TS — Exercices sur l'échantillonnage

Pour les intervalles de fluctuation asymptotique au seuil 95 %, on utilisera la formule :

$$\left[p-u_{0,05}\sqrt{\frac{p(1-p)}{n}}\ ;\ p+u_{0,05}\sqrt{\frac{p(1-p)}{n}}\right]$$

ou, éventuellement, $\left[p-1,96\sqrt{\frac{p(1-p)}{n}}\ ;\ p+1,96\sqrt{\frac{p(1-p)}{n}}\right]$.

Intervalle de fluctuation

1 Dans une maternité, on admet qu'il naît en moyenne 51 % de garçons. On fait le point sur la fréquence de garçons toutes les 100 naissances.
Déterminer l'intervalle de fluctuation asymptotique de cette fréquence au seuil de 95 %.
On donnera des valeurs décimales approchées des bornes à 10^{-3} près.

2 22 % des Français sont d'accord pour supprimer les panneaux indiquant la présence de radars sur les routes.
Déterminer l'intervalle de fluctuation asymptotique au seuil de 95 % de la fréquence des personnes désirant supprimer les panneaux avertissant des radars dans un échantillon de 2000 personnes.

3 Une boîte contient 20 % de réglisses et 80 % de guimauves.
Déterminer l'intervalle de fluctuation asymptotique au seuil de 99 % de la fréquence de réglisses dans un échantillon de 100 bonbons tirés au hasard et avec remise dans la boîte.

4 La proportion d'un caractère dans une population est $p=0,6$.
Déterminer l'intervalle de fluctuation asymptotique au seuil de 95 % de la fréquence de ce caractère dans les échantillons de taille 100 prélevés au hasard et avec remise.

5 L'amplitude de l'intervalle de fluctuation asymptotique de la fréquence d'un caractère au seuil de 95 %, dans un échantillon de taille n, avec $p=0,5$, est égale à 0,049.
Quelle est la taille de cet échantillon ?

Prise de décision

6 On lance 100 fois une pièce de monnaie et on obtient 38 « pile » et 62 « face ».
Peut-on considérer que la pièce est équilibrée au seuil de décision de 5 % ?

7 On lance 240 fois un dé cubique non truqué. On obtient 52 fois le numéro 6.
Peut-on considérer que le dé est non truqué au seuil de décision de 1 % ?

8 La proportion de naissances d'enfants prématurés est de 6 %. Des chercheurs suggèrent que les femmes ayant eu un travail pénible pendant leur grossesse sont plus susceptibles d'avoir un enfant prématuré. On réalise une enquête auprès d'un échantillon aléatoire de 400 naissances correspondant à des femmes ayant eu pendant leur grossesse un travail pénible. Les chercheurs décident *a priori* que si la proportion d'enfants nés prématurés dans cet échantillon est supérieure à la borne supérieure de l'intervalle de fluctuation asymptotique au seuil de 0,95 alors leur hypothèse sera acceptée. On sait que dans l'échantillon, il y a 50 enfants prématurés.
1°) Déterminer l'intervalle de fluctuation asymptotique I associé à cette situation au seuil de 95 %.
2°) Quelle est la conclusion ?

9 On considère une maladie bénigne qui guérit naturellement en moins de cinq jours dans 60 % des cas. On veut tester un médicament censé abréger la durée de la maladie. Pour cela, on administre le médicament à 1000 personnes. Pour 63 % d'entre elles, la guérison a eu lieu en moins de cinq jours.
1°) Déterminer l'intervalle de fluctuation asymptotique I au seuil de 95 % pour un échantillon de taille 1000 associé à cette situation, après avoir justifié les conditions de validité.
2°) Que peut-on penser de l'efficacité de ce médicament ?

Intervalle de confiance

Pour les intervalles de confiance au niveau 95 %, on utilisera, sauf mention contraire, la formule :

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

10 Un candidat à une élection fait effectuer un sondage. Sur 100 personnes interrogées, 63 déclarent vouloir voter pour lui. On suppose que les électeurs ne changent pas d'avis le jour du vote. On notera p le pourcentage de voix obtenues par le candidat. On suppose que $0,5\leqslant p\leqslant 0,7$.
Déterminer l'intervalle de confiance de p au niveau de confiance de 95 %.
Énoncer le résultat en langage courant.

11 Lors d'une épidémie de grippe, 13 des élèves d'une classe qui compte 34 élèves ont contracté la maladie. On suppose que cette classe constitue un échantillon représentatif de l'ensemble des 850 élèves du lycée.
1°) Donner la fréquence d'élèves malades dans la classe.
2°) Donner cet intervalle de confiance au niveau de confiance de 95 % de la proportion d'élèves malades de la grippe dans le lycée.
3°) Louis, élève de la classe, estime qu'il est possible que plus de la moitié des élèves du lycée aient eu la grippe.
Que peut-on en penser ?

12 Une urne contient des boules blanches et des boules noires. On effectue 100 tirages avec remise dans cette urne et on obtient 39 boules noires.
Peut-on affirmer, avec une probabilité d'au moins 0,95, qu'il y a plus de boules blanches que de boules noires dans l'urne ?

13 Pour décider la construction d'un grand stade, une municipalité veut sonder la population pour estimer si plus de 50 % des électeurs y sont favorables.
1°) La municipalité réalise un sondage aléatoire de taille 100 et obtient 54 avis favorables.
a) Quelle est la fréquence f d'avis favorables sur ce sondage ?
b) La municipalité peut-elle décider la construction du stade en prétextant que plus de 50 % de la population y est favorable ?
2°) On suppose que la municipalité réalise un sondage de taille n et que la fréquence f des votes favorables reste égale à 0,54.
Si p est la proportion (inconnue) d'avis favorables dans la population, donner l'expression d'un intervalle de confiance de p au niveau de 95 %.
3°) La municipalité ne construira le stade que si la proportion p d'avis favorables dépasse 50 %.
a) Démontrer que pour avoir $p > 0,5$, au niveau de confiance de 95 %, il suffit d'avoir : $0,54 - \frac{1}{\sqrt{n}} > 0,5$ (i).
b) Déterminer le plus petit entier naturel n_0 à partir duquel l'inégalité (i) est vérifiée.
c) Si le sondage, avec la même fréquence observée, portait sur 650 personnes, peut-on penser que le stade aurait été construit ?

14 Une enquête statistique a été menée dans différentes maternités pour étudier l'impact du tabac sur le poids des bébés à la naissance. On s'intéresse à trois populations : les mères non fumeuses, les mères fumeuses qui s'arrêtent de fumer au début de la grossesse et les mères fumeuses qui continuent à fumer pendant la grossesse.
Un échantillon aléatoire de taille 100 est constitué dans chacune des trois populations ; les proportions de bébés pesant moins de 2,5 kg à la naissance sont respectivement 5 %, 6 % et 11 %.
1°) Pour chacune des trois populations, déterminer un intervalle de confiance au niveau de 95 % de la proportion de bébés pesant moins de 2,5 kg à la naissance.
On utilisera l'intervalle de confiance « standard » au niveau de 95 % :
$\left[f - 1,96\sqrt{\frac{f(1-f)}{n}} \,;\, f + 1,96\sqrt{\frac{f(1-f)}{n}} \right]$; on arrondira les bornes au millième.
2°) En exploitant les intervalles de confiance précédents, peut-on considérer qu'il existe une différence significative de la proportion de bébés de poids inférieur à 2,5 kg à la naissance, dans chacune des trois populations ?

15 Pour comparer les cotes de popularité de trois personnalités X, Y et Z, un journal a réalisé un sondage : sur 1 320 personnes, 27 % ont voté pour X, 38,5 % pour Y et 34,5 % pour Z.
1°) Estimer par un intervalle la cote de popularité de X, Y et Z au niveau de confiance de 95 %, dans la population. Un classement peut-il être publié ?
2°) Avec ces mêmes taux, pour quelle taille du sondage, un classement aurait-il été statistiquement fondé au niveau de confiance de 95 % ?

Comparaison de proportions

16 Un chercheur veut tester un traitement. Pour cela, il considère deux groupes de malades ayant la même pathologie.
Le groupe A est composé de 150 malades à qui il donne le traitement. Pour 121 patients, l'effet de la maladie s'estompe.
Le groupe B est composé de 150 malades à qui il administre un placebo. Pour 94 malades, l'effet de la maladie s'estompe.
1°) Déterminer, à 10^{-2} près, f_A et f_B les fréquences des personnes de chaque groupe pour lesquelles les effets de la maladie s'estompent.
2°) Déterminer, à 10^{-2} près, l'intervalle de confiance au niveau de confiance 0,95 pour chaque groupe.
3°) On considère que le traitement a un effet si les intervalles de confiance au niveau de confiance 0,95 sont disjoints.
Que peut conclure le chercheur sur l'effet du traitement qu'il a mis en place ?

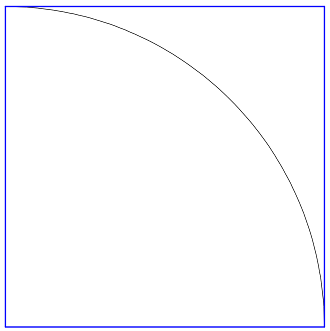
17 Une importante société de vente au détail de matériel informatique veut juger de l'impact d'une campagne publicitaire menée dans les médias pour une tablette numérique particulière.
Dans un échantillon, considéré comme prélevé au hasard et avec remise, de 200 ventes de tablettes effectuées avant la campagne publicitaire, on observe que 44 ventes concernent cette tablette.
Dans un nouvel échantillon de même taille et prélevé dans les mêmes conditions après la campagne publicitaire, on observe que 66 ventes concernent cette tablette ;
1°) Déterminer un intervalle de confiance, avec un niveau de confiance de 95 %, pour la proportion des ventes de cette tablette avant la campagne publicitaire (arrondir les bornes au millième).
2°) Même question après la campagne publicitaire.
3°) La différence entre les deux fréquences observées est-elle significative ?
4°) Le responsable de la campagne publicitaire affirme : « Les deux échantillons montrent que les ventes de cette tablette ont augmenté de 50 % ; donc la campagne a été très efficace ». Que peut-on en penser ?

18 Une entreprise veut comparer les performances de deux de ses commerciaux.
En un mois, le premier commercial a démarché 160 clients et il a conclu un contrat avec 65 de ces clients.
Dans le même temps, le second commercial a démarché 220 clients et conclu un contrat avec 100 de ces clients.
Peut-on conclure, au niveau de confiance 95 %, que ces deux commerciaux ont la même efficacité ?

19 Deux entreprises A et B fabriquent le même type d'ampoules électriques.
Sur un échantillon de 100 ampoules de l'entreprise A, on a trouvé 60 ampoules d'une durée de vie supérieure à 100 h.
Sur un échantillon de 200 ampoules de l'entreprise B, on a trouvé 140 ampoules d'une durée de vie supérieure à 100 h.
Peut-on conclure au niveau de confiance de 99 % qu'il y a une différence significative entre les durées de vie des ampoules fabriquées par les entreprises A et B ?

20 Dans une expérience à grande échelle, on sépare 2000 enfants en deux groupes de 1000.
Un groupe reçoit un sérum pour la prévention de la maladie, tandis que l'autre ne le reçoit pas.
Le nombre d'enfants de chaque groupe qui contracte la maladie est respectivement égal à 30 et à 50.
Peut-on conclure au niveau de confiance 95 % que la différence observée entre les deux groupes est significative ?

21 On considère un quart de cercle de rayon 1 inscrit dans un carré de côté 1. On note p la probabilité qu'un point choisi au hasard dans ce carré soit dans le quart de disque.



Sur 2 500 points choisis au hasard dans ce carré, 1 977 sont situés dans le quart de disque.

1°) Déterminer l'intervalle de confiance de p au niveau 95 %. Écrire ses bornes avec quatre décimales.
2°) En déduire un encadrement de π au niveau de confiance de 95 %.

Corrigé

Intervalle de fluctuation

On peut concevoir un programme sur calculatrice qui permet de donner les bornes des intervalles de fluctuation (éventuellement avec un menu : seuil, intervalle simplifié ou non…).

$\boxed{1}$

Déterminons l'intervalle de fluctuation asymptotique de la fréquence de naissances de garçons dans la maternité au seuil de 95 %.

$n = 100$
$p = 0,51$
$1 - p = 0,49$

On vérifie les conditions d'application de la formule du cours :

$n \geqslant 30$
$np \geqslant 5$
$n(1-p) \geqslant 5$

L'intervalle de fluctuation asymptotique de la fréquence de naissances de garçons dans la maternité au seuil de 95 % est : $\text{I} = \left[p - 1,96\sqrt{\dfrac{p(1-p)}{n}} \,;\, p + 1,96\sqrt{\dfrac{p(1-p)}{n}} \right]$.

On calcule les bornes de l'intervalle I.

$p - 1,96\sqrt{\dfrac{p(1-p)}{n}} \approx 0,412$ (valeur décimale approchée au millième par défaut)

$p + 1,96\sqrt{\dfrac{p(1-p)}{n}} \approx 0,608$ (valeur décimale approchée au millième par excès)

$\text{I} = [0,412 \,;\, 0,608]$ (le signe $=$ est faux à cause des valeurs approchées que l'on a pris à la place des bornes)

$\boxed{2}$

$n = 2000$
$p = 0,22$
$1 - p = 0,78$

Déterminons l'intervalle de fluctuation asymptotique de la fréquence des personnes désirant supprimer les panneaux avertissant des radars dans un échantillon de 2000 personnes au seuil de 95 %.

On vérifie les conditions d'application de la formule du cours :

$n \geqslant 30$
$np \geqslant 5$
$n(1-p) \geqslant 5$

$$p - 1,96\sqrt{\frac{p(1-p)}{n}} = 0,22 - 1,96\sqrt{\frac{0,22 \times 0,78}{2000}} = 0,2018...$$

$$p + 1,96\sqrt{\frac{p(1-p)}{n}} = 0,22 + 1,96\sqrt{\frac{0,22 \times 0,78}{2000}} = 0,2381...$$

On peut donner comme intervalle de fluctuation asymptotique au seuil de 95 % l'intervalle $\text{I} = [0,201 \,;\, 0,239]$.

$\boxed{3}$

$n = 100$
$p = 0,2$
$1 - p = 0,8$

Déterminons l'intervalle de fluctuation asymptotique de la fréquence de réglisses dans un échantillon de 100 bonbons tirés au hasard et avec remise dans la boîte au seuil de 99 %.

On vérifie les conditions d'application de la formule du cours :

$n \geqslant 30$
$np \geqslant 5$
$n(1-p) \geqslant 5$

On calcule les bornes de l'intervalle de fluctuation asymptotique au seuil de 99 %.

$$p - 2,58\sqrt{\frac{p(1-p)}{n}} \approx 0,096$$

$$p + 2,58\sqrt{\frac{p(1-p)}{n}} \approx 0,304$$

$\text{I} = [0,096 \,;\, 0,304]$

$\boxed{4}$

$n = 100$
$p = 0,6$
$1 - p = 0,4$

Déterminons l'intervalle de fluctuation asymptotique de la fréquence de ce caractère dans les échantillons de taille 100 prélevés au hasard et avec remise au seuil de 95 %.

On vérifie les conditions d'application de la formule du cours :

$n \geqslant 30$
$np \geqslant 5$
$n(1-p) \geqslant 5$

$\mathrm{I} = [0,503\,;\,0,697]$

$\boxed{5}$

$p = 0,5$

Déterminons la taille n de l'échantillon.

L'amplitude de l'intervalle de fluctuation à 95 % est égale à :

$$\left(p + 1,96\sqrt{\frac{p(1-p)}{n}}\right) - \left(p - 1,96\sqrt{\frac{p(1-p)}{n}}\right) = 2 \times 1,96\sqrt{\frac{p(1-p)}{n}}$$

$$= 2 \times 1,96\sqrt{\frac{0,5 \times 0,5}{n}}$$

$$= 2 \times 1,96 \times 0,5 \times \frac{1}{\sqrt{n}}$$

$$= \frac{1,96}{\sqrt{n}}$$

On a : $\dfrac{1,96}{\sqrt{n}} = 0,049$ (1).

(1) $\Leftrightarrow \sqrt{n} = \dfrac{1,96}{0,049}$

$\Leftrightarrow \sqrt{n} = 40$

$\Leftrightarrow n = 1600$

On vérifie que les trois conditions pour déterminer l'intervalle de fluctuation asymptotique sont vérifiées (notamment $n \geqslant 30$ ainsi que les deux autres conditions).

Prise de décision sur échantillon à partir d'une fréquence

$\boxed{6}$ Lancers d'une pièce de monnaie

$n = 100$

On raisonne avec la fréquence de pile (on pourrait raisonner également avec la fréquence de face mais il est inutile de faire les deux).

La fréquence de pile est $f = 0,38$.

On veut tester l'hypothèse « la probabilité d'obtenir pile en un lancer est $p = 0,5$ » avec un risque de 5 % (au plus).

On vérifie les conditions :

$n \geqslant 30$
$np \geqslant 5$
$n(1-p) \geqslant 5$

On peut donc donner l'intervalle de fluctuation asymptotique au seuil de 95 %.

- avec la formule générale :

$\mathrm{I} = [0,402\,;\,0,598]$ (bornes données sous la forme de valeurs décimales approchées au millième)

- avec la formule simplifiée :

$$\mathrm{I} = \left[p - \frac{1}{\sqrt{n}}\,;\,p + \frac{1}{\sqrt{n}}\right]$$

$\mathrm{I} = [0,4\,;\,0,6]$

0,38 $\notin$ I donc on rejette l'hypothèse « la pièce est équilibrée » au seuil de décision 0,95 (car 0,38 n'appartient pas à I).

Quelques remarques :

• On peut aussi travailler avec la fréquence de « face ». On obtient le même résultat.

• On pourrait aussi prendre l'intervalle de fluctuation asymptotique plus précis :
$$\mathrm{J} = \left[p - 1,96\sqrt{\frac{p(1-p)}{n}}\,;\,p + 1,96\sqrt{\frac{p(1-p)}{n}}\right].$$
Cet intervalle est inclus dans J. Comme 0,38 n'appartient pas à J, 0,38 n'appartient pas non plus à I.

$\boxed{7}$ Lancers d'un dé cubique

$n = 240$

On veut tester l'hypothèse « la probabilité d'obtenir le numéro 6 en un lancer est $p = \frac{1}{6}$ » avec un risque d'erreur de 1 % (au plus).

On utilise la notion d'intervalle asymptotique de la fréquence.

On vérifie les conditions de validité :

$n \geqslant 30$
$np \geqslant 5$
$n(1-p) \geqslant 5$

On peut donc donner l'intervalle de fluctuation asymptotique au seuil de 99 % de la fréquence du numéro 6 dans un échantillon de taille 240.

$$\mathrm{I} = \left[\frac{1}{6} - 2,58\sqrt{\frac{\frac{1}{6}\times\frac{5}{6}}{240}} \; ; \; \frac{1}{6} + 2,58\sqrt{\frac{\frac{1}{6}\times\frac{5}{6}}{240}} \right]$$

Les bornes de I sont : 0,1046015… et 0,2287318… .

$f = \frac{52}{240} = 0,2166\underline{6}...$

$f \in \mathrm{I}$ donc on ne rejette pas l'hypothèse que le dé est équilibré avec un risque de 1 %.

$\boxed{8}$ Naissances prématurées

$n = 400$

1°) On veut tester l'hypothèse : « La proportion de naissances prématurées est $p = 0,06$ ».

Déterminons l'intervalle de fluctuation asymptotique I associé à cette situation au seuil de 0,95.

$n \geqslant 30$
$np \geqslant 5$
$n(1-p) \geqslant 5$

On détermine un intervalle de fluctuation asymptotique au seuil de 95 % de la fréquence de naissance prématurées dans un échantillon de taille 400.

- avec la formule simplifiée :

$$\mathrm{I} = \left[p - \frac{1}{\sqrt{n}} \; ; \; p + \frac{1}{\sqrt{n}} \right]$$

$\mathrm{I} = [0,01 \; ; 0,11]$

- avec la formule non simplifiée :

$\mathrm{I} = [0,0367 \; ; 0,0833]$ (assez différent du résultat avec la formule simplifiée)

2°) **Décision**

La fréquence observée de naissances prématurées dans l'échantillon est égale à $f = \frac{50}{400} = 0,125$.

0,125 $\notin$ I et 0,125 est strictement supérieure à la borne de droite de l'intervalle I donc l'hypothèse des scientifiques peut être acceptée (cf. énoncé).

$\boxed{9}$ Efficacité d'un médicament

$n = 1000$

$n \geqslant 30$

1°) On veut tester l'hypothèse : « La proportion de personnes guéries en moins de cinq jours est $p = 0,6$ ».

Déterminons l'intervalle de fluctuation asymptotique I associé à cette situation au seuil de 0,95.

$$\mathrm{I} = \left[0,6 - 1,96\sqrt{\frac{0,6\times 0,4}{1000}} \; ; \; 0,6 + 1,96\sqrt{\frac{0,6\times 0,4}{1000}} \right]$$ (on prend cette formule pour davantage de précision)

I = [0,569 ; 0,631] (bornes données en valeurs décimales approchées au millième)

2°) **Décision**

La fréquence observées de personnes ayant pris le médicament et guéries en moins de cinq jours est $f = 0,63$.

$f \in$ I donc on ne rejette pas l'hypothèse au seuil de 95 % et on peut en conclure que le médicament n'est pas efficace.

Intervalle de confiance d'une proportion

10 Élections

Un candidat à une élection fait effectuer un sondage. Sur 100 personnes interrogées, 63 déclarent vouloir voter pour lui. On suppose que les électeurs ne changent pas d'avis le jour du vote. On notera p le pourcentage de voix obtenues par le candidat. On suppose que $0,5 \leqslant p \leqslant 0,7$.

Déterminons l'intervalle de confiance de p au niveau de confiance 0,95.

$n = 100$

$f = 0,63$

On vérifie la condition de validité avant de commencer le calcul : $n \geqslant 30$.

On utilise la formule permettant de déterminer un intervalle de confiance de p au niveau 95 %.

$\mathrm{I} = \left[f - \frac{1}{\sqrt{n}} \; ; \; f + \frac{1}{\sqrt{n}} \right]$ (on adopte cette formule en l'absence d'indication de l'énoncé ; on évite d'utiliser la formule plus compliquée)

$$\mathrm{I} = \left[0,63 - \frac{1}{\sqrt{100}} \; ; \; 0,63 + \frac{1}{\sqrt{100}} \right]$$

$\mathrm{I} = [0,53 \; ; \; 0,73]$ (on parle de « fourchette de sondage »)

On vérifie les conditions de validité après les calculs : $np_{\text{min}} \geqslant 5$ et $n(1 - p_{\text{max}}) \geqslant 5$ (avec ici : $p_{\text{min}} = 0,53$ et $p_{\text{max}} = 0,73$).

Énonçons le résultat en langage courant.

On peut dire que la proportion de personnes qui vont voter pour le candidat est comprise entre 53 % et 73 % au niveau de confiance 95 %.
Donc le candidat a de grandes chances d'être élu.

11

1°) **Donnons la fréquence d'élèves malades dans la classe.**

$$f = \frac{13}{34}$$

2°) **Déterminons l'intervalle de confiance au niveau de 95 % de la proportion d'élèves malades de la grippe dans le lycée.**

$n = 34$

On vérifie la condition de validité avant de commencer les calculs : $n \geqslant 30$.

L'intervalle de confiance au niveau de 95 % de la proportion d'élèves malades de la grippe dans le lycée est $\mathrm{I} = \left[\frac{13}{34} - \frac{1}{\sqrt{34}} \; ; \; \frac{13}{34} + \frac{1}{\sqrt{34}} \right]$.

Avec la calculatrice, on trouve :

$$\frac{13}{34} - \frac{1}{\sqrt{34}} = 0,21085435...$$

$$\frac{13}{34} + \frac{1}{\sqrt{34}} = 0,553851526...$$

En prenant des valeurs décimales approchées des bornes à 10^{-3} près (par défaut pour la borne pour la borne inférieure et par excès pour la borne supérieure), l'intervalle de confiance au niveau de 95 % de la proportion d'élèves malades de la grippe dans le lycée est $[0,210 \; ; \; 0,554]$.

On vérifie les conditions de validité après les calculs : $np_{\text{min}} \geqslant 5$ et $n(1 - p_{\text{max}}) \geqslant 5$.

On n'utilise pas la valeur 850 fournie par l'énoncé !

3°) **Louis, élève de la classe, estime qu'il est possible que plus de la moitié des élèves du lycée aient eu la grippe.**
Que peut-on en penser ?

La proportion de personnes ayant eu la grippe est comprise entre 21 % et 55,4 % au niveau de confiance 0,95.
Donc il est effectivement possible que plus de la moitié des élèves du lycée aient eu la grippe mais rien ne permet de l'affirmer avec certitude.

$\boxed{12}$ Une urne contient des boules blanches et des boules noires. On effectue 100 tirages avec remise dans cette urne et on obtient 39 boules noires.
Peut-on affirmer, avec une probabilité d'au moins 0,95, qu'il y a plus de boules blanches que de boules noires dans l'urne ?

$n=100$

La fréquence observée de boules noires dans l'échantillon est $f=\frac{39}{100}$.

On vérifie la condition de validité avant de commencer le calcul : $n\geqslant 30$.

L'intervalle de confiance de la proportion de boules noires au niveau de confiance de 95 % est $\left[0,29\,;0,49\right]$ (valeurs approchées des bornes à 10^{-2} près).

On vérifie les conditions de validité après les calculs : $np_{\min}\geqslant 5$ et $n\left(1-p_{\max}\right)\geqslant 5$.

Donc on peut affirmer qu'il y a plus de boules blanches que de boules noires dans l'urne avec un niveau de confiance de 95 %.

$\boxed{13}$

1°) a) $f=\frac{54}{100}=0,54$

b) On ne connaît l'avis que d'un échantillon et non de la population totale. La municipalité ne peut donc pas décider de construire le stade.
Dans la question 2°), on va estimer la proportion d'avis favorables dans la population.

2°) IC à 95 %

$$\mathrm{I}=\left[f-\frac{1}{\sqrt{n}}\,;\,f+\frac{1}{\sqrt{n}}\right]$$
$$\mathrm{I}=\left[0,44\,;0,64\right]$$

Par suite, on ne peut pas dire si la municipalité décidera de construire le stade ou non.
Notre échantillon est trop petit pour apporter une conclusion (amplitude de l'intervalle de confiance trop large).

3°) Pour avoir $p>0,5$, il suffit que toutes les valeurs de l'intervalle de confiance soient strictement supérieures à 0,5 c'est-à-dire que $f-\frac{1}{\sqrt{n}}>0,5$ soit $0,54-\frac{1}{\sqrt{n}}>0,5$ (i).

On cherche $n\in\mathbb{N}^*$ tel que (i) soit vérifiée.

$$(i)\Leftrightarrow 0,54-\frac{1}{\sqrt{n}}>0,5$$
$$\Leftrightarrow 0,04>\frac{1}{\sqrt{n}}$$
$$\Leftrightarrow \sqrt{n}>25$$
$$\Leftrightarrow n>625$$

L'entier n_0 cherché est 626.

c) $650>626$ donc, d'après la question précédente, on peut penser que le stade aurait été construit.

Comparaison de deux proportions à partir d'échantillons

$\boxed{14}$

1°)

Mères non fumeuses : $\mathrm{I}=\left[0,007\,;0,093\right]$
Mères fumeuses qui arrêtent de fumer à la naissance : $\mathrm{I}'=\left[0,013\,;0,107\right]$
Mères fumeuses : $\mathrm{I}''=\left[0,049\,;0,171\right]$

2°) Les intervalles ne sont pas disjoints donc on peut considérer qu'il n'existe pas de différence significative entre les 3 populations au seuil de 95 %.

$\boxed{15}$

1°)

$\mathrm{I}_{\mathrm{X}}=\left[0,246\,;0,294\right]$
$\mathrm{I}_{\mathrm{Y}}=\left[0,358\,;0,412\right]$
$\mathrm{I}_{\mathrm{Z}}=\left[0,319\,;0,371\right]$

L'intervalle de confiance de X nous permet de dire que ce candidat ne sera pas élu.
En revanche, les intervalles de confiance de Y et Z ne sont pas disjoints. Par conséquent, aucun classement n'est possible entre ces deux candidats au seuil de 95 %.

2°) Classement possible si :

$$0,27+1,96\sqrt{\frac{0,27\times 0,73}{n}}<0,345-1,96\sqrt{\frac{0,345\times 0,655}{n}}$$
$$1,96\times\frac{\sqrt{0,1971}+\sqrt{0,225975}}{\sqrt{n}}<0,075$$

$$\sqrt{n}>1,96\times\frac{\sqrt{0,1971}+\sqrt{0,225975}}{0,075}$$
$$n>\left(1,96\times\frac{\sqrt{0,1971}+\sqrt{0,225975}}{0,075}\right)^2$$

$n>577$

et si

$$0,345+1,96\sqrt{\frac{0,345\times 0,655}{n}}<0,385-1,96\sqrt{\frac{0,385\times 0,015}{n}}$$

$$1,96\times\frac{\sqrt{0,225975}+\sqrt{0,236775}}{\sqrt{n}}<0,04$$

$$\sqrt{n}>1,96\times\frac{\sqrt{0,225975}+\sqrt{0,236775}}{0,04}$$

$$n>\left(1,96\times\frac{\sqrt{0,225975}+\sqrt{0,236775}}{0,04}\right)^2$$

$$n>2221,8$$

Il faut donc (ou il suffit d') avoir un échantillon supérieur à 2222 personnes.

$\boxed{16}$

1°)

$f_{\text{A}}=\dfrac{121}{150}$ $(=0,8066666666\ldots$ soit environ 0,81)

$f_{\text{B}}=\dfrac{94}{150}$ (environ 0,63)

2°) $\text{I}_{\text{A}}=\left[f_{\text{A}}-1,96\sqrt{\dfrac{f_{\text{A}}(1-f_{\text{A}})}{n}}\;;\;f_{\text{A}}+1,96\sqrt{\dfrac{f_{\text{A}}(1-f_{\text{A}})}{n}}\right]$

$\text{I}_{\text{A}}=[0,72\;;0,88]$

$\text{I}_{\text{B}}=[0,55\;;0,71]$

3°) $\text{I}_{\text{A}}\cap\text{I}_{\text{B}}=\varnothing$

Donc on peut juger que le traitement a un effet au seuil de 5 %.

$\boxed{17}$

1°)

$$\text{I}_{\text{A}}=\left[f_{\text{A}}-1,96\sqrt{\frac{f_{\text{A}}(1-f_{\text{A}})}{n}}\;;\;f_{\text{A}}+1,96\sqrt{\frac{f_{\text{A}}(1-f_{\text{A}})}{n}}\right]$$

$f_{\text{A}}=\dfrac{44}{200}$ et $n=200$

$\text{I}_{\text{A}}=[0,163\;;0,277]$

2°)

$$\text{I}_{\text{B}}=\left[f_{\text{B}}-1,96\sqrt{\frac{f_{\text{B}}(1-f_{\text{B}})}{n}}\;;\;f_{\text{A}}+1,96\sqrt{\frac{f_{\text{B}}(1-f_{\text{B}})}{n}}\right]$$

$f_{\text{B}}=\dfrac{66}{200}$ $\quad n=200$

$\text{I}_{\text{B}}=[0,265\;;0,395]$

3°) Les deux intervalles I_{A} et I_{B} ne sont pas disjoints.
Au seuil de 5 %, la différence entre les deux fréquences observées f_{A} et f_{B} avant et après la campagne publicitaire n'est donc pas significative.

4°) $\dfrac{f_{\text{B}}-f_{\text{A}}}{f_{\text{A}}}=\dfrac{0,33-0,22}{0,22}=0,5=50\ \%$

On applique la formule d'un pourcentage d'évolution : $\dfrac{\text{valeur finale}-\text{valeur initiale}}{\text{valeur initiale}}$.

Le responsable de la campagne a raison d'affirmer que pour les deux échantillons les ventes de cette tablette ont augmenté de 50 %.
Cependant, comme la différence entre f_{A} et f_{B} n'est pas significative, on juge que les deux proportions p_{A} et p_{B} sont égales (avec un petit risque d'erreur).
Le responsable de la campagne ne peut donc pas affirmer, sur la seule base de ces deux échantillons, que la campagne publicitaire a été très efficace.

$\boxed{18}$

Commercial 1 : $\text{I}_1=[0,327\;;0,406]$

Commercial 2 : $\text{I}_2=[0,387\;;0,455]$

Les deux intervalles n'étant pas disjoints, on ne peut donc pas conclure qu'il y a une différence significative au seuil de 5 %.
On ne peut rien conclure sur l'efficacité des deux commerciaux.

19

On applique la formule de l'intervalle $\mathrm{IC} = \left[f - \dfrac{1,29}{\sqrt{n}} \,;\, f + \dfrac{1,29}{\sqrt{n}} \right]$ au niveau de confiance de 99 %.

Entreprise A : $\mathrm{I}_{\mathrm{A}} = [0,471 \,;\, 0,6]$

Entreprise B : $\mathrm{I}_{\mathrm{B}} = [0,571 \,;\, 0,7]$

Les deux intervalles n'étant pas disjoints, on ne peut donc pas conclure qu'il y a une différence significative entre les durées de vie des ampoules des entreprises A et B.

20

$f_{\mathrm{A}} = \dfrac{30}{1000} = 0,03$

$f_{\mathrm{B}} = \dfrac{50}{1000} = 0,05$

$\mathrm{I}_{\mathrm{A}} = [-0,001 \,;\, 0,062]$ (on ne s'étonnera pas que la borne inférieure est négative ; la formule donnant un intervalle plus précis permet d'avoir une borne de gauche positive ou nulle)

$\mathrm{I}_{\mathrm{B}} = [0,018 \,;\, 0,082]$

On notera bien qu'il est tout à fait possible d'avoir une borne négative dans un intervalle de confiance.

$\mathrm{I}_{\mathrm{A}} \cap \mathrm{I}_{\mathrm{B}} \neq \varnothing$

La différence n'est pas donc pas significative.