Université Abou Bakr Belkaid

Département de Biologie

L3 AACQ S2 (2019-2020)

Chapitre II : Estimation

[1]

1. D^r $A.BOUKENKOUL$

Table des matières

Chapitre 1

Correlation et régression

Remarque 1.0.1. *Ce chapitre a été vu en classe*

Chapitre 2

Estimation des paramètres inconnus

2.1 Estimer une proportion dans une population

2.2 Estimer une moyenne dans une population

2.2.1 Estimer la moyenne d'une population (l'ecart type σ connu)

Considérons toujours l'exemple de Mendel avec le croisement des pois à gousses vertes avec des pois à gousses jaunes.
Cette lignée comportait 580 pois dont 428 à gousses vertes et 152 à gousses jaunes.
Mendel s'attendait à 25% de pois à gousses jaunes. Toute fin, avec ces résultats on a : 26,2% de pois à gousses jaunes ; Comment peut-on expliquer cette différence ? Est-elle importante pour suggérer que les 25% de Mendel sont incorrectes ?

Soit Ω une population de moyenne μ (inconnu) et l'ecart type σ (connu).
<u>Conditions requises pour estimer μ quand σ est connu :</u>

1. l'échantillon est un échantillon aléatoire simple (ie : tous les échantillons de la même taille ont la même chance d 'être sélectionnés).
2. la valeur de l'ecart type σ de Ω est connue.
3. On a soit $n \geq 30$ soit la population Ω est normalement distribuée.

Exemple 2.2.1. *Dans un échantillon de 106 températures corporelles, on a :* $\bar{x} = 36,78$ *et* $S = 0,34$
Utilisez cet échantillon pour trouver la meilleure estimation ponctuelle de la moyenne μ *de la population de toutes les températures corporelles.*

Solution 1. $\bar{x} = 36,78$ *moyenne de l'échantillon, comme* $\bar{x}$ *est la meilleure estimation ponctuelle de* μ *moyenne de* Ω*. On conclut que* μ *de* Ω *de toutes les températures corporelles est* $36,78c$

Intervalle de confiance :

Marge d'erreur :

Définition 2.2.1. *Lorsque* σ *est connu la marge d'erreur est donnée par :*

$$E = Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

et l'intervalle de confiance par : $IC = \bar{x} - E \leq \mu \leq \bar{x} + E$ *ou* $\bar{x} \mp E$ *ou* $[\bar{x} - E\,;\, \bar{x} + E]$.

Exemple 2.2.2. *Pour l'échantillon des températures corporelles (précédent) ; on a :* $n = 106$; $\bar{x} = 36,78$*. Supposons que l'échantillon est un échantillon aléatoire simple et que* $\sigma = 0,34$*. A l'aide d'un intervalle de confiance à 95% trouvez :*

1. *La marge d'erreur* E*.*
2. *L'intervalle de confiance pour* μ*.*

Solution 2. *Les conditions requises sont satisfaites*

1. *Marge d'erreur :*

$$E = Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

à un niveau de confiance de 95% on a un risque $\alpha = 5\%$ *et une valeur critique* $Z_{\frac{\alpha}{2}} = 1,96$*, donc :*

$$E = 1,96 \frac{0,34}{\sqrt{106}} = 0,06472649$$

2. *Avec* $\bar{x} = 36,78$ *et* $E = 0,06472649$*, l'IC est :*

$$\bar{x} - E \leq \mu \leq \bar{x} + E$$
$$36,78 - 0,06472649 \leq \mu \leq 36,78 + 0,06472649$$
$$36,72 \leq \mu \leq 36,84$$

Interprétation 1. *L'intervalle de confiance est* $]36,72,\ 36,84[$
ie si on séléctionnait de nombreux échantillons de taille 106 et qu'on construisait les intervalles de confiance correspendants, 95% d'entre eux contiendraient effectivement la vraie valeur μ *de la moyenne de* Ω*. Notons que* $37 \notin IC$*, et que 37 c est la valeur communément admise pour le corps humain. À partir de ces résultats, il semble peu-probable que 37 c soit la température moyenne du corps humain pour la population totale.*

Taille d'échantillon pour estimer la moyenne μ :

On admet la formule suivantes :

$$n = (\frac{Z_{\frac{\alpha}{2}}.\sigma}{E})^2$$

où :

- $Z_{\frac{\alpha}{2}}$: le score-z critique (valeur critique).
- E :La marge d'erreur voulue.
- σ : Ecart type de la population Ω.

Remarque 2.2.1. *Il faut toujours arrondir* n *à l'entier immediatement superieur.*

Exemple 2.2.3. *Trouvez la taille d'échantillon minimale requise pour estimer la moyenne* μ *inconnue de la population sachant que : Marge d'erreur : 125\$, Niveau de confiance : 95%, l'ecart type* $\sigma = 500\$$.

Solution 3. *On applique la formule suivante :*
$n = (\frac{Z_{\frac{\alpha}{2}}.\sigma}{E})^2$ *on a :* $E = 125$, $\sigma = 500$ *et* $Z_{\frac{\alpha}{2}} = 1,96$, *donc :*

$$n = (\frac{1,96 \times 500}{125})^2 = 62$$

2.2.2 Estimer la moyenne d'une population (σ inconnu) :

Conditions requises pour estimer μ quand σ est inconnu

- L'échantillon est un échantillon aléatoire simple.
- Il faut soit $n \geq 30$ soit Ω est normalement distribuée.

La loi de Student :

Si la distribution de Ω est essentiellement normale (Approximativement en cloche), alors la distribution :

$$t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

est essentiellement une loi t de Student pour tous les échantillons de taille n. Ces valeurs critiques sont notées $t_{\frac{\alpha}{2}}$.

Définition 2.2.2. *(Nombre de degrés de liberté (d.d.l))*
Le nombre de d.d.l pour un échantillon est le nombre de valeurs d'échantillon qui peuvent varier aprés avoir imposé certaines restrictions sur l'ensemble des données.

Exemple 2.2.4. *Dix (10) étudients passent un test avec une moyenne de 80. On peut donner n'importe qu'elle valeur aux 9 premiers scores mais le 10 éme score est déterminé. La somme des scores doit faire 800, donc le 10 éme doit valoir 800(-) moins la somme des 9 premiers scores.*
Parceque les 9 premiers peuvent être librement choisis, on dit qu'il y a 9 degrés de liberte disposibles.

Remarque 2.2.2. *On prend comme d.d.l = n - 1 (n taille de l'échantillon)*

Exemple 2.2.5. *Un échantillon de taille n = 15 (échantillon aléatoire simple) est séléctionné à partir d'une loi normal. Trouvez la valeur critique* $t_{\frac{\alpha}{2}}$ *correspondant au niveau de confiance 95%*

Solution 4. *n = 15, donc : d.d.l = n - 1 = 14.*
NC = 95%, donc : $\alpha = 5\%$ *d'ou* $t_{\frac{\alpha}{2}} = 2,145$ *(d'près la table de Student)*

Marge d'erreur pour l'estimation de μ (avec σ inconnu et intervalle de confiance) :

On utilisera :
$E = t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$ où $t_{\frac{\alpha}{2}}$ à (n -1) degrés de liberte

Définition 2.2.3. *L'intervalle de confiance est donné par :*

$$\begin{gathered} \bar{x} - E \leq \mu \leq \bar{x} + E \\ avec : E = t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \\ ou : \bar{x} \mp E \\ ou : [\bar{x} - E\,;\, \bar{x} + E]. \end{gathered}$$

Exemple 2.2.6. *On a pris un exemple qui montrait la constrution d'un intervalle de confiance pour estimer μ.*
Nous avons utilisé l'échantillon des températures corporelles avec $n = 106$ et $\bar{x} = 36,78$.
Nous avons aussi supposé que σ était connu et valait 0,34.
En réalité σ n'est pas connu. Utilisez ces informations pour trouver à 95%.

1. *La marge d'erreur.*
2. *L'IC pour μ.*

Solution 5.

1. *L'échantillon est un échantillon aléatoire simple, n = 106, ($n \geq 30$), donc les conditions requises sont satisfaites. On utilise la loi de Student.*
2. *Pour $n - 1 = 105$, le degré de liberte = 105 et $\alpha = 0,05$. La table de t de Student donne la valeur : $t_{\frac{\alpha}{2}} = 1,984$*
3. *On a : $E = t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} = 1,984.\frac{0,34}{\sqrt{106}} = 0,06551906$*
4. *IC peut se calculer comme suit : $\bar{x} = 36,78$ et $E = 0,06551906$*

$$\begin{gathered} \bar{x} - E < \mu < \bar{x} + E \\ 36,72 < \mu < 36,84 \end{gathered}$$

Interprétation 2. *IC = $[36,72\,;\,36,84]$*
On peut être sûr à 95% que les limites 36,72 c et 36,84 c contiennent effectivement la vraie valeur de μ dans Ω.
Notons que $37 \notin IC$ et que 37 est la valeur communément admise pour le corps humain.
À partir de ces résultats, il semble probable que la valeur 37 pour la températures du corps humain est fausse.

Proprétés importantes de la loi t de Student

1. La loi t de Student est différente pour différentes tailles d'échantillon.
2. La loi t de Student a une moyenne de $t = 0$ (tout comme la loi normale a une moyenne $Z = 0$).
3. L'ecart type (σ_t) de la loi t varie avec la taille de l'échantillon mais il est plus grand que 1 $(\sigma_t > 1)$ (contrairement à la loi normale pour laquelle $\sigma = 1$).

Exemple 2.2.7. *Supposons que vous devriez construire un intervalle de confiance pour estimer la moyenne μ d'une population. Utilisez les données fournies pour déterminer si la marge d'erreur E doit être calculée avec une valeur critique $Z_{\frac{\alpha}{2}}$ (de la loi normale), avec une valeur critique $t_{\frac{\alpha}{2}}$ (de la loi t de Student) ou non calculée par ces résultats.*

1. $n = 105$; $\bar{x} = 100$; $s = 15$ *et la population a une distribution asymétrique.*
2. $n = 8$; $\bar{x} = 100$; $s = 15$ *et la population suit une loi normale.*
3. $n = 8$; $\bar{x} = 100$; $s = 15$ *et la population a une distribution trés asymétrique.*

Solution 6.
1. *Parceque l'écart type σ de la population n'est pas connu et $n > 30$. La marge d'erreur E est calculée avec $t_{\frac{\alpha}{2}}$.*
2. *Puisque σ est inconnu et Ω suit la loi normale. La marge d'ereur E est calculée avec $t_{\frac{\alpha}{2}}$.*
3. *Comme l'échantillon est petit $(n \leq 30)$ et que la population ne suit pas une loi normale, la marge d'erreur ne devrait pas être calculée avec $t_{\frac{\alpha}{2}}$*

EXERCICES

2.2.3 Exercices

Exercice 2.2.1. *On a mesuré le poids de raisin par souche sur 10 souches prises au hasard dans une vigne. On a obtenu les résultats suivants (en kg) : 2,4 ; 3,2 ; 3,6 ; 4,1 ; 4,3 ; 4,7 ; 5,4 ; 5,9 ; 6,5 ; 6,9.*

1. *Déterminez une estimation ponctuelle non biaisée de la moyenne de la population dont ces souches sont extraites.*
2. *Donnez un intervalle de confiance de la moyenne de la population au risque de 0,05, en supposant que le poids de raisin par souche suit une loi normale au niveau de la vigne.*

Exercice 2.2.2. *Les données suivantes ont été obtenues sur des échantillons d'individus d'une région d'Europe. Le caractère étudié est le poids du cerveau exprimé en grammes pour des sujets de 20 à 49 ans.*

Hommes

centres de classes	1170	1220	1270	1320	1370	1420	1470	Total
effectifs	4	36	45	50	61	49	19	265

Femmes :

centres de classes	1070	1120	1170	1220	1270	1370	Total	
effectifs	12	22	45	54	52	20	10	215

Déterminez un intervalle de confiance au risque de 1%

1. *Pour la moyenne de la population des hommes ;*
2. *Pour la moyenne de la population des femmes.*