

Dans un deuxième temps on cherche à vérifier cette hypothèse :

Pour cela on passe à une phase "d'expérimentation" en faisant effectuer par les stagiaires un ou plusieurs tirages de 10 balles avec remise.

Remarque pour le formateur :

Noter bien la différence entre équiprobabilité de tirage de chaque balle qui dans le tirage avec remise est réalisée ($p\{X\} = 1/20$), mais par contre l'équiprobabilité des numéros dans notre exemple n'est pas réalisée (quand on connaît le contenu du sac) et c'est ce que l'on veut faire découvrir aux stagiaires.

Les résultats des tirages (nombres de zéros pour une série de 10 tirages) sont inscrits au fur et à mesure sur un tableau.

- Au vu de ce tableau quelles conclusions peut on tirer?
- Quelle fréquence d'apparition des zéros observe-t-on?
- Cette observation remet-elle en cause l'hypothèse de départ?
- Si $f\{0\}$ est très différente de $1/10$, on remet en cause l'hypothèse de départ, on décide de rejeter l'hypothèse de départ.

Remarque pour les stagiaires sur les notions de fréquence et de probabilité :

- La fréquence d'apparition de 0 est trouvée à partir de l'échantillon (l'expérience de 10 tirages).
- La probabilité d'apparition de 0 est calculée quand on connaît la population (ici le contenu du sac).
- Les fréquences observées estiment donc les probabilités.

Le formateur peut maintenant dévoiler le contenu du sac et donner la probabilité d'apparition de 0 soit :

$$p\{0\} = 10/20 = 1/2$$

Remarque pour le formateur :

L'expérience de tirer 10 fois avec remise, en comptant le nombre de 0 obtenu, donne une loi Binomiale B(10,1/2).

Rappel sur la démarche faite au cours de cet exercice :

- On formule une hypothèse sur la probabilité d'apparition du zéro après tirage avec remise.
- On fait une "expérience" par laquelle on obtient une fréquence d'apparition de 0.
- D'après cette observation (résultat de l'expérience des tirages) on décide le rejet ou l'acceptation de l'hypothèse formulée initialement.

2.2 Etape 2 : Utilisation d'une statistique

ECHANTILLONNAGE; POSITIONNER DES MEDIANES SUR UN GRAPHIQUE :

Suite à une enquête sur les CPA à l'INRA (Cessation progressive d'activité) on se demande si les agents du Centre d'Antibes sont plus âgés que ceux du Centre de Nancy?

Plus précisément on se pose la question suivante :

- L'âge médian (on raisonne sur les médianes) des agents d'Antibes est-il supérieur à l'âge médian des agents de Nancy?

Le formateur demande alors aux stagiaires comment répondre à cette question?

Réponse souhaitée : Nous allons faire un échantillonnage des deux populations.

Le formateur distribue alors les fichiers des âges des agents d'Antibes et de Nancy (Annexe 1), ainsi qu'une table de nombres au hasard (Annexe 2).

A l'aide de ce matériel on propose à chaque stagiaire (ou groupes de stagiaires) de réaliser un échantillonnage des deux populations :

- Un échantillon pour Antibes comportant 3 "individus" après tirages au sort à l'aide de la table de nombres au hasard .

- Un échantillon pour Nancy comportant 4 "individus".

Chaque stagiaire (ou groupe) reporte ses résultats sur le tableau 1 de la feuille d'Exercice 1, et place aussi les médianes des deux échantillons sur un axe.(pour l'instant ne pas remplir le tableau 2 des rangs).

Rappel pour le calcul de la Médiane d'un échantillon :

C'est la valeur de la variable correspondant au milieu de la distribution. C'est donc la valeur telle que la moitié des mesures lui sont inférieures, et la moitié des mesures lui sont supérieures. Si la taille (l'effectif total) de l'échantillon est un nombre pair (ex : 10) on prend comme valeur de la médiane la moyenne entre la 5ème et la 6ème donnée.

Pour aider les stagiaires, le formateur utilise le Transparent 1 à partir de l'exemple choisi ici.

Le formateur peut ensuite compléter son transparent en y reportant les âges Médians trouvés par les différents groupes, ceci pour appréhender les différences d'échantillonnage et donc de position des médianes.

L'HYPOTHESE NULLE (H0) et L'HYPOTHESE ALTERNATIVE (H1) :

Le formateur pose alors la question suivante :

- Comment, chaque stagiaire au vu de ses propres résultats, va décider que $mA = mN$ est une hypothèse vraisemblable ou au contraire que : $mA \geq mN$?

A ce niveau dire que $mA = mN$ est appelée l'hypothèse nulle encore notée H0.

Et que $mA \geq mN$ est appelée l'hypothèse alternative notée H1

Remarque :

Pour l'hypothèse alternative (H1) et dans cet exemple, on ne se pose pas la question $mA \leq mN$? En effet le problème n'est pas posé initialement (on réalise un test unilatéral). Le cas $mA \geq mN$ est exclu du modèle.

LE CALCUL D'UNE STATISTIQUE; LA REGLE DE DECISION :

En vue d'établir une règle de décision, on propose de résumer les données et les graphiques en utilisant les positions relatives (les rangs) des individus de chaque échantillon, et de calculer ensuite la somme des rangs

de chaque échantillon.

Ce travail est réalisé par chaque stagiaire sur la feuille d'Exercice 1 à partir de son exemple en remplissant le tableau 2.

Le formateur peut aider à remplir le tableau 2 de l'Exercice à l'aide du Transparent 1.(listes des rangs des deux centres et calcul des sommes des rangs WA et WN)
où WA est la somme des rangs pour Antibes
et WN la somme des rangs pour Nancy

Remarques :

1. Si on trouve des ex-aequo, on prend pour chaque valeur le rang moyen (voir Transparent 1). Dans notre exemple on peut expliquer les ex-aequo par "l'imprécision" de la mesure (âge des agents en année).
2. La statistique que nous utilisons est la somme des rangs. C'est un critère de dominance des âges d'Antibes (A) sur ceux de Nancy (N).
3. On fait remarquer aux stagiaires que : quel que soit les échantillons on trouve : WA + WN = 28 (voir transparent 2, colonnes 2 et 3). En effet quelque soit la combinaison on a toujours :

   Somme des rangs = 1+2+3+4+5+6+7 = 28

   La somme des rangs d'Antibes (WA) et la somme des rangs de Nancy (WN) sont donc complètement liées, l'une ou l'autre caractérise de façon identique l'arrangement des âges d'Antibes et de Nancy.

Partant de cette dernière remarque on peut donc travailler indifféremment sur WA ou WN. Dans notre exemple nous choisissons de travailler sur WA car l'effectif de l'échantillon d'Antibes est plus faible, les calculs seront donc plus rapides.

Question :

Que décider au vu de la valeur de WA? Va-t-on faire le choix de $H0(mA = mN)$ ou de $H1(mA \geq mN)$?

Le formateur inscrit alors au tableau la liste des différents WA calculés par chaque groupe de stagiaires.
par exemple :

WA1=15
WA2=13
WA3=16
...

Avec les éléments dont nous disposons nous ne pouvons pas encore répondre à cette question. Nous allons donc examiner toutes les valeurs que peut prendre WA sous l'Hypothèse nulle H0, c'est à dire si les individus d'Antibes et de Nancy appartiennent à une même population d'âge. Pour cela nous allons faire la liste des différentes combinaisons possibles des rangs d'Antibes par rapport à Nancy que nous classons par ordre décroissant des WA.
Le formateur distribue alors l'Exercice 2 qui comporte cette liste (pour d'éventuelles explications il peut s'aider des 3 premières colonnes du Transparent 2).

LA VALEUR CRITIQUE :

A partir de cet exercice faire remarquer que WA est compris entre deux valeurs 6 et 18. ($6 \leq WA \leq 18$). On peut alors amener les stagiaires à trouver une règle de décision en choisissant une valeur critique de WA (notée wc) que l'on compare à la valeur du WA trouvée par l'expérience.

La règle de décision sera :

si $WA \geq wc$ on rejette H0
si $WA < wc$ on accepte H0

2.3 Etape 3 : Le risque de 1ère espèce α)

La feuille d'Exercice 2 donne la liste de toutes les combinaisons possibles des rangs respectifs d'Antibes et de Nancy dans la 1ère colonne, et la valeur de la Somme des Rangs pour Antibes (WA) correspondant à chaque combinaison dans la 3e colonne.

Le formateur (à l'aide du Transparent 2) va guider les stagiaires dans la construction de l'exercice 2 en posant les questions suivantes :

- Quelle est la probabilité d'observer chacune des 35 combinaisons sous H0?

Réponse : 1/35

- Quelle est la probabilité d'observer chacune des valeurs prises par WA (sous H0)?

Réponse :

On remplit la colonne 4 dans laquelle les probabilités sont exprimées en fractions sachant que :

$$P\{WA = wc\} = \frac{\text{nb de cas où on a } wa \text{ comme valeur de } WA}{\text{nb de cas possibles}}$$

On remplit ensuite la colonne 5 où $P\{WA = wa\}$ est exprimée en %.(Faire remarquer ici que sous H0, les valeurs extrêmes ont les probabilités les plus faibles).

Demander ensuite aux stagiaires de remplir la colonne 6 qui donne $P\{WA \geq wa\}$ en %.
Pour cela il suffit de cumuler les $P\{WA = wa\}$ sous H0 (Cf colonne 5).

On demande ensuite quelle est la Probabilité d'avoir sous H0 : un $WA \geq 15$?

Réponse : $\approx$ 20 %

On constate que sous H0, $WA \geq 15$ a une probabilité d'apparition de 20 %, et que par conséquent on rejette H0 avec un risque d'erreur de 20 %. Si on décide de rejeter H0 quand $WA \geq 15$ ce risque, s'appelle le risque (ou erreur) de 1ère espèce, il est noté α.
Le risque α :
C'est le risque de conclure H1 alors que H0 est vraie. C'est donc la probabilité de conclure à une différence qui en réalité n'existe pas.

Remarque :

- C'est l'expérimentateur qui doit fixer (choisir) ce risque α.
- Dans notre exemple que peut on penser de α= 20 %? - Prendre un autre exemple avec rejet si WA $\geq$ 17 on obtient $P\{WA \geq 17\} = 5,7\%$, on a donc $\alpha \approx 6\%$. Comparer α dans ces deux exemples.

Enfin, on peut joindre à cette partie un exercice graphique qui consiste à représenter (sur un diagramme en bâtons ou en densités) la distribution de WA (Exercice 3 pour les stagiaires; Transparents 3 et 3B, 3C pour le formateur).
A l'aide de ces graphiques, discussion à partir des échantillons réalisés par les stagiaires sur la décision qu'ils prennent.

Remarque (pour les stagiaires) :

Nous venons de construire une table donnant les probabilités d'apparition de toutes les combinaisons possibles pour la statistique que nous utilisons. Il faut noter ici que ces tables sont disponibles dans la littérature (voir annexe 3 donnant la table pour le test de Wilcoxon).

2.4 Etape 4 (risque de 2ème ESPECE β; PUISSANCE d'un test 1-β)

On sait, (car nous avons accès à toute la population grâce aux fichiers), que les médianes des deux populations sont respectivement :
mA= 46 pour Antibes et mN= 41 pour Nancy
Les populations sont différentes. L'âge médian des agents d'Antibes est plus élevé que l'âge médian des agents de Nancy. Dans ce cas H0 est fausse et H1 est vraie.

Le formateur demande alors qui avait conclu H0? Ou si personne n'avait conclu H0, présenter le document sur lequel on a effectué 1000 tirages de couples d'échantillons de taille 3 et 4 dans les populations d'Antibes et de Nancy (Complément au scénario en ANNEXE 4). D'après ce document on observe que dans 12 % des cas (122 cas sur 1000) on est amené à conclure H0 (en ayant choisi α= 6 %) alors que l'on est sous H1.
Cette dernière phrase peut encore s'exprimer ainsi :
Dans 122 cas sur 1000 (12 %) on n'a pas décelé une différence qui en réalité existe.
C'est un autre risque nommé risque (ou erreur) de 2ème espèce (noté β).

Le risque β :

C'est le risque de conclure H0 alors que H1 est vraie. C'est donc la probabilité de ne pas déceler une différence qui en réalité existe.

Remarque :

- L'expérimentateur choisit α mais il ne peut pas connaître β. On peut faire remarquer ici que l'on a estimé β à l'aide de 1000 tirages. Or comme 1000 tirages $\approx \infty$ on a une bonne estimation de β, (mais il ne s'agit que d'une estimation car nous ne pouvons faire tous les tirages possibles).
Le formateur fournit aux stagiaires le tableau de l' Exercice 4 qu'il complète avec eux (à l'aide du transparent 4).
(voir tableau page suivante qui reproduit le transparent 4)

	Ce que l'on décide	
Situation Réelle	Accepter H_o	Rejeter H_o (ou accepter H_1
H_o vraie	décision correcte $1-\alpha$ ou niveau de confiance	erreur de 1ère espèce α
H_o fausse ou (H_1 vraie)	erreur de 2ème espèce β	Décision correcte $1-\beta$ ou puissance du test

Résultats possibles d'un test d'hypothèse (voir transparent 4)

Dans l'exemple que nous avons choisi, β = 12 %, et par suite $1-\beta$ = 88 %. On a donc 88 chances sur 100 de mettre en évidence une différence qui existe. Cette probabilité est appelée puissance du test. (Remarque : l'évaluation de $1-\beta$ est effectuée par de nombreux logiciels).

La Puissance du test : $1-\beta$

C'est la probabilité de déceler une différence qui existe.

Remarque (pour le formateur) :

Pour illustrer ces notions le formateur pourra s'aider du logiciel belge, de Stat ITCF, ainsi que d'une brochure de l'ITCF intitulée PUISSANCE D'UNE EXPERIENCE (voir les références dans la bibliographie).

Illustration anecdotique sur l'importance des risques α et β.

Nous savons que dans certains états des USA la peine de mort n'est pas abolie. Dans ces conditions les jurés d'un tribunal américain se trouvent face au problème suivant :

- Faut-il accepter le risque d'innocenter un coupable, soit conclure H0 alors que H1 est vraie? C'est le risque β!
- Ou au contraire peut-on accepter de condamner un innocent, soit conclure H1 alors que H0 est vraie?! C'est le risque β! Le débat sur l'exemplarité des peines et sur l'erreur judiciaire est donc ouvert!!!

Remarque : Le formateur a tout le loisir de choisir d'autres exemples selon sa sensibilité, cela va de soi!

Risque du vendeur et de l'acheteur dans les contrôles de qualité.
Choix du Risque dans les contrôles des pièces industrielles (aviation, nucléaire ...)

2.5 Etape 5 : (Bilan des Etapes; Utilisation d'un autre test, le test*t* Quelques conditions de validité d'un test).

A) Faire le bilan des étapes précédentes :.

Nous venons dans les étapes précédentes de "construire" un test appelé Test de Rang.

Dans ce but qu'avons nous fait? (voir transparent 5)

1 - Définition des Hypothèses H0 et H1.
2 - Choix d'un résumé des données (ici rangs et somme des rangs).
3 - Chercher toutes les valeurs possibles (de la somme des rangs) sous H0 et leur probabilité (P{WA=wa}) (ou plus simplement dans la pratique on les trouve dans une table).
4 - Choisir le risque de 1ère espèce α.
5 - Faire le calcul de la statistique selon le résultat de l'expérience (ici WA) et décider le rejet ou l'acceptation de l'Hypothèse de départ en fonction soit d'une valeur critique (wc) ou d'une probabilité critique.

Enfin on peut proposer, en guise d'évaluation, mais aussi pour apprendre à utiliser une table statistique, l'exercice suivant, basé sur l'utilisation du test de Signe et Rang de Wilcoxon (ou encore test de rang à un échantillon), (Exercice 5 à distribuer aux stagiaires) :

Pour une expérience, sur un grand nombre de poules pondeuses, on utilise une machine permettant de distribuer des rations individuelles d'aliment à chaque animal. On désire distribuer à chaque poule 1 000 g d'aliment. Pour vérifier si la machine destinée à cette opération est correctement réglée, on prend au hasard 12 mangeoires préalablement remplies par la machine, et on les pèse; on obtient les poids suivants :

1017 1001 1008 995 1006 1011 1009 1009 1003 998 990 1007

Question : au vu de ces résultats doit-on régler la machine?

Réponse : voir corrigé en annexe 5

B) Evaluation : Utilisation du test *t* de Student : cette évaluation est facultative et ne doit être présentée que si les stagiaires ont entendu parler de ce test.

Illustration à partir d'un nouveau jeu de données : (distribuer aux stagiaires l'exercice 6; et les guider en réalisant avec eux les 5 étapes du test).

Poids (en g) de 2 lots de poulets à 6 semaines :

Lot 1	Lot 2
445	418
521	366
453	357
422	354
403	356
414	428
424	473
425	380
425	385
462	449

$\widehat{m1} = 439{,}4$ $\widehat{m2} = 396{,}6$ ou $\widehat{m}$: moyenne de l'échantillon

$\widehat{s1} = 33{,}85$ $\widehat{s2} = 42{,}69$ et $\widehat{s}$: écart-type estimé

Remarque : on peut utiliser des notations majuscules au lieu des chapeautés $\widehat{m1}$, $\widehat{m2}$. Mais pas $m1$ et $m2$ qui ressemblent aux médianes théoriques du paragraphe précédent.

1 - Définition des hypothèses H0 et H1 :

les poids des poulets des lots 1 et 2 ne diffèrent pas et sont issus d'une population unique (c'est l'hypothèse H0), les différences constatées sont alors dues à un phénomène aléatoire. ou : Les poulets du lot 1 sont plus lourds que ceux du lot 2 (c'est l'hypothèse H1).

2 - Choix d'un résumé des données :

On résume les données par la différence des moyennes des deux échantillons et on divise cette différence par une estimation de son écart-type.

Rappel de la formule :

$t = \frac{|\widehat{m1} - \widehat{m2}|}{Sd}$ avec $Sd = S\sqrt{\frac{1}{n1} + \frac{1}{n2}}$ et $S = \sqrt{\frac{(n1-1)\widehat{s_1}^2 + (n2-1)\widehat{s_2}^2}{n1+n2-2}}$

ou :

t : statistique de Student.
$\widehat{m1}$ et $\widehat{m2}$: moyennes par échantillon.
Sd : estimée de l'écart-type de la différence
S : écart-type estimé de la population supposée unique (H0).
n1 et n2 : effectifs par lot.
$\widehat{s1}$ et $\widehat{s2}$: écart-type par lot.

Dans notre cas n1 = n2 = 10. On peut alors utiliser la formule simplifiée de Sd, d'où :
$Sd = \sqrt{\frac{\widehat{s_1}^2 + \widehat{s_2}^2}{n}}$ $t = (\widehat{m_1} - \widehat{m_2}) \frac{\sqrt{n}}{\sqrt{\widehat{s_1}^2 + \widehat{s_2}^2}}$

Remarque : Ici on peut utiliser la formule de t simplifiée (quand n1 = n2) pour faire comprendre les influences respectives de n et de $\widehat{s1}$ et $\widehat{s2}$ sur la valeur du t.
-t augmente quand $\sqrt{n}$ augmente donc quand n augmente.
-t augmente quand $\widehat{s1}$ et $\widehat{s2}$ diminuent.

3 - Chercher toutes les valeurs possibles de t sous H0, et leur probabilité (voir les remarques qui suivent en A et B sur la distribution de t) :

Ces informations sont fournies par la table de distribution de t que l'on donnera aux stagiaires.(voir annexe 6)

4 - Choix du risque de 1ère espèce α
Dans cet exercice on peut décider que $\alpha = 5\ \%$

5 - Faire le calcul de la statistique et prendre une décision :
Dans notre exemple on trouve $t = 2{,}48$
On lit dans la table de distribution de t la valeur critique de t pour $\alpha = 5\ \%$ et un degré de liberté de 18 (ddl = n1+n2-2 = 10+10-2=18).
Cette valeur est : $t_{0,05} = 2{,}10$
On a donc $t > t_{0,05}$. On rejette H0. On peut donc conclure que l'on observe une différence significative entre les lots au seuil de 5%.

Conclusion de l'essai : la moyenne des poids du lot 1 est donc significativement plus élevée que celle du lot 2, sachant que nous avons 5 chances sur 100 de nous tromper en choisissant cette hypothèse.

Remarque : si nous choisissons $\alpha = 1\ \%$ nous trouvons dans la table $t_{0,01} = 2{,}88$. Dans ce cas $t < t_{0,01}$. Donc au seuil $\alpha = 1\ \%$ on ne rejette pas H0. Conclusion : Dans les conditions de notre expérience et en choisissant $\alpha = 1\ \%$, nous n'avons pas mis en évidence de différence hautement significative entre les lots 1 et 2.

Remarque pour le formateur : Il faut noter que dans cet exemple, les poulets du lot 1 sont de sexe mâle tandis que ceux du lot 2 sont des femelles, le test à donc porté sur l'effet du sexe.

Remarques sur la Distribution de t et le test de t :

A) Distribution de t :

Dans la mesure où la variable étudiée (ici les poids des poulets) suit la loi normale (loi de Laplace Gauss), on sait que la statistique t suit la loi de Student. Or quand les effectifs des échantillons sont importants (donc si $n \Longrightarrow \infty$) la distribution de t suit la loi normale (voir transparent 7)

A partir d'une même population si l'on effectue le tirage de deux échantillons, on se place sous H0. Cette simulation nous permet de trouver les différentes valeurs de t sous H0, et leur probabilité d'apparition. Sur le graphique (transparent 7), on voit que les valeurs de t les plus probables sont voisines de 0, et les moins probables sont celles qui diffèrent le plus de 0. On peut donc dire, qu'à partir d'une population unique on a plus de chance de tirer des échantillons semblables que très différents.

B) Test t :

La valeur de t calculée à partir de deux échantillons permet de mesurer la probabilité qu'ont ces deux échantillons d'appartenir à une population unique (H0).
Dans l'exemple traité t= 2,48. On trouve une probabilité associée de 1,10 % que $t \geq 2{,}48$ sous H0 (voir annexe 7). Cela veut dire que la valeur calculée de $t \geq 2{,}48$ n'a qu'une probabilité de 1,10 % d'apparition dans le cas ou les deux échantillons sont issus d'une population unique.
Enfin, on peut également donner une signification plus "intuitive" de la formule du t.
Rappel de la formule de t :

$$t = \frac{\left|\widehat{m1} - \widehat{m2}\right|}{Sd}$$

D'après cette formule on voit que t tient compte de deux éléments :

- la différence des moyennes $\left|\widehat{m1} - \widehat{m2}\right|$
- L'écart-type de cette différence (Sd) elle-même liée aux écart- types des deux échantillons, et aux effectifs

(voir la formule de Sd plus haut, et le transparent 6).

On peut imaginer les situations A et B suivantes (voir transparent 8)

Dans les situations A et B, les différences des moyennes sont identiques on a bien : $\left|\widehat{m1} - \widehat{m2}\right| = \left|\widehat{m'1} - \widehat{m'2}\right|$

Cependant, en observant les représentations graphiques, on voit que la "séparation" entre les 2 échantillons est moins évidente dans la situation B que dans la situation A. Il y a en B un plus grand recouvrement des 2 courbes, qui est du à une variabilité plus importante à l'"intérieur des échantillons" (variabilité intra échantillon) dans l'exemple B que dans l'exemple A.

En résumé :

Comme S'd > Sd on aura, pour la même différence entre les moyennes, $t' < t$ le test sera donc moins "puissant" dans la situation B.

C) Les Conditions d'utilisation d'un test statistique :

Remarques très importantes : (pour les stagiaires)

Il est important de noter ici que le "statisticien amateur" risque d'utiliser de manière intempestive des techniques statistiques par ailleurs excellentes, parce qu'il n'en connaît pas assez les conditions d'application. En effet, les tests statistiques ne sont valides que dans des conditions déterminées, car ils ont été établis en partant de prémisses définies. Celles-ci constituent ce que les statisticiens ont coutume d'appeler le "modèle mathématique", à partir duquel sont calculées les équations sur lesquelles se basent les tests en question. En construisant un modèle mathématique, le statisticien utilise des paramètres qui ont certaines propriétés et qui remplissent certaines conditions. Les déductions ainsi obtenues ne seront applicables à des situations pratiques que si ces conditions sont satisfaites.

ATTENTION à l'application pure et simple d'une formule, et à la mentalité "Je presse le bouton, puis la machine fera le reste"!!! Il importe donc d'éviter tout "automatisme" dans le traitement statistique des données.

A) Les Conditions d'utilisation du test de Wilcoxon :

-1) Indépendance des données (on ne doit pas pouvoir déduire xi+1 à partir de xi).
-2) Même forme de distribution pour les deux populations.
-3) Distribution symétrique (cas du test pour un échantillon)

B) Les Conditions d'utilisation du test de *t* de Student :
(Utiliser un tableau pour écrire ces conditions et le Transparent 10)
Le test *t*s'appuie sur un modèle mathématique plus complexe que le test de Wilcoxon, les conditions de validités seront donc plus contraignantes.

-1) Normalité de la distribution de l'erreur expérimentale :

En d'autres termes Les résultats que l'on obtiendrait, en l'absence de tout traitement, doivent être dispersés suivant la loi normale (Transparent 7).

Rappel : l'erreur en statistique ne signifie pas "méprise", "bévue" ou "faute", c'est seulement la résultante de toutes les variations aléatoires ou variations non contrôlées.

Remarque : cette hypothèse peut quelquefois être vérifiée dans une expérience à blanc (essai d'uniformité). Plus simplement par une bonne connaissance de la variable étudiée. Les tests utilisés (ici le t) supportent quelque écart par rapport à cette hypothèse.
Remède : transformation éventuelle des données (ex : log(x) permet de "normaliser" certaines distributions dissymétriques).

-2) Homogénéité de la Variance (ou Invariance de la Variance, ou Homoscédasticité) :
(voir transparent 9)
Il faut que tous les groupes expérimentaux aient la même variance. En d'autres termes : les résultats obtenus sous l'action d'un traitement doivent être dispersés suivant une loi analogue à la population d'origine, seule la moyenne a éventuellement varié. Donc, l'écart-type σ n'a pas changé, c'est à dire que l'effet du traitement est de faire glisser la courbe de distribution (des poids par exemple), le long de l'axe de la variable, en laissant sa forme inchangée.(voir transparent 9).
Remarques : Si cette condition n'est pas remplie, il y a une perte d'information (qui peut être très grande) sur les effets des traitements. Il y a donc une diminution considérable de l'efficacité de l'analyse. Donc attention cette condition est importante, toujours jeter un coup d'oeil sur les variances des groupes expérimentaux (traitements, blocs etc...).
Remèdes éventuels : Transformation de la variable.

-3) Indépendance des erreurs :

Il faut que les variations non contrôlées n'aient pas de liaison entre elles, donc qu'elles se produisent strictement au hasard. En d'autres termes : la probabilité pour une observation donnée d'avoir une "erreur" d'une certaine valeur et d'un signe donné ne doit pas dépendre des valeurs ni des signes des "erreurs" des autres observations.

Remarques : Cette condition est essentielle et sa non observation entraîne les effets les plus graves (biais dans l'estimation de l'"erreur", dans les estimations des effets traitements et blocs.).
Remède : Le seul possible, c'est de prêter une attention particulière au dispositif expérimental et de respecter la casualisation du matériel (répartition au hasard). Il faut donc une étude sérieuse du dispositif expérimental et de l'organisation même de l'essai.
Quelques règles en vue de respecter l'indépendance des erreurs :
- Tout individu doit avoir une chance égale d'être soumis à un traitement quelconque (les traitements doivent être assignés au hasard entre tous les individus qui constituent le matériel expérimental).
- Au cours de l'expérience tous les individus doivent avoir une chance égale d'être influencés par les diverses conditions du milieu non contrôlées.
- La casualisation du matériel (répartition au hasard) doit se faire à l'aide de moyens objectifs (tirage au sort, tables de nombres au hasard etc...), qui sont indépendants de la connaissance ou de la volonté de l'expérimentateur.

Pour terminer passons à l'interprétation des résultats :
Enfin, si notre expérimentateur, qui est agent de l'INRA, est Technicien de la Recherche, il pourra dégager des éléments d'analyse, si il est AI il participera à l'analyse des données et si ses fonctions le lui permettent (IE et IR) il interprétera les phénomènes observés (ici les formateurs utiliseront un matériel pédagogique qu'ils peuvent se procurer, auprès du service du personnel de leur centre, qui sont les fiches types de fonctions diffusées par l'INRA, et en donner un exemplaire à chaque stagiaire correspondant à sa catégorie, afin que ces derniers méditent sur les techniques d'interprétation des résultats en fonction de leur catégorie)!!!
Dernière remarque : ce dernier paragraphe n'était qu'une boutade en cette période de concours internes (Février-Mars 91)!!!

2.6 Mots-clé et points stratégiques

- Probabilité
- Hypothèses
- Hypothèse nulle (H0)
- Hypothèse alternative (H1)
- Règle de décision
- Valeur critique
- Risque de 1ère espèce α
- Risque de 2ème espèce β

- Puissance du test $1 - \beta$
- Test unilatéral ou test bilatéral
- Conditions d'utilisation et validité d'un test

2.7 Liste du vocabularaire

- Hypothèse nulle
- Hypothèse alternative
- Risque d'erreur de première espèce
- Risque d'erreur de deuxième espèce
- Valeur critique
- Puissance d'un test

2.8 Références bibliographiques

R Tomassone - Mathématique et informatique; BIOMETRIE INRA Juillet 1987

P Nemenyi; S K Dixon; N B White Jr et M L Hedström - Statistics from scratch; Pilot édition 1977.

P Dagnelie - Théorie et méthodes statistiques vol 1 et 2; Les presses agronomiques de Gembloux 1973 réimpression 1985.

A Vessereau - Méthodes statistiques en biologie et en agronomie; J.B Baillière (Lavoisier) 1988.

K Mather - Analyse statistique en biologie; Gauthier-Villars 1965.

R Heller - Manuel de statistique biologique; Gauthier-Villars 1968.

G.W Snedecor - Statistical methods; The Iowa State College Press 1956.

L. Lison - Statistique appliquée à la biologie expérimentale Gauthier- Villars 1968.

G. Philippeau - Puissance d'une expérience; I.T.C.F 1974.

2.9 Liste des Transparents, Exercices, Annexes

I) Liste des Transparents :

1. Calcul des Médianes et des Sommes des Rangs
2. Tableau des Probabilités cumulées des Wa sous H0
3. Distribution des Wa sous H0

   3 B) idem avec $Wa \geq 15; \alpha = 20\%$

   3 C) idem avec $Wa \geq 17; @alpha = 6\%$
4. Résultats possibles d'un test d'hypothèse
5. Les Etapes d'un Test d'Hypothèse
6. Formule du t de Student
7. Distribution Normale et distribution de t
8. 2 situations différentes pour une même différence de moyennes
9. Homogénéité de la Variance

10. Les Conditions de validité du test t

II) Liste des Exercices :

1. Calcul des Médianes et des Sommes des Rangs (Cf Transp 1)
2. Tableau des Probabilités cumulées des Wa sous H0 (Cf Transp 2)
3. Distribution des Wa sous H0 (Cf Transp 3)
4. Résultats possibles d'un test d'hypothèse (Cf Transp 4)
5. Utilisation du test de Wilcoxon (Cf Annexe 5)
6. Utilisation du test t de Student (Cf Annexe 7)

III) Liste des Annexes :

1. Fichier des âges des agents d'Antibes et de Nancy
2. Table de nombres au hasard
3. Table pour le test de Wilcoxon
4. Complément au scénario 10 (Claire Chabanet).
5. Corrigé de l'Exercice 5 sur le test de Wilcoxon
6. Table de t
7. Corrigé de l'Exercice 6 sur le test t
8. Des simulations pour étudier l'effet de l'écart aux conditions d'utilisation sur la validité d'un test. Ce texte n'est pas joint à ce document. Il est disponible auprès de l'auteur Claire Chabanet, Biométrie, Centre de Jouy-en-Josas.