Lois de probabilité

Section de techniciens supérieurs
Informatique et réseaux pour l'industrie et les services techniques
Lycée Charles PONCET

Mars 2013

Table des matières

Le symbole ➡ indique les exemples à traiter, des démonstrations à trouver.
Le symbole ☛ indique des points importants, des pièges possibles, des notations particulières, etc.

1 Compléments sur les variables aléatoires discrètes

1.1 Fonction de répartition d'une variable aléatoire discrète

X est une variable aléatoire discrète définie sur l'univers fini Ω dont on connait la loi de probabilité.

Définition 1.1.1
On appelle fonction de répartition *de la variable aléatoire discrète* X *la fonction* F *définie par :*

$$\begin{aligned} F : \ & \mathbb{R} \longrightarrow \mathbb{R} \\ & x \longmapsto F(x) = P(X \leqslant x). \end{aligned}$$

Théorème 1.1.1 (Propriétés des fonctions de répartition des variables aléatoires discrètes)

i. Quel que soit $x \in \mathbb{R}$, $0 \leqslant F(x) \leqslant 1$.

ii. Quel que soit $x \in \mathbb{R}$, $P(X > x) = 1 - P(X \leqslant x) = 1 - F(x)$.

iii. Quels que soient $(x\ ;\ y) \in \mathbb{R}^2$, $P(x < X \leqslant y) = P(X \leqslant y) - P(X \leqslant x) = F(y) - F(x)$.

iv. La fonction F *est croissante.*

v. Si X *prend les valeurs* $x_1, x_2, \dots, x_n$ *avec* $x_1 < x_2 < \cdots < x_n$, *alors pour tout* $x < x_1$, $F(x) = 0$ *et pour tout* $x \geqslant x_n$, $F(x) = 1$.

⇒ Démontrer les cinq propriétés du théorème 1.1.1.

Remarques

- Une fonction de répartition d'une variable aléatoire discrète est une fonction en escalier.
- Si une variable aléatoire X prend les valeurs $x_1, x_2, \dots, x_n$, alors la fonction de répartition F de X est continue sur les intervalles $[x_i\ ;\ x_{i+1}[$ avec $i \in \{1\ ;\ 2\ ;\ \cdots\ ;\ n-1\}$, mais n'est pas continue en x_i (F est continue par morceaux).

1.2 Exemple

Une urne contient trois boules numérotées de 1 à 3, indiscernables au toucher. On procède à deux tirages successifs d'une boule, en remettant à chaque tirage la boule dans l'urne.

1. Définir l'univers Ω.
2. On s'intéresse à la somme des numéros inscrits sur les deux boules tirées, on définit ainsi une variable aléatoire X.
   Quelles sont les valeurs x_i prises par X ? Définir la loi de probabilité de X, c'est-à-dire, calculer les probabilités $P(X = x_i)$.
3. Calculer l'espérance mathématique, la variance et l'écart-type de X.
4. Déterminer et représenter la fonction de répartition de la variable aléatoire X.

2 Loi binomiale

2.1 Exemple

Une urne contient 2 boules rouges et 4 boules vertes indiscernables au toucher. On tire au hasard une boule que l'on remet dans l'urne après avoir noté sa couleur (on suppose qu'il y a équiprobabilité). On effectue 6 tirages.
On veut calculer la probabilité de tirer 4 boules rouges.

1. Le tirage d'une boule avec remise est une épreuve.
   Quelle est la probabilité de tirer une boule rouge pour une épreuve (évènement R) ?
   Quelle est la probabilité de ne pas tirer une boule rouge pour une épreuve (évènement $N = \overline{R}$) ?

$p = P(R)$ est la probabilité du *succès* et $q = 1 - p = P(N)$ est la probabilité de l'*échec*.

Cette épreuve est dite « *épreuve de* BERNOULLI [1] *de paramètre* p ».

2. On repète 6 fois cette épreuve de façon identique. Ces 6 épreuves sont des évènements indépendants.
Calculer la probabilité de la suite $(R\,;\ R\,;\ N\,;\ R\,;\ N\,;\ R)$. Cette suite comporte 4 succès, c'est-à-dire 4 boules rouges.
3. Combien y-a-t-il de suites de 6 évènements (R ou N) qui comportent 4 fois l'évènement R et 2 fois l'évènement N ?
4. Soit X la variable aléatoire égale au nombre de boules rouges tirées au cours des 6 épreuves. Déduire de ce qui précède $P(X = 4)$.
5. Par un raisonnement identique, donner la loi de probabilité de X. On dit que X suit la loi binomiale de paramètres n et p, avec, ici, $n = 6$ et $p = \frac{1}{3}$.

2.2 Définition de la loi binomiale $\mathscr{B}(n\,;\ p)$

Définition 2.2.1
Soit n *un entier naturel non nul et* p *un nombre réel de* $[0\,;\ 1]$.
On dit que la loi de probabilité d'une variable aléatoire X *discrète est une* loi binomiale *de paramètres* n *et* p*, notée* $\mathscr{B}(n\,;\ p)$*, si :*

$$P(X = k) = C_n^k\, p^k\, (1-p)^{n-k} \ \text{ avec } \ k \in \{0\,;\ 1\,;\ 2\,;\ \cdots\,;\ n\}.$$

Remarque : Les probabilités obtenues correspondent aux termes successifs de la formule du binôme de NEWTON donc :

$$\sum_{k=0}^{n} P(X = k) = \sum_{k=0}^{n} C_n^k\, p^k\, (1-p)^{n-k} = (p + (1-p))^n = 1.$$

2.3 Valeurs caractéristiques de la loi binomiale $\mathscr{B}(n\,;\ p)$

Théorème 2.3.1
Soit n *un entier naturel non nul et* p *un nombre réel de* $[0\,;\ 1]$.
Si une variable aléatoire X *suit la loi binomiale de paramètres* n *et* p*,* $\mathscr{B}(n\,;\ p)$*, alors :*

$$E(X) = np,\ \ V(X) = np(1-p) \ \text{ et } \ \sigma(X) = \sqrt{np(1-p)}.$$

$\Rightarrow$ Calculer la variance et l'écart-type de la variable aléatoire X définie dans l'exemple du paragraphe 2.1.

3 Loi de POISSON [2]

On peut généraliser les définitions et propriétés d'une variable aléatoire discrète X dans le cas où X prend une infinité *dénombrable* de valeurs. Par exemple lorsque $X(\Omega) = \mathbb{N} = \{0\,;\ 1\,;\ 2\,;\ \cdots\,;\ n\,;\ \cdots\}$.
Les formules avec le symbole « $\sum_{k=0}^{n}$ » restent valables mais s'écrivent avec le symbole « $\sum_{k=0}^{+\infty}$ » (on suppose alors que les séries convergent).

1. BERNOULLI, famille de mathématiciens originaire d'Anvers et qui vers la fin du XVIe siècle se réfugia à Bâle. Jakob BERNOULLI (1654 – 1705). Ses travaux concernent, en particulier, l'analyse, la géométrie différentielle et le calcul des probabilités.
2. Siméon, *Denis* POISSON (1781 – 1840), mathématicien français. Ses travaux concernent, en particulier, le calcul des variations, les équations aux dérivées partielles, et le calcul des probabilités. C'est l'un des créateurs de la physique mathématique.

3.1 Définition de la loi de POISSON $\mathscr{P}(\lambda)$

C'est une loi où l'ensemble des valeurs est $\mathbb{N}$.

Définition 3.1.1
Soit λ un nombre réel strictement positif. La loi de probabilité d'une variable aléatoire X à valeurs dans $\mathbb{N}$ est la loi de POISSON *de paramètre λ, notée $\mathscr{P}(\lambda)$, si :*

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

➡ Soit X une variable aléatoire qui suit la loi de POISSON de paramètre $\lambda = 4$.
Calculer $P(X = 2)$, $P(X < 2)$ et $P(X \geqslant 2)$.

3.2 Valeurs caractéristiques de la loi de POISSON $\mathscr{P}(\lambda)$

Théorème 3.2.1
Soit λ un nombre réel strictement positif. Si une variable aléatoire X suit la loi de POISSON *de paramètre λ, notée $\mathscr{P}(\lambda)$, alors :*

$$E(X) = \lambda,\ V(X) = \lambda \ \textit{ et } \ \sigma(X) = \sqrt{\lambda}.$$

4 Variable aléatoire continue

4.1 Exemple

Un jeu consiste à trouver un nombre compris entre 1 et 10.

<u>Première situation</u>

On considère la variable aléatoire discrète X égale à un nombre *entier* compris entre 1 et 10.

Soit x_0 le nombre choisi. La probabilité de trouver x_0 est $P(X = x_0) = \dfrac{1}{10}$.

Le nombre de valeurs prises par X est fini. On sait étudier une telle variable aléatoire discrète.

<u>Deuxième situation</u>

On considère la variable aléatoire X égale à un nombre *réel* compris entre 1 et 10.

Soit x_0 le nombre choisi. La probabilité de trouver x_0 est $P(X = x_0) = 0$.

L'ensemble des valeurs prises par X est un intervalle de $\mathbb{R}$, c'est un ensemble infini non dénombrable de nombres réels.

On ne peut pas dans ce cas définir la loi de probabilité de X avec des nombres $P(X = x_i)$.

4.2 Définition d'une variable aléatoire continue

Définition 4.2.1
Soit Ω un univers. On dit qu'une variable aléatoire X est continue *si l'ensemble des valeurs prises par X est un intervalle I de $\mathbb{R}$.*

4.3 Fonction de répartition d'une variable aléatoire continue

Définition 4.3.1
On appelle fonction de répartition *de la variable aléatoire X la fonction F définie par :*

$$\begin{aligned} F :\ & \mathbb{R} \longrightarrow \mathbb{R} \\ & x \longmapsto F(x) = P(X \leqslant x). \end{aligned}$$

Théorème 4.3.1 (Propriétés des fonctions de répartition des variables aléatoires continues)

i. *Quel que soit* $x \in \mathbb{R}$, $0 \leqslant F(x) \leqslant 1$.

ii. *Quel que soit* $x \in \mathbb{R}$, $P(X > x) = 1 - P(X \leqslant x) = 1 - F(x)$.

iii. *Quels que soient* $(x\ ;\ y) \in \mathbb{R}^2$, $P(x < X \leqslant y) = P(X \leqslant y) - P(X \leqslant x) = F(y) - F(x)$.

iv. *La fonction* F *est croissante.*

v. $\lim\limits_{x \to -\infty} F(x) = 0$ *et* $\lim\limits_{x \to +\infty} F(x) = 1$.

vi. *La fonction* F *est continue sur* $\mathbb{R}$.

Si X prend ses valeurs dans l'intervalle I, soit $x_0 \in I$. On a alors :

$$\begin{aligned}
\lim_{\substack{x \to x_0 \\ x > x_0}} F(x) &= \lim_{\substack{x \to x_0 \\ x > x_0}} P(X \leqslant x) \\
&= \lim_{\substack{x \to x_0 \\ x > x_0}} \left[P(X \leqslant x_0) + P(x_0 < X \leqslant x)\right] \\
&= P(X \leqslant x_0) + P(X = x_0) \\
&= F(x_0).
\end{aligned}$$

F est donc continue à gauche en x_0. On démontre de même que F est continue à droite en x_0. F est donc continue pour tout $x_0 \in \mathbb{R}$ donc sur $\mathbb{R}$.

4.4 Densité de probabilité

Définition 4.4.1

Soit f *une fonction continue sur* $\mathbb{R}$ *et soit* a *un nombre réel.*

Si les intégrales $\displaystyle\int_a^{+\infty} f(x)\,dx$ *et* $\displaystyle\int_{-\infty}^a f(x)\,dx$ *convergent vers* A *et* B, *on dit que l'intégrale* $\displaystyle\int_{-\infty}^{+\infty} f(x)\,dx$ *converge et on pose* $\displaystyle\int_{-\infty}^{+\infty} f(x)\,dx = A + B$.

Définition 4.4.2

Une fonction f *définie sur* $\mathbb{R}$ *est une* densité de probabilité *si :*

- *pour tout nombre réel* x, $f(x) \geqslant 0$;
- f *est continue sur* $\mathbb{R}$ *sauf éventuellement en un nombre fini de points ;*
- $\displaystyle\int_{-\infty}^{+\infty} f(x)\,dx = 1$.

Théorème 4.4.1

Soit X *une variable aléatoire continue et* F *la fonction de répartition de* X.

- *La dérivée de* F *sur* $\mathbb{R}$ *est une densité de probabilité de* X *notée* f.
- *La fonction* F *est une primitive de* f *sur* $\mathbb{R}$ *et elle est définie pour tout réel* x *par :*

$$F(x) = \int_{-\infty}^{x} f(t)\,dt.$$

⇒ Soit f la fonction définie sur $\mathbb{R}$ par : $f(x) = \begin{cases} 0 & \text{si } x < 0 \\ \dfrac{1}{3} & \text{si } 0 \leqslant x < 3. \\ 0 & \text{si } x \geqslant 3 \end{cases}$

Représenter f et montrer que f est une densité de probabilité d'une variable aléatoire X.
Définir et représenter la fonction de répartition F de X.

4.5 Valeurs caractéristiques d'une variable aléatoire continue

Espérance mathématique

Définition 4.5.1
Soit X *une variable aléatoire continue de densité de probabilité* f*. L'*espérance mathématique *de* X *est le nombre réel, noté* $E(X)$*, défini par :*

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

➾ Calculer l'espérance mathématique de la variable aléatoire X dont la densité de probabilité est définie dans le paragraphe 4.4.

Variance et écart-type

Définition 4.5.2
Soit X *une variable aléatoire continue de densité de probabilité* f*. La* variance *de* X *est le nombre réel positif, noté* $V(X)$*, défini par :*

$$V(X) = \int_{-\infty}^{+\infty} x^2 f(x)\, dx - (E(X))^2.$$

*L'*écart-type *de* X*, noté* $\sigma(X)$*, est la racine carrée de la variance :*

$$\sigma(X) = \sqrt{V(X)}.$$

➾ Calculer la variance et l'écart-type de la variable aléatoire X dont la densité de probabilité est définie dans le paragraphe 4.4.

5 Loi normale ou loi de LAPLACE[3]–GAUSS[4]

C'est une loi où l'ensemble des valeurs est $\mathbb{R}$.

5.1 Définition de la loi normale $\mathcal{N}(m\,;\ \sigma)$

Définition 5.1.1
Soit m *un nombre réel et* σ *un nombre réel strictement positif.*
La loi de probabilité d'une variable aléatoire continue X *est la* loi normale *ou* loi de LAPLACE–GAUSS *de paramètres* m *et* σ*, notée* $\mathcal{N}(m\,;\ \sigma)$*, si la densité de probabilité est la fonction numérique* f *définie sur* $\mathbb{R}$ *par :*

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{t-m}{\sigma}\right)^2}.$$

☛ La fonction f est représentée par la figure 1 pour $m = 2$ et $\sigma = 2$.

3. Pierre Simon, *marquis* de LAPLACE (1749 – 1827). Astronome, mathématicien, physicien et homme politique français. Ses travaux ont porté, en particulier, sur les formes quadratiques, les équations différentielles et aux dérivées partielles, les développements asymptotiques, le calcul des probabilités et la mécanique céleste.

4. Carl Friedrich GAUSS (1777 – 1855). Astronome, mathématicien et physicien allemand. Ses travaux ont porté, en particulier, sur l'arithmétique, l'algèbre, les séries hypergéométriques, la géométrie différentielle des surfaces, la géométrie non euclidienne et le calcul des probabilités

Fonction de répartion associée à la loi normale $\mathcal{N}(m\,;\,\sigma)$

Définition 5.1.2
Soit X *une variable aléatoire continue qui suit la loi normale* $\mathcal{N}(m\,;\,\sigma)$*. La fonction de répartition de* X *est la fonction numérique* F *définie pour tout réel* x *par :*

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-m}{\sigma}\right)^2}\,dt.$$

☛ On ne peut pas calculer cette intégrale à l'aide des fonctions élémentaires.
L'interprétation graphique de F est représentée par la figure 2.

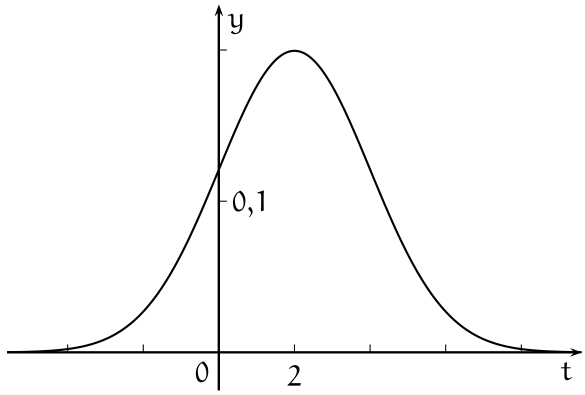

FIGURE 1 – représentation graphique de f pour $m = 2$ et $\sigma = 2$

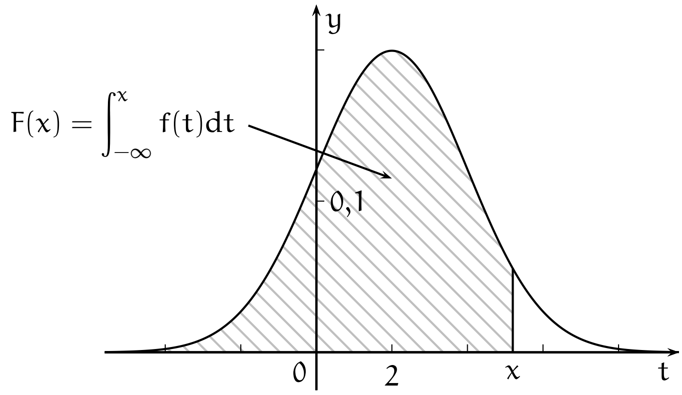

FIGURE 2 – interprétation graphique de F

5.2 Valeurs caractéristiques de la loi normale $\mathcal{N}(m\,;\,\sigma)$

Théorème 5.2.1
Si une variable aléatoire X *suit la loi normale* $\mathcal{N}(m\,;\,\sigma)$*, alors :*

$$E(X) = m,\ V(X) = \sigma^2 \ \textit{et}\ \sigma(X) = \sigma.$$

m *est appelé* moyenne *et* σ *l'*écart-type *de la loi normale* $\mathcal{N}(m\,;\,\sigma)$.

5.3 Définition de la loi normale centrée réduite $\mathcal{N}(0\,;\,1)$

Définition 5.3.1
Si les paramètres d'une loi normale sont $m = 0$ *et* $\sigma = 1$*, alors on dit que c'est la* loi normale centrée réduite*, notée* $\mathcal{N}(0\,;\,1)$.

Conséquences

– La densité de probabilité associée à la loi normale centrée réduite $\mathcal{N}(0\,;\,1)$ est la fonction f définie pour tout réel x par :

$$f(x) = \frac{1}{\sqrt{2\pi}}\,e^{-\frac{x^2}{2}}.$$

– f est donc paire et sa représentation graphique (dans le plan muni d'un repère orthogonal) est symétrique par rapport à l'axe des ordonnées (cf. figure 3).

Théorème 5.3.1
Si X *est une variable aléatoire qui suit la loi normale* $\mathcal{N}(m\,;\,\sigma)$ *alors la variable aléatoire* T *définie par :*

$$T = \frac{X - m}{\sigma}$$

suit la loi normale centrée réduite $\mathcal{N}(0\,;\,1)$.

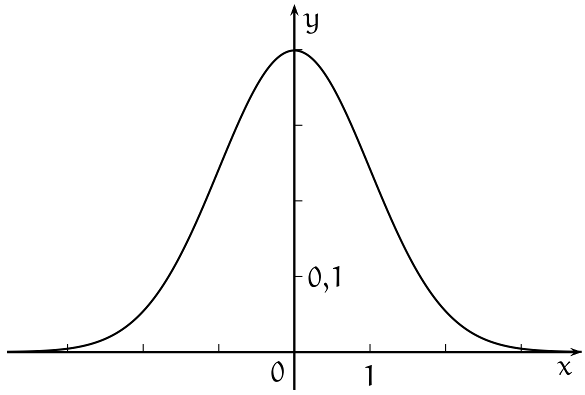

FIGURE 3 – représentation graphique de f

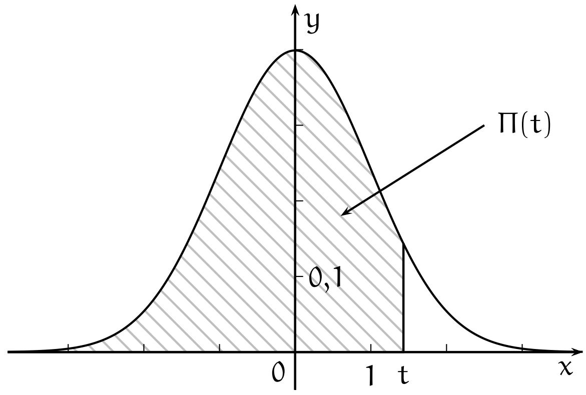

FIGURE 4 – interprétation graphique de Π

☛ Il suffit de choisir $u = \dfrac{t-m}{\sigma}$ comme nouvelle variable d'intégration et de poser $T = \dfrac{X-m}{\sigma}$, c'est-à-dire $X = \sigma T + m$, pour passer de la fonction de répartition de X à celle de T.

Fonction de répartion associée à la loi normale centrée réduite $\mathcal{N}(0\,;\,1)$

Définition 5.3.2
Soit T une variable aléatoire continue qui suit la loi normale centrée réduite $\mathcal{N}(0\,;\,1)$. La fonction de répartition Π de T est la fonction numérique définie pour tout réel x par :

$$\Pi(t) = P(T \leqslant t) = \int_{-\infty}^{t} f(x)\,dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}}\,dx.$$

☛ On ne peut pas calculer cette intégrale à l'aide des fonctions élémentaires.
L'interprétation graphique de Π est représentée par la figure 4.

Exercice d'application

1. En utilisant la table de la loi normale centrée réduite, calculer $P(T \leqslant 2{,}47)$ et $P(T > 2{,}47)$.
2. Calculer $P(T \leqslant -1{,}23)$ en utilisant la symétrie de la représentation graphique de f (cf. figure 5).
3. Soit X la variable aléatoire qui suit la loi normale de moyenne $m = 2{,}09$ et d'écart-type $\sigma = 0{,}13$.
Calculer $P(X \leqslant 2{,}35)$ et $P(1{,}895 \leqslant X \leqslant 2{,}285)$ en utilisant le changement de variable $T = \dfrac{X-m}{\sigma}$.

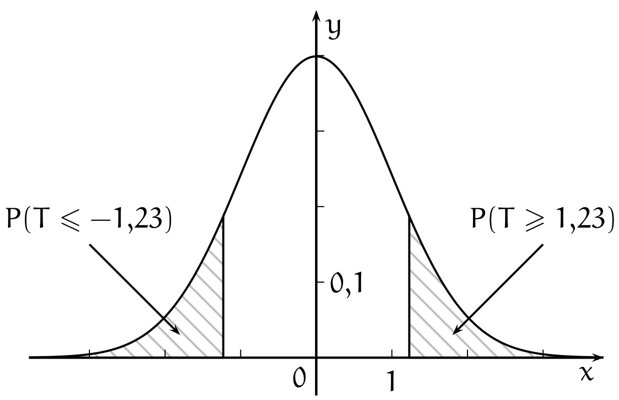

FIGURE 5 – calcul de $P(T \leqslant -1{,}23)$

6 Approximations de la loi binomiale

6.1 Approximation de la loi binomiale par la loi de POISSON

Exemple

Dans une chaine de fabrication, 5 % des pièces sont défectueuses.
On prélève une pièce, on examine si elle est défectueuse et on la replace parmi les autres. On repète 120 fois cette expérience.
On désigne par X la variable aléatoire qui, à chaque tirage de 120 pièces, associe le nombre de pièces défectueuses.

1. Justifier que X suit une loi binomiale dont on précisera les paramètres n et p et calculer $P(X = 5)$.
2. Soit Y la variable aléatoire qui suit la loi de POISSON $\mathscr{P}(\lambda)$ avec $\lambda = np$. Calculer $P(Y = 5)$.

Théorème 6.1.1
Soit X *une variable aléatoire qui suit la loi binomiale* $\mathscr{B}(n\ ;\ p)$.
Pour n *« assez grand » (*$n > 30$*) et pour* p *« voisin » de* 0 *(*$p \leqslant 0{,}1$*) tels que* $np(1-p) \leqslant 10$*, on peut approcher la loi binomiale* $\mathscr{B}(n\ ;\ p)$ *par la loi de* POISSON $\mathscr{P}(\lambda)$ *avec* $\lambda = np$*. On a alors :*

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

6.2 Approximation de la loi binomiale par la loi normale

Exemple

On lance 300 fois une pièce de monnaie truquée ce qui constitue une partie. La probabilité d'obtenir « face » est 0,65.
On désigne par X la variable aléatoire qui à chaque partie associe le nombre de « faces » obtenues.

1. Justifier que X suit une loi binomiale dont on précisera les paramètres n et p.
2. Calculer à l'aide de la calculatrice $P(X > 210)$?

Théorème 6.2.1
Soit X *une variable aléatoire qui suit la loi binomiale* $\mathscr{B}(n\ ;\ p)$.
Pour n *« assez grand » (*$n \geqslant 50$*) et pour* p *ni « voisin » de* 0*, ni « voisin » de* 1*, tels que* $np(1-p) > 10$*, on peut approcher la loi binomiale* $\mathscr{B}(n\ ;\ p)$ *par la loi normale* $\mathscr{N}(m\ ;\ \sigma)$ *avec* $m = np$ *et* $\sigma = \sigma(X) = \sqrt{np(1-p)}$*. On a alors :*

$$P(X \leqslant k) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{k} e^{-\frac{1}{2}\left(\frac{t-m}{\sigma}\right)^2}\,dt.$$

➡ On reprend l'exemple précédent (paragraphe 6.2).
Montrer qu'une approximation de la loi binomiale par une loi normale se justifie.
Calculer $P(X > 210)$ à l'aide de l'approximation (utiliser la table du formulaire).

☛ Pour améliorer la manière dont on passe d'une variable aléatoire discrète à une variable aléatoire continue, on peut appliquer la *correction de continuité* qui consiste à remplacer tout entier k par un intervalle d'extrémités $k - \frac{1}{2}$ et $k + \frac{1}{2}$.
Cette correction est surtout utile lorsque l'on cherche $P(X < k)$ alors que la table donne $P(X \leqslant k)$. Dans ce cas, on prend $P\left(X \leqslant k - \frac{1}{2}\right)$.

☛ Calculer $P(X > 210{,}5)$ (lorsque X suit la loi normale définie précédemment).

Table des figures