Estimations ponctuelles – Fiche de cours

1. Echantillonnage

Un échantillon de taille *n* est une liste de résultats obtenus pour *n* répétitions identiques et indépendantes.

On va constituer des échantillons pour essayer d'estimer le plus précisément possible une moyenne ou une proportion dans une population.

2. Estimation ponctuelle

a. Définition

Un estimateur est une fonction des variables observées :

$$T(X_1, X_2, \ldots X_n)$$

b. Sommes de variables aléatoires

Soit X une variable aléatoire de moyenne E(X), de variance V(X) et d'écart type σ_X

Soient n variables aléatoires X_1, X_2, ...X_n distribuées identiquement suivant la même loi de probabilité

On appelle M_n la moyenne de l'échantillon définie par :

$$M_n = \frac{X_1 + X_2 + \ldots X_n}{n}$$

On appelle $E(M_n)$ espérance mathématique de la moyenne de l'échantillon définie par :

$$E(M_n) = E(X) \quad \text{ou} \quad \mu = m$$

On appelle $V(M_n)$ la variance de la moyenne de l'échantillon définie par :

$$V(M_n) = \frac{V(X_1) + V(X_2) + \ldots V(X_n)}{n^2} \quad ; \quad V(M_n) = \frac{V(X)}{n} \quad \text{ou} \quad s^2 = \frac{\sigma^2}{n}$$

On appelle $\sigma(M_n)$ l'écart type de la moyenne de l'échantillon définie par :

$$\sigma(M_n) = \frac{\sigma(X)}{\sqrt{n}}$$

c. Qualités d'un estimateur

- **biais**

Dans une série statistique, un biais est une erreur systématique ; un estimateur sans biais est centré autour de la vraie valeur

- **convergence**

Un estimateur est convergent lorsqu'il tend vers une même valeur pour un nombre d'échantillon grand ; sa variance tend vers 0

d. Estimation d'une moyenne

- **dans une population**

La moyenne estimée dans une population de N individus vaut :

$$\mu = \frac{x_1 + x_2 + \ldots + x_N}{N}$$

- **dans un échantillon**

La moyenne estimée dans un échantillon de n individus vaut :

$$m = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

e. Estimation d'une variance

- **dans une population**

La variance estimée dans une population de N individus vaut :

$$\sigma^2 = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \ldots + (x_N - \mu)^2}{N}$$

- **dans un échantillon**

La variance estimée dans un échantillon de n individus vaut :

$$s^2 = \frac{(x_1 - m)^2 + (x_2 - m)^2 + \ldots + (x_n - m)^2}{n-1}$$

3. Intervalle de confiance

a. Définition

Un intervalle de confiance au risque α contient la vraie valeur d'une population dans $(1-\alpha)$ des échantillons

b. Théorème central limite

Lors de l'échantillonnage d'une loi de probabilité sous condition de $n \geq 30$, la moyenne des échantillons tend vers une loi normale

La loi binomiale converge vers la loi normale lorsque :

$$n \geq 30 \quad ; \quad np \geq 5 \quad ; \quad n(1-p) \geq 5$$

La loi normale va servir à évaluer la largueur de l'intervalle de confiance par rapport au risque α souhaité

- pour $\alpha = 1\%$ $z_{\frac{\alpha}{2}} = 2{,}58$ - pour $\alpha = 4{,}5\%$ $z_{\frac{\alpha}{2}} = 2$
- pour $\alpha = 5\%$ $z_{\frac{\alpha}{2}} = 1{,}96$ - pour $\alpha = 31\%$ $z_{\frac{\alpha}{2}} = 1$

c. Intervalle de confiance d'une moyenne

Il est possible d'estimer au risque α la vraie valeur de la moyenne d'une population lors d'un échantillonnage de taille n, avec l'intervalle :

$$IC_{1-\alpha} = \left[m - z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} ; m + z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \right]$$

d. Intervalle de confiance d'une proportion

Il est possible d'estimer au risque α la vraie valeur de la proportion d'une population suivant une loi binomiale lors d'un échantillonnage de taille n, avec l'intervalle :

$$IC_{1-\alpha} = \left[p - z_{\frac{\alpha}{2}} \frac{\sqrt{p(1-p)}}{\sqrt{n}} ; p + z_{\frac{\alpha}{2}} \frac{\sqrt{p(1-p)}}{\sqrt{n}} \right]$$

e. Facteurs influençant l'intervalle de confiance

- si α est grand, alors la largeur de l'intervalle de confiance augmente (estimation moins précise)
- si n est grand alors la largeur de l'intervalle de confiance diminue (estimation plus précise)