

Méthodes d'évaluation expérimentale

A. Gouaïch¹

¹Département d'Informatique
Université de Montpellier

Module Serious Games

Sommaire

- 1 Introduction
 - Objectifs du cours
 - Prérequis
- 2 Section Statistique descriptive et statistique inférentielle
- 3 Les différents types d'études expérimentales
- 4 Les variables
- 5 Mesures de la tendance centrale et la variabilité
- 6 Normalisation des mesures avec Z-score
- 7 Distribution des moyennes des échantillons
- 8 Test de Student
- 9 Mesures répétées
 - Exemple

Lignes directrices

- 1 Introduction
 - Objectifs du cours
 - Prérequis
- 2 Section Statistique descriptive et statistique inférentielle
- 3 Les différents types d'études expérimentales
- 4 Les variables
- 5 Mesures de la tendance centrale et la variabilité
- 6 Normalisation des mesures avec Z-score
- 7 Distribution des moyennes des échantillons
- 8 Test de Student
- 9 Mesures répétées
 - Exemple

Les Objectifs du Cours

Les sous-titres sont facultatifs.

- Savoir conduire des évaluations de vos jeux (playtest)
- Savoir collecter, résumer, interpréter et présenter les résultats des expérimentations

Les Objectifs du Cours

Les sous-titres sont facultatifs.

- Savoir conduire des évaluations de vos jeux (playtest)
- Savoir collecter, résumer, interpréter et présenter les résultats des expérimentations

Les Objectifs du Cours

Les sous-titres sont facultatifs.

- Savoir conduire des évaluations de vos jeux (playtest)
- Savoir collecter, résumer, interpréter et présenter les résultats des expérimentations

Lignes directrices

- 1 Introduction
 - Objectifs du cours
 - **Prérequis**
- 2 Section Statistique descriptive et statistique inférentielle
- 3 Les différents types d'études expérimentales
- 4 Les variables
- 5 Mesures de la tendance centrale et la variabilité
- 6 Normalisation des mesures avec Z-score
- 7 Distribution des moyennes des échantillons
- 8 Test de Student
- 9 Mesures répétées
 - Exemple

Définitions

Définition

Les statistiques offrent un ensemble de procédures et de méthodes mathématiques pour **organiser**, **résumer** et **interpréter** de l'**information**.

Définitions

La population est l'ensemble des tous les individus qui font l'objet de l'étude

Définition

Un échantillon sous-ensemble d'individus issus de la population et qui la représentent.

Définition

Une variable est une propriété/caractéristique ou une condition qui possède plusieurs valeurs selon les individus

Définition

Une donnée est la valeur mesurée d'une variable sur un individu

Définition

Un paramètre est une valeur numérique attachée à la population. Généralement, cette valeur est dérivée des données de la population.

Définition

Une statistique, est une valeur numérique attachée à un échantillon. Généralement cette valeur est dérivée des données de l'échantillon.

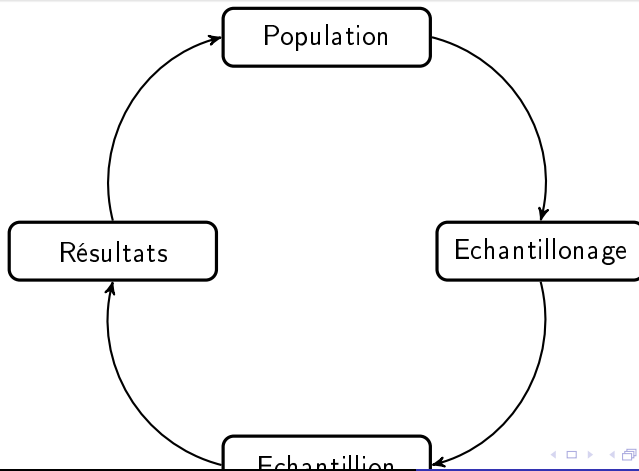
Statistiques descriptives

- Les statistiques descriptives offrent un ensemble de procédure pour résumer/ organiser/ simplifier les données
- Exemple de paramètres/statistiques étudiés : moyenne, écart type, mode, médiane, min/max, intervalle de confiance
- Statistique inférentielle : nous offre un ensemble de procédures qui permettent d'étudier des échantillons et de généraliser les résultats obtenus à la population.
- Erreur d'échantillonnage :
 - C'est la différence qui va exister entre la statistique d'un échantillon et le paramètre de la population.
 - Cette différence est inhérente et naturelle pour un processus d'échantillonnage.

Statistiques descriptives

- Les statistiques descriptives offrent un ensemble de procédure pour résumer/ organiser/ simplifier les données
- Exemple de paramètres/statistiques étudiés : moyenne, écart type, mode, médiane, min/max, intervalle de confiance
- Statistique inférentielle : nous offre un ensemble de procédures qui permettent d'étudier des échantillons et de généraliser les résultats obtenus à la population.
- Erreur d'échantillonnage :
 - C'est la différence qui va exister entre la statistique d'un échantillon et le paramètre de la population.
 - Cette différence est inhérente et naturelle pour un processus d'échantillonnage.

Picture



Utilisation des stats dans une étude expérimentale

Schéma général d'utilisation des statistiques dans une étude expérimentale

- ① Etape 1 : fixer la population et établir l'objectif de l'étude
 - ① Avoir des échantillons de la population
 - ② Récolter les données
- ② Etape 2 : Statistique descriptive
 - ① Calculer les statistiques descriptives (moyenne, variance, déviation standard)
- ③ Etape 3 : Statistique inférentielle
 - ① Nous devons répondre à la question suivante avec un certain degré de confiance :
 - la différence des statistiques des échantillons est causée par :
 - l'erreur d'échantillonnage

Etudier une variable individuellement

- Objectif : mesurer un paramètre de la population en utilisant une statistique de l'échantillon
- Déduire une moyenne μ de la population à partir de la moyenne d'un échantillon M
- Estimer la variance de la population à partir de la variance de l'échantillon
- Estimer un intervalle de confiance dans lequel se trouve la moyenne de la population à partir de la moyenne d'un échantillon

Étudier les relations entre des variables

Étudier la corrélation entre deux variables

- mesurer deux variables par individu
- calculer l'indice de corrélation entre ces deux variables
- attention la corrélation ne signifie pas causalité
- Pour déduire la causalité : méthode expérimentale

Étudier les relations de causalité entre 2 (ou plus) variables

Deux méthodes sont disponibles dans cette catégorie :

- ① méthode expérimentale
- ② méthode non expérimentale (quasi expérimentale)

Présentation de la méthode expérimentale

- L'objectif est de démontrer une relation de cause à effet entre deux variables
- Le changement de valeur d'une variable va causer/produire des changements sur une autre variable
- Exemple : force et accélération en physique.
- Deux propriétés importantes des méthodes expérimentales :
 - manipulation : pouvoir manipuler une variable en observant une autre
 - exemple : manipuler la force à appliquer sur un objet avec un ressort
 - le contrôle : nous devons assurer que tous les paramètres externes pouvant influencer les données sont contrôlés durant l'expérimentation afin de minimiser leurs effets

Présentation de la méthode expérimentale

- L'objectif est de démontrer une relation de cause à effet entre deux variables
- Le changement de valeur d'une variable va causer/produire des changements sur une autre variable
- Exemple : force et accélération en physique.
- Deux propriétés importantes des méthodes expérimentales :
 - manipulation : pouvoir manipuler une variable en observant une autre
 - exemple : manipuler la force à appliquer sur un objet avec un ressort
 - le contrôle : nous devons assurer que tous les paramètres externes pouvant influencer les données sont contrôlés durant l'expérimentation afin de minimiser leurs effets

Présentation de la méthode non expérimentale (quasi expérimentale)

- Deux sous catégories :
 - étude sur des groupes non équivalents (homme vs femme ; tranches d'âge)
 - Pré – Post traitement
- Ici les conditions expérimentales rendent le contrôle plus difficile.
- D'autres paramètres peuvent être à l'origine des effets mesurés.
- Par exemple : la différence inhérente entre les groupes ; le temps qui s'écoule entre les mesures.
- Nous allons utiliser pour ces méthodes les mêmes outils et la même terminologie que pour les méthodes expérimentales.

Variable discrète/continue

Définition

variable discrète : valeurs distinctes et séparées sans valeurs intermédiaires entre deux variables.

Définition

variable continues : il y a une infinité de valeurs entre les valeurs observées.

Variable discrète/continue

Définition

variable discrète : valeurs distinctes et séparées sans valeurs intermédiaires entre deux variables.

Définition

variable continues : il y a une infinité de valeurs entre les valeurs observées.

Les échelles de mesure

- L'échelle nominale : ensemble simple de valeurs distinctes
- l'échelle ordinale : un ensemble ordonné de valeurs
- l'échelle intervalle : ensemble ordonné avec une distance entre les valeurs
- l'échelle ratio : une échelle intervalle où le zéro signifie l'absence de la quantité mesurée.

La moyenne

- Distinguer la moyenne de la population et la moyenne de l'échantillon
- Moyenne de population : $\mu = \frac{\sum_{i=1}^N X_i}{N}$
- moyenne d'échantillon : $M = \frac{\sum_{i=1}^n X_i}{n}$

Mediane

- La mediane : point de mesure où 50 % des valeurs sont inférieures
- Le mode : c'est le score avec la plus grande fréquence

La variabilité

- La variabilité mesure la dispersion autour de la moyenne

Pour la population

Définition

déviati3n d'une mesure c'est la différence par rapport à la moyenne
 $d_i = X_i - \mu$

Définition

La variance est la moyenne des déviations au carré (SS : sum of squares)

$$\frac{\sum (d_i)^2}{N} = \frac{SS}{N}$$

Définition

La déviation standard $\sigma = \sqrt{\text{variance}} = \sqrt{\frac{SS}{N}}$

Pour la population

Définition

déviatiun d'une mesure c'est la différence par rapport à la moyenne
 $d_i = X_i - \mu$

Définition

La variance est la moyenne des déviations au carré (SS : sum of squares)

$$\frac{\sum (d_i)^2}{N} = \frac{SS}{N}$$

Définition

La déviation standard $\sigma = \sqrt{\text{variance}} = \sqrt{\frac{SS}{N}}$

Pour la population

Définition

déviations d'une mesure c'est la différence par rapport à la moyenne

$$d_i = X_i - \mu$$

Définition

La variance est la moyenne des déviations au carré (SS : sum of squares)

$$\frac{\sum (d_i)^2}{N} = \frac{SS}{N}$$

Définition

La déviation standard $\sigma = \sqrt{\text{variance}} = \sqrt{\frac{SS}{N}}$

Pour un échantillon (taille n)

- variance $s^2 = \frac{SS}{\text{degré liberté}} = \frac{SS}{n-1}$
- déviation standard $= \sqrt{\frac{SS}{\text{degré liberté}}} = \sqrt{\frac{SS}{n-1}}$

Degré de liberté

- Le degré de liberté : Pour un échantillon de taille n le degré de liberté (ddl) est de $n-1$
- Il indique le nombre de mesures qui peuvent varier indépendamment sans changer la moyenne.

Résumé les mesures

Nous savons résumer un ensemble de mesures par un couple :
tendance centrale et variabilité

Pour une population le couple (μ, σ) et pour un échantillon le couple (M, s) résument les informations connues sur la population et l'échantillon

Ecriture standard

Notre objectif maintenant est de trouver une écriture standard pour pouvoir comparer différentes distributions de mesures.

Cette écriture est le z-score et se calcule comme suit :

- pour une population : $z_i = \frac{X_i - \mu}{\sigma}$
- pour un échantillon : $z_i = \frac{X_i - M}{s}$

On peut voir que le z score représente la déviation de la moyenne mesurée en utilisant la déviation standard comme unité de longueur.

Z-score peut être positif ou négatif : $z\text{-score} \in]-\infty, +\infty[$

En transformant toutes les mesures avec leurs z-score nous obtenons une distribution standardisée qui est centrée autour de 0 et la valeur ± 1 correspond à $\pm \sigma$

Avantages du z-score

Nous pouvons faire des comparaisons entre les distributions standardisées

Nous pouvons faire des transformations de scores en conservant certaines propriétés (comme la fréquence et l'ordre entre les fréquences)

Exemple :

Un examen est noté sur un barème de 56 points, comment le transformer rapidement sur un barème de 20 points avec une moyenne de 10 et un écart type de 5 points ?

- Etape 1, calculer la moyenne et l'écart type des notes
- Etape 2, calculer le z-score pour chaque étudiant
- Etape 3 : faire une transformation $z = \frac{x' - 10}{5}$ donc $x' = (5 * z) + 10$

Calculer le z-score pour un échantillon

La logique est semblable au calcul du z-score pour une population. Nous remplaçons simplement la moyenne de la population μ par la moyenne de l'échantillon M et la déviation standard de la population σ par la déviation standard de l'échantillon s .

Exercice

Etudiant #	Note
1	3
2	5
3	5
4	7
5	1
6	1
7	6
8	6
9	7
10	9

- Quels sont les z-scores de la population ?
- Nous allons prendre l'échantillon des étudiants avec des numéros pairs
- Quels sont les z-scores pour cet échantillon ?

Erreur d'échantillonnage

- L'erreur d'échantillonnage est une erreur intrinsèque
- C'est la différence entre la statistique de l'échantillon et le paramètre de la population
- Cette erreur suit une loi !
- C'est le théorème central limite qui le prouve.
- Les moyennes M_i des échantillons de taille n suivent une loi normale de moyenne μ et d'écart type $\frac{\sigma}{\sqrt{n}}$

Erreur Standard

Définition

L'erreur standard pour un échantillon de taille n : $\sigma_M = \frac{\sigma}{\sqrt{n}}$

Test de Student

Nous savons calculer le z-score

Nous savons calculer l'erreur standard $\sigma_M = \frac{\sigma}{\sqrt{n}}$ pour un échantillon

Problème : la formule de σ_M utilise σ la déviation standard de la population, comment l'obtenir ?

Solution : nous allons simplement estimer σ en utilisant la variance de l'échantillon $s^2 = \frac{SS}{ddl}$

Erreur standard estimée devient donc : $s_M = \frac{s}{\sqrt{n}}$

Formule du t-score

La formule du t-score utilise le même principe que pour z-score en remplaçant σ_M par son estimateur s_M

$$t_M = \frac{M - \mu}{s_M}$$

- μ est la moyenne pour la population
- M est la moyenne pour l'échantillon
- s_M est l'erreur standard estimée pour l'échantillon

Hypothèse H0 et test de Student

Le t-score est donné par :

$$t_M = \frac{M - \mu}{s_M}$$

- M est obtenu des données
- s_M est obtenu des données
- μ est obtenu par hypothèse H0 !

t-test pour mesures répétées

Chaque individu donne deux mesures pour les deux conditions d'expérience A et B (within-subject)

Nous allons étudier la différence entre les deux mesures pour un sujet $D_i = A_i - B_i$

Hypothèse H0

H0 : L'hypothèse 0 est formulée : En moyenne, pour la population, la différence entre les mesures vaut 0

Le t-score pour D est :

$$t_D = \frac{M_D - \mu_D}{s_{M_D}}$$

t score

Sous $H_0, \mu_D = 0$, t score devient :

$$t_D = \frac{M_D}{s_{M_D}}$$

- M_D moyenne des différences sur l'échantillon
- s_{M_D} erreur standard estimée $= \sqrt{\frac{s^2}{n}}$
- s^2 est la variance estimée $= \frac{SS}{ddl}$

Prendre une décision

- L'étape suivante est de comparer la valeur de t_D avec la table de la loi de student
- Attention, il faut lire la bonne valeur dans la table :
 - Le bon α
 - Le bon degré de liberté
- Si t_D est supérieur à la valeur de la table, alors on peut rejeter H_0 avec un risque d'erreur de α
- Sinon, la différence n'est pas statistiquement significative ; elle peut être causée simplement par le hasard

Lignes directrices

- 1 Introduction
 - Objectifs du cours
 - Prérequis
- 2 Section Statistique descriptive et statistique inférentielle
- 3 Les différents types d'études expérimentales
- 4 Les variables
- 5 Mesures de la tendance centrale et la variabilité
- 6 Normalisation des mesures avec Z-score
- 7 Distribution des moyennes des échantillons
- 8 Test de Student
- 9 **Mesures répétées**
 - **Exemple**

Exemple

A	B
6	9
8	9
7	10
7	11
8	11
6	9
5	11
10	11
8	11

Etape 1 : D

A	B	D
6	9	-3
8	9	-1
7	10	-3
7	11	-4
8	11	-3
6	9	-3
5	11	-6
10	11	-1
8	11	-3

Etape 2 : Moyenne M_D

$$M_D = -3$$

Etape 3 : SS

A	B	D	$(D - M_D)^2$
6	9	-3	0
8	9	-1	4
7	10	-3	0
7	11	-4	1
8	11	-3	0
6	9	-3	0
5	11	-6	9
10	11	-1	4
8	11	-3	0

$$SS = 18$$

Etape 4 : variance & erreur std

$$s^2 = \frac{SS}{n-1} = \frac{18}{8} = 2.25$$

$$s_{M_D} = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{2.25}{9}} = 0.5$$

Etape 5 : t-score

$$t = \frac{M_D}{s_{M_D}} = \frac{-3}{0.5} = -6$$

Etape 6 : Décision

Région critique pour un ddl = 8 et $\alpha = 0.01 = 3.355$

$|t| = 6 > 3.355$ donc on rejete H_0 !

La différence est significative il y'a un donc un effet qui n'est pas simplement du au hasard (avec un risque d'erreur de 0.01)