



FACULTÉ DE MÉDECINE ET DE PHARMACIE
UNIVERSITÉ HASSAN II DE CASABLANCA



Biostatistiques

Résumé

Module : Médecine communautaire + Biostatistiques

Basé sur : Les QCMs

-> Ce résumé est un complément de cours, il contient suffisamment d'informations, mais ne remplace pas le polycopié du professeur.

-> Merci d'envoyer toutes vos remarques via l'adresse mail suivante :
mahdikettani1@gmail.com

-> Bon courage et bonne lecture !

Auteur : Kettani El Mahdi, étudiant de la promotion médecine 2019

اللهم أستودعك ما قرأت و ما حفظت و ما تعلمت، فردّه عند حاجتي إليه، إنك على كل شيء قدير

Biostatistique

Rôle de la Biostatistique :

- > Ne supprime pas la variabilité biologique
- > Organise les données provenant des observations
- > Décrit des phénomènes par des paramètres résumant les observations
- > Prédire la probabilité de survenue d'événements
- > Compare des paramètres entre plusieurs populations

Variabilité biologique :

- > Existe indépendamment des erreurs de mesure
- > N'est pas réservé aux mesures quantitatives
- > Entraîne une variabilité au niveau des échantillons
- > Variabilité biologique : Moléculaire - Génomique - Cellulaire - Fonctionnelle - Populationnelle

Plan d'échantillonnage Probabiliste : (Respecte l'inférence statistique)

- > Aléatoire Simple
- > Systématique
- > En grappe
- > Stratifié (+ raffiné)

Plan d'échantillonnage Non Probabiliste : (Peu coûteux - Rapide - Facile à appliquer)

- > Accidentels
- > Volontaire
- > Quotas (+ utilisé)
- > Boucle de neige

Mesure de Dispersion :

- > Etendue (Différence des valeurs extrêmes $E = X_{\max} - X_{\min}$)
- > Variance
- > Ecrit Type
- > Coefficient de variation
- > Intervalle interquartile et semi-interquartile

Description de la distribution d'une variable quantitative :

- > Etendue
- > Intervalle interquartile

Description d'une variable quantitative :

- > Mesure de Tendance centrale
- > Mesure de dispersion

Méthode pour vérifier la normalité d'une distribution :

- > Histogramme
- > Boîte à moustache
- > Test de Kolmogorov Smirnov

Fréquence de dépression Casablanca comparé à celle Nationale :

- > Comparaison d'un pourcentage à une valeur de référence
- > H_1 est unilatérale
- > H_0 c'est $P_{\text{casa}} = P_{\text{nat}}$ avec P_{casa} : P de dépression à Casa et P_{nat} : P de dépression au Maroc

Mesure à prendre dans un plan d'échantillonnage :

- > Disponibilité des ressources financières et humaines
- > Durée
- > Sujet de question d'étude
- > Disponibilité des données

Variable :

1) Quantitative :

- a) Continue (Mesurable)
- b) Discontinue ou Discrète (Dénombrable)
- c) Temporelle (Age de grossesse, temps de sommeil...)

2) Qualitative :

- a) Nominale (sans ordre. Exp : Vert, Rouge, Bleu...)
- b) Ordinale (avec ordre. Exp : Bac +1 Bac +2 Bac +3)
- c) Binaire ou Dichotomique (Exp : Sexe : homme ou femme)

1) Qualitative :

a) Nominale :

- > < 5 camembert = diagramme en secteur
- > > 5 Battons Horizontale

b) Ordinale :

- > bâtons verticale

2) Quantitative :

a) Discontinue :

- > Histogramme

b) Continue :

- > Histogramme
- > Polygone des effectifs et des fréquences

La Médiane :

Exp 1: 1,1,1,1,2,2,3,4,4,4,5 -> Médiane = 2

Exp 2: 1,2,2,3,5,10 -> Médiane = $2+3/2 = 2,5$

La Moyenne :

Exp : 1,1,2,3,3,3,5,8,11,19,25 -> Moyenne = $1+1+2+3+3+3+5+8+11+19+25 / 11 = 81/11 = 7,36$

Le Mode :

Valeur de la valeur la + élevé

Exp 1: 5,2,3,10,4 -> Mode = 10 (Unimodale)

Exp 2: 5,2,3,7,4,7,2 -> Mode = 7 (Bimodale)

Exp 3: 5,2,12,4,12,12,3,12,1,12 = Mode 12 (Plurimodal)

Intervalle de confiance :

1) Moyenne : (Ecart-Type / Variance)

$m = + \text{ ou } - 1,96 \times S_m$

$S_m = S/\sqrt{n} = \sqrt{S^2/n}$

$$S_m = S/\sqrt{n}$$

2) Pourcentage : (Effectif / Pourcentage)

$m = + \text{ ou } - 1,96 \times S_p$

$S_p = \sqrt{(Pe \times (1 - Pe))/n}$

$Pe = \text{Effectif} / n$

$$S_p = \sqrt{p_e(1-p_e)/n}$$

Condition pour calculer l'intervalle de confiance :

n_{pi} , n_{ps} , $n(1-p_i)$, $n(1-p_s)$ soient supérieurs ou égaux à 5

Calcul du ddl :

$ddl = (Cologne - 1) \times (Ligne - 1)$

Exp :

Calcul de l'effectif théorique :

	1	2	3	Total
a				
b		?		Y
c				
Total		X		Z

$$? = E_T = (X + Y) / Z$$

	1	2	3
a			
b			

$$ddl = (2 - 1) \times (3 - 1) = 2$$

Ecart-Type = $\sqrt{\text{variance}}$
S = racine de S au carré

Test statistique

Variable Qualitative

Test Ecart Réduit

Test Khi2

Peut passer de là à là

Variable Quantitative

Test Student

Test Z

Condition :

- > Variable en %
- > np ≥ 5
- > nq ≥ 5 (avec q = 1- p)

1) Comparaison entre 2 échantillon :

$$\varepsilon = \frac{Po1 - Po2}{sdp}$$

$$Sdp = \sqrt{P(1-P) \left(\frac{1}{n1} + \frac{1}{n2} \right)}$$

$$P = \frac{r1}{n1} + \frac{r2}{n2} = \frac{n1 Po1 + n2 Po2}{n1 + n2}$$

Pour un risque α = 5%

> 2 alors différence significative

< 2 alors différence non significative

1) Comparaison entre 1 échantillon et % théorique:

$$\varepsilon = \frac{P - po}{Sp}$$

$$Sp = \sqrt{PQ/n}$$

Pour un risque α = 5%

> 2 alors différence significative

< 2 alors différence non significative

NB :

Différence significative = Rejet de H0, Prise de H1

Différence non significative = Garder H0

Condition :

- > Variable en effectif
- > Effectif calculé Ci ≥ 5
- 1) 1 échantillon :

$$\chi^2 = \sum (O_i - C_i)^2 / C_i$$

> χ² Théorique alors différence significative

< χ² Théorique alors différence non significative

χ² Théorique ddl = K-1

	1%	2%	3%	4%	5%	6%	7%
1							
2							
3							
4							

NB :

Ci = nombre qu'on a calculé théoriquement

Oi = nombre qu'on a trouvé réellement

2) 2 échantillon :

$$ddl = (L-1) \times (C-1)$$

$$Ci = (Total L + Total C) / Total général$$

Si np, nq < 5, χ² n'est plus valable, on utilise alors le test χ² corrigé par correction de Yates

$$\chi^2_c = \frac{(|O_1 - C_1| - 1/2)^2}{C_1} + \frac{(|O_2 - C_2| - 1/2)^2}{C_2}$$

Elle est valable juste s'il y a 2 échantillon et np ou nq voisin de 5

Si Ci = 1, χ² corrigé n'est plus valide, on utilise Test exact de Fisher

Condition :

- > n < 30
- > Distribution normal
- > Variable = ou presque
- > Dépend de S (écart-type) et m (moyenne)

1) Comparaison entre 2 échantillon :

$$t_0 = \frac{m_1 - m_2}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

avec $s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$

≥ t x (n1+n2-2; α) différence significative

< t x (n1+n2-2; α) différence non significative

Condition :

- > n ≥ 30
- 1) Comparaison entre 2 échantillon :

$$z_0 = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

≥ Za/2 différence significative

< Za/2 différence non significative