Échantillonnage et estimation – Exercices

Loi binomiale, révision

1 Une classe compte 30 élèves dont 20 filles. À chaque cours de physique, le professeur interroge au hasard un élève sans se rappeler quels élèves il a précédemment interrogés.

1. Quelle est la probabilité que sur 12 cours consécutifs, exactement 4 filles soient interrogées ?

2. Quelle est la probabilité que sur 12 cours consécutifs au plus 5 garçons soient interrogés ?

2 Dans une fabrication d'objets en série, 8 % de ces objets présentent un défaut. Un carton contient 15 objets qui présentent ou non le défaut indépendamment les uns des autres. Calculer à 10^{-3} près la probabilité que dans un carton :

a. les dix objets soient sans défaut ;

b. 5 objets soient sans défaut ;

c. au moins 5 objets soient sans défaut.

Variable aléatoire fréquence

3 Soit X une variable aléatoire suivant la loi binomiale de paramètres $n = 100$ et $p = 0{,}5$.

1. Calculer $P(X = 40)$ à 10^{-4} près.

2. En déduire la probabilité de l'événement $(F = 0{,}4)$ où F désigne la variable aléatoire fréquence associée à la variable aléatoire X.

4 On considère une population présentant un caractère présent avec une proportion $p = 0{,}35$. Soit F la variable aléatoire fréquence associé aux échantillons de taille $n = 40$.

1. Calculer $P(F \geq 0{,}5)$ à 10^{-4} près.

2. Calculer $P(0{,}025 \leq F \leq 0{,}45)$ à 10^{-4} près.

3. Calculer $P(0{,}36 \leq F \leq 0{,}71)$ à 10^{-4} près.

Échantillonnage

5 Soit X une variable aléatoire qui suit une loi binomiale de paramètres 320 et 0,16.

En utilisant les résultats ci-dessous, déterminer l'intervalle de fluctuation au seuil de 95 % d'une fréquence correspondant à la réalisation de X sur un échantillon aléatoire de taille 320.

$P(X \leq 37) \approx 0{,}015$	$P(X \leq 63) \approx 0{,}967$
$P(X \leq 38) \approx 0{,}023$	$P(X \leq 64) \approx 0{,}976$
$P(X \leq 39) \approx 0{,}034$	$P(X \leq 65) \approx 0{,}983$

6 On donne la proportion p d'un caractère dans une population et la taille n d'un échantillon. Déterminer un intervalle de fluctuation asymptotique au seuil indiqué dans chacun des cas suivants. Attention à la validité des conditions sur n et p.

a. $n = 30$ et $p = 0{,}5$, au seuil 95 %

b. $n = 100$ et $p = 0{,}44$, au seuil 90 %

c. $n = 100$ et $p = 0{,}02$, au seuil 95 %

d. $n = 10000$ et $p = 0{,}36$, au seuil 99 %

e. $n = 5000$ et $p = 0{,}0003$, au seuil de 90 %

7 Il y a 23 % d'élèves boursiers dans les établissements d'enseignement secondaire en France. Déterminer un intervalle de fluctuation asymptotique à 95 % de la fréquence des boursiers dans les lycées de 1200 élèves. Arrondir à 10^{-3}.

8 22 % des français sont d'accord pour supprimer les panneaux indiquant la présence de radars sur les routes. Déterminer un intervalle de fluctuation asymptotique au seuil de confiance 95 % de la proportion de personnes désirant supprimer les panneaux dans un échantillon de 2000 personnes.

Prise de décision

9 Une étude américaine, semble-t-il sérieuse, a émis l'hypothèse que la proportion de personnes non réceptive à l'hypnose est de 22 %. Lors d'une émission télévisée retransmise en direct, 40 spectateurs présents sur le plateau ont été choisis au hasard par un hypnotiseur pour participer à son numéro. Malgré de nombreuses tentatives de ce dernier, 13 spectateurs restent insensibles et donc non réceptifs à cette séance d'hypnose en groupe.

1. Préciser la fréquence observée de personnes non réceptives à l'hypnose lors de ce numéro.

2. Déterminer l'intervalle de fluctuation au seuil de 95 % de la fréquence observée de personnes non réceptives à l'hypnose dans un groupe de 40 personnes choisies au hasard.

3. À partir de cette émission télévisée, peut-on dire que l'hypothèse faite par cette étude n'est pas réaliste ?

10 Une entreprise fabrique des composants électroniques en grande quantité. Une étude interne affirme que la probabilité qu'un composant électronique choisi au hasard dans cette production soit défectueux est égale à 2 %.

1. Déterminer l'intervalle de fluctuation au seuil de 95 % de la fréquence de composants électroniques défectueux sur des échantillons de taille 1000.

2. Un client a acheté 1000 pièces parmi lesquelles 23 étaient défectueuses. Peut-il remettre en cause l'enquête interne ?

3. Même question pour l'achat de 10 000 pièces parmi lesquelles 230 étaient défectueuses.

11 En 1976, dans un comté du Texas, Rodrigo Partida était condamné à huit ans de prison. Il attaqua ce jugement au motif que la désignation des jurés de ce comté était discriminante à l'écart des américains d'origine mexicaine. En effet, 79 % de la population de ce comté est d'origine mexicaine, et sur les 870 personnes convoquées pour être jurés lors d'une certain période de référence, il n'y eut que 339 personnes d'origine mexicaine. Peut-on dire que la constitution du jury a été faite de façon aléatoire ?

12 *(2013, Amérique du Nord).* Une boulangerie industrielle utilise une machine pour fabriquer des pains de campagne pesant en moyenne 400 grammes. Pour être vendus aux clients, ces pains doivent peser au moins 385 grammes. Un pain dont la masse est strictement inférieure à 385 grammes est un pain non-commercialisable, un pain dont la masse est supérieure ou égale à 385 grammes est commercialisable.

La masse d'un pain fabriqué par la machine peut être modélisée par une variable aléatoire X suivant la loi normale d'espérance $\mu = 400$ et d'écart-type $\sigma = 11$.
Les probabilités seront arrondies au millième le plus proche.

Partie A – On pourra utiliser le tableau suivant dans lequel les valeurs sont arrondies au millième le plus proche.

x	380	385	390	395	400	405	410	415	420
$P(X \leq x)$	0,035	0,086	0,182	0,325	0,5	0,675	0,818	0,914	0,965

1. Calculer $P(390 \leq X \leq 410)$.
2. Calculer la probabilité p qu'un pain choisi au hasard dans la production soit commercialisable.
3. Le fabricant trouve cette probabilité p trop faible. Il décide de modifier ses méthodes de production afin de faire varier la valeur de σ sans modifier celle de μ.
   Pour quelle valeur de σ la probabilité qu'un pain soit commercialisable est-elle égale à 96 %? On arrondira le résultat au dixième.
   On pourra utiliser le résultat suivant : lorsque Z est une variable aléatoire qui suit la loi normale d'espérance 0 et d'écart-type 1, on a $P(Z \leq -1{,}751) \approx 0{,}040$.

Partie B – Les méthodes de production ont été modifiées dans le but d'obtenir 96 % de pains commercialisables.
Afin d'évaluer l'efficacité de ces modifications, on effectue un contrôle qualité sur un échantillon de 300 pains fabriqués.

1. Déterminer l'intervalle de fluctuation asymptotique au seuil de 95 % de la proportion de pains commercialisables dans un échantillon de taille 300.
2. Parmi les 300 pains de l'échantillon, 283 sont commercialisables.
   Au regard de l'intervalle de fluctuation obtenu à la question 1, peut-on décider que l'objectif a été atteint ?

Partie C – Le boulanger utilise une balance électronique. Le temps de fonctionnement sans dérèglement, en jours, de cette balance électronique est une variable aléatoire T qui suit une loi exponentielle de paramètre λ.

1. On sait que la probabilité que la balance électronique ne se dérègle pas avant 30 jours est de 0,913. En déduire la valeur de λ arrondie au millième.
   Dans toute la suite on prendra $\lambda = 0{,}003$.
2. Quelle est la probabilité que la balance électronique fonctionne encore sans dérèglement après 90 jours, sachant qu'elle a fonctionné sans dérèglement 60 jours ?
3. Le vendeur de cette balance électronique a assuré au boulanger qu'il y avait une chance sur deux pour que la balance ne se dérègle pas avant un an. A-t-il raison ? Sinon, pour combien de jours est-ce vrai ?

13 *(2014, Liban).* Une fabrique de desserts glacés dispose d'une chaîne automatisée pour remplir des cônes de glace.

Partie A – Les cônes de glace sont emballés individuellement puis conditionnés en lots de 2 000 pour la vente en gros.
On considère que la probabilité qu'un cône présente un défaut quelconque avant son conditionnement en gros est égale à 0,003.
On nomme X la variable aléatoire qui, à chaque lot de 2 000 cônes prélevés au hasard dans la production, associe le nombre de cônes défectueux présents dans ce lot.
On suppose que la production est suffisamment importante pour que les tirages puissent être supposés indépendants les uns des autres.

1. Quelle est la loi suivie par X ? Justifier la réponse et préciser les paramètres de cette loi.
2. Si un client reçoit un lot contenant au moins 12 cônes défectueux, l'entreprise procède alors à un échange de celui-ci.
   Déterminer la probabilité qu'un lot ne soit pas échangé ; le résultat sera arrondi au millième.

Partie B – Chaque cône est rempli avec de la glace à la vanille. On désigne par Y la variable aléatoire qui, à chaque cône, associe la masse (exprimée en grammes) de crème glacée qu'il contient.
On suppose que Y suit une loi normale $\mathcal{N}(110; \sigma^2)$ d'espérance $\mu = 110$ et d'écart-type σ.
Une glace est considérée comme commercialisable lorsque la masse de crème glacée qu'elle contient appartient à l'intervalle $[104; 116]$.
Déterminer une valeur approchée à 10^{-1} près du paramètre σ telle que la probabilité de l'évènement « la glace est commercialisable » soit égale à 0,98.

Partie C – Une étude réalisée en l'an 2000 a permis de démontrer 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 ?

14 *(2015, Antilles-Guyane).* 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 ?

15 *(2015, Amérique du Sud).* Les trois parties A, B et C peuvent être traitées de façon indépendante.

Partie A – Le chikungunya est une maladie virale transmise d'un être humain à l'autre par les piqûres de moustiques femelles infectées.
Un test a été mis au point pour le dépistage de ce virus. Le laboratoire fabriquant ce test fournit les caractéristiques suivantes :

- la probabilité qu'une personne atteinte par le virus ait un test positif est de 0,98 ;
- la probabilité qu'une personne non atteinte par le virus ait un test positif est de 0,01.

On procède à un test de dépistage systématique dans une population « cible ». Un individu est choisi au hasard dans cette population. On définit les événements :

- M : « l'individu choisi est atteint du chikungunya » ;
- T : « le test de l'individu choisi est positif ».

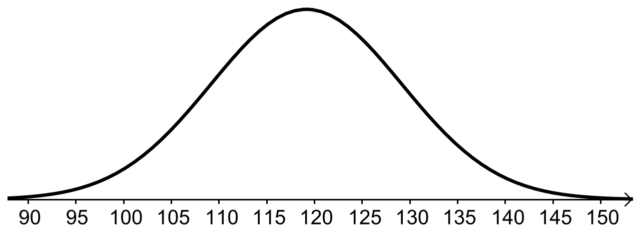
On note p $(0 \leq p \leq 1)$ la proportion de personnes atteintes par la maladie dans la population cible.

1. **a.** Illustrer la situation par un arbre pondéré.
   b. Exprimer $P(M \cap T)$ et $P(\overline{M} \cap T)$ puis $P(T)$ en fonction de p.
2. **a.** Démontrer que la probabilité de M sachant T est donnée par la fonction f définie sur $[0;1]$ par :
   $$f(p) = \frac{98p}{97p+1}.$$
   b. Étudier les variations de la fonction f.
3. On considère que le test est fiable lorsque la probabilité qu'une personne ayant un test positif soit réellement atteinte du chikungunya est supérieure à 0,95.
   En utilisant les résultats de la question **2.**, à partir de quelle proportion p de malades dans la population le test est-il fiable ?

Partie B – En juillet 2014, l'institut de veille sanitaire d'une île, en s'appuyant sur les données remontées par les médecins, publie que 15 % de la population est atteinte par le virus.
Comme certaines personnes ne consultent pas forcément leur médecin, on pense que la proportion est en réalité plus importante.
Pour s'en assurer, on se propose d'étudier un échantillon de 1000 personnes choisies au hasard dans cette île. La population est suffisamment importante pour considérer qu'un tel échantillon résulte de tirages avec remise.
On désigne par X la variable aléatoire qui, à tout échantillon de 1000 personnes choisies au hasard, fait correspondre le nombre de personnes atteintes par le virus et par F la variable aléatoire donnant la fréquence associée.

1. **a.** Sous l'hypothèse $p = 0{,}15$, déterminer la loi de X.
   b. Dans un échantillon de 1000 personnes choisies au hasard dans l'île, on dénombre 197 personnes atteintes par le virus.
   Quelle conclusion peut-on tirer de cette observation à propos du chiffre de 15 % publié par l'institut de veille sanitaire ?
   Justifier. (On pourra s'aider du calcul d'un intervalle de fluctuation au seuil de 95 %).
2. On considère désormais que la valeur de p est inconnue. En utilisant l'échantillon de la question **1.b.**, proposer un intervalle de confiance de la valeur de p, au niveau de confiance de 95 %.

Partie C – Le temps d'incubation, exprimé en heures, du virus peut être modélisé par une variable aléatoire T suivant une loi normale d'écart type $\sigma = 10$.
On souhaite déterminer sa moyenne μ.
La représentation graphique de la fonction densité de probabilité de T est donnée ci-dessous.



1. **a.** Conjecturer une valeur approchée de μ.
   b. On donne $P(T < 110) = 0{,}18$. Hachurer sur le graphique un domaine dont l'aire correspond à la probabilité donnée.
2. On note T' la variable aléatoire égale à $\frac{T-\mu}{10}$.
   a. Quelle loi la variable aléatoire T' suit-elle ?
   b. Déterminer une valeur approchée à l'unité près de la moyenne μ de la variable aléatoire T et vérifier la conjecture de la question 1.

Estimation et intervalles de confiance

16 En lançant 200 fois un dé, on obtient 28 fois le numéro 1. Donner un intervalle de confiance pour la proportion du nombre de 1, au niveau de confiance de 95 %.

17 Une semaine avant une élection un sondage est effectué sur 1024 personnes choisis au hasard parmi les 42821 inscrites sur les listes ; 532 déclarent voter pour le candidat A. Le candidat A a-t-il raison de penser qu'il va être élu ?

18 En interrogeant 100 personnes sur un sujet pour lequel il faut répondre « oui » ou « non » on obtient 51 « oui » et 49 « non ».

1. Donner les intervalles de confiance pour les proportions de « oui » et « non » au niveau de confiance 95 %.
2. En supposant que les proportions restent inchangées, déterminer le nombre minimal de personnes qu'il faut interroger pour que les intervalles de confiance des deux intervalles ne se recouvrent pas.

19 Une banque désire savoir si son site est bien adapté au besoin de plus de 5 millions de clients. Elle commande à un institut de sondage une enquête afin d'estimer la proportion de ses clients satisfaits. Elle impose un niveau de confiance de 95 % avec une amplitude de 0,04.
Combien de personnes doit au minimum interroger l'institut de sondage ?

20 *(2014, Métropole, Bac ES).*

Partie A – Chaque jour, Antoine s'entraine au billard américain pendant une durée comprise entre 20 minutes et une heure. On modélise la durée de son entrainement, en minutes, par une variable aléatoire X qui suit la loi uniforme sur l'intervalle $[20; 60]$.

1. Calculer la probabilité p pour que l'entrainement dure plus de 30 minutes.
2. Calculer l'espérance de X. Interpréter ce résultat.

Partie B – *Dans cette partie les probabilités seront, si besoin, arrondies au millième.*
Les boules de billard américain avec lesquelles Antoine s'entraine sont dites de premier choix si leur diamètre est compris entre 56,75 mm et 57,25 mm ; sinon elles sont dites de second choix.
On note D la variable aléatoire qui, à chaque boule prélevée au hasard dans la production de l'entreprise, associe son diamètre, en millimètres.
On suppose que D suit la loi normale d'espérance 57 et d'écart-type 0,11.

1. Déterminer la probabilité p_1 que la boule prélevée ait un diamètre inférieur à 57 mm.
2. Déterminer la probabilité p_2 que la boule prélevée soit une boule de premier choix.
3. En déduire la probabilité p_3 que la boule prélevée soit une boule de second choix.

Partie C – Le président de la fédération française de billard (FFB) souhaite estimer le niveau de satisfaction de ses 14 000 licenciés quant à l'organisation des tournois.
Antoine estime que les 80 adhérents de son club constituent un échantillon représentatif des licenciés de la FFB. Il est chargé de faire une étude au sein de son club : les 80 adhérents ont répondu, et 66 ont déclaré qu'ils étaient satisfaits.

1. Quelle est, sur cet échantillon, la fréquence observée f de personnes satisfaites de la FFB ?
2. Déterminer un intervalle de confiance au niveau de confiance 0,95 de la proportion p de licenciés satisfaits de la FFB. Les bornes de l'intervalle seront arrondies au millième.

21 *(2014, Liban, Bac ES). QCM avec une seule réponse exacte.*
Un fumeur est dit fumeur régulier s'il fume au moins une cigarette par jour.
En 2010, en France, la proportion notée p de fumeurs réguliers, âgés de 15 à 19 ans, était de 0,236. On a $p = 0,236$.

1. La probabilité que, sur un groupe de 10 jeunes âgés de 15 à 19 ans choisis au hasard et de manière indépendante, aucun ne soit fumeur régulier est, à 10^{-3} près :
   a. 0,136 **b.** 0 **c.** 0,068 **d.** 0,764
2. Un intervalle de fluctuation asymptotique au seuil de 0,95 de la fréquence de fumeurs réguliers dans un échantillon de 500 jeunes âgés de 15 à 19 ans est : *(les bornes de chaque intervalle sont données à 10^{-3} près)*
   a. $[0,198; 0,274]$ **b.** $[0,134; 0,238]$
   c. $[0,191; 0,281]$ **d.** $[0,192; 0,280]$
3. La taille n de l'échantillon choisi afin que l'amplitude de l'intervalle de fluctuation au seuil de 0,95 soit inférieure à 0,01, vaut :
   a. 200 **b.** 400 **c.** 21 167 **d.** 27 707
4. Dans un échantillon de 250 jeunes fumeurs réguliers, âgés de 15 à 19 ans, 99 sont des filles.
   Au seuil de 95 %, un intervalle de confiance de la proportion de filles parmi les fumeurs réguliers âgés de 15 à 19 ans est : *(les bornes de chaque intervalle sont données à 10^{-2} près)*
   a. $[0,35; 0,45]$ **b.** $[0,33; 0,46]$
   c. $[0,39; 0,40]$ **d.** $[0,30; 0,50]$

22 *(2014, Amérique du Nord).* 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 $\frac{X-50}{\sigma'}$.
   a. Préciser la loi que suit la variable aléatoire Z.
   b. Déterminer une valeur approchée du réel u tel que $P(Z \leq u) = 0,06$.
   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.