

Rapport de relecture

13-echantillonnage.tex

Cours et exercices sur l'échantillonnage, Seconde générale

Synthèse

La relecture du fichier 13-echantillonnage.tex (568 lignes) a permis d'identifier les éléments suivants :

- **5 erreurs majeures** dont un sens d'inégalité inversé entraînant une conclusion finale fautive dans l'exercice « taux de passage en 1^{re} S et ES », une probabilité de référence incohérente avec l'énoncé (0,918 au lieu de 0,797), une incohérence entre énoncé et corrigé d'un programme Python (inégalités strictes vs larges), et une borne d'intervalle approchée fautive de 0,002 ;
- **4 incohérences** de rédaction (accord, espacement, parenthèses non fermées, expression maladroite « et la borne inférieure ») ;
- **4 coquilles** mineures (« caondidat-e », arrondis non conformes à la règle annoncée, etc.).

Ce chapitre, signalé dès l'attention liminaire (l. 3-9) comme étant hors-programme officiel pour aller au-delà des simulations informatiques attendues, présente la notion d'échantillon, d'intervalle de fluctuation au seuil de 95 % (avec la formule classique $I_f = [p - 1/\sqrt{n}; p + 1/\sqrt{n}]$ valable sous les conditions $n \geq 25$ et $0,2 \leq p \leq 0,8$), de prise de décision (le dé tétraédrique de l'exemple n'apparaît pas truqué à 100 lancers mais le devient à 10 000), suivis de huit exercices très variés (le dé d'Al, le Dédale, l'INSEE et la répartition femmes/hommes, une élection locale, le taux de réussite au bac, la planète Vailox, la pièce de Double Face, les taux de passage en 1^{re} S et ES, et un exercice de lecture de programme Python). L'immense majorité des calculs est exacte (vérifications systématiques avec Python) : intervalles [0,6584; 0,7216] pour le Dédale, [0,4829; 0,5461] pour 1000 personnes, [0,4647; 0,5353] pour les 800 électeurs, [0,7523; 0,8417] pour le bac, [0,4375; 0,5625] et [0,3375; 0,4625] pour la pièce truquée, $n \leq 204$ pour Vailox.

Tous les calculs (intervalles de fluctuation, inversions $1/\sqrt{n}$, encadrements pour déterminer la taille maximale d'un échantillon, frontières d'arrondi) mentionnés dans ce rapport ont été vérifiés au moyen de calcul numérique (Python avec `math.sqrt` et `Fraction`).

1 Erreurs mathématiques majeures

E1 Probabilité incohérente avec l'énoncé : $p = 0,918$ au lieu de $p = 0,797$ (l. 288)

L'énoncé de l'exercice « taux de réussite au bac » (l. 276) précise : « le taux de réussite au bac est en moyenne égal à 79,7 % ». Le début du corrigé écrit toutefois :

Indiqué : « Notons $p = 0,918$ la probabilité qu'un-e caondidat-e ait son bac. »

Correct : $p = 0,797$ (en cohérence avec le 79,7 % de l'énoncé). La valeur 0,918 apparaît sans aucun support dans l'énoncé. La formule de l'intervalle qui suit (l. 293) utilise heureusement la bonne valeur 0,797, et le calcul [0,7523; 0,8417] est correct.

E2 Sens d'inégalité inversé pour la 1^{re} ES (l. 430)

Dans l'exercice « taux de passage en 1^{re} S et ES », pour la 1^{re} ES, la résolution aboutit à $\sqrt{n} \leq \frac{1}{0,12}$, soit $n \leq \frac{1}{0,12^2} \approx 69,444$. La conclusion s'énonce :

Indiqué : « soit $n \leq \frac{1}{0,12^2}$, soit $n \leq 69,444$, donc $n \geq 70$ ».

Correct : $n \leq 69$ (puisque n est entier et $n \leq 69,444$, on obtient $n \leq 69$, et non $n \geq 70$). Le

sens d'inégalité est inversé dans la dernière étape.

E3 Conclusion finale fautive « entre 70 et 204 » (l. 433)

À la suite de E2, la conclusion globale s'énonce :

Indiqué : « Ainsi, le nombre d'élèves de Seconde dans le lycée Casimir doit être compris entre 70 et 204. »

Correct : pour que les deux taux observés (45 % et 48 %) soient acceptables au seuil de 95 %, il faut que n vérifie simultanément $n \leq 204$ (1^{re} S) et $n \leq 69$ (1^{re} ES), soit $n \leq 69$ (intersection des deux conditions, en prenant le minimum). En pratique, on ajoutera $n \geq 25$ pour la validité de la formule, donc $25 \leq n \leq 69$. La fourchette « entre 70 et 204 » correspond au contraire à des valeurs où la fréquence de 1^{re} ES n'est pas acceptable, ce qui est l'inverse du résultat cherché.

E4 Incohérence entre énoncé et corrigé du programme Python (l. 451 vs 498-504)

Le programme Python proposé dans l'énoncé contient (l. 451) :

Indiqué (énoncé, l. 451) : « `if (f>a) and (f<b)` » (inégalités strictes).

Indiqué (corrigé, l. 498) : « `if (f>=a) and (f<=b)` » (inégalités larges).

Correct : reproduire fidèlement le code de l'énoncé dans le corrigé. De plus, l'explication ligne 504 dit « Si $f > a$ ET $f < b$, c'est-à-dire si $f \in [a; b]$ », mais la conjonction stricte « $f > a$ ET $f < b$ » équivaut à $f \in]a; b[$ (intervalle ouvert), et non $[a; b]$ (fermé). Soit on adopte les strictes (et alors l'intervalle est ouvert), soit les larges (et alors l'intervalle est fermé), mais il faut être cohérent.

E5 Borne inférieure approchée fautive (l. 233)

Dans l'exercice INSEE pour 500 personnes, le calcul devrait donner $p - 1/\sqrt{500} \approx 0,5145 - 0,04472 \approx 0,4698$. La rédaction écrit :

Indiqué : « $I_f \approx [0,4678; 0,5592]$ ».

Correct : $I_f \approx [0,4698; 0,5592]$ (la borne inférieure 0,4678 donne $p - 0,4678 = 0,0467$, valeur différente de $1/\sqrt{500} \approx 0,04472$). La borne supérieure 0,5592 est correcte. La conclusion $f = 0,55 \in I_f$ reste valide quelle que soit la borne (puisque $0,55 \in [0,4678; 0,5592]$ comme $\in [0,4698; 0,5592]$).

2 Incohérences et rédaction

I1 Arrondis non conformes à la règle annoncée (l. 80)

La règle d'arrondi des bornes annoncée plus loin l. 296 est : « On arrondit toujours la borne inférieure *par excès* et la borne supérieure *par défaut*. ». Or, dans l'introduction l. 80, la rédaction écrit :

Indiqué : « $I \approx [0,066; 0,267]$ ».

Correct : avec $p = 1/6 = 0,1666\dots$ et $\delta = 0,1$, on obtient $p - \delta = 0,0666\dots$ et $p + \delta = 0,2666\dots$. La règle annoncée donnerait $[0,067; 0,266]$ (et non $[0,066; 0,267]$, qui inverse les deux arrondis). Les valeurs 0,067 et 0,266 sont d'ailleurs utilisées dans la formule SI ligne 86, ce qui confirme cette interprétation.

I2 « fréquence observé » : accord manquant (l. 147)

Indiqué : « car la fréquence **observé** est plus grande que la probabilité ».

Correct : « fréquence **observée** » (accord avec le nom féminin « fréquence »).

I3 Espacement absent après guillemet fermant (l. 178)

Indiqué : « La probabilité d'obtenir la face « 3 »**avec** un dé équilibré » (pas d'espace avant « avec »).

Correct : ajouter une espace insécable ou un espace fin après « » : « « 3 » **avec** un dé équilibré ».

I4 Parenthèse ouvrante non refermée naturellement (l. 521-526)

La phrase l. 521 ouvre une parenthèse : « ... compris entre 0 et 1 (la fonction `random()` du module *random* a cet objectif... », puis enchaîne avec une remarque qui occupe les l. 523-529, et la parenthèse fermante se trouve curieusement insérée à la fin d'une URL l. 526 (« `random.html` »). La structure est confuse pour le lecteur.

Correct : refermer la parenthèse avant la remarque, par exemple « ... a cet objectif). », et reformuler la remarque ensuite.

I5 Verbe manquant « et la borne inférieure » au lieu de « est la borne inférieure » (l. 542)

Indiqué : « $a = 0,3 - \frac{1}{\sqrt{n}}$ **et** la borne inférieure de l'intervalle de fluctuation ».

Correct : « $a = 0,3 - \frac{1}{\sqrt{n}}$ **est** la borne inférieure » (substitution « et » → « est », vraisemblablement coquille de saisie).

3 Coquilles et saisie**C1 « caondidat.e » au lieu de « candidat.e » (l. 288)**

Indiqué : « la probabilité qu'un.e **caondidat.e** ait son bac ».

Correct : « candidat.e » (lettre o intercalée par inadvertance entre le « a » et le « n »).

C2 Borne supérieure 0,5354 vs 0,5353 (l. 266)

Indiqué : « $I_f \approx [0,4647; \mathbf{0,5353}]$ ».

Correct : $0,5 + 1/\sqrt{800} \approx 0,53536\dots$, ce qui arrondit en standard à 0,5354. La valeur 0,5353 est cohérente avec la règle « borne supérieure par défaut » annoncée ligne 296 ; pas une erreur, mais à mentionner pour cohérence interne.

C3 Tournure familière dans une chaîne Python (l. 471)

Indiqué : « `print("Y'a comme qui dirait un bins quelque part...")` ».

Correct : choix stylistique (chaîne affichée à l'écran) ; à conserver ou non selon le degré de familiarité souhaité. Pas une erreur.

C4 « success » en anglais (l. 444, 462)

La fonction « **success** » est en anglais alors que tout le reste du chapitre est en français. Le choix peut être pédagogique (le mot « succès » pourrait poser problème avec l'accent dans certains éditeurs Python), mais il vaut la peine de signaler le terme.

4 Récapitulatif

Code	Objet	Ligne(s)	Sévérité
E1	$p = 0,918$ au lieu de $p = 0,797$ (taux de réussite au bac)	288	Majeure
E2	« $n \leq 69,444$ donc $n \geq 70$ » : sens d'inégalité inversé	430	Majeure
E3	Conclusion « n entre 70 et 204 » fausse ($n \leq 69$ correct)	433	Majeure
E4	Énoncé Python avec inégalités strictes mais corrigé avec inégalités larges, et description intervalle ouvert/fermé incohérente	451, 498, 504	Majeure
E5	Borne inférieure 0,4678 au lieu de 0,4698 pour $p = 0,5145$, $n = 500$	233	Majeure
I1	Arrondis [0,066 ; 0,267] ne respectant pas la règle l. 296	80	Moyenne
I2	« fréquence observé » au lieu de « observée »	147	Moyenne
I3	Espacement absent après guillemet fermant : « « 3 » avec »	178	Moyenne
I4	Parenthèse non fermée naturellement (intercalée par une remarque)	521-526	Moyenne
I5	« $a = \dots$ et la borne inférieure » au lieu de « est la borne inférieure »	542	Moyenne
C1	« caondidat.e » au lieu de « candidat.e »	288	Mineure
C2	Borne sup 0,5353 vs 0,5354 (arrondi par défaut cohérent avec la règle)	266	Mineure
C3	Tournure familière « Y'a comme qui dirait » dans print	471	Mineure
C4	Fonction « success » en anglais	444, 462	Mineure

Méthodologie : tous les résultats numériques et symboliques mentionnés dans ce rapport ont été recalculés et vérifiés au moyen de calcul numérique (Python avec `math.sqrt` et `Fraction`) avant d'être consignés. Les vérifications portent en particulier sur les intervalles [0,6584 ; 0,7216] (Dédale, $p = 0,69$, $n = 1000$), [0,4829 ; 0,5461] (INSEE, $p = 0,5145$, $n = 1000$), [0,4647 ; 0,5354] (élections, $p = 0,5$, $n = 800$), [0,7523 ; 0,8417] (bac, $p = 0,797$, $n = 500$), [0,4375 ; 0,5625] et [0,3375 ; 0,4625] (Double Face, $n = 256$), $f = 117/256 \approx 0,457$, ainsi que les seuils $n \leq 1/0,07^2 \approx 204,08$ (Vailox et 1^{re} S) et $n \leq 1/0,12^2 \approx 69,44$ (1^{re} ES). Les cinq anomalies majeures relevées concernent principalement la rédaction (probabilité affichée incohérente avec l'énoncé, sens d'inégalité inversé entraînant une

conclusion fausse, incohérence stricte/large dans le programme Python). Une fois ces points corrigés, l'ensemble du chapitre forme un cours complet et didactique pour l'introduction à l'échantillonnage en Seconde.