Tests paramétriques

Frédéric Bertrand[1]

[1]IRMA, Université de Strasbourg
Strasbourg, France

M1 MCB2a FC 01-06-2010

Première partie I

Introduction générale à la notion de tests

Partie 1 : Introduction générale à la notion de tests

1. À quoi sert un test ?
   - Hypothèses et erreurs
   - Tests bilatéral et unilatéral
2. Contruction d'un test
   - Statistique de test
   - Région critique et région d'acceptation
3. Mise en oeuvre pratique
   - La démarche à suivre pour la mise en place d'un test
   - Choix d'un test

Definition

Un test est un mécanisme qui permet de trancher entre deux hypothèses à la vue des résultats d'un échantillon, en quantifiant le risque associé à la décision prise.

Les deux hypothèses

Soient $\mathcal{H}_0$ et $\mathcal{H}_1$ deux hypothèses, dont une et une seule est vraie.

$\mathcal{H}_0$ joue le plus souvent un rôle prédominant par rapport à $\mathcal{H}_1$.

En effet $\mathcal{H}_0$ est l'**hypothèse de référence** alors que $\mathcal{H}_1$ est l'**hypothèse alternative**.

Exemple d'hypothèses

Vous pouvez avoir comme hypothèse nulle :

$$\boxed{\mathcal{H}_0 : \text{La moyenne de la population est égale à } \mu_0}$$

et, dans ce cas, une hypothèse alternative **pourrait être**

$$\boxed{\mathcal{H}_1 : \text{La moyenne de la population est différente de } \mu_0}$$

ou encore

$$\boxed{\mathcal{H}_1 : \text{La moyenne de la population est strictement plus grande que } \mu_0.}$$

Exemple d'hypothèses (suite)

Une manière condensée d'écrire ces hypothèses est :

$$\boxed{\mathcal{H}_0 : \mu = \mu_0}$$

contre

$$\boxed{\mathcal{H}_1 : \mu \neq \mu_0.}$$

et

$$\boxed{\mathcal{H}_0 : \mu = \mu_0}$$

contre

$$\boxed{\mathcal{H}_1 : \mu > \mu_0.}$$

Décision

La décision d'un test consiste à choisir entre $\mathcal{H}_0$ et $\mathcal{H}_1$.

Risques

Il y a donc quatre cas possibles qui sont détaillés dans le tableau ci-dessous :

	$\mathcal{H}_0$ vraie	$\mathcal{H}_1$ vraie
$\mathcal{H}_0$ décidée	$1-\alpha$	β
$\mathcal{H}_1$ décidée	α	$1-\beta$

où α **et** β **sont les risques d'erreur de première et de deuxième espèce.**

Definition

L'**erreur de première espèce** est le fait de décider que l'hypothèse alternative $\mathcal{H}_1$ est vraie alors qu'en fait, en réalité, c'est l'hypothèse nulle $\mathcal{H}_0$ qui est vraie.

Le **risque d'erreur** associé à cette décision est noté généralement α.

Il s'agit donc de la probabilité de décider **à tort** que l'hypothèse alternative $\mathcal{H}_1$ est vraie.

Definition

L'**erreur de deuxième espèce** est le fait de décider que l'hypothèse nulle $\mathcal{H}_0$ est vraie alors qu'en fait, en réalité, c'est l'hypothèse alternative $\mathcal{H}_1$ qui est vraie.

Le **risque d'erreur** associé à cette décision est noté généralement β.

Il s'agit donc de la probabilité de décider **à tort** que l'hypothèse nulle $\mathcal{H}_0$ est vraie.

Lien entre les risques

La situation idéale serait que ces deux erreurs soient nulles mais ce n'est pas possible.

Pire encore, toutes autres choses étant fixées, ces deux erreurs sont **antagonistes** :

Si vous diminuez α alors β augmente et inversement si diminuez β alors α augmente.

Niveau de significativité

Depuis les travaux de Neyman et Pearson, l'erreur de première espèce est limitée à un niveau dit **niveau de significativité**.

Le fait d'imposer α faible conduit à une règle de décision plus stricte.

En effet, dans ce cas, la décision consiste à abandonner $\mathcal{H}_0$ dans des cas rarissimes, et à conserver, plus souvent, à tort $\mathcal{H}_0$.

Niveaux usuels

Les valeurs les plus courantes pour α sont 10%, 5% ou 1%.

Definition

La **puissance d'un test** est égale à $1 - \beta$ ou encore la puissance est la probabilité de rejeter $\mathcal{H}_0$ à raison.

À retenir

Généralement la puissance doit au moins être égale à **0,80** pour être considérée comme satisfaisante.

Remarque

Le calcul de la puissance d'un test est généralement assez complexe : il faut souvent faire appel à un des fonctions ou à des logiciels spécialisés.

`Gpower3` est un logiciel gratuit qui permet de réaliser la plupart des calculs de puissance.
`http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3`

Partie 1 : Introduction générale à la notion de tests

1. À quoi sert un test ?
   - Hypothèses et erreurs
   - Tests bilatéral et unilatéral
2. Contruction d'un test
   - Statistique de test
   - Région critique et région d'acceptation
3. Mise en oeuvre pratique
   - La démarche à suivre pour la mise en place d'un test
   - Choix d'un test

Introduction

Avant d'appliquer tout test statistique, il s'agit de bien définir le problème posé.

En effet, selon les hypothèses formulées, vous appliquerez soit un **test bilatéral**, soit un **test unilatéral**.

Definition

Un **test bilatéral** s'applique quand vous cherchez une différence entre deux paramètres, ou entre un paramètre et une valeur donnée sans se préoccuper du signe ou du sens de la différence.

Dans ce cas, la zone de rejet de l'hypothèse principale se fait de part et d'autre de la distribution de référence.

Definition

Un **test unilatéral** s'applique quand vous cherchez à savoir si un paramètre est supérieur (ou inférieur) à un autre ou à une valeur donnée.

La zone de rejet de l'hypothèse principale est située d'un seul côté de la distribution de probabilité de référence.

Exemple de test

Certains tests comme le test du Khi-carré ou le test de Fisher dans une analyse de la variance sont pratiquement toujours unilatéraux.

Partie 1 : Introduction générale à la notion de tests

1. À quoi sert un test ?
   - Hypothèses et erreurs
   - Tests bilatéral et unilatéral
2. Contruction d'un test
   - Statistique de test
   - Région critique et région d'acceptation
3. Mise en oeuvre pratique
   - La démarche à suivre pour la mise en place d'un test
   - Choix d'un test

Statistique de test

Le risque d'erreur de première espèce α étant fixé, il faut choisir une **variable de décision** encore appelée **statistique de test**.

Cette variable est construite afin d'apporter de l'information sur le problème posé, à savoir le choix entre les deux hypothèses.

Sa loi doit être parfaitement déterminée dans au moins une des deux hypothèses (le plus souvent dans $\mathcal{H}_0$) afin de ne pas introduire de nouvelles inconnues dans le problème.

Partie 1 : Introduction générale à la notion de tests

1. À quoi sert un test ?
   - Hypothèses et erreurs
   - Tests bilatéral et unilatéral
2. Contruction d'un test
   - Statistique de test
   - Région critique et région d'acceptation
3. Mise en oeuvre pratique
   - La démarche à suivre pour la mise en place d'un test
   - Choix d'un test

Definition

La **région critique** notée W (W pour wrong), ou encore appelée **zone de rejet** est égale à l'ensemble des valeurs de la variable de décision qui conduisent à écarter $\mathcal{H}_0$ au profit de $\mathcal{H}_1$.

La **région critique** correspond donc aux intervalles dans lesquels les différences sont trop grandes pour être le fruit du hasard d'échantillonnage.

Remarque

Dans la plupart des situations que vous rencontrerez dans la suite, la région critique W peut être reliée au risque d'erreur de première espèce α par $\mathbb{P}_{\mathcal{H}_0}[W] = \alpha$.

Definition

La **région d'acceptation** notée $\overline{W}$, ou encore appelée **zone d'acceptation** est la région complémentaire de la région critique W. Elle correspond à l'intervalle dans lequel les différences observées entre les réalisations et la théorie sont attribuables aux fluctuations d'échantillonnage.

Remarque

Dans la plupart des situations que vous rencontrerez dans la suite, la région d'acceptation $\overline{W}$ peut être reliée au risque d'erreur de première espèce α par $\mathbb{P}_{\mathcal{H}_0}\left[\overline{W}\right] = 1 - \alpha$.

Partie 1 : Introduction générale à la notion de tests

1. À quoi sert un test ?
   - Hypothèses et erreurs
   - Tests bilatéral et unilatéral
2. Contruction d'un test
   - Statistique de test
   - Région critique et région d'acceptation
3. Mise en oeuvre pratique
   - La démarche à suivre pour la mise en place d'un test
   - Choix d'un test

Comment réaliser un test et conclure à l'aide d'une région critique ?

1. Choix des deux hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$.
2. Détermination de la variable de décision.
3. Allure de la région critique en fonction de $\mathcal{H}_1$: test bilatéral ou unilatéral.
4. Calcul de la région critique en fonction de α.
5. Calcul de la variable de décision observée sur l'échantillon.

Comment réaliser un test et conclure à l'aide d'une région critique ?

1. Choix des deux hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$.
2. Détermination de la variable de décision.
3. Allure de la région critique en fonction de $\mathcal{H}_1$: test bilatéral ou unilatéral.
4. Calcul de la région critique en fonction de α.
5. Calcul de la variable de décision observée sur l'échantillon.

Comment réaliser un test et conclure à l'aide d'une région critique ?

1. Choix des deux hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$.
2. Détermination de la variable de décision.
3. Allure de la région critique en fonction de $\mathcal{H}_1$: test bilatéral ou unilatéral.
4. Calcul de la région critique en fonction de α.
5. Calcul de la variable de décision observée sur l'échantillon.

Comment réaliser un test et conclure à l'aide d'une région critique ?

1. Choix des deux hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$.
2. Détermination de la variable de décision.
3. Allure de la région critique en fonction de $\mathcal{H}_1$: test bilatéral ou unilatéral.
4. Calcul de la région critique en fonction de α.
5. Calcul de la variable de décision observée sur l'échantillon.

Comment réaliser un test et conclure à l'aide d'une région critique ?

1. Choix des deux hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$.
2. Détermination de la variable de décision.
3. Allure de la région critique en fonction de $\mathcal{H}_1$: test bilatéral ou unilatéral.
4. Calcul de la région critique en fonction de α.
5. Calcul de la variable de décision observée sur l'échantillon.

Comment réaliser un test et conclure à l'aide d'une région critique ?

6. Conclusion du test.
   Si la valeur calculée en 5 appartient à la région construite en 4, le test est significatif au niveau α.
   Vous rejetez l'hypothèse nulle $\mathcal{H}_0$ et vous décidez que l'hypothèse alternative $\mathcal{H}_1$ est vraie.
   Le risque associé à cette décision est un risque d'erreur de première espèce qui vaut α.

   Si la valeur calculée en 5 n'appartient pas à la région construite en 4, le test n'est pas significatif au niveau α.
   Vous conservez l'hypothèse nulle $\mathcal{H}_0$ par défaut.
   Le risque associé à cette décision est un risque d'erreur de deuxième espèce qui vaut β. Pour l'évaluer, il faudrait calculer la puissance $1 - \beta$ du test.

Comment réaliser un test et conclure à l'aide d'une région critique ?

7. Calcul de la puissance $1 - \beta$ du test lorsque celui-ci n'est pas significatif.

Comment réaliser un test et conclure à l'aide d'une p-valeur ?

1. Choix des deux hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$.
2. Détermination de la variable de décision.
3. Calcul de la p-valeur à partir des données de l'échantillon.

Comment réaliser un test et conclure à l'aide d'une p-valeur ?

1. Choix des deux hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$.
2. Détermination de la variable de décision.
3. Calcul de la p-valeur à partir des données de l'échantillon.

Comment réaliser un test et conclure à l'aide d'une p-valeur ?

1. Choix des deux hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$.
2. Détermination de la variable de décision.
3. Calcul de la p-valeur à partir des données de l'échantillon.

Comment réaliser un test et conclure à l'aide d'une *p*-valeur ?

4. Conclusion du test.
   Si la *p*-valeur est inférieure ou égale à α, le test est significatif au niveau α.
   Vous rejetez l'hypothèse nulle $\mathcal{H}_0$ et vous décidez que l'hypothèse alternative $\mathcal{H}_1$ est vraie.
   Le risque associé à cette décision est un risque d'erreur de première espèce qui vaut α.

   Si la *p*-valeur est strictement supérieure à α, le test n'est pas significatif au niveau α.
   Vous conservez l'hypothèse nulle $\mathcal{H}_0$ par défaut.
   Le risque associé à cette décision est un risque d'erreur de deuxième espèce qui vaut β. Pour l'évaluer, il faudrait calculer la puissance $1 - \beta$ du test.

Comment réaliser un test et conclure à l'aide d'une p-valeur ?

5. Calcul de la puissance $1 - \beta$ du test lorsque celui-ci n'est pas significatif.

Partie 1 : Introduction générale à la notion de tests

1. À quoi sert un test ?
   - Hypothèses et erreurs
   - Tests bilatéral et unilatéral
2. Contruction d'un test
   - Statistique de test
   - Région critique et région d'acceptation
3. Mise en oeuvre pratique
   - La démarche à suivre pour la mise en place d'un test
   - Choix d'un test

Introduction

Plusieurs tests de conception très différente sont souvent disponibles pour soumettre à une épreuve de vérité une hypothèse.

Definition

Le **test le plus puissant** est le test qui fournit l'erreur β la plus petite, pour une même valeur de α ou encore qui fournit la plus grande valeur de la puissance $1 - \beta$.

Intérêt pratique

Il peut détecter les plus petites différences entre les populations sans pour autant augmenter α.

Conditions d'utilisations

La majorité des tests statistiques repose sur le respect d'un certain nombre de conditions. Selon le degré de respect de ces conditions d'utilisation, la validité des résultats se trouve plus ou moins affectée et elle l'est d'autant plus que le test est moins robuste.

Definition

La **robustesse d'un test** équivaut à sa tolérance vis-à-vis du respect des conditions d'application du test.

À retenir

Vous pouvez disposer de plusieurs tests pour vérifier une même hypothèse. En fonction du contexte, il faudra penser à utiliser le plus puissant d'entre eux. Vous apprendrez bientôt les différentes caractéristiques des tests les plus fréquemment utilisés.

Remarque

1. Les tests peu puissants augmentent la probabilité de commettre une erreur de deuxième espèce. Or, cette erreur peut s'avérer particulièrement grave.

Exemple d'erreur de deuxième espèce

En effet, en médecine, considérez une analyse statistique qui permettrait de décider si un patient est sain ($\mathcal{H}_0$) ou malade ($\mathcal{H}_1$). Classer comme malade un sujet bien portant (erreur de première espèce), peut avoir des conséquences aussi graves que classer comme bien portant un sujet malade (erreur de deuxième espèce).

2. Pour évaluer la puissance d'un test, vous pourrez être amené à utiliser des **courbes de puissance** ou encore appelées **abaques**.

Deuxième partie II

Tests de comparaison avec une norme

Problématique

Un problème fréquent est de comparer la moyenne d'un caractère d'une population avec une norme.

Nous supposons que ce caractère est distribué normalement au sein de la population.

Partie 2 : Tests de comparaison avec une norme

4 Espérance
- Espérance d'une loi normale de variance connue
- Espérance d'une loi normale de variance inconnue

5 Variance
- Variance d'une loi normale d'espérance connue
- Variance d'une loi normale d'espérance inconnue

6 Grands échantillons
- Espérance d'une loi quelconque, cas des grands échantillons
- Variance d'une loi quelconque, cas des grands échantillons

7 Proportion
- Proportion, tirage avec remise
- Proportion, tirage sans remise

Test unilatéral I

Soit X une variable aléatoire qui suit une loi normale d'espérance μ et de variance σ^2 connue.

Hypothèses du test

Vous souhaitez choisir entre les deux hypothèses suivantes :

$$\boxed{\mathcal{H}_0 : \mu = \mu_0}$$

contre

$$\boxed{\mathcal{H}_1 : \mu > \mu_0 \text{ ou } \mu < \mu_0.}$$

Test unilatéral II

Conditions d'application du test

Il faut que l'échantillon $x_1, \ldots, x_n$ soit des réalisations indépendantes de la variable aléatoire X qui suit une loi normale.

Statistique du test

La variable aléatoire $Z = \dfrac{\widehat{\mu}_n - \mu_0}{\sigma/\sqrt{n}}$ suit une loi normale $\mathcal{N}(0,1)$.

Test unilatéral III

Décision et conclusion du test

La valeur critique du test, notée c_α, est lue dans la table de la loi normale centrée et réduite. Si la valeur de la statistique calculée sur l'échantillon, notée z_{obs}, est supérieure ou égale à c_α (ou inférieure ou égale à c_α), alors le test est significatif. Vous rejetez $\mathcal{H}_0$ et vous décidez que $\mathcal{H}_1$ est vraie avec un risque de première espèce $\alpha = 5\%$. Si la valeur de la statistique calculée sur l'échantillon, notée z_{obs}, est strictement inférieure à c_α, alors le test n'est pas significatif. Vous conservez $\mathcal{H}_0$ avec un risque de deuxième espèce β.

Test unilatéral IV

À retenir

Microsoft Excel propose une fonction pour réaliser ce test : `TEST.Z`.

Test bilatéral I

Hypothèses du test

Vous souhaitez choisir entre les deux hypothèses suivantes :

$$\boxed{\mathcal{H}_0 : \mu = \mu_0}$$

contre

$$\boxed{\mathcal{H}_1 : \mu \neq \mu_0.}$$

Test bilatéral II

Conditions d'application du test

Il faut que l'échantillon $x_1, \ldots, x_n$ soit des réalisations indépendantes de la variable aléatoire X qui suit une loi normale.

Statistique du test

La variable aléatoire $Z = \dfrac{\widehat{\mu}_n - \mu_0}{\sigma/\sqrt{n}}$ suit une loi normale $\mathcal{N}(0,1)$.

Test bilatéral III

Décision et conclusion du test

La valeur critique du test, notée c_α, est lue dans la table de la loi normale centrée et réduite. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée z_{obs}, est supérieure ou égale à c_α, alors le test est significatif. Vous rejetez $\mathcal{H}_0$ et vous décidez que $\mathcal{H}_1$ est vraie avec un risque de première espèce $\alpha = 5\%$. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée z_{obs}, est strictement inférieure à c_α, alors le test n'est pas significatif. Vous conservez $\mathcal{H}_0$ avec un risque de deuxième espèces β.

Test bilatéral IV

À retenir

Microsoft Excel propose une fonction pour réaliser ce test :
`2 * MIN(TEST.Z(matrice,mu0,sigma), 1 - TEST.Z(matrice,mu0,sigma))`.

Example

Dans l'atmosphère, le taux d'un gaz nocif, pour un volume donné, suit une loi normale d'espérance μ et de variance σ^2 égale à 100. 30 prélèvements ont été effectués et les valeurs de ces 30 prélèvements sont les suivantes :

$$52,0;\ 60,2;\ 68,8;\ 46,8;\ 62,2;\ 53,5;\ 50,9;\ 44,9;\ 73,2;\ 60,4;$$
$$61,9;\ 67,8;\ 30,5;\ 52,5;\ 40,4;\ 29,6;\ 58,3;\ 62,6;\ 53,6;\ 64,6;$$
$$54,4;\ 53,8;\ 49,8;\ 57,4;\ 63,1;\ 53,4;\ 59,4;\ 48,6;\ 40,7;\ 51,9.$$

Pouvez-vous conclure, avec un risque $\alpha = 5\%$, que l'espérance μ est inférieure à 50, qui est le seuil tolérable admis ?

Partie 2 : Tests de comparaison avec une norme

4 Espérance
- Espérance d'une loi normale de variance connue
- Espérance d'une loi normale de variance inconnue

5 Variance
- Variance d'une loi normale d'espérance connue
- Variance d'une loi normale d'espérance inconnue

6 Grands échantillons
- Espérance d'une loi quelconque, cas des grands échantillons
- Variance d'une loi quelconque, cas des grands échantillons

7 Proportion
- Proportion, tirage avec remise
- Proportion, tirage sans remise

Test unilatéral I

Soit X une variable aléatoire qui suit une loi normale d'espérance μ et de variance σ^2 inconnues.

Remarque

Le test unilatéral se déduit aisément du test bilatéral.

À retenir

Microsoft Excel propose une fonction pour réaliser ce test : `TEST.STUDENT`.

Test bilatéral I

Hypothèses du test

Ce sont les mêmes que précédemment.

Conditions d'application du test

Il faut que l'échantillon $x_1, \ldots, x_n$ soit des réalisations indépendantes de la variable aléatoire X qui suit une loi normale.

Test bilatéral II

Statistique du test

La variable aléatoire $T_{n-1} = \dfrac{\widehat{\mu}_n - \mu_0}{S_{n,c}/\sqrt{n}}$ suit une loi de Student $t(n-1)$.

Test bilatéral III

Décision et conclusion du test

La valeur critique du test, notée c_α, est lue dans une table de la loi de Student. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $t_{n-1,obs}$, est supérieure ou égale à c_α, alors le test est significatif. Vous rejetez $\mathcal{H}_0$ et vous décidez que $\mathcal{H}_1$ est vraie avec un risque de première espèce $\alpha = 5\%$. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $t_{n-1,obs}$, est strictement inférieure à c_α, alors le test n'est pas significatif. Vous conservez $\mathcal{H}_0$ avec un risque de deuxième espèce β.

Test bilatéral IV

À retenir

Microsoft Excel propose une fonction pour réaliser ce test : `TEST.STUDENT`.

Test bilatéral V

Remarque

Dans le cas où la condition d'application n'est pas vérifiée, il vous faut alors utiliser un test non paramétrique : le test des signes ou le test des rangs signés de Wilcoxon. Ce dernier demande aussi une condition d'application mais moins restrictive, à savoir que la variable dont est issu l'échantillon doit être distribuée symétriquement. Notez aussi que les hypothèses associées à ce test ne sont pas les mêmes.

Test bilatéral VI

À retenir

Microsoft Excel propose une fonction pour réaliser ce test : `TEST.STUDENT`.

Example

Le jardinier aimerait savoir si les glycines blanches qu'il a plantées sur son terrain suivent bien les spécificités de la notice qu'il a reçu lorsqu'il a commandé ses graines sur internet. Il était indiqué sur la notice que chaque gousse de glycines blanches à maturité doit mesurer 15 cm de long. Comment peut-il s'assurer que les gousses qu'il a dans son jardin suivent bien cette spécifité ?

Partie 2 : Tests de comparaison avec une norme

4 Espérance
- Espérance d'une loi normale de variance connue
- Espérance d'une loi normale de variance inconnue

5 Variance
- Variance d'une loi normale d'espérance connue
- Variance d'une loi normale d'espérance inconnue

6 Grands échantillons
- Espérance d'une loi quelconque, cas des grands échantillons
- Variance d'une loi quelconque, cas des grands échantillons

7 Proportion
- Proportion, tirage avec remise
- Proportion, tirage sans remise

Test unilatéral I

Soit X une variable aléatoire qui suit une loi normale d'espérance μ connue et de variance σ^2 inconnue.

Remarque

Le test unilatéral se déduit aisément du test bilatéral.

Test bilatéral I

Hypothèses du test

Vous souhaitez choisir entre les deux hypothèses suivantes :

$$\boxed{\mathcal{H}_0 : \sigma^2 = \sigma_0^2}$$

contre

$$\boxed{\mathcal{H}_1 : \sigma^2 \neq \sigma_0^2.}$$

Test bilatéral II

Conditions d'application du test

Il faut que l'échantillon $x_1, \ldots, x_n$ soit des réalisations indépendantes de la variable aléatoire X qui suit une loi normale.

Statistique du test

La variable aléatoire $\dfrac{n\widehat{\sigma}_n^2}{\sigma^2}$ suit une loi du Khi-deux $\chi^2(n)$.

Test bilatéral III

Décision et conclusion du test

Les valeurs critiques du test sont lues dans une table de la loi du Khi-deux. Si la valeur de la statistique calculée sur l'échantillon, notée $\chi^2_{obs}(n)$, n'est pas entre ces deux valeurs alors le test est significatif. Vous rejetez $\mathcal{H}_0$ et vous décidez que $\mathcal{H}_1$ est vraie avec un risque de première espèce $\alpha = 5\%$. Si la valeur de la statistique calculée sur l'échantillon, notée $\chi^2_{obs}(n)$, est entre ces deux valeurs, alors le test n'est pas significatif. Vous conservez $\mathcal{H}_0$ avec un risque de deuxième espèce β.

Partie 2 : Tests de comparaison avec une norme

4 Espérance
- Espérance d'une loi normale de variance connue
- Espérance d'une loi normale de variance inconnue

5 Variance
- Variance d'une loi normale d'espérance connue
- Variance d'une loi normale d'espérance inconnue

6 Grands échantillons
- Espérance d'une loi quelconque, cas des grands échantillons
- Variance d'une loi quelconque, cas des grands échantillons

7 Proportion
- Proportion, tirage avec remise
- Proportion, tirage sans remise

Test unilatéral I

Soit X une variable aléatoire qui suit une loi normale d'espérance μ et de variance σ^2 inconnues.

Remarque

Le test unilatéral se déduit aisément du test bilatéral.

Test bilatéral I

Hypothèses du test

Ce sont les mêmes que précédemment.

Conditions d'application du test

Il faut que l'échantillon $x_1, \ldots, x_n$ soit des réalisations indépendantes de la variable aléatoire X qui suit une loi normale.

Statistique du test

La variable aléatoire $\dfrac{nS_{n,c}^2}{\sigma^2}$ suit une loi du Khi-deux $\chi^2(n-1)$.

Test bilatéral II

Décision et conclusion du test

Les valeurs critiques du test sont lues dans une table de la loi du Khi-deux. Si la valeur de la statistique calculée sur l'échantillon, notée $\chi^2_{obs}(n-1)$, n'est pas entre ces valeurs, alors le test est significatif. Vous rejetez $\mathcal{H}_0$ et vous décidez que $\mathcal{H}_1$ est vraie avec un risque de première espèce $\alpha = 5\%$. Si la valeur de la statistique calculée sur l'échantillon, notée $\chi^2_{obs}(n-1)$, est entre ces deux valeurs, alors le test n'est pas significatif. Vous conservez $\mathcal{H}_0$ avec un risque de deuxième espèce β.

Example

Vous venez d'acquérir dans votre laboratoire une nouvelle balance et vous souhaitez comparer la régularité du travail de cette dernière pour des très petites pesées à la norme habituelle du descriptif pour laquelle la variance σ^2 est égale à $4g^2$. Vous prélevez un échantillon d'effectif égal à 30 dont les valeurs sont données ci-dessous :

$$2,53;\ 1,51;\ 1,52;\ 1,44;\ 4,32;\ 2,36;\ 2,41;\ 2,06;\ 1,57;\ 1,68;$$
$$3,09;\ 0,54;\ 2,32;\ 0,19;\ 2,66;\ 2,20;\ 1,04;\ 1,02;\ 0,74;\ 1,01;$$
$$0,35;\ 2,42;\ 2,66;\ 1,11;\ 0,56;\ 1,75;\ 1,51;\ 3,80;\ 2,22;\ 2,88.$$

Au risque $\alpha = 5\%$, pouvez-vous considérer que la variance de l'échantillon est conforme à la norme souhaitée ?

Partie 2 : Tests de comparaison avec une norme

4 Espérance
- Espérance d'une loi normale de variance connue
- Espérance d'une loi normale de variance inconnue

5 Variance
- Variance d'une loi normale d'espérance connue
- Variance d'une loi normale d'espérance inconnue

6 Grands échantillons
- Espérance d'une loi quelconque, cas des grands échantillons
- Variance d'une loi quelconque, cas des grands échantillons

7 Proportion
- Proportion, tirage avec remise
- Proportion, tirage sans remise

Dans les tests précédents sur les espérances, il est possible de remplacer l'hypothèse de normalité par une hypothèse portant sur la taille de l'échantillon analysé. Généralement, un échantillon d'effectif supérieur ou égal à 30 permet une telle approximation.

Partie 2 : Tests de comparaison avec une norme

4 Espérance
- Espérance d'une loi normale de variance connue
- Espérance d'une loi normale de variance inconnue

5 Variance
- Variance d'une loi normale d'espérance connue
- Variance d'une loi normale d'espérance inconnue

6 Grands échantillons
- Espérance d'une loi quelconque, cas des grands échantillons
- Variance d'une loi quelconque, cas des grands échantillons

7 Proportion
- Proportion, tirage avec remise
- Proportion, tirage sans remise

Dans les tests précédents sur les variances, il est possible de remplacer l'hypothèse de normalité par une hypothèse portant sur la taille de l'échantillon analysé. Généralement, un échantillon d'effectif supérieur ou égal à 30 permet une telle approximation.

Partie 2 : Tests de comparaison avec une norme

4 Espérance
- Espérance d'une loi normale de variance connue
- Espérance d'une loi normale de variance inconnue

5 Variance
- Variance d'une loi normale d'espérance connue
- Variance d'une loi normale d'espérance inconnue

6 Grands échantillons
- Espérance d'une loi quelconque, cas des grands échantillons
- Variance d'une loi quelconque, cas des grands échantillons

7 Proportion
- Proportion, tirage avec remise
- Proportion, tirage sans remise

Test unilatéral I

Remarque

Le test unilatéral se déduit aisément du test bilatéral.

Test bilatéral I

Hypothèses du test

Vous souhaitez choisir entre les deux hypothèses suivantes :

$$\boxed{\mathcal{H}_0 : \pi_A = \pi_0}$$

contre

$$\boxed{\mathcal{H}_1 : \pi_A \neq \pi_0.}$$

Test bilatéral II

Conditions d'application du test

Il faut que l'échantillon $x_1, \ldots, x_n$ soit des réalisations indépendantes.

Statistique du test

La variable aléatoire $n\widehat{\pi}_{n,A} = n_A$ ($\widehat{\pi}_{n,A}$ a été défini dans le chapitre 6) suit une loi binomiale $\mathcal{B}(n, \pi_0)$.

Test bilatéral III

Décision et conclusion du test

La valeur critique du test, notée c_α, est lue dans une table de la loi binomiale. Si la valeur de la statistique calculée sur l'échantillon, notée $n_A(obs)$, est supérieure ou égale à c_α, alors le test est significatif. Vous rejetez $\mathcal{H}_0$ et vous décidez que $\mathcal{H}_1$ est vraie avec un risque de première espèce $\alpha = 5\%$. Si la valeur de la statistique calculée sur l'échantillon, notée $n_A(obs)$, est strictement inférieure à c_α, alors le test n'est pas significatif. Vous conservez $\mathcal{H}_0$ avec un risque de deuxième espèce β.

Test bilatéral IV

Vous pouvez trouver dans la littérature une variation de ce test, qui va être présentée ci-dessous, avec cette fois-ci d'autres conditions d'application et une autre statistique que celle présentée précedemment. Cette variation utilise l'approximation de la loi binomiale par une loi normale. Cette approximation a longtemps été utilisée mais avec des logiciels comme `R`, il est préférable d'avoir recours au test exact présenté ci-dessus.

Test unilatéral I

Remarque

Le test unilatéral se déduit aisément du test bilatéral.

Test bilatéral I

Hypothèses du test

Ce sont les mêmes que précédemment.

Conditions d'application du test

Il faut que l'échantillon $x_1, \ldots, x_n$ soit des réalisations indépendantes. De plus, il faut que les trois inégalités $n \geqslant 50$, $n\pi_0 \geqslant 16$ et $n(1-\pi_0) \geqslant 16$ soient vérifiées.

Test bilatéral II

Statistique du test

La variable aléatoire $Z = \dfrac{\widehat{\pi}_{n,A} - \pi_0}{\sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}}$ suit une loi normale $\mathcal{N}(0,1)$.

Test bilatéral III

Décision et conclusion du test

La valeur critique du test, notée c_α, est lue dans la table de la loi normale centrée et réduite. Si la valeur de la statistique calculée sur l'échantillon, notée z_{obs}, est supérieure ou égale à c_α, alors le test est significatif. Vous rejetez $\mathcal{H}_0$ et vous décidez que $\mathcal{H}_1$ est vraie avec un risque de première espèce $\alpha = 5\%$. Si la valeur de la statistique calculée sur l'échantillon, notée z_{obs}, est strictement inférieure ou égale à c_α, alors le test n'est pas significatif. Vous conservez $\mathcal{H}_0$ avec un risque de deuxième espèce β.

Example

Dans le « Ouest-France » du samedi 23 janvier 2010, vous pouvez lire :
« Plus de garçons que de filles ! Avec 507 bébés mâles comptabilisés à Saint-Lô en 2009, contre 481 fillettes, les naissances masculines sont toujours plus nombreuses. »
Pensez-vous que les garçons sont plus nombreux significativement que les filles, avec un risque $\alpha = 5\%$?

Partie 2 : Tests de comparaison avec une norme

4 Espérance
- Espérance d'une loi normale de variance connue
- Espérance d'une loi normale de variance inconnue

5 Variance
- Variance d'une loi normale d'espérance connue
- Variance d'une loi normale d'espérance inconnue

6 Grands échantillons
- Espérance d'une loi quelconque, cas des grands échantillons
- Variance d'une loi quelconque, cas des grands échantillons

7 Proportion
- Proportion, tirage avec remise
- Proportion, tirage sans remise

Si l'effectif de l'échantillon prélevé est inférieur à 10% de l'effectif total de la population, il est possible de considérer que le tirage a lieu sans remise et d'utiliser les résultats précédents.

Troisième partie III

Tests de comparaison entre deux populations indépendantes

Partie 3 : Tests de comparaison entre deux populations indépendantes

8 Espérances
- Espérances de deux lois normales de variances connues
- Espérances de deux lois normales de même variance inconnue
- Espérances de deux lois normales de variances différentes inconnues

9 Variances
- Variances de deux lois normales d'espérances inconnues

10 Grands échantillons
- Espérances de deux lois quelconques, cas des grands échantillons

11 Proportions
- Égalité de deux proportions, tirage avec remise
- Égalité de deux proportions, tirage sans remise

Test unilatéral I

Soit X une variable aléatoire qui suit une loi normale $\mathcal{N}(\mu_1, \sigma_1)$ et Y une variable aléatoire qui suit une loi normale $\mathcal{N}(\mu_2, \sigma_2)$ avec σ_1 et σ_2 connues.

Remarque

Le test unilatéral se déduit aisément du test bilatéral.

Test bilatéral I

Hypothèses du test

Vous souhaitez choisir entre les deux hypothèses suivantes :

$$\boxed{\mathcal{H}_0 : \mu_1 = \mu_2}$$

contre

$$\boxed{\mathcal{H}_1 : \mu_1 \neq \mu_2.}$$

Test bilatéral II

Conditions d'application du test

Il faut que l'échantillon $x_1, \ldots, x_n$ soit des réalisations indépendantes de la variable aléatoire X qui suit une loi normale et que le second échantillon aléatoire $y_1, \ldots, y_{n_2}$ soit aussi des réalisations indépendantes de la variable aléatoire Y qui suit une loi normale. De plus, les effectifs n_1 et n_2 peuvent ne pas être égaux.

Test bilatéral III

Statistique du test

La variable aléatoire $Z = \dfrac{\widehat{\mu}_{n_1} - \widehat{\mu}_{n_2}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ suit une loi normale $\mathcal{N}(0,1)$.

Test bilatéral IV

Décision et conclusion du test

La valeur critique du test, notée c_α, est lue dans la table de la loi normale centrée et réduite. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée z_{obs} est supérieure ou égale à c_α, alors le test est significatif. Vous rejetez $\mathcal{H}_0$ et vous décidez que $\mathcal{H}_1$ est vraie avec un risque de première espèce α. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée z_{obs} est strictement inférieure à c_α, alors le test n'est pas significatif. Vous conservez $\mathcal{H}_0$ avec un risque de deuxième espèce β.

Test bilatéral V

À retenir

Pour réaliser ce test sous Excel, il est possible d'utiliser la fonction `LOI.Z`.

Partie 3 : Tests de comparaison entre deux populations indépendantes

8 Espérances
- Espérances de deux lois normales de variances connues
- Espérances de deux lois normales de même variance inconnue
- Espérances de deux lois normales de variances différentes inconnues

9 Variances
- Variances de deux lois normales d'espérances inconnues

10 Grands échantillons
- Espérances de deux lois quelconques, cas des grands échantillons

11 Proportions
- Égalité de deux proportions, tirage avec remise
- Égalité de deux proportions, tirage sans remise

Test unilatéral I

Remarque

Le test unilatéral se déduit aisément du test bilatéral.

Test bilatéral I

Hypothèses du test

Ce sont les mêmes que précédemment.

Conditions d'application du test

Il faut que l'échantillon $x_1, \ldots, x_n$ soit des réalisations indépendantes de la variable aléatoire X qui suit une loi normale et que le second échantillon aléatoire $y_1, \ldots, y_{n_2}$ soit aussi des réalisations indépendantes de la variable aléatoire Y qui suit une loi normale. De plus, les effectifs n_1 et n_2 peuvent ne pas être égaux. Maintenant il faut distinguer deux cas : soit $\sigma_1^2 = \sigma_2^2$, soit $\sigma_1^2 \neq \sigma_2^2$.

Test bilatéral II

Le test de Student.

Cas où $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

Une question que vous devez vous poser est : « comment vérifier cette hypothèse ? ». Pour tester cette hypothèse , vous allez utiliser le test de Fisher-Snedecor qui est présenté plus loin.

Test bilatéral III

Statistique du test

La variable aléatoire $T_{n_1+n_2-2} = \dfrac{\widehat{\mu}_{n_1} - \widehat{\mu}_{n_2}}{\widehat{\sigma}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ où

$\widehat{\sigma} = \sqrt{\dfrac{n_1 S_{n_1}^2 + n_2 S_{n_2}^2}{n_1 + n_2 - 2}}$ suit une loi de Student $t(n_1 + n_2 - 2)$.

Test bilatéral IV

Décision et conclusion du test

La valeur critique du test, notée c_α, est lue dans une table de la loi de Student. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $t_{n_1+n_2-2,obs}$ est supérieure ou égale à c_α, alors le test est significatif. Vous rejetez $\mathcal{H}_0$ et vous décidez que $\mathcal{H}_1$ est vraie avec un risque de première espèce α. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $t_{n_1+n_2-2,obs}$ est strictement inférieure à c_α, alors le test n'est pas significatif. Vous conservez $\mathcal{H}_0$ avec un risque de deuxième espèce β.

Test bilatéral V

À retenir

La fonction sous Excel qui permet de réaliser ce test est la fonction `TEST.STUDENT`.

Partie 3 : Tests de comparaison entre deux populations indépendantes

8 Espérances
- Espérances de deux lois normales de variances connues
- Espérances de deux lois normales de même variance inconnue
- Espérances de deux lois normales de variances différentes inconnues

9 Variances
- Variances de deux lois normales d'espérances inconnues

10 Grands échantillons
- Espérances de deux lois quelconques, cas des grands échantillons

11 Proportions
- Égalité de deux proportions, tirage avec remise
- Égalité de deux proportions, tirage sans remise

Test unilatéral I

Remarque

Le test unilatéral se déduit aisément du test bilatéral.

Test bilatéral I

Hypothèses du test

Ce sont les mêmes que précédemment.

Conditions d'application du test

Il faut que l'échantillon $x_1, \ldots, x_n$ soit des réalisations indépendantes de la variable aléatoire X qui suit une loi normale et que le second échantillon aléatoire $y_1, \ldots, y_{n_2}$ soit aussi des réalisations indépendantes de la variable aléatoire Y qui suit une loi normale. De plus, les effectifs n_1 et n_2 peuvent ne pas être égaux. Maintenant il faut distinguer deux cas : soit $\sigma_1^2 = \sigma_2^2$, soit $\sigma_1^2 \neq \sigma_2^2$.

Test bilatéral II

Le test de Welch

Cas où $\sigma_1^2 \neq \sigma_2^2$.

Une question que vous devez vous poser est : « comment vérifier cette hypothèse ? ». Pour tester cette hypothèse , vous allez utiliser le test de Fisher-Snedecor qui est présenté ci-dessous.

Test bilatéral III

Statistique du test

La variable aléatoire $T_\nu = \dfrac{\widehat{\mu}_{n_1} - \widehat{\mu}_{n_2}}{\sqrt{\dfrac{S_{n_1}^2}{n_1 - 1} + \dfrac{S_{n_2}^2}{n_2 - 1}}}$ suit une loi de Student $t(\nu)$, où ν est l'entier le plus proche de

$$\frac{\left(\dfrac{S_{n_1}^2}{n_1 - 1} + \dfrac{S_{n_2}^2}{n_2 - 1}\right)^2}{\dfrac{S_{n_1}^4}{(n_1 - 1)n_1^2} + \dfrac{S_{n_2}^4}{(n_2 - 1)n_2^2}}.$$

Test bilatéral IV

Décision et conclusion du test

La valeur critique du test, notée c_α, est lue dans une table de la loi de Student. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $t_{\nu,obs}$ est supérieure ou égale à c_α, alors le test est significatif. Vous rejetez $\mathcal{H}_0$ et vous décidez que $\mathcal{H}_1$ est vraie avec un risque de première espèce α. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $t_{\nu,obs}$ est strictement inférieure à c_α, alors le test n'est pas significatif. Vous conservez $\mathcal{H}_0$ avec un risque de deuxième espèce β.

Test bilatéral V

À retenir

La fonction sous Excel qui permet de réaliser ce test est la fonction `TEST.STUDENT`.

Test bilatéral VI

Remarque

Le test de Student est assez robuste mais si vous vous éloignez trop de la condition de normalité qui est nécessaire pour procéder à ce test, il est préférable d'utiliser un test non paramétrique, comme par exemple le test de Mann-Whitney-Wilcoxon qui a aussi certaines conditions d'application à savoir l'échantillon $X_1, \ldots, X_n$ est aléatoire et les variables X_i sont indépendantes. De plus, les deux distributions des deux variables doivent avoir la même forme. Cette dernière condition est très importante. En effet, dans le cas où elle ne serait pas respectée, le test de Mann-Whitney-Wilcoxon peut aboutir à des conclusions inexactes. Il vous est donc conseillée de faire une analyse graphique des deux distributions avant d'appliquer ce test.

Le jus d'orange contre l'acide ascorbique I

Example

D'après Couty-Fredon, Debord, Fredon, « Mini Manuel de Probabilités et Statistique », aux éditions Dunod, 2007.
Pour comparer l'effet de la vitamine C du jus d'orange et de l'acide ascorbique de synthèse, on a donné, pendant 6 semaines, du jus d'orange à un groupe de 10 cobayes et de la vitamine de synthèse à un groupe de 10 autres cobayes. Puis on a mesuré la longueur des odontoblastes des incisives.

Le jus d'orange contre l'acide ascorbique II

Example

Les résultats suivants ont été relevés :

1. jus d'orange :

$$8,2;\ 9,4;\ 9,6;\ 9,7;\ 10,0;\ 14,5;\ 15,2;\ 16,1;\ 17,6;21,5;$$

2. acide ascorbique

$$4,2;\ 5,2;\ 5,8;\ 6,4;\ 7,0;\ 7,3;\ 10,1;\ 11,2;\ 11,3;\ 11,5.$$

Testez, au risque $\alpha = 5\%$, l'hypothèse $\mathcal{H}_0$ « l'effet des deux produits est le même » contre $\mathcal{H}_1$ « l'effet d'orange accélère la croissance ».

Traitement pour la pression artérielle I

Example

D'après Couty-Fredon, Debord, Fredon, « Mini Manuel de Probabilités et Statistique », aux éditions Dunod, 2007.
Chez un groupe de 10 sujets, les effets d'un traitement destiné à diminuer la pression artérielle ont été expérimentés. Les résultats (valeur de la tension artérielle systolique en cmHg) ont été relevés sur les 10 sujets et sont présentés dans le tableau ci-dessous.

Traitement pour la pression artérielle II

Example

Sujet n°	1	2	3	4	5	6	7	8	9	10
Avant traitement	15	18	17	20	21	18	17	15	19	16
Après traitement	12	16	17	18	17	15	18	14	16	18

Le traitement a-t-il une action significative, au risque $\alpha = 5\%$?

Partie 3 : Tests de comparaison entre deux populations indépendantes

8 Espérances
- Espérances de deux lois normales de variances connues
- Espérances de deux lois normales de même variance inconnue
- Espérances de deux lois normales de variances différentes inconnues

9 Variances
- Variances de deux lois normales d'espérances inconnues

10 Grands échantillons
- Espérances de deux lois quelconques, cas des grands échantillons

11 Proportions
- Égalité de deux proportions, tirage avec remise
- Égalité de deux proportions, tirage sans remise

Test unilatéral I

Remarque

Le test unilatéral se déduit aisément du test bilatéral.

Test bilatéral I

Hypothèses du test

Vous souhaitez choisir entre les deux hypothèses suivantes :

$$\boxed{\mathcal{H}_0 : \sigma_1^2 = \sigma_2^2}$$

contre

$$\boxed{\mathcal{H}_1 : \sigma_1^2 \neq \sigma_2^2.}$$

Test bilatéral II

Conditions d'application du test

Il faut que l'échantillon $x_1, \ldots, x_n$ soit des réalisations indépendantes de la variable aléatoire X qui suit une loi normale et que le second échantillon aléatoire $y_1, \ldots, y_{n_2}$ soit aussi des réalisations indépendantes de la variable aléatoire Y qui suit une loi normale. De plus n_1 et n_2 sont pas forcément égaux.

Test bilatéral III

Statistique du test

La variable aléatoire $F = \frac{S^2_{n_1,c}}{S^2_{n_2,c}}$, où $S^2_{n_1,c} = \frac{n_1 S^2_{n_1}}{n_1 - 1}$ et $S^2_{n_2,c} = \frac{n_2 S^2_{n_2}}{n_2 - 1}$, suit une loi de Fisher $F(n_1 - 1, n_2 - 1)$.

Test bilatéral IV

Décision et conclusion du test

La valeur critique du test, notée c_α, est lue dans la table de la loi de Fisher.
Si la valeur de la statistique calculée sur l'échantillon, notée f_{obs} est supérieure ou égale à c_α, alors le test est significatif. Vous rejetez $\mathcal{H}_0$ et vous décidez que $\mathcal{H}_1$ est vraie avec un risque de première espèce α. Si la valeur de la statistique calculée sur l'échantillon, notée f_{obs} est strictement inférieure à c_α, alors le test n'est pas significatif. Vous conservez $\mathcal{H}_0$ avec un risque de deuxième espèce β.

Test bilatéral V

À retenir

La fonction sous Excel qui permet de réaliser ce test est la fonction `TEST.F`.

Test bilatéral VI

Remarque

Vous pouvez constater qu'il manque un cas : comparaison de deux variances de lois normales d'espérances connues. Ce cas est extrêment rare dans la pratique. C'est pourquoi il ne sera pas exposé dans le cadre de cet ouvrage. Si jamais vous devez l'utiliser, il vous faudra donc le faire vous-même avec une ligne de commande en vous inspirant du cas où les espérances sont inconnues.

Example

Cet exemple provient de *Couty-Fredon, Debord, Fredon « Manuel de probabilités et statistique », aux éditions Dunod, 2007.*
Le Coucou est un oiseau qui fait couver ses oeufs par des oiseaux d'autres espèces que la sienne, de tailles très différentes. Une hypothèse a été émise comme quoi le Coucou puisse adapter la taille de ses oeufs à la taille du nid dans lequel il pond. Une étude faite par le biologiste O.H. Latter et publiée sous le titre « The Egg Of Cuculus Canorus : An Enquiry Into The Dimensions Of The Cuckoo'S Ego And The Relation Of The Variations To The Size Of The Eggs Of The Foster-Parent, With Notes On Coloration. » dans la revue *Biometrica, 1902, Volume 1, pages 164-176* sur la taille des oeufs déposés dans les nids de petite taille (Roitelet) ou de plus grande taille (Fauvette) a donné les valeurs (en millimètres) suivantes :

Example

1. Dans les nids de Roitelet : 19,5 ; 22,1 ; 21,5 ; 20,9 ; 22,0 ; 21,0 ; 22,3 ; 21,0 ; 20,3 ; 20,9 ; 22,0 ; 22,0 ; 20,8 ; 21,2 ; 21,0.
2. Dans les nids de Fauvette : 22,0 ; 23,9 ; 20,9 ; 23,8 ; 25,0 ; 24,0 ; 23,8 ; 21,7 ; 22,8 ; 23,1 ; 23,5 ; 23,0 ; 23,0 ; 23,1.

Deux questions vous sont alors posées :

1. Pouvez-vous dire, au risque $\alpha = 5\%$, que les deux populations ont la même variance ?
2. Tester également l'hypothèse comme quoi le Coucou adapte la taille de ses oeufs à la taille du nid dans lequel il pond.

Partie 3 : Tests de comparaison entre deux populations indépendantes

8 Espérances
- Espérances de deux lois normales de variances connues
- Espérances de deux lois normales de même variance inconnue
- Espérances de deux lois normales de variances différentes inconnues

9 Variances
- Variances de deux lois normales d'espérances inconnues

10 Grands échantillons
- Espérances de deux lois quelconques, cas des grands échantillons

11 Proportions
- Égalité de deux proportions, tirage avec remise
- Égalité de deux proportions, tirage sans remise

Test unilatéral I

Remarque

Le test unilatéral se déduit aisément du test bilatéral.

Test bilatéral I

Hypothèses du test

Ce sont les mêmes que précédemment.

Conditions d'application du test

Les effectifs n_1 et n_2 sont tous les deux supérieurs à 30.

Test bilatéral II

Statistique du test

La variable aléatoire $Z = \dfrac{\widehat{\mu}_{n_1} - \widehat{\mu}_{n_2}}{\sqrt{\dfrac{S_{n_1}^2}{n_1 - 1} + \dfrac{S_{n_2}^2}{n_2 - 1}}}$ suit approximativement une loi normale $\mathcal{N}(0, 1)$.

Test bilatéral III

Décision et conclusion du test

La valeur critique du test, notée c_α, est lue dans la table de la loi normale centrée et réduite.
Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée z_{obs} est supérieure ou égale à c_α, alors le test est significatif. Vous rejetez $\mathcal{H}_0$ et vous décidez que $\mathcal{H}_1$ est vraie avec un risque de première espèce α. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée z_{obs} est strictement inférieure à c_α, alors le test n'est pas significatif. Vous conservez $\mathcal{H}_0$ avec un risque de deuxième espèce β.

Test bilatéral IV

À retenir

Il n'existe pas de fonction sous Excel qui réalise ce test mais il est facile de le réaliser avec la fonction `LOI.Z`.

Partie 3 : Tests de comparaison entre deux populations indépendantes

8 Espérances
- Espérances de deux lois normales de variances connues
- Espérances de deux lois normales de même variance inconnue
- Espérances de deux lois normales de variances différentes inconnues

9 Variances
- Variances de deux lois normales d'espérances inconnues

10 Grands échantillons
- Espérances de deux lois quelconques, cas des grands échantillons

11 Proportions
- Égalité de deux proportions, tirage avec remise
- Égalité de deux proportions, tirage sans remise

Dans deux populations, un caractère à deux modalités (A et $\overline{A}$) est observé. Chaque individu présente ou non la modalité A. Les fréquences d'apparition de A dans les deux populations sont les nombres inconnus $\pi_{A,1}$ et $\pi_{A,2}$. Quelques notations vont être introduites :

1. $n_{A,1}$ le nombre de personnes de l'échantillon 1 qui présentent la modalité A ;
2. $n_{A,2}$ le nombre de personnes de l'échantillon 2 qui présentent la modalité A ;
3. $n_{\overline{A},1}$ le nombre de personnes de l'échantillon 1 qui ne présentent pas la modalité A ;

4. $n_{\overline{A},2}$ le nombre de personnes de l'échantillon 2 qui ne présentent pas la modalité A ;
5. $n_{\bullet,1}$ le nombre de personnes présentent dans l'échantillon 1 ;
6. $n_{\bullet,2}$ le nombre de personnes présentent dans l'échantillon 2 ;
7. $n_{A,\bullet}$ le nombre de personnes qui présentent la modalité A dans l'ensemble des deux échantillons ;
8. $n_{\overline{A},\bullet}$ le nombre de personnes qui ne présentent pas la modalité A dans l'ensemble des deux échantillons ;
9. $n = n_{\bullet,1} + n_{\bullet,2}$ la somme des deux effectifs des échantillons.

Test unilatéral I

Remarque

Le test unilatéral se déduit aisément du test bilatéral.

Test bilatéral I

Hypothèses du test

Vous souhaitez choisir entre les deux hypothèses suivantes :

$$\boxed{\mathcal{H}_0 : \pi_{A,1} = \pi_{A,2}}$$

contre

$$\boxed{\mathcal{H}_1 : \pi_{A,1} \neq \pi_{A,2}.}$$

Conditions d'application du test

Les effectifs n_1 et n_2 peuvent ne pas être égaux.

Test bilatéral II

Statistique du test

La variable aléatoire $n_{A,1}$ suit une loi hypergéométrique $\mathcal{H}\left(n, n_{\bullet,1}, \frac{n_{A,\bullet}}{n}\right)$.

Décision et conclusion du test

Il faut raisonner ici avec la p-valeur. Donc il faut avoir recours pour ce test à une table spécifique.

Test bilatéral III

À retenir

Il n'y a pas de fonction qui permette de réaliser ce test sous Excel, il est néanmoins possible d'y parvenir avec une table.

Vous pouvez trouver dans la littérature une variation de ce test, qui va être présentée ci-dessous, avec cette fois-ci d'autres conditions d'application et une autre statistique que celle présentée précedemment. Cette variation utilise l'approximation de la loi hypergéométrique par une loi normale. Cette approximation a longtemps été utilisée mais avec certains logiciels comme R, il est préférable d'avoir recours au test exact de Fisher présenté ci-dessus.

Test unilatéral I

Remarque

Le test unilatéral se déduit aisément du test bilatéral.

Test bilatéral I

Hypothèses du test

Ce sont les mêmes que précédemment.

Conditions d'application du test

Les effectifs n_1 et n_2 ne sont pas forcément égaux. De plus, chaque case du tableau de contingence doit présenter un effectif théorique supérieur ou égal à 5.

Test bilatéral II

Statistique du test

La variable aléatoire $Z = \dfrac{n_{A,1} - \dfrac{n_{A,\bullet} \times n_{\bullet,1}}{n}}{\sqrt{\dfrac{n_{\bullet,1} \times n_{\bullet,2} \times n_{A,\bullet} \times n_{\overline{A},\bullet}}{n^2(n-1)}}}$ suit une loi normale $\mathcal{N}(0,1)$.

Test bilatéral III

Décision et conclusion du test

La valeur critique du test, notée c_α, est lue dans la table de la loi normale centrée et réduite. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée z_{obs}, est supérieure ou égale à c_α, alors le test est significatif. Vous rejetez $\mathcal{H}_0$ et vous décidez que $\mathcal{H}_1$ est vraie avec un risque de première espèce $\alpha = 5\%$. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée z_{obs}, est strictement inférieure ou égale à c_α, alors le test n'est pas significatif. Vous conservez $\mathcal{H}_0$ avec un risque β.

Test bilatéral IV

À retenir

La fonction sous Excel qui permet de réaliser ce test est la fonction `LOI.Z`.

Partie 3 : Tests de comparaison entre deux populations indépendantes

8 Espérances
- Espérances de deux lois normales de variances connues
- Espérances de deux lois normales de même variance inconnue
- Espérances de deux lois normales de variances différentes inconnues

9 Variances
- Variances de deux lois normales d'espérances inconnues

10 Grands échantillons
- Espérances de deux lois quelconques, cas des grands échantillons

11 Proportions
- Égalité de deux proportions, tirage avec remise
- Égalité de deux proportions, tirage sans remise

Si l'effectif des deux échantillons prélevés est inférieur à 10% de l'effectif total de chaque population, il est possible de considérer que le tirage a lieu sans remise et d'utiliser les résultats précédents.

Quatrième partie IV

Tests à deux échantillons appariés

Partie 4 : Tests à deux échantillons appariés

12 Espérances
- Comparaison des espérances si la différence est normale

13 Grands échantillons
- Comparaison des espérances, cas des grands échantillons

Definition

Deux populations sont appariées (associées par paires) lorsque chaque valeur $x_{i,1}$ est associée à une valeur de $x_{i,2}$.
Soit μ_1 la moyenne dans la première population et μ_2 la moyenne dans la seconde population.
Soit D_i la différence entre $X_{i,1}$ et $X_{i,2}$ et D la variable aléatoire qui suit une loi normale $\mathcal{N}(\mu_D, \sigma_D)$.

Remarque

Il est équivalent de chercher à savoir si $\mu_1 = \mu_2$ ou de déterminer si $\mu_D = \mu_1 - \mu_2 = 0$.

Test unilatéral I

Remarque

Le test unilatéral se déduit aisément du test bilatéral.

Test bilatéral I

Hypothèses du test

Vous souhaitez choisir entre les deux hypothèses

$$\boxed{\mathcal{H}_0 : \mu_1 = \mu_2 \text{ (ou } \mu_D = 0)}$$

contre

$$\boxed{\mathcal{H}_1 : \mu_1 \neq \mu_2 \text{ (ou } \mu_D \neq 0).}$$

Test bilatéral II

Conditions d'application du test

Il faut que l' échantillon $d_1, \ldots, d_n$ soit des réalisations indépendantes de la variable aléatoire D qui suit une loi normale.

Statistique du test

La variable aléatoire $T_{n-1} = \frac{\widehat{\mu}_D}{S_{D,c}/\sqrt{n}}$ suit une loi de Student $t(n-1)$, où $\widehat{\mu}_D = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} D_i$ et $S_{D,c}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(D_i - \widehat{\mu}_D)^2$.

Test bilatéral III

Décision et conclusion du test

La valeur critique du test, notée c_α, est lue dans une table de la loi de Student. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $t_{n-1,obs}$ est supérieure ou égale à c_α, alors le test est significatif. Vous rejetez $\mathcal{H}_0$ et vous décidez que $\mathcal{H}_1$ est vraie avec un risque de première espèce α. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $t_{n-1,obs}$ est strictement inférieure à c_α, alors le test n'est pas significatif. Vous conservez $\mathcal{H}_0$ avec un risque de deuxième espèce β.

Test bilatéral IV

À retenir

La fonction sous Excel qui permet de réaliser ce test est la fonction `TEST.STUDENT`.

Partie 4 : Tests à deux échantillons appariés

12 Espérances
- Comparaison des espérances si la différence est normale

13 Grands échantillons
- Comparaison des espérances, cas des grands échantillons

Lorsque l'effectif des deux échantillons est important, supérieur ou égal à 30, il est possible d'utiliser le test précédent sans que l'hypothèse de normalité soit vérifiée.