

Plaquette 3 - Concentration des fréquences

Concentration, loi des grands nombres — Terminale spé maths

Corrections détaillées

Boîte à outils

Fréquence. Si $X_n \sim B(n; p)$ et $F_n = \frac{X_n}{n}$:

$$E(F_n) = p \quad V(F_n) = \frac{p(1-p)}{n}$$

Inégalité de concentration (Bienaymé-Tchebychev appliquée à F_n) :

$$P(|F_n - p| \geq \delta) \leq \frac{p(1-p)}{n\delta^2}$$

Si p est inconnu : $p(1-p) \leq \frac{1}{4}$ pour tout $p \in [0; 1]$, d'où

$$P(|F_n - p| \geq \delta) \leq \frac{1}{4n\delta^2}$$

Taille d'échantillon : pour garantir un risque au plus α , il suffit que $n \geq \frac{1}{4\alpha\delta^2}$.

Effets des paramètres : doubler n divise le majorant par 2 ; diviser δ par 2 demande 4 fois plus d'essais.

Démarche de test : si une fréquence observée est « très improbable » sous l'hypothèse $p = p_0$ (majorant petit), on peut remettre en cause cette hypothèse.

Correction de l'exercice 1 - Espérance et variance d'une fréquence

a) $X_n \sim B(n; p)$, donc $E(X_n) = np$ et $V(X_n) = np(1-p)$. Par transformation affine ($F_n = \frac{1}{n}X_n$) :

$$E(F_n) = \frac{1}{n} \times np = p \quad V(F_n) = \frac{1}{n^2} \times np(1-p) = \frac{p(1-p)}{n}$$

b) On applique **Bienaymé-Tchebychev** à F_n , d'espérance p :

$$- P(|F_n - p| \geq \delta) \leq \frac{V(F_n)}{\delta^2} = \frac{p(1-p)}{n\delta^2}.$$

— **Interprétation** : pour δ fixé, le majorant tend vers 0 quand $n \rightarrow +\infty$: la fréquence se **concentre** autour de la probabilité.

Réponse. $E(F_n) = p$, $V(F_n) = \frac{p(1-p)}{n}$, puis Bienaymé-Tchebychev.

Correction de l'exercice 2 - Un majorant universel

a) **Forme canonique** :

$$p(1-p) = p - p^2 = \frac{1}{4} - \left(p - \frac{1}{2}\right)^2 \leq \frac{1}{4}$$

car un carré est positif. L'égalité a lieu pour $p = \frac{1}{2}$.

b) Dans l'inégalité de concentration, on remplace $p(1-p)$ par un majorant :

$$- P(|F_n - p| \geq \delta) \leq \frac{p(1-p)}{n\delta^2} \leq \frac{1}{4n\delta^2}.$$

- **Intérêt** : dans un sondage, p est justement **inconnu** ; cette majoration permet de garantir la précision **avant** de connaître le résultat.

Réponse. a) $p(1-p) = \frac{1}{4} - \left(p - \frac{1}{2}\right)^2$ · b) $\frac{1}{4n\delta^2}$.

Correction de l'exercice 3 - Mille lancers de pièce

- a) **Inégalité de concentration** avec $p = 0,5$, $n = 1\,000$, $\delta = 0,05$:

$$P(|F - 0,5| \geq 0,05) \leq \frac{0,25}{1\,000 \times 0,0025} = \frac{0,25}{2,5} = 0,1$$

- b) Par passage au contraire :

$$- P(0,45 < F < 0,55) \geq 0,9.$$

- **Lecture** : dans au moins 90 % des séries de 1 000 lancers, on obtient entre 450 et 550 « pile » (bornes exclues).

- **Réalité** : la probabilité exacte d'être hors de cet intervalle est d'environ 0,002 ; l'inégalité est prudente.

Réponse. a) $\leq 0,1$ · b) $]0,45 ; 0,55[$.

Correction de l'exercice 4 - Combien d'essais pour une précision d'un point ?

Avec le majorant universel :

$$P(|F_n - p| \geq 0,01) \leq \frac{1}{4n \times 0,01^2} = \frac{1}{0,0004n} = \frac{2\,500}{n}$$

- Il suffit que $\frac{2\,500}{n} \leq 0,05$, soit $n \geq 50\,000$.

- **Interprétation** : pour une précision d'un point de pourcentage avec 95 % de garantie, l'inégalité exige 50 000 observations.

- **Remarque** : des méthodes plus fines (utilisant la forme de la loi) montrent qu'environ 10 000 suffisent ; Bienaymé-Tchebychev reste une garantie sans hypothèse.

Réponse. $n \geq 50\,000$.

Correction de l'exercice 5 - Fréquence des six

- a) Épreuve de Bernoulli de paramètre $p = \frac{1}{6}$ répétée 600 fois :

$$E(F) = \frac{1}{6} \quad V(F) = \frac{\frac{1}{6} \times \frac{5}{6}}{600} = \frac{5}{21\,600} \approx 0,000\,231$$

- b) **Concentration** :

$$- P\left(F - \frac{1}{6} \geq 0,05\right) \leq \frac{0,000\,231}{0,05^2} \approx 0,093.$$

- **Lecture** : dans plus de 90 % des séries, la fréquence du 6 est entre 0,117 et 0,217 (soit entre 70 et 130 six).

- **Remarque** : ici $p(1-p) = \frac{5}{36} \approx 0,139$, bien moins que $\frac{1}{4}$; utiliser la vraie valeur de p donne un meilleur majorant.

Réponse. a) $E(F) = \frac{1}{6}$, $V(F) \approx 0,000\,231$ · b) $\leq 0,093$.

Correction de l'exercice 6 - Un intervalle de fluctuation

On cherche δ tel que le majorant vaille 0,05 :

$$\frac{p(1-p)}{n\delta^2} = 0,05 \Leftrightarrow \delta^2 = \frac{0,25}{400 \times 0,05} = 0,0125 \Leftrightarrow \delta \approx 0,112$$

- Alors $P(|F - 0,5| \geq 0,112) \leq 0,05$, donc $P(0,388 < F < 0,612) \geq 0,95$.
- **Intervalle** : $]0,388; 0,612[$, soit entre 156 et 244 « pile » environ.
- **Comparaison** : l'intervalle de fluctuation « classique » (loi binomiale exacte) serait plus étroit, environ $[0,45; 0,55]$.

Réponse. $]0,388; 0,612[$ environ.

Correction de l'exercice 7 - Un sondage de 2 000 personnes

a) **Majorant universel** avec $n = 2\,000$ et $\delta = 0,04$:

$$P(|F - p| \geq 0,04) \leq \frac{1}{4 \times 2\,000 \times 0,0016} = \frac{1}{12,8} \approx 0,078$$

b) Avec une garantie d'au moins 92 %, p est dans $]0,48; 0,56[$. Cet intervalle contient des valeurs inférieures à 0,5 : on **ne peut pas** conclure que le projet est majoritaire.

- **Leçon** : un sondage donnant 52 % contre 48 % est souvent « dans la marge d'erreur » ; il faudrait un échantillon plus grand ou un écart plus net.

Réponse. a) $\leq 0,078$ · b) non, 0,5 reste dans la marge.

Correction de l'exercice 8 - Taux de défaut

On utilise ici la **vraie valeur** de p (connue) :

$$P(|F - 0,02| \geq 0,01) \leq \frac{0,02 \times 0,98}{5\,000 \times 0,0001} = \frac{0,0196}{0,5} = 0,0392$$

- Au plus 3,9 % de chances d'observer moins de 1 % ou plus de 3 % de défauts.
- **Comparaison** : avec le majorant universel $\frac{1}{4}$, on trouverait $\frac{0,25}{0,5} = 0,5$, beaucoup moins utile. Quand p est petit, $p(1-p) \ll \frac{1}{4}$: il faut utiliser p s'il est connu.

Réponse. $\leq 0,0392$.

Correction de l'exercice 9 - Une pièce suspecte

a) Sous l'hypothèse $p = 0,5$, l'écart observé vaut 0,06. **Concentration** :

$$P(|F - 0,5| \geq 0,06) \leq \frac{0,25}{2\,000 \times 0,0036} = \frac{0,25}{7,2} \approx 0,035$$

b) Un tel écart aurait **moins de** 3,5 % de chances de se produire avec une pièce équilibrée. On a de bonnes raisons de **douter** que la pièce soit équilibrée.

- **Nuance** : ce n'est pas une preuve absolue ; mais l'hypothèse « pièce équilibrée » rend l'observation très improbable.
- **Démarche** : c'est le principe d'un **test statistique** — confronter une observation à ce que prévoit une hypothèse.

Réponse. a) $\leq 0,035$ · b) la pièce est probablement déséquilibrée.

Correction de l'exercice 10 - Majorant contre valeur exacte

a) Calcul :

$$P(|F - 0,5| \geq 0,1) \leq \frac{0,25}{100 \times 0,01} = 0,25$$

b) $|F - 0,5| \geq 0,1 \Leftrightarrow X \leq 40$ ou $X \geq 60$. Par **symétrie** de $B(100; 0,5)$:

— $P = 2 \times 0,0284 \approx 0,057$.

— Le majorant (0,25) est plus de **quatre fois** trop grand.

— **Leçon** : l'inégalité est une **garantie** simple et universelle ; pour une loi connue, le calcul exact est bien plus précis.

Réponse. a) $\leq 0,25$ · b) exact $\approx 0,057$.

Correction de l'exercice 11 - Précision et taille d'échantillon

a) - Doubler n **divise** b par 2.

$$b\left(n, \frac{\delta}{2}\right) = \frac{1}{4n\left(\frac{\delta}{2}\right)^2} = \frac{4}{4n\delta^2} = \frac{1}{n\delta^2} = \frac{1}{4} b(n, \delta)$$

— Diviser δ par 2 **multiplie** b par 4.

b) Pour garder le même majorant avec une marge deux fois plus petite, il faut compenser le facteur 4 : n doit être **multiplié par 4**.

— Il faut interroger 4 000 personnes.

— **Leçon** : la précision d'un sondage s'améliore comme $\frac{1}{\sqrt{n}}$; gagner un facteur 2 en précision coûte un facteur 4 en taille.

Réponse. a) $\div 2$; $\times 4$ · b) 4 000 personnes.

Correction de l'exercice 12 - Naissances de garçons

On a $p(1 - p) \approx 0,512 \times 0,488 \approx 0,25$.

a) $n = 100$:

$$P(|F - p| \geq 0,05) \leq \frac{0,25}{100 \times 0,0025} \approx 1$$

L'inégalité ne dit **rien** : dans une petite maternité, observer 45 % ou 58 % de garçons une année n'a rien d'étonnant.

b) $n = 10\,000$: le majorant vaut $\frac{0,25}{10\,000 \times 0,0025} = 0,01$.

— À l'échelle d'un pays, la fréquence est quasiment sûre d'être entre 46,2 % et 56,2 % — et en pratique bien plus près de 51,2 %.

— **Leçon** : la stabilité des fréquences n'apparaît que sur de **grands** effectifs.

Réponse. a) majorant ≈ 1 (inutile) · b) $\leq 0,01$.

Correction de l'exercice 13 - Estimer avec une marge de 2 points

a) **Majorant universel** : il suffit que

$$\frac{1}{4n \times 0,02^2} \leq 0,1 \Leftrightarrow n \geq \frac{1}{4 \times 0,0004 \times 0,1} = 6\,250$$

b) Si $p \leq 0,1$, alors $p(1 - p) \leq 0,1 \times 0,9 = 0,09$ (la fonction $p \mapsto p(1 - p)$ est croissante sur

$\left[0; \frac{1}{2}\right]$). Il suffit que

$$- \frac{0,09}{n \times 0,0004} \leq 0,1 \Leftrightarrow n \geq 2\,250.$$

— **Leçon** : une information sur p (proportion rare) permet de diviser presque par trois la taille de l'échantillon.

Réponse. a) $n \geq 6\,250$ · b) $n \geq 2\,250$.

Correction de l'exercice 14 - La roulette et le rouge

a) Calcul :

$$p(1-p) = \frac{18}{37} \times \frac{19}{37} = \frac{342}{1\,369} \approx 0,2498$$

(très proche de $\frac{1}{4}$, car p est proche de $\frac{1}{2}$).

b) **Concentration** :

$$- P(|F - p| \geq 0,02) \leq \frac{0,2498}{10\,000 \times 0,0004} \approx 0,062.$$

— Dans au moins 93,8 % des cas, la fréquence des « rouge » sur 10 000 tours est entre 0,466 et 0,506.

— **Pour le casino** : sur un grand nombre de tours, la fréquence reste proche de $\frac{18}{37} < \frac{1}{2}$, ce qui assure son avantage.

Réponse. a) $\approx 0,25$ · b) $\leq 0,062$.

Correction de l'exercice 15 - Ce que la concentration ne dit pas

a) **Non** : les lancers sont **indépendants**, la pièce n'a pas de mémoire. $P(\text{face}) = 0,5$ à chaque lancer.

b) Après 1 010 lancers, si les 1 000 derniers donnent environ 500 « pile », la fréquence vaut environ

$$\frac{8 + 500}{1\,010} \approx 0,503$$

— L'excès initial de 3 « pile » n'est pas **compensé** ; il est **dilué** dans un grand nombre de lancers.

— **Leçon** : la loi des grands nombres ne corrige pas les écarts passés, elle les rend **négligeables** ; croire à une compensation est l'« erreur du joueur ».

Réponse. a) Non (0,5) · b) l'écart est dilué, pas compensé.