Échantillonnage - intervalle de fluctuation asymptotique

Ex 1 : ***QCM***

Soit X_n une variable aléatoire suivant une loi binomiale $B(n;p)$, et α un réel tel que $0<\alpha<1$.
Si X est une variable aléatoire suivant la loi normale centrée réduite $N(0,1)$, on appelle u_α l'unique réel tel que : $P(-u_\alpha\leqslant X\leqslant u_\alpha)=1-\alpha$
L'intervalle de fluctuation asymptotique I_n de la fréquence $\frac{X_n}{n}$ au seuil $1-\alpha$ est :

1) $I_n=\left[p-u_\alpha\frac{p(1-p)}{n};p+u_\alpha\frac{p(1-p)}{n}\right]$

2) $I_n=\left[p-u_\alpha\frac{\sqrt{p(1-p)}}{n};p+u_\alpha\frac{\sqrt{p(1-p)}}{n}\right]$

3) $I_n=\left[p-u_\alpha\frac{\sqrt{1-p}}{\sqrt{n}};p+u_\alpha\frac{\sqrt{1-p}}{\sqrt{n}}\right]$

4) $I_n=\left[p-u_\alpha\sqrt{\frac{p(1-p)}{n}};p+u_\alpha\sqrt{\frac{p(1-p)}{n}}\right]$

5) $I_n=\left[p-u_\alpha\frac{\sqrt{p(1-p)}}{\sqrt{n}};p+u_\alpha\frac{\sqrt{p(1-p)}}{\sqrt{n}}\right]$

Ex 2 : ***Prendre une déci ion à partir d'u intervalle de fluctuation***

Un professeur de mathématiques (Monsieur ..X, souhaitant rester anonyme) affirme que 87 % des élèves assistant à ses cours trouvent qu'il n'a aucun sens de l'humour .
Sa femme, inquiétée par l'équilibre psychologique de son mari, souhaite vérifier la véracité de ce qu'il affirme.
À partir d'un échantillon de $n=150$ élèves qu'elle choisit parmi les enfants, petits-enfants … de ses copines, ayant eu ou ayant son mari comme professeur (la population des élèves et anciens élèves est suffisamment grande pour considérer qu'il s'agit de tirages avec remise), sa femme obtient une fréquence F d'élèves trouvant qu'il n'a aucun sens de l'humour.
Elle souhaite savoir pour quelles valeurs de F, elle peut mettre en doute le pourcentage annoncé par son mari au seuil de 8 %.

1) On fait l'hypothèse que le professeur dit vrai et que le taux d'élèves trouvant qu'il n'a aucun sens de l'humour est $p=0,87$. Quelle est la loi suivie par la variable aléatoire X_n comptant le nombre d'élèves trouvant qu'il n'a aucun sens de l'humour dans l'échantillon de taille n ?

2) a) Déterminer l'intervalle de fluctuation asymptotique au seuil de 92 %, I_n , correspondant à la fréquence aléatoire F_n .
b) Énoncer la règle de décision permettant de rejeter ou non l'hypothèse $p=0,87$ selon la valeur de la fréquence F d'élèves trouvant qu'il n'a aucun sens de l'humour.
c) Sur 150 élèves et anciens élèves interrogés, 70 % trouvent qu'il n'a aucun sens de l'humour . Peut-on remettre en cause, au seuil de 8 %, le pourcentage donné par le professeur ?
d) Que peut-on dire si le pourcentage d'élèves trouvant qu'il n'a aucun sens de l'humour est de 70 %, mais sur un échantillon de 50 élèves ?

Ex 3 : ***Prendre une décision à partir d'un intervalle de fluctuation***

Un professeur (souhaitant encore rester anonyme) aimerait connaître la proportion d'élèves aimant les mathématiques parmi les lycéens du Lycée Lyautey.
Pour cela, il sélectionne de manière aléatoire un échantillons de 400 lycéens.
Dans un premier temps, il veut s'assurer que l'échantillon choisi est représentatif des élèves du lycée.
Pour cela, il demande à la direction du lycée, la proportion de garçons parmi les élèves du lycée . Cette proportion est de 51 %.
Dans l'échantillon de $n=400$ lycéens, il y a 196 garçons.

1) La proportion de garçons dans le lycée est $p=0,51$. Soit X_n la variable aléatoire comptant le nombre de garçons dans l'échantillon de taille n . Quelle est la loi suivie par X_n ?

2) a) Déterminer l'intervalle de fluctuation asymptotique au seuil de 98 %, I_n correspondant à la fréquence aléatoire F_n .
b) À quelle condition usuelle doit satisfaire la taille n de l'échantillon aléatoire pour pouvoir utiliser l'approximation $P(F_n\in I_n)\approx 0,98$?

3) Énoncer la règle de décision permettant de rejeter ou non l'hypothèse « l'échantillon choisi représente bien la proportion de garçons du lycée ». Quel est le risque de se tromper en rejetant cette hypothèse ?

4) Que peut-on décider au sujet de l'échantillon choisi ?

Intervalle de confiance

Ex 4 : ***Estimer une proportion au niveau de confiance de 95 %***

Sur la planète mars, en 2253, on considère le cas d'un sondage précédant la sortie d'un nouveau modèle de spountaz.
Un institut effectue un sondage dans la population susceptible d'acheter un spountaz dans les six mois suivant la sortie du nouveau modèle.
Il constitue un échantillon de 1500 martiens, que l'on suppose choisis de manière aléatoire parmi les acheteurs potentiels.
Les résultats sont les suivants :
- 803 ont déclaré vouloir acheter le nouveau modèle présenté.
- 697 ont déclaré ne pas vouloir le faire.

1) Calculer la fréquence des intentions d'achat de l'échantillon.

2) Déterminer l'intervalle de confiance de la fréquence au seuil de 95 %.

3) Six mois après le lancement du nouveau modèle, les ventes réelles représentent 55 % des achats de spountaz. Ces résultats sont-ils compatibles avec les estimations établies à la question 2).

Ex 5 : ***Estimer une proportion au niveau de confiance de 95 %***

Deux cartons contiennent 200 stylos chacun, dont des stylos rouges.
A l'aide d'un tirage avec remise, on extrait un échantillon de 35 stylos dans chacun des cartons et l'on trouve 5 stylos rouges dans le premier carton, et 8 stylos rouges dans le second carton.

1) Calculer, pour chaque carton, la fréquence des stylos rouges dans les échantillons de 35 stylos.

2) Déterminer, pour chaque carton, l'intervalle de confiance de la fréquence au seuil de 95 %.

3) Lorsque l'on ouvre les cartons, on trouve 30 stylos rouges dans le premier carton et 32 dans le second.
Ces données sont-elles compatibles avec les estimations établies à la la question 2).

Ex 6 : ***Estimer l'efficacité d'un produit***

Monsieur ..X, fameux professeur de mathématiques bien connu dans cette fiche d'exercices et inventeur loufoque, a imaginé un détecteur de tricheurs.
Pour s'assurer de l'efficacité de son détecteur de tricheurs, il le teste en approchant le boîtier, 50 fois d'élèves tricheurs, et 50 fois d'élèves non tricheurs.
Il résume dans le tableau suivant le nombre de fois où l'alarme s'est déclenchée.

Alarme déclenchée	Oui	non
Élève tricheur	37	13
Élève non tricheur	7	43

1) Déterminer un intervalle de confiance au seuil de 95 % de la proportion de tests positifs lorsque l'élève triche.

2) Déterminer un intervalle de confiance au seuil de 95 % de la proportion de tests négatifs lorsque l'élève ne triche pas.

3) Interpréter les résultats.

Problèmes divers

Ex 7 : ***Un article douteux***

Un article scientifique, écrit par Monsieur ..X, suggère que l'exposition du fœtus à des intégrales et des dérivées augmente les chances d'être bon en math en TS.
Un étude menée en France et au Maroc auprès de 30000 élèves de TS estime que la proportion d'élèves très bons en math est de 3,1 %.
Pour évaluer l'association entre l'exposition à des intégrales et des dérivées pendant la grossesse et la chance d'être très bons en math en TS, 1335 mères ayant choisi de lire des mathématiques tous les jours pendant leur grossesse ont été choisies de façon aléatoire.
On a observé 45 élèves très bons en math en TS.
En utilisant l'intervalle de fluctuation asymptotique au seuil de 95 %, que peut-on conclure au sujet de cet article ?

Ex 8 : ***D'autres intervalles de fluctuation***

Vérifier que les intervalles $]-\infty;1,31[$ et $[-2,67;1,32]$ peuvent être considérés comme des intervalles de fluctuation au seuil de 90 % d'une variable aléatoire X suivant la loi normale $N(0;1)$.

Ex 9 : ***Avec un tableur***

1) Lors d'un sondage sur un échantillon de 44 femmes utilisant un produit de beauté, 30 se disent satisfaites des résultats obtenus.
a) Estimer la proportion de femmes satisfaites par le produit de beauté.
b) Déterminer l'intervalle de confiance au seuil de 95 %.

2) Avant d'acheter ce même produit, Sophia voudrait mener une enquête approfondie.
Elle se demande donc quelle devrait être la taille minimale de l'échantillon pour pouvoir avoir une précision respectivement de 10 %,5 %,2 %,1 % et 0,5 % sur l'estimation des femmes satisfaites au seuil de 95 %.

a) Calculer la taille minimale de l'échantillon dans les cinq cas , en considérant que la précision de l'estimation est donnée par l'amplitude de l'intervalle de confiance au seuil de 95 %.

b) Représenter la taille minimale de l'échantillon en fonction de la précision.

EN ROUTE VERS LE BAC

Ex 10 : ***Baccalauréat S Antilles-Guyane 9 septembre 2015 - ex 2***

Arbre – probabilités totales – probabilité conditionnelle – loi binomiale – intervalle de fluctuation asymptotique

Dans un supermarché, on réalise une étude sur la vente de bouteilles de jus de fruits sur une période d'un mois.

- 40 % des bouteilles vendues sont des bouteilles de jus d'orange ;
- 25 % des bouteilles de jus d'orange vendues possèdent l'appellation « pur jus ».

Parmi les bouteilles qui ne sont pas de jus d'orange, la proportion des bouteilles de « pur jus » est notée x, où x est un réel de l'intervalle $[0 ; 1]$.
Par ailleurs, 20 % des bouteilles de jus de fruits vendues possèdent l'appellation « pur jus ».
On prélève au hasard une bouteille de jus de fruits passée en caisse. On définit les évènements suivants :
R : la bouteille prélevée est une bouteille de jus d'orange ;
J : la bouteille prélevée est une bouteille de « pur jus ».

Partie A

1. Représenter cette situation à l'aide d'un arbre pondéré.
2. Déterminer la valeur exacte de x.
3. Une bouteille passée en caisse et prélevée au hasard est une bouteille de « pur jus ».
   Calculer la probabilité que ce soit une bouteille de jus d'orange.

Partie B

Afin d'avoir une meilleure connaissance de sa clientèle, le directeur du supermarché fait une étude sur un lot des 500 dernières bouteilles de jus de fruits vendues.
On note X la variable aléatoire égale au nombre de bouteilles de « pur jus » dans ce lot.
On admettra que le stock de bouteilles présentes dans le supermarché est suffisamment important pour que le choix de ces 500 bouteilles puisse être assimilé à un tirage au sort avec remise.

1. Déterminer la loi suivie par la variable aléatoire X. On en donnera les paramètres.
2. Déterminer la probabilité pour qu'au moins 75 bouteilles de cet échantillon de 500 bouteilles soient de « pur jus ». On arrondira le résultat au millième.

Partie C

Un fournisseur assure que 90 % des bouteilles de sa production de pur jus d'orange contiennent moins de 2 % de pulpe. Le service qualité du supermarché prélève un échantillon de 900 bouteilles afin de vérifier cette affirmation. Sur cet échantillon, 766 bouteilles présentent moins de 2 % de pulpe.

1. Déterminer l'intervalle de fluctuation asymptotique de la proportion de bouteilles contenant moins de 2 % de pulpe au seuil de 95 %.
2. Que penser de l'affirmation du fournisseur ?*

Ex 11 : ***Baccalauréat S Nouvelle-Calédonie 17 nov 2014 - ex 1 partie C***

Intervalle de fluctuation asymptotique

Une fabrique de desserts glacés dispose d'une chaîne automatisée pour remplir des cônes de glace.

Une étude réalisée en l'an 2000 a permis de montrer que le pourcentage de Français consommant régulièrement des glaces était de 84 %.
En 2010, sur 900 personnes interrogées, 795 d'entre elles déclarent consommer des glaces.
Peut-on affirmer, au niveau de confiance de 95 % et à partir de l'étude de cet échantillon, que le pourcentage de Français consommant régulièrement des glaces est resté stable entre les années 2000 et 2010 ?

Ex 12 : *Baccalauréat S Amérique du nord 30 mai 2014 - ex 1*

Loi normale – loi binomiale – intervalle de confiance

Dans cet exercice, tous les résultats demandés seront arrondis à 10^{-3} près.

Une grande enseigne de cosmétiques lance une nouvelle crème hydratante.

Partie A : Conditionnement des pots

Cette enseigne souhaite vendre la nouvelle crème sous un conditionnement de 50 mL et dispose pour ceci de pots de contenance maximale 55 mL.
On dit qu'un pot de crème est non conforme s'il contient moins de 49 mL de crème.

1. Plusieurs séries de tests conduisent à modéliser la quantité de crème, exprimée en mL, contenue dans chaque pot par une variable aléatoire X qui suit la loi normale d'espérance $\mu = 50$ et d'écart-type $\sigma = 1,2$.
   Calculer la probabilité qu'un pot de crème soit non conforme.
2. La proportion de pots de crème non conformes est jugée trop importante. En modifiant la viscosité de la crème, on peut changer la valeur de l'écart-type de la variable aléatoire X, sans modifier son espérance $\mu = 50$. On veut réduire à 0,06 la probabilité qu'un pot choisi au hasard soit non conforme.
   On note σ' le nouvel écart-type, et Z la variable aléatoire égale à $\dfrac{X-50}{\sigma'}$
   a. Préciser la loi que suit la variable aléatoire Z.
   b. Déterminer une valeur approchée du réel u tel que $p(Z \leqslant u) = 0,06$.
   c. En déduire la valeur attendue de σ'.
3. Une boutique commande à son fournisseur 50 pots de cette nouvelle crème. On considère que le travail sur la viscosité de la crème a permis d'atteindre l'objectif fixé et donc que la proportion de pots non conformes dans l'échantillon est 0,06.
   Soit Y la variable aléatoire égale au nombre de pots non conformes parmi les 50 pots reçus.
   a. On admet que Y suit une loi binomiale. En donner les paramètres.
   b. Calculer la probabilité que la boutique reçoive deux pots non conformes ou moins de deux pots non conformes.

Partie B : Campagne publicitaire

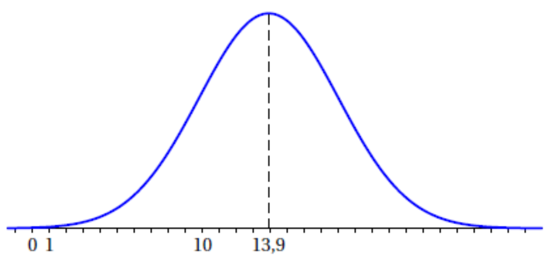
Une association de consommateurs décide d'estimer la proportion de personnes satisfaites par l'utilisation de cette crème.
Elle réalise un sondage parmi les personnes utilisant ce produit. Sur 140 personnes interrogées, 99 se déclarent satisfaites.
Estimer, par intervalle de confiance au seuil de 95 %, la proportion de personnes satisfaites parmi les utilisateurs de la crème.

Ex 13 : *Baccalauréat S Pondichéry Avril 2016 - ex 1*

Loi normale – $P(\mu-2\sigma \leqslant T \leqslant \mu+2\sigma)$ – probabilités totales – intervalle de confiance

Partie A

Des études statistiques ont permis de modéliser le temps hebdomadaire, en heures, de connexion à internet des jeunes en France âgés de 16 à 24 ans par une variable aléatoire T suivant une loi normale de moyenne $\mu = 13,9$ et d'écart type σ.
La fonction densité de probabilité de T est représentée ci-dessous :



1. On sait que $p(T \geqslant 22) = 0,023$.
   En exploitant cette information :
   a. hachurer sur le graphique donné en annexe, deux domaines distincts dont l'aire est égale à 0,023 ;
   b. déterminer $P(5,8 \leqslant T \leqslant 22)$. Justifier le résultat. Montrer qu'une valeur approchée de σ au dixième est 4,1.
2. On choisit un jeune en France au hasard.
   Déterminer la probabilité qu'il soit connecté à internet plus de 18 heures par semaine.
   Arrondir au centième.

Partie B

Dans cette partie, les valeurs seront arrondies au millième.
La Hadopi (Haute Autorité pour la diffusion des Œuvres et la Protection des droits sur Internet) souhaite connaître la proportion en France de jeunes âgés de 16 à 24 ans pratiquant au moins une fois par semaine le téléchargement illégal sur internet. Pour cela, elle envisage de réaliser un sondage.
Mais la Hadopi craint que les jeunes interrogés ne répondent pas tous de façon sincère. Aussi, elle propose le protocole ($\mathscr{P}$) suivant :

On choisit aléatoirement un échantillon de jeunes âgés de 16 à 24 ans.
Pour chaque jeune de cet échantillon :

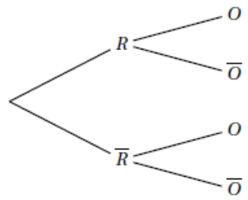
- le jeune lance un dé équilibré à 6 faces ; l'enquêteur ne connaît pas le résultat du lancer ;
- l'enquêteur pose la question : « Effectuez-vous un téléchargement illégal au moins une fois par semaine ? » ;

■ si le résultat du lancer est pair alors le jeune doit répondre à la question par « Oui » ou « Non » de façon sincère ;
■ si le résultat du lancer est « 1 » alors le jeune doit répondre « Oui » ;
■ si le résultat du lancer est « 3 ou 5 » alors le jeune doit répondre « Non ».

Grâce à ce protocole, l'enquêteur ne sait jamais si la réponse donnée porte sur la question posée ou résulte du lancer de dé, ce qui encourage les réponses sincères.

On note p la proportion inconnue de jeunes âgés de 16 à 24 ans qui pratiquent au moins une fois par semaine le téléchargement illégal sur internet.

1. *Calculs de probabilités*
   On choisit aléatoirement un jeune faisant parti du protocole ($\mathscr{P}$).
   On note : R l'évènement « le résultat du lancer est pair »,
   O l'évènement « le jeune a répondu Oui ».
   Reproduire et compléter l'arbre pondéré ci-dessous :



   En déduire que la probabilité q de l'évènement « le jeune a répondu Oui » est :
   $$q = \frac{1}{2}p + \frac{1}{6}.$$
2. *Intervalle de confiance*
   a. À la demande de l'Hadopi, un institut de sondage réalise une enquête selon le protocole ($\mathscr{P}$). Sur un échantillon de taille 1 500, il dénombre 625 réponses « Oui ».
   Donner un intervalle de confiance, au niveau de confiance de 95 %, de la proportion q de jeunes qui répondent « Oui » à un tel sondage, parmi la population des jeunes français âgés de 16 à 24 ans.
   b. Que peut-on en conclure sur la proportion p de jeunes qui pratiquent au moins une fois par semaine le téléchargement illégal sur internet ?