Université de Picardie Jules Verne **2017-2018**
IAE Amiens

Licence mention Gestion parcours Management et Marketing Vente - Semestre 5
Statistiques appliquées

Notion de variable aléatoire

1. Variable aléatoire

Considérons une population Ω sur laquelle on définit un caractère quantitatif X.

X est une application de Ω dans $\mathbb{R}$ qui, à tout individu ω, associe un réel $x = X(\omega) \in X(\Omega) = \Omega_X$ ensemble des valeurs du caractère.

Cette application modélise le caractère de façon déterministe en ce sens que, si on connaît l'individu ω, on connaît aussitôt la valeur x. Son étude relève de la statistique descriptive qui conduit, par exemple, au tableau des couples (x_i, f_i) où x_i est une valeur observée et f_i sa fréquence.

Supposons maintenant que l'on tire au hasard un individu ω dans cette population Ω pour consigner la valeur x du caractère. Ne pouvant pas prévoir quel individu précis sera tiré, on ne peut pas prévoir non plus la valeur précise de x qui sera consigner. On aimerait donc disposer d'un moyen d'attribuer une probabilité aux éléments de Ω_X. L'idée est de transporter sur Ω_X la probabilité sur Ω construite pour modéliser la situation aléatoire correspondant au tirage aléatoire d'un individu.

De même qu'un caractère quantitatif peut être discret ou continu, on parlera de variable aléatoire discrète ou continue (dans le deuxième cas, on parlera de variable aléatoire à densité).

Exemple introductif.

Un joueur lance un dé équilibré à 6 faces numérotées de 1 à 6, et on observe le numéro obtenu. Si le joueur obtient 1, 3 ou 5, il gagne 1 euro ; s'il obtient 2 ou 4, il gagne 5 euros ; s'il obtient 6, il perd 10 euros.

Prenons $\Omega = \{1,2,3,4,5,6\}$, $\mathcal{A} = \mathcal{P}(\Omega)$ et P l'équiprobabilité sur $(\Omega, \mathcal{A})$

Soit X l'application de Ω dans $\mathbb{R}$ qui à tout $\omega \in \Omega$ associe le gain correspondant.

On a $X(1) = X(3) = X(5) = 1$, $X(2) = X(4) = 5$ et $X(6) = -10$. Ainsi, $\Omega_X = \{1, 5, -10\}$.

On peut s'intéresser à la probabilité de gagner 1 euro, c'est-à-dire d'avoir $X(\omega) = 1$, ce qui se réalise si et seulement si $\omega \in \{1,3,5\}$. La probabilité cherchée est donc égale à $P(\{1,3,5\}) = 3/6 = 1/2$. On écrira aussi $P_X(\{1\}) = P(X = 1) = 1/2$.

On pourra donc considérer l'événement : $(X = 1) = \left\{\omega \in \Omega \ / \ X(\omega) = 1\right\} = \{1,3,5\}$.

On aura de même $P(X = 5) = 1/3$ et $P(X = -10) = 1/6$.

On a ainsi construit un nouvel ensemble $\Omega_X = \{-10, 1, 5\}$, et muni cet ensemble de la probabilité P_X définie par les valeurs ci-dessus : $P_X(\{-10\}) = 1/6$, $P_X(\{1\}) = 1/2$ et $P_X(\{5\}) = 1/3$.

1.1. Cas d'un nombre fini de valeurs

Définitions. **Notations**.

On considère une expérience aléatoire et $(\Omega, \mathcal{A}, P)$ est un espace probabilisé.**fini** adapté.

On appelle variable aléatoire réelle (v.a.r.) toute application X de Ω dans $\mathbb{R}$.

L'ensemble $\Omega_X = X(\Omega)$ des valeurs prises par X est appelé univers-image. Ω_X est alors fini.

On désigne par P_X la probabilité-image définie par :

pour tout $x \in \Omega_X$, $P_X(\{x\}) = P\left(\left\{\omega \in \Omega / X(\omega) = x\right\}\right) = P(X = x)$.

Les événements $\left\{\omega \in \Omega / X(\omega) = x\right\}$, $\left\{\omega \in \Omega / X(\omega) \leq x\right\}$ et $\left\{\omega \in \Omega / a < X(\omega) \leq x\right\}$ sont notés respectivement $(X = x)$, $(X \leq x)$ et $(a < X \leq x)$.

Loi de probabilité.

Si X est une v.a.r. dont l'univers-image fini est $\Omega_X = \{x_1, \ldots, x_n\}$, alors X est appelée v.a.r. discrète finie et on appelle loi de probabilité de X la donnée des couples (x_i, p_i), avec $p_i = P(X = x_i)$.

On pourra toujours s'assurer que l'on a bien $p_i \geq 0$ et $\sum_{i=1}^{n} p_i = p_1 + p_2 + \cdots + p_n = 1$.

Exemple.

Reprenons l'exemple introductif. La loi de probabilité de X est donnée par

x_i	-10	1	5	
p_i	1/6	1/2	1/3	1

Espérance mathématique. **Variance**. **Ecart-type**.

Soit X une v.a.r. discrète dont l'univers-image fini est $\Omega_X = \{x_1, \ldots, x_n\}$.

On appelle espérance mathématique de X le réel $E(X) = \sum_{i=1}^{n} p_i x_i = p_1 x_1 + p_2 x_2 + \cdots + p_n x_n$. Ce nombre représente la valeur moyenne de X.

On appelle variance de X le réel $V(X) = \sum_{i=1}^{n} p_i (x_i - E(X))^2 = E(X^2) - (E(X))^2 = \sum_{i=1}^{n} p_i x_i^2 - (E(X))^2$.

On appelle écart-type de X le réel $\sigma(X) = \sqrt{V(X)}$. Ce nombre mesure l'écart à la moyenne des valeurs prises par X.

Exemple.

Reprenons l'exemple introductif.

On a $E(X) = \frac{1}{6} \times (-10) + \frac{1}{2} \times 1 + \frac{1}{3} \times 5 = \frac{3}{6} = \frac{1}{2}$.

On a $E(X^2) = \frac{1}{6} \times (-10)^2 + \frac{1}{2} \times 1^2 + \frac{1}{3} \times 5^2 = \frac{153}{6} = \frac{51}{2}$,

d'où $V(X) = E(X^2) - (E(X))^2 = \frac{51}{2} - \left(\frac{1}{2}\right)^2 = \frac{101}{4}$ et $\sigma(X) = \sqrt{V(X)} = \frac{\sqrt{101}}{2} = 5,025$.

Remarque.

En statistique descriptive des variables quantitatives discrètes, nous avons des formules suivantes analogues.

Moyenne : $\bar{x} = \frac{1}{n} \sum n_i x_i = \sum \frac{n_i}{n} x_i = \sum f_i x_i$. Variance : $s_x^2 = \frac{1}{n} \sum n_i x_i^2 - \bar{x}^2 = \sum f_i x_i^2 - \bar{x}^2$.

Ecart-type : $s_x = \sqrt{s_x^2}$.

Lois classiques : **Binomiale et Hypergéométrique**.

Répétition d'expériences.

On répète n fois dans les mêmes conditions une même expérience aléatoire (les répétitions étant indépendantes entre elles) au cours de laquelle un événement A a une probabilité p d'être réalisé.

La variable aléatoire X, égale au nombre de fois où l'événement A est réalisé, suit la loi Binomiale $\mathcal{B}(n, p)$, ce qui signifie que :

X est à valeurs dans $\Omega_X = \{0, 1, \ldots, n\}$ et, pour tout $k \in \{0, 1, \ldots, n\}$, $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

On a de plus : $E(X) = np$ et $V(X) = np(1-p)$.

Tirages dans une population à 2 catégories.

On considère une population de N individus à deux catégories, avec N_1 individus de catégorie 1 et N_2 individus de catégorie 2 (on a $N_1 + N_2 = N$).

On effectue n tirages d'un individu dans la population et on désigne par X la variable aléatoire égale au nombre d'individus de catégorie 1 obtenus.

Si on effectue les tirages **avec remise**, alors X suit la loi Binomiale $\mathcal{B}\left(n, \frac{N_1}{N}\right)$.

Si on effectue les tirages **sans remise** ou **simultanés** (dans ce cas, on doit avoir $n \leq N$), alors X suit la loi Hypergéométrique $\mathcal{H}\left(N, n, \frac{N_1}{N}\right)$, ce qui signifie que

X est à valeurs dans $\Omega_X = \{\max(0, n - (N - N_1)), \min(N_1, n)\}$

et pour tout $k \in \{0, 1, \ldots, n\}$, $P(X = k) = \frac{\binom{N_1}{k}\binom{N-N_1}{n-k}}{\binom{N}{n}}$.

On a de plus : $E(X) = np$ et $V(X) = np(1-p)\frac{N-n}{N-1}$ en posant $p = \frac{N_1}{N}$.

Exemple.

Un magasin de vêtements a constitué un stock d'un certain type de pantalons. On admet que le pourcentage de pantalons du stock présentant un défaut est égal à 5%. On choisit au hasard un lot de 3 pantalons dans le stock. On suppose que le stock est suffisamment important pour que ce choix puisse être assimilé à 3 tirages indépendants avec remise.

On appelle X la variable aléatoire qui dénombre les pantalons présentant un défaut dans le lot de 3 pantalons prélevés.

Alors X suit la loi Binomiale $\mathcal{B}(3; 0,05)$, c'est-à-dire X est à valeurs dans $\Omega_X = \{0, 1, \ldots, n\} = \{0, 1, 2, 3\}$ et pour tout $k \in \{0, 1, 2, 3\}$, $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{3}{k} (0,05)^k (0,95)^{3-k}$.

En effet, on peut considérer que l'on répète $n = 3$ fois l'expérience aléatoire "choisir au hasard un pantalon" au cours de laquelle l'événement A "le pantalon présente un défaut" a la probabilité $p = 0,05$ d'être réalisé.

On peut aussi considérer que l'on effectue $n = 3$ tirages d'un individu dans la population des N pantalons, avec N_1 pantalons avec défaut et N_2 pantalons sans défaut (on a $N_1 + N_2 = N$). On ne connaît ni N_1 ni N, donc on ne peut pas supposer les tirages sans remise qui conduiraient à la loi Hypergéométrique. On doit donc supposer les tirages avec remise et $\dfrac{N_1}{N} = 0,05$. Ceci est justifié par le résultat suivant.

Propriété.

Lorsque N est très grand devant n (en pratique $N > 10n$), on peut approcher la loi Hypergéométrique $\mathcal{H}\left(N, n, \dfrac{N_1}{N}\right)$ par la loi Binomiale $\mathcal{B}\left(n, \dfrac{N_1}{N}\right)$. Cela signifie que $\dfrac{\binom{N_1}{k}\binom{N-N_1}{n-k}}{\binom{N}{n}} \simeq \binom{n}{k}\left(\dfrac{N_1}{N}\right)^k \left(1 - \dfrac{N_1}{N}\right)^{n-k}$. Et donc confondre tirages avec ou sans remise.

1.2. Cas d'un nombre infini dénombrable de valeurs

Définitions.

On appelle variable aléatoire réelle discrète infinie toute v.a.r. dont l'univers-image Ω_X (ensemble des valeurs prises par X) est infini dénombrable. Dans ce cours, on se limitera à $\Omega_X = \mathbb{N} = \left\{k \,/\, k \in \mathbb{N}\right\}$ ou $\Omega_X = \mathbb{N}^*$.

On appelle loi de probabilité de X la donnée des couples (k, p_k), avec $p_k = P(X = k)$, pour $k \in \Omega_X$.

On pourra toujours s'assurer que l'on a bien $p_k \geq 0$ et $\sum_{k \in \Omega_X} p_k = 1$.

Attention ! Cette dernière somme contient une infinité de termes ! Il s'agit de la somme d'une série numérique, qui peut être infinie ou ne pas exister. Nous n'aborderons pas pour le moment les problèmes liés à cette notion.

Espérance mathématique. **Variance**. **Ecart-type**.

Soit X une v.a.r. discrète dont l'univers-image infini est Ω_X.

Sous reserve d'existence des sommes indiquées ci-dessous, on a :

- espérance mathématique de X : $E(X) = \sum_{k \in \Omega_X} k p_k$.
- variance de X : $V(X) = \sum_{k \in \Omega_X} (k - E(X))^2 p_k$ et $V(X) = E(X^2) - (E(X))^2 = \sum_{k \in \Omega_X} k^2 p_k - (E(X))^2$.
- écart-type de X : $\sigma(X) = \sqrt{V(X)}$ - intervalle moyen de X : $[E(X) - \sigma(X), E(X) + \sigma(X)]$.

Lois classiques : **Poisson et Géométrique**.

Loi de Poisson.

Une variable aléatoire X suit la loi de Poisson $\mathcal{P}(\lambda)$ de paramètre $\lambda > 0$ si et seulement si :

X est à valeurs dans $\Omega_X = \mathbb{N}$ et, pour tout $k \in \mathbb{N}$, $P(X = k) = e^{-\lambda} \dfrac{\lambda^k}{k!}$.

On a de plus : $E(X) = \lambda$ et $V(X) = \lambda$.

Loi Géométrique.

On répète dans les mêmes conditions une même expérience aléatoire (les répétitions étant indépendantes entre elles) au cours de laquelle un événement A a une probabilité p d'être réalisé.

La variable aléatoire X, égale au nombre d'expériences nécessaires pour avoir la première réalisation de A (i.e. au temps d'attente de l'événement A), suit la loi Géométrique $\mathcal{G}(p)$, ce qui signifie que :

X est à valeurs dans $\Omega_X = \mathbb{N}^*$ et, pour tout $k \in \mathbb{N}^*$, $P(X = k) = p(1-p)^{k-1}$.

On a de plus : $E(X) = \dfrac{1}{p}$ et $V(X) = \dfrac{1-p}{p^2}$.

1.3. Cas d'un nombre infini non dénombrable de valeurs

Lorsque Ω_X est infini non-dénombrable (par exemple $\Omega_X = [a,b]$ ou $\mathbb{R}$), la théorie montre que $P(X = x) = 0$ pour tout x de $\mathbb{R}$. On est alors amené à donner la loi de probabilité de X par sa fonction de répartition F_X définie par $F_X(x) = P(X \leq x)$.

Définition.

On appelle *densité* ou *densité de probabilité sur* $\mathbb{R}$ toute fonction f de $\mathbb{R}$ dans $\mathbb{R}$ vérifiant :

- f est positive ;
- f est continue sur $\mathbb{R}$ sauf éventuellement en un nombre fini ou dénombrable de points ;
- l'intégrale $\int_{-\infty}^{+\infty} f(x)dx$ est convergente et $\int_{-\infty}^{+\infty} f(x)dx = 1$.

Définition. Une variable aléatoire X est dite *continue* s'il existe une densité f_X telle que pour tout $x \in \mathbb{R}$, $F_X(x) = \int_{-\infty}^{x} f_X(t)dt$.

f_X est alors appelée *densité de X*. On dit aussi que X est une *variable aléatoire réelle à densité*.

Propriétés.

Soit X une variable aléatoire continue dont f_X est une densité, et soit F_X sa fonction de répartition.

Alors F_X est une fonction de $\mathbb{R}$ dans $[0,1]$, continue et croissante sur $\mathbb{R}$, dérivable sur $\mathbb{R}$, sauf éventuellement en un nombre fini ou dénombrable de points, de dérivée f_X, et telle que $\lim\limits_{x\to-\infty} F(x) = 0$ et $\lim\limits_{x\to+\infty} F(x) = 1$. De plus, pour tous réels x, a et b tels que $a < b$, on a :

$$P(X = x) = 0 \text{ ; } P(X \leq x) = P(X < x) \text{ ; } P(a \leq X \leq b) = P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(t)dt.$$

Espérance mathématique. **Variance**. **Ecart-type**.

Sous les mêmes hypothèses que ci-dessus, et sous réserve que les intégrales suivantes existent, on a :

- $E(X) = \int_{-\infty}^{+\infty} x f_X(x)dx$ (ce nombre représente la valeur moyenne de X) ;
- $V(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f_X(x)dx$; $V(X) = E(X^2) - (E(X))^2 = \int_{-\infty}^{+\infty} x^2 f_X(x)dx - (E(X))^2$;
- $\sigma(X) = \sqrt{V(X)}$; intervalle moyen de X : $[E(X) - \sigma(X), E(X) + \sigma(X)]$.

Lois classiques : **Uniforme**, **Exponentielle et Normale**

Loi Uniforme sur $[a,b]$.

$$\Omega_X = [a,b],\ f_X(x) = \begin{cases} \dfrac{1}{b-a} \text{ si } x \in [a,b] \\ 0 \text{ si } x \notin [a,b] \end{cases} \quad F_X(x) = \begin{cases} 0 \text{ si } x < a \\ \dfrac{x-a}{b-a} \text{ si } x \in [a,b] \\ 1 \text{ si } x > b \end{cases}$$

$$E(X) = \frac{a+b}{2} \text{ et } V(X) = \frac{(b-a)^2}{12}.$$

Loi Exponentielle de paramètre θ, **avec** $\theta > 0$.

$$\Omega_X = [0,+\infty[,\ f_X(x) = \begin{cases} \theta e^{-\theta x} \text{ si } x \geq 0 \\ 0 \text{ si } x < 0 \end{cases},\ F_X(x) = \begin{cases} 1 - e^{-\theta x} \text{ si } x \geq 0 \\ 0 \text{ si } x < 0 \end{cases},\ E(X) = \frac{1}{\theta} \text{ et } V(X) = \frac{1}{\theta^2}.$$

1.4 Loi Normale $\mathcal{N}(\mu, \sigma)$.

1.4.1. Cas de la loi $\mathcal{N}(0, 1)$ (loi normale centrée réduite).

On a $\Omega_X = \mathbb{R}$ et $f_X(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ pour tout $x \in \mathbb{R}$.

On a de plus $E(X) = 0$ et $V(X) = 1$.

La fonction de répartition de X est donnée par $\phi(x) = F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \int_{-\infty}^{x} \dfrac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ pour tout $x \in \mathbb{R}$. Cette fonction n'étant pas exprimable par des fonctions usuelles, on utilise généralement une valeur approchée de cette fonction. Ses valeurs sont tabulées pour les $x \geq 0$ (table 1 en page 9). Pour les $x < 0$, on peut utiliser la formule $\phi(-x) = 1 - \phi(x)$ (formule valable pour tout réel x)

La table 1 ne donne des valeurs approchées de $\phi(x)$ que pour des x compris entre 0 et 3. On pourra admettre que $\phi(x) \simeq 1$ pour tout $x \geq 3$, ou utiliser les valeurs supplémentaires données.

TABLE 1

Fonction de répartition de la loi normale réduite

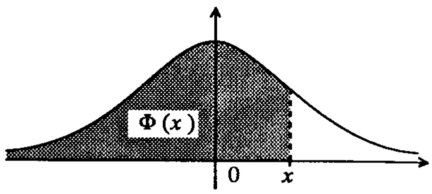
Si U suit la loi normale réduite, pour $x \geqslant 0$, la table donne la valeur :

$\phi(x) = P(U \leqslant x)$.

La valeur x s'obtient par addition des nombres inscrits en marge.

Pour $x < 0$, on a :

$\phi(x) = 1 - \phi(-x)$.



x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,500 0	0,504 0	0,508 0	0,512 0	0,516 0	0,519 9	0,523 9	0,527 9	0,531 9	0,535 9
0,1	0,539 8	0,543 8	0,547 8	0,551 7	0,555 7	0,559 6	0,563 6	0,567 5	0,571 4	0,575 3
0,2	0,579 3	0,583 2	0,587 1	0,591 0	0,594 8	0,598 7	0,602 6	0,606 4	0,610 3	0,614 1
0,3	0,617 9	0,621 7	0,625 5	0,629 3	0,633 1	0,636 8	0,640 6	0,644 3	0,648 0	0,651 7
0,4	0,655 4	0,659 1	0,662 8	0,666 4	0,670 0	0,673 6	0,677 2	0,680 8	0,684 4	0,687 9
0,5	0,691 5	0,695 0	0,698 5	0,701 9	0,705 4	0,708 8	0,712 3	0,715 7	0,719 0	0,722 4
0,6	0,725 7	0,729 1	0,732 4	0,735 7	0,738 9	0,742 2	0,745 4	0,748 6	0,751 7	0,754 9
0,7	0,758 0	0,761 1	0,764 2	0,767 3	0,770 4	0,773 4	0,776 4	0,779 4	0,782 3	0,785 2
0,8	0,788 1	0,791 0	0,793 9	0,796 7	0,799 5	0,802 3	0,805 1	0,807 8	0,810 6	0,813 3
0,9	0,815 9	0,818 6	0,821 2	0,823 8	0,826 4	0,828 9	0,831 5	0,834 0	0,836 5	0,838 9
1,0	0,841 3	0,843 8	0,846 1	0,848 5	0,850 8	0,853 1	0,855 4	0,857 7	0,859 9	0,862 1
1,1	0,864 3	0,866 5	0,868 6	0,870 8	0,872 9	0,874 9	0,877 0	0,879 0	0,881 0	0,883 0
1,2	0,884 9	0,886 9	0,888 8	0,890 7	0,892 5	0,894 4	0,896 2	0,898 0	0,899 7	0,901 5
1,3	0,903 2	0,904 9	0,906 6	0,908 2	0,909 9	0,911 5	0,913 1	0,914 7	0,916 2	0,917 7
1,4	0,919 2	0,920 7	0,922 2	0,923 6	0,925 1	0,926 5	0,927 9	0,929 2	0,930 6	0,931 9
1,5	0,933 2	0,934 5	0,935 7	0,937 0	0,938 2	0,939 4	0,940 6	0,941 8	0,942 9	0,944 1
1,6	0,945 2	0,946 3	0,947 4	0,948 4	0,949 5	0,950 5	0,951 5	0,952 5	0,953 5	0,954 5
1,7	0,955 4	0,956 4	0,957 3	0,958 2	0,959 1	0,959 9	0,960 8	0,961 6	0,962 5	0,963 3
1,8	0,964 1	0,964 9	0,965 6	0,966 4	0,967 1	0,967 8	0,968 6	0,969 3	0,969 9	0,970 6
1,9	0,971 3	0,971 9	0,972 6	0,973 2	0,973 8	0,974 4	0,975 0	0,975 6	0,976 1	0,976 7
2,0	0,977 2	0,977 8	0,978 3	0,978 8	0,979 3	0,979 8	0,980 3	0,980 8	0,981 2	0,981 7
2,1	0,982 1	0,982 6	0,983 0	0,983 4	0,983 8	0,984 2	0,984 6	0,985 0	0,985 4	0,985 7
2,2	0,986 1	0,986 4	0,986 8	0,987 1	0,987 5	0,987 8	0,988 1	0,988 4	0,988 7	0,989 0
2,3	0,989 3	0,989 6	0,989 8	0,990 1	0,990 4	0,990 6	0,990 9	0,991 1	0,991 3	0,991 6
2,4	0,991 8	0,992 0	0,992 2	0,992 5	0,992 7	0,992 9	0,993 1	0,993 2	0,993 4	0,993 6
2,5	0,993 8	0,994 0	0,994 1	0,994 3	0,994 5	0,994 6	0,994 8	0,994 9	0,995 1	0,995 2
2,6	0,995 3	0,995 5	0,995 6	0,995 7	0,995 9	0,996 0	0,996 1	0,996 2	0,996 3	0,996 4
2,7	0,996 5	0,996 6	0,996 7	0,996 8	0,996 9	0,997 0	0,997 1	0,997 2	0,997 3	0,997 4
2,8	0,997 4	0,997 5	0,997 6	0,997 7	0,997 7	0,997 8	0,997 9	0,997 9	0,998 0	0,998 1
2,9	0,998 1	0,998 2	0,998 2	0,998 3	0,998 4	0,998 4	0,998 5	0,998 5	0,998 6	0,998 6

Quelques valeurs supplémentaires

x	$\phi(x)$
2,99	0,998605
3,00	0,998650
3,50	0,999767
4,00	0,999968
4,50	0,999997
5,00	1,000000

Exemple de calcul.

Soit X une v.a.r. de loi normale $\mathcal{N}(0,1)$. On a :

$P(X \leq 1,23) = \phi(1,23) = 0,8907$

$P(-0,06 \leq X \leq 1,23) = \phi(1,23) - \phi(-0,06) = \phi(1,23) - (1 - \phi(0,06)) = \phi(1,23) + \phi(0,06) - 1$
$= 0,8907 + 0,5239 - 1 = 0,4146$.

1.4.2. Loi normale $\mathcal{N}(\mu,\sigma)$, avec μ réel et $\sigma > 0$.

On a $f_X(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ pour tout $x \in \mathbb{R}$. Le paramètre μ est un paramètre de localisation (point où f_X atteint son maximum), et le paramètre σ un paramètre d'échelle, caractérisant l'applatissement de la courbe en cloche représentative de f_X.

Proposition.

X suit la loi normale $\mathcal{N}(\mu,\sigma)$ si et seulement si $Y = \dfrac{X-\mu}{\sigma}$ suit la loi normale $\mathcal{N}(0,1)$.

On a alors $E(X) = \mu$ et $V(X) = \sigma^2$. Ainsi, μ est la moyenne de X et σ est l'écart-type de X.

Exemple de calcul.

Soit X une v.a.r. de loi normale $\mathcal{N}(2,10)$. Alors $Y = \dfrac{X-2}{10}$ suit la loi normale $\mathcal{N}(0,1)$. On a

$P(X \leq 14,3) = P\left(\dfrac{X-2}{10} \leq \dfrac{14,3-2}{10}\right) = P(Y \leq 1,23) = \phi(1,23) = 0,8907$

$P(1,4 \leq X \leq 14,3) = P\left(\dfrac{1,4-2}{10} \leq \dfrac{X-2}{10} \leq \dfrac{14,3-2}{10}\right) = P(-0,06 \leq Y \leq 1,23)$
$= \phi(1,23) - \phi(-0,06) = \cdots = 0,4146$.

1.4.3. Approximation de la loi Binomiale par la loi Normale

Soit X une variable aléatoire de loi Binomiale $\mathcal{B}(n,p)$. Si n est "grand" et p "pas trop petit", alors X suit approximativement la loi Normale $\mathcal{N}\left(np, \sqrt{np(1-p)}\right)$ (de même moyenne et même écart-type que la Binomiale).

En pratique, ce résultat s'applique dès que $np \geq 10$ et $n(1-p) \geq 10$. (Autres conditions possibles : $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$).

2. Exercices

Exercice 1.

Le tableau suivant présente la loi de probabilité de la variable aléatoire X qui à chaque année associe les bénéfices d'une société (x_i= bénéfice en milliers d'euros) pour cette année.

x_i	-100	0	50	100	150	200
p_i	$0,10$	$0,20$	$0,20$	$0,25$	$0,10$	p_6

1) Quelle est la valeur de p_6 ? Comment peut-on interpréter cette valeur ?
2) Calculer la probabilité que cette société réalise des bénéfices au cours d'une année.
3) Calculer la probabilité que cette société réalise un bénéfice d'au moins 100 000 euros en une année.

Exercice 2.

Considérons les ventes chez un concessionnaire automobile. Les données de vente journalière sur une durée de 300 jours sont les suivantes : au cours de 54 jours, aucune automobile n'a été vendue ; au cours de 117 jours, une automobile a été vendue chaque jour ; au cours de 72 jours, 2 automobile ont été vendues chaque jour ; au cours de 42 jours, 3 automobiles ont été vendues chaque jour ; au cours de 12 jours, 4 au tomobiles ont été vendues chaque jour ; au cours de 3 jours, 5 automobiles ont été vendues chaque jour.

1) On désigne par X la variable aléatoire qui à chaque jour associe le nombre d'automobiles vendues ce jour là. Quelles valeurs peut prendre la variable aléatoire X ? Est-elle discrète ou à densité ?
2) Construire le tableau de la loi de probabilité de X. Les conditions requises pour l'existence d'une loi de probabilité discrète sont-elles respectées ?
3) Représenter graphiquement cette loi de probabilité.
4) Calculer l'espérance mathématique et la variance de X.

Exercice 3.

Un pâtissier vend des gâteaux au chocolat. Chaque jour, un certain nombre de clients lui en achète un. Ce nombre est modélisé par une variable aléatoire X dont la loi de probabilité est donnée par :

x_i	1	2	3	4	5	6
p_i	$0,1$	$0,1$	$0,3$	p_4	p_5	$0,1$

1) Déterminer les valeurs de p_4 et p_5, sachant que les événements $(X = 4)$ et $(X = 5)$ sont équiprobables.

2) Calculer l'espérance mathématique et l'écart-type de X.

3) Déterminer et représenter graphiquement la fonction de répartition de X.

4) Un gateau est vendu 30 euros pièce. Le coût de fabrication d'un gateau pour le pâtissier est de 10 euros. Un gateau invendu en fin de journee représente une perte nette car le pâtissier ne vend que des gateaux ultra-frais.

a) Ce matin, le pâtissier a fabriqué 3 gâteaux. Quelle est l'espérance de son gain ?

b) Combien de gâteaux doit-il fabriquer pour maximiser l'espérance de son gain ?

Exercice 4.

Une société informatique envisage l'extension de l'usine afin de pouvoir commencer la production d'un nouvel ordinateur. Le président de la société doit déterminer si l'extension doit être réalisée à moyenne ou grande échelle. La demande pour le nouveau produit est incertaine ; elle peut être faible, moyenne ou élevée. Les estimations en probabilité pour la demande sont respectivement 0.20, 0.50 et 0.30. Soit X le profit annuel en milliers d'euros. Les prévisionnistes de la firmes ont développé les prévisions de profits suivantes pour les projets d'expansion à moyenne et à grande échelle.

		Expansion à moyenne échelle		Expansion à grande échelle	
		x_i	p_i	x_i	p_i
Demande	Faible	50	0,20	0	0,20
	Moyenne	150	0,50	100	0,50
	Elevée	200	0,30	300	0,30

1) Calculer l'espérance mathématique du profit pour les deux alternatives d'expansion. Quelle décision est préférable en termes de maximisation du profit ?

2) Calculer la variance du profit pour les deux alternatives d'expansion. Quelle décision est préférable en termes de minimisation des risques ou de l'incertitude ?

Exercice 5.

Considérons les décisions d'achat des trois prochains clients qui entrent dans un magasin de prêt-à-porter. Sur la base de son expérience passée, le gérant du magasin estime la probabilité qu'un client fasse un achat à 0.40. On cherche à calculer la probabilité que deux des trois clients suivants fassent un achat. Pour cela, on considère la variable aléatoire X égale au nombre de clients qui vont faire un achat parmi les trois prochains clients

1) Justifier que X suit une loi binomiale dont on précisera les paramètres.

2) En déduire la probabilité que deux des trois prochains clients fassent un achat ?

Exercice 6.

Supposons que, dans un groupe de 10 individus, 6 préfèrent Coca et 4 Pepsi. On sélectionne un échantillon aléatoire de trois de ces individus.

1) Quelle est la probabilité que deux individus préfèrent Coca ?

2) Quelle est la probabilité qu'une majorité (2 ou 3 individus) préfère Pepsi ?

Exercice 7.

Dans une fabrication en série, 8% des articles présentent des défauts.

1) Quelle est la probabilité pour que dans un contrôle portant sur 40 articles, il y ait 4 articles défectueux ? Quelle est la probabilité pour qu'il y ait 4 articles défectueux au plus ?

2) Quelle est la probabilité pour qu'il y ait 6 articles défectueux sur un lot de 100 ?

Exercice 8.

Une cargaison de 10 unités comporte de ux unités défectueuses et huit unités non défectueuses. Lors de l'inspection de la cargaison, un échantillon aléatoire d'unités est sélectionné et testé. Si une unité défectueuse est trouvée, la cargaison des 10 unités sera refusée.

1) Si un échantillon de 3 unités est sélectionné, quelle est la probabilité que la cargaison soit refusée ?

2) Si un échantillon de 4 unités est sélectionné, quelle est la probabilité que la cargaison soit refusée ?

3) Si un échantillon de 5 unités est sélectionné, quelle est la probabilité que la cargaison soit refusée ?

4) Si la direction désire avoir une probabilité de rejet d'une cargaison, qui comprend deux unités défectueuses sur 10, égale à 0,90, quelle taille d'échantillon peut-on lui recommander ?

Exercice 9.

Dans un atelier, le nombre d'accidents au cours d'une année est une variable aléatoire qui suit une loi de Poisson de paramètre 5. Calculer la probabilité des événements suivants :

1) Il n'y a pas d'accidents au cours d'une année.

2) Il y a exactement 4 accidents au cours d'une année.

3) Il a plus de 6 accidents au cours d'une année.

Exercice 10.

On suppose que le pourcentage de gauchers dans une population est de 1%. Soit X la variable aléatoire prenant comme valeur le nombre de gauchers dans un echantillon de 200 personnes choisies au hasard.

1) Justifier que X suit approximativement une loi de Poisson dont on précisera la moyenne et la variance

2) Quelle est la probabilité pour qu'il y ait moins de 4 gauchers dans l'échantillon ?

Exercice 11.

Les appels téléphoniques arrivent à un taux de 60 par heure au bureau des réservations d'une compagnie aérienne.

1) On désigne par X la variable aléatoire égale au nombre d'appels reçus dans un intervalle de 5 minutes, soit 300 secondes.

a) Justifier que l'on peut considérer que X suit une loi de Poisson de paramètre 5.

b) Calculer la probabilité de recevoir trois appels dans un intervalle de 5 minutes.

2) Calculer la probabilité de recevoir exactement 10 appels en 15 minutes.

3) Supposons qu'il n'y ait aucun appel en attente pour le moment. Si l'agent met cinq minutes pour répondre à l'appel en cours, combien de personnes attendront pendant ce temps en moyenne ? Quelle est la probabilité que personne n'attende ?

4) S'il n'y a aucun appel en cours, quelle est la probabilité que l'agent puisse prendre trois minutes de repos sans être dérangé ?

Exercice 12.

Tous les jours, Guy joue à un jeu en ligne sur un site, avec trois amis. Lorsqu'il se connecte sur le site, la durée X (en seconde) qu'il faut pour réunir les quatre joueurs est une variable aléatoire qui suit une loi uniforme sur l'intervalle $[20; 120]$.

1) Déterminer la probabilité que les quatre joueurs soient réunis au bout de 60 secondes.

2) Calculer l'espérance mathématique de D. Interpréter ce résultat.

Exercice 13.

On modélise le temps d'attente, exprimé en minutes, à un guichet, par une variable aléatoire X qui suit la loi exponentielle de paramètre $\lambda = 0,7$.

1) Calculer la probabilité qu'un client attende au plus cinq minutes à ce guichet.

2) Calculer la probabilité qu'un client attende au moins cinq minutes à ce guichet

3) Calculer le temps d'attente moyen à ce guichet.

Exercice 14.

Une entreprise de jouets en peluche souhaite commercialiser un nouveau produit et à cette fin, effectue divers tests permettant de rejeter les peluches ne répondant pas aux normes en vigueur.

On considère que la vie d'une peluche se termine lorsqu'elle subit un dommage majeur (déchirure, arrachage ...). On admet que la durée de vie en années d'une peluche, notée X, suit une loi exponentielle de paramètre λ.

1) On sait que $P(X \leq 4) = 0,5$. Interpréter ce résultat dans le contexte de cet exercice.

2) En déduire la valeur exacte de λ.

Exercice 15.

1) Soit X une variable aléatoire de loi normale $\mathcal{N}(0;1)$.

a) Calculer $P(X \leq 2)$ et $P(0 \leq X \leq 1)$.

b) Déterminer le réel x tel que $P(X \leq x) = 0,975$.

2) Soit X une variable aléatoire de loi normale $\mathcal{N}(4;2)$.

a) Calculer $P(X \leq 6)$ et $P(0 \leq X \leq 6)$.

b) Déterminer le réel x tel que $P(X \leq x) = 0,975$.

Exercice 16.

Un bar débite la bière en chopes dont le contenu effectif est une variable aléatoire X supposée gaussienne de moyenne $\mu = 25$ cl et d'écart-type $\sigma = 2$ cl.

Déterminer la probabilité que votre chope de bière contienne :

a) moins de 23 cl de bière.

b) plus de 26 cl de bière.

c) entre 24 et 26 cl de bière.

Exercice 17.

Soit X une variable aléatoire de loi normale $\mathcal{N}(\mu;\sigma)$. Calculer $P(X \leq \mu + \sigma)$, $P(\mu - \sigma \leq X \leq \mu + \sigma)$, $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma)$ et $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma)$.

Exercice 18.

Une machine remplit automatiquement des boîtes de sucre en poudre de telle façon que le poids de sucre effectivement contenu dans une boîte soit une variable aléatoire normale de paramètres μ et σ exprimés en grammes. On souhaite régler la machine de sorte que le poids de sucre contenu dans une boîte dépasse 980 grammes avec une probabilité égale à 0,95.

1) Lorsque $\sigma = 30$, quelle valeur faut-il donner à la moyenne μ ?

2) Lorsque $\mu = 1000$, quelle valeur doit avoir l'écart-type σ ?

Exercice 19.

Une société envisage la mise en place de nouveaux équipements. Dans le cadre de ce projet, elle a défini trois tâches A, B et C. On sait que la tâche A dure 10 semaines et que les tâches B et C ont des durées aléatoires indépendantes ; la durée de B obéissant à la loi normale de moyenne 18 et d'écart-type 4, la durée de C obéissant à la loi normale de moyenne 20 et d'´ecart-type 5.

1) Quelle est la probabilité que la tâche B (resp. C) dure entre 18 et 25 semaines ?

2) Quelle est la probabilité que la durée totale D des trois tâches ne dépasse pas 55 semaines ?

3) Déterminer un intervalle centré sur l'espérance mathématique dans lequel se trouve la durée totale avec une probabilité de 0,75.

Exercice 20.

La demande mensuelle d'un produit obéit à une loi normale. Elle a une probabilité 0,1 d'être inférieure à 15 000 unités (resp. supérieure à 25 000 unités).

1) Quels sont les paramètres de cette loi ?

2) Calculer la probabilité qu'en un mois la demande dépasse 23 000 unités.

3) Quel doit être le stock pour ne risquer une rupture qu'avec une probabilité d'environ 0.001 ?

Exercice 21.

Un balladeur MP3 fabriqué par la compagnie Multisonic est garanti contre tout défaut de fabrication pour une période de 2 ans. D'après l'expérience de la compagnie, les chances d'observer une non-conformité majeure durant les 26 mois (respectivement 52 mois) suivant l'achat sont de 1 sur 100 (respectivement 975 sur 1000). Supposons que le temps X requis après l'achat pour qu'une non-conformité majeure se présente soit distribué normalement.

1) Déterminer les paramètres de cette gaussienne.

2) Quelle est la probabilité que l'appareil présente une non-conformité majeure avant la fin de la période de garantie ?

3) Quelle devrait être la période de garantie si Multisonic ne souhaitait remplacer que 0,05% des appareils vendus ?

Exercice 22.

Un fabricant souhaite lancer une nouvelle console de jeu pour Noël. Les études marketing montrent que parmi les 2000 joueurs de la région, 40% ont déclaré avoir l'intention d'acheter le jeu. On appelle X la variable aléatoire égale au nombre de personnes allant effectivement acheter le jeu.

1) Quelle est la loi de X ?

2) En approchant la loi de X par une loi normale dont on précisera les paramètres, déterminer le stock que doit avoir le magasin pour que la probabilité de rupture de stock soit inférieure a 0,1.

Exercice 23.

Une enquête marketing a pour but de vérifier si les cibles potentielles seraient tentées par un nouveau produit. Il a été montré que 56% des gens sont favorables au nouveau produit.

1) Pour aller plus loin, on interroge à nouveau 200 personnes. Quelle est la loi du nombre de clients potentiels parmi les 200 ? Par quelle loi peut-on l'approcher ? Calculer $P(X \geq 100)$ et $P(100 \leq X \leq 150)$.

2) On souhaite maintenant demander des précisions à un grand nombre de personnes favorables au produit, disons 100 personnes. Déterminer la taille n de l'échantillon de personnes à interroger pour que l'échantillon contienne au moins 100 personnes favorables, avec une probabilité supérieure ou égale à 95%.