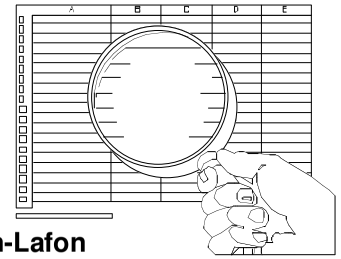
1

Analyse quantitative de données

- Statistiques
- Analyse Exploratoire de Données



Michel Beaudouin-Lafon
Université Paris-Sud

Cours de Wendy E. Mackay (I.N.R.I.A.)

2

Statistiques

- C'est un domaine TRÈS vaste
- Il est difficile de faire des suppositions correctes
- Les erreurs sont courantes
- On peut faire des tests simples

3

Analyse statistique

- Fournit les caractéristiques mathématiques des données
- Décrit comment des ensembles de données se rapportents les uns aux autres
- Estime la probabilité que les hypothèses soients correctes

4

Mesures standards en IHM

- Temps pour executer une opération
- Fréquences de certaines actions
- Nombres d'erreurs
- Réactions qualitatitves

  par exemple : "être perdu"

1

Analyse de données

- Le type des donnés et le dispositif expérimental déterminent les statistiques applicables
- Echelles de données
  - Nominale — Maillots de footballers
  - Ordinale — Classements
  - Intervalle — Température
  - Ratio — Poids
- Statistiques descriptives ou inférentielles

2

Statistiques descriptives

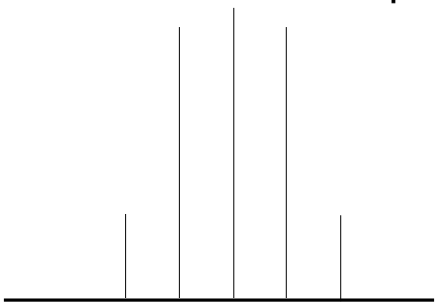
- Plus simple mais moins puissant
- Réduit la masse des données
  - Intervalle des valeurs
  - Distribution de fréquences des valeurs
  - Mesures de tendances centrales
    - moyenne
    - médiane
    - mode
  - Mesures de variance autour du centre
    - variance
    - déviation standard

3

Mesures de tendances centrales

Normale, Gaussian, "Bell"

- Moyenne = Médiane = Mode
- Calculation < 50% > le plus souvent

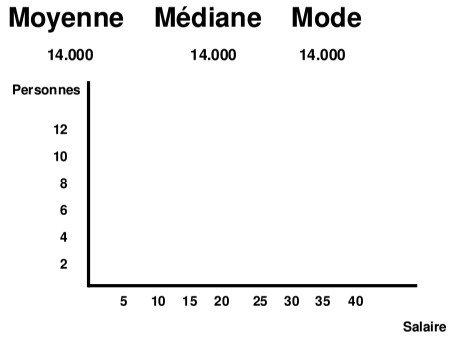


4

Mesures de tendances centrales

Salaires

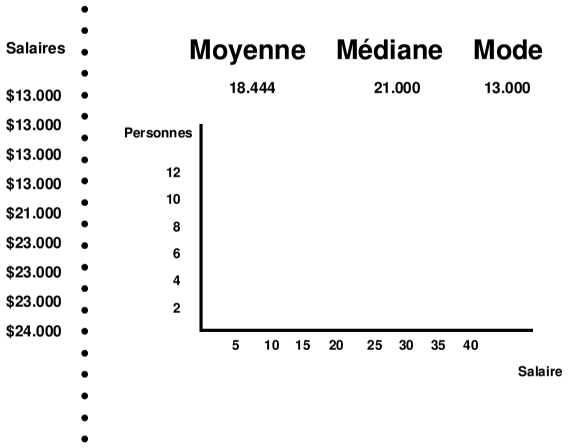
$12.000
$13.000
$13.000
$14.000
$14.000
$14.000
$15.000
$15.000
$16.000



1

Mesures de tendances centrales

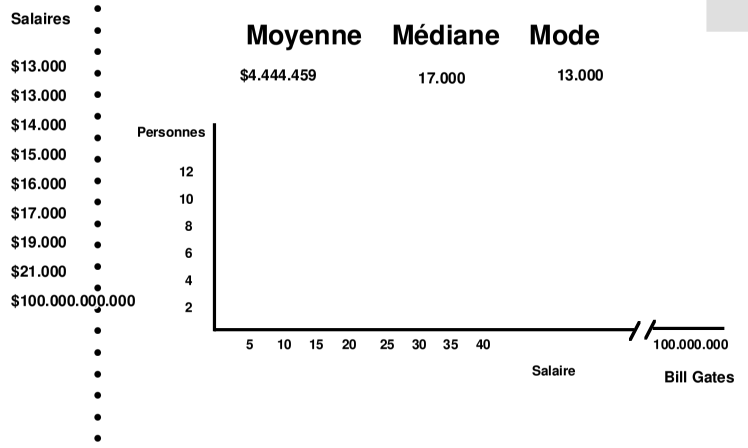
Bi-modal



2

Mesures de tendances centrales

"Skew"



3

Statistiques inférentielles

Complexe, plus puissant

Fondé sur la théorie des probabilités

Exemple : Comparer les moyennes

T test

Exemple : Mesurer une correlation

facteur rho de Pearson

Exemple : Analyse de régression

4

Exemple :

Différences entre moyennes

Données :

Deux ensembles de données qui mesurent le temps de sélection d'un item dans un menu

Question :

Y a-t-il une différence réelle entre ces deux ensembles de données, ou bien les différences sont-elles seulement dues au hasard?

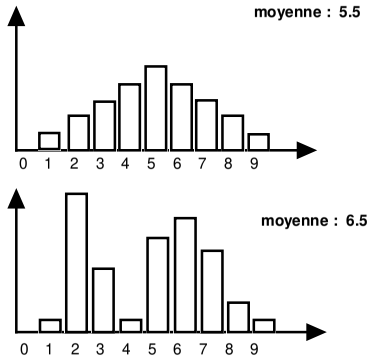
Hypothèse nulle :

Il n'y a pas de différence entre les moyennes

1

Exemple:

Différence significative entre les moyennes ?



2

Signification statistique

- **Fournit une estimation quantitative de la probabilité que deux distributions soient différentes**
- **Signification ≠ importance**
  - **Si le nombre de sujet est grand, une petite différence peut produire un résultat significatif**

3

Inspection visuelle des données

- **Les données sont toujours variables !**
- **Attention à ne pas tirer de conclusion erronnée**
- **Distinguer une variation aléatoire d'une variation systématique**
- **MAIS :**
  L'anlyse exploratoire de données utilise des techniques visuelles pour comprendre les données

4

T-Test

- **Le plus utilisé en IHM**
- **Donne le niveau de confiance que la différence entre deux moyennes est due à une différence réelle et non pas à une variation aléatoire**
- **NE DIT RIEN de la taille de la différence entre les moyennes**

1

Exemples de T-Test :

- **Comparer 2 ensembles d'observations indépendantes**
  - Deux groupes avec des sujets différents
  - Le nombre de sujets par groupe peut varier
- **Observations par couples**
  - Un groupe avec des conditions différentes
  - Les couples de données d'un sujet sont comparés
- **Non-directionnel ou directionnel**
  - "Two-tailed test"
    - Différences dans les deux directions
  - "One-tailed test"
    - Différences dans une seule direction

2

Suppositions pour le T-Test

- **Les données sont distribuées aléatoirement**
  - Distribution normale
- **Les variances des populations sont égales**
  - Le test est cependant robuste si ce n'est pas le cas
- **Les observations individuelles sont indépendantes**
- **p-value: la probability que les moyennes ne soient pas différentes**
- **Choisir le critère au préalable (p=.05)**
- **Accepter or rejeter l'hypothèse nulle en fonction de la p-value**

3

Types d'erreur

Théorie de "Signal Detection"

		Réalité: Signal	Réalité: Bruit
Prédiction	Signal	Hit	False Alarm
	Bruit	Miss	Correct Rejection

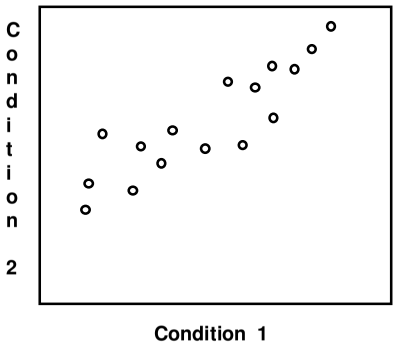
4

Niveau de signification et types d'erreur

- **Erreur de type 1**
  - Rejeter l'hypothèse nulle alors qu'elle est vraie
- **Erreur de type 2**
  - Accepter l'hypothèse nulle alors qu'elle est fausse
- **Un fort niveau de signification augmente les chances d'une erreur de type 2**
- **Un faible niveau de signification augmente les chances d'une erreur de type 1**

1

Corrélation



2

Corrélation

- **Mesure à quel point deux données sont corrélées**
- **Fondée sur un coefficent de corrélation**
  - **+1 Corrélation positive parfaite**
  - **0 Pas de corrélation**
  - **-1 Corrélation négative parfaite**
- **Corrélation ≠ Cause !!**
- **Pas fiable avec des petits groupes**

3

Régression

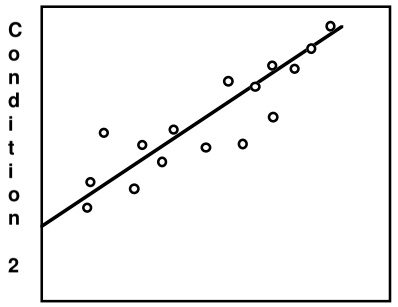
- **Calcule la droite qui approche au mieux les données**
- **Utiliser la valeur d'une variable pour prédire l'autre**

4

Exemple :

Condition 1	Condition 2
5	6
4	4
6	6
4	4
5	5
3	3
5	5
4	4
5	5
6	6
6	6
7	7
6	6
7	7

y=.99
x =1.32
r2 = .668



1

Erreurs dans l'interprétation

- TRES fréquentes : ce n'est pas facile !
- Signification ≠ Importance
- Les suppositions sont souvent fausses :
  - Les données ne suivent pas une distribution normale
  - Puissance du test
  - Valeur de la "p-value" ≠ taille de l'effet
    - Signification vs. force de la relation

2

P-value

Critère alpha = 0,05

- 0,05 = 1/20
- S'il n'y a pas de différence et que je fait cette expérience 20 fois,
- un test donnera un résultat significatif
  les 19 autres donnera un résultat non-significatif

3

Analyse exploratoire de données

Tukey & Mosteller

- Processus interactif et itératif
- Explorer les donnés pour découvrir leurs propriétés :
  - Le "doux" et le "rugueux"
- Principes clés
  - Etre ouvert
  - Etre sceptique

4

Analyse exploratoire de données

- Importance de la visalisation
  - Mais base mathématique solide
- Mesures de :
  - Position : moyenne, médiane, mode
  - Etalement : variance, déviation standard
  - Forme : courbe en cloche, biais, points isolés
- Représentation graphique des données
  - Graphes "Stem-and-leaf"
  - "Box-and-whisker"

1

Box-and-whisker plot

2

Études utilisateurs

Toujours avec les vrais utilisateurs !

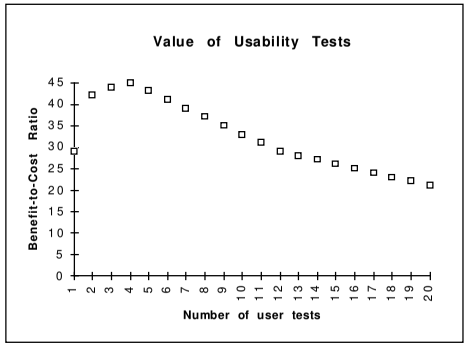
Problèmes

- **Comment définir une étape ?**
- **Et si l'utilisateur ne peut pas faire un tâche ?**
- **Que faire si le système est mal installé ?**
- **Quelle est l'importance des connaissances préalables de l'utilisateur ?**

3

Intérêt des tests d'utilisabilité

Rapport coût-bénéfice

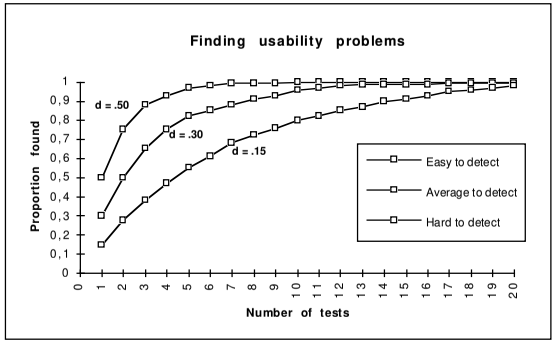


Nielsen & Landauer (1993)

4

Les problèmes d'utilisabilité

Combien de tests sont nécessaires ?



Nielsen & Landauer (1993)

d = détectabilité

Probabilité qu'un test identifiera un problème

1

Résultats

Types d'observation

Fonctionalité

Pas d'aide, etc.

Generale

Difficile à utiliser, Trop d'icônes

Specifique

Il faut quitter pour effacer un postit

Bugs

L'import de postit ne fonctionne pas

Suggestions

Une vue generale sur tous les postits crées

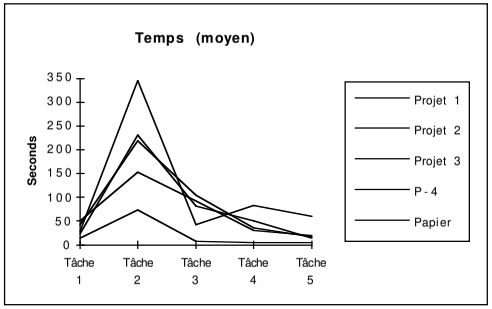
Décisions

Tous les postits sont dans le meme mode

2

Résultats

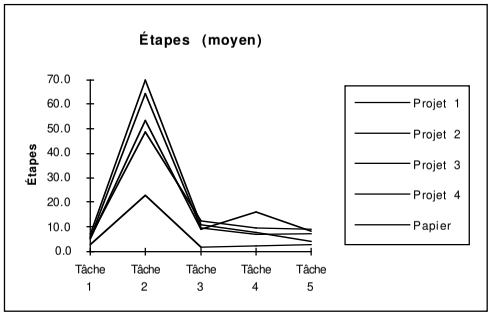
Temps



3

Résultats

Étapes



4