Leçon d'oral du CAPES de mathématiques

Estimation, ponctuelle ou par intervalle de confiance, d'un paramètre, tests d'hypothèse

Table des matières

1 Introduction et problématique **3**

2 Prérequis et objectifs **3**

2.1 Prérequis 3

2.2 Objectifs 3

3 Cadre probabiliste et échantillonnage **4**

3.1 Population, modèle et échantillon 4

3.2 Statistique et paramètre 4

4 Estimation ponctuelle **5**

4.1 Définition 5

4.2 Biais et variance 5

4.3 Exemple : estimation d'une moyenne 5

4.4 Convergence et consistance 6

5 Estimation par intervalle de confiance **6**

5.1 Principe 6

5.2 Intervalle de confiance d'une moyenne : variance connue 7

5.3 Intervalle de confiance d'une moyenne : variance inconnue 7

5.4 Intervalle de confiance d'une proportion 8

5.5 Exemple numérique 8

5.6 Choisir la taille de l'échantillon 9

6 Tests d'hypothèse **9**

6.1 Principe général 9

6.2 Test bilatéral et test unilatéral 9

6.3 Risque de première espèce 10

6.4 Risque de seconde espèce et puissance 10

6.5 Région critique et valeur critique 10

6.6 La valeur-p 10
6.7 Exemple : test sur une moyenne 11
6.8 Lien entre intervalle de confiance et test 11

7 Limites et précautions d'interprétation **12**
7.1 Les hypothèses du modèle 12
7.2 Significativité statistique et importance pratique 12
7.3 Multiplicité des tests 12
7.4 Interprétation des intervalles 12

8 Pistes pédagogiques pour l'oral du CAPES **12**
8.1 Objectifs d'apprentissage 12
8.2 Activité possible : estimation d'une proportion 13
8.3 Outils numériques 13

9 Conclusion **13**

1 Introduction et problématique

La statistique inférentielle vise à tirer des conclusions sur une population à partir d'observations effectuées sur un échantillon. Contrairement à la statistique descriptive, elle cherche notamment à estimer des paramètres inconnus et à confronter des hypothèses à des données.

On suppose que les observations sont issues d'un modèle probabiliste dépendant d'un paramètre inconnu θ. À partir des données observées, on souhaite répondre à trois questions :

— Quelle valeur peut-on raisonnablement attribuer à θ ? C'est le problème de l'estimation ponctuelle.
— Quelle plage de valeurs est compatible avec les données, au niveau de confiance choisi ? C'est le problème de l'estimation par intervalle de confiance.
— Les données permettent-elles de remettre en cause une hypothèse formulée sur θ ? C'est le problème des tests d'hypothèse.

Ces trois démarches sont complémentaires. Elles nécessitent de distinguer les fluctuations d'échantillonnage, les hypothèses du modèle et les risques d'erreur liés à toute décision statistique.

La problématique de la leçon est donc la suivante :

Comment exploiter un échantillon pour estimer un paramètre inconnu et décider, avec une maîtrise explicite des risques, si une hypothèse statistique doit être rejetée ?

Les notions présentées s'appuient notamment sur les ouvrages de Saporta [1], Dagnelie [2] et Casella et Berger [3].

2 Prérequis et objectifs

2.1 Prérequis

La compréhension de cette leçon nécessite de maîtriser les notions suivantes :

— **Probabilités :** univers, événements, probabilités conditionnelles, indépendance et variables aléatoires.
— **Lois de probabilité :** loi de Bernoulli, loi binomiale, loi normale et loi normale centrée réduite.
— **Statistique descriptive :** effectifs, fréquences, moyenne, variance et écart-type.
— **Échantillonnage :** échantillon aléatoire, fluctuations d'échantillonnage, moyenne empirique et fréquence observée.
— **Outils mathématiques :** espérance, variance, fonction de répartition, quantiles et calcul d'inégalités.

Pour les développements plus théoriques, on suppose également connues les notions de convergence en probabilité, de biais d'un estimateur et de loi de Student.

2.2 Objectifs

À l'issue de cette leçon, on doit être capable de :

— distinguer un paramètre inconnu, un estimateur aléatoire et l'estimation obtenue à partir d'un échantillon ;
— caractériser un estimateur à l'aide de son biais, de sa variance et, lorsque cela est pertinent, de sa convergence ;

— construire et interpréter un intervalle de confiance pour une moyenne ou une proportion, en précisant les hypothèses nécessaires ;
— comprendre l'influence de la taille de l'échantillon, de la variabilité des données et du niveau de confiance sur la précision d'une estimation ;
— formuler une hypothèse nulle et une hypothèse alternative, puis choisir une statistique de test adaptée ;
— distinguer les risques de première et de seconde espèce, comprendre la notion de puissance et interpréter une valeur-p ;
— relier, dans les situations appropriées, un test bilatéral à un intervalle de confiance ;
— interpréter les conclusions statistiques en tenant compte des hypothèses du modèle et des limites des méthodes employées.

L'objectif général est de comprendre comment les données d'un échantillon permettent d'estimer un paramètre inconnu et de prendre une décision statistique en maîtrisant les risques d'erreur associés.

3 Cadre probabiliste et échantillonnage

3.1 Population, modèle et échantillon

On considère une population décrite par une loi de probabilité dépendant d'un paramètre inconnu $\theta \in \Theta$.

On prélève un échantillon de taille n :

$$X_1, \dots, X_n.$$

Dans le cadre usuel de l'échantillonnage aléatoire simple, les variables $X_1, \dots, X_n$ sont supposées indépendantes et de même loi, notée P_θ.

Le paramètre θ peut être, par exemple :

— une proportion p ;
— une espérance μ ;
— une variance σ^2 ;
— un paramètre d'une loi de probabilité.

Les observations effectivement recueillies sont notées $x_1, \dots, x_n$. Avant leur observation, les données sont aléatoires ; après leur observation, elles sont connues.

Cette distinction est essentielle pour comprendre les intervalles de confiance et les tests.

3.2 Statistique et paramètre

Définition 1. *Une statistique est une fonction mesurable de l'échantillon qui ne dépend pas du paramètre inconnu :*

$$T_n = T(X_1, \dots, X_n).$$

Une statistique peut servir à estimer un paramètre ou à construire une procédure de décision.

Par exemple, la moyenne empirique

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

est une statistique utilisée pour estimer l'espérance μ.

La variance empirique corrigée est

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2, \qquad n \geq 2.$$

Lorsque les observations sont indépendantes, de même loi, d'espérance μ et de variance finie σ^2, on a

$$\mathrm{E}(\overline{X}_n) = \mu, \qquad \mathrm{Var}(\overline{X}_n) = \frac{\sigma^2}{n}.$$

La moyenne empirique se rapproche donc de l'espérance lorsque la taille de l'échantillon augmente, sous les hypothèses appropriées.

4 Estimation ponctuelle

4.1 Définition

Définition 2. *Un estimateur d'un paramètre θ est une statistique T_n destinée à fournir une approximation de ce paramètre. Sa réalisation $T_n(x_1, \ldots, x_n)$ constitue une estimation ponctuelle.*

Il faut distinguer l'estimateur, qui est une variable aléatoire, et l'estimation, qui est la valeur obtenue après observation de l'échantillon.

4.2 Biais et variance

Définition 3. *Le biais d'un estimateur T_n de θ est défini par*

$$\mathrm{Biais}_\theta(T_n) = \mathrm{E}_\theta(T_n) - \theta.$$

L'estimateur est sans biais si

$$\mathrm{E}_\theta(T_n) = \theta$$

pour tout paramètre admissible θ.

Le biais mesure l'écart entre l'espérance de l'estimateur et la valeur recherchée.

La variance mesure la dispersion de l'estimateur autour de son espérance.

Propriété 1. *Lorsque l'espérance et la variance de T_n existent,*

$$\mathrm{E}_\theta[(T_n - \theta)^2] = \mathrm{Var}_\theta(T_n) + \big(\mathrm{Biais}_\theta(T_n)\big)^2.$$

Le membre de gauche est l'erreur quadratique moyenne. Cette décomposition montre qu'un estimateur peut être évalué à partir de son biais et de sa variance.

4.3 Exemple : estimation d'une moyenne

Soit $X_1, \ldots, X_n$ un échantillon indépendant et identiquement distribué, d'espérance μ et de variance finie σ^2.

On utilise l'estimateur

$$\widehat{\mu}_n = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

On a

$$\mathrm{E}(\widehat{\mu}_n) = \mu, \qquad \mathrm{Var}(\widehat{\mu}_n) = \frac{\sigma^2}{n}.$$

Cet estimateur est donc sans biais. Son erreur quadratique moyenne vaut

$$\mathrm{E}[(\widehat{\mu}_n - \mu)^2] = \frac{\sigma^2}{n}.$$

Elle tend vers zéro lorsque n tend vers l'infini.

4.4 Convergence et consistance

Un estimateur T_n est convergent en probabilité vers θ si, pour tout $\varepsilon > 0$,

$$P_\theta(|T_n - \theta| > \varepsilon) \longrightarrow 0 \quad \text{lorsque } n \to +\infty.$$

On parle alors de consistance de l'estimateur.

Sous les hypothèses précédentes, la loi faible des grands nombres assure que

$$\overline{X}_n \xrightarrow{\mathbb{P}} \mu.$$

La convergence en probabilité exprime que la probabilité d'une erreur supérieure à une tolérance fixée tend vers zéro.

Il convient de ne pas confondre absence de biais et consistance : ces propriétés sont distinctes, même si elles peuvent être réunies pour un même estimateur.

5 Estimation par intervalle de confiance

5.1 Principe

Une estimation ponctuelle ne donne pas, à elle seule, d'indication sur la précision de l'estimation. On cherche donc à construire un intervalle aléatoire $I(X_1, \ldots, X_n)$ tel que

$$P_\theta(\theta \in I(X_1, \ldots, X_n)) = 1 - \alpha,$$

ou au moins une probabilité supérieure ou égale à $1 - \alpha$.

Le nombre $1 - \alpha$ est appelé niveau de confiance nominal et α le risque de non-couverture associé à la procédure.

Définition 4. *Un intervalle de confiance de niveau $1 - \alpha$ pour θ est un intervalle aléatoire construit à partir de l'échantillon dont la probabilité de contenir la vraie valeur du paramètre est égale à $1 - \alpha$, pour tout paramètre admissible, ou au moins égale à cette valeur dans le cas d'une couverture conservative.*

Une fois l'échantillon observé, les bornes de l'intervalle sont fixées. Dans l'interprétation fréquentiste, le paramètre θ reste fixe : c'est l'intervalle qui varie d'un échantillon à l'autre.

Ainsi, un intervalle de confiance à 95 % ne signifie pas qu'après observation il existe une probabilité de 95 % que le paramètre fixe appartienne à cet intervalle. Il signifie que la procédure utilisée couvre le paramètre dans 95 % des échantillons, lorsque le modèle est correct.

5.2 Intervalle de confiance d'une moyenne : variance connue

On suppose que

$$X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2),$$

où σ est connue et μ inconnue.

Alors

$$Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1).$$

On note $z_{1-\alpha/2}$ le quantile d'ordre $1-\alpha/2$ de la loi normale centrée réduite. On a

$$P\left(-z_{1-\alpha/2} \leq Z \leq z_{1-\alpha/2}\right) = 1-\alpha.$$

Après réarrangement, on obtient l'intervalle de confiance

$$\boxed{\left[\overline{X}_n - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},\ \overline{X}_n + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right].}$$

Au niveau de confiance de 95 %,

$$z_{0{,}975} \simeq 1{,}96.$$

L'intervalle s'écrit alors

$$\boxed{\left[\overline{X}_n - 1{,}96\frac{\sigma}{\sqrt{n}},\ \overline{X}_n + 1{,}96\frac{\sigma}{\sqrt{n}}\right].}$$

Sa largeur est

$$L = 2z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

Elle diminue lorsque la taille de l'échantillon augmente.

5.3 Intervalle de confiance d'une moyenne : variance inconnue

Lorsque les observations proviennent d'une loi normale et que la variance est inconnue, on utilise la statistique de Student :

$$T = \frac{\overline{X}_n - \mu}{S_n/\sqrt{n}},$$

où S_n^2 est la variance empirique corrigée.

Cette statistique suit une loi de Student à $n-1$ degrés de liberté :

$$T \sim t_{n-1}.$$

En notant $t_{n-1;1-\alpha/2}$ le quantile correspondant, un intervalle de confiance de niveau $1-\alpha$ est

$$\boxed{\left[\overline{X}_n - t_{n-1;1-\alpha/2}\frac{S_n}{\sqrt{n}},\ \overline{X}_n + t_{n-1;1-\alpha/2}\frac{S_n}{\sqrt{n}}\right].}$$

La normalité de la population est ici une hypothèse importante pour obtenir ce résultat exact à taille finie.

Lorsque l'échantillon est suffisamment grand, des approximations asymptotiques peuvent être utilisées dans d'autres situations, sous des hypothèses adaptées.

5.4 Intervalle de confiance d'une proportion

On considère une population dans laquelle la proportion d'individus présentant un caractère donné est p.

On prélève un échantillon aléatoire simple de taille n. Si X_i vaut 1 lorsque le caractère est présent et 0 sinon, alors

$$X_i \sim \mathcal{B}(p), \qquad \widehat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

On a

$$\mathrm{E}(\widehat{p}) = p, \qquad \mathrm{Var}(\widehat{p}) = \frac{p(1-p)}{n}.$$

Sous des conditions d'approximation normale, on utilise

$$\frac{\widehat{p} - p}{\sqrt{p(1-p)/n}} \approx \mathcal{N}(0,1).$$

En pratique, l'intervalle de Wald obtenu en remplaçant p par $\widehat{p}$ est

$$\boxed{\left[\widehat{p} - z_{1-\alpha/2}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}},\ \widehat{p} + z_{1-\alpha/2}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}\right].}$$

Cette formule est une approximation et peut donner des résultats peu satisfaisants lorsque l'échantillon est petit ou lorsque $\widehat{p}$ est proche de 0 ou de 1.

Pour ces situations, on peut préférer un intervalle de Wilson ou un intervalle exact binomial, en précisant la méthode employée.

5.5 Exemple numérique

Une enquête auprès de $n = 400$ personnes révèle que 240 d'entre elles déclarent utiliser un service donné.

La proportion observée est

$$\widehat{p} = \frac{240}{400} = 0,60.$$

Avec l'approximation de Wald au niveau de confiance de 95 %, la marge d'erreur est

$$1,96\sqrt{\frac{0,60 \times 0,40}{400}} \simeq 0,048.$$

On obtient l'intervalle approximatif

$$[0,552\,;\,0,648].$$

Cet intervalle donne une estimation de la proportion dans la population comprise entre environ 55,2 % et 64,8 %, selon la méthode retenue et sous les hypothèses du modèle.

Cette conclusion suppose notamment que le plan d'échantillonnage et les observations permettent l'utilisation de l'approximation.

5.6 Choisir la taille de l'échantillon

Pour estimer une moyenne avec variance connue, on souhaite obtenir une marge d'erreur au plus égale à ε. Il faut donc imposer

$$z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \varepsilon.$$

Cela conduit à

$$\boxed{n \geq \left(\frac{z_{1-\alpha/2}\sigma}{\varepsilon}\right)^2.}$$

Cette relation permet de dimensionner une enquête à partir de la précision souhaitée, du niveau de confiance et de la dispersion supposée des observations.

Pour une proportion, on peut utiliser la borne $p(1-p) \leq 1/4$ afin d'obtenir un dimensionnement prudent fondé sur l'approximation normale :

$$n \geq \frac{z^2_{1-\alpha/2}}{4\varepsilon^2}.$$

Cette formule ne dispense pas de vérifier les conditions d'application de l'approximation.

6 Tests d'hypothèse

6.1 Principe général

Un test d'hypothèse est une procédure permettant de prendre une décision à partir d'un échantillon.

On formule deux hypothèses :

— l'hypothèse nulle H_0, considérée comme référence ;

— l'hypothèse alternative H_1, qui exprime une différence ou un effet recherché.

Par exemple, pour une moyenne inconnue μ, on peut poser

$$H_0 : \mu = \mu_0 \qquad \text{contre} \qquad H_1 : \mu \neq \mu_0.$$

Un test repose sur une statistique de test dont la loi est connue sous H_0, exactement ou approximativement.

On détermine une région critique : si la valeur observée de la statistique appartient à cette région, on rejette H_0 ; sinon, on ne rejette pas H_0.

6.2 Test bilatéral et test unilatéral

Pour un paramètre réel θ, trois formes classiques d'hypothèses alternatives sont possibles :

$$\begin{aligned}
&\text{Test bilatéral :} && H_0 : \theta = \theta_0, \quad H_1 : \theta \neq \theta_0,\\
&\text{Test unilatéral à droite :} && H_0 : \theta = \theta_0, \quad H_1 : \theta > \theta_0,\\
&\text{Test unilatéral à gauche :} && H_0 : \theta = \theta_0, \quad H_1 : \theta < \theta_0.
\end{aligned}$$

Le choix dépend de la question scientifique posée et doit être effectué avant l'analyse des résultats.

6.3 Risque de première espèce

Le risque de première espèce, noté α, est la probabilité de rejeter H_0 alors que H_0 est vraie.

Un test de niveau α vérifie

$$P_\theta(\text{rejet de } H_0) \leq \alpha \qquad \text{pour tout } \theta \in \Theta_0,$$

où Θ_0 désigne l'ensemble des paramètres compatibles avec l'hypothèse nulle.

Dans le cas simple où H_0 fixe entièrement la loi, cette condition se réduit au contrôle de la probabilité de rejet sous cette loi.

Le niveau de 5 % est fréquemment utilisé, mais il ne constitue pas une valeur universelle.

6.4 Risque de seconde espèce et puissance

Le risque de seconde espèce, noté $\beta(\theta)$, est la probabilité de ne pas rejeter H_0 alors que le paramètre vaut une valeur θ compatible avec l'alternative :

$$\beta(\theta) = P_\theta(\text{non-rejet de } H_0), \qquad \theta \in \Theta_1.$$

La puissance du test est

$$\boxed{\pi(\theta) = 1 - \beta(\theta).}$$

Elle représente la probabilité de rejeter H_0 lorsque l'alternative considérée est vraie.

À niveau fixé, augmenter la taille de l'échantillon permet souvent d'améliorer la puissance, à condition que les autres hypothèses du modèle soient satisfaites.

6.5 Région critique et valeur critique

La région critique est l'ensemble des valeurs de la statistique de test qui conduisent au rejet de H_0.

Pour un test bilatéral fondé sur une loi normale centrée réduite, au niveau α, la région critique est

$$\boxed{|Z| > z_{1-\alpha/2}.}$$

Pour un test unilatéral à droite, elle est

$$\boxed{Z > z_{1-\alpha}.}$$

Pour un test unilatéral à gauche, elle est

$$\boxed{Z < z_{\alpha}.}$$

Le choix des seuils dépend de la loi de la statistique sous l'hypothèse nulle et du niveau du test.

6.6 La valeur-p

La valeur-p est la probabilité, calculée sous H_0, d'obtenir une statistique de test au moins aussi incompatible avec H_0 que celle observée, selon la définition de l'alternative.

Pour un test bilatéral fondé sur une loi normale centrée réduite, la valeur-p est

$$p_{\text{val}} = 2(1 - \Phi(|z_{\text{obs}}|)),$$

où Φ est la fonction de répartition de la loi normale centrée réduite.

La règle de décision usuelle est :

$$\begin{cases} p_{\text{val}} \leq \alpha & \Longrightarrow \text{rejet de } H_0, \\ p_{\text{val}} > \alpha & \Longrightarrow \text{non-rejet de } H_0. \end{cases}$$

La valeur-p n'est ni la probabilité que H_0 soit vraie, ni la probabilité que les résultats soient dus au hasard. Elle s'interprète toujours relativement au modèle et à l'hypothèse nulle.

6.7 Exemple : test sur une moyenne

On suppose que les observations sont indépendantes et suivent une loi normale de moyenne μ et d'écart-type connu $\sigma = 2$.

On prélève un échantillon de taille $n = 25$ et l'on obtient une moyenne observée $\overline{x} = 10,8$.

On souhaite tester au niveau de 5 % :

$$H_0 : \mu = 10 \qquad \text{contre} \qquad H_1 : \mu \neq 10.$$

Sous H_0, la statistique de test est

$$Z = \frac{\overline{X}_n - 10}{2/\sqrt{25}} \sim \mathcal{N}(0,1).$$

La valeur observée vaut

$$z_{\text{obs}} = \frac{10,8 - 10}{2/5} = 2.$$

Pour un test bilatéral de niveau 5 %, la région critique est définie par

$$|Z| > 1,96.$$

Comme $|2| > 1,96$, on rejette H_0 au niveau de 5 %.

La valeur-p est approximativement

$$p_{\text{val}} = 2(1 - \Phi(2)) \simeq 0,0455.$$

La conclusion est que les données sont incompatibles avec l'hypothèse $\mu = 10$ au seuil de 5 %, dans le cadre du modèle retenu.

Cela ne prouve pas que l'hypothèse nulle est fausse et ne mesure pas l'importance pratique de l'écart observé.

6.8 Lien entre intervalle de confiance et test

Dans certaines situations, un intervalle de confiance et un test bilatéral sont deux formulations d'une même procédure.

Par exemple, pour une moyenne normale avec variance connue, le test de niveau α de

$$H_0 : \mu = \mu_0 \qquad \text{contre} \qquad H_1 : \mu \neq \mu_0$$

rejette H_0 si et seulement si μ_0 n'appartient pas à l'intervalle de confiance bilatéral de niveau $1 - \alpha$:

$$\left[\overline{X}_n - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}},\ \overline{X}_n + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right].$$

Ce lien repose sur la construction particulière du test et de l'intervalle. Il ne faut pas le généraliser sans vérifier que les deux procédures sont effectivement associées.

7 Limites et précautions d'interprétation

7.1 Les hypothèses du modèle

Les résultats précédents reposent sur des hypothèses qui doivent être discutées :

- représentativité et mode de sélection de l'échantillon ;
- indépendance éventuelle des observations ;
- adéquation du modèle probabiliste aux données ;
- validité des approximations asymptotiques ;
- prise en compte des biais de mesure et des données manquantes, lorsqu'ils existent.

Un calcul correct ne compense pas un échantillonnage biaisé ou un modèle inadapté.

7.2 Significativité statistique et importance pratique

Une différence peut être statistiquement significative sans être importante en pratique. Inversement, une différence qui n'est pas statistiquement significative ne doit pas être interprétée comme une preuve d'égalité.

L'interprétation doit prendre en compte la taille de l'effet, la précision de l'estimation, la puissance du test et le contexte scientifique ou professionnel.

7.3 Multiplicité des tests

Lorsque de nombreux tests sont réalisés, le risque d'obtenir au moins un résultat significatif par hasard peut augmenter.

Il convient alors d'envisager des procédures adaptées de contrôle du risque global ou du taux de fausses découvertes, selon l'objectif de l'étude.

7.4 Interprétation des intervalles

Un intervalle étroit indique une meilleure précision dans le cadre de la procédure utilisée, mais ne garantit pas l'absence de biais.

La largeur d'un intervalle dépend notamment de la taille de l'échantillon, de la variabilité des observations et du niveau de confiance choisi.

Un niveau de confiance plus élevé conduit généralement à un intervalle plus large.

8 Pistes pédagogiques pour l'oral du CAPES

8.1 Objectifs d'apprentissage

Au lycée, selon les programmes et les niveaux concernés, les situations d'échantillonnage permettent d'introduire les fluctuations d'une proportion observée, la variabilité des estimations et la prise de décision à partir de données.

Les objectifs pédagogiques peuvent être les suivants :

- distinguer population, échantillon et fréquence observée ;
- comprendre qu'une estimation varie d'un échantillon à l'autre ;
- interpréter une marge d'erreur et un intervalle de confiance ;
- utiliser une simulation pour illustrer la fréquence de couverture d'une procédure ;

— comprendre la logique d'un test et les risques associés à une décision.

Il est important d'adapter le formalisme aux programmes en vigueur et de distinguer les activités de simulation des démonstrations mathématiques générales.

8.2 Activité possible : estimation d'une proportion

On souhaite estimer la proportion p de personnes possédant un certain caractère dans une population.

On simule plusieurs échantillons indépendants de taille n, puis on calcule pour chacun la fréquence observée

$$\widehat{p} = \frac{X}{n}, \qquad X \sim \mathcal{B}(n, p).$$

On peut alors :

1. représenter la distribution des fréquences observées ;
2. comparer leur dispersion pour différentes valeurs de n ;
3. construire un intervalle selon une méthode donnée ;
4. répéter l'expérience et mesurer la fréquence de couverture des intervalles ;
5. discuter les limites de l'approximation normale.

La simulation permet de donner un sens fréquentiste à la notion de niveau de confiance, sans remplacer les justifications théoriques.

8.3 Outils numériques

Un tableur, un logiciel de calcul formel ou un langage de programmation peut servir à simuler des échantillons, calculer des intervalles et visualiser les distributions d'une statistique.

L'outil numérique doit soutenir le raisonnement : les hypothèses, la loi utilisée, la méthode de calcul et l'interprétation des résultats restent à expliciter.

9 Conclusion

L'estimation ponctuelle fournit une valeur destinée à approcher un paramètre inconnu. L'étude du biais, de la variance et de la convergence permet d'évaluer les propriétés de l'estimateur.

L'intervalle de confiance complète cette démarche en donnant une indication probabiliste sur la couverture d'une procédure d'estimation. Son interprétation repose sur la répétition théorique de l'échantillonnage et sur la validité du modèle.

Le test d'hypothèse répond à une autre question : il formalise une décision à partir des données, en contrôlant notamment le risque de première espèce. La puissance et le risque de seconde espèce permettent d'étudier les limites de cette décision.

Ces trois outils ont un objectif commun : tirer des conclusions raisonnées à partir de données incertaines, sans confondre résultat numérique, hypothèse de modèle et certitude scientifique.

Références

[1] G. Saporta, *Probabilités, analyse des données et statistique*, 3e édition, Technip, 2011.

Référence générale pour les probabilités, l'échantillonnage, l'estimation et les méthodes statistiques.

[2] P. Dagnelie, *Statistique théorique et appliquée*, tome 2 : Inférence statistique à une et à deux dimensions, 2e édition, De Boeck, 2013.

Référence pour les estimateurs, les intervalles de confiance et les tests d'hypothèse.

[3] G. Casella et R. L. Berger, *Statistical Inference*, 2e édition, Duxbury, 2002.

Ouvrage de référence pour les fondements théoriques de l'inférence statistique, les propriétés des estimateurs et les procédures de test.

[4] B. Illowsky et S. Dean, *Introductory Statistics 2e*, OpenStax, Rice University, 2023. `https://openstax.org/details/books/introductory-statistics-2e`

Ressource pédagogique en accès libre pour les intervalles de confiance, les tests d'hypothèse et leur interprétation.