

Licence mention Mathématiques - Semestre 3
Statistique

Rappels sur les variables aléatoires réelles discrètes et à densité

1. Variables aléatoires réelles discrètes

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.

Exemple : le nombre d'enfants X d'une famille d'une population Ω de 1000 familles amiénoises a donné :

Nombre d'enfants x_i	0	1	2	3	4	5	6	7	8
Nombre de familles n_i	162	240	297	208	62	16	8	5	2
Fréquence f_i	0.162	0.240	0.297	0.208	0.062	0.016	0.008	0.005	0.002

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 , ce qui conduit au tableau des couples (x_i, p_i) où x_i est une valeur observée et p_i sa probabilité.

Prenant $\mathcal{A} = \mathcal{P}(\Omega)$ et P l'équiprobabilité sur $(\Omega, \mathcal{A})$, on écrira par exemple $P(X = 0) = \frac{162}{1000} = 0.162$.

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).

1.1. Cas d'un nombre fini de valeurs

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 pourra alors considérer l'événement $(X = 1) = \{1, 3, 5\}$ et écrire $P(X = 1) = 1/2$.

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

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.

Les événements $\{\omega \in \Omega / X(\omega) = x\}$, $\{\omega \in \Omega / X(\omega) \leq x\}$ et $\{\omega \in \Omega / a < X(\omega) \leq x\}$ 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 $\Omega_X = \{x_1, \dots, x_n\}$, 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 $\sum_{i=1}^n p_i = p_1 + p_2 + \dots + p_n = 1$.

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

Fonction de répartition.

Si X est une v.a.r., alors on appelle fonction de répartition de X la fonction F_X de $\mathbb{R}$ dans $[0, 1]$ définie par, pour tout réel x , $F_X(x) = P(X \leq x)$.

Si X est discrète finie, alors F_X est une fonction en escalier dont les sauts se présentent aux points x_i de Ω_X , chaque saut étant égal à $p_i = P(X = x_i)$.

Exemple. Reprenons l'exemple introductif. On a :

- pour tout $x < -10$, $F_X(x) = P(X \leq x) = P(\emptyset) = 0$;
- pour tout $-10 \leq x < 1$, $F_X(x) = P(X \leq x) = P(X = -10) = 1/6$;
- pour tout $1 \leq x < 5$,
 $F_X(x) = P(X \leq x) = P((X = -10) \cup (X = 1)) = P(X = -10) + P(X = 1) = 1/6 + 1/2 = 2/3$;
- pour tout $x \geq 5$, $F_X(x) = P(X \leq x) = P(\Omega) = 1$.

Remarque. On a donc $F(x_i) = P(X \leq x_i) = \sum_{k=1}^i p_k = p_1 + p_2 + \dots + p_i$, probabilité que X prenne une valeur inférieure ou égale à x_i .

En statistique descriptive des variables quantitatives discrètes, les fréquences cumulées étaient données par : $F_i = \sum_{k=1}^i f_k = f_1 + f_2 + \dots + f_i$, et représentaient la fréquence des valeurs inférieures ou égales à x_i .

Interprétant les probabilités p_i comme les limites des fréquences f_i sur un échantillon, on observe que les formules et leur interprétation sont analogues.

Espérance mathématique. Variance. Ecart-type.

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

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

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

On a $V(X) = E(X^2) - (E(X))^2 = \sum_{i=1}^n p_i x_i^2 - (E(X))^2$.

On appelle écart-type de X le réel $\sigma(X) = \sqrt{V(X)}$.

On appelle intervalle moyen de X l'intervalle $[E(X) - \sigma(X), E(X) + \sigma(X)]$.

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 les formules suivantes.

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 : $s_x = \sqrt{s_x^2}$.

Interprétant les p_i comme les limites des f_i , les formules et leur interprétation sont analogues.

Changement d'origine et d'échelle

Il arrive souvent que l'on effectue une transformation affine $X \mapsto aX + b$ sur une variable aléatoire X , a étant le paramètre de changement d'échelle, et b le paramètre de changement d'origine. On a :

$$E(aX + b) = aE(X) + b ; \quad \text{Var}(aX + b) = a^2 \text{Var}(X) ; \quad \sigma(aX + b) = |a| \sigma(X).$$

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

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 $k \in \{0, 1, \dots, n\}$,

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

Exemples.

1) On lance 10 fois de suite une pièce de monnaie équilibrée. Quelle est la probabilité d'obtenir 3 fois (exactement) Pile ?

On peut considérer que l'on répète $n = 10$ fois dans les mêmes conditions l'expérience aléatoire "lancer la pièce" (les répétitions étant indépendantes entre elles) au cours de laquelle l'événement A "obtenir Pile" a la probabilité $p = \frac{1}{2}$ d'être réalisé. La variable aléatoire X , égale au nombre de fois où l'événement A est réalisé, i.e. au nombre de Pile obtenus, suit la loi Binomiale $\mathcal{B}(n, p) = \mathcal{B}\left(10, \frac{1}{2}\right)$, ce qui signifie que :

X est à valeurs dans $\Omega_X = \{0, 1, \dots, n\} = \{0, 1, \dots, 10\}$ et pour tout $k \in \{0, 1, \dots, 10\}$,

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{10}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{10-k} = \binom{10}{k} \left(\frac{1}{2}\right)^{10}.$$

On en déduit que la probabilité d'obtenir 3 fois Pile est $P(X = 3) = \binom{10}{3} \left(\frac{1}{2}\right)^{10} \simeq 0,1172$.

2) On tire successivement 4 boules d'une urne contenant 3 boules blanches et 7 boules noires. Quelle est la probabilité d'obtenir (exactement) 2 boules blanches ?

On peut considérer que l'on effectue $n = 4$ tirages d'un individu dans la population des $N = 10$ boules, avec $N_1 = 3$ boules blanches et $N_2 = 7$ boules noires (on a $N_1 + N_2 = N$). Si on désigne par X la variable aléatoire égale au nombre de boules blanches obtenues, alors :

- dans le cas de tirages avec remise, X suit la loi Binomiale $\mathcal{B}\left(n, \frac{N_1}{N}\right) = \mathcal{B}\left(4, \frac{3}{10}\right)$; ainsi, X est à valeurs dans $\Omega_X = \{0, 1, \dots, n\} = \{0, 1, \dots, 4\}$ et pour tout $k \in \{0, 1, \dots, 4\}$,

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{4}{k} \left(\frac{3}{10}\right)^k \left(\frac{7}{10}\right)^{4-k},$$
 donc $P(X = 2) = \binom{4}{2} \left(\frac{3}{10}\right)^2 \left(\frac{7}{10}\right)^2 \simeq 0,2646$.

- dans le cas de tirages sans remise, X suit la loi Hypergéométrique $\mathcal{H}\left(N, n, \frac{N_1}{N}\right) = \mathcal{H}\left(10, 4, \frac{3}{10}\right)$; ainsi, X est à valeurs dans $\Omega_X = \{\max(0, n - (N - N_1)), \min(N_1, n)\} = \{0, 1, 2, 3\}$ et pour tout $k \in \{0, 1, 2, 3\}$,

$$P(X = k) = \frac{\binom{N_1}{k} \binom{N-N_1}{n-k}}{\binom{N}{n}} = \frac{\binom{3}{k} \binom{7}{4-k}}{\binom{10}{4}},$$
 donc $P(X = 2) = \frac{\binom{3}{2} \binom{7}{2}}{\binom{10}{4}} = 0,3$.

Remarque.

Lorsque N est très grand devant n (en pratique $N > 10n$), on peut approcher la loi Hypergéométrique $\mathcal{H}\left(N, n, \frac{N_1}{N}\right)$ par la loi Binomiale $\mathcal{B}\left(n, \frac{N_1}{N}\right)$. Cela signifie que

$$\frac{\binom{N_1}{k} \binom{N-N_1}{n-k}}{\binom{N}{n}} \simeq \binom{n}{k} \left(\frac{N_1}{N}\right)^k \left(1 - \frac{N_1}{N}\right)^{n-k}.$$

1.2. Cas d'un nombre infini 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 $\Omega_X = \mathbb{N} = \{k / k \in \mathbb{N}\}$ 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 dans ce cours 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 réserve d'existence des sommes indiquées ci-dessous, on a :

- espérance mathématique de X : $E(X) = \sum_{k \in \Omega_X} kp_k$.

- 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 $k \in \mathbb{N}$, $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$.

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 : $E(X) = \frac{1}{p}$ et $V(X) = \frac{1-p}{p^2}$.

Remarque.

Lorsque n est grand et p assez petit (en pratique $n \geq 30$, $p \leq 0,1$ et $np \leq 10$), on peut approcher la loi Binomiale $\mathcal{B}(n, p)$ par la loi de Poisson $\mathcal{P}(np)$. Cela signifie que $\binom{n}{k} p^k (1-p)^{n-k} \simeq e^{-np} \frac{(np)^k}{k!}$.

2. Variables aléatoires réelles à densité

2.1. Généralités et premiers exemples

On considère une expérience aléatoire, $(\Omega, \mathcal{A}, P)$ un espace probabilisé adapté et X une variable aléatoire réelle.

On a traité dans le paragraphe 1 le cas où l'univers-image Ω_X (et donc l'univers Ω) était fini ou infini dénombrable. La loi de probabilité de X était alors donnée par les quantités $P(X = x_i)$ pour tout x_i de Ω_X .

Lorsque Ω_X (et donc Ω) 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 sur $\mathbb{R}$ 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 $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$. De plus, pour tous réels x , a et b tels que $a < b$, on a :

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

Espérance mathématique. Variance. Ecart-type.

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

- $E(X) = \int_{-\infty}^{+\infty} xf_X(x)dx$ (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

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

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

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

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

Loi Normale de paramètres μ et σ , avec $\mu \in \mathbb{R}$ et $\sigma > 0$. Voir le paragraphe suivant.

2.2. Loi Normale $\mathcal{N}(\mu, \sigma)$.**2.2.1. Cas de la loi $\mathcal{N}(0, 1)$ (loi normale centrée réduite).**

On a $\Omega_X = \mathbb{R}$ et $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ 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 $\phi(x) = F_X(x) = \int_{-\infty}^x f_X(t)dt = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ 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). Pour les $x < 0$, on peut utiliser la formule $\phi(-x) = 1 - \phi(x)$ (formule valable pour tout réel x)

La table 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 $x \geq 3$.

On peut démontrer que ϕ est une bijection strictement croissante de $\mathbb{R}$ sur $]0; 1[$. Ainsi, pour tous réels x et y , on a : $\phi(x) = \phi(y) \Leftrightarrow x = y$ et $\phi(x) < \phi(y) \Leftrightarrow x < y$.

TABLE 1**Fonction de répartition
de la loi normale réduite**

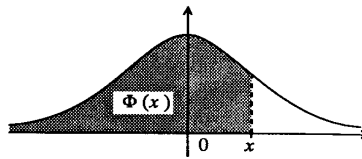
Si U suit la loi normale réduite, pour $x \geq 0$, la table donne la valeur :

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

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

2.2.2. Loi normale $\mathcal{N}(\mu, \sigma)$, avec $\mu \in \mathbb{R}$ et $\sigma > 0$.

On a $\Omega_X = \mathbb{R}$ et $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$ 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'aplatissement de la courbe en cloche représentative de f_X .

Proposition.

X suit la loi normale $\mathcal{N}(\mu, \sigma)$ si et seulement si $Y = \frac{X - \mu}{\sigma}$ 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 $Y = \frac{X - 2}{\sqrt{10}}$ suit la loi normale $\mathcal{N}(0, 1)$. On a :

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

$$P(1,4 \leq X \leq 14,3) = P\left(\frac{1,4 - 2}{\sqrt{10}} \leq \frac{X - 2}{\sqrt{10}} \leq \frac{14,3 - 2}{\sqrt{10}}\right) = P(-0,06 \leq Y \leq 1,23)$$

$$= \phi(1,23) - \phi(-0,06) = \dots = 0,4146.$$

2.2.3. 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 $\mathcal{N}(np, \sqrt{np(1-p)})$ (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 $n(1-p) \geq 5$).

Exemple.

On jette $n = 12000$ fois un dé équilibré. On cherche la probabilité que le nombre de 6 obtenus soit compris entre 1800 et 2100. Le nombre X de 6 obtenus suit une loi Binomiale $\mathcal{B}(n, p)$ avec $n = 12000$ et $p = \frac{1}{6}$. Comme $np = 2000 \geq 10$ et $n(1-p) = 10000 \geq 10$, on peut utiliser l'approximation par la loi

Normale $\mathcal{N}(np, \sqrt{np(1-p)}) = \mathcal{N}\left(2000, \sqrt{12000 \frac{1}{6} \frac{5}{6}}\right) = \mathcal{N}\left(2000, \frac{100}{\sqrt{6}}\right)$:

Ainsi, $Z = \frac{X-2000}{\frac{100}{\sqrt{6}}}$ suit approximativement la loi Normale $\mathcal{N}(0, 1)$. On a alors :

$$P(1800 \leq X \leq 2100) = P\left(\frac{1800-2000}{\frac{100}{\sqrt{6}}} \leq \frac{X-2000}{\frac{100}{\sqrt{6}}} \leq \frac{2100-2000}{\frac{100}{\sqrt{6}}}\right) = P(-4,90 \leq Z \leq 2,45)$$

$$\simeq \phi(2,45) - \phi(-4,90) = \phi(2,45) + \phi(4,90) - 1 \simeq 0,9929 + 1 - 1 = 0,9929.$$

Le calcul direct serait possible mais plus fastidieux :

$$P(1800 \leq X \leq 2100) = \sum_{k=1800}^{2100} P(X = k) = \sum_{k=1800}^{2100} \binom{12000}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{12000-k} \simeq 0,993.$$

Correction de continuité

Approchant une loi discrète (à valeurs entières) par une loi continue (avec valeurs réelles), il est utile de distinguer inégalités stricte et large. Pour se faire, on effectue une correction de continuité. Ainsi, pour approcher $P(X = k)$, on effectue le calcul suivant :

$$P(X = k) = P\left(k - \frac{1}{2} < X \leq k + \frac{1}{2}\right) \simeq \phi\left(\frac{\left(k + \frac{1}{2}\right) - np}{\sqrt{np(1-p)}}\right) - \phi\left(\frac{\left(k - \frac{1}{2}\right) - np}{\sqrt{np(1-p)}}\right).$$

On a alors, pour tous entiers m et m' :

$$P(m \leq X \leq m') = P\left(m - \frac{1}{2} < X \leq m' + \frac{1}{2}\right) \simeq \phi\left(\frac{\left(m' + \frac{1}{2}\right) - np}{\sqrt{np(1-p)}}\right) - \phi\left(\frac{\left(m - \frac{1}{2}\right) - np}{\sqrt{np(1-p)}}\right).$$

Le calcul de l'exemple précédent est alors légèrement modifié.

Exemple.

Reprenons l'exemple précédent. On a :

$$P(1800 \leq X \leq 2100) = P\left(1800 - \frac{1}{2} \leq X \leq 2100 + \frac{1}{2}\right) = P(1799,5 \leq X \leq 2100,5)$$

$$= P\left(\frac{1799,5 - 2000}{\frac{100}{\sqrt{6}}} \leq \frac{X - 2000}{\frac{100}{\sqrt{6}}} \leq \frac{2100,5 - 2000}{\frac{100}{\sqrt{6}}}\right) = P(-4,91 \leq Z \leq 2,46)$$

$$\simeq \phi(2,46) - \phi(-4,91) = \phi(2,46) + \phi(4,91) - 1 \simeq 0,9931 + 1 - 1 = 0,9931.$$

3. Exercices

Exercice 1.

- 1) Un joueur a une chance sur trois de gagner une partie. Il joue cinq parties. Calculer la probabilité qu'il gagne : a) trois parties. b) cinq parties. c) au plus une partie. d) au moins deux parties.
- 2) Dans une production de pièces très nombreuse, il y a 3% de pièces défectueuses. Soit X la variable aléatoire égale au nombre de pièces défectueuses dans un échantillon de 10 pièces prélevées au hasard.
- Quelle est la loi de probabilité de X . Expliquer.
 - En déduire le nombre moyen de pièces défectueuses obtenu dans un lot de 10 pièces.
 - Calculer la probabilité d'obtenir au moins deux pièces défectueuses.

Exercice 2.

Une entreprise fabrique des conserves alimentaires dont l'étiquette annonce une masse de 250 grammes. Les masses obtenues pour un échantillon de 500 conserves prises au hasard (dans la production) sont données dans le tableau suivant :

Masse (en g)	]235; 240]	]240; 245]	]245; 250]	]250; 255]	]255; 260]
Nombre de conserves	35	65	215	135	50

- Pour la présentation des différents calculs effectués dans les questions suivantes, on construira dès le début un unique tableau présentant l'ensemble des résultats demandés ou utiles.
 - Préciser la population étudiée, la variable étudiée et sa nature, la taille de l'échantillon.
 - Construire l'histogramme des effectifs correspondant aux résultats présentés ci-dessus.
 - Donner la moyenne et l'écart-type de la distribution ; préciser les données utilisées pour le calcul.
- Dans l'échantillon de 500 conserves, on choisit successivement, au hasard et avec remise, 30 conserves. On désigne par X la variable aléatoire qui à un tel prélèvement associe le nombre de conserves de masse inférieure ou égale à 250 g.
 - Justifier que X suit une loi binomiale dont on précisera les paramètres.
 - Calculer l'espérance mathématique $E(X)$ de X . Interpréter ce résultat par une phrase.
 - Calculer $P(X = 15)$ et $P(X = 20)$ (arrondir au millième). Interpréter à l'aide d'une phrase.
- On admet que la variable aléatoire Y qui à chaque conserve produite associe sa masse (en g) suit la loi normale de moyenne 250 et d'écart type 5.
 - Calculer la probabilité qu'une conserve prise au hasard ait une masse comprise entre 245 g et 255 g.
 - Déterminer le nombre réel a tel que $P(250 - a \leq Y \leq 250 + a) = 0,95$.

Exercice 3.

Une machine fabrique un très grand nombre de pièces d'un même modèle. On admet que le pourcentage de pièces conformes dans la production d'une journée est de 90%. On prélève au hasard un lot de 200 pièces dans la production pour vérification de l'épaisseur. La production est suffisamment importante pour que l'on puisse assimiler ce prélèvement à un tirage avec remise. On désigne par X la variable aléatoire associant à chaque lot de 200 pièces le nombre de pièces non conformes du lot.

- Justifier que Y suit une loi binomiale et préciser ses paramètres.
- Calculer la probabilité qu'il y ait exactement 20 pièces non conformes dans ce lot.
- Justifier que l'on peut approcher la loi de X par une loi normale dont on précisera les paramètres, puis calculer une valeur approchée de la probabilité demandée dans la question 2).

Exercice 4.

Dans une entreprise, le nombre d'appels reçus à chaque minute par le standard téléphonique est une variable aléatoire X de loi de Poisson de paramètre 0,7. Calculer la probabilité qu'entre 8h59 et 9h00, ce standard :

- ne reçoive aucun appel ;
- reçoive exactement un appel ;
- reçoive au moins un appel ;
- reçoive plus d'un appel ;
- reçoive entre 2 et 5 appels.

Exercice 5.

Les probabilités demandées ci-dessous seront données sous forme de fractions irréductibles.

Une entreprise de vente à distance communique sur la qualité et la ponctualité de la livraison des produits commandés sur son site web.

Un client a commandé un produit et la livraison est garantie pour le lundi 24 juin entre 10h et 11h. On suppose que le livreur peut se présenter à n'importe quel moment entre 10h et 11h. Le temps d'attente en minutes à partir de 10h peut donc être modélisé par une loi Uniforme sur l'intervalle $[0; 60]$.

- 1) Quelle est la probabilité que le livreur arrive avant 10h10 ?
- 2) a) Calculer la probabilité que le livreur arrive entre 10h20 et 10h40.
b) Calculer la probabilité que le livreur arrive entre 10h30 et 10h50. Quelle remarque peut-on faire ?
- 3) Sachant que le client attend toujours le livreur à 10h15, quelle est la probabilité que le livreur arrive dans les dix prochaines minutes ?

Exercice 6.

Une liaison aérienne entre deux villes est assurée par un avion de 140 places. La réservation est obligatoire. L'expérience a montré que la probabilité pour qu'une personne confirme sa réservation et retire son billet est égale à 0,8. On suppose que chaque personne se comporte indépendamment des autres.

La compagnie accepte n réservations, avec $n \geq 140$. On désigne par X la variable aléatoire égale au nombre de réservations confirmées.

- 1) Justifier que X suit une loi binomiale. Préciser son espérance et sa variance.
On admettra pour la suite de l'exercice que X suit la loi normale $\mathcal{N}(0,8n ; 0,4\sqrt{n})$.
- 2) On cherche le nombre n de réservations que la compagnie peut accepter sachant qu'elle s'accorde un risque de 5% de ne pas pouvoir satisfaire toutes les personnes ayant réservé.
 - a) Justifier que ce risque se traduit par l'inégalité $P(X \leq 140) \geq 0,95$.
 - b) Démontrer que n est solution de l'inéquation $0,8n + 0,658\sqrt{n} - 140 \leq 0$.
 - c) Quel est le nombre maximum de réservations acceptables ?