Échantillon de taille n d'une variable aléatoire.

Échantillon de taille n.

Définition 1. Soit $X : \Omega \to \mathbb{R}$ une variable aléatoire. Un *échantillon de taille n* de la variable aléatoire X est une n-liste de variables aléatoires $(X_1, X_2, \ldots, X_n)$ indépendantes suivant toutes la même loi que X.

Exemples.

1. On lance 20 fois un dé parfaitement équilibré. On note, pour $i \in [\![1,20]\!]$, X_i la variable aléatoire égale à 3 si le résultat est strictement supérieur à 4 et -1 sinon. $(X_1, \ldots, X_n)$ est un échantillon de la variable aléatoire X dont la distribution est $\left(\left(3, \frac{1}{3}\right), \left(-1, \frac{2}{3}\right)\right)$.
2. Soit $(X_0, X_1, \ldots, X_n)$ un échantillon de taille $n+1$ d'une variable aléatoire $X \hookrightarrow \mathscr{B}(p)$ ($p \in]0,1[$). On note, pour $i \in [\![1,n]\!]$, $Y_i = X_i + X_{i-1}$. $(Y_1, \ldots, Y_n)$ ne constitue pas un échantillon De la variable aléatoire $X_0 + X_1$ car Y_1 et Y_2 ne sont pas indépendantes.
3. Soient $X_1 \hookrightarrow \mathscr{B}(p)$ et $X_2 \hookrightarrow \mathscr{U}([\![0,1]\!])$. (X_1, X_2) ne constitue pas un échantillon de taille 2 car X_1 et X_2 ne suivent pas la même loi (si $p \neq \frac{1}{2}$).

Remarques.

1. On parle aussi de variables indépendantes et identiquement distribuées.

Proposition 1. Soient : $X : \Omega \to \mathbb{R}$ une variable aléatoire, $(X_1, X_2, \ldots, X_n)$ un échantillon de taille n de la loi de X, $S_n = X_1 + X_2 + \cdots + X_n$ la variable aléatoire somme de l'échantillon, $M_n = \frac{S_n}{n}$ la variable aléatoire moyenne de l'échantillon. $\mathbb{E}(S_n) = n\mathbb{E}(X)$, $\mathbb{V}(S_n) = n\mathbb{V}(X)$, $\sigma(S_n) = \sqrt{n}\sigma(X)$. $\mathbb{E}(M_n) = \mathbb{E}(X)$, $\mathbb{V}(M_n) = \frac{1}{n}\mathbb{V}(X)$, $\sigma(M_n) = \frac{1}{\sqrt{n}}\sigma(X)$.

Démonstration.

* $\mathbb{E}(S_n) = \mathbb{E}(X_1 + X_2 + \cdots + X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}(X_n) = \mathbb{E}(X) + \mathbb{E}(X) + \cdots + \mathbb{E}(X) = n\mathbb{E}(X)$.
* $\mathbb{V}(S_n) = \mathbb{V}(X_1 + X_2 + \cdots + X_n) = \mathbb{V}(X_1) + \mathbb{V}(X_2) + \cdots + \mathbb{V}(X_n)$ car les variables X_1, X_2,..., X_n sont indépendantes. $\mathbb{V}(S_n) = \mathbb{V}(X) + \mathbb{V}(X) + \cdots + \mathbb{V}(X) = n\mathbb{V}(X)$.
* $\sigma(S_n) = \sqrt{\mathbb{V}(S_n)} = \sqrt{n\mathbb{V}(X)} = \sqrt{n}\sigma(X)$.
* $\mathbb{E}(M_n) = \mathbb{E}\left(\frac{S_n}{n}\right) = \frac{1}{n}\mathbb{E}(S_n) = \frac{1}{n}\mathbb{E}(X)$.
* $\mathbb{V}(M_n) = \mathbb{V}\left(\frac{S_n}{n}\right) = \left(\frac{1}{n}\right)^2 \mathbb{V}(S_n) = \frac{1}{n^2} \times n\mathbb{V}(X) = \frac{1}{n}\mathbb{V}(X)$.
* $\sigma(M_n) = \sigma\left(\mathbb{V}(M_n)\right) = \sqrt{\frac{1}{n}\mathbb{V}(X)} = \frac{1}{\sqrt{n}}\sigma(X)$.

EXERCICE 1. On étudie la marche aléatoire d'une particule se déplaçant sur les points d'abscisses entières d'un axe gradué d'origine O. La particule est à l'origine au temps 0 et se déplace de chaque unité de temps d'une unité vers la droite avec la probabilité $\frac{1}{2}$ ou d'une unité vers la gauche avec la probabilité $\frac{1}{2}$. On suppose les déplacements de la particule indépendants les uns des autres. Pour tout entier naturel k non nul on note X_k la variable aléatoire qui vaut 1 si le k-ième déplacement a lieu vers la droite et qui vaut -1 dans le cas contraire. Pour tout entier naturel n non nul, on note S_n l'abscisse de la particule à l'instant n.

1. Donnez, pour tout entier naturel non nul k l'espérance et la variance de X_k.
2. Pour tout entier naturel non nul n, exprimer S_n en fonction de certaines des variables X_k.
3. En déduire l'espérance et la variance de S_n.

Exercice 1.

1. $X_k \hookrightarrow \mathscr{U}(\{-1; 1\})$. $\mathbb{E}(X) = 0$ et $\mathbb{V}(X) = 1$.
2. $S_n = \sum_{k+1}^{n} X_k$.
3. $\mathbb{E}(S_n) = \sum_{k=1}^{n} \mathbb{E}(X_k) = 0$ et Puisque les X_k sont indépendantes : $\mathbb{V}(S_n) = \sum_{k=1}^{n} \mathbb{V}(X_k) = n$.

EXERCICE 2. Une machine remplit des pots de yaourts dont la masse en grammes est une variable aléatoire Y d'espérance 125 et d'écart-type 1,2. On suppose que les remplissages des pots sont indépendants.

1. Exprimez la masse d'un lot de 16 comme la variable aléatoire somme S_{16} d'un échantillon de 16 pots.
2. Déterminez l'espérance de S_{16}.
3. Montrez que l'écart-type de S_{16} est 4,8.

Exercice 2.

1. Soit X_i la masse du i-ème pot d'un échantillon de 16 pots. D'après l'énoncé les X_i suivent la même loi que Y et de plus sont indépendantes deux à deux. Donc en notant $S_{16} = X_1 + \cdots + X_{16}$ nous avons bien S_{16} variable aléatoire somme d'un échantillon de taille 16 de la variable aléatoire Y.
2. $\mathbb{E}(S_{16}) = 16 \times \mathbb{E}(Y) = 16 \times 125 = 2000$ g.
3. $\mathbb{V}(S_16) = \frac{1}{16}\mathbb{V}(Y) = \frac{1}{16} \times 1{,}2^2 = 0{,}09$ donc $\sigma(X) = 0{,}3$ g.

Application à la loi binomiale.

Proposition 2. Soit $X \hookrightarrow \mathscr{B}(n,p)$. $\mathbb{E}(X) = np$, $\mathbb{V}(X) = np(1-p)$ et $\sigma(X) = \sqrt{np(1-p)}$.

Démonstration. $X = Y_1 + Y_2 + \cdots + Y_n$ où $(Y_1, Y_2, \ldots, Y_n)$ est un échantillon de taille n d'une variable aléatoire $Y \hookrightarrow \mathscr{B}(p)$ et on applique la précédente proposition en sachant que $\mathbb{E}(Y) = p$, $\mathbb{V}(Y) = p(1-p)$.

Exemples.

1. Si $X \hookrightarrow \mathscr{B}\left(100, \frac{1}{4}\right) = 100 \times \frac{1}{4} = 25$ et $\mathbb{V}(X) = 100 \times \frac{1}{4}\left(1 - \frac{1}{4}\right) = 18{,}75$.

EXERCICE 3. Une étude statistique a été réalisée sur le temps d'attente en secondes subi par la clientèle avant d'être mise en communication avec un standardiste. La variable aléatoire T qui a tout client associe son temps d'attente a pour espérance 18 et pour écart-type 7. On estime que la probabilité qu'un client ait une attente de plus de 20 s est de 0,4.

1. Au cours d'une même semaine un client passe 5 appels indépendants les uns des autres. On note X la variable aléatoire comptant le nombre de fois que l'appel à dépassé 20 s. Déterminez l'espérance et l'écart-type de X.
2. Dans le but de diminuer le temps d'attente, on effectue une enquête sur un échantillon de 100 clients. Soit Y la variable aléatoire exprimant le temps d'attente moyen en secondes pour un échantillon de 100 clients. Déterminez l'espérance et l'écart-type de Y.

EXERCICE 4. On interroge de manière indépendante 300 personnes lors d'un sondage sur la réalisation d'un projet d'urbanisme. On note X_i la variable aléatoire qui vaut 0 si la i-ième personne interrogée est opposée au projet et 1 si elle y est favorable ; la probabilité qu'elle soit opposée au projet est 0,365.

1. Donnez la loi de probabilité des X_i.
2. On note $S = X_1 + \cdots + X_{300}$. Donnez en justifiant la loi de probabilité de la variable aléatoire S puis calculez et interprétez $\mathbb{P}(S > 100)$.

Exercice 4.

1. $\forall i \in [\![1,300]\!]), X_i \hookrightarrow \mathscr{B}(0{,}635)$.

2. Les variables aléatoires X_i, pour $i \in [\![1,300]\!]$, suivent toutes la même loi de Bernoulli et sont supposées indépendantes donc la variable aléatoire somme de l'échantillon $(X_1,\ldots,X_{300})$ suit une loi binomiale : $X_1 + \cdots + X_{300} \hookrightarrow \mathscr{B}(300{,}0{,}635)$.

$$\begin{aligned}\mathbb{P}(S > 100) &= 1 - \mathbb{P}\left(\overline{S > 100}\right) \\ &= 1 - \mathbb{P}(S \leqslant 99)\end{aligned}$$

En utilisant la fonction de répartition de la loi binomiale programmée dans la calculatrice on obtient :

$$\mathbb{P}(S > 100) =$$

Exercices.

EXERCICE 5. Une machine produit des joints dont l'épaisseur définit une variable aléatoire X d'espérance 2 mm et d'écart-type 0,1 mm. Soit M_n la variable aléatoire moyenne d'un échantillon de taille n de la variable aléatoire X.

1. Déterminez l'espérance de M_n.
2. Calculez l'écart-type de M_{100} puis de $M_{10\,000}$.

EXERCICE 6. Dans une population d'adultes la variable aléatoire X associe à chaque individu sa glycémie en milligrammes par 100 millilitres. On suppose que X a pour espérance 92 et pour écart-type 7. Soit M_{400} la variable aléatoire d'un échantillon de taille 400 de la variable aléatoire X. Déterminez l'espérance et l'écart-type de M_{400}.

Exercice 6. $\mathbb{E}(M_{400}) = 400 \times \mathbb{E}(X) = 400 \times 92 = 9200$.
$\sigma(M_{400}) = \sqrt{400}\sigma(X) = 200 \times 7 = 1400$ mg/dℓ.

EXERCICE 7. Sur des paquets de cartes à collectionner il est écrit que 20 % des paquets comportent une et une seule carte rare.

1. Déterminez la loi de probabilité de la variable aléatoire C qui à chaque paquet associe le nombre de cartes rares.
2. On achète 12 paquets : le nombre de cartes rares de ce lot est modélisé par un échantillon de taille 12 de la variable aléatoire C, noté $(C_1,\ldots,C_{12})$. Expliquez pourquoi il est plausible d'affirmer que les variables aléatoires formant cet échantillon sont indépendantes.
3. Déterminez la loi de probabilité de $L = C_1 + \cdots + C_{12}$.
4. A-t-on plus de 50 % de chance d'avoir récupéré 4 cartes rares ?

Exercice 7.

1. $C \hookrightarrow \mathscr{B}(0{,}20)$.
2.
3. $L \hookrightarrow \mathscr{B}(12; 0{,}2)$.
4. $\mathbb{P}(L = 4) = \binom{12}{4} 0{,}2^4 (1 - 0{,}2)^{12-4}$.

EXERCICE 8. La masse en grammes d'un paquet de café est donnée par un variable aléatoire C d'espérance 250 et de variance 4. Soit L la variable aléatoire donnant la masse en grammes d'un lot de deux paquets. On suppose les masses des deux paquets indépendantes entre elles. Déterminez $\mathbb{E}(L)$ et $\mathbb{V}(L)$.

EXERCICE 9. Afin de mieux sécuriser les achats en ligne, certaines cartes bancaires ont un cryptogramme dynamique : ce code de sécurité qui figure au dos de la carte change chaque heure de manière « aléatoire » : chaque cryptogramme est composé de trois chiffres et les combinaisons de trois chiffres identiques, comme 111 ou 888 sont exclues.

1. Combien de cryptogrammes différents la carte bancaire peut-elle générer ?
2. Combien de cryptogrammes comporte deux fois le même chiffre ?

3. On note D la variable aléatoire valant 1 si le cryptogramme comportent deux fois le même chiffre et 0 sinon. Établissez la loi de probabilité de la variable aléatoire D.
4. $(D_1, \ldots, D_{24})$ est un échantillon de taille 24 de la variable aléatoire D et $S = D_1 + \cdots + D_{24}$. Calculez $\mathbb{P}(S \geqslant 2)$.

EXERCICE 10. On réalise des semis dans des pots séparés et cultivés dans les mêmes conditions et on constate qu'ils survivent au bout de 2 mois dans 78 % des cas.

1. On note V la variable aléatoire qui à chaque semis associe 1 s'il est vivant au bout de deux mois et 0 sinon. Établissez la loi de V.
2. 500 semis ont été plantés. on les modélise par un échantillon $(V_1, \ldots V_{500})$ de taille 500 de V et on note $S = V_1 + \cdots + V_{500}$. Déterminez la loi de probabilité de S.
3. Calculez la probabilité $\mathbb{P}(S > 400)$ et interprétez ce résultat.

EXERCICE 11. Le jeu des mille euros est un jeu radiophonique au cours duquel une équipe de deux candidats doit répondre successivement à six questions de difficulté progressive : trois questions bleues, deux questions blanches et une question rouge. On estime que la probabilité de répondre juste à une question bleue est 0,95, celle de répondre juste à une question blanche est de 0,90 et celle de répondre juste à une question rouge est de 0,80. Les questions sont indépendantes entre elles. On note X la variable aléatoire qui, à une équipe, associe le nombre de réponses justes aux questions bleues, Y la variable aléatoire qui, à une équipe, associe le nombre de réponses justes aux questions blanches, et Z la variable aléatoire qui, à une équipe, associe 1 si elle répond juste à la question rouge et 0 sinon.

1. Déterminez les lois suivies par X, Y et Z.
2. Déterminez l'espérance et la variance de chacune de ces variables aléatoires.
3. On appelle R la variable aléatoire qui, à une équipe, associe le nombre total de réponses justes. Déterminez l'espérance et la variance de R.
4. Les candidats qui ont 6 bonnes réponses justes accèdent directement à la question Banco, et ceux qui ont 5 réponses justes passent par un repêchage pour accéder au Banco. On estime que la probabilité de répondre correctement à la question de repêchage est de 0,8.
   (a) Illustrez la situation par un arbre pondéré.
   (b) Quelle est la probabilité d'accéder directement à la question Banco ?
   (c) Quelle est la probabilité de passer par un repêchage ?
   (d) Quelle est la probabilité d'accéder au Banco ?
5. La probabilité de répondre juste à la question Banco est de 0,4. Quelle est la probabilité pour une équipe de candidats qui se présente au jeu, de remporter le Banco ?

EXERCICE 12.

0ieme bordas indice page 414 exo 59 à 69