

Corrigé et barème – Échantillonnage et estimation

Intervalles de confiance et fluctuation d'échantillonnage — programme officiel Terminale générale
(option Maths complémentaires)

Durée : 120 min – Calculatrice interdite – Sur 20 points · Maths complémentaires (option Tle) Terminale

Exercice 1 – Un taux de retard contesté

(4 pts)

- a) $n = 325 \geq 30$; $np = 325 \times 0,16 = 52 \geq 5$; $n(1 - p) = 325 \times 0,84 = 273 \geq 5$: les trois conditions sont remplies. (1)
- b) $\sqrt{\frac{0,16 \times 0,84}{325}} = \sqrt{\frac{0,1344}{325}} \approx 0,020\,34$; marge $1,96 \times 0,020\,34 \approx 0,039\,9$; $I_{325} \approx [0,120 ; 0,200]$. (1)
- c) $f = \frac{74}{325} \approx 0,227\,7 > 0,200$, donc $f \notin I_{325}$: on rejette au seuil de 5 % l'annonce des 16 %, le taux réel de retard semble supérieur. (1)
- d) 22,77 % est une fréquence observée sur un seul échantillon, soumise à la fluctuation ; elle estime le taux réel sans le donner. Il faut l'assortir d'un intervalle de confiance : marge $\frac{1}{\sqrt{325}} \approx 0,055$, soit $[0,172 ; 0,283]$ au niveau 95 %. (1)

Erreur fréquente — Rejeter une annonce ne fournit pas la valeur exacte de la proportion : il faut ensuite estimer, avec un intervalle de confiance.

Exercice 2 – Estimation dans une médiathèque

(4 pts)

- a) $f = \frac{294}{784} = 0,375$; $n = 784 \geq 30$, $nf = 294 \geq 5$, $n(1 - f) = 490 \geq 5$: conditions remplies. (1)
- b) $\sqrt{784} = 28$, donc la marge vaut $\frac{1}{28} \approx 0,035\,7$, soit environ 3,6 points. (1)
- c) Bornes : $0,375 - 0,035\,7 \approx 0,339$ et $0,375 + 0,035\,7 \approx 0,411$, d'où $J' \approx [0,339 ; 0,411]$. (1)
- d) Un tiers vaut $\frac{1}{3} \approx 0,333$, valeur strictement inférieure à la borne basse 0,339 : toutes les valeurs de l'intervalle dépassent un tiers, l'affirmation est donc compatible avec l'enquête au niveau de confiance 95 %. (1)

Le point clé — Comparer un seuil à l'intervalle entier, et non à la seule fréquence observée, est ce qui donne sa valeur à la conclusion.

Exercice 3 – Dimensionner une enquête

(4 pts)

- a) $\sqrt{1\,521} = 39$, donc la marge vaut $\frac{1}{39} \approx 0,025\,6$, soit environ 2,6 points. (1)
- b) $\frac{1}{\sqrt{n}} \leq 0,012 \iff n \geq \frac{1}{0,000144} \approx 6\,944,4$, donc $n \geq 6\,945$ (vérification : $\frac{1}{\sqrt{6\,945}} \approx 0,012\,0$). (1)
- c) $n \geq \frac{1,96^2 \times 0,44 \times 0,56}{0,012^2} = \frac{3,8416 \times 0,246\,4}{0,000144} \approx 6\,573,4$, donc $n \geq 6\,574$: environ 371 personnes de moins, l'écart restant faible car $f = 0,44$ est proche de 0,5. (1)
- d) La marge varie en $\frac{1}{\sqrt{n}}$: doubler n la divise par $\sqrt{2} \approx 1,41$ seulement. Pour la diviser par 2, il faut quadrupler l'échantillon, soit $4 \times 1\,521 = 6\,084$ personnes. (1)

Repère — Chaque point de précision supplémentaire coûte de plus en plus cher : la précision progresse comme la racine carrée de l'effectif.

Exercice 4 – Simulation en Python

(4 pts)

- a) La fonction `simule` 196 tirages indépendants, chacun étant un succès avec la probabilité 0,45, et renvoie la fréquence de succès : k suit la loi binomiale $\mathcal{B}(196; 0,45)$, d'espérance $196 \times 0,45 = 88,2$. (1)
- b) $\sqrt{196} = 14$, donc $\frac{1}{\sqrt{196}} = \frac{1}{14} \approx 0,0714$. La variable c compte, parmi les 5 000 échantillons simulés, ceux dont l'intervalle de confiance simplifié contient la proportion réelle 0,45. (1)
- c) $\frac{4783}{5000} = 0,9566$: environ 95,7 % des échantillons simulés produisent un intervalle contenant 0,45, résultat cohérent avec le niveau de confiance de 95 % — et légèrement supérieur, la marge $\frac{1}{\sqrt{n}}$ étant calculée pour le cas le plus défavorable $f = 0,5$. (1)
- d) Chaque exécution simule de nouveaux échantillons : le comptage est lui-même une variable aléatoire qui fluctue. Pour stabiliser le résultat, il faut augmenter le nombre N de simulations, la fréquence obtenue se rapprochant alors de la probabilité réelle (loi des grands nombres). (1)

Attention — Un résultat de simulation s'annonce toujours avec un « environ » : deux exécutions ne donnent jamais le même nombre.

Exercice 5 – Synthèse : une annonce à vérifier puis à estimer

(4 pts)

- a) Conditions : $441 \geq 30$, $np = 123,48 \geq 5$, $n(1-p) = 317,52 \geq 5$. $\sqrt{\frac{0,28 \times 0,72}{441}} = \sqrt{\frac{0,2016}{441}} \approx 0,02138$; marge $1,96 \times 0,02138 \approx 0,0419$; $I_{441} \approx [0,238 ; 0,322]$. (1)
- b) $f = \frac{140}{441} \approx 0,3175$ appartient à $[0,238 ; 0,322]$: on ne rejette pas l'annonce au seuil de 5 %, l'écart avec 28 % s'expliquant par la fluctuation d'échantillonnage. (1)
- c) $\sqrt{441} = 21$, marge $\frac{1}{21} \approx 0,0476$; $J' \approx [0,270 ; 0,365]$. Cet intervalle contient 0,28 : les deux approches concordent, l'annonce de la plateforme reste une valeur plausible. (1)
- d) $\frac{1}{\sqrt{n}} \leq 0,02 \iff n \geq \frac{1}{0,0004} = 2500$, donc 2 500 comptes, soit $2500 - 441 = 2059$ comptes supplémentaires à examiner. (1)

À retenir — Fluctuation et confiance sont deux lectures du même échantillon : la première teste une valeur annoncée, la seconde délimite les valeurs plausibles.