I FLUCTUATION D'ÉCHANTILLONNAGE

1 INTERVALLE DE FLUCTUATION AU SEUIL DE 95%

On s'intéresse à un caractère de proportion p connue au sein d'une population.
On considère la variable aléatoire F_n qui à chaque échantillon aléatoire de taille n associe la fréquence du caractère étudié.

DÉFINITION

On appelle intervalle de fluctuation de F_n au seuil de 95%, tout intervalle $[\alpha;\beta]$ tel que la probabilité $P\left(F_n \in [\alpha;\beta]\right) \geqslant 0{,}95$

EXEMPLE

En première partie de soirée une série a attiré près de 6,2 millions de téléspectateurs soit 34 % de part d'audience.
Déterminons un intervalle de fluctuation de la part d'audience de cette série pour un échantillon de taille 100.
Soit X la variable aléatoire qui correspond au nombre de téléspectateurs qui ont regardé cette série dans un échantillon de 100 personnes ayant regardé la télévision en première partie de soirée.
Le nombre de téléspectateurs en première partie de soirée est suffisamment important pour considérer que la variable X suit la loi binomiale de paramètres $n = 100$ et $p = 0{,}34$.
Le plus petit entier a tel que $P(X \leqslant a) > 0{,}025$ est 25 et, le plus petit entier b tel que $P(X \leqslant b) \geqslant 0{,}975$ est 43.
Un intervalle de fluctuation à 95% de la fréquence des téléspectateurs qui ont regardé cette série dans un échantillon de taille 100 est :

$$I = \left[\frac{25}{100};\frac{43}{100}\right] \text{ soit } I = [0{,}25;0{,}43]$$

2 INTERVALLE DE FLUCTUATION ASYMPTOTIQUE AU SEUIL DE 95%

On appelle intervalle de fluctuation asymptotique au seuil de 95% de la variable aléatoire F_n, l'intervalle :

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

INTERPRÉTATION

L'intervalle I_n contient la fréquence F_n avec une probabilité proche de 0,95 pourvu que n soit suffisamment grand.
En pratique, on utilise l'intervalle de fluctuation asymptotique au seuil 0,95 dès que :

$$n \geqslant 30\,,\ np \geqslant 5 \text{ et } n(1-p) \geqslant 5.$$

EXEMPLE

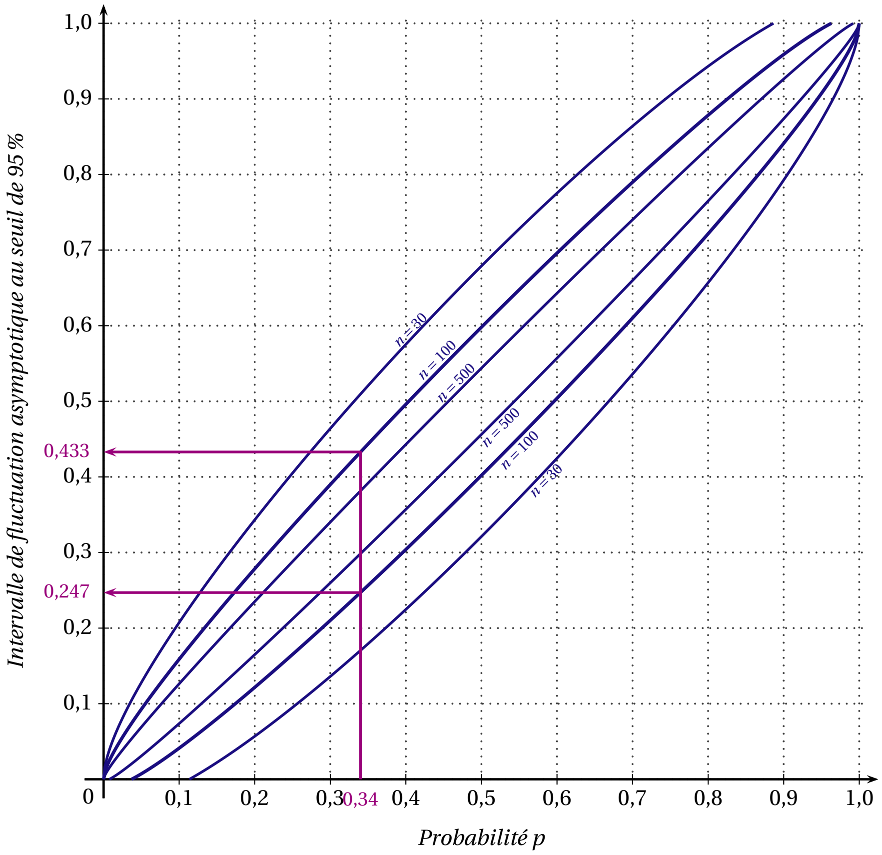
Avec $p = 0{,}34$ et $n = 100$ on a $np = 34$ et $n(1-p) = 66$, les critères d'approximation sont vérifiés.
L'intervalle de fluctuation asymptotique à 0,95 sur un échantillon de taille 100 est :

$$I_{100} = \left[0{,}34 - 1{,}96 \times \sqrt{\frac{0{,}34 \times 0{,}66}{100}}; 0{,}34 + 1{,}96 \times \sqrt{\frac{0{,}34 \times 0{,}66}{100}}\right]$$

Soit avec des valeurs approchées à 10^{-3} près des bornes de l'intervalle, $I_{100} \approx [0{,}247;0{,}433]$.

REMARQUE

On a tracé ci-dessous, sur l'intervalle $]0;1[$, les courbes représentatives des fonctions $f_{\text{inf}} = p - 1{,}96 \times \sqrt{\dfrac{p(1-p)}{n}}$ et $f_{\text{sup}} = p + 1{,}96 \times \sqrt{\dfrac{p(1-p)}{n}}$ associées aux bornes des intervalles de fluctuation asymptotique au seuil de 95 % pour les valeurs de $n = 30$, $n = 100$ et $n = 500$.



3 DÉCISION À PARTIR DE LA FRÉQUENCE D'UN ÉCHANTILLON

Quand les critères d'approximation sont vérifiés, l'intervalle de fluctuation asymptotique I_n permet de déterminer des seuils de décision :

— pour accepter ou rejeter l'hypothèse selon laquelle p est la proportion d'un caractère dans la population;

— pour déterminer si un échantillon issu de la population est représentatif.

On formule l'hypothèse que la proportion d'un caractère dans la population est p.
On prélève dans la population un échantillon de taille n et on note f la fréquence observée du caractère étudié.
Lorsque $n \geqslant 30$, $np \geqslant 5$ et $n(1-p) \geqslant 5$ on pose :

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

— Si la fréquence observée f n'appartient pas à l'intervalle I_n, alors on rejette l'hypothèse selon laquelle p est la proportion du caractère étudié dans la population avec un risque d'erreur de 5 %.

— Si la fréquence observée f appartient à l'intervalle I_n, alors l'hypothèse selon laquelle p est la proportion du caractère étudié dans la population est acceptée.

EXEMPLE

Dans un forum on a constaté que 28 personnes sur 100 ont regardé la série dont la part d'audience a été estimée à 34 %. Ce résultat remet-il en question l'estimation de la part d'audience de la série ?
La fréquence observée de la part d'audience dans l'échantillon de taille 100 est : $f = \dfrac{28}{100} = 0{,}28$.
L'intervalle de fluctuation asymptotique au seuil de 95 % de la part d'audience de la série dans les échantillons de taille 100 est $I_{100} = [0{,}247;0{,}433]$.
Comme $0{,}28 \in [0{,}247;0{,}433]$, l'estimation d'une part d'audience de 34 % pour la série n'est pas remise en cause.

II INTERVALLE DE CONFIANCE

On cherche à estimer avec un certain niveau de confiance, la proportion p **inconnue** d'un caractère au sein d'une population à partir d'un échantillon de taille n.

1 DÉFINITION

Soit f la fréquence observée d'un caractère dans un échantillon de taille n.
Sous les conditions usuelles d'approximation $n \geqslant 30$, $np \geqslant 5$ et $n(1-p) \geqslant 5$, l'intervalle $\left[f - \dfrac{1}{\sqrt{n}}; f + \dfrac{1}{\sqrt{n}}\right]$ est un intervalle de confiance au niveau de confiance 0,95 de la proportion inconnue p dans la population.

REMARQUES

- En pratique, les conditions de validité de la formule peuvent être vérifiées à posteriori.
- La précision de l'intervalle de confiance est donnée par son amplitude $\dfrac{2}{\sqrt{n}}$. Plus la taille de l'échantillon est grande, plus les intervalles de confiance obtenus sont précis.
- La différence entre deux fréquences f_1 et f_2 observées sur deux échantillons est considérée comme significative quand les intervalles de confiance correspondants sont disjoints.
  Dans ce cas, on considère que les deux proportions p_1 et p_2 sont différentes. Dans le cas contraire, on ne peut pas conclure.

EXEMPLE

On interroge au hasard 100 clients ayant effectué des achats à la sortie d'une grande surface. Le temps d'attente aux caisses a été jugé raisonnable par 52 personnes interrogées.

1. Peut-on considérer que plus de 50% des clients de cette grande surface estiment que le temps d'attente aux caisses est raisonnable ?
   Soit $f = \dfrac{52}{100} = 0{,}52$ la fréquence des clients qui estiment que le temps d'attente aux caisses est raisonnable.
   Les bornes de l'intervalle de confiance au niveau de confiance 95% de la proportion des clients qui estiment que le temps d'attente aux caisses est raisonnable sont :
   $$f - \frac{1}{\sqrt{n}} = 0{,}52 - \frac{1}{\sqrt{100}} = 0{,}42 \text{ et } f + \frac{1}{\sqrt{n}} = 0{,}52 + \frac{1}{\sqrt{100}} = 0{,}62$$
   On a : $n = 100$, $0{,}42 \leqslant p \leqslant 0{,}62$, $100 \times 0{,}42 \leqslant np \leqslant 100 \times 0{,}62$ et $100 \times (1-0{,}62) \leqslant n\left(1-p\right) \leqslant 100 \times (1-0{,}42)$.
   Soit $n \geqslant 30$, $42 \leqslant np \leqslant 62$ et $38 \leqslant n\left(1-p\right) \leqslant 58$. Les conditions d'approximation d'un intervalle de confiance au niveau de confiance 95 % sont vérifiées.
   Un intervalle de confiance au niveau de confiance 95 % de la proportion de clients qui estiment que le temps d'attente aux caisses est raisonnable est $[0{,}42;0{,}62]$.

La borne inférieure de l'intervalle de confiance est 0,42, il est donc possible que moins de 50% des clients trouvent que le temps d'attente aux caisses est raisonnable.

2. Déterminer le nombre minimal de clients qu'il faut interroger pour estimer la proportion p de clients qui trouvent le temps d'attente aux caisses raisonnable avec une précision inférieure à 0,1.

   La précision de l'estimation de p est $\dfrac{2}{\sqrt{n}}$. Pour tout entier naturel n :

$$\begin{aligned}\frac{2}{\sqrt{n}} < 0{,}1 &\iff \frac{1}{\sqrt{n}} < 0{,}05\\ &\iff \sqrt{n} > \frac{1}{0{,}05}\\ &\iff \sqrt{n} > 20\\ &\iff n > 400\end{aligned}$$

   Il faut interroger plus de 400 clients pour obtenir une estimation de la proportion p de clients qui trouvent le temps d'attente aux caisses raisonnable avec une précision inférieure à 0,1.

3. À fréquence observée égale à 0,52, quel nombre de clients aurait-il fallu interroger pour estimer que plus de 50% des clients trouvent que le temps d'attente aux caisses est raisonnable ?

   La borne inférieure de l'intervalle de confiance au niveau de confiance 0,95 sur un échantillon de taille n est $0{,}52 - \dfrac{1}{\sqrt{n}}$ d'où n est solution de l'inéquation :

$$\begin{aligned}0{,}52 - \frac{1}{\sqrt{n}} \geqslant 0{,}5 &\iff -\frac{1}{\sqrt{n}} \geqslant -0{,}02\\ &\iff \frac{1}{\sqrt{n}} \leqslant 0{,}02\\ &\iff \sqrt{n} \geqslant \frac{1}{0{,}02}\\ &\iff \sqrt{n} \geqslant 50\\ &\iff n \geqslant 2500\end{aligned}$$

Avec une fréquence observée égale à 0,52, il faudrait un échantillon de taille supérieure à 2500 pour que la proportion p de clients qui trouvent le temps d'attente aux caisses raisonnable appartienne à un intervalle de confiance dont la borne inférieure est supérieure à 0,5.

EXERCICE 1

Une entreprise produit en grande quantité des emballages alimentaires de forme cubique. Elle utilise pour cela la technique du thermoformage, qui consiste à chauffer une plaque de plastique puis à la former à l'aide d'un moule. Lors du refroidissement, la pièce rétrécit légèrement mais conserve la forme du moule.

A. LOI NORMALE

La mesure de l'arête d'une boîte est modélisée par une variable aléatoire X qui suit la loi normale d'espérance $\mu = 17$ et d'écart type $\sigma = 0{,}185$.

1. Calculer $P(X \geqslant 16{,}5)$.
2. Calculer $P(X \geqslant 17{,}5)$.
3. Une boîte est jugée conforme lorsque la mesure de son arête, exprimée en centimètres, appartient à l'intervalle $[16{,}6 ; 17{,}4]$.
   a) Calculer $P(16{,}6 \leqslant X \leqslant 17{,}4)$.
   b) Déterminer la probabilité qu'une boîte prélevée au hasard dans la production soit non conforme.

B. LOI BINOMIALE

L'entreprise conditionne ces boîtes par lots de 200. On prélève au hasard une boîte dans la production.
On note p la probabilité de l'évènement : « la boîte prélevée au hasard dans la production est non conforme ».
On prélève au hasard 200 boîtes 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 Y qui, à un lot de 200 boîtes, associe le nombre de boîtes non conformes qu'il contient. On admet que Y suit une loi binomiale de paramètres 200 et p, et, qu'en moyenne chaque lot de 200 boîtes en contient 6 non conformes.

1. Justifier que $p = 0{,}03$.
2. Calculer la probabilité qu'il y ait au moins deux boîtes non conformes dans ce lot de 200 boîtes.

C. INTERVALLE DE FLUCTUATION

Dans le cadre d'un fonctionnement correct du thermoformage, on admet que la proportion p de boîtes non conformes dans la production est 3 %.

1. Déterminer les bornes de l'intervalle de fluctuation asymptotique à 95 % pour un échantillon de taille 200.
2. On contrôle le bon fonctionnement du thermoformage en prélevant au hasard dans la production des échantillons de 200 boîtes. Au cours de l'un de ces contrôles, un technicien a compté 10 boîtes non conformes.
   Doit-il prendre la décision d'effectuer des réglages sur la thermoformeuse ? Justifier la réponse.

EXERCICE 2

Une entreprise fabrique en grande quantité des tubes en aluminium.
La longueur des tubes est exprimée en millimètres. Un tube est dit « conforme pour la longueur » lorsque celle-ci appartient à l'intervalle $[245 ; 255]$. *Dans cet exercice, les résultats approchés sont à arrondir à* 10^{-3}

PARTIE A

Dans cette partie, on considère que 5 % des tubes fabriqués ne sont pas conformes pour la longueur.
On prélève au hasard 50 tubes dans le stock. Le stock est suffisamment important pour que l'on puisse assimiler ce prélèvement à un tirage avec remise de 50 tubes.
On considère la variable aléatoire X qui, à tout prélèvement de 50 tubes, associe le nombre de tubes qui ne sont pas conformes pour la longueur.

1. Calculer l'espérance mathématique $E(X)$ de la variable aléatoire X.

2. Calculer la probabilité $P(X = 2)$. Interpréter le résultat.
3. Calculer la probabilité que dans un tel prélèvement deux tubes au moins ne sont pas conformes pour la longueur.

PARTIE B

On désigne par Y la variable aléatoire qui, à chaque tube pris au hasard dans la production d'une journée, associe sa longueur.
On admet que la variable aléatoire Y suit la loi normale de moyenne 250 et d'écart type 2,5.

1. Calculer la probabilité qu'un tube prélevé au hasard dans la production d'une journée soit conforme pour la longueur.
2. Le contrôle de conformité mis en place rejette les tubes dont la longueur est inférieure à 245 millimètres.
   Quelle est la probabilité pour qu'un tube prélevé au hasard dans la production d'une journée soit rejeté par le contrôle de conformité ?

PARTIE C

Le cahier des charges établit que la proportion dans la production de 2 % de tubes refusés par le contrôle de conformité est acceptable.
On veut savoir si une machine de production est correctement réglée. Pour cela on prélève au hasard dans la production un échantillon de taille 250 dans lequel 6 tubes se révèlent être non conformes.

1. Donner l'intervalle de fluctuation asymptotique à 95 % de la fréquence des tubes non conformes dans un échantillon de taille 250.
2. La machine de production doit-elle être révisée ? Justifier votre réponse.

EXERCICE 3

Une entreprise fabrique, en grande quantité, des batteries Lithium-ion pour smartphone.

PARTIE A

Ces batteries sont produites par trois ateliers. L'atelier A produit 15% des batteries, l'atelier B en produit 20% et l'atelier C fournit le reste de la production.
Le contrôle de qualité mis en place a permis d'établir que sur l'ensemble de la production 3% des batteries sont défectueuses, 6% des batteries produites dans l'atelier A sont défectueuses et 4% des batteries produites dans l'atelier B sont défectueuses.
On prélève au hasard une batterie parmi la production totale de l'entreprise et, on définit les évènements suivants :

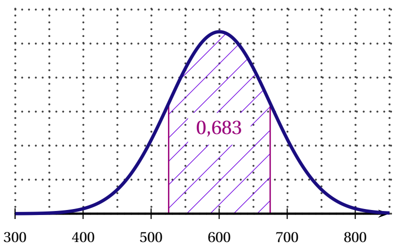
— A : « la batterie provient de l'atelier A » ;
— B : « la batterie provient de l'atelier B » ;
— C : « la batterie provient de l'atelier C » ;
— D : « la batterie est défectueuse ».

1. a) En utilisant l'énoncé, donner les probabilités $P(D)$, $P_A(D)$, $P_B(D)$ et calculer $P(C)$.
   b) Représenter la situation à l'aide d'un arbre pondéré.
2. Calculer $P(A \cap D)$ et formuler une interprétation de ce résultat.
3. a) Montrer $P(C \cap D) = 0{,}013$.
   b) En déduire la probabilité qu'une batterie produite par l'atelier C soit défectueuse.
4. La batterie est défectueuse. Quelle est la probabilité qu'elle ait été produite par l'atelier A ?
5. On prélève au hasard 10 batteries dans la production d'une journée. La production est suffisamment importante pour assimiler ce prélèvement à 10 tirages avec remise.
   Quelle est la probabilité, arrondie à 10^{-3} près, que parmi les 10 batteries, il y en ait au moins une qui soit défectueuse ?

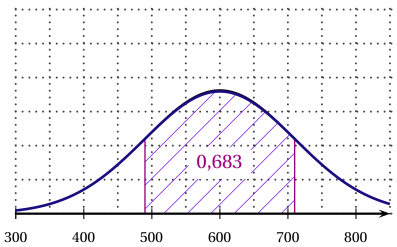
PARTIE B

Le nombre de cycles de charge d'une batterie est appelé durée de vie de la batterie.
La durée de vie des batteries Lithium-ion mises en vente par cette entreprise est modélisée par la variable aléatoire X suivant la loi normale de moyenne $\mu = 600$ et d'écart-type $\sigma = 74{,}6$.

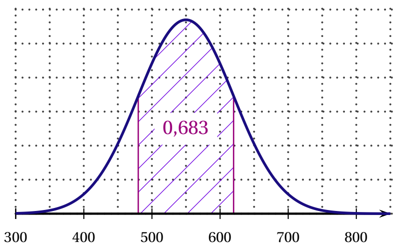
1. La fonction densité associée à X est représentée sur un seul de trois graphiques ci-dessous.
   Quel est ce graphique ? Expliquer le choix.



Graphique 1



Graphique 2



Graphique 3

2. a) Déterminer $P(550 \leqslant X \leqslant 1000)$ en donnant le résultat arrondi au centième.
   b) Quelle est la probabilité que la durée de vie d'une batterie soit inférieure à 500 cycles de charge ?

PARTIE C

Le service commercial affirme que 91% des batteries proposées à la vente ont une durée de vie supérieure à 500 cycles de charge.
Pour vérifier cette affirmation, un laboratoire indépendant a reconstitué la vie de 100 batteries en simulant des cycles de charge et de décharge pour déterminer leur durée de vie en fonction de différents facteurs.
Sur ce lot, on a constaté que 13 batteries ont eu une durée de vie inférieure à 500 cycles de charge.

Le résultat de ce test remet-il en question l'affirmation du service commercial ?

EXERCICE 4

Sauf mention contraire, les résultats seront donnés sous forme décimale arrondis à 10^{-4} *près.*

Une usine fabrique en grande quantité des lames de parquet en chêne. Les bois proviennent de deux fournisseurs A et B.

PARTIE A

Dans le stock de cette usine, 75 % des bois proviennent du fournisseur A.
On constate que 9 % des lames obtenues à partir des bois du fournisseur A et 13 % des lames obtenues à partir des bois du fournisseur B présentent un léger défaut qui ne justifie pas le déclassement des lames.
On prélève au hasard une lame. On considère les évènements suivants :

— A : « la lame prélevée est obtenue à partir de bois du fournisseur A » ;
— B : « la lame prélevée est obtenue à partir de bois du fournisseur B » ;
— D : « la lame prélevée a un léger défaut ».

1. Calculer la probabilité $P(B \cap D)$.
2. Calculer la probabilité que la lame a un léger défaut.
3. Calculer la probabilité qu'une lame ayant un léger défaut provienne de bois du fournisseur A.

PARTIE B

On prélève au hasard 40 lames dans le stock, pour vérification. On admet que la probabilité qu'une lame prélevée au hasard dans ce stock ait un défaut est égale à 0,1.
Le stock est suffisamment important pour assimiler le lot de 40 lames à un tirage avec remise de 40 lames.
On considère la variable aléatoire X qui, à tout prélèvement de 40 lames dans ce stock, associe le nombre de lames ayant un défaut.

1. Justifier que la variable aléatoire X suit une loi binomiale.
2. Calculer l'espérance mathématique $E(X)$. Interpréter le résultat.
3. Déterminer la probabilité de trouver quatre lames qui ont un défaut.
4. Déterminer la probabilité qu'au moins deux lames ont un défaut.

PARTIE C

Pour satisfaire la commande d'un client, on prélève au hasard dans le stock 400 lames.
On admet que la loi de la variable aléatoire Z qui, à tout prélèvement de 400 lames dans ce stock, associe le nombre de lames ayant un défaut peut être approchée par la loi normale de moyenne 40 et d'écart type 6.

1. Déterminer, la probabilité que dans un prélèvement de 400 lames, il y ait plus de 50 lames ayant un défaut.
2. Déterminer l'intervalle de fluctuation asymptotique à 95 % de la proportion de lames ayant un défaut. En déduire le nombre de lames ayant un défaut que le client peut trouver avec une probabilité proche de 0,95.

PARTIE D

Le fabriquant souhaite évaluer la proportion inconnue p de clients satisfaits par son produit. Pour cela, il effectue un sondage auprès d'un échantillon de 200 clients. Sa clientèle est suffisamment importante pour considérer que cet échantillon résulte d'un tirage aléatoire avec remise.
Lors de ce sondage, 156 clients se sont déclarés satisfaits par son produit.

1. Donner une estimation ponctuelle f de la proportion p de clients satisfaits.
2. Déterminer un intervalle de confiance centré sur f de la proportion p avec le coefficient de confiance 95 %. Arrondir les bornes de l'intervalle à 10^{-2}.
3. Ce fabriquant peut-il être certain que plus de 70% de sa clientèle est satisfaite par son produit ?

EXERCICE 5 (*D'après sujet bac Amérique du Sud 2017*)

Une entreprise d'élevage de poissons en bassin a constaté qu'une partie de sa production est infectée par une nouvelle bactérie.
Un laboratoire a réalisé deux prélèvements, l'un au mois de janvier et l'autre au mois de juin, afin d'étudier l'évolution de l'infection.

Les trois parties de cet exercice peuvent être traitées de façon indépendante.

PARTIE A

Au mois de janvier, lors du premier test, le laboratoire a prélevé au hasard 1 000 poissons parmi l'ensemble des poissons du bassin.
La fréquence de poissons infectés par la bactérie dans cet échantillon est $f_1 = 5\,\%$.
Au mois de juin, le laboratoire a prélevé de nouveau 1 000 poissons.
Pour ce second test, la fréquence de poissons infectés est $f_2 = 10\,\%$.
La fréquence de poissons infectés dans les deux échantillons ayant doublé en cinq mois, le laboratoire préconise d'arrêter la vente des poissons de l'entreprise.
On note p_1 la proportion de poissons infectés parmi tous les poissons du bassin au mois de janvier et p_2 la proportion de poissons infectés parmi tous les poissons du bassin au mois de juin.

1. Déterminer les intervalles de confiance au niveau de confiance 95 % de la proportion p_1 puis de la proportion p_2.
   On arrondira les bornes des intervalles à 10^{-3}.
2. Quel argument pourrait donner l'entreprise pour éviter l'arrêt de la vente ?

PARTIE B

Pour déterminer la fréquence de poissons infectés dans un prélèvement, le laboratoire dispose d'un test de dépistage dont les résultats sont les suivants :

— sur des poissons infectés par la bactérie, le test est positif dans 60 % des cas ;

— sur des poissons non infectés par la bactérie, le test est positif dans 10 % des cas.

Pour un poisson prélevé au hasard, on note :

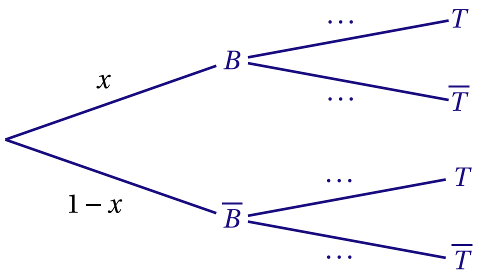
— B l'évènement : « le poisson est infecté par la bactérie » ;

— T l'évènement : « le test du poisson est positif » ;

— $\overline{B}$ et $\overline{T}$ les évènements contraires de B et T.

On note x la probabilité qu'un poisson soit infecté par la bactérie.

1. Recopier et compléter l'arbre pondéré traduisant cette situation.



2. a) Démontrer que $p(T) = 0{,}5x + 0{,}1$.

   b) Le laboratoire a constaté que 12,5 % des poissons d'un prélèvement ont eu un test positif.
   Quelle estimation de la proportion de poissons infectés le laboratoire va-t-il proposer pour ce prélèvement ?

PARTIE C

Un traitement antibiotique permet de guérir les poissons infectés par la bactérie.
Le temps de guérison d'un poisson infecté, exprimé en jours, peut être modélisé par une variable aléatoire X suivant la loi normale de moyenne $\mu = 21$ et d'écart-type $\sigma = 5$.
Les résultats seront arrondis au millième.

1. Déterminer la probabilité $p\,(14 < X < 28)$.
2. Déterminer la probabilité qu'un poisson infecté ne soit pas encore guéri après 5 semaines de traitement antibiotique.

EXERCICE 6

(*D'après sujet bac France métropolitaine, La Réunion septembre 2017*)

Les trois parties de cet exercice sont indépendantes.

En janvier 2015, le directeur d'un musée d'art contemporain commande une enquête concernant les habitudes des visiteurs.

PARTIE A

Le musée dispose d'un site internet. Pour acheter son billet, une personne intéressée peut se rendre au guichet d'entrée du musée ou commander un billet en ligne.
Trois types de visites sont proposés :

— La visite individuelle sans location d'audioguide.

— La visite individuelle avec location d'audioguide.

— La visite en groupe d'au moins 10 personnes. Dans ce cas, un seul billet est émis pour le groupe.

Le site internet permet uniquement d'acheter les billets individuels avec ou sans audioguide. Pour la visite de groupe, il est nécessaire de se rendre au guichet d'entrée du musée.
Sur l'année 2015 l'enquête a révélé que :

— 55 % des billets d'entrée ont été achetés au guichet du musée;

— parmi les billets achetés au guichet du musée, 51 % des billets correspondent à des visites individuelles sans location d'audioguide, et 37 % à des visites avec location d'audioguide;

— 70 % des billets achetés en ligne correspondent à des visites individuelles sans location d'audioguide.

On choisit au hasard un billet d'entrée au musée acheté en 2015. On considère les évènements suivants :

— E : « le billet a été acheté en ligne »;

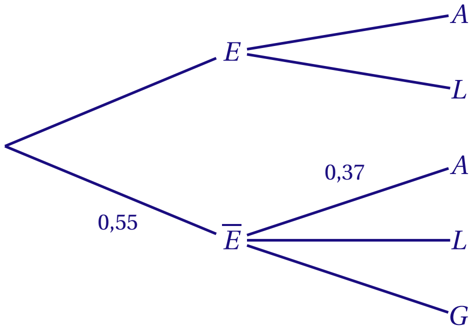
— A : « le billet correspond à une visite individuelle avec location d'audioguide »;

— L : « le billet correspond à une visite individuelle sans location d'audioguide »;

— G : « le billet correspond à une visite de groupe ».

On rappelle que si E et F sont deux évènements, $p(E)$ désigne la probabilité de l'évènement E et $p_F(E)$ désigne la probabilité de l'évènement E sachant que l'évènement F est réalisé. On note $\overline{E}$ l'évènement contraire de E.

1. Recopier et compléter l'arbre pondéré suivant qui représente la situation décrite dans l'énoncé :



2. Montrer que la probabilité que le billet ait été acheté en ligne et corresponde à une visite individuelle avec location d'audioguide est égale à 0,135.
3. Montrer que $p(A) = 0{,}3385$.
4. Le billet choisi correspond à une visite individuelle avec location d'audioguide. Quelle est la probabilité que ce billet ait été acheté au guichet du musée?
   On arrondira le résultat au millième.

PARTIE B

Pour gérer les flux des visiteurs, une partie de l'enquête a porté sur la durée d'une visite de ce musée. Il a été établi que la durée D d'une visite, en minutes, suit la loi normale de moyenne $\mu = 90$ et d'écart-type $\sigma = 15$.

1. Déterminer $p\,(90 \leqslant D \leqslant 120)$ puis interpréter ce résultat dans le contexte de l'exercice.
2. Le directeur précise qu'il augmentera la capacité d'accueil de l'espace restauration du musée si plus de 2 % des visiteurs restent plus de 2 heures et 30 minutes par visite.
   Quelle sera alors sa décision?

PARTIE C

Sur l'ensemble des musées d'art contemporain, 22 % des visiteurs sont de nationalité étrangère.
Sur un échantillon aléatoire de 2 000 visiteurs du musée considéré précédemment, 490 visiteurs sont de nationalité étrangère.
Que peut en conclure le directeur de ce musée? Argumenter.

EXERCICE 7 (*D'après sujet bac Polynésie 2017*)

Les trois parties de cet exercice sont indépendantes.

PARTIE A

D'après le « bilan des examens du permis de conduire » pour l'année 2014 publiée par le Ministère de l'Intérieur en novembre 2015, 20 % des personnes qui se sont présentées à l'épreuve pratique du permis de conduire avaient suivi la filière de l'apprentissage anticipé de la conduite (AAC). Parmi ces candidats, 75 % ont été reçus à l'examen. Pour les candidats n'ayant pas suivi la filière AAC, le taux de réussite à l'examen était seulement de 56,6 %.
On choisit au hasard l'un des candidats à l'épreuve pratique du permis de conduire en 2014.
On considère les évènements suivants :

— A « le candidat a suivi la filière AAC » ;

— R « le candidat a été reçu à l'examen ».

On rappelle que si E et F sont deux évènements, la probabilité de l'évènement E est notée $P(E)$ et celle de E sachant F est notée $P_F(E)$. De plus $\overline{E}$ désigne l'évènement contraire de E.

1. a) Donner les probabilités $P(A)$, $P_A(R)$ et $P_{\overline{A}}(R)$.
   b) Traduire la situation par un arbre pondéré.
2. a) Calculer la probabilité $P(A \cap R)$.
   b) Interpréter ce résultat dans le cadre de l'énoncé.
3. Justifier que $P(R) = 0{,}6028$.
4. Sachant que le candidat a été reçu à l'examen, calculer la probabilité qu'il ait suivi la filière AAC.
   On donnera une valeur approchée à 10^{-4} près de cette probabilité.

PARTIE B

Un responsable d'auto-école affirme que pour l'année 2016, la probabilité d'être reçu à l'examen est égale à 0,62.
Ayant des doutes sur cette affirmation, une association d'automobilistes décide d'interroger 400 candidats à l'examen parmi ceux de 2016. Il s'avère que 220 d'entre eux ont effectivement obtenu le permis de conduire.

1. Déterminer un intervalle de fluctuation asymptotique au seuil de 95 % de la fréquence de candidats reçus dans un échantillon aléatoire de 400 candidats.
2. Peut-on émettre des doutes sur l'affirmation du responsable de cette auto-école ?
   Justifier votre réponse.

PARTIE C

Selon une enquête menée en 2013 par l'association « Prévention Routière », le coût moyen d'obtention du permis de conduire atteignait environ 1 500 €. On décide de modéliser le coût d'obtention du permis de conduire par une variable aléatoire X qui suit la loi normale d'espérance $\mu = 1500$ et d'écart-type $\sigma = 410$.

1. Déterminer une valeur approchée à 10^{-2} près de la probabilité que le coût du permis de conduire soit compris entre 1 090 € et 1 910 €.
2. Déterminer $P(X \leqslant 1155)$.
   On donnera le résultat sous forme approchée à 10^{-2} près.
3. a) Par la méthode de votre choix, estimer la valeur du nombre réel a arrondi à l'unité, vérifiant $P(X \geqslant a) = 0{,}2$.
   b) Interpréter ce résultat dans le cadre de l'énoncé.

EXERCICE 8 (*D'après sujet bac France métropolitaine, La Réunion 2017*)

Dans cet exercice, tous les résultats seront arrondis au millième près.

1. Un supermarché dispose de plusieurs caisses. Un client qui se présente à une caisse doit attendre un certain temps T_l avant d'être pris en charge par le caissier. On considère que ce temps d'attente T_1 exprimé en minute, est une variable aléatoire qui suit la loi uniforme sur l'intervalle $[0;12]$.

   a) Quelle est la probabilité qu'un client attende au moins 5 minutes avant d'être pris en charge ?

   b) Quel est le temps moyen d'attente à une caisse ?

2. Le gérant du magasin décide de mettre à disposition des clients des caisses automatiques, de façon à réduire le temps d'attente pour les clients ayant un panier contenant peu d'articles.

   Le temps d'attente T_2, exprimé en minute, à chacune de ces caisses automatiques est modélisé par une variable aléatoire qui suit la loi normale de moyenne 5 et d'écart type 1,5.

   Calculer la probabilité que le temps d'attente à une caisse automatique soit compris entre 0,75 minute et 6 minutes.

3. Ces caisses automatiques tombent souvent en panne. On donne les informations suivantes.

   — Le nombre de caisses automatiques est $n = 10$.

   — La probabilité qu'une caisse automatique tombe en panne pendant une journée donnée est $p = 0{,}1$.

   — Une panne constatée sur une caisse automatique n'influence pas les autres caisses automatiques.

   Soit X la variable aléatoire correspondant au nombre de caisses automatiques qui tombent en panne pendant une journée donnée.

   a) Quelle est la loi de probabilité suivie par X ? Préciser ses paramètres.

   b) Calculer la probabilité pour qu'aucune caisse automatique ne tombe en panne pendant une journée donnée.

4. Sur la devanture de son magasin, le gérant du supermarché affiche :

   « Plus de 90 % des clients de notre magasin sont satisfaits par la mise en place de nos caisses automatiques. »

   Une association de consommateurs souhaite examiner cette affirmation. Pour cela, elle réalise un sondage : 860 clients sont interrogés, et 763 d'entre eux se disent satisfaits par la mise en place de ces caisses automatiques.

   Cela remet-il en question l'affirmation du gérant ?