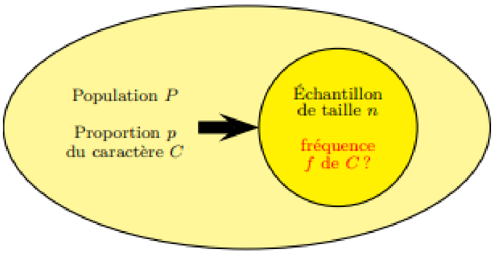
Chapitre P3 : Echantillonnage & Estimation

1 – Echantillonnage

a. Introduction



● Soit une population, dont la **proportion** p d'un caractère est connue.

Ex : Dans la population française, il y a 13 % de gauchers.

● On prélève un échantillon de **taille** n dans cette population et on s'intéresse à la **fréquence** f du caractère dans cet échantillon.

Ex : Votre classe est un échantillon de taille $n =$ …… et la fréquence de gauchers est $f =$ ……………………………

● Si l'on évalue la fréquence f de ce caractère dans plusieurs échantillons de même taille n, on obtiendra des résultats légèrement différents. Ce phénomène s'appelle **fluctuation d'échantillonnage.**

Ex : Dans une autre classe, la fréquence de gauchers pourra être différente.

● Plus la taille n des échantillons est grande, moins les variations de f seront importantes : La fréquence f observée dans les échantillons se rapproche de la proportion p.

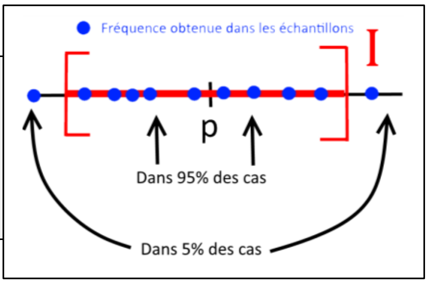
Ex : Résultats obtenus sur 10 classes de 40 élèves puis sur 10 amphithéâtres de 250 élèves :

Classe	1	2	3	4	5	6	7	8	9	10
Gauchers	10,00%	32,50%	10,00%	7,50%	10,00%	15,00%	10,00%	2,50%	7,50%	12,50%

Amphi	1	2	3	4	5	6	7	8	9	10
Gauchers	13,60%	11,60%	10,80%	18,40%	12,00%	11,20%	13,40%	10,40%	17,50%	14,40%

● Afin de rendre compte des fluctuations naturelles de la fréquence observée dans les différents échantillons on utilise un intervalle :

Définition 1 : On dit que I est **intervalle de fluctuation au seuil de 95%**, s'il contient la fréquence f observée dans les échantillons de taille n, dans au moins 95% de cas.



Propriété 1 : Si $n \geq 30$ alors $I = \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]$ est un intervalle de fluctuation au seuil de 95%.

Exemple 1 : Calculons l'intervalle de fluctuation de la fréquence de gauchers dans des échantillons de taille :

● $n = 40 : I = \left[0.13 - \frac{1}{\sqrt{40}} ; 0.13 + \frac{1}{\sqrt{40}}\right] = [0.13 - 0.158 \,; 0.13 + 0.158] = [0; 0.288]$

● $n = 250 : I = \left[0.13 - \frac{1}{\sqrt{250}} ; 0.13 + \frac{1}{\sqrt{250}}\right] = [0.13 - 0.063 \,; 0.13 + 0.063] = [0.067; 0.193]$

Remarque : Plus n est grand, plus l'intervalle sera précis.

b. Intervalle de fluctuation asymptotique

Propriété 2 : Si $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$ alors l'intervalle $I = \left[p - 1.96\sqrt{\frac{p(1-p)}{n}}\ ;\ p + 1.96\sqrt{\frac{p(1-p)}{n}}\right]$ est un intervalle de fluctuation (dit asymptotique) au seuil de 95 % de la fréquence F.

Démonstration : D'où vient cette formule ?

● Soit une population, dont la proportion d'un caractère est p. On prélève un échantillon de taille n.

● On considère X le nombre d'individu présentant ce caractère dans cet échantillon. X suit la loi binomiale : $S =$ « L'individu choisi présente le caractère » et $P(S) = p$. Comme la population est d'une taille importante par rapport à celle de l'échantillon on peut considérer que les n individus sont choisis de manière indépendante. On a donc $X \sim B(n; p)$.

● La fréquence du caractère dans l'échantillon est donc donnée par la variable $F = \frac{X}{n}$

● Or nous avons vu que si n est pris suffisamment grand (tel que $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$) alors on peut approximer la loi de X par la loi normale $\mathcal{N}(\mu; \sigma)$ de paramètres $\mu = np$ et $\sigma = \sqrt{np(1-p)}$.

● De plus, on sait que $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 0.95$ (à 10^{-2} près).

En fait, nous avons même $P(\mu - 1.96\sigma \leq X \leq \mu + 1.96\sigma) \approx 0.95$ (à 10^{-5} près).

● En divisant l'inégalité $\mu - 1.96\sigma \leq X \leq \mu + 1.96\sigma$ par « n », et en remplaçant μ et σ, on obtient :

$$P\left(\frac{np - 1.96\sqrt{np(1-p)}}{n} \leq \frac{X}{n} \leq \frac{np + 1.96\sqrt{np(1-p)}}{n}\right) \approx 0.95 \text{ d'où } P\left(p - 1.96\sqrt{\frac{p(1-p)}{n}} \leq F \leq p + 1.96\sqrt{\frac{p(1-p)}{n}}\right) \approx 0.95.$$

Exemple 1 : Calculons l'intervalle de fluctuation asymptotique de la fréquence de gauchers dans des échantillons de taille $n = 250$

$$I = \left[0.13 - 1.96\sqrt{\frac{0.13(1-0.13)}{250}}\ ;\ 0.13 + 1.96\sqrt{\frac{0.13(1-0.13)}{250}}\right] = [0.088; 0.172]$$

Dans 95% des cas, la fréquence de gauchers dans des échantillons de taille 250 est entre 8.8% et 17.2%.

Remarque : L'intervalle de fluctuation asymptotique est plus précis que l'intervalle de la propriété 1.

c. Règle de décision

Soit une population dans laquelle on suppose l'hypothèse : H : « La proportion d'un caractère est p ».

On prélève un échantillon de taille n et on note f la fréquence observée.

Règle de décision : Si I est un intervalle de fluctuation au seuil de 95 %. Alors :

● Si $f \in I$, alors on accepte l'hypothèse avec un seuil de confiance de 95 %

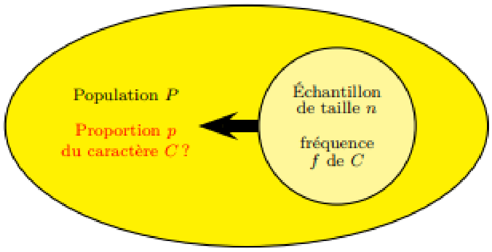
● Si $f \notin I$, alors on rejette l'hypothèse avec un risque de 5 %

Exemple 3 : On suppose l'hypothèse que le pourcentage de gauchers dans le monde est aussi de 13%. On prélève un échantillon de 250 personnes dans le monde et on dénombre 28 gauchers. Doit-on accepter ou rejeter l'hypothèse ?

$f = \frac{28}{250} = 0.112 \in [0.088; 0.172]$. On accepte donc l'hypothèse avec un seuil de confiance de 95%.

2 – Estimation

a. Intervalle de confiance



● Soit une population, dont la **proportion** p d'un caractère est inconnue.

Ex : Avant le vote (2016), quelle est la proportion de britannique qui vont voter pour le BREXIT ?

● On prélève un échantillon de **taille** n dans cette population et on s'intéresse à la **fréquence** f du caractère dans cet échantillon.

Ex : On effectue un sondage sur 1000 britanniques choisis au hasard dans la population parmi lesquels 512 se sont prononcés en faveur du BREXIT. Dans ce cas on a $n = 1000$ et $f = 0.512$.

● Problématique : Peut –on en conclure que le BREXIT aura lieu ?

Pour répondre à cette question, on utilise un intervalle dit de confiance.

Définition 1 : On dit que I est **intervalle de confiance au seuil de 95%**, si la proportion p appartient à cet intervalle avec une probabilité de 95%.

Propriété 3 : Si $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$ alors l'intervalle $I = \left[f - 1.96\sqrt{\frac{f(1-f)}{n}} \; ; \; f + 1.96\sqrt{\frac{f(1-f)}{n}}\right]$ est un intervalle de confiance de la proportion p.

Remarque : Il s'agit de la même formule que pour l'intervalle de fluctuation mais avec f au lieu de p.

Exemple 3 : Calculons l'intervalle de confiance de la proportion de britanniques qui voteront pour le BREXIT :

$$I = \left[0.512 - 1.96\sqrt{\frac{0.512(1-0.512)}{1000}} \; ; \; 0.512 + 1.96\sqrt{\frac{0.512(1-0.512)}{1000}}\right] = [0.481; 0.543]$$

Avec un seuil de confiance à 95%, on peut dire la proportion de britanniques qui se prononcera pour le BREXIT sera comprise entre 48% et 54% (si les intentions de votes restent les mêmes).

b. Comparaison de deux proportions.

On souhaite comparer deux proportions p_1 et p_2 du même caractère dans deux populations différentes :

● On prélève un échantillon dans la première population dont on mesure la fréquence f_1, puis on en déduit l'intervalle de confiance I_1 pour p_1

● On prélève un échantillon dans la deuxième population dont on mesure la fréquence f_2, puis on en déduit l'intervalle de confiance I_2 pour p_2

Règle de décision : On compare les deux intervalles I_1 et I_2

● Si I_1 et I_2 sont disjoints, on considère que la différence entre les deux proportions p_1 et p_2 est significative.

● Sinon on considère que les deux proportions sont égales avec un seuil de confiance de 95 %.

Exemple 3 : Un autre sondage par région a été réalisé. En Angleterre, sur 1500 personnes interrogées 788 Sont en faveur du BREXIT alors qu'en Ecosse seulement 590 des 1500 sondés sont favorables au BREXIT.

● Angleterre : $f_1 = \frac{788}{1500} \approx 0.525$; $I_1 = \left[0.525 - 1.96\sqrt{\frac{0.525(1-0.525)}{1500}}\ ;\ 0.525 + 1.96\sqrt{\frac{0.525(1-0.525)}{1500}}\right]$

On obtient, $I_1 = [0.500; 0.550]$

● Ecosse : $f_2 = \frac{590}{1500} \approx 0.393$; $I_1 = \left[0.393\ - 1.96\sqrt{\frac{0.393\ (1-0.393\ )}{1500}}\ ;\ 0.393\ + 1.96\sqrt{\frac{0.393\ (1-0.393\ )}{0.393}}\right]$

On obtient $I_2 = [0.368; 0.418]$

● I_1 et I_2 sont disjoints, on peut en conclure que les deux proportions sont significativement différentes.

Exercice 1 (Bac STI2D Nouvelle Calédonie 2014)

Un groupe agricole vend des sachets de graines donnant des plantes résistantes aux maladies. Le directeur de ce groupe affirme que 92 % des sachets sont efficaces et donnent des plantes résistantes. *Dans cet exercice, les valeurs approchées seront arrondies à 10^{-2} près.*

1. On prélève au hasard un échantillon de 100 sachets.
   (a) Déterminer l'intervalle de fluctuation asymptotique à 95 % de la fréquence de sachets efficaces sur un échantillon de taille 100.
   (b) Dans le prélèvement de 100 sachets, 88 donnent des plantes résistantes. Peut-on rejeter l'hypothèse du directeur ?
2. On considère la variable aléatoire X qui, à tout prélèvement de 100 sachets, associe le nombre de sachets donnant des plantes résistantes. On admet que la variable aléatoire X suit la loi binomiale de paramètres $n = 100$ et $p = 0,92$.
   (a) Déterminer l'espérance et l'écart type de X (arrondi à $0,01$ près).
   (b) La variable aléatoire X peut être approchée par la variable aléatoire Y qui suit la loi normale d'espérance 92 et d'écart type $2,7$.
   En utilisant la variable aléatoire Y, calculer la probabilité que le nombre de sachets donnant des plantes résistantes soit compris entre 89 et 94.

Exercice 2 (Bac STI2D Métropole 2013)

Une entreprise fabrique en grande série des barres de pâte d'amande. La masse annoncée sur leur emballage est de 125 grammes. La machine qui fabrique les barres de pâte d'amande est préréglée afin que ces dernières respectent la masse de 125 grammes avec une certaine tolérance. Une barre de pâte d'amande est dite conforme lorsque sa masse est comprise dans un intervalle de tolérance de [124 ;127,5].

1. On désigne par X la variable aléatoire qui, à une barre de pâte d'amande prélevée au hasard dans la production, associe sa masse en grammes. Le service qualité estime que la variable aléatoire X suit la loi normale $\mathcal{N}(125,5; 0,75)$.
   (a) Calculer la probabilité qu'une barre de pâte d'amande prélevée au hasard ait une masse supérieure à 125,5 grammes.
   (b) Calculer la probabilité qu'une barre prélevée au hasard soit conforme.
   (c) En déduire la probabilité qu'une barre prélevée au hasard soit non conforme.

a	$p(X \leqslant a)$
122,00	0,0000015
122,25	0,0000073
122,50	0,0000317
122,75	0,0001229
123,00	0,0004291
123,25	0,0013499

a	$p(X \leqslant a)$
123,50	0,0038304
123,75	0,0098153
124,00	0,0227501
124,25	0,0477904
124,50	0,0912112
124,75	0,1586553

a	$p(X \leqslant a)$
125,00	0,2524925
125,25	0,3694413
125,50	0,5000000
125,75	0,6305587
126,00	0,7475075
126,25	0,8413447

a	$p(X \leqslant a)$
126,50	0,9087888
126,75	0,9522096
127,00	0,9772499
127,25	0,9901847
127,50	0,9961696
127,75	0,9986501

2. Lors d'un contrôle, le responsable qualité prélève de façon aléatoire un échantillon de 300 barres de pâte d'amande dans la production et constate que 280 barres sont conformes. On admet que, lorsque la machine est correctement réglée, la proportion de barres de pâte d'amande conformes dans l'ensemble de la production est de 97 %.
   On souhaite savoir si le réglage de la machine peut être jugé satisfaisant.
   (a) Donner l'intervalle de fluctuation asymptotique à 95 % de la fréquence des barres de pâte d'amande de masse conforme obtenue sur un échantillon de taille 300 (les bornes de l'intervalle seront arrondis à 10^{-3} près).
   (b) Le résultat obtenu lors du contrôle qualité remet-il en question le réglage de la machine ?

Exercice 3 (Bac STI2D Polynésie 2013)

Les trois parties sont indépendantes, les résultats seront arrondis, si nécessaire, à 10^{-3} près.

Une entreprise produit en grande quantité des pièces détachées destinées à l'industrie. L'objectif de cet exercice est d'étudier l'exploitation de divers outils mathématiques pour analyser la qualité de cette production.

A. Loi normale

Une pièce est conforme lorsque sa longueur, exprimée en millimètres, appartient à l'intervalle [74,4 ; 75,6]. On note L la variable aléatoire qui, à chaque pièce prélevée au hasard dans la production, associe sa longueur. On suppose que la variable aléatoire L suit la loi normale d'espérance 75 et d'écart type $0,25$.

1. Calculer $P(74,4 \leqslant L \leqslant 75,6)$.
2. Quelle valeur doit-on donner à h pour avoir $P(75 - h \leqslant L \leqslant 75 + h) = 0,95$?

B. Loi binomiale

Les pièces produites par l'entreprise sont livrées par lots de 20. On note D l'événement : « une pièce prélevée au hasard dans la production n'est pas conforme ». On suppose que $P(D) = 0,02$. On prélève au hasard 20 pièces dans la production. La production est assez importante pour que l'on puisse assimiler ce prélèvement à un tirage aléatoire avec remise. On considère la variable aléatoire X qui, à un lot de 20 pièces, associe le nombre de pièces non conformes qu'il contient.

1. Justifier que la variable X suit la loi binomiale de paramètres 20 et $0,02$.
2. Calculer la probabilité $P(X = 0)$.
3. Calculer la probabilité qu'il y ait au moins une pièce non conforme dans ce lot de 20 pièces.
4. Calculer l'espérance mathématiques, $E(X)$, de cette variable aléatoire et interpréter le résultat.

C. Intervalle de fluctuation

Le cahier des charges établit que la proportion de 2 % de pièces non conformes dans la production est acceptable.

1. Donner l'intervalle de fluctuation asymptotique à 95 % de la fréquence des pièces non conformes dans un échantillon de taille 80.

On veut savoir si la machine est correctement réglée. Pour cela on prélève au hasard dans la production un échantillon de taille 80 dans lequel 3 pièces sont non conformes.

2. Quelle est la fréquence des pièces non conformes dans l'échantillon prélevé ?
3. La machine de production doit-elle être révisée ? Justifier votre réponse.

Exercice 1 **(11 p 288, TSTI2D, hachette éducation)**

Dans un échantillon de taille 1020, 51 individus présentent le caractère C.
On note p la proportion d'individus présentant ce caractère dans la population entière. On suppose que p est compris entre 0,005 et 0,995.
Déterminer l'intervalle de confiance de p au niveau 95%.

Exercice 2 **(19 p289, TSTI2D, hachette éducation)**

Une entreprise commercialisant des téléphones portables souhaite estimer la proportion p de ses clients satisfaits de leur acquisition.
Elle réalise un sondage : les personnes interrogées sont choisies au hasard parmi la clientèle (suffisamment importante pour que ce choix soit assimilé à un tirage avec remise). L'entreprise an nonce qu'au niveau de confiance 95%, le taux de satisfaction est 0,93.

1. (a) Soit n le nombre de personnes interrogées lors du sondage. On suppose que p est compris entre $\dfrac{5}{n}$ et $1-\dfrac{5}{n}$. Déterminer la longueur L_n de l'intervalle de confiance de p au niveau de confiance 95% en fonction de n.
   (b) Résoudre l'inéquation $L_n \leq 0,06$.
   (c) En déduire le nombre minimal de personnes interrogées si ce taux de satisfaction de 88% donne, au niveau de confiance 95%, une estimation de la proportion p à $0,06$ près.
2. Combien de personnes a-t-on interrogées si, au niveau de confiance 95%, ce taux de satisfaction de 88% est une estimation de p à 0,01 près ?

Exercice 3 **(24 p 291, TSTI2D, hachette éducation)**

Deux entreprises concurrentes A et B commercialisent de la crème au caramel.

— Un sondage effectué sur 200 personnes choisies au hasard parmi les consommateurs révèle que 56% des personnes interrogées préfèrent le goût de la crème de marque A.
— Un autre sondage, effectué sur 200 personnes choisies au hasard parmi les consommateurs, révèle que 68% d'entre elles préfèrent celui de la crème de marque B.

On suppose que le choix des 200 personnes peut être assimilé à un choix avec remise.
On appelle p (respectivement p') la proportion de personnes préférant la me de marque A (respectivement la crème de marque B). On suppose que p et p ' sont comprises entre 0,025 et 0,975.

1. Déterminer l'intervalle de confiance au niveau 95% de chacune des proportions p et p' suite à ce sondage (bornes de l'intervalle arrondies à 10^{-4} près).
2. Peut-on, au niveau de confiance 95%, dire qu'une des deux marques a la préférence des consommateurs ?
3. Si on avait obtenu les mêmes fréquences avec deux sondages effectués sur des échantillons de 500 per sonnes, aurait-on conclu de la même façon ?

Exercices avec prise d'initiatives

Exercice 4 **(40 p 298, TSTI2D, hachette éducation)**

Dans tout cet exercice, la production est supposée suffisamment importante pour que l'on assimile le choix d'un échantillon à un tirage avec remise. Un sous-traitant est chargé de concevoir des pièces pour un constructeur automobile.

1. Un sondage est réalisé pour tester la qualité de la production du sous-traitant. On suppose que la proportion de pièces non conformes dans la production est comprise entre 0,02 et 0,98 ; sur un échantillon de 250 pièces, 12 ne répondent pas au cahier des charges.
   (a) Déterminer l'intervalle de confiance au niveau confiance 95% obtenu à partir de cet échantillon (bornes de l'intervalle arrondies à 10^{-3} près).
   (b) Le client veut un taux de pièces non conformes inférieur où égal à 6%. Compte tenu du sondage effectué, peut-on affirmer avec un niveau de confiance de 95% que cette condition est respectée ?
2. Le sous-traitant procède à de nouveaux réglages et fait un nouveau sondage sur 250 pièces.
   Déterminer, à l'aide de la calculatrice, le nombre maximum de pièces non conformes dans cet échantillon pour pouvoir affirmer que, au niveau de confiance de 95%, le taux de pièces non conformes est inférieur ou égal à 6% dans la production. Expliquer votre démarche.

Exercice 5 **(41 p 298, TSTI2D, hachette éducation)**

Deux sociétés pharmaceutiques, Mégavax et Antimal, veulent commercialiser un médicament. Elles se sont toutes les deux livrées à des études statistiques :

— Mégavax a testé son médicament sur un échantillon de 1000 malades et le traitement s'est révélé efficace pour 810 d'entre eux.
— Antimal a testé son médicament sur un échantillon de 500 malades et le traitement s'est révélé efficace pour 420 d'entre eux.

Chacune des deux entreprises prétend avoir le meilleur médicament, Antimal arguant d'un taux d'efficacité plus élevé, Mégavax expliquant que les taux sont voisins, mais que son étude a été menée sur davantage de malades et est donc plus probante.
Qu'en pensez-vous ?

Echantillonnage & Estimation – Fiche d'exercices

Ex 1 QCM (Tiré de divers sujet bac)

Dans le cadre du fonctionnement correct d'une chaîne de production de pièces détachées, la proportion de pièces détachées conformes doit être 96 %.

On contrôle la production de la chaîne en prélevant de manière aléatoire un échantillon de 150 pièces détachées.

On rappelle que l'intervalle de fluctuation asymptotique à 95 % de la fréquence p des pièces détachées conformes sur un échantillon de taille n est :

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

En utilisant un intervalle de fluctuation asymptotique à 95 %, on prendra la décision d'effectuer des réglages sur la chaîne de production si le nombre de pièces détachées non conformes trouvées dans l'échantillon prélevé est :

a. 8 **b.** 9 **c.** 10 **d.** 11

Ex 2 Paquets de sucre (Tiré du bac Polynésie 2015)

Une entreprise achète du sucre et le revend après conditionnement à des grossistes pour le marché de la grande distribution.

Les résultats seront arrondis à 10^{-3} près.

On arrondit à 0.08 la probabilité p pour qu'un paquet conditionné dans l'usine soit refusé. Ainsi $p = 0.8$

On contrôle la masse d'un échantillon de 100 paquets de sucre dans le stock global de l'entreprise. Après contrôle, 10 paquets sont refusés.

Rappel : Lorsque la proportion p dans la population est connue, l'intervalle de fluctuation asymptotique à 95 % d'une fréquence obtenue sur un échantillon de taille n est :

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

L'échantillon est-il représentatif de la production de l'usine ? Justifier.

Ex 3 Jus de fruits (Tiré du bac Métropole 2015)

Dans l'ensemble de l'exercice, les résultats seront arrondis à 10^{-4} près.

L'usine OCEFRAIS embouteille des jus de fruits. L'étiquette de la bouteille indique 1,5 litre de jus de fruits. Le volume de la bouteille est de 1,55 litre.

Une bouteille est dite conforme si elle contient entre 1,46 litre et 1,54 litre de jus de fruits.

Selon l'usine OCEFRAIS, la probabilité qu'une bouteille soit non conforme est 0,0077.

Un supermarché achète un lot de 10 000 bouteilles.

a. Déterminer l'intervalle de fluctuation asymptotique à 95 % de la fréquence observée de bouteilles non conformes dans un tel lot.

b. Dans le lot de 10 000 bouteilles, on a compté 90 bouteilles non conformes. Le gérant du supermarché trouve le nombre de bouteilles non conformes anormalement élevé.
L'usine OCEFRAIS a-t-elle des raisons de s'inquiéter ?

Ex 4 Batterie au Lithium (Tiré du bac Nouvelle Caledonie 2016)

Une usine fabrique des batteries au lithium-ion pour des vélos électriques. Le cahier des charges indique qu'une batterie mesure 15 cm de large.

Dans le cadre d'un fonctionnement correct des machines de la chaîne de production, on admet que la proportion p de batteries non conformes est 1,2 %.

Le responsable de l'usine affirme qu'il ne vend pas de lot de 2 000 batteries qui en contienne plus de 40 non conformes. Quelle est la fiabilité de cette affirmation ?

Justifier.

Ex 5 Plats préparés sous vide (Tiré du bac Antilles-Guyanne 2015)

L'entreprise BUENPLATO produit en grande quantité des plats préparés sous vide. Le fabricant annonce sur les étiquettes de ses produits une proportion de produits non conformes de 12 %. On prélève au hasard dans la production un échantillon de taille 1200 dans lequel 150 plats se révèlent être non conformes.

a. Quelle est la fréquence de plats non conformes dans l'échantillon prélevé ?

b. Déterminer l'intervalle de fluctuation avec un niveau de confiance de 95 % de la fréquence de plats non conformes dans un échantillon de taille 1200.

Rappel : Lorsque la proportion p dans la population est connue, l'intervalle de fluctuation asymptotique à 95 % d'une fréquence obtenue sur un échantillon de taille n est :

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

c. L'échantillon est-il représentatif de la production du fabricant ? Justifier.

Ex 6 Ascenseurs (Tiré du bac Polynésie 2018)

Une entreprise assure la maintenance d'un parc de 75 ascenseurs qui fonctionnent de façon indépendante.

On considère dans cette partie que la probabilité qu'un ascenseur du parc tombe en panne un jour donné est 0,08 .

Depuis quelque temps, l'entreprise constate de nombreuses pannes parmi les 75 ascenseurs. Ainsi, sur une période de 30 jours, il a été relevé 263 pannes en tout.

L'entreprise doit-elle remettre en cause, au seuil de 95 %, le modèle selon lequel la probabilité qu'un ascenseur tombe en panne un jour donné est 0,08 ? Justifier la réponse.

Ex 7 Poulet (Tiré du bac Polynésie 2018)

Dans un élevage de poulets fermiers, les volailles sont commercialisées après 90 jours d'élevage. Un poulet de 90 jours sera dit conforme si sa masse est comprise entre 2,8 kg et 3,2 kg.

1. L'avicultrice a constaté que la masse M, exprimée en kg, de ses poulets de 90 jours suit une loi normale de moyenne 3 et d'écart type 0,1.
   a. Déterminer au centième près la probabilité qu'un poulet de 90 jours prélevé au hasard soit conforme.
   b. Déterminer au millième près la probabilité que la masse d'un poulet de 90 jours prélevé au hasard soit supérieure à 3,3 kg.
2. On admet dans cette question que 95 % des poulets de 90 jours sont conformes.
   Un rôtisseur achète tous les samedis 100 de ces poulets. On admet que le nombre de poulets de l'élevage est suffisamment important pour que cet achat puisse être assimilé à un prélèvement avec remise.
   On appelle X la variable aléatoire égale au nombre de poulets non conformes, c'est-à-dire dont la masse n'est pas dans l'intervalle $[2,8 ; 3,2]$.
   a. Justifier que X suit une loi binomiale. Préciser ses paramètres.
   b. Calculer l'espérance mathématique de X. Que représente ce nombre?
3. Lors de son dernier achat, le rôtisseur a compté 9 poulets non conformes. Il se plaint auprès de l'éleveur.
   Avec un tableur, on a calculé les probabilités $P(X \leqslant a)$ pour a allant de 0 à 13.

a	0	1	2	3	4	5	6
$P(X \leqslant a)$	0,005 9	0,037 1	0,118 3	0,257 8	0,436 0	0,616 0	0,766 0

a	7	8	9	10	11	12	13
$P(X \leqslant a)$	0,872 0	0,936 9	0,971 8	0,988 5	0,995 7	0,998 5	0,999 5

   a. Déterminer l'intervalle de fluctuation à 95 % de la fréquence de poulets non conformes.
   b. Le rôtisseur a-t-il eu raison de se plaindre?

Ex 8 Diabète (Tiré du bac Polynésie 2017)

Les deux parties de cet exercice peuvent être traitées de manière indépendante.
Les résultats seront arrondis, si nécessaire, à 10^{-3} près.

En 2016, l'Organisation Mondiale de la Santé (OMS) affirme que 5,1 millions de personnes en France souffraient de diabète, soit 8% de la population.
Chaque personne dispose d'un dossier médical régulièrement actualisé.

Partie A

Dans le cadre de la semaine nationale de prévention du diabète qui s'est tenue en 2016, une campagne de sensibilisation de cette maladie a été menée.
Sur 85 dossiers médicaux prélevés au hasard, on a compté 3 cas de diabète.

1. Quelle est la fréquence de cas de diabète dans l'échantillon prélevé ?
2. Déterminer l'intervalle de fluctuation avec un niveau de confiance de 95% de la fréquence de cas de diabète sur cet échantillon de 85 dossiers.
   Rappel : Lorsque la proportion p dans la population est connue, l'intervalle de fluctuation asymptotique à 95% d'une fréquence obtenue sur un échantillon de taille n est :
   $$I = \left[p - 1,96\sqrt{\frac{p(1-p)}{n}} ; p + 1,96\sqrt{\frac{p(1-p)}{n}}\right]$$
3. L'échantillon est-il représentatif de la population française ? Justifier.

Partie B

Dans le corps humain, la régulation du taux de glycémie est assurée grâce à un équilibre permanent entre différentes substances principalement hormonales.
Le tableau suivant présente trois états de la glycémie :

Hypoglycémie	À jeun : inférieur à 0,70 g/l
Glycémie normale	À jeun : entre 0,70 g/l et 1,10 g/l
Hyperglycémie	À jeun : supérieur à 1,10 g/l

On note N la variable aléatoire qui, à chaque dossier médical prélevé au hasard dans la population, associe le taux de glycémie à jeun en g/l de la personne.
On suppose que N suit la loi normale de moyenne 0,9 et d'écart type 0,1.
Dans le cadre de cet exercice, on considère qu'une personne souffre de diabète si cette personne ne présente pas une glycémie normale à jeun.

1. Déterminer la probabilité pour que le dossier prélevé soit celui d'une personne en hypoglycémie.
2. Déterminer la probabilité pour que le dossier prélevé soit celui d'une personne en hyperglycémie.
3. Déterminer la probabilité que le dossier prélevé soit celui d'une personne souffrant de diabète.

Ex 9 Thyroïde (Tiré du bac Antilles-Guyanne 2017)

En 2012, l'Agence Nationale de Sécurité du Médicament (ANSM) s'est inquiétée de la forte augmentation des ventes du médicament qui traite l'hypothyroïdie. Pour obtenir un état des lieux de l'utilisation de ce médicament en France, l'ANSM a effectué un sondage sur 530 877 personnes.
Dans cet échantillon, 21 771 personnes ont déclaré qu'elles utilisaient ce médicament.

1. Quelle est la fréquence des utilisateurs du médicament dans l'échantillon étudié ?
2. Déterminer un intervalle de confiance avec un niveau de confiance de 95% de la proportion d'utilisateurs de ce médicament dans la population française.

Rappel : Lorsqu'une fréquence f est mesurée dans un échantillon de taille n, l'intervalle de confiance à 95 % de la proportion dans la population est donné par :

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

Ex 10 Pneumatiques (Tiré du bac Antilles-Guyanne 2016)

Un manufacturier de pneumatiques produit des pneus d'avions en grande quantité.

Il s'engage à livrer des produits spécifiques aux avionneurs de masse maximum garantie de 124 kg. Ces pneus doivent supporter une charge nominale de 10 tonnes, des vitesses pouvant aller jusqu'à 420 km.h^{-1} et des températures instables allant de -40 °C (en altitude) à 250 °C (au moment du décollage).

Un pneu trop lourd entraîne une augmentation de la consommation du kérosène. Lorsque la masse d'un pneu reçu par une compagnie aérienne dépasse 121,9 kg cela entraîne des pénalités financières pour le manufacturier.

Sur la chaîne de fabrication, on prélève de façon aléatoire un échantillon de 36 pneus et on constate que 2 d'entre eux ont une masse qui dépasse 121,9 kg.

1. Quelle est la fréquence des pneus dans l'échantillon prélevé dont la masse dépasse 121,9 kg ?
2. Déterminer l'intervalle de confiance avec un niveau de confiance de 95% de la proportion de pneus dont la masse dépasse 121,9 kg dans la production.

On rappelle que lorsqu'une fréquence f est mesurée dans un échantillon de taille n, l'intervalle de confiance à 95% de la proportion dans la population est donné par :

$$I=\left[f-1{,}96\sqrt{\frac{f(1-f)}{n}}\,;f+1{,}96\sqrt{\frac{f(1-f)}{n}}\right].$$

3. Donner une interprétation du résultat précédent.

Ex 11 Porte blindée (Tiré du bac Métropole 2018)

Un industriel commercialise des portes blindées. Il projette de lancer un nouveau modèle de portes blindées : les portes « SECUR ». Équipées d'un digicode et d'une caméra, elles seront donc plus sécurisées que celles déjà existantes sur le marché.

Les résultats seront arrondis à 10^{-4} près.

L'industriel envisage de commercialiser les portes blindées de modèle « SECUR » au tarif M déterminé précédemment. Il souhaite estimer la proportion de personnes susceptibles d'acheter son nouveau modèle. Une enquête est réalisée sur un échantillon de 984 personnes intéressées par l'achat d'une porte blindée. Sur cet échantillon, 123 personnes se disent favorables à l'achat du modèle « SECUR ».

1. Déterminer l'intervalle de confiance, au niveau de confiance 95%, de la proportion de personnes favorables à l'achat du nouveau modèle.

On rappelle que pour une fréquence f observée dans un échantillon de taille n, l'intervalle de confiance au niveau de confiance 95% de la proportion p du caractère étudié dans la population est donné par : $\left[f-1{,}96\sqrt{\frac{f(1-f)}{n}}\,;f+1{,}96\sqrt{\frac{f(1-f)}{n}}\right]$.

2. Pour que l'industriel prenne le risque d'investir dans les portes « SECUR », il faudrait qu'au minimum 20% des personnes souhaitant s'équiper d'une porte blindée soient favorables à ce nouveau modèle. A-t-il intérêt à réaliser son projet ?