

# L'ENSEIGNEMENT DE LA BIOSTATISTIQUE EN PAYS EN VOIE DE DÉVELOPPEMENT

*C. Senterre <sup>(a,b)</sup>, M. Dramaix <sup>(a)</sup>, A. Levêque <sup>(b)</sup>*

*<sup>(a)</sup> Département de Biostatistique,.*

*<sup>(b)</sup> Unité Epidémiologie, Traumatismes et Maladies chroniques.*

*Ecole de Santé Publique de l'Université Libre de Bruxelles*

# Contexte....

- Les programmes universitaires de formation en sciences de la santé incluent généralement un cours de biostatistique et ce que ces formations soient dispensées en Europe ou en Afrique.
- Les formations en Santé Publique ne dérogent pas à cette règle.

ULB

ESP

# Contexte....

- « *la bonne compréhension de certains concepts paraît plus importante, dans une première approximation, que la connaissance des formules mathématiques sur lesquels ils s'appuient* » (Huguier & Flahaut, 2003).
  
- → A l'ESP: enseignement au moyen d'exemples
  - Issus d'études épidémiologiques
  - Issus de la littérature



# Contexte....

- Une fois diplômé, le praticien de santé publique constatera que « *la majeure partie du temps consacré à une étude [...] doit porter sur le choix et la pertinence de la méthode employée, puis sur la signification et l'interprétation des résultats* »(Ancelle, 2002).
- → Attention accordée
  - Au choix des méthodes,
  - a l'interprétation et à la présentation des résultats.



# Contexte....

- Si cette approche pédagogique est tout à fait acquise chez nous, il n'en est pas toujours de même dans certains pays d'Afrique où l'enseignement reste souvent « fort théorique ».

# Un exemple?

Model to predict

$$E(Y_i) = \beta_0 + \beta_1 X_i$$

$$\varepsilon_i = Y_i - E(Y_i)$$

Find  $\hat{\beta}_0$  and  $\hat{\beta}_1$  which minimize  $\sum_{i=1}^n \varepsilon_i^2$

Normal Equations

$$\begin{cases} \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) = 0 \\ \sum_{i=1}^n X_i (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) = 0 \\ n\bar{Y} - n\hat{\beta}_0 - n\hat{\beta}_1 \bar{X} = 0 \\ \sum_{i=1}^n X_i Y_i - n\hat{\beta}_0 \bar{X} - \sum_{i=1}^n X_i^2 \hat{\beta}_1 = 0 \end{cases}$$

$$\sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2$$

$$\frac{\partial}{\partial \beta_0} \left( \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2 \right) = -2 \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)$$

$$\frac{\partial}{\partial \beta_1} \left( \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2 \right) = -2 \sum_{i=1}^n X_i (Y_i - \beta_0 - \beta_1 X_i)$$

*Solution*

$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$ : the estimator of the intercept  
and

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n X_i Y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^n X_i^2 - n\bar{X}^2}$$
 : the estimator of the slope

Predictor of E(Y)

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

# Contexte....

- Mais **une volonté de changement existe!**
- Collaborations étroites avec
  - l'Institut Régional de Santé Publique de l'Université d'Abomey-Calavi au Bénin
  - l'Ecole de Santé Publique de l'Université Nationale du Rwanda
  - la Faculté des Sciences et des Techniques de l'Université Cheikh Anta Diop au Sénégal
  - ...

ULB

ESP

# Approche....

- Méthodes d'apprentissage = identiques ESP  
(opérationnalisation théorie via situations réelles de santé publique + intégration dimension épidémiologique.
- Compréhension des concepts + bons choix méthodologiques face à leur « quotidien » car apprenants = professionnels de terrain (médecins, techniciens supérieurs en soins infirmiers et obstétricaux, assistants médicaux, pharmaciens, biologistes, etc.) + présentation (diffusion, communication, publication)

ULB

ESP



## La régression linéaire simple

- **Objectif :**  
Prédire y (var dép) en fonction de x (var prédictrice)
- **Modèle :**  

$$y_{obs} = y_{préd} + \text{erreur}$$

$$y_{préd} = \alpha + \beta x \quad (\rightarrow \text{DROITE})$$
- **Estimation** de « a » de  $\alpha$  et « b » de  $\beta$  :  
→ moindres carrés

## HDL en le BMI

Récapitulatif des modèles<sup>b</sup>

Modèle	R	R-deux	R-deux ajusté	Erreur standard de l'estimation
1	,231 <sup>a</sup>	,053	,050	15,028

a. Valeurs prédites : (constantes), bmi

b. Variable dépendante : hdl

5,3% de la variabilité du HDL est due au BMI

## HDL en le BMI (output SPSS)

Coefficients<sup>a</sup>

Modèle	Coefficients non standardisés		Coefficients standardisés		Sig.
	A	Erreur standard	Bêta	t	
1 (Constante)	75,839	5,867		12,926	,000
bmi	-,812	,219	-,231	-3,713	,000

a. Variable dépendante : hdl

HDL prédit = 75,839 – 0,812 BMI

Quand le BMI augmente de 1 kg/m<sup>2</sup>,  
le HDL diminue de 0.812 unités de mesure

**• Objectif :**

Prédire  $y$  (var dép) en fonction des  $x_j$   
(prédicteurs)

**Exercice de régression linéaire multiple (base de données BIRHN)**

BIRNH = The Belgian Interuniversity Research on Nutrition and Health<sup>1</sup>.

**Les objectifs de l'analyse :**

Il s'agit d'étudier la pression artérielle systolique (PAS) selon différents prédicteurs que sont l'âge (AGE), le sexe (SEXE), l'indice de masse corporelle (BMI), le niveau d'éducation (GRETU), le statut tabagique (TABAC2), la claudication (CLAUD2) et le fait d'avoir des antécédents de maladies cardiovasculaires (FREE).

**Les étapes de l'analyse :**

1. Description de l'échantillon
2. Description de la variable dépendante
3. Etude bivariée de la variable dépendante selon chaque prédicteurs pris un à un.
4. Etude bivariée des prédicteurs entre eux.
5. Etude multivariée de la variable dépendante selon l'ensemble des prédicteurs retenus.
6. Analyse des résidus pour décider si le modèle final choisi peut être conservé ou non.

**L'analyse :**

# La suite....

- Ces expériences sont très enrichissantes
- Des améliorations devraient être apportées, comme par exemple, l'implication d'enseignants locaux afin que ces institutions puissent, petit à petit, s'appropriier les contenus dispensés et devenir autonomes dans l'enseignement des méthodes biostatistiques de base et avancées