

Plaquette 3 – Estimer une proportion inconnue

Échantillonnage et fluctuation — Seconde

Corrections détaillées

Boîte à outils

Deux problèmes symétriques.

Problème	On connaît	On cherche
tester	p annoncée et f observée	f est-elle compatible avec p ?
estimer	f observée seule	une fourchette pour p

Fourchette d'estimation (intervalle de confiance au niveau de 95 %). À partir d'une fréquence f observée sur un échantillon de taille n :

$$J = \left[f - \frac{1}{\sqrt{n}} ; f + \frac{1}{\sqrt{n}} \right]$$

La **marge d'erreur** est $\frac{1}{\sqrt{n}}$, l'**amplitude** $\frac{2}{\sqrt{n}}$.

Ce que signifie « au niveau de 95 % ». Si l'on répétait le sondage un grand nombre de fois, environ 95 % des fourchettes obtenues contiendraient la vraie valeur de p . Pour **un** sondage donné, on ne sait pas si l'on est dans les 95 % ou dans les 5 %.

Trois formulations à ne pas confondre.

Formulation	Statut
« p vaut f »	faux — f n'est qu'une estimation
« p appartient à J »	acceptable, au niveau de confiance 95 %
« la probabilité que $p \in J$ vaut 0,95 »	abus de langage : p est fixé, ce n'est pas lui qui varie

Ce qui ne dépend pas de la taille de la population. La marge d'erreur dépend de n , la **taille de l'échantillon**, et non du nombre total d'habitants. Un sondage de 1000 personnes a la même précision dans une ville de 50 000 habitants et dans un pays de 60 millions.

Correction de l'exercice 1 – Première fourchette d'estimation

— **Fréquence observée** : $f = \frac{180}{400} = 0,45$ ✓

— **Marge d'erreur** : $\frac{1}{\sqrt{400}} = \frac{1}{20} = 0,05$ ✓

— **Fourchette** :

$$J = [0,45 - 0,05 ; 0,45 + 0,05] = [0,40 ; 0,50]$$

— **Formulation** : la proportion de lecteurs est comprise entre 40 % et 50 %, au niveau de confiance 95 % ✓

— **Amplitude** : 0,10, soit 10 points ✓

— **Marge en points** : ± 5 points ✓

— **Ce qu'on ne peut pas dire** : « exactement 45 % des gens lisent un quotidien » ✓

— **Raison** : 0,45 est une **estimation**, entachée d'une incertitude de 5 points ✓

- **Effectifs correspondants aux bornes** : 160 et 200 lecteurs sur 400 ✓
- **Contrôle** : $0,40 \times 400 = 160$ ✓ et $0,50 \times 400 = 200$ ✓
- **Peut-on affirmer que moins de la moitié lisent ?** Non : 0,50 est la borne haute de la fourchette ✓
- **Pour trancher** : il faudrait un échantillon plus grand ✓
- **Avec $n = 2500$ et la même fréquence** : $J = [0,43 ; 0,47]$, et l'on pourrait conclure ✓
- **Contrôle** : $\frac{1}{\sqrt{2500}} = 0,02$ ✓

Réponse. $J = [0,40 ; 0,50]$: entre 40 et 50 % de lecteurs, au niveau de confiance 95 %.

Correction de l'exercice 2 - Sondage électoral

- **Fréquence observée** : $f = 0,52$ ✓
- **Marge d'erreur** : $\frac{1}{\sqrt{1000}} \approx 0,0316$ ✓
- **Fourchette** :

$$J \approx [0,52 - 0,0316 ; 0,52 + 0,0316] = [0,4884 ; 0,5516]$$
- **Test décisif** : la fourchette contient 0,50 ✓
- **Conclusion** : on **ne peut pas** affirmer l'élection au premier tour ✓
- **Formulation correcte** : « le candidat est en tête, mais l'écart n'est pas significatif » ✓
- **Marge en points** : $\pm 3,2$ points ✓
- **Écart de la fréquence au seuil** : $0,52 - 0,50 = 2$ points ✓
- **Comparaison** : $2 < 3,2$, donc l'avance est **dans la marge d'erreur** ✓
- **Taille nécessaire pour conclure** : il faut $\frac{1}{\sqrt{n}} < 0,02$, soit $n > 2500$ ✓
- **Avec $n = 2500$ et $f = 0,52$** : $J = [0,50 ; 0,54]$, tout juste à la limite ✓
- **Avec $n = 4900$** : marge $\frac{1}{70} \approx 0,0143$, donc $J \approx [0,5057 ; 0,5343]$ ✓
- **Conclusion pour $n = 4900$** : la majorité serait alors significative ✓
- **Pratique journalistique** : c'est pourquoi les instituts parlent de « scrutin trop serré » sous 2 ou 3 points d'écart ✓

Réponse. Non : $J \approx [0,488 ; 0,552]$ contient 0,50, l'avance de 2 points est dans la marge d'erreur.

Correction de l'exercice 3 - Deux sondages, même fréquence

- **Premier institut** : $\frac{1}{\sqrt{100}} = 0,1$ ✓
- **Fourchette** : $[0,26 ; 0,46]$ ✓
- **Second institut** : $\frac{1}{\sqrt{900}} = \frac{1}{30} \approx 0,0333$ ✓
- **Fourchette** : $\approx [0,3267 ; 0,3933]$ ✓
- **Amplitudes** : 0,20 contre $\approx 0,0667$ ✓
- **Rapport** : $\frac{0,20}{0,0667} = 3$ ✓
- **Cohérence** : $\sqrt{\frac{900}{100}} = 3$ ✓
- **Conclusion** : le second sondage est **trois fois plus précis** ✓
- **Inclusion** : la petite fourchette est entièrement contenue dans la grande ✓

- **Coût de la précision** : 9 fois plus de personnes interrogées pour 3 fois plus de précision ✓
- **Lecture du premier sondage** : « entre 26 et 46 % » est une information très faible ✓
- **Lecture du second** : « entre 33 et 39 % » est exploitable ✓
- **Ce que les deux ont en commun** : la même **estimation ponctuelle** de 36 % ✓
- **Ce qui les sépare** : la seule taille de l'échantillon ✓
- **Règle à retenir** : une fréquence publiée sans sa taille d'échantillon est inutilisable ✓

Réponse. $[0,26; 0,46]$ contre $\approx [0,327; 0,393]$: le second est trois fois plus précis.

Correction de l'exercice 4 - Retrouver la taille de l'échantillon

- **Équation** : $\frac{1}{\sqrt{n}} = 0,025$ ✓
- **Passage aux inverses** : $\sqrt{n} = \frac{1}{0,025} = 40$ ✓
- **Élévation au carré** :

$$n = 1600$$
- **Vérification** : $\frac{1}{\sqrt{1600}} = \frac{1}{40} = 0,025$ ✓
- **Fourchette publiée** : $[0,395; 0,445]$ ✓
- **Contrôle** : $0,42 - 0,025 = 0,395$ ✓ et $0,42 + 0,025 = 0,445$ ✓
- **Effectif de réponses positives** : $0,42 \times 1600 = 672$ personnes ✓
- **Contrôles pour d'autres marges** : ± 1 point donne $n = 10\,000$ ✓
- **Détail** : $\frac{1}{0,01} = 100$ et $100^2 = 10\,000$ ✓
- **Pour ± 2 points** : $n = 2500$ ✓
- **Pour ± 3 points** : $\frac{1}{0,03} \approx 33,33$, donc $n \approx 1111$ ✓
- **Pour ± 4 points** : $n = 625$ ✓
- **Lecture** : les tailles usuelles des sondages (1000 à 2000) correspondent à des marges de 2 à 3 points ✓
- **Attention** : les instituts utilisent parfois une formule plus fine tenant compte de f , ce qui donne des marges légèrement différentes ✓
- **Notre formule** : elle majore la marge, elle est donc **prudente** ✓

Réponse. $n = 1600$ personnes.

Correction de l'exercice 5 - La taille de la population n'intervient pas

- **Marge d'erreur dans les deux cas** : $\frac{1}{\sqrt{1000}} \approx 0,0316$ ✓
- **Conclusion immédiate** : la marge est **identique** ✓
- **Réponse** : l'affirmation est **fausse** ✓
- **Ce qui compte** : la taille n de l'**échantillon**, pas celle de la population ✓
- **Formule** : la marge s'écrit $\frac{1}{\sqrt{n}}$, où N (taille de la population) n'apparaît pas ✓
- **Image intuitive** : pour goûter une soupe, une cuillerée suffit, que la marmite soit petite ou grande ✓
- **Condition indispensable** : la soupe doit être **bien mélangée** — c'est le caractère **aléatoire** du tirage ✓
- **Vraie source d'erreur d'un sondage** : un échantillon **mal tiré**, pas un échantillon trop petit face à la population ✓

- **Exemple de biais** : n'interroger que des personnes possédant un téléphone fixe ✓
- **Autre exemple** : n'interroger que des volontaires, qui ont souvent un avis plus tranché ✓
- **Taux de sondage pour la ville** : $\frac{1000}{20\,000} = 5\%$ ✓
- **Taux pour le pays** : $\frac{1000}{60\,000\,000} \approx 0,0017\%$ ✓
- **Malgré cet écart énorme** : la précision est la même ✓
- **Nuance technique** : si l'échantillon représente une part importante de la population (plus de 10 %), la précision est en réalité **meilleure** que prévue ✓
- **Conclusion** : la taille de la population ne dégrade **jamais** la précision ✓

Réponse. L'affirmation est fausse : la marge $\frac{1}{\sqrt{n}} \approx 0,032$ ne dépend que de n , pas de la population.

Correction de l'exercice 6 - Comparer deux candidats

- **Marge d'erreur** : $\frac{1}{\sqrt{1600}} = \frac{1}{40} = 0,025$ ✓
- **Fourchette pour A** : $[0,435; 0,485]$ ✓
- **Fourchette pour B** : $[0,415; 0,465]$ ✓
- **Recouvrement** : l'intervalle $[0,435; 0,465]$ est commun aux deux ✓
- **Conclusion** : l'écart **n'est pas significatif** ✓
- **Écart observé** : $46 - 44 = 2$ points ✓
- **Comparaison** : $2 < 2 \times 2,5 = 5$ points ✓
- **Formulation correcte** : « les deux candidats sont au coude à coude » ✓
- **Formulation trompeuse** : « A devance B » ✓
- **Écart minimal pour conclure** : plus de 5 points, avec ces tailles ✓
- **Taille nécessaire pour rendre 2 points significatifs** : il faut $2 \times \frac{1}{\sqrt{n}} < 0,02$, soit $\frac{1}{\sqrt{n}} < 0,01$ ✓
- **Résolution** : $n > 10\,000$ personnes ✓
- **Effectifs observés** : 736 pour A et 704 pour B ✓
- **Contrôles** : $0,46 \times 1600 = 736$ ✓ et $0,44 \times 1600 = 704$ ✓
- **Écart en personnes** : 32 sur 1600, soit très peu ✓
- **Leçon** : dans la presse, un écart inférieur à la marge doit se lire comme une **égalité statistique** ✓

Réponse. Non : $[0,435; 0,485]$ et $[0,415; 0,465]$ se recouvrent — il faudrait plus de 5 points d'écart.

Correction de l'exercice 7 - Contrôle d'un lot de pièces

- **Fréquence observée** : $f = \frac{50}{625} = 0,08$ ✓
- **Marge d'erreur** : $\frac{1}{\sqrt{625}} = \frac{1}{25} = 0,04$ ✓
- **Fourchette** :

$$J = [0,08 - 0,04; 0,08 + 0,04] = [0,04; 0,12]$$
- **Objectif** : taux de défaut $< 0,10$ ✓
- **Test** : la fourchette contient 0,10 ✓
- **Conclusion** : on **ne peut pas affirmer** que l'objectif est atteint ✓

- **Nuance** : l'estimation ponctuelle (8 %) est favorable, mais l'incertitude est trop grande ✓
- **Écart au seuil** : $0,10 - 0,08 = 2$ points ✓
- **Marge** : 4 points, soit le double ✓
- **Taille nécessaire pour conclure** : il faut $\frac{1}{\sqrt{n}} < 0,02$, soit $n > 2500$ ✓
- **Avec $n = 2500$ et $f = 0,08$** : $J = [0,06 ; 0,10]$, encore à la limite ✓
- **Avec $n = 10\,000$ et $f = 0,08$** : $J = [0,07 ; 0,09]$, l'objectif serait alors **démontré** ✓
- **Nombre de pièces défectueuses attendu** : $0,08 \times 10\,000 = 800$ ✓
- **Réserve sur la validité** : $f = 0,08 < 0,2$, la formule sort de son domaine usuel ✓
- **Conséquence** : la marge réelle est plus petite, l'estimation obtenue est donc **prudente** ✓

Réponse. $J = [0,04 ; 0,12]$: on ne peut pas conclure, l'objectif de 10 % est dans la fourchette.

Correction de l'exercice 8 - Lire une fourchette à l'envers

- **Fréquence observée** : c'est le **centre** de la fourchette ✓
- **Calcul** :

$$f = \frac{0,61 + 0,67}{2} = 0,64$$

- **Marge d'erreur** : la **demi-amplitude** ✓

- **Calcul** : $\frac{0,67 - 0,61}{2} = 0,03$ ✓

- **Équation en n** : $\frac{1}{\sqrt{n}} = 0,03$ ✓

- **Résolution** : $\sqrt{n} = \frac{1}{0,03} \approx 33,33$ ✓

- **Donc** :

$$n \approx 33,33^2 \approx 1111$$

- **Vérification** : $\frac{1}{\sqrt{1111}} \approx 0,0300$ ✓

- **Arrondi plausible** : l'institut a probablement interrogé 1111 ou 1100 personnes ✓

- **Contrôle avec $n = 1100$** : $\frac{1}{\sqrt{1100}} \approx 0,0302$, soit 3,0 points ✓

- **Effectif de satisfaits** : $0,64 \times 1111 \approx 711$ personnes ✓

- **Contrôle de cohérence** : les bornes sont bien symétriques autour de 0,64 ✓

- **Détails** : $0,64 - 0,03 = 0,61$ ✓ et $0,64 + 0,03 = 0,67$ ✓

- **Peut-on affirmer qu'une majorité est satisfaite ?** Oui : la borne basse 0,61 dépasse 0,50 ✓

- **Marge de sécurité** : 11 points au-dessus du seuil de majorité ✓

Réponse. $f = 0,64$ et $n \approx 1111$ personnes.

Correction de l'exercice 9 - Le niveau de confiance

- **Niveau de confiance** : 95 % ✓

- **Fourchettes correctes attendues** :

$$0,95 \times 200 = 190$$

- **Fourchettes manquant la vraie valeur** : $0,05 \times 200 = 10$ ✓

- **Contrôle** : $190 + 10 = 200$ ✓
- **Nature de ces nombres** : ce sont des **espérances**, pas des certitudes ✓
- **Fluctuation prévisible** : on pourrait observer 8 ou 13 échecs ✓
- **Interprétation du niveau de confiance** : c'est la proportion de fourchettes **correctes à long terme** ✓
- **Ce qu'on ignore** : pour un sondage particulier, on ne sait pas s'il fait partie des 190 ou des 10 ✓
- **Amplitude de chaque fourchette** : $\frac{2}{\sqrt{400}} = 0,1$ ✓
- **Conséquence d'un niveau de 99 %** : il faudrait des fourchettes **plus larges** ✓
- **Compromis fondamental** : plus on veut être sûr, moins on est précis ✓
- **Cas extrême** : la fourchette $[0; 1]$ est correcte à 100 % et totalement inutile ✓
- **Autre extrême** : la fourchette réduite au point f est presque toujours fausse ✓
- **Choix du 95 %** : c'est une convention, jugée un bon compromis ✓
- **Contrôle avec 1000 sondages** : on attendrait 950 fourchettes correctes et 50 fausses ✓

Réponse. Environ 190 fourchettes contiendront la vraie proportion, et 10 la manqueront.

Correction de l'exercice 10 - Estimer un effectif de population

- **Fréquence observée** : $f = 0,33$ ✓
- **Marge d'erreur** : $\frac{1}{\sqrt{900}} = \frac{1}{30} \approx 0,0333$ ✓
- **Fourchette pour la proportion** :

$$J \approx [0,2967; 0,3633]$$
- **Passage aux effectifs** : on multiplie les bornes par 48 000 ✓
- **Borne basse** : $0,2967 \times 48\,000 \approx 14\,240$ ✓
- **Borne haute** : $0,3633 \times 48\,000 \approx 17\,440$ ✓
- **Fourchette en effectifs** : environ $[14\,240; 17\,440]$ utilisateurs ✓
- **Estimation ponctuelle** : $0,33 \times 48\,000 = 15\,840$ ✓
- **Contrôle** : 15 840 est bien le centre de la fourchette ✓
- **Amplitude en effectifs** : $17\,440 - 14\,240 = 3\,200$ ✓
- **Contrôle** : $\frac{2}{30} \times 48\,000 = 3\,200$ ✓
- **Lecture pour la mairie** : entre 14 000 et 17 500 usagers, soit une incertitude de 3 200 personnes ✓
- **Conséquence budgétaire** : si le coût annuel par usager est de 120 €, le budget va de 1,71 à 2,09 million d'euros ✓
- **Contrôles** : $14\,240 \times 120 \approx 1\,708\,800$ € et $17\,440 \times 120 \approx 2\,092\,800$ € ✓
- **Recommandation** : pour resserrer, il faudrait interroger davantage — 3 600 personnes diviseraient la marge par 2 ✓

Réponse. Entre 14 240 et 17 440 utilisateurs environ (estimation ponctuelle 15 840).

Correction de l'exercice 11 - Deux échantillons cumulés

- **Effectif positif du premier** : $0,55 \times 400 = 220$ ✓
- **Effectif positif du second** : $0,59 \times 500 = 295$ ✓
- **Effectif cumulé** : $220 + 295 = 515$ ✓
- **Taille cumulée** : $400 + 500 = 900$ ✓

— **Fréquence globale :**

$$f = \frac{515}{900} \approx 0,5722$$

- **Piège :** la moyenne simple $\frac{0,55 + 0,59}{2} = 0,57$ n'est pas la bonne valeur ✓
- **Écart :** 0,5722 contre 0,57 ✓
- **Explication :** le second échantillon, plus grand, pèse davantage ✓
- **Marge du premier seul :** $\frac{1}{\sqrt{400}} = 0,05$ ✓
- **Marge du second seul :** $\frac{1}{\sqrt{500}} \approx 0,0447$ ✓
- **Marge cumulée :** $\frac{1}{\sqrt{900}} \approx 0,0333$ ✓
- **Gain de précision :** la marge passe de 0,05 à 0,0333, soit un tiers de moins ✓
- **Fourchette cumulée :** $\approx [0,5389; 0,6056]$ ✓
- **Conclusion utile :** la borne basse 0,539 dépasse 0,5, donc la majorité est **significative** ✓
- **Avec le premier seul :** $J = [0,50; 0,60]$, la conclusion aurait été impossible ✓
- **Leçon :** réunir les échantillons est le moyen le plus économique de gagner en précision ✓

Réponse. $f \approx 0,572$ sur 900 personnes, avec une marge réduite à $\approx 0,033$: la majorité devient significative.

Correction de l'exercice 12 - Quand la formule ne s'applique pas

- **Fréquence observée :** $f = \frac{27}{30} = 0,90$ ✓
- **Première condition :** $n = 30 \geq 25$, remplie ✓
- **Seconde condition :** $0,2 \leq f \leq 0,8$ ✓
- **Test :** $0,90 > 0,8$, condition **non remplie** ✓
- **Conclusion :** la formule sort de son domaine de validité ✓
- **Application naïve :** $\frac{1}{\sqrt{30}} \approx 0,1826$, donc $J \approx [0,7174; 1,0826]$ ✓
- **Absurdité visible :** la borne haute dépasse 1 ✓
- **Correction minimale :** tronquer à $[0,717; 1]$ ✓
- **Problème restant :** une amplitude de 28 points est inexploitable ✓
- **Raison mathématique :** quand f approche 1, la fluctuation réelle devient plus petite que $\frac{1}{\sqrt{n}}$ ✓
- **Conséquence :** la formule **surestime** l'incertitude dans ce cas ✓
- **Ce qu'on peut dire malgré tout :** la proportion de satisfaits est **très probablement supérieure à 70 %** ✓
- **Bonne pratique :** augmenter n ✓
- **Avec $n = 400$ et $f = 0,90$:** $J = [0,85; 0,95]$, cette fois exploitable ✓
- **Réserve maintenue :** $f = 0,9$ reste hors du domaine $[0,2; 0,8]$, mais l'ordre de grandeur devient fiable ✓
- **Autre limite de l'exemple :** 30 personnes, c'est aussi trop peu pour garantir un tirage représentatif ✓

Réponse. Non : $f = 0,90 > 0,8$ sort du domaine de validité, et l'intervalle obtenu dépasserait 1.

Correction de l'exercice 13 - Marge et pourcentage relatif

- **Marge d'erreur** : $\frac{1}{\sqrt{2500}} = \frac{1}{50} = 0,02$ ✓
- **Fourchette** :

$$J = [0,04 ; 0,08]$$
- **Marge en points** : ± 2 points ✓
- **Marge en relatif** : $\frac{0,02}{0,06} \approx 33,3\%$ ✓
- **Lecture frappante** : le score réel peut être **le tiers** ou **le double** de l'annonce ✓
- **Contrôle** : $\frac{0,04}{0,06} \approx 0,67$ et $\frac{0,08}{0,06} \approx 1,33$ ✓
- **Comparaison avec un gros parti à 40 %** : la marge reste ± 2 points ✓
- **Marge relative correspondante** : $\frac{0,02}{0,40} = 5\%$ ✓
- **Conclusion** : la marge **absolue** est la même, la marge **relative** est six fois plus grande pour le petit parti ✓
- **Contrôle** : $\frac{33,3}{5} \approx 6,7$ ✓
- **Conséquence médiatique** : une variation de 6 à 7 % d'un sondage à l'autre est **sans signification** ✓
- **Alors qu'une variation** de 40 à 41 % l'est tout autant, mais paraît moins spectaculaire ✓
- **Réserve technique** : pour $f = 0,06$, la formule sort du domaine $[0,2 ; 0,8]$ et **surestime** la marge ✓
- **Marge affinée par les instituts** : de l'ordre de $\pm 0,9$ point pour $f = 0,06$ et $n = 2500$ ✓
- **Conclusion nuancée** : la marge réelle est plus petite, mais reste importante en **relatif** ✓

Réponse. Marge ± 2 points, soit $J = [0,04 ; 0,08]$ — une incertitude **relative** de 33 %.

Correction de l'exercice 14 - Estimer une durée moyenne ? Non.

- **Ce que la formule estime** : une **proportion**, c'est-à-dire un nombre de $[0 ; 1]$ ✓
- **Ce que l'élève cherche** : une **moyenne**, exprimée en minutes ✓
- **Incompatibilité d'unités** : ajouter $\frac{1}{\sqrt{n}}$, sans unité, à une durée en minutes n'a pas de sens ✓
- **Exemple concret** : une moyenne de 22 minutes sur 100 trajets donnerait $[21,9 ; 22,1]$ ✓
- **Absurdité** : cette précision de 6 secondes ne dépend ni de la dispersion des durées ni de rien de réel ✓
- **Test de l'absurde** : la même formule donnerait la même marge pour des trajets très réguliers et très irréguliers ✓
- **Ce qui manque** : la formule d'un intervalle pour une moyenne dépend de l'**écart-type** de la série ✓
- **Formule correcte (hors programme de Seconde)** : la marge est proportionnelle à $\frac{\sigma}{\sqrt{n}}$ ✓
- **Ordre de grandeur avec $\sigma = 8$ min et $n = 100$** : $\frac{8}{10} = 0,8$ minute ✓
- **Marge à 95 %** : environ $2 \times 0,8 = 1,6$ minute ✓

- **Fourchette plausible** : [20,4 ; 23,6] minutes ✓
- **Comparaison** : $\pm 1,6$ minute contre $\pm 0,1$ minute par la formule détournée — un facteur 16 ✓
- **Ce qu'on peut faire en Seconde** : estimer la **proportion** de trajets de plus de 30 minutes ✓
- **Exemple** : si 22 trajets sur 100 dépassent 30 minutes, $f = 0,22$ et $J = [0,12 ; 0,32]$ ✓
- **Conclusion** : on transforme la question sur une moyenne en question sur une **proportion** ✓

Réponse. La formule ne vaut que pour une **proportion** (sans unité) ; pour une moyenne, la marge dépend de l'écart-type.

Correction de l'exercice 15 - Sondage biaisé

- **Calcul brut de la marge** : $\frac{1}{\sqrt{5000}} \approx 0,0141$, soit 1,4 point ✓
- **Le calcul est juste** : arithmétiquement, oui ✓
- **Mais la conclusion est fausse** : l'échantillon n'est **pas aléatoire** ✓
- **Nature de l'échantillon** : ce sont des **volontaires**, auto-sélectionnés ✓
- **Biais principal** : ceux qui répondent sont ceux que le sujet intéresse, souvent les plus favorables ✓
- **Second biais** : les visiteurs du site ne représentent pas la population générale ✓
- **Conséquence** : la fourchette [0,856 ; 0,884] n'estime **rien** de la population ✓
- **Ce qu'elle estime au mieux** : l'opinion des visiteurs qui ont voulu répondre ✓
- **Point crucial** : augmenter n ne corrige **pas** un biais ✓
- **Illustration** : avec 500 000 réponses, la marge tomberait à 0,0014 ✓
- **Mais** : l'estimation resterait tout aussi fausse, simplement plus **précise dans l'erreur** ✓
- **Condition indispensable oubliée** : le tirage doit être **au hasard** dans la population visée ✓
- **Exemple historique** : le sondage du *Literary Digest* de 1936, sur plus de deux millions de réponses, s'est trompé de 19 points ✓
- **Cause de cet échec** : l'échantillon avait été tiré parmi les propriétaires de téléphone et d'automobile ✓
- **Leçon** : un petit échantillon **bien tiré** vaut mieux qu'un énorme échantillon **biaisé** ✓

Réponse. Le calcul $\pm 1,4$ point est exact, mais inutilisable : l'échantillon est auto-sélectionné, donc biaisé — et augmenter n ne corrige pas un biais.

Correction de l'exercice 16 - Synthèse : enquête sur la restauration scolaire

a) **Fréquence observée** :

$$f = \frac{144}{225} = 0,64$$

— **Marge d'erreur** : $\frac{1}{\sqrt{225}} = \frac{1}{15} \approx 0,0667$ ✓

— **Fourchette** :

$$J \approx [0,64 - 0,0667 ; 0,64 + 0,0667] = [0,5733 ; 0,7067]$$

— **Formulation** : entre 57,3 % et 70,7 % de satisfaits, au niveau de confiance 95 % ✓

— **Amplitude** : 13,3 points ✓

b) **Passage aux effectifs** : on multiplie par 1250 ✓

- **Borne basse** : $0,5733 \times 1250 \approx 717$ élèves ✓
 - **Borne haute** : $0,7067 \times 1250 \approx 883$ élèves ✓
 - **Estimation ponctuelle** : $0,64 \times 1250 = 800$ élèves ✓
 - **Contrôle** : 800 est bien le centre de $[717; 883]$ ✓
 - **Amplitude en effectifs** : $883 - 717 = 166$ élèves ✓
- c) **Question de la majorité** : le seuil est 0,50 ✓
- **Test** : $0,5733 > 0,50$, donc $0,50 \notin J$ ✓
 - **Conclusion** : **oui**, on peut affirmer qu'une majorité est satisfaite ✓
 - **Marge de sécurité** : $0,5733 - 0,50 \approx 7,3$ points ✓
 - **Effectif minimal de satisfaits** : 717, soit bien plus que la moitié de 1250 (625) ✓
 - **Contrôle** : $717 - 625 = 92$ élèves de marge ✓
 - **Formulation à donner au proviseur** : « au moins 57 % des élèves sont satisfaits, donc une majorité l'est » ✓
- d) **Taille nécessaire** ✓
- **Condition** : $\frac{1}{\sqrt{n}} \leq 0,03$ ✓
 - **Résolution** : $\sqrt{n} \geq \frac{1}{0,03} \approx 33,33$ ✓
 - **Donc** :
- $$n \geq 1111 \text{ élèves}$$
- **Comparaison à l'effectif du lycée** : 1111 sur 1250, soit 88,9 % des élèves ✓
 - **Verdict** : c'est presque un **recensement** ✓
 - **Conséquence** : autant interroger tout le monde, ce qui donnerait la proportion **exacte** ✓
 - **Nuance mathématique importante** : quand l'échantillon représente une grande part de la population, la précision réelle est **meilleure** que $\frac{1}{\sqrt{n}}$ ✓
 - **Cas limite** : avec $n = 1250 = N$, la marge réelle est **nulle** ✓
 - **Recommandation pratique** : viser 400 élèves, soit une marge de 5 points et 32 % du lycée ✓
 - **Contrôle** : $\frac{1}{\sqrt{400}} = 0,05$ et $\frac{400}{1250} = 32 \%$ ✓
 - **Bilan** : dans une petite population, l'enquête exhaustive devient rapidement plus simple qu'un sondage précis ✓
- Réponse.** a) $J \approx [0,573; 0,707]$ · b) entre 717 et 883 élèves · c) oui, $0,50 \notin J$ · d) $n \geq 1111$, soit 89 % du lycée — mieux vaut recenser.