

Plaquette 5 - Problèmes de synthèse :
Bienaymé-Tchebychev, échantillons et sondages
Concentration, loi des grands nombres — Terminale spé maths
 Corrections détaillées

Boîte à outils

Inégalité de Markov (pour $X \geq 0$ et $a > 0$) :

$$P(X \geq a) \leq \frac{E(X)}{a}$$

Inégalité de Bienaymé-Tchebychev. Pour toute variable aléatoire X d'espérance μ et de variance V , et tout $\delta > 0$:

$$P(|X - \mu| \geq \delta) \leq \frac{V}{\delta^2}$$

Formulation équivalente, avec $\delta = k\sigma$:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Échantillon de taille n . Si $X_1, \dots, X_n$ sont **indépendantes** et de **même loi** que X (espérance μ , variance V), alors pour la **somme** S_n et la **moyenne** $M_n = \frac{S_n}{n}$:

$$E(S_n) = n\mu \quad V(S_n) = nV \quad E(M_n) = \mu \quad V(M_n) = \frac{V}{n}$$

Inégalité de concentration. Appliquée à M_n :

$$P(|M_n - \mu| \geq \delta) \leq \frac{V}{n\delta^2}$$

Loi des grands nombres. Le majorant $\frac{V}{n\delta^2}$ tend vers 0 : pour tout $\delta > 0$,

$$\lim_{n \rightarrow +\infty} P(|M_n - \mu| \geq \delta) = 0$$

La moyenne d'un grand échantillon se **concentre** autour de l'espérance.

Cas d'une fréquence. Si F_n est la fréquence de succès dans n épreuves de paramètre p , alors $E(F_n) = p$, $V(F_n) = \frac{p(1-p)}{n}$ et

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

car $p(1-p) \leq \frac{1}{4}$ — majoration **universelle**, utilisable même quand p est inconnu.

Correction de l'exercice 1 - Inégalité de Markov

— **Inégalité de Markov** :

$$P(X \geq 10) \leq \frac{4}{10} = 0,4 \quad P(X \geq 20) \leq \frac{4}{20} = 0,2$$

— **Lecture** : au plus 40 % des valeurs atteignent 10, au plus 20 % atteignent 20.

— **Caractère grossier de la majoration** : l'inégalité n'utilise **que** l'espérance, sans rien savoir de la dispersion. Elle est donc souvent très large.

- **Cas d'égalité** : elle est atteinte quand X ne prend que les valeurs 0 et 10, avec $P(X = 10) = 0,4$ — alors $E(X) = 4$ ✓ et $P(X \geq 10) = 0,4$ exactement.
- **Exemple où elle est très pessimiste** : si X vaut toujours 4, alors $P(X \geq 10) = 0$, très loin de 0,4.
- **Condition d'emploi** : X doit être **positive** — l'inégalité est fausse sinon.

Réponse. $P(X \geq 10) \leq 0,4$ et $P(X \geq 20) \leq 0,2$.

Correction de l'exercice 2 - Bienaymé-Tchebychev de base

- **Majoration** :

$$P(|X - 50| \geq 10) \leq \frac{25}{100} = 0,25$$

- **Minoration de l'événement contraire** :

$$P(40 < X < 60) = 1 - P(|X - 50| \geq 10) \geq 1 - 0,25 = 0,75$$

- **Lecture** : au moins 75 % des valeurs sont strictement entre 40 et 60.
- **En unités d'écart type** : $\sigma = 5$, donc $10 = 2\sigma$ et l'on retrouve la forme $\frac{1}{k^2} = \frac{1}{4}$ ✓
- **Comparaison utile** : pour une loi **normale**, la vraie probabilité serait environ 0,954 — Tchebychev garantit seulement 0,75, mais sans **aucune** hypothèse sur la loi. C'est tout son intérêt.
- À 3σ : $P(|X - 50| \geq 15) \leq \frac{25}{225} \approx 0,111$, donc au moins 88,9 % des valeurs dans $]35 ; 65[$.

Réponse. $\leq 0,25$; $P(40 < X < 60) \geq 0,75$.

Correction de l'exercice 3 - Application à une binomiale

- **Paramètres** : $E(X) = 50$ et $V(X) = 100 \times 0,5 \times 0,5 = 25$.

- **Majoration** :

$$P(|X - 50| \geq 15) \leq \frac{25}{225} = \frac{1}{9} \approx 0,1111$$

- **Comparaison** : la valeur exacte est 0,0035, soit **32 fois plus petite** que le majorant.
- **Conclusion** : l'inégalité est **valide** mais très **pessimiste** — elle donne une garantie universelle, pas une bonne approximation.
- **Pourquoi ?** Tchebychev ne connaît que μ et V ; elle doit couvrir le pire des cas parmi **toutes** les lois de mêmes paramètres, y compris des lois très étalées. La binomiale, elle, est fortement concentrée.
- **Utilité réelle** : l'inégalité sert à **démontrer** la loi des grands nombres et à obtenir des tailles d'échantillon garanties sans hypothèse de loi, pas à calculer des probabilités précises.
- **Contrôle de l'ordre de grandeur** : $15 = 3\sigma$, et $\frac{1}{3^2} = \frac{1}{9}$ ✓

Réponse. Majorant $\frac{1}{9} \approx 0,111$, contre 0,0035 en réalité : majoration valide mais très large.

Correction de l'exercice 4 - Moyenne d'un échantillon

- **Loi de la moyenne** : $E(M_n) = 20$ et $V(M_n) = \frac{16}{400} = 0,04$.
- **Inégalité de concentration** :

$$P(|M_n - 20| \geq 0,5) \leq \frac{0,04}{0,25} = 0,16$$

- **Lecture** : au moins 84 % des échantillons de taille 400 donnent une moyenne entre 19,5 et 20,5.
- **Écart type de la moyenne** : $\sigma(M_n) = \sqrt{0,04} = 0,2$, alors que $\sigma(X) = 4$ — la moyenne est **20 fois** moins dispersée que la variable individuelle, car $\sqrt{400} = 20$ ✓
- **Effet de la taille** : en passant à $n = 1600$, $V(M_n) = 0,01$ et le majorant tombe à 0,04 — quadrupler n divise le majorant par 4.
- **Formule directe** : $\frac{V}{n\delta^2} = \frac{16}{400 \times 0,25} = \frac{16}{100} = 0,16$ ✓

Réponse. $\leq 0,16$; au moins 84 % des moyennes dans $]19,5 ; 20,5[$.

Correction de l'exercice 5 - Taille d'échantillon garantie

- **Condition** :

$$\frac{9}{n \times 0,09} \leq 0,05$$

- **Simplification** : $\frac{100}{n} \leq 0,05$, donc $n \geq \frac{100}{0,05} = 2000$.
- **Conclusion** : $n \geq 2000$.
- **Vérification** : pour $n = 2000$, le majorant vaut $\frac{9}{2000 \times 0,09} = \frac{9}{180} = 0,05$ ✓ — la borne est **atteinte** exactement.
- **Contrôle du sens** : pour $n = 1999$, le majorant vaut $\approx 0,05003 > 0,05$ — insuffisant ✓
- **Sensibilité à la précision** : diviser δ par 2 (viser 0,15) exigerait $n \geq 8000$, soit **quatre fois** plus. La taille croît en $\frac{1}{\delta^2}$: gagner un chiffre de précision coûte cent fois plus d'observations.
- **Formule générale** : $n \geq \frac{V}{\alpha\delta^2}$, où α est le risque accepté.

Réponse. $n \geq 2000$.

Correction de l'exercice 6 - Fréquence dans un sondage

- **Variance de la fréquence** : $V(F_n) = \frac{p(1-p)}{n}$, avec p **inconnu**.
- **Majoration universelle** : $p(1-p) \leq \frac{1}{4}$ pour tout $p \in [0; 1]$ (maximum en $p = 0,5$).
- **Application** :

$$P(|F_n - p| \geq 0,05) \leq \frac{1}{4 \times 1000 \times 0,0025} = \frac{1}{10} = 0,1$$

- **Lecture** : au moins 90 % des sondages de 1000 personnes donnent une fréquence à moins de 5 points de la vraie proportion.
- **Justification de la majoration** : $p(1-p) = \frac{1}{4} - \left(p - \frac{1}{2}\right)^2 \leq \frac{1}{4}$ ✓ — c'est un simple développement de la forme canonique.
- **Gain si p est connu** : avec $p = 0,2$, on a $p(1-p) = 0,16$ et le majorant tombe à $\frac{0,16}{2,5} = 0,064$ — nettement mieux.
- **Intérêt pratique** : cette majoration permet de dimensionner un sondage **avant** de connaître p , ce qui est précisément la situation réelle.

Réponse. $\leq 0,1$ (majoration $p(1-p) \leq \frac{1}{4}$).

Correction de l'exercice 7 - Dimensionner un sondage

— **Condition :**

$$\frac{1}{4n \times 0,0004} \leq 0,05$$

— **Calcul :** $4 \times 0,0004 = 0,0016$, donc $\frac{1}{0,0016n} \leq 0,05$, soit

$$n \geq \frac{1}{0,0016 \times 0,05} = \frac{1}{0,00008} = 12\,500$$

— **Conclusion :** il faut interroger **au moins 12 500 personnes.**

— **Vérification :** $\frac{1}{4 \times 12\,500 \times 0,0004} = \frac{1}{20} = 0,05 \checkmark$

— **Formule mémorisable :** $n \geq \frac{1}{4\alpha\delta^2}$, avec ici $\alpha = 0,05$ et $\delta = 0,02$.

— **Commentaire réaliste :** les instituts de sondage interrogent environ 1000 personnes et annoncent une marge de ± 3 points. Tchebychev, sans hypothèse de loi, exigerait $n \geq \frac{1}{4 \times 0,05 \times 0,0009} \approx 5556$ pour la même marge. L'écart vient de ce que les instituts utilisent l'**approximation normale**, bien plus fine mais moins universelle.

— **Coût de la précision :** viser $\delta = 0,01$ demanderait 50 000 personnes — quatre fois plus pour deux fois plus de précision.

Réponse. $n \geq 12\,500$ personnes.

Correction de l'exercice 8 - Loi des grands nombres en acte

— **Paramètres d'un dé :** $\mu = 3,5$ et $V = \frac{35}{12} \approx 2,9167$.

— **Pour** $n = 1000$:

$$P \leq \frac{35/12}{1000 \times 0,01} = \frac{2,9167}{10} \approx 0,2917$$

— **Pour** $n = 10\,000$:

$$P \leq \frac{2,9167}{100} \approx 0,02917$$

— **Lecture :** en multipliant n par 10, le majorant est divisé par 10 — il tend vers 0, ce qui **est** la loi des grands nombres.

— **Contrôle des écarts types :** $\sigma(M_{1000}) = \sqrt{\frac{2,9167}{1000}} \approx 0,054$ et $\sigma(M_{10\,000}) \approx 0,017$.

— **Cohérence :** avec $n = 10\,000$, l'écart demandé 0,1 vaut près de 6σ — d'où un majorant de l'ordre de $\frac{1}{36} \approx 0,028 \checkmark$

— **Taille pour un majorant de 0,01 :** $n \geq \frac{2,9167}{0,01 \times 0,01} \approx 29\,167$ lancers.

Réponse. $\approx 0,292$ pour $n = 1000$; $\approx 0,029$ pour $n = 10\,000$.

Correction de l'exercice 9 - Somme et non moyenne

— **Paramètres de la somme :** $E(S_n) = 900 \times 3,5 = 3150$ et $V(S_n) = 900 \times \frac{35}{12} = 2625$.

— **Majoration :**

$$P(|S_n - 3150| \geq 100) \leq \frac{2625}{10\,000} = 0,2625$$

- **Lecture** : au moins 73,75 % des séries de 900 lancers donnent une somme entre 3050 et 3250.
- **Écart type** : $\sigma(S_n) = \sqrt{2625} \approx 51,23$, donc $100 \approx 1,95\sigma$ — et $\frac{1}{1,95^2} \approx 0,263$ ✓
- **Attention à ne pas confondre** : pour la **somme**, la variance est **multipliée** par n ; pour la **moyenne**, elle est **divisée** par n .
- **Lien entre les deux** : l'événement $|S_n - 3150| \geq 100$ équivaut à $|M_n - 3,5| \geq \frac{100}{900} \approx 0,111$, et $\frac{35/12}{900 \times 0,1111^2} = \frac{2,9167}{11,111} \approx 0,2625$ ✓ — les deux approches donnent bien le même majorant.
- **Conséquence** : la somme **s'écarte** de plus en plus en valeur absolue (en $\sqrt{n}$), alors que la moyenne se **resserre** (en $\frac{1}{\sqrt{n}}$).

Réponse. $\leq 0,2625$.

Correction de l'exercice 10 - Contrôle de production

- **Paramètres** : $\mu = 100$, $\sigma = 0,4$, donc $V = 0,16$.
- **Majoration** :

$$P(|X - 100| \geq 1) \leq \frac{0,16}{1} = 0,16$$

- **Lecture** : au plus 16 % de rejets garantis, donc au moins 84 % de pièces conformes.
- **En unités d'écart type** : $1 = 2,5\sigma$, et $\frac{1}{2,5^2} = \frac{1}{6,25} = 0,16$ ✓
- **Comparaison loi normale** : si la longueur suivait une loi normale, la proportion de rejets serait d'environ 1,2 % — Tchebychev, sans hypothèse, ne garantit que 16 %.
- **Amélioration du procédé** : pour garantir au plus 4 % de rejets, il faut $\frac{\sigma^2}{1} \leq 0,04$, soit $\sigma \leq 0,2$ mm — il faut **diviser l'écart type par 2**.
- **Usage industriel** : cette majoration donne une garantie **minimale** exploitable même quand la loi de fabrication est mal connue, ce qui est souvent le cas en début de série.

Réponse. Au plus 16 % de rejets (au moins 84 % de conformes).

Correction de l'exercice 11 - Minoration d'un intervalle

- **Symétrie de l'intervalle** : $]6; 18[$ est centré en 12, de demi-amplitude $\delta = 6$.
- **Majoration de l'événement contraire** :

$$P(|X - 12| \geq 6) \leq \frac{4}{36} = \frac{1}{9} \approx 0,1111$$

- **Minoration** :

$$P(6 < X < 18) \geq 1 - \frac{1}{9} = \frac{8}{9} \approx 0,8889$$

- **Lecture** : au moins 88,9 % des valeurs sont dans $]6; 18[$.
- **En unités d'écart type** : $6 = 3\sigma$, d'où $1 - \frac{1}{3^2} = \frac{8}{9}$ ✓
- **Règle générale** : à $k\sigma$, Tchebychev garantit au moins $1 - \frac{1}{k^2}$ — soit 0 pour $k = 1$ (aucune information), 0,75 pour $k = 2$, $\frac{8}{9}$ pour $k = 3$, 0,96 pour $k = 5$.

- **Point important** : pour $k \leq 1$ l'inégalité ne dit **rien** d'utile, le majorant dépassant 1. Elle n'est informative qu'au-delà d'un écart type.

Réponse. $P(6 < X < 18) \geq \frac{8}{9} \approx 0,889$.

Correction de l'exercice 12 - Fréquence avec p connu

- **Variance exacte** : $V(F_n) = \frac{0,6 \times 0,4}{500} = \frac{0,24}{500} = 0,00048$.

- **Majoration avec p connu** :

$$P \leq \frac{0,00048}{0,0016} = 0,3$$

- **Majoration universelle** ($p(1-p) \leq 0,25$) :

$$P \leq \frac{0,25}{500 \times 0,0016} = \frac{0,25}{0,8} = 0,3125$$

- **Comparaison** : 0,3 contre 0,3125 — connaître p améliore peu le résultat ici, car 0,6 est proche de 0,5.
- **Cas où l'écart est net** : avec $p = 0,05$, on aurait $p(1-p) = 0,0475$ et un majorant de $\frac{0,0475}{0,8} \approx 0,0594$, soit **cinq fois** mieux que 0,3125.
- **Conclusion pratique** : la majoration universelle est presque optimale pour p proche de $\frac{1}{2}$, et très pessimiste pour p extrême.
- **Écart type** : $\sigma(F_n) = \sqrt{0,00048} \approx 0,0219$, donc $0,04 \approx 1,83\sigma$ — un majorant de l'ordre de $\frac{1}{1,83^2} \approx 0,30$ ✓

Réponse. 0,3 avec p connu, contre 0,3125 en universel.

Correction de l'exercice 13 - Nombre de lancers

- **Condition** :

$$\frac{35/12}{n \times 0,0025} \leq 0,1$$

- **Calcul** : $\frac{2,91667}{0,0025n} \leq 0,1$, soit

$$n \geq \frac{2,91667}{0,00025} = 11\,666,67$$

- **Conclusion** : $n \geq 11\,667$ lancers.
- **Vérification** : pour $n = 11\,667$, le majorant vaut $\frac{2,91667}{11\,667 \times 0,0025} = \frac{2,91667}{29,1675} \approx 0,09999 \leq 0,1$ ✓
- **Pour** $n = 11\,666$: $\approx 0,100006 > 0,1$ — insuffisant ✓
- **Forme exacte** : $n \geq \frac{35}{12 \times 0,1 \times 0,0025} = \frac{35}{0,003} = \frac{35\,000}{3} \approx 11\,666,67$.
- **Commentaire** : plus de 11 000 lancers pour garantir 0,05 de précision — la convergence de la loi des grands nombres est **certaine** mais **lente**, ce qui justifie l'usage de méthodes plus fines (approximation normale) dans la pratique.

Réponse. $n \geq 11\,667$ lancers.

Correction de l'exercice 14 - Comparaison de deux échantillons

- **Majoration universelle** : $\frac{1}{4n\delta^2}$ avec $\delta = 0,03$, donc $\delta^2 = 0,0009$.
- **Pour** $n = 400$:

$$\frac{1}{4 \times 400 \times 0,0009} = \frac{1}{1,44} \approx 0,6944$$
- **Pour** $n = 2500$:

$$\frac{1}{4 \times 2500 \times 0,0009} = \frac{1}{9} \approx 0,1111$$
- **Comparaison** : le majorant est divisé par 6,25, exactement le rapport $\frac{2500}{400} \checkmark$
- **Lecture critique** : avec 400 personnes, le majorant 0,694 ne garantit presque **rien** — il faut se méfier des sondages de faible effectif annonçant ± 3 points.
- **Taille nécessaire pour un majorant de 0,05** : $n \geq \frac{1}{4 \times 0,05 \times 0,0009} \approx 5556$ personnes.
- **Marge atteignable avec** $n = 400$ pour un risque de 0,05 :

$$\delta \geq \sqrt{\frac{1}{4 \times 400 \times 0,05}} = \sqrt{0,0125} \approx 0,1118, \text{ soit } \pm 11 \text{ points — bien loin des 3 points annoncés.}$$

Réponse. $\approx 0,694$ pour $n = 400$ contre $\approx 0,111$ pour $n = 2500$.

Correction de l'exercice 15 - Estimation d'une variance

- **Inégalité** : $0,05 \leq \frac{V}{16}$, donc $V \geq 0,8$.
- **Attention au sens** : Tchebychev fournit une **minoration** de V ici, car la probabilité observée est **inférieure ou égale** au majorant.
- **Interprétation** : puisque 5 % des valeurs s'écartent d'au moins 4, la variance ne peut pas être trop petite — il faut $V \geq 0,8$.
- **Vérification par un exemple** : si X vaut 4 ou 12 avec la probabilité 0,025 chacune, et 8 le reste du temps, alors $V = 16 \times 0,05 = 0,8 \checkmark$ — la borne est **atteinte**.
- **Ce que l'on ne peut pas conclure** : aucune **majoration** de V . La variable pourrait être très dispersée à l'intérieur de $]4; 12[$ sans changer la probabilité donnée.
- **Contrôle de cohérence** : avec $V = 0,8$, l'inégalité donne $P \leq \frac{0,8}{16} = 0,05$, exactement la valeur observée $\checkmark$ — c'est le cas d'égalité, atteint par la loi à trois valeurs ci-dessus.

Réponse. $V \geq 0,8$ (borne atteinte par une loi à trois valeurs).

Correction de l'exercice 16 - Synthèse : sondage électoral

a) **Majoration universelle** :

$$P(|F_n - p| \geq 0,025) \leq \frac{1}{4 \times 1600 \times 0,000625} = \frac{1}{4} = 0,25$$

— Au moins 75 % des sondages sont à moins de 2,5 points de la vraie proportion.

b) **Condition** : $\frac{1}{4n \times 0,0001} \leq 0,01$, soit

$$n \geq \frac{1}{4 \times 0,0001 \times 0,01} = \frac{1}{0,000004} = 250\,000$$

— **Vérification** : $\frac{1}{4 \times 250\,000 \times 0,0001} = \frac{1}{100} = 0,01 \checkmark$

— **Commentaire** : 250 000 personnes — chiffre irréaliste, qui montre la **prudence extrême** de Tchebychev. L'approximation normale ramènerait ce nombre à environ 16 600.

c) **Analyse critique.** Il faudrait pouvoir exclure $p \leq 0,5$, donc un écart d'au moins 0,02.

— Le majorant vaut $\frac{1}{4 \times 1600 \times 0,0004} = \frac{1}{2,56} \approx 0,3906$.

— **Conclusion** : la probabilité de se tromper de plus de 2 points peut atteindre 39 % — on **ne peut pas** affirmer que le candidat est majoritaire.

— **Précision garantie à 5 % de risque** : $\delta \geq \sqrt{\frac{1}{4 \times 1600 \times 0,05}} = \sqrt{0,003125} \approx 0,0559$, soit $\pm 5,6$ points. L'intervalle $[0,464 ; 0,576]$ contient 0,5.

— **Leçon** : un écart de 52 % contre 48 % sur 1600 personnes n'est **pas** concluant. C'est exactement pourquoi les instituts parlent de résultats « dans la marge d'erreur ».

Réponse. a) $\leq 0,25$ · b) $n \geq 250\,000$ · c) non : marge garantie de $\pm 5,6$ points.