Exercice 1

Les enfants sont dits prématurés lorsque la durée gestationnelle est inférieure ou égale à 259 jours. La proportion de ces naissances 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 que les autres.

Il est décidé de réaliser 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.

Finalement le nombre d'enfants prématurés est de 50. Quelle est donc la conclusion ?

Solution proposée

Sous l'hypothèse que la proportion de prématurés dans l'échantillon est la même que dans la population générale, on détermine l'intervalle de fluctuation asymptotique au seuil 0,95 :
$n = 400 \geq 30$, $\quad np = 24 \geq 5$ et $n(1-p) = 376 \geq 5$

$$\left[0{,}06 - 1{,}96 \times \frac{\sqrt{0{,}06 \times 0{,}94}}{\sqrt{400}} \; ; \; 0{,}06 + 1{,}96 \times \frac{\sqrt{0{,}06 \times 0{,}94}}{\sqrt{400}}\right] \approx [0{,}036 \; ; 0{,}084]$$

On calcule la valeur observée de proportion de prématurés dans l'échantillon et on obtient $\frac{50}{400} = 0{,}125$.

Cette valeur n'appartient pas à l'intervalle de fluctuation asymptotique au seuil de 95 %, donc avec la règle de décision choisie, on rejette l'hypothèse posée. Les chercheurs concluent donc que la proportion d'enfants prématurés est plus élevée chez les femmes ayant eu un travail pénible pendant leur grossesse.

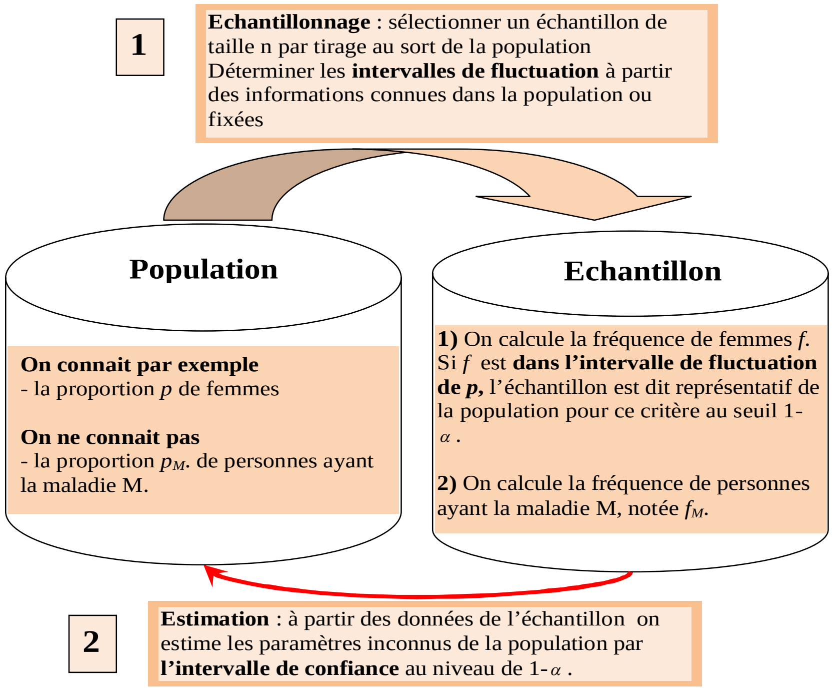
Exercice 2

Dans un lycée qui comporte 1 500 demi-pensionnaires, un sondage sur la restauration a été réalisé en interrogeant 338 élèves qui déjeunent à la cantine.

Ce lycée comporte, parmi les demi-pensionnaires, 54,8 % de filles et 174 d'entre elles ont participé à ce sondage. De plus il y a 11,8 % de sportifs de haut niveau et 47 d'entre eux ont participé à ce sondage, parmi les demi-pensionnaires.

1) Cet échantillon peut-il être considéré comme représentatif de la population du lycée ?
2) A la question « Préféreriez-vous vous servir vous-même ? », 181 élèves répondent oui. Au vu de ce résultat, si vous étiez consulté au sujet de ce changement, quelle décision prendriez-vous ?
3) L'administration n'est pas convaincue de devoir changer. Afin de l'aider à décider, on suppose maintenant qu'on réalise un sondage de taille n et que la fréquence d'élèves préférant se servir eux-mêmes reste identique. Déterminez la taille minimum de l'échantillon qui pourrait davantage convaincre l'administration de changer.

Solution proposée



1)

✓ Si on choisit un élève au hasard dans le lycée, la probabilité que ce soit une fille est :

$$p_1 = 54{,}8\ \% = 0{,}548$$

$(n = 338 \geq 30,\ np_1 \approx 185 \geq 5 \text{ et } n(1-p_1) \approx 153 \geq 5)$

L'intervalle de fluctuation au seuil de 95 % associé est donc :

$$I_1 = \left[0{,}548 - 1{,}96 \times \frac{\sqrt{0{,}548 \times 0{,}452}}{\sqrt{338}} \ ; \ 0{,}548 + 1{,}96 \times \frac{\sqrt{0{,}548 \times 0{,}452}}{\sqrt{338}}\right] \approx [0{,}4949\ ; 0{,}6011\ ]$$

La proportion de filles dans l'échantillon est de : $f_1 = \dfrac{174}{338} \approx 0{,}5148$.

On a donc $f_1 \in I_1$.

✓ Si on choisit un élève au hasard dans le lycée, la probabilité que ce soit un sportif de haut niveau est :

$$p_2 = 11{,}8\ \% = 0{,}118$$

$(n = 338 \geq 30,\ np_2 \approx 39 \geq 5 \text{ et } n(1-p_2) \approx 298 \geq 5)$

L'intervalle de fluctuation au seuil de 95 % associé est donc :

$$I_2 = \left[0{,}118 - 1{,}96 \times \frac{\sqrt{0{,}118 \times 0{,}882}}{\sqrt{338}} \ ; \ 0{,}118 + 1{,}96 \times \frac{\sqrt{0{,}118 \times 0{,}882}}{\sqrt{338}}\right] \approx [0{,}0836\ ; 0{,}1524\ ]$$

La proportion de sportifs de haut niveau l'échantillon est de : $f_2 = 13{,}9\ \% = 0{,}139$

On a donc $f_2 \in I_2$.

- ✓ On peut donc dire que cet échantillon peut être considéré comme représentatif sur les deux critères considérés (sexe et SHN).

2) La fréquence d'élèves qui répondent oui à la question est de :

$$f_3 = \frac{181}{338} \approx 0{,}5355$$

$(n = 338 \geq 25$ et $0{,}2 \leq f_3 \leq 0{,}8)$

L'intervalle de confiance calculé au niveau de confiance 0,95 est donc :

$$I_3 = \left[0{,}5355 - \frac{1}{\sqrt{338}}\ ;\ 0{,}5355 + \frac{1}{\sqrt{338}}\right] \approx [0{,}4811\ ; 0{,}5899\ ]$$

Cet intervalle est une estimation au niveau de confiance 0,95 de la proportion d'élèves demi-pensionnaires préférant se servir eux-mêmes à la cantine.

3) Il suffit que la borne inférieure de l'intervalle de confiance soit supérieure à 50 %, ce qui équivaut à :

$$0{,}5355 - \frac{1}{\sqrt{n}} > 0{,}5$$

$$n > \left(\frac{1}{0{,}0355}\right)^2$$

Il faut donc interroger au moins 794 élèves pour pouvoir convaincre l'administration de changer son mode de fonctionnement, à condition que la fréquence d'élèves préférant se servir eux-mêmes dans ce nouvel échantillon reste identique.

Exercice 3

1) A un concours de recrutement national pour un emploi administratif, se présentent 1 438 femmes et 704 hommes.
   500 personnes sont admises dont 188 hommes.
   Pensez-vous que le jury a respecté la parité dans son mode de recrutement ?
2) L'année suivante, on compte 1356 présentes et 698 présents. Il y a de nouveau 500 reçus dont 341 femmes.
   Pensez-vous que le jury a respecté la parité dans son mode de recrutement ?

Solution proposée

La parité doit être considérée, ici, par rapport à la proportion réelle de femmes qui se présentent au concours et non par rapport à une proportion théorique de 50% d'hommes et 50% de femmes.

1) La proportion de femmes parmi les personnes se présentant au concours est de :

$$p_1 = \frac{1438}{2142} \approx 0{,}671$$

$(n = 500 \geq 30,\ np_1 \approx 336 \geq 5$ et $n(1-p_1) \approx 165 \geq 5)$

L'intervalle de fluctuation au seuil de 95 % associé est donc :

$$I_1 = \left[0{,}671 - 1{,}96 \times \frac{\sqrt{0{,}671 \times 0{,}329}}{\sqrt{500}}\ ;\ 0{,}671 + 1{,}96 \times \frac{\sqrt{0{,}671 \times 0{,}329}}{\sqrt{500}}\right] \approx [0{,}629\ ; 0{,}713\ ]$$

La proportion de femmes parmi les personnes reçues est de :

$$f_1 = \frac{312}{500} = 0{,}624$$

On a donc $f_1 \notin I_1$. On rejette donc l'hypothèse que le jury ait respecté la parité avec un risque d'erreur de 5 %.

Cet exercice, en terminale S, peut être l'occasion de faire varier le seuil afin d'en comprendre le sens.
On peut, ici, faire évoluer le seuil jusqu'à ce que f soit dans I.

On cherche u_α tel que : $u_\alpha \times \frac{\sqrt{0{,}671 \times 0{,}329}}{\sqrt{500}} > 0{,}671 - 0{,}624$ c'est à dire que, pour $u_\alpha = 2{,}237$,
f_1 est dans I mais pour $u_\alpha = 2{,}236$, f_1 n'appartient toujours pas à I.

On a alors si $u_\alpha = 2{,}236$, $P[f_1 \in I] \approx 0{,}975$.
On peut donc rejeter l'hypothèse que le jury ait respecté la parité avec un risque d'erreur inférieur à 3 %.

2) La proportion de femmes parmi les personnes se présentant au concours est de :

$$p_2 = \frac{1356}{2054} \approx 0{,}660$$

$(n = 500 \geq 30,\ np_2 = 330 \geq 5 \text{ et } n(1-p_2) = 170 \geq 5)$
L'intervalle de fluctuation au seuil de 95 % associé est donc :

$$I_2 = \left[0{,}66 - 1{,}96 \times \frac{\sqrt{0{,}66 \times 0{,}34}}{\sqrt{500}}\ ;\ 0{,}66 + 1{,}96 \times \frac{\sqrt{0{,}66 \times 0{,}34}}{\sqrt{500}}\right] \approx [0{,}618\ ; 0{,}702\ ]$$

La proportion de femmes parmi les personnes reçues est de :

$$f_2 = \frac{341}{500} = 0{,}682$$

On a donc $f_2 \in I_2$. On ne peut donc pas rejeter l'hypothèse que le jury ait respecté la parité.

Cet exercice, en terminale S, peut être l'occasion de faire varier le seuil afin d'en comprendre le sens.
On peut, ici, faire évoluer le seuil jusqu'à ce que f ne soit plus dans I.

On cherche u_α tel que : $u_\alpha \times \frac{\sqrt{0{,}66 \times 0{,}34}}{\sqrt{500}} < 0{,}682 - 0{,}66$ c'est à dire $u_\alpha = 1{,}038$, f_2 n'est plus dans I.
On a alors si $u_\alpha = 1{,}038$, $P[f \in I] \approx 0{,}701$.

On ne pourrait donc rejeter l'hypothèse que le jury ait respecté la parité qu'au niveau de confiance 0,701 et on prendrait donc un risque minimum de 29,9 % de se tromper. Il apparait donc cohérent de ne pas rejeter l'hypothèse que le jury ait été impartial.

Exercice 4

Un vol Nice-Paris est assuré par un Airbus de 140 places. La réservation est obligatoire. L'expérience a montré que la probabilité qu'une personne confirme sa réservation et retire son billet est de 0,8. On suppose que les comportements des voyageurs sont indépendants les uns des autres. La compagnie fait du surbooking et se demande combien de réservations elle a intérêt à accepter afin d'avoir 99 % de chances de ne dédommager personne.

Solution proposée

On appelle n le nombre de billets vendus par la compagnie. Soit X_n la variable aléatoire égale au nombre de passagers confirmant et se présentant à l'embarquement. X_n suit alors une loi binomiale de paramètres n et $p = 0{,}8$.
Comme $n \geq 140 \geq 30$, $np \geq 112 \geq 5$ et $n(1-p) \geq 28 \geq 5$, on peut donc utiliser l'intervalle de fluctuation asymptotique au seuil de 99 % $\left(u_{0,01} \approx 2{,}58\right)$:

$$I_n = \left[0{,}8 - 2{,}58 \times \frac{\sqrt{0{,}8 \times 0{,}2}}{\sqrt{n}}\ ;\ 0{,}8 + 2{,}58 \times \frac{\sqrt{0{,}8 \times 0{,}2}}{\sqrt{n}}\right]$$

$$I_n \subset \left[0\,;\frac{140}{n}\right] \implies 0{,}8 + 2{,}58 \times \frac{0{,}4}{\sqrt{n}} \leq \frac{140}{n} \implies 0{,}8n + 1{,}032\sqrt{n} - 140 \leq 0$$

On se ramène à une inéquation du second degré en posant $x = \sqrt{n}$ ou, puisque n est un entier, on se sert de la table de la calculatrice pour trouver le dernier entier $n_0 \geq 140$ pour lequel cette expression soit négative.
On trouve $n_0 = 158$.

Ainsi la compagnie doit vendre au maximum 158 places pour avoir 99 % de chances de ne dédommager personne.

Exercice 5

Sur une chaîne d'embouteillage dans une brasserie, la quantité X (en cL) de liquide fournie par la machine pour remplir chaque bouteille de contenance 110 cL peut être modélisée par une variable aléatoire de loi normale de moyenne μ et d'écart-type $\sigma = 2$.
La législation impose qu'il y ait moins de 0,1 % de bouteilles contenant moins d'un litre.

1) À quelle valeur de la moyenne μ peut-on régler la machine pour respecter cette législation ?
2) Quelle est alors la probabilité qu'une bouteille déborde lors du remplissage ?
3) Le directeur de la coopérative veut qu'il y ait moins de 1 % de bouteilles qui débordent au risque de ne plus suivre la législation.
   Quelle peut être alors la valeur de μ ?
4) Peut-on satisfaire à ces deux conditions ?

Solution proposée

1) Il s'agit de déterminer la valeur de μ telle que $P(X < 100) < 0{,}001$. On détermine d'abord les valeurs z de la loi normale centrée réduite, telle que $P(Z < z) < 0{,}001$. On trouve (logiciels ou calculatrices), $z \leq -3{,}09$.

   Comme $Z = \dfrac{X-\mu}{2}$, il ne reste plus, pour trouver μ, qu'à résoudre $-3{,}09 \geq \dfrac{100-\mu}{2}$.
   On trouve $\mu \geq 106{,}18$.

2) Avec $\mu \approx 106{,}18$, on obtient $P(X > 110) \approx 0{,}028$.

3) Il s'agit cette fois de déterminer la valeur de μ telle que $P(X > 110) < 0{,}01$. On détermine d'abord les valeurs z de la loi normale centrée réduite, telle que $P(Z > z) < 0{,}01$ ou encore $P(Z < z) > 0{,}99$ On trouve (logiciels ou calculatrices), $z \geq 2{,}33$.
   Comme $Z = \dfrac{X-\mu}{2}$, il ne reste plus, pour trouver μ, qu'à résoudre $2{,}33 \leq \dfrac{110-\mu}{2}$.
   On trouve $\mu \leq 105{,}34$.

4) Satisfaire ces deux conditions à la fois signifie d'après les questions précédentes, qu'il faut avoir à la fois $\mu \geq 106{,}18$ et $\mu \leq 105{,}34$, ce qui n'est pas possible.

Exercice 6

La durée de vie d'un certain type d'appareil est modélisée par une variable aléatoire suivant une loi normale de moyenne et d'écart-type inconnus. Les spécifications impliquent que 80 % de la production des appareils ait une durée de vie entre 120 et 200 jours et que 5% de la production ait une durée de vie inférieure à 120 jours.

1) Quelles sont les valeurs de μ et σ ?
2) Quelle est alors la probabilité d'avoir un appareil dont la durée de vie soit comprise entre 200 jours et 230 jours ?

Solution proposée

1) On note X la variable durée de vie. Les spécifications se traduisent par :

$$P(120 \leq X \leq 200) = 0{,}8 \text{ et } P(X < 120) = 0{,}05$$

En notant toujours $Z = \dfrac{X - \mu}{\sigma}$ la variable centrée réduite, on obtient :

$$P\left(\frac{120-\mu}{\sigma} \leq Z \leq \frac{200-\mu}{\sigma}\right) = 0{,}8 \text{ et } P\left(Z < \frac{120-\mu}{\sigma}\right) = 0{,}05$$

Soit $P\left(Z \leq \dfrac{200-\mu}{\sigma}\right) = 0{,}85$ et $P\left(Z < \dfrac{120-\mu}{\sigma}\right) = 0{,}05$

En utilisant un logiciel ou une calculatrice, on obtient : $\mu \approx 120 + 1{,}65\sigma$ et $\mu \approx 200 - 1{,}04\sigma$.
La résolution du système donne $\mu \approx 169$ et $\sigma \approx 30$.

2) $P(200 \leq X \leq 230) \approx 0{,}13$.
La probabilité que la durée de vie de l'appareil soit comprise entre 200 et 230 jours est d'environ 0,13.

Exercice 7

Dans le but d'évaluer la prise en charge de la bronchiolite du nourrisson dans un hôpital de la région Aquitaine, une étude rétrospective a été mise en place.

1) Il est recommandé de coucher l'enfant de manière très inclinée (couchage en proclive) dans le cadre de la prise en charge de la bronchiolite. On évalue cette pratique à partir d'un échantillon de 134 dossiers. 97 des enfants ont été couchés en proclive.
   Déterminer un intervalle de confiance au niveau de confiance de 95 % de la proportion d'enfants dont le couchage respecte la recommandation.

2) Une étude plus fine permet de comparer les pratiques entre les différents services ayant admis des enfants (cf. tableau).

   Tableau : Répartition des cas suivant le type de services et le respect de la recommandation de couchage en proclive ; évaluation de la prise en charge de la bronchiolite en Aquitaine, une année donnée.

Couchage proclive	Service des urgences	Service hospitalier	Total
Oui	45	52	97
Non	29	8	37
Total	74	60	134

   a. Déterminer un intervalle de confiance au seuil de 95 % de la proportion de couchage en proclive pour chaque type de service.
   b. Peut-on conclure selon vous au seuil de 95 % que la pratique de couchage n'est pas identique selon le service ?

Solution proposée

1) Intervalle de confiance au niveau de confiance 95 % ($n = 134 \geq 25$ et $0{,}2 \leq f \leq 0{,}8$):

$$\left[\frac{97}{134} - \frac{1}{\sqrt{134}}\ ;\ \frac{97}{134} + \frac{1}{\sqrt{134}}\right] \approx [0{,}637\ ;\ 0{,}811]$$

   On peut estimer, au niveau de confiance 0,95, que l'intervalle [0,637 ; 0,811] contient la vraie proportion d'enfants dont le couchage respecte la recommandation dans cet hôpital.

2) a. En service des urgences ($n = 74 \geq 25$ et $0{,}2 \leq f \leq 0{,}8$), intervalle de confiance au niveau de confiance 95 % :

$$\left[\frac{45}{74} - \frac{1}{\sqrt{74}}\ ;\ \frac{45}{74} + \frac{1}{\sqrt{74}}\right] \approx [0{,}491\ ;\ 0{,}725]$$

   En service hospitalier ($n = 60 \geq 25$ et $0{,}2 \leq f \leq 0{,}8$), intervalle de confiance au niveau de confiance 95 % :

$$\left[\frac{52}{60} - \frac{1}{\sqrt{60}}\ ;\ \frac{52}{60} + \frac{1}{\sqrt{60}}\right] \approx [0{,}737\ ;\ 0{,}996]$$

   b. Les deux intervalles de confiance sont disjoints, on en conclut que les pratiques diffèrent très certainement entre les deux services.