

*UE4 : Biostatistiques*

# Chapitre 5 : Intervalles de confiance

Professeur Philippe CINQUIN

Année universitaire 2011/2012

Université Joseph Fourier de Grenoble - Tous droits réservés.

# Plan

- A) Introduction
- B) Statistiques descriptives
- C) Probabilités
- D) Estimation
- E) Intervalles de confiance
- F) Problème récapitulatif
- G) Résumé des objectifs

# Intervalle de confiance à 95%

- **Soit  $N(m,s)$** 
  - Trouver  $\delta$  tel que
  - **$P(m-\delta \leq N \leq m+\delta) = 0.95$**
- $P(m-\delta \leq N \leq m+\delta) = P(-\delta \leq N-m \leq \delta)$   
 $= P(-\delta/s \leq (N-m)/s \leq \delta/s)$
- Or,  $Z = (N-m)/s$  est une loi normale  $N(0,1)$
- Donc  $P(m-\delta \leq N \leq m+\delta) = 0.95 \Leftrightarrow \delta/s = 2$
- ou encore  **$P(m-2s \leq N \leq m+2s) = 0.95$**

# Intervalle de confiance à $(1-\alpha)$

- **Soit  $N(m, s)$** 
  - Trouver  $\delta$  tel que
  - $P(m-\delta \leq N \leq m+\delta) = 1-\alpha$
- $P(m-\delta \leq N \leq m+\delta) = P(-\delta \leq N-m \leq \delta)$   
 $= P(-\delta/s \leq (N-m)/s \leq \delta/s) = P(-\varepsilon \leq (N-m)/s \leq \varepsilon) = 1-\alpha$
- $Z = (N-m)/s$  étant une loi normale  $N(0,1)$ , le  $\varepsilon$  associé au  $\alpha$  se lit directement dans la table « de l'écart réduit »

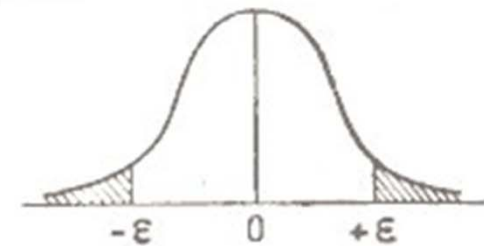
# Exemple

- Quel est l'intervalle de confiance à 99% du poids de naissance d'un nouveau-né, si le poids de naissance suit une loi normale  $N(3200, 400)$  ?
- $0.99 = 1 - 0.01$ . Il faut donc trouver dans la « table de l'écart réduit » la valeur de  $\varepsilon$  associée à  $\alpha = 0.01$  ( $= 0.00 + 0.01$ ). On lit  $\varepsilon = 2.576$

TABLE I

Table de l'écart-réduit (loi normale) (\*).

La table donne la probabilité  $\alpha$  pour que l'écart-réduit égale ou dépasse, en valeur absolue, une valeur donnée  $\varepsilon$ , c'est-à-dire la probabilité extérieure à l'intervalle  $(-\varepsilon, +\varepsilon)$ .



$\alpha$	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,00	$\infty$	2,576	2,326	2,170	2,054	1,960	1,881	1,812	1,751	1,695
0 10	1,645	1,598	1,555	1,514	1,476	1,440	1,405	1,372	1,341	1,311

# Exemple

- Donc l'intervalle de confiance à 99% du poids de naissance d'un nouveau-né, si le poids de naissance suit une loi normale  $N(3200, 400)$  est :  
[3200 - 2.576 x 400 ; 3200 + 2.576 x 400] =  
[3200 - 1030 ; 3200 + 1030] =  
[2170 ; 4230]

# Intervalle de confiance de la moyenne d'un échantillon

- Soient  $X$  une variable aléatoire définie sur une population, et  $n$  mesures  $\{x_i, i=1, \dots, n\}$  réalisées sur un échantillon extrait de cette population.

- Les paramètres  $\mu$  et  $\sigma^2$  inconnus de  $X$  seront estimés par

$$m = \sum x_i / n \text{ et } s^2 = \sum (x_i - m)^2 / (n - 1)$$

- Comment définir un intervalle  $[a, b]$  tel que

$$P(a \leq \mu \leq b) = 1 - \alpha ?$$

# Moyenne empirique

- **Remarque** :  $m$  est une réalisation de la variable aléatoire “ **moyenne empirique dans des échantillons de taille  $n$**  ”
- Exemple : soient 8 échantillons de 200 P1, dont on **mesure** la taille moyenne. On obtient 8 valeurs :  $m_1, m_2, m_3, m_4, m_5, m_6, m_7, m_8$ . Les  $m_i$  n'ont aucune raison d'être toutes égales : ce sont des réalisations de la variable « moyenne empirique de la taille dans des échantillons d'effectif 200 ».



# Moyenne empirique (suite)

- Supposons que l'on connaisse la moyenne  $\mu$  et la variance  $\sigma^2$  de la variable aléatoire « taille » dans la population de l'ensemble des P1 de France.
  - Que peut-on en déduire sur la distribution des  $m_i$  ?
- Réciproquement, si à partir d'un échantillon j'estime  $\mu$  par  $m$  et  $\sigma$  par  $s$ , que puis-je dire sur la « fiabilité » de l'estimation de  $\mu$  ?
  - Intuitivement, on se dit que plus  $\sigma$  est grand, moins « fiable » sera l'estimation de  $\mu$  par  $m$ .

# Moyenne empirique (fin)

- On démontre que la v. a. « moyenne empirique dans un échantillon de taille  $n$  » suit une loi :
  - de moyenne  $\mu$ ,
  - de variance  $\sigma_e^2 = \sigma^2/n$ 
    - NB si  $n = 1$ ,  $\sigma_e^2 = \sigma^2$  (pas étonnant, un échantillon de taille 1, c'est un individu!)
    - NB si  $n$  tend vers l'infini,  $\sigma_e^2$  tend vers 0 : (pas étonnant, on connaît parfaitement la moyenne si on a pour échantillon l'univers entier!).
- Si  $n > 30$ , cette loi est normale, (*sinon, elle suit une loi « de Student », voir le cours de J. Labarère, mais c'est hors programme pour l'intervalle de confiance de la moyenne*)

# Intervalle de confiance de la moyenne

- donc  $P(\mu - \varepsilon \sigma_e \leq m \leq \mu + \varepsilon \sigma_e) = 1 - \alpha$ , où  $\varepsilon$  et  $\alpha$  sont lus dans la “ table de l'écart réduit ”.

- Or, *(puisque'on connaît m, on veut un intervalle centré sur lui, permettant de prédire où est  $\mu$ )*

$$P(\mu - \varepsilon \sigma_e \leq m \leq \mu + \varepsilon \sigma_e) = P(-\varepsilon \sigma_e \leq m - \mu \leq \varepsilon \sigma_e) =$$

$$P(-\varepsilon \leq (m - \mu) / \sigma_e \leq \varepsilon) = P(|m - \mu| / \sigma_e \leq \varepsilon) =$$

$$P(|\mu - m| / \sigma_e \leq \varepsilon) =$$

$$\mathbf{P(m - \varepsilon \sigma_e \leq \mu \leq m + \varepsilon \sigma_e) = 1 - \alpha,}$$

- **L'intervalle de confiance à  $1 - \alpha$ , de la moyenne est donc**

$$\mathbf{[m - \varepsilon \sigma_e ; m + \varepsilon \sigma_e]}$$

- **En pratique, il faut se souvenir que  $\sigma$  n'est pas connu. Nous devons donc l'estimer par  $s$ .**
- **L'intervalle de confiance de la moyenne est donc**

$$\left[ m - \varepsilon \frac{s}{\sqrt{n}} \leq \mu \leq m + \varepsilon \frac{s}{\sqrt{n}} \right]$$

# Mentions légales

L'ensemble de cette œuvre relève des législations française et internationale sur le droit d'auteur et la propriété intellectuelle, littéraire et artistique ou toute autre loi applicable.

Tous les droits de reproduction, adaptation, transformation, transcription ou traduction de tout ou partie sont réservés pour les textes ainsi que pour l'ensemble des documents iconographiques, photographiques, vidéos et sonores.

Cette œuvre est interdite à la vente ou à la location. Sa diffusion, duplication, mise à disposition du public (sous quelque forme ou support que ce soit), mise en réseau, partielles ou totales, sont strictement réservées à l'université Joseph Fourier (UJF) Grenoble 1 et ses affiliés.

L'utilisation de ce document est strictement réservée à l'usage privé des étudiants inscrits à l'Université Joseph Fourier (UJF) Grenoble 1, et non destinée à une utilisation collective, gratuite ou payante.