# 6

# DISTRIBUTIONS DE PROBABILITÉ CONTINUES

## STATISTIQUES APPLIQUÉES

***Procter & Gamble**** 

***Cincinnati, État de l'Ohio***

La société Procter&Gamble (P&G) fabrique et commercialise divers produits comme des détergents, des couches-culottes, des produits pharmaceutiques, des dentifrices, du savon, des bains de bouche et du papier toilette. À travers le monde, cette société possède des marques dominantes dans plus de catégories de produits que n'importe quelle autre société de biens de consommation. Depuis sa fusion avec Gillette, P&G fabrique et commercialise également des rasoirs, des lames et beaucoup d'autres produits de soin.

Leader dans l'application des méthodes statistiques dans le processus de décision, P&G emploie des personnes ayant différentes formations académiques : ingénierie, statistiques, recherche opérationnelle, commerce. L'aide à la décision et l'analyse des risques, les simulations avancées, l'amélioration de la qualité et les méthodes quantitatives (par exemple, programmation linéaire, analyse de la régression, analyse probabiliste) sont les principales fonctions de ces personnes.

Le département d'industrie chimique de P&G est l'un des principaux fabricants d'alcools gras, issus de substances naturelles, comme l'huile de noix de coco, et du pétrole. La division a souhaité évaluer les opportunités et les risques économiques liés à l'expansion de leurs installations de production; dans ce but, la direction a fait appel à ses spécialistes en décision probabiliste et en analyse des risques. Après avoir structuré et modélisé le problème, ces spécialistes ont indiqué que le différentiel de coût entre les matières premières dérivées de la noix de coco et celles dérivées du pétrole était l'élément clé de la rentabilité. Les coûts futurs étaient inconnus, mais les analystes ont été capables de les modéliser par les variables aléatoires continues suivantes : $x$, le prix de l'huile de coco par livre d'alcool gras et $y$, le prix de la matière première dérivée du pétrole par livre d'alcool gras.

Puisque la clé de la rentabilité était la différence entre ces deux variables aléatoires, une troisième variable aléatoire, $d = x - y$, a été utilisée pour l'analyse. Les spécialistes ont déterminé la distribution de probabilité des variables $x$ et $y$, puis en ont déduit celle de la différence, $d$. Selon la loi de probabilité de $d$, la probabilité que la différence de prix soit inférieure ou égale à 0,0655 dollar est égale à 0,9 et la probabilité que la différence de prix soit inférieure ou égale à 0,035 dollar est égale à 0,5. De plus, la probabilité que la différence de prix soit inférieure ou égale à 0,0045 dollar n'est que de 0,1.**

Le département d'industrie chimique pensait que le fait de quantifier l'impact de la différence de prix entre les matières premières permettrait de faire un choix. En effet, les probabilités obtenues ont été utilisées dans une analyse d'impact de la différence de prix des matières premières, qui a fourni suffisamment d'informations pour guider la direction dans sa décision.

L'utilisation de variables aléatoires continues et de leurs distributions de probabilité a permis à P&G d'analyser les risques économiques associés à sa production d'alcools gras. Dans ce chapitre, vous vous familiariserez avec les variables aléatoires continues et leurs distributions de probabilité, en particulier avec l'une des plus importantes distributions de probabilité en statistiques, la distribution normale.

* Les auteurs remercient Joel Kahn de Procter &Gamble, de leur avoir fourni ce Statistiques appliquées.

** Les différences de prix citées ici ont été modifiées pour des raisons de confidentialité des données.

Dans le chapitre précédent, nous avons traité des variables aléatoires discrètes et de leurs distributions de probabilité. Dans ce chapitre, nous étudierons les variables aléatoires continues. Plus particulièrement, nous étudierons trois distributions de probabilité continues : la loi uniforme, la loi normale et la loi exponentielle.

Une différence fondamentale distingue le calcul des probabilités des variables aléatoires discrètes et continues. Pour une variable aléatoire discrète, la fonction de probabilité $f(x)$ fournit la probabilité que la variable aléatoire prenne une valeur particulière. Pour une variable aléatoire continue, la *fonction de densité de probabilité*, également notée $f(x)$, est l'équivalent de la fonction de probabilité. Contrairement à la fonction de probabilité des variables aléatoires discrètes, la fonction de densité de probabilité des variables aléatoires continues ne fournit pas directement les probabilités. Cependant, l'aire située sous le graphique de $f(x)$ dans un intervalle particulier donne la probabilité que la variable aléatoire continue $X$ prenne une valeur dans cet intervalle. Ainsi, lorsqu'on calcule des probabilités pour des variables aléatoires continues, on calcule la probabilité que la variable aléatoire prenne n'importe quelle valeur dans un intervalle particulier.

Une des implications de cette définition de la probabilité pour les variables aléatoires continues est que la probabilité que la variable aléatoire prenne une valeur particulière est nulle, puisque l'aire sous le graphique de $f(x)$ à un point donné est nulle. Dans la section 6.1, nous appliquerons ces concepts à une variable aléatoire continue distribuée selon une loi uniforme.

Une grande partie du chapitre est consacrée à des exemples d'application de la loi normale. La loi normale est très importante : elle est très utilisée en inférence statistique. Le chapitre se termine par une discussion sur la loi exponentielle, utile dans des applications impliquant des temps d'attente et des durées de service.

## 6.1 LA LOI UNIFORME

Considérons la variable aléatoire $X$ qui représente la durée du vol en avion entre Chicago et New York. Supposons que la durée du vol soit comprise entre 120 et 140 minutes. Puisque la variable aléatoire $X$ peut prendre n'importe quelle valeur dans cet intervalle de temps, $X$ est une variable aléatoire continue et non pas discrète. Supposons que les données actuelles sur la durée du vol nous permettent de conclure que la probabilité que la durée du vol appartienne à un intervalle d'une minute, compris entre 120 et 140 minutes, est la même que la probabilité que la durée du vol appartienne à un autre intervalle d'une minute compris entre 120 et 140 minutes. Puisque tous les intervalles d'une minute, compris entre 120 et 140, sont équiprobables, on dit que la variable aléatoire $X$ suit une **loi uniforme**. La fonction de densité de probabilité, qui définit la loi uniforme de cette variable aléatoire $X$, correspond à :

$$f(x) = \begin{cases} 1/20 & \text{si } 120 \leq x \leq 140 \\ 0 & \text{sinon} \end{cases}$$

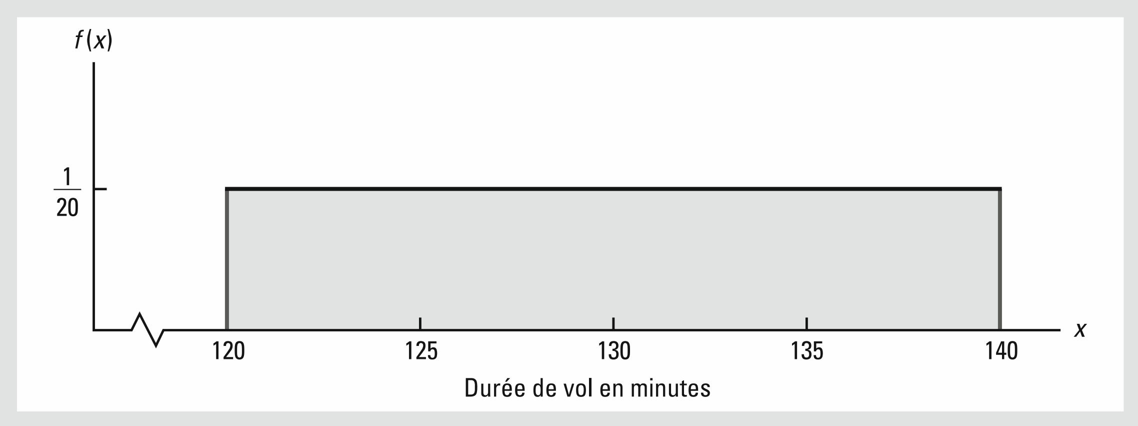



**Figure 6.1** ***Distribution de probabilité uniforme pour la durée de vol***

Lorsque la probabilité est proportionnelle à la longueur de l'intervalle, la variable aléatoire est distribuée de façon uniforme.

La figure 6.1 est une représentation graphique de cette fonction de densité. De façon plus générale, la fonction de densité uniforme pour une variable aléatoire $X$ est obtenue en utilisant la formule suivante :

► **Fonction de densité de probabilité uniforme**

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{sinon} \end{cases} \tag{6.1}$$

Dans l'exemple de la durée du vol entre Chicago et New York, $a = 120$ et $b = 140$.

Comme nous l'avons dit en introduction, pour une variable aléatoire continue, la probabilité correspond à la vraisemblance que cette variable aléatoire prenne une valeur appartenant à un intervalle particulier. Dans l'exemple relatif à la durée du vol, on peut se demander quelle est la probabilité que celle-ci soit comprise entre 120 et 130 minutes, c'est-à-dire quelle est la valeur de $P(120 \leq x \leq 130)$. Puisque la durée du vol doit être comprise entre 120 et 140 minutes et que les probabilités sont uniformément distribuées sur cet intervalle, on pressent que $P(120 \leq x \leq 130) = 0{,}50$. Dans le paragraphe suivant, nous montrerons que cette probabilité est égale à l'aire située sous le graphique de $f(x)$, entre 120 et 130 (cf. figure 6.2).

## 6.1.1 L'aire comme mesure des probabilités

Considérons l'aire sous le graphique de $f(x)$, entre 120 et 130, représenté à la figure 6.2. La partie considérée du graphique est rectangulaire. Par conséquent, son aire est simplement égale à la largeur multipliée par la hauteur. Avec la largeur de l'intervalle égale à 10 ($130 - 120 = 10$) et la hauteur égale à la valeur de la fonction de densité, $f(x) = 1/20$, nous avons une aire de 0,50 ($10 \times (1/20) = 10/20 = 0{,}50$).

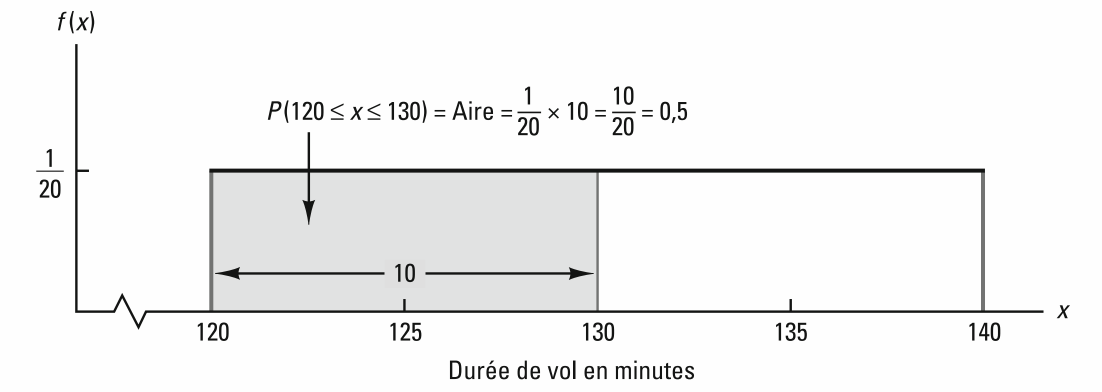



**Figure 6.2** ***L'aire fournit la probabilité que la durée du vol soit comprise entre 120 et 130 minutes***

Quelle remarque pouvez-vous faire concernant l'aire sous le graphique de $f(x)$ et la probabilité ? Elles sont identiques ! Ce résultat est généralisable à toutes les variables aléatoires continues. Une fois la fonction de densité $f(x)$ identifiée, la probabilité que $X$ prenne une valeur comprise entre $x_1$ et $x_2$ est égale à l'aire sous le graphique de $f(x)$ comprise entre $x_1$ et $x_2$.

Étant donnée la distribution uniforme de la durée de vol, en utilisant l'interprétation de l'aire en termes de probabilité, on peut répondre à un certain nombre de questions en matière de probabilité concernant la durée de vol. Par exemple, quelle est la probabilité que la durée du vol soit comprise entre 128 et 136 minutes ? La largeur de l'intervalle est égale à 8 ($136-128=8$). Avec une hauteur uniforme de 1/20, $P(128 \leq x \leq 136) = 8 \quad (1/20) = 0,40$.

Notez que $P(120 \leq x \leq 140) = 20 \quad (1/20) = 1$. En d'autres termes, l'aire totale sous le graphique de $f(x)$ est égale à 1. Cette propriété est valable pour toutes les lois continues et correspond à la condition associée à une fonction de probabilité discrète selon laquelle la somme des probabilités doit être égale à 1. Pour une fonction de densité continue, on doit également avoir $f(x) \geq 0$ pour toute valeur de $X$. Cette condition est analogue à la condition $f(x) \geq 0$ associée aux fonctions de probabilité discrètes.

Deux différences majeures subsistent entre le traitement des variables aléatoires continues et celui des variables aléatoires discrètes.

1. On ne parle plus de la probabilité d'une variable aléatoire prenant une valeur particulière. Au contraire, on parle de la probabilité qu'une variable aléatoire prenne une valeur appartenant à un intervalle donné.
2. La probabilité qu'une variable aléatoire prenne une valeur dans un intervalle donné, entre $x_1$ et $x_2$, est égale à l'aire située sous le graphique de la fonction de densité entre $x_1$ et $x_2$. Ceci implique que la probabilité qu'une variable aléatoire prenne une valeur particulière est nulle, puisque l'aire sous le graphique de $f(x)$ à un point donné est nulle. Ceci signifie également que la

probabilité qu'une variable aléatoire continue prenne une valeur dans un intervalle donné est la même que les bornes de l'intervalle soient incluses ou non.

> Pour voir que la probabilité d'une valeur isolée est nulle, référez-vous à la figure 6.2 et calculez la probabilité d'une valeur isolée, par exemple $x = 125$.
> $P(x = 125) = P(125 \leq x \leq 125) = 0 \quad (1/20) = 0$.

Le calcul de l'espérance mathématique et de la variance d'une variable aléatoire continue est analogue à celui d'une variable aléatoire discrète. Cependant, puisque les calculs contiennent des intégrales, nous laissons le soin à des ouvrages plus avancés de les développer.

Pour la loi uniforme continue introduite dans cette section, les formules de l'espérance mathématique et de la variance sont :

$$E(x) = \frac{a+b}{2}$$

$$Var(x) = \frac{(b-a)^2}{12}$$

Dans ces formules, $a$ est la plus petite valeur et $b$ la plus grande valeur que la variable aléatoire puisse prendre.

En appliquant ces formules à l'exemple de la durée de vol entre Chicago et New York, nous obtenons :

$$E(x) = \frac{120+140}{2} = 130$$

$$Var(x) = \frac{(140-120)^2}{12} = 33,33$$

L'écart type de la durée du vol, $\sigma$, est égal à la racine carrée de la variance, soit 5,77 minutes.

## REMARQUES

Pour voir plus clairement pourquoi la hauteur de la fonction de densité n'est pas une probabilité, considérons une variable aléatoire distribuée uniformément de la façon suivante :

$$f(x) = \begin{cases} 2 \text{ si } 0 \leq x \leq 0,5 \\ 0 \quad \text{sinon} \end{cases}$$

La hauteur de la fonction de densité $f(x)$ est égale à 2 pour les valeurs de $X$ comprises entre 0 et 0,5. Or, nous savons que les probabilités ne peuvent jamais être supérieures à 1. Aussi, $f(x)$ ne peut être interprétée comme la probabilité que $X = x$.

**EXERCICES**

## Méthode

**1.** La variable aléatoire $X$ est uniformément distribuée entre 1,0 et 1,5.

- **a)** Représenter graphiquement la fonction de densité de probabilité.
- **b)** Calculer $P(x = 1,25)$.
- **c)** Calculer $P(1,0 \leq x \leq 1,25)$.
- **d)** Calculer $P(1,2 < x < 1,5)$.

**2.** La variable aléatoire $X$ est uniformément distribuée entre 10 et 20.

- **a)** Représenter graphiquement la fonction de densité de probabilité.
- **b)** Calculer $P(x < 15)$.
- **c)** Calculer $P(12 \leq x \leq 18)$.
- **d)** Calculer $E(X)$.
- **e)** Calculer $Var(X)$.

## Applications

**3.** Delta Airlines évalue le temps du vol entre Cincinnati et Tampa à 2 heures et 5 minutes. Supposons que les temps de vol soient uniformément distribués entre 2 heures et 2 heures et 20 minutes.

- **a)** Représenter graphiquement la fonction de densité de probabilité pour les temps de vol.
- **b)** Quelle est la probabilité que le vol n'ait pas plus de 5 minutes de retard ?
- **c)** Quelle est la probabilité que le vol ait plus de 10 minutes de retard ?
- **d)** Quel est le temps de vol moyen ?

**4.** La plupart des langages informatiques ont une fonction qui génère des nombres aléatoires. La fonction RAND d'Excel peut être utilisée pour générer des nombres aléatoires entre 0 et 1. Soit $X$ une variable aléatoire continue générée par la fonction RAND, dont la fonction de densité est :

$$f(x) = \begin{cases} 1 & \text{si } 0 \leq x \leq 1 \\ 0 & \text{sinon} \end{cases}$$

- **a)** Représenter graphiquement la fonction de densité de probabilité.
- **b)** Quelle est la probabilité de générer un nombre aléatoire compris entre 0,25 et 0,75 ?
- **c)** Quelle est la probabilité de générer un nombre aléatoire inférieur ou égal à 0,30 ?
- **d)** Quelle est la probabilité de générer un nombre aléatoire supérieur à 0,60 ?
- **e)** Générer 50 nombres aléatoires en entrant =RAND() dans 50 cellules d'une feuille de calcul Excel.
- **f)** Calculer la moyenne et l'écart type des nombres aléatoires générés à la question (e).

**5.** En octobre 2012, Apple a lancé une version plus petite de son iPad, connu sous le nom de iPad Mini. Pesant moins de 11 onces, il est environ 50 % plus léger que l'iPad standard. Les tests réalisés ont montré que la batterie de l'iPad Mini avait une durée d'autonomie moyenne de

10,25 heures (*The Wall Street Journal*, 31 octobre 2012). Supposez que la durée d'autonomie de la batterie d'un iPad Mini est uniformément distribuée entre 8,5 et 12 heures.

- **a)** Donner l'expression mathématique de la fonction de densité de probabilité de la durée d'autonomie de la batterie.
- **b)** Quelle est la probabilité que la durée d'autonomie de la batterie soit inférieure ou égale à 10 heures ?
- **c)** Quelle est la probabilité que la durée d'autonomie de la batterie soit supérieure ou égale à 11 heures ?
- **d)** Quelle est la probabilité que la durée d'autonomie de la batterie soit comprise entre 9,5 et 11,5 heures ?
- **e)** Parmi une cargaison de 100 iPad Mini, combien devraient avoir une durée d'autonomie d'au moins 9 heures ?

**6.** Un sondage Daily Tracking de la société Gallup a révélé que les dépenses courantes quotidiennes moyennes des Américains gagnant plus de 90 000 dollars par an s'élevaient à 136 dollars (*USA Today*, 30 juillet 2012). Les dépenses courantes quotidiennes ne tiennent pas compte des achats de logement, de véhicule et des factures courantes mensuelles. Soit $X$ la variable aléatoire correspondant aux dépenses courantes quotidiennes. Supposez qu'elle suive une loi uniforme dont la fonction de densité est donnée par $f(x) = 0{,}00625$ pour $a \leq x \leq b$ .

- **a)** Quelles sont les valeurs de $a$ et de $b$ ?
- **b)** Quelle est la probabilité que les consommateurs de ce groupe aient des dépenses courantes quotidiennes comprises entre 100 et 200 dollars ?
- **c)** Quelle est la probabilité que les consommateurs de ce groupe aient des dépenses courantes quotidiennes supérieures ou égales à 150 dollars ?
- **d)** Quelle est la probabilité que les consommateurs de ce groupe aient des dépenses courantes quotidiennes inférieures ou égales à 80 dollars ?

**7.** Supposez que nous nous intéressions à l'acquisition d'une parcelle de terrain et que nous sachions qu'une autre personne est également intéressée.[1] Le vendeur a annoncé que l'offre la plus élevée, supérieure à 10 000 dollars, serait acceptée. Supposez que l'offre concurrente $X$ est une variable aléatoire uniformément distribuée entre 10 000 et 15 000 dollars.

- **a)** Supposez que vous offriez 12 000 dollars. Quelle est la probabilité que votre offre soit acceptée ?
- **b)** Supposez que vous offriez 14 000 dollars. Quelle est la probabilité que votre offre soit acceptée ?
- **c)** Quel montant devez-vous offrir pour maximiser la probabilité d'obtention du terrain ?
- **d)** Supposez que vous connaissiez quelqu'un qui soit prêt à vous donner 16 000 dollars pour le terrain. Offririez-vous un montant inférieur à celui de la question (c) ? Pourquoi ?

## 6.2 LA LOI NORMALE

La loi la plus importante pour décrire une variable aléatoire continue est la **loi normale**. La loi normale a été utilisée dans de nombreuses applications pratiques, dans lesquelles

[1] Cet exercice est basé sur un problème suggéré par le professeur Roger Myerson de l'Université de Northwestern.

les variables aléatoires étaient la taille et le poids d'individus, les résultats des tests d'intelligence, des mesures scientifiques, le niveau des précipitations, etc. Elle est également très utilisée dans le domaine de l'inférence statistique, principal sujet de la suite de cet ouvrage. Dans de telles applications, la loi normale fournit une description des résultats possibles obtenus grâce à un échantillon.

Abraham de Moivre, un mathématicien français, a publié en 1733 *La Doctrine de la Chance*. Il y développa la loi normale.

## 6.2.1 La courbe normale

La loi normale est représentée par une courbe en forme de cloche (cf. figure 6.3). La fonction de densité de probabilité qui définit la courbe en forme de cloche de la loi normale est la suivante :

► **Fonction de densité de probabilité normale**

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \tag{6.2}$$

où

$\mu$ correspond à la moyenne
$\sigma$ correspond à l'écart type
$\pi \cong 3{,}14159$
$e \cong 2{,}71828$

La courbe normale a deux paramètres, $\mu$ et $\sigma$. Ils déterminent la position et la forme de la distribution.

Plusieurs remarques sur les caractéristiques de la loi normale s'imposent.

1. Il existe une famille entière de lois normales. Elles se différencient par leur moyenne $\mu$ et leur écart type $\sigma$.
2. Le point le plus élevé de la courbe normale correspond à la moyenne, qui est également la médiane et le mode de la distribution.
3. La moyenne de la distribution peut être négative, nulle ou positive. Trois courbes normales ayant le même écart type mais trois moyennes différentes (–10, 0 et 20) sont représentées ci-dessous.
4. La distribution normale est symétrique : la courbe à gauche de la moyenne correspond à l'image inversée de la courbe à droite de la moyenne. Les queues de la courbe s'étendent à l'infini de chaque côté et théoriquement, ne touchent jamais l'axe horizontal. La distribution étant symétrique, son coefficient d'asymétrie est nul.
5. L'écart type détermine la largeur et le degré d'aplatissement de la courbe. Plus l'écart type est grand, plus la courbe sera large, aplatie, traduisant ainsi une plus grande dispersion des données. Deux distributions normales de même moyenne mais avec des écarts type différents sont représentées ici.

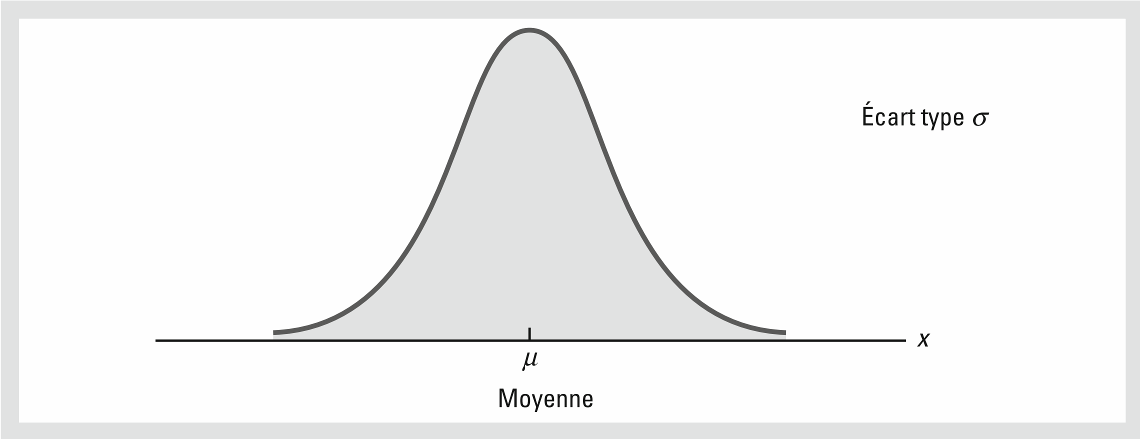
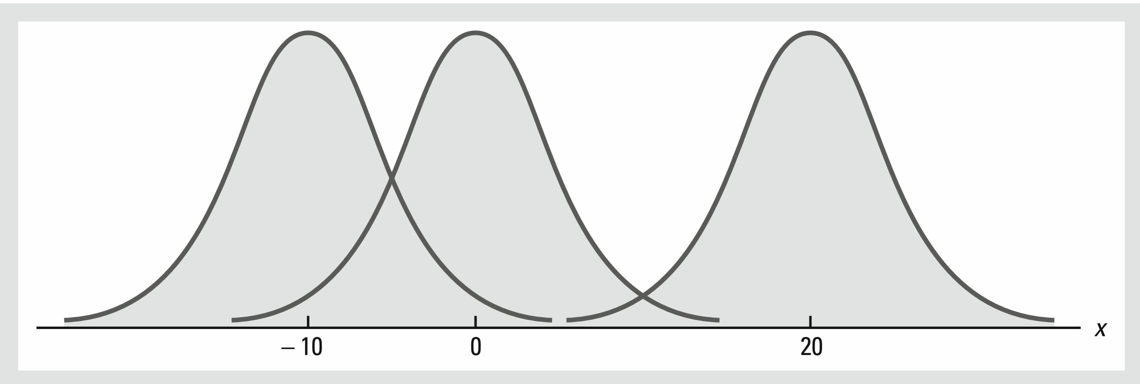



**Figure 6.3** ***Courbe en forme de cloche de la loi normale***

**6.** Les probabilités d'une variable aléatoire normale sont données par l'aire sous la courbe. L'aire totale située sous la courbe d'une distribution de probabilité normale est égale à 1. Puisque la distribution est symétrique, l'aire sous la courbe à gauche de la moyenne est égale à 0,5 et l'aire sous la courbe à droite de la moyenne à 0,5 également.




**7.** En règle générale,

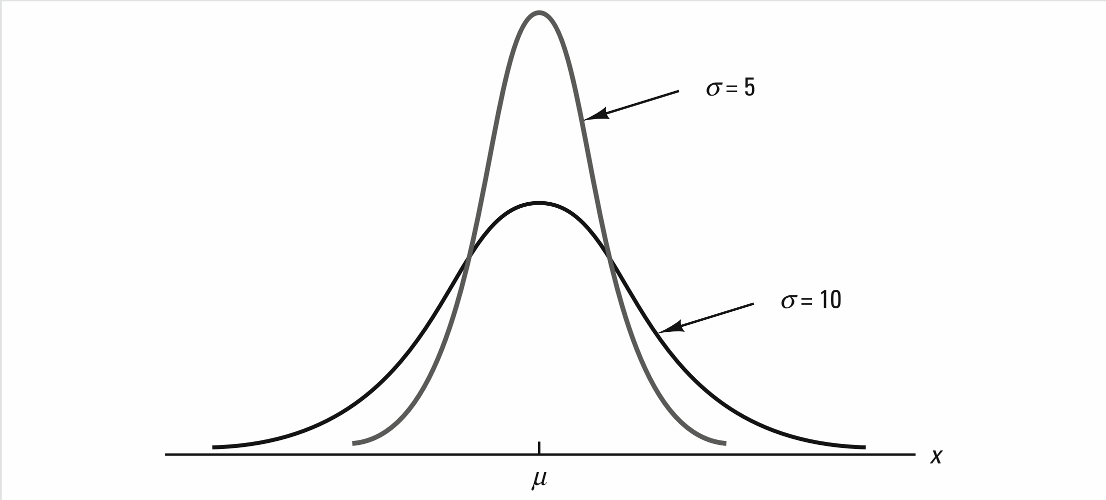
**a.** 68,3% des valeurs d'une variable aléatoire normale sont comprises dans l'intervalle $[\mu-\sigma\,;\,\mu+\sigma]$.

**b.** 95,4% des valeurs d'une variable aléatoire normale sont comprises dans l'intervalle $[\mu-2\sigma\,;\,\mu+2\sigma]$.

**c.** 99,7% des valeurs d'une variable aléatoire normale sont comprises dans l'intervalle $[\mu-3\sigma\,;\,\mu+3\sigma]$.

Ces pourcentages sont à la base de la règle empirique présentée à la section 3.3.

La figure 6.4 illustre graphiquement les propriétés (a), (b) et (c).




## 6.2.2 La loi normale centrée réduite

Une variable aléatoire qui a une distribution de probabilité normale de moyenne nulle et d'écart type égal à 1, suit ce que l'on appelle une **loi normale centrée réduite**. La lettre $Z$ est habituellement utilisée pour désigner cette variable aléatoire normale particulière. La figure 6.5 représente la loi normale centrée réduite. Elle a la même apparence générale que d'autres distributions normales, mais avec $\mu = 0$ et $\sigma = 1$.

Puisque $\mu = 0$ et $\sigma = 1$, l'expression de la fonction de densité normale centrée réduite est plus simple que l'expression (6.2).

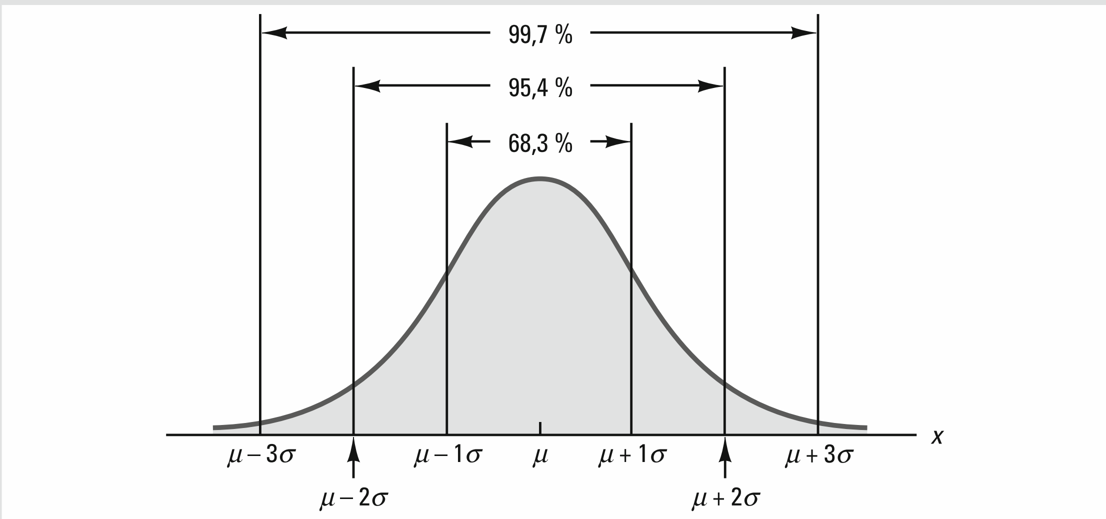



**Figure 6.4** ***Aire sous la courbe d'une loi normale***

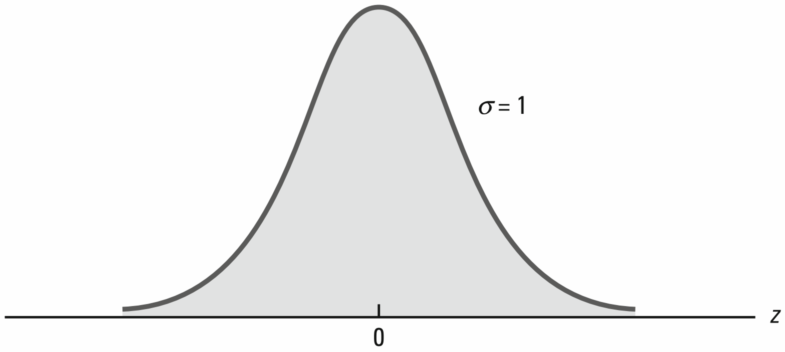



**Figure 6.5** *La loi normale centrée réduite*

► **Fonction de densité normale centrée réduite**

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

Comme pour les autres variables aléatoires continues, les probabilités d'une loi normale sont obtenues en calculant l'aire sous la courbe de la fonction de densité. Ainsi, pour trouver la probabilité qu'une variable aléatoire normale prenne une valeur appartenant à un intervalle donné, nous devons calculer l'aire sous la courbe normale dans cet intervalle.

La hauteur de la courbe de la fonction de densité normale varie et des calculs avancés sont nécessaires pour obtenir l'aire qui correspond à la probabilité.

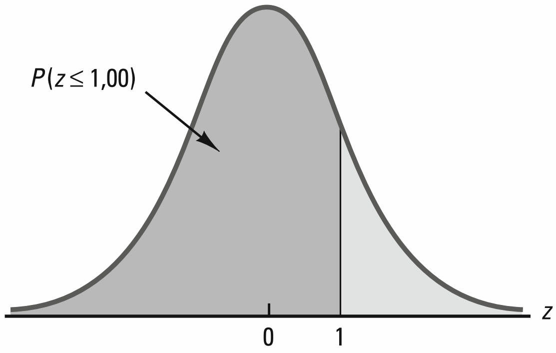
Pour la loi normale centrée réduite, les aires sous la courbe normale ont été calculées et sont disponibles dans des tables utilisées pour calculer les probabilités. Ces tables de probabilité sont reproduites sur les deux pages intérieures de la couverture du livre. La table sur la page de gauche contient les aires ou les probabilités cumulées pour des valeurs $z$ inférieures ou égales à la moyenne (égale à zéro). La table sur la page de droite contient les aires ou les probabilités cumulées pour des valeurs $z$ supérieures ou égales à la moyenne (égale à zéro).

Les trois types de probabilités qu'il peut être nécessaire de calculer sont (1) la probabilité que la variable aléatoire centrée réduite $Z$ soit inférieure ou égale à une certaine valeur ; (2) la probabilité que $Z$ soit comprise entre deux valeurs données ; et (3) la probabilité que $Z$ soit supérieure ou égale à une certaine valeur. Pour illustrer l'utilisation de la table des probabilités cumulées d'une distribution normale centrée réduite pour calculer ces trois types de probabilités, considérons les exemples suivants.

Pour commencer, voyons comment calculer la probabilité que la valeur $z$ d'une variable aléatoire normale centrée réduite $Z$ soit inférieure à 1 ; c'est-à-dire $P(z \leq 1)$. La probabilité cumulée correspond à l'aire sous la courbe normale à gauche de $z = 1$ sur le graphique suivant.

> Puisque la variable aléatoire normale centrée réduite est continue, $P(z \le 1) = P(z < 1)$.

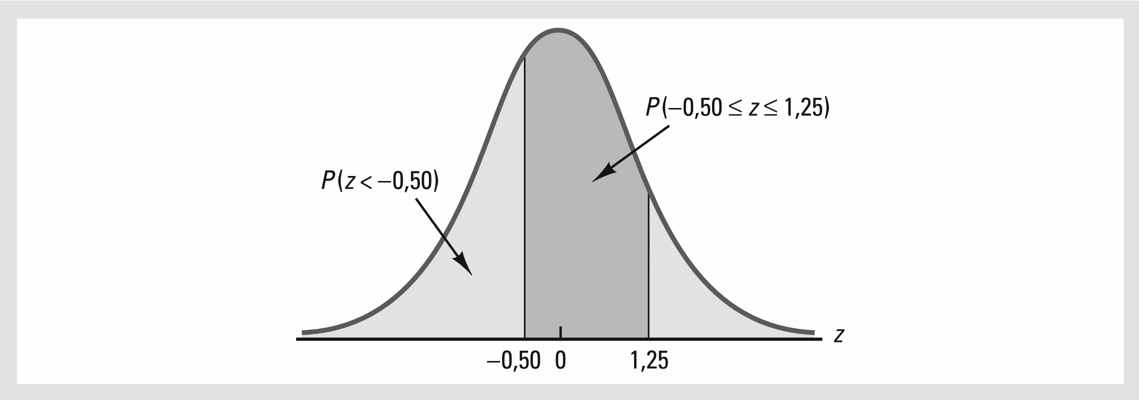
Référez-vous à la page de droite de la table des probabilités normales centrées réduites sur la page de couverture intérieure du livre. La probabilité cumulée correspon-




dant à $z = 1$ est située dans la table à l'intersection de la ligne intitulée 1,0 et de la colonne intitulée 0,00. À cette intersection se trouve la valeur 0,8413 ; ainsi, $P(z \le 1) = 0{,}8413$. L'extrait suivant de la table de probabilité illustre ces étapes.

| **z** | **0,00** | **0,01** | **0,02** |
|---|---|---|---|
| . | | | |
| . | | | |
| . | | | |
| **0,9** | 0,8159 | 0,8186 | 0,8212 |
| **1,0** | **0,8413** | 0,8438 | 0,8461 |
| **1,1** | 0,8643 | 0,8665 | 0,8686 |
| **1,2** | 0,8849 | 0,8869 | 0,8888 |
| . | | | |
| . | | | |
| . | | | |
| | $P(z \le 1{,}00)$ | | |

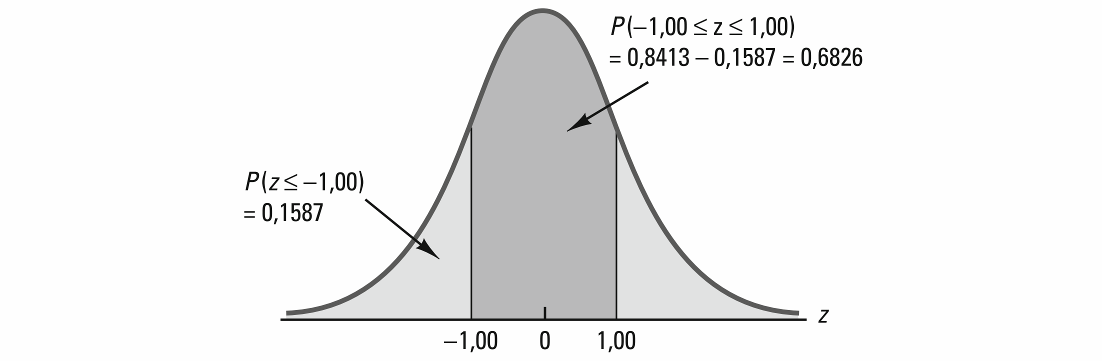
Pour illustrer le second type de calcul de probabilités, nous montrons comment calculer la probabilité que la valeur de la variable aléatoire normale centrée réduite soit comprise entre –0,50 et 1,25 ; c'est-à-dire $P(-0{,}50 \le z \le 1{,}25)$. Le graphique suivant illustre cette aire ou probabilité.



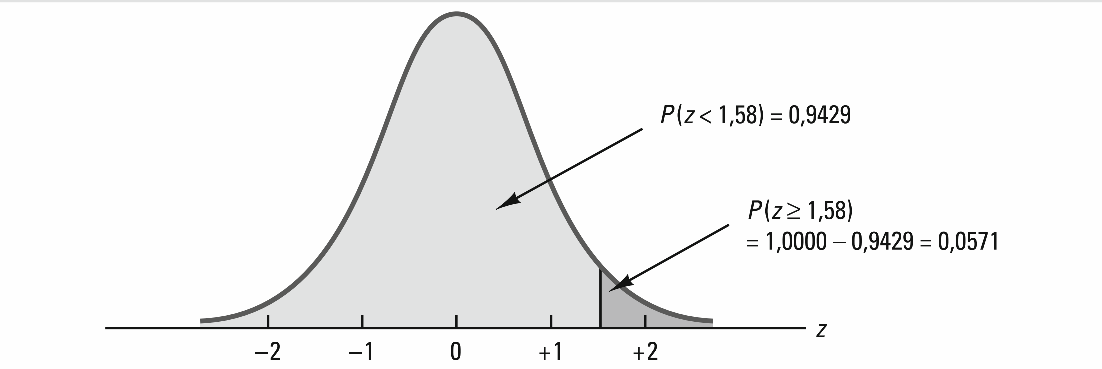

Trois étapes sont nécessaires au calcul de cette probabilité. Tout d'abord, nous trouvons l'aire sous la courbe normale à gauche de $z = 1{,}25$. Ensuite, nous trouvons l'aire sous la courbe normale à gauche de $z = -0{,}50$. Enfin, nous soustrayons l'aire à gauche de $z = -0{,}50$ à l'aire à gauche de $z = 1{,}25$ pour trouver $P(-0{,}50 \le z \le 1{,}25)$.

Pour trouver l'aire sous la courbe normale à gauche de $z = 1{,}25$, nous nous intéressons à la cellule de la table située à l'intersection de la ligne 1,2 et de la colonne 0,05. Puisque cette cellule contient la valeur 0,8944, $P(z \le 1{,}25) = 0{,}8944$. De même, pour trouver l'aire sous la courbe à gauche de $z = -0{,}50$ nous nous intéressons à la cellule de la table de probabilité située à l'intersection de la ligne –0,5 et de la colonne 0,00. La valeur de cette cellule est égale à 0,3985 : $P(z \le -0{,}5) = 0{,}3085$. Ainsi, $P(-0{,}50 \le z \le 1{,}25) = P(z \le 1{,}25) - P(z \le -0{,}50) = 0{,}8944 - 0{,}3085 = 0{,}5859$.

Considérons un autre exemple de calcul de la probabilité que $Z$ soit dans un intervalle entre deux valeurs données. Souvent il est intéressant de calculer la probabilité qu'une variable aléatoire normale prenne une valeur à l'intérieur d'un intervalle s'écartant d'un certain nombre d'écarts type de la moyenne. Supposons que l'on veuille calculer la probabilité qu'une variable aléatoire centrée réduite soit comprise dans l'intervalle d'un écart type autour de la moyenne, c'est-à-dire que $P(-1 \le z \le 1)$. Pour calculer cette probabilité nous devons trouver l'aire sous la courbe entre -1 et 1. Précédemment nous avons trouvé que $P(z \le 1) = 0{,}8413$. En se référant de nouveau à la table de probabilité située sur la couverture intérieure du livre, nous trouvons que l'aire sous la courbe à gauche de $z = -1$ est égale à 0,1587, ainsi $P(z \le -1) = 0{,}1587$. Donc, $P(-1 \le z \le 1) = P(z \le 1) - P(z \le -1) = 0{,}8413 - 0{,}1587 = 0{,}6826$. Cette probabilité est illustrée graphiquement par la figure suivante.



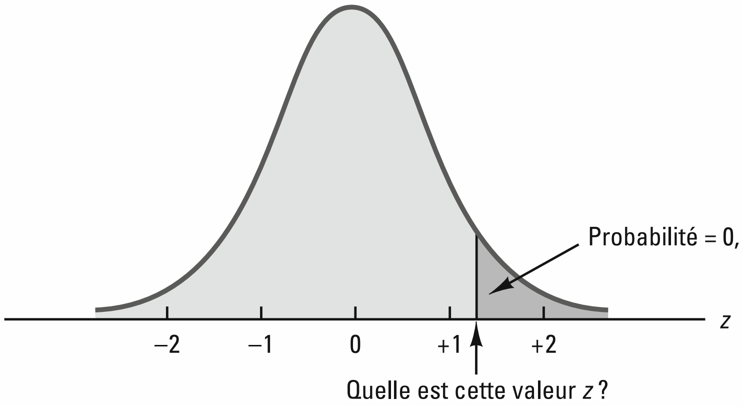

Pour illustrer comment calculer le troisième type de probabilité, supposons que nous voulions calculer la probabilité d'obtenir une valeur $z$ supérieure ou égale à 1,58 ; c'est-à-dire, $P(z \geq 1,58)$. La valeur située à l'intersection de la ligne 1,5 et de la colonne 0,08 dans la table des probabilités normales cumulées est égale à 0,9429 ; ainsi, $P(z < 1,58) = 0,9429$. Cependant, puisque l'aire totale sous la courbe normale est égale à 1, $P(z \geq 1,58) = 1 - P(z < 1,58) = 1 - 0,9429 = 0,0571$. La probabilité est illustrée par la figure suivante.




Dans les illustrations précédentes, nous avons montré comment calculer les probabilités étant données des valeurs $z$ spécifiques. Dans certaines situations, nous connaissons la probabilité et nous recherchons la valeur $z$ correspondante. Supposons que nous voulions trouver une valeur $z$ telle que la probabilité d'obtenir une valeur $z$ plus importante soit égale à 0,10. La figure suivante illustre cette situation.

Ce problème est l'inverse des exemples précédents. Précédemment, on spécifiait la valeur $z$ à laquelle on s'intéressait et cherchait la probabilité ou l'aire correspondante. Dans cet exemple, la probabilité ou l'aire est donnée et on cherche la valeur $z$ qui lui correspond. Pour cela, on utilise la table des probabilités de la loi normale centrée réduite d'une manière un peu différente.

> Étant donnée une probabilité, on peut utiliser la table des probabilités de la loi normale centrée réduite de manière inverse pour trouver la valeur z correspondante.




Rappelons que la table fournit l'aire sous la courbe à gauche d'une valeur particulière de la variable aléatoire normale *Z*. Nous savons que l'aire dans la queue droite de la courbe est égale à 0,10. Par conséquent, l'aire sous la courbe à gauche de la valeur *z* inconnue doit être égale à 0,9. En recherchant dans le corps de la table, nous trouvons que 0,8997 est la valeur de la probabilité cumulée la plus proche de 0,9. La partie de la table contenant cette valeur est reproduite ci-dessous.

| **z** | **0,06** | **0,07** | **0,08** | **0,09** |
|---|---|---|---|---|
| . . . | | | | |
| **1,0** | 0,8554 | 0,8577 | 0,8599 | 0,8621 |
| **1,1** | 0,8770 | 0,8790 | 0,8810 | 0,8830 |
| **1,2** | 0,8962 | 0,8980 | $\boxed{0{,}8997}$ | 0,9015 |
| **1,3** | 0,9131 | 0,9147 | 0,9162 | 0,9177 |
| **1,4** | 0,9279 | 0,9292 | 0,9306 | 0,9319 |
| . . . | | | | |
| | | Valeur de la probabilité cumulée la plus proche de 0,9 | | |

La valeur *z* associée à cette probabilité est 1,28 (elle se trouve à l'intersection de la colonne 1,2 et de la ligne 0,08). Ainsi, une aire d'environ 0,9 (en fait 0,8997) se situe à gauche de $z = 1,28$.[2] En utilisant les termes de la question posée à l'origine, il y a une probabilité d'environ 0,10 que *z* soit supérieur à 1,28.

Les exemples illustrent l'utilisation de la table des probabilités cumulées de la loi normale centrée réduite pour trouver les probabilités associées aux valeurs d'une variable

[2] On peut extrapoler les valeurs de la table pour obtenir une meilleure approximation de la valeur *z* qui correspond à une aire de 0,9. Pour une décimale supplémentaire, cette extrapolation donne une valeur *z* égale à 1,282. Cependant, dans la plupart des cas, l'utilisation de la valeur la plus proche de la probabilité souhaitée, contenue dans la table, est suffisamment précise.

aléatoire normale centrée réduite $Z$. Deux types de questions peuvent être posés. Le premier type spécifie une valeur ou des valeurs de $Z$ et implique l'utilisation de la table pour déterminer l'aire ou la probabilité correspondante. Le second type de question spécifie une aire ou une probabilité et implique l'utilisation de la table pour déterminer la valeur $z$ correspondante. Ainsi, la manière d'utiliser la table des probabilités de la loi normale centrée réduite varie selon la question posée. Dans la plupart des cas, représenter la loi normale centrée réduite et griser l'aire appropriée aide à visualiser le problème et à trouver la bonne réponse.

### 6.2.3 Calcul des probabilités d'une loi normale quelconque

Nous avons tant discuté de la loi normale centrée réduite parce que les probabilités de toute loi normale sont calculées à partir de cette loi centrée réduite. En effet, lorsqu'on a une distribution normale de moyenne $\mu$ et d'écart type $\sigma$, on commence par la convertir en distribution normale centrée réduite, pour répondre aux questions en matière de probabilités. Ensuite, on peut utiliser la table des probabilités normales centrées réduites et les valeurs appropriées de $Z$ pour trouver les probabilités souhaitées. La formule utilisée pour convertir toute variable aléatoire normale $X$, de moyenne $\mu$ et d'écart type $\sigma$, en une variable aléatoire normale centrée réduite, est :

► **Conversion en distribution normale centrée réduite**

$$z = \frac{x-\mu}{\sigma} \tag{6.3}$$

> La formule de la variable aléatoire normale centrée réduite est identique à celle introduite dans le chapitre 3, pour calculer la valeur centrée réduite $z$ pour un ensemble de données.

Si la variable aléatoire $X$ est égale à sa moyenne, alors la valeur de la variable aléatoire $Z$ est $z = (\mu - \mu)/\sigma = 0.$ En d'autres termes, si la variable aléatoire $X$ est égale à sa moyenne $\mu$, $Z$ est égale à sa moyenne 0. Maintenant, supposons que la variable aléatoire $X$ soit égale à sa moyenne plus un écart type, c'est-à-dire $x = \mu + \sigma.$ En appliquant la formule (6.3), la valeur correspondante de $Z$ est $z = [(\mu+\sigma)-\mu]/\sigma = \sigma/\sigma = 1.$ En d'autres termes, si $x = \mu + \sigma$ , $z = 1.$ De façon générale, on peut interpréter $z$ comme le nombre d'écarts type qui séparent la variable aléatoire $X$ de sa moyenne $\mu$.

Pour illustrer le fait que cette conversion nous permet de calculer des probabilités associées à toute distribution normale, supposons que la distribution normale soit de moyenne $\mu = 10$ et d'écart type $\sigma = 2.$ Quelle est la probabilité que la variable aléatoire $X$ soit comprise entre 10 et 14 ? En utilisant la formule (6.3), on voit que pour $x = 10,\quad z = (x-\mu)/\sigma = (10-10)/2 = 0$ et pour $x = 14,\quad z = (14-10)/2 = 4/2 = 2.$ Ainsi, la probabilité que la variable aléatoire $X$ soit comprise entre 10 et 14, est équivalente à la probabilité que la variable aléatoire $Z$ soit comprise entre 0 et 2. En d'autres termes, la probabilité que nous recherchons est la probabilité que la variable aléatoire $X$ soit comprise entre sa moyenne et deux écarts type au-dessus de sa moyenne. En utilisant $z = 2$ et la table des probabilités normales centrées réduites

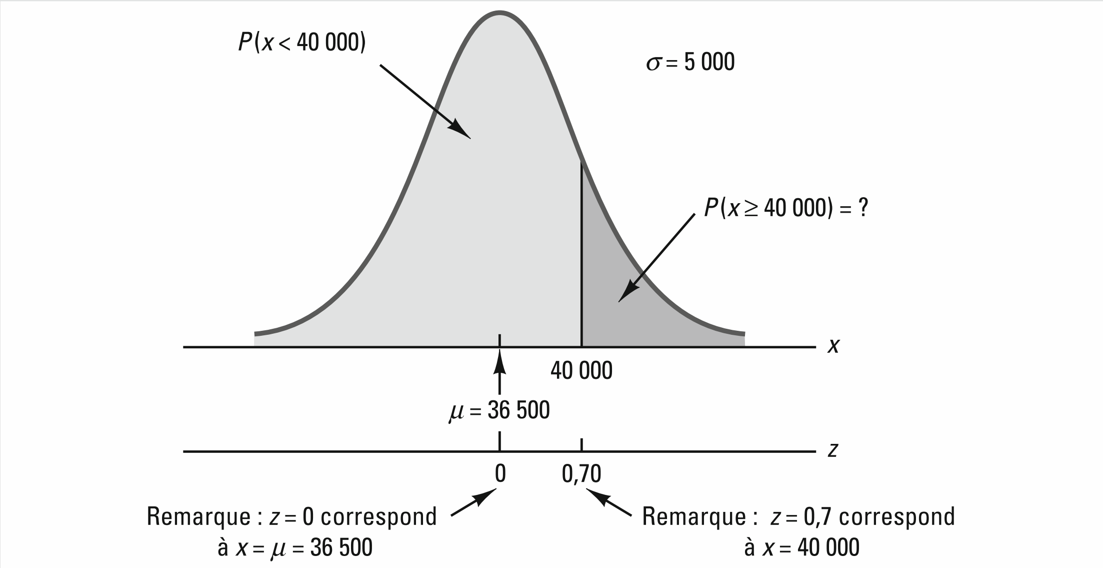



**Figure 6.6** ***Distribution du kilométrage pour le problème de la société Grear Tire***

en couverture du livre, on trouve que $P(z \leq 2) = 0{,}9772$. Puisque $P(z \leq 0) = 0{,}5$, $P(0 \leq z \leq 2) = P(z \leq 2) - P(z \leq 0) = 0{,}9772 - 0{,}5 = 0{,}4772$. Par conséquent, la probabilité que la variable aléatoire $X$ soit comprise entre 10 et 14 est égale à 0,4772.

## 6.2.4 Le problème de la société Grear Tire

Considérons à présent une application de la distribution de probabilité normale. Supposons que la société Grear Tire ait conçu un nouveau pneu radial, ceinturé d'acier, qui pourrait être vendu dans une chaîne nationale de magasins discount. Puisque le pneu est un nouveau produit, les responsables de Grear Tire pensent que la garantie du kilométrage effectué par le pneu serait un facteur déterminant dans la commercialisation du produit. Avant de définir le nombre de kilomètres garantis, les responsables de Grear veulent obtenir des informations en termes de probabilités sur le nombre de kilomètres que peut effectuer le pneu.

À partir des tests de route effectués avec les pneus, les ingénieurs de Grear ont estimé le kilométrage moyen du pneu à 36 500 km, avec un écart type de 5 000 km. De plus, les données collectées indiquent que l'on peut raisonnablement supposer que la distribution est normale. Quel est le pourcentage de pneus qui peuvent effectuer plus de 40 000 km ? En d'autres termes, quelle est la probabilité que le kilométrage effectué par un pneu excède 40 000 km ? On peut répondre à cette question en calculant l'aire de la partie grisée de la figure 6.6.

Pour $x = 40\,000$,

$$z = \frac{x - \mu}{\sigma} = \frac{40000 - 36500}{5000} = \frac{3500}{5000} = 0{,}70$$

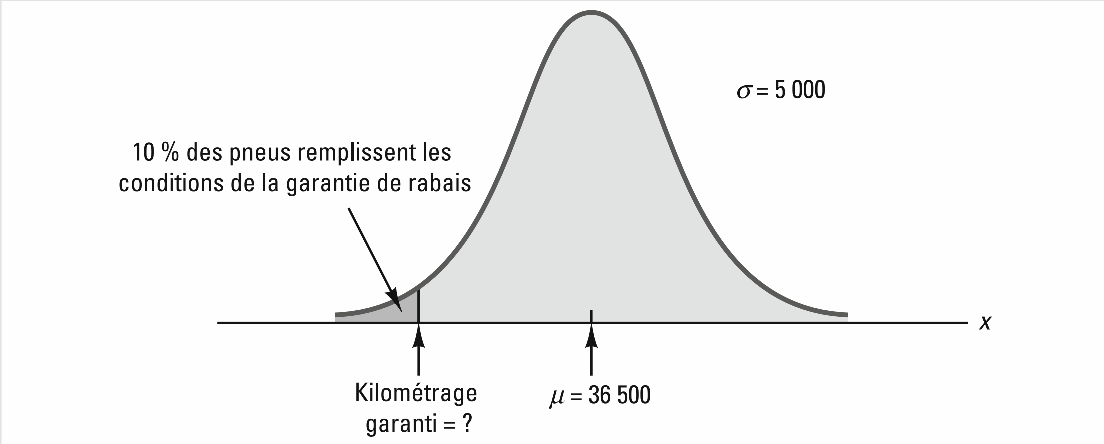



**Figure 6.7** ***Garantie de rabais de la société Grear***

En nous référant au bas de la figure 6.6, nous voyons qu'une valeur de la variable aléatoire $X$ égale à 40 000 correspond à une valeur de la variable normale centrée réduite $Z$ égale à 0,70. En utilisant la table de probabilité centrée réduite, nous constatons que l'aire sous la courbe normale à gauche de $z = 0,70$ est égale à 0,7580. Ainsi, $1 - 0,7580 = 0,2420$ est la probabilité que $z$ soit supérieur à 0,70 et donc que $x$ soit supérieur à 40 000. On peut conclure qu'environ 24,2 % des pneus auront un kilométrage supérieur à 40 000 km.

Supposons maintenant que Grear étudie la mise en place d'une garantie qui offre le remplacement des pneus à tarif réduit si les pneus originaux ne dépassent pas le kilométrage garanti. Quelle devrait être le kilométrage garanti pour qu'au plus 10 % des pneus n'effectuent pas le nombre de kilomètres garantis ? Cette question est interprétée graphiquement à la figure 6.7.

Selon la figure 6.7, l'aire sous la courbe à gauche du kilométrage garanti inconnu doit être égale à 0,10. Nous devons donc trouver la valeur $z$ qui correspond à une aire de 0,10 dans la queue inférieure de la distribution normale centrée réduite. En utilisant la table des probabilités normales centrées réduites, nous constatons que $z = -1,28$ est la valeur de la variable aléatoire normale centrée réduite correspondant au kilométrage garanti souhaité. Pour trouver le kilométrage $x$ correspondant à $z = -1,28$, nous avons :

$$z = \frac{x - \mu}{\sigma} = -1,28$$

$$x - \mu = -1,28\sigma$$

$$x = \mu - 1,28\sigma$$

Le kilométrage garanti que nous devons trouver se situe à 1,28 écart type en-dessous de la moyenne. Ainsi, $x = \mu - 1,28\sigma$.

Avec $\mu = 36\ 500$ et $\sigma = 5\ 000$,

$$x = 36500 - (1,28 \times 5000) = 30\,100$$

Ainsi, une garantie de 30 100 km satisfait la condition selon laquelle 10 % des pneus n'effectueraient pas le nombre de kilomètres garantis. Aux vues de ces informations, l'entreprise fixera peut-être sa garantie de kilométrage à 30 000 km.

Avec une garantie fixée à 30 000 km, le pourcentage réel de pneus qui ne respectent pas la garantie s'élève à 9,68 %.

De nouveau, nous constatons le rôle majeur des distributions de probabilité dans le processus d'aide à la décision. Une fois la distribution de probabilité établie pour une application particulière, elle peut être utilisée rapidement et facilement pour obtenir des informations probabilistes sur le problème. Les probabilités ne permettent pas de prendre directement une décision mais fournissent des informations qui aident le responsable à mieux comprendre et mesurer les risques et les incertitudes liés au problème. En fin de compte, cette information peut aider le responsable à prendre la bonne décision.

## EXERCICES

### *Méthode*

**8.** En vous référant à la figure 6.4, dessiner la courbe normale d'une variable aléatoire $X$ de moyenne $\mu$ égale à 100 et d'écart type $\sigma$ égal à 10. Inscrire les valeurs 70, 80, 90, 100, 110, 120 et 130 sur l'axe des abscisses.

**9.** Une variable aléatoire est normalement distribuée, avec une moyenne $\mu$ égale à 50 et un écart type $\sigma$ égal à 5.

- **a)** Dessiner la courbe normale de la fonction de densité. Inscrire les valeurs 35, 40, 45, 50, 55, 60 et 65 sur l'axe des abscisses. La figure 6.4 montre que la courbe normale touche presque l'axe des abscisses lorsqu'elle est à trois écarts type de part et d'autre de la moyenne (dans ce cas, aux points d'abscisse 35 et 65).
- **b)** Quelle est la probabilité que la variable aléatoire prenne une valeur comprise entre 45 et 55 ?
- **c)** Quelle est la probabilité que la variable aléatoire prenne une valeur comprise entre 40 et 60 ?

**10.** Représenter une distribution normale centrée réduite. Inscrire les valeurs –3, –2, –1, 0, 1, 2 et 3 sur l'axe des abscisses. Utiliser ensuite la table des probabilités de la loi normale centrée réduite pour calculer les probabilités suivantes :

- **a)** $P(z \leq 1,5)$
- **b)** $P(z \leq 1)$
- **c)** $P(1 \leq z \leq 1,5)$
- **d)** $P(0 < z < 2,5)$

**11.** Étant donné que $Z$ est une variable aléatoire normale centrée réduite, calculer les probabilités suivantes :

**a)** $P(z \leq -1)$
**b)** $P(z \geq -1)$
**c)** $P(z \geq -1{,}5)$
**d)** $P(z \geq -2{,}5)$
**e)** $P(-3 < z \leq 0)$

**12.** Étant donné que $Z$ est une variable aléatoire normale centrée réduite, calculer les probabilités suivantes :

**a)** $P(0 \leq z \leq 0{,}83)$
**b)** $P(-1{,}57 \leq z \leq 0)$
**c)** $P(z > 0{,}44)$
**d)** $P(z \geq -0{,}23)$
**e)** $P(z < 1{,}20)$
**f)** $P(z \leq -0{,}71)$

**13.** Étant donné que $Z$ est une variable aléatoire normale centrée réduite, calculer les probabilités suivantes :

**a)** $P(-1{,}98 \leq z \leq 0{,}49)$
**b)** $P(0{,}52 \leq z \leq 1{,}22)$
**c)** $P(-1{,}75 \leq z \leq -1{,}04)$

**14.** Étant donné que $Z$ est une variable aléatoire normale centrée réduite, trouver la valeur $z$ de $Z$ dans les cas suivants :

**a)** L'aire à gauche de $z$ est égale à 0,9750.
**b)** L'aire entre 0 et $z$ est égale à 0,4750.
**c)** L'aire à gauche de $z$ est égale à 0,7291.
**d)** L'aire à droite de $z$ est égale à 0,1314.
**e)** L'aire à gauche de $z$ est égale à 0,67.
**f)** L'aire à droite de $z$ est égale à 0,33.

**15.** Étant donné que $Z$ est une variable aléatoire normale centrée réduite, trouver la valeur $z$ de $Z$ dans les cas suivants :

**a)** L'aire à gauche de $z$ est égale à 0,2119.
**b)** L'aire entre $-z$ et $z$ est égale à 0,9030.
**c)** L'aire entre $-z$ et $z$ est égale à 0,2052.
**d)** L'aire à gauche de $z$ est égale à 0,9948.
**e)** L'aire à droite de $z$ est égale à 0,6915.

**16.** Étant donné que $Z$ est une variable aléatoire normale centrée réduite, trouver la valeur $z$ de $Z$ dans les cas suivants :

a) L'aire à droite de $z$ est égale à 0,01.
b) L'aire à droite de $z$ est égale à 0,025.
c) L'aire à droite de $z$ est égale à 0,05.
d) L'aire à droite de $z$ est égale à 0,10.

## Applications

**17.** Le coût moyen des vols domestiques aux États-Unis a atteint un niveau record de 385 dollars par billet (site Internet du bureau des statistiques sur le transport, 2 novembre 2012). Les tarifs considérés incluent le prix pratiqué par les compagnies aériennes et toutes les taxes additionnelles. Supposez que ces tarifs domestiques soient distribués selon une loi normale ayant un écart type de 110 dollars.

a) Quelle est la probabilité qu'un tarif domestique soit supérieur ou égal à 550 dollars ?
b) Quelle est la probabilité qu'un tarif domestique soit inférieur ou égal à 250 dollars ?
c) Quelle est la probabilité qu'un tarif domestique soit compris entre 300 et 500 dollars ?
d) Quel est le montant des 3 % des tarifs domestiques les plus élevés ?

**18.** Le rendement moyen des actions domestiques sur les trois années 2009-2011 était de 14,4 % (*AAII Journal*, février 2012). Supposez que le rendement sur trois ans soit normalement distribué parmi les actions, avec un écart type de 4,4 %.

a) Quelle est la probabilité qu'une action domestique particulière ait eu un rendement sur les trois années considérées d'au moins 20 % ?
b) Quelle est la probabilité qu'une action domestique particulière ait eu un rendement sur les trois années considérées d'au plus 10 % ?
c) Quel aurait dû être le rendement pour qu'une action domestique fasse partie des 10 % les plus rentables sur la période considérée ?

**19.** Dans un article sur le coût des soins médicaux, le magazine *Money* rapportait qu'une visite aux urgences d'un hôpital pour quelque chose d'aussi banal qu'un mal de gorge coûtait en moyenne 328 dollars (*Money*, janvier 2009). Supposez que le coût de ce type de visite aux urgences soit normalement distribué avec un écart type de 92 dollars. Répondre aux questions suivantes.

a) Quelle est la probabilité que le coût soit supérieur à 500 dollars ?
b) Quelle est la probabilité que le coût soit inférieur à 250 dollars ?
c) Quelle est la probabilité que le coût soit compris entre 300 et 400 dollars ?
d) Si le coût d'un patient représente moins de 8 % des charges de ce service médical, quel est le coût de la visite de ce patient aux urgences ?

**20.** Le prix moyen d'un gallon d'essence est de 3,73 dollars aux États-Unis et 3,40 dollars en Russie (*Bloomberg Business*, 5-11 mars 2012). Supposez que ces moyennes correspondent aux moyennes de la population dans les deux pays et que les distributions de probabilité sont normalement distribuées avec un écart type de 0,25 dollar aux États-Unis et de 0,20 dollar en Russie.

**a)** Quelle est la probabilité qu'une station-service sélectionnée aléatoirement sur le territoire américain pratique un prix inférieur à 3,50 dollars le gallon ?

**b)** Quel pourcentage de stations-service russes pratique un prix inférieur à 3,50 dollars le gallon ?

**c)** Quelle est la probabilité qu'une station-service sélectionnée aléatoirement en Russie pratique un prix supérieur au prix moyen pratiqué aux États-Unis ?

**21.** Pour devenir membre de Mensa, association internationale des personnes ayant un quotient intellectuel élevé, une personne doit obtenir une note au test de QI se situant parmi les 2 % des notes de la population les plus élevées. L'association compte 110 000 membres dans 100 pays à travers le monde (site Internet de Mensa International, 8 janvier 2013). Si les notes sont normalement distribuées, avec une moyenne de 100 et un écart type de 15, quelle note doit obtenir une personne pour devenir membre de l'association Mensa ?

**22.** Le temps passé à regarder la télévision a atteint un nouveau record lorsque la société Nielsen a estimé le temps moyen passé à regarder la télévision à 8,35 heures par jour par ménage (*USA Today*, 11 novembre 2009). Utiliser une distribution de probabilité normale avec un écart type de 2,5 heures pour répondre aux questions suivantes relatives au nombre d'heures quotidiennes qu'un ménage passe à regarder la télévision.

**a)** Quelle est la probabilité qu'un ménage passe entre 5 et 10 heures par jour devant sa télévision?

**b)** À combien devrait s'élever le nombre d'heures passées à regarder la télévision par un ménage pour qu'il soit parmi les 3 % regardant le plus la télévision ?

**c)** Quelle est la probabilité qu'un ménage regarde la télévision plus de 3 heures par jour ?

**23.** Le temps nécessaire pour passer l'examen de fin d'année dans un lycée est normalement distribué avec une moyenne de 80 minutes et un écart type de 10 minutes. Répondre aux questions suivantes :

**a)** Quelle est la probabilité de finir l'examen en au plus une heure ?

**b)** Quelle est la probabilité qu'un étudiant finisse l'examen en plus de 60 minutes mais moins de 75 minutes ?

**c)** Supposez que la classe contienne 60 élèves et que la durée de l'examen soit fixée à 90 minutes. Combien d'étudiants ne seront pas capables de finir l'examen dans le temps imparti ?

**24.** L'Association Américaine de l'Automobile (AAA) rapportait que les familles qui ont prévu de voyager durant le week-end de la fête du travail, dépenseraient en moyenne 749 dollars (The Asssociated Press, 12 août 2012). Supposez que le montant dépensé soit normalement distribué avec un écart type de 225 dollars.

**a)** Quelle est la probabilité que les dépenses d'une famille durant ce week-end soient inférieures à 400 dollars ?

**b)** Quelle est la probabilité que les dépenses d'une famille durant ce week-end soient supérieures ou égales à 800 dollars ?

**c)** Quelle est la probabilité que les dépenses d'une famille durant ce week-end soient comprises entre 500 et 1 000 dollars ?

**d)** Quelles sont les dépenses des 5 % des familles qui ont les projets de voyage les plus onéreux ?

**25.** New York est la ville la plus chère des États-Unis en termes d'hébergement. Le prix moyen d'une chambre d'hôtel est de 204 dollars par nuit (*USA Today*, 30 avril 2012). Supposez que les prix des chambres soient normalement distribués avec un écart type de 55 dollars.

**a)** Quelle est la probabilité qu'une chambre d'hôtel coûte au moins 225 dollars par nuit ?

**b)** Quelle est la probabilité qu'une chambre d'hôtel coûte au plus 140 dollars par nuit ?

**c)** Quelle est la probabilité qu'une chambre d'hôtel coûte entre 200 et 300 dollars par nuit ?

**d)** Quel est le prix des 20 % des chambres les plus chères de New York ?

## 6.3 APPROXIMATION NORMALE DES PROBABILITÉS BINOMIALES

Dans la section 5.5, nous avons présenté la loi discrète binomiale. Rappelons qu'une expérience binomiale est une séquence de $n$ tirages identiques et indépendants, qui ont deux issues possibles, un succès et un échec. La probabilité d'un succès est la même pour tous les tirages et est notée $p$. La variable aléatoire binomiale correspond au nombre de succès obtenus en $n$ tirages, et les questions probabilistes se rapportent à la probabilité de $x$ succès en $n$ tirages.

Lorsque le nombre de tirages devient important, la fonction de probabilité binomiale devient difficile à calculer, que ce soit à la main ou avec une calculatrice. Dans les cas où $np \geq 5$ et $n(1-p) \geq 5$, la loi normale permet d'estimer facilement des probabilités binomiales. Pour ce faire, on pose $\mu = np$ et $\sigma = \sqrt{np(1-p)}$ afin de définir la courbe normale.

Illustrons l'approximation normale de la loi binomiale en supposant qu'une société fait des erreurs, d'après les données collectées, dans 10 % de ses factures. Un échantillon de 100 factures est sélectionné ; nous voulons calculer la probabilité que 12 factures contiennent des erreurs. C'est-à-dire, nous voulons trouver la probabilité binomiale de 12 succès en 100 tirages. En appliquant l'approximation normale de la loi binomiale à ce cas, on pose $\mu = np = 100 \times 0{,}1 = 10$ et $\sigma = \sqrt{np(1-p)} = \sqrt{100 \times 0{,}1 \times 0{,}9} = 3$. Une distribution normale avec $\mu = 10$ et $\sigma = 3$ est représentée à la figure 6.8.

Rappelons qu'avec une loi continue, les probabilités correspondent à l'aire sous la fonction de densité. Par conséquent, la probabilité d'une valeur isolée est nulle. Pour estimer la probabilité binomiale de 12 succès, on doit calculer l'aire sous la courbe normale comprise entre 11,5 et 12,5. Les 0,5 que l'on ajoute et soustrait à 12 sont appelés **facteur de correction de la continuité**. Ce facteur de correction est introduit car on utilise une loi continue pour approcher une loi discrète. Ainsi, $P(x = 12)$ pour la loi binomiale *discrète* est estimée par $P(11{,}5 \leq x \leq 12{,}5)$ pour la loi normale *continue*.

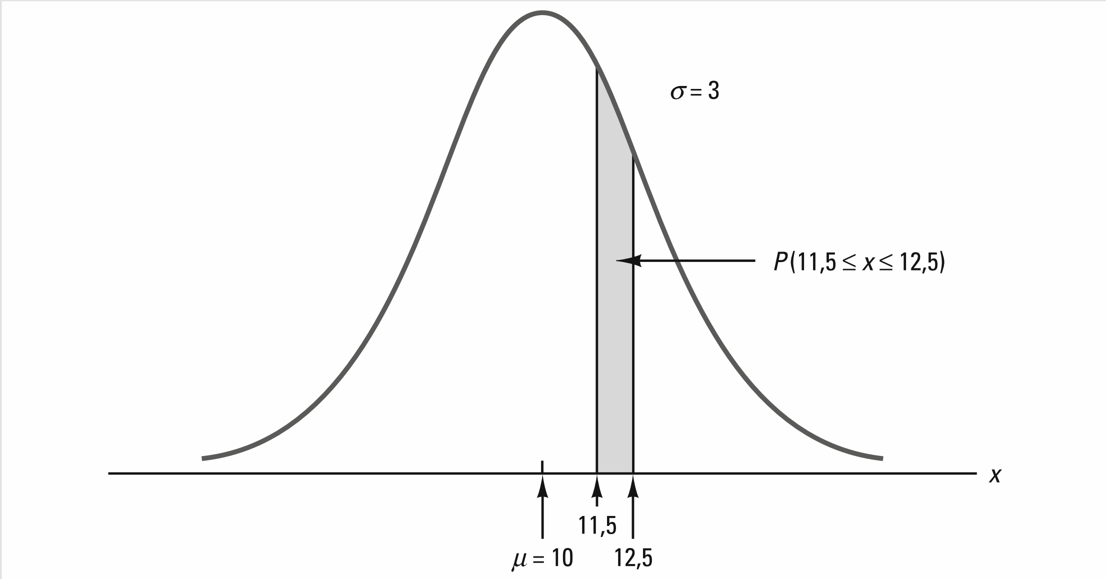



**Figure 6.8** *Approximation normale de la loi binomiale avec $n = 100$ et $p = 0{,}10$, donnant la probabilité de 12 erreurs*

En convertissant la loi normale en loi normale centrée réduite pour calculer $P(11{,}5 \leq x \leq 12{,}5)$ nous avons

$$z = \frac{x - \mu}{\sigma} = \frac{12{,}5 - 10}{3} = 0{,}83 \qquad \text{pour } x = 12{,}5$$

et

$$z = \frac{x - \mu}{\sigma} = \frac{11{,}5 - 10}{3} = 0{,}50 \qquad \text{pour } x = 11{,}5$$

Grâce à la table des probabilités normales centrées réduites, nous trouvons que l'aire sous la courbe (figure 6.8) à gauche de 12,5 est égale à 0,7967. De manière similaire, l'aire sous la courbe à gauche de 11,5 est égale à 0,6915. Par conséquent, l'aire comprise entre 11,5 et 12,5 est égale à 0,1052 ($0{,}7967 - 0{,}6915 = 0{,}1052$). L'approximation normale de la probabilité de 12 succès en 100 tirages est égale à 0,1052.

Considérons un autre exemple. Supposons que l'on veuille calculer la probabilité d'au plus 13 erreurs dans l'échantillon de 100 factures. La figure 6.9 représente l'aire sous la courbe normale qui estime cette probabilité. Notez que le facteur de correction de la continuité impose l'utilisation de la valeur 13,5 pour calculer la probabilité désirée. La valeur $z$ correspondant à $x = 13{,}5$ est

$$z = \frac{13{,}5 - 10}{3} = 1{,}17$$

Selon la table des probabilités normales centrées réduites, l'aire sous la courbe normale à gauche de 1,17 est égale à 0,8790. L'aire sous la courbe normale estimant la probabilité d'au plus 13 erreurs est représentée par la partie grisée de la figure 6.9.

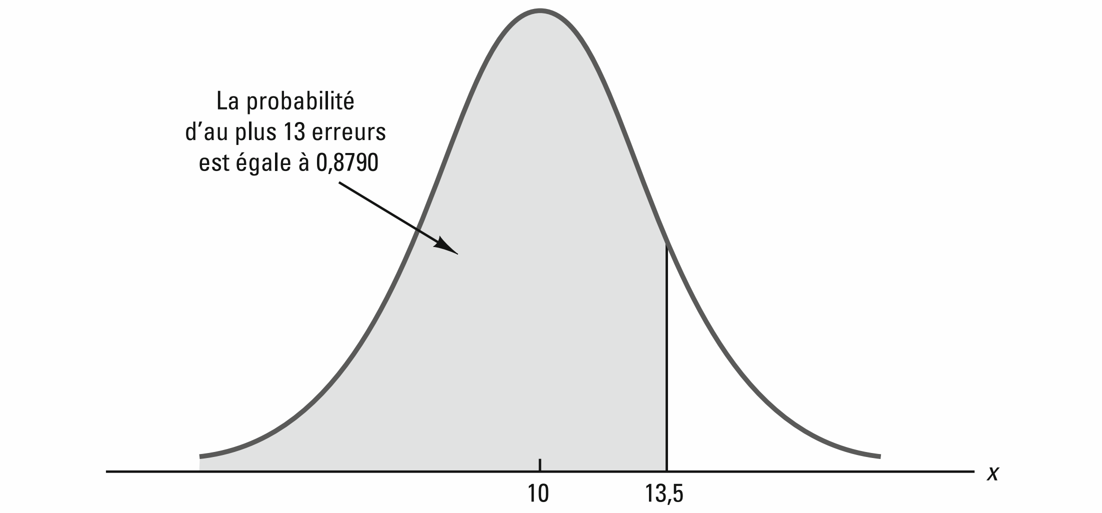



**Figure 6.9** *Approximation normale de la loi binomiale avec $n = 100$ et $p = 0,10$, donnant la probabilité d'au plus 13 erreurs*

## EXERCICES

### *Méthode*

**26.** Une loi binomiale a les caractéristiques suivantes : $p = 0,2$ et $n = 100$.

**a)** Quelle est la moyenne ? Quel est l'écart type ?

**b)** Dans cette situation, les probabilités binomiales peuvent-elles être estimées par la loi normale ? Expliquez.

**c)** Quelle est la probabilité d'exactement 24 succès ?

**d)** Quelle est la probabilité que le nombre de succès soit compris entre 18 et 22 ?

**e)** Quelle est la probabilité que le nombre de succès soit inférieur ou égal à 15 ?

**27.** Une loi binomiale a les caractéristiques suivantes : $p = 0,6$ et $n = 200$.

**a)** Quelle est la moyenne ? Quel est l'écart type ?

**b)** Dans cette situation, les probabilités binomiales peuvent-elles être estimées par la loi normale ? Expliquez.

**c)** Quelle est la probabilité que le nombre de succès soit compris entre 100 et 110 ?

**d)** Quelle est la probabilité que le nombre de succès soit supérieur ou égal à 130 ?

**e)** Quel est l'avantage d'utiliser la loi normale pour estimer les probabilités binomiales ? Utiliser la question (d) pour répondre.

## *Applications*

**28.** Bien que les études prouvent que fumer génère de graves problèmes de santé, 20 % des adultes américains fument. Considérez un groupe de 250 adultes.

**a)** Quelle est l'espérance mathématique du nombre d'adultes qui fument ?

**b)** Quelle est la probabilité que moins de 40 adultes fument ?

**c)** Quelle est la probabilité qu'entre 55 et 60 adultes fument ?

**d)** Quelle est la probabilité qu'au moins 70 adultes fument ?

**29.** Selon une enquête du comité de surveillance du centre des impôts, 82 % des contribuables ont déclaré qu'il était très important que le service de recouvrement des impôts s'assure que les contribuables à hauts revenus ne trichent pas dans leur déclaration (*The Wall Street Journal*, 11 février 2009).

**a)** Pour un échantillon de huit contribuables, quelle est la probabilité qu'au moins six d'entre eux déclarent qu'il est très important de s'assurer que les contribuables à hauts revenus ne trichent pas ? Utiliser l'approximation normale de la loi binomiale pour répondre à cette question.

**b)** Pour un échantillon de 80 contribuables, quelle est la probabilité qu'au moins 60 d'entre eux déclarent qu'il est très important de s'assurer que les contribuables à hauts revenus ne trichent pas ? Utiliser l'approximation normale de la loi binomiale pour répondre à cette question.

**c)** Lorsque le nombre de tirages dans une application de la loi binomiale devient important, quel est l'avantage d'utiliser l'approximation normale de la loi binomiale pour calculer les probabilités ?

**d)** Lorsque le nombre de tirages dans une application de la loi binomiale devient important, les développeurs de logiciels statistiques préfèrent-ils utiliser la fonction de distribution binomiale présentée à la section 5.4 ou l'approximation normale de cette loi présentée à la section 6.3 ? Expliquer.

**30.** Les jeux vidéo sont très populaires. Plus de 70 % des ménages y jouent. Parmi les joueurs, 18 % ont moins de 18 ans, 53 % ont entre 18 et 59 ans et 29 % ont plus de 59 ans (*The Wall Street Journal*, 6 mars 2012).

**a)** Sur un échantillon de 800 joueurs, combien de personnes en moyenne ont moins de 18 ans ?

**b)** Sur un échantillon de 600 joueurs, quelle est la probabilité qu'au plus 100 joueurs aient moins de 18 ans ?

**c)** Sur un échantillon de 800 joueurs, quelle est la probabilité qu'au moins 200 joueurs aient plus de 59 ans ?

**31.** Selon une enquête du bureau des affaires nationales (*USA Today*, 12 novembre 2009), 79 % des employeurs octroient à leurs employés deux jours de congés payés lors de Thanksgiving (le jeudi et le vendredi sont des jours chômés). Quatre-vingt-dix pourcent des employeurs octroient un jour de congé payé à leurs employés (le jour de Thanksgiving). Deux pourcent des employeurs n'octroient pas de congés payés à cette occasion. Considérez un échantillon de 120 employeurs.

**a)** Quelle est la probabilité qu'au moins 85 des employeurs octroient deux jours de congés payés ?

**b)** Quelle est la probabilité qu'entre 90 et 100 employeurs octroient deux jours de congés payés ? C'est-à-dire que vaut $P(90 \le x \le 100)$ ?

**c)** Quelle est la probabilité que moins de 20 employeurs octroient un jour de congé payé ?

# 6.4 LA LOI EXPONENTIELLE

La **loi exponentielle** peut être utilisée pour décrire des variables aléatoires telles que le temps entre les arrivées à une station de lavage, le temps nécessaire pour charger un camion, la distance entre les défauts majeurs sur une autoroute, etc. La fonction de densité exponentielle s'écrit :

► **Fonction de densité de probabilité exponentielle**

$$f(x) = \frac{1}{\mu} e^{-x/\mu} \quad \text{pour } x \ge 0, \quad \mu \ge 0 \tag{6.4}$$

où $\mu$ est la valeur espérée ou moyenne

Comme exemple de la loi exponentielle, supposons que le temps de chargement d'un camion sur les docks de Schips suive une telle distribution. Si le temps moyen de chargement d'un camion est de 15 minutes ( $\mu = 15$ ), la fonction de densité appropriée s'écrit :

$$f(x) = \frac{1}{15} e^{-x/15}$$

La figure 6.10 représente cette fonction de densité.

## 6.4.1 *Calcul des probabilités d'une loi exponentielle*

Comme pour toute loi continue, l'aire sous la courbe dans un intervalle donné fournit la probabilité que la variable aléatoire prenne une valeur appartenant à cet intervalle. Dans l'exemple des docks de Schips, la probabilité qu'un camion soit chargé en au plus 6 minutes, $P(x \le 6)$, correspond à l'aire sous la courbe, représentée par la figure 6.10, comprise entre $x = 0$ et $x = 6$. De même, la probabilité qu'un camion soit chargé en au plus 18 minutes $P(x \le 18)$ correspond à l'aire sous la courbe comprise entre $x = 0$ et $x = 18$. Notez aussi que la probabilité que le temps de chargement du camion soit compris entre 6 et 18 minutes $P(6 \le x \le 18)$ correspond à l'aire sous la courbe comprise entre $x = 6$ et $x = 18$.

Dans les exemples sur les files d'attente, la distribution exponentielle est souvent utilisée pour le temps de service.

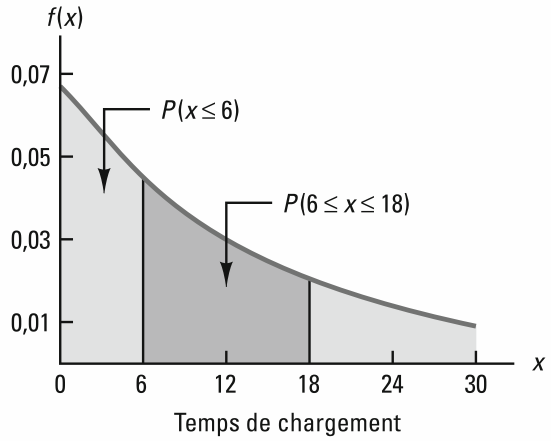



**Figure 6.10** ***Loi exponentielle pour l'exemple des docks de Schips***

Pour calculer les probabilités exponentielles comme celles décrites ci-dessus, on utilise la formule suivante. Elle fournit la probabilité cumulée d'obtenir une valeur inférieure ou égale à une valeur donnée de la variable aléatoire exponentielle, notée $x_0$.

► **Loi exponentielle : probabilités cumulées**

$$P\left(x \le x_0\right) = 1 - e^{-x_0/\mu} \tag{6.5}$$

Pour l'exemple des docks de Schips, $x$ = temps de chargement (en minutes) et $\mu = 15$ minutes, ce qui implique :

$$P\left( x \le x_0\right) = 1 - e^{-x_0/15}$$

Par conséquent, la probabilité que le temps de chargement d'un camion prenne, au plus, 6 minutes est égale à

$$P\left(x \le 6\right) = 1 - e^{-6/15} = 0,3297$$

La probabilité de charger un camion en au plus 18 minutes est égale à :

$$P\left(x \le 18\right) = 1 - e^{-18/15} = 0,6988$$

Ainsi, la probabilité que le temps de chargement d'un camion soit compris entre 6 et 18 minutes est égale à 0,3691 ($0,6988 - 0,3297 = 0,3691$). Les probabilités pour tout autre intervalle peuvent être calculées de la même façon.

Dans l'exemple précédent, le temps moyen de chargement d'un camion est de 15 minutes. Une propriété de la loi exponentielle implique que la moyenne et l'écart type de la distribution sont *égaux*. Ainsi, l'écart type du temps de chargement d'un camion est $\sigma = 15$ minutes. La variance est égale à $\sigma^2 = (15)^2 = 225$.

Une propriété de la loi exponentielle est l'égalité de la moyenne et de l'écart type.

## 6.4.2 Relation entre les distributions de Poisson et exponentielle

Dans la section 5.5, nous avons introduit la loi de Poisson en tant que loi de probabilité discrète, utile pour examiner le nombre d'occurrences d'un événement dans un intervalle de temps ou d'espace donné. Rappelons que la fonction de probabilité de Poisson s'écrit :

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}$$

où $\mu$ est l'espérance mathématique ou le nombre moyen d'occurrences dans un intervalle.

La loi exponentielle, continue, est liée à la loi de Poisson, discrète. Si la distribution de Poisson fournit une bonne description du nombre d'occurrences par intervalle, la distribution exponentielle fournit une description de la longueur de l'intervalle entre les occurrences.

> Si les arrivées suivent une loi de Poisson, le temps écoulé entre deux arrivées doit suivre une loi exponentielle.

Pour illustrer cette relation, supposons que le nombre de voitures qui arrivent à une station de lavage en une heure est décrit par une distribution de Poisson de moyenne égale à 10 voitures par heure. La fonction de probabilité de Poisson qui donne la probabilité de $x$ arrivées en une heure est :

$$f(x) = \frac{10^x e^{-10}}{x!}$$

Puisque le nombre moyen d'arrivées par heure est égal à 10, le temps moyen entre deux arrivées est :

$$\frac{1 \text{ heure}}{10 \text{ voitures}} = 0{,}1 \text{ heure/voiture}$$

Ainsi, la distribution exponentielle, qui décrit le temps entre les arrivées, a une moyenne égale à 0,1 heure par voiture ; la fonction de densité exponentielle est alors

$$f(x) = \frac{1}{0{,}1} e^{-x/0{,}1} = 10e^{-10x}$$

**REMARQUES**

Comme nous pouvons le voir sur la figure 6.10, la distribution exponentielle est asymétrique à droite. Le coefficient d'asymétrie pour des distributions exponentielles est égal à 2. La distribution exponentielle est une parfaite illustration d'une distribution asymétrique.

**EXERCICES**

## *Méthode*

**32.** Considérer la fonction de densité de probabilité exponentielle suivante :

$$f(x) = \frac{1}{8}e^{-x/8} \qquad \text{pour } x \geq 0$$

- **a)** Trouver $P(x \leq 6)$.
- **b)** Trouver $P(x \leq 4)$.
- **c)** Trouver $P(x \geq 6)$.
- **d)** Trouver $P(4 \leq x \leq 6)$.

**33.** Considérer la fonction de densité de probabilité exponentielle suivante :

$$f(x) = \frac{1}{3}e^{-x/3} \qquad \text{pour } x \geq 0$$

- **a)** Écrire la formule pour $P(x \leq x_0)$.
- **b)** Trouver $P(x \leq 2)$.
- **c)** Trouver $P(x \geq 3)$.
- **d)** Trouver $P(x \leq 5)$.
- **e)** Trouver $P(2 \leq x \leq 5)$.

## *Applications*

**34.** La durée d'autonomie de la batterie du Motorola Droid Razr Maxx est de 20 heures lorsque l'appareil est utilisé pour téléphoner (*The Wall Street Journal*, 7 mars 2012). La durée d'autonomie de la batterie tombe à 7 heures lorsque le téléphone est principalement utilisé pour surfer sur Internet. Supposez que la durée d'autonomie de la batterie pour les deux usages suive une loi exponentielle.

- **a)** Quelle est la fonction de densité de probabilité de la durée d'autonomie du téléphone lorsqu'il est utilisé pour téléphoner ?
- **b)** Quelle est la probabilité que la durée d'autonomie de la batterie d'un téléphone Droid Razr Maxx sélectionné aléatoirement soit inférieure ou égale à 15 heures lorsqu'il est utilisé principalement pour téléphoner ?
- **c)** Quelle est la probabilité que la durée d'autonomie de la batterie d'un téléphone Droid Razr Maxx sélectionné aléatoirement soit supérieure à 20 heures lorsqu'il est utilisé principalement pour téléphoner ?
- **d)** Quelle est la probabilité que la durée d'autonomie de la batterie d'un téléphone Droid Razr Maxx sélectionné aléatoirement soit inférieure ou égale à 5 heures lorsqu'il est utilisé principalement pour surfer sur Internet ?

**35.** Le temps qui s'écoule entre l'arrivée de deux véhicules à un carrefour particulier suit une loi exponentielle avec une moyenne de 12 secondes.

**a)** Représenter cette distribution de probabilité exponentielle.

**b)** Quelle est la probabilité que le temps qui s'écoule entre l'arrivée de deux véhicules soit inférieur ou égal à 12 secondes ?

**c)** Quelle est la probabilité que le temps qui s'écoule entre l'arrivée de deux véhicules soit inférieur ou égal à 6 secondes ?

**d)** Quelle est la probabilité que le temps qui s'écoule entre l'arrivée de deux véhicules soit supérieur ou égal à 30 secondes ?

**36.** La société Comcast est la plus importante société de télévision par câble, le deuxième fournisseur Internet et le quatrième fournisseur de services de téléphonie aux États-Unis. Généralement connue pour la qualité et la fiabilité de ses services, la société connaît périodiquement des interruptions de service involontaires. Le 14 janvier 2009, une telle interruption s'est produite pour les clients de Comcast vivant en Floride. Lorsque les abonnés ont appelé le service client, un message enregistré leur disait que la société était consciente du problème d'interruption du service et qu'elle espérait rétablir la situation dans les deux heures. Supposez que deux heures correspondent au temps moyen nécessaire pour effectuer la réparation et que le temps de réparation suive une loi exponentielle.

**a)** Quelle est la probabilité que le service de télévision par câble soit restauré en une heure au maximum ?

**b)** Quelle est la probabilité que la réparation prenne entre une et deux heures ?

**c)** Pour un client qui appelle le service client de Comcast à 13 heures, quelle est la probabilité que le service de télévision ne soit pas restauré à 17 heures ?

**37.** Le magasin de café italien Collina à Houston au Texas annonce que la préparation des commandes prend environ 25 minutes (site Internet de Collina, 27 février 2008). Supposez que le temps nécessaire pour qu'une commande soit prête, suive une loi exponentielle de moyenne égale à 25 minutes.

**a)** Quelle est la probabilité que la préparation d'une commande prenne moins de 20 minutes ?

**b)** Si un client vient chercher sa commande 30 minutes après l'avoir passée, quelle est la probabilité que la commande ne soit pas prête?

**c)** Un client particulier vit à 15 minutes du magasin. Si le client passe commande à 17h20, quelle est la probabilité que le client puisse venir au magasin, retirer sa commande et être de retour chez lui à 18h ?

**38.** Les pompiers de Boston reçoivent des appels d'urgence au taux moyen de 1,6 appel par heure (site Internet Mass.gov, novembre 2012). Supposez que le nombre d'appels par heure suive une loi de Poisson.

**a)** Quelle est la durée moyenne en minutes entre deux appels reçus par les pompiers de Boston ?

**b)** En utilisant la moyenne obtenue à la question (a), déterminer la fonction de densité de probabilité de la durée en minutes entre deux appels d'urgence.

**c)** Quelle est la probabilité qu'il s'écoule moins d'une heure entre deux appels d'urgence ?

**d)** Quelle est la probabilité qu'il s'écoule au moins 30 minutes entre deux appels d'urgence ?

**e)** Quelle est la probabilité qu'il s'écoule plus de 5 minutes mais moins de 20 minutes entre deux appels d'urgence ?

## RÉSUMÉ

Ce chapitre a étendu la discussion des distributions de probabilité au cas des variables aléatoires continues. La différence majeure entre les distributions de probabilités discrètes et continues se situe au niveau de la méthode de calcul des probabilités. La fonction de probabilité pour des variables aléatoires discrètes $f(x)$ fournit la probabilité que la variable aléatoire *X* prenne différentes valeurs. Avec des distributions continues, la fonction de densité de probabilité $f(x)$ ne fournit pas directement les probabilités. Celles-ci sont déterminées par l'aire sous la courbe de la fonction de densité $f(x)$. Puisque l'aire sous la courbe pour un point isolé est nulle, la probabilité qu'une variable aléatoire continue prenne une valeur isolée est nulle.

Trois lois continues – les lois uniforme, normale et exponentielle – ont été traitées en détail. La loi normale est fréquemment utilisée en inférence statistique et sera beaucoup utilisée dans la suite de cet ouvrage.

## GLOSSAIRE

**Fonction de densité de probabilité.** Fonction utilisée pour calculer les probabilités d'une variable aléatoire continue. L'aire sous le graphique d'une fonction de densité de probabilité comprise dans un intervalle donné représente la probabilité.

**Loi uniforme.** Distribution de probabilité continue pour laquelle la probabilité que la variable aléatoire prenne une valeur dans un intervalle est la même pour chaque intervalle de même longueur.

**Loi normale.** Distribution de probabilité continue. Sa fonction de densité est en forme de cloche et est déterminée par la moyenne $\mu$ et l'écart type $\sigma$.

**Loi normale centrée réduite.** Distribution normale de moyenne nulle et d'écart type égal à 1.

**Facteur de correction de continuité.** Valeur de 0,5 ajoutée ou soustraite à la valeur de *X* lorsque la loi normale est utilisée pour estimer la loi binomiale discrète.

**Loi exponentielle.** Distribution de probabilité continue utile pour calculer les probabilités relatives au temps nécessaire pour achever une tâche.

## FORMULES CLÉ

### Fonction de densité de probabilité uniforme

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } a \le x \le b \\ 0 & \text{sinon} \end{cases} \tag{6.1}$$

### Fonction de densité de probabilité normale

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \tag{6.2}$$

### Conversion en distribution normale centrée réduite

$$z = \frac{x-\mu}{\sigma} \tag{6.3}$$

### Fonction de densité de probabilité exponentielle

$$f(x) = \frac{1}{\mu} e^{-x/\mu} \quad \text{pour } x \geq 0,\ \mu \geq 0 \tag{6.4}$$

### Loi exponentielle : Probabilités cumulées

$$P(x \leq x_0) = 1 - e^{-x_0/\mu} \tag{6.5}$$

## EXERCICES SUPPLÉMENTAIRES

**39.** Un cadre commercial est muté de Chicago à Atlanta et doit vendre sa maison de Chicago rapidement. Son employeur a offert d'acheter la maison 210 000 dollars mais son offre expire à la fin de la semaine. Le cadre n'a pas, pour le moment, de meilleure offre mais a les moyens de laisser la maison en vente un mois de plus. Après avoir consulté son agent immobilier, le cadre pense que le prix qu'il pourra obtenir en laissant sa maison en vente un mois de plus, est uniformément distribué entre 200 000 et 225 000 dollars.

- **a)** S'il laisse sa maison en vente un mois de plus, quelle est l'expression mathématique de la fonction de densité du prix de vente ?
- **b)** S'il laisse sa maison en vente un mois de plus, quelle est la probabilité qu'il obtienne au moins 215 000 dollars pour la maison ?
- **c)** S'il laisse sa maison en vente un mois de plus, quelle est la probabilité qu'il obtienne moins de 210 000 dollars ?
- **d)** Le cadre doit-il laisser sa maison en vente un mois de plus ? Pourquoi ?

**40.** La NCAA estime que le montant annuel d'une bourse d'études sportives dans une université d'État s'élève à 19 000 dollars (*The Wall Street Journal*, 12 mars 2012). Supposez que ce montant suive une loi normale avec un écart type de 2 100 dollars.

- **a)** Considérez les 10 % des bourses les plus faibles. Quel est leur montant moyen ?
- **b)** Quel est le pourcentage de bourses d'études sportives dont le montant est supérieur ou égal à 22 000 dollars ?
- **c)** Considérez les 3 % des bourses les plus élevées. Quel est leur montant moyen ?

**41.** Motorola a utilisé la loi normale pour déterminer la probabilité de défauts et le nombre moyen de défauts dans un processus de production. Supposez qu'un processus de production soit conçu pour produire des pièces dont le poids moyen est égal à 10 onces. Calculer la probabilité d'un défaut et le nombre moyen de défauts dans un lot de 1 000 pièces, dans les situations suivantes :

**a)** L'écart type du processus est égal à 0,15 et le contrôle du processus est fixé à plus ou moins un écart type. Les pièces dont le poids est inférieur à 9,85 ou supérieur à 10,15 onces, sont considérées comme défectueuses.

**b)** Grâce à des améliorations du processus, l'écart type est réduit à 0,05. Supposez que le contrôle du processus reste le même : les pièces dont le poids est inférieur à 9,85 ou supérieur à 10,15 onces, sont considérées comme défectueuses.

**c)** Quel est l'avantage de réduire la variabilité du processus et de fixer les limites de contrôle du processus à un plus grand nombre d'écarts type par rapport à la moyenne ?

**42.** Début 2012, les difficultés économiques ont pesé sur le système social français. Un indicateur de ces difficultés fut le nombre croissant d'individus qui ont eu recours aux services de prêteurs sur gage : il est passé à 658 par jour (*Bloomberg Businessweek*, 5-11 mars 2012). Supposez que le nombre de personnes qui ont eu recours aux services d'un prêteur sur gage par jour en 2012 suive une loi normale de moyenne égale à 658.

**a)** Supposez que vous appreniez qu'au cours de 3 % de ces jours, au plus 610 individus ont eu recours aux services d'un prêteur sur gage. Quel est l'écart type du nombre d'individus ayant eu recours aux services d'un prêteur sur gage ?

**b)** Un jour donné, quelle est la probabilité qu'entre 600 et 700 individus aient eu recours aux services d'un prêteur sur gage ?

**c)** Au cours des 3 % des jours les plus chargés, combien d'individus ont eu recours aux services d'un prêteur sur gage ?

**43.** Le port de Louisiane du Sud, situé à 54 miles de la Nouvelle Orléans et de Baton Rouge sur le fleuve Mississipi, est le plus grand port de fret de marchandises du monde. Le corps des ingénieurs de l'armée américaine rapporte que le port traite en moyenne 4,5 millions de tonnes de marchandises par semaine (*USA Today*, 25 septembre 2012). Supposez que le nombre de tonnes de marchandises traitées par semaine suive une loi normale avec un écart type de 0,82 million de tonnes.

**a)** Quelle est la probabilité que le port traite moins de 5 millions de tonnes de marchandises en une semaine ?

**b)** Quelle est la probabilité que le port traite au moins 3 millions de tonnes de marchandises en une semaine ?

**c)** Quelle est la probabilité que le port traite entre 3 et 4 millions de tonnes de marchandises en une semaine ?

**d)** Supposez que 85 % du temps, le port est en mesure de traiter le volume de marchandises hebdomadaire sans allonger ses heures d'ouverture. Quel est le nombre de tonnes de marchandises hebdomadaire qui nécessiterait une augmentation de la durée d'ouverture du port ?

**44.** La société Ward Doering Auto Sales étudie l'opportunité d'offrir un contrat de service spécial qui couvrirait tous les coûts d'entretien des voitures en leasing. De par son expérience, le responsable estime que les coûts annuels sont normalement distribués, avec une moyenne de 150 dollars et un écart type de 25 dollars.

**a)** Si la société fixe le prix du contrat de service à 200 dollars par an, quelle est la probabilité que les coûts d'entretien du véhicule d'un client excèdent le prix du contrat fixé ?

**b)** Quel est le profit moyen de Ward par contrat ?

**45.** Le minibar d'une chambre d'hôtel révèle généralement si l'hôtel est un hôtel haut de gamme ou non. Les études PKF Hospitality ont indiqué que les consommations des minibars fournissaient un revenu annuel moyen de 368 dollars par chambre (*USA Today*, 9 février 2012). Considérez un hôtel haut de gamme de San Antonio au Texas qui a au total 330 chambres, chacune disposant d'un minibar. Supposez que le revenu mensuel total du service minibar de l'hôtel suive une loi normale avec un écart type de 2 200 dollars.

**a)** En utilisant le revenu annuel moyen de 368 dollars par minibar, quel est le revenu mensuel total moyen pour le service minibar de cet hôtel ?

**b)** Quelle est la probabilité que le service minibar génère un revenu mensuel supérieur à 12 000 dollars à cet hôtel ?

**c)** Quelle est la probabilité que le service minibar génère un revenu mensuel inférieur à 7 500 dollars à cet hôtel ?

**d)** L'hôtel étudie la possibilité de proposer des boissons plus haut de gamme pour rendre le minibar plus attractif. Les nouvelles offres du minibar sont supposées augmenter le revenu annuel moyen jusqu'à 420 dollars par minibar. Supposez que le revenu mensuel total du nouveau service de minibar de l'hôtel suive une loi normale avec un écart type de 2 500 dollars. Répondre aux questions (b) et (c) pour le service amélioré de minibar. Soutenez-vous la stratégie de montée en gamme du service de minibar de l'hôtel ? Pourquoi ?

**46.** Supposez que les notes obtenues au test d'admission d'un collège soient normalement distribuées, avec une moyenne de 450 et un écart type de 100.

**a)** Quel est le pourcentage de personnes qui ont une note comprise entre 400 et 500 ?

**b)** Supposez que quelqu'un ait une note de 630. Quel est le pourcentage de personnes qui ont une meilleure note ? Une moins bonne note ?

**c)** Si une université particulière n'admet pas les personnes qui ont une note inférieure à 480, quel est le pourcentage de personnes qui, ayant fait ce test, pourront être admises à l'université ?

**47.** Selon Salary Wizard, le salaire de base moyen d'un responsable commercial de Houston au Texas s'élève à 88 592 dollars et celui d'un responsable commercial de Los Angeles en Californie à 97 417 dollars (site Internet de Salary Wizard, 27 février 2008). Supposez que les salaires soient normalement distribués, que l'écart type pour les responsables commerciaux de Houston soit égal à 19 900 dollars et que l'écart type pour les responsables commerciaux de Los Angeles soit égal à 21 800 dollars.

**a)** Quelle est la probabilité qu'un responsable commercial de Houston ait un salaire de base supérieur à 100 000 dollars ?

**b)** Quelle est la probabilité qu'un responsable commercial de Los Angeles ait un salaire de base supérieur à 100 000 dollars ?

**c)** Quelle est la probabilité qu'un responsable commercial de Los Angeles ait un salaire de base inférieur à 75 000 dollars ?

**d)** Combien un responsable commercial de Los Angeles devrait-il toucher pour avoir un salaire supérieur à celui que touchent 99 % des responsables commerciaux de Houston ?

**48.** Une machine remplit des récipients d'un produit particulier. L'écart type des poids de remplissage est, d'après les données historiques, égal à 0,6 once. Si seulement 2 % des récipients contiennent moins de 18 onces, quel est le poids moyen de remplissage de la machine ? C'est-à-dire, quelle est la valeur de $\mu$ ? Supposez que les poids de remplissage suivent une loi normale.

**49.** Considérez un questionnaire à choix multiples de 50 questions. Quatre réponses sont possibles à chaque question. Supposez qu'un étudiant qui a fait ses devoirs à la maison et suivi les cours, ait une probabilité de 0,75 de répondre correctement à une question.

**a)** Un étudiant doit répondre correctement à au moins 43 questions pour obtenir la note A. Quel est le pourcentage d'étudiants qui ayant suivi les cours et fait leurs devoirs, obtiendront un A à ce questionnaire à choix multiples ?

**b)** Un étudiant qui répond correctement à un nombre de questions compris entre 35 et 39, obtiendra un C. Quel est le pourcentage d'étudiants qui ayant suivi les cours et fait leurs devoirs, obtiendront un C à cet examen?

**c)** Un étudiant doit répondre correctement à au moins 30 questions pour réussir l'examen. Quel est le pourcentage d'étudiants qui ayant suivi les cours et fait leurs devoirs, réussiront l'examen ?

**d)** Supposez qu'un étudiant n'a ni suivi les cours, ni fait ses devoirs. De plus, supposez que l'étudiant devine simplement la réponse de chaque question. Quelle est la probabilité que cet étudiant réponde correctement à au moins 30 questions et réussisse l'examen ?

**50.** Un joueur de blackjack, dans un casino de Las Vegas, a appris que la maison lui fournirait une chambre gratuitement s'il jouait pendant quatre heures avec une mise moyenne de 50 dollars. Sa stratégie de jeu assure une probabilité égale à 0,49 de gagner une partie et le joueur sait qu'environ 60 parties sont jouées en une heure. Supposez qu'il joue pendant quatre heures avec une mise de 50 dollars par partie.

**a)** Quel est le gain espéré du joueur ?

**b)** Quelle est la probabilité que le joueur perde au moins 1 000 dollars ?

**c)** Quelle est la probabilité que le joueur gagne ?

**d)** Supposez que le joueur débute avec 1 500 dollars. Quelle est la probabilité qu'il fasse banqueroute ?

**51.** L'association de contrôle et d'audit des systèmes d'information a enquêté auprès d'employés de bureau pour déterminer quel usage ils feraient de leur ordinateur professionnel pour effectuer leurs courses de Noël (*USA Today*, 11 novembre 2009). Supposez que le nombre d'heures qu'un employé pense passer à effectuer des achats de Noël sur son ordinateur professionnel suive une loi exponentielle.

**a)** L'étude a rapporté qu'il y a une probabilité de 0,53 qu'un employé utilise son ordinateur professionnel pour effectuer des achats de Noël au plus durant 5 heures.

Est-ce que le temps moyen passé à effectuer des achats de Noël sur l'ordinateur professionnel est plus proche de 5,8, 6,2, 6,6 ou 7 heures ?

**b)** En utilisant le temps moyen déterminé à la question (a), quelle est la probabilité qu'un employé passe plus de 10 heures à effectuer des achats de Noël sur son ordinateur professionnel ?

**c)** Quelle est la probabilité qu'un employé utilise son ordinateur professionnel entre 4 et 8 heures pour effectuer des achats de Noël ?

**52.** Le site web de Bed and Breakfast Inns d'Amérique du Nord reçoit approximativement 7 visites par minute. Supposez que le nombre de visiteurs sur le site web, par minute, suive une loi de Poisson.

**a)** Quel est le temps moyen écoulé entre deux visites sur le site web ?

**b)** Écrire la fonction de densité de probabilité exponentielle pour le temps écoulé entre deux visites sur le site web.

**c)** Quelle est la probabilité que personne ne se connecte au site web pendant une période d'une minute ?

**d)** Quelle est la probabilité que personne ne se connecte au site web pendant une période de 12 secondes ?

**53.** L'enquête sur les communautés américaines a montré que les habitants de la ville de New York ont les temps de trajet domicile-travail les plus longs, comparativement aux autres villes américaines (site Internet du bureau du recensement américain, août 2008). Selon les dernières statistiques disponibles, le temps moyen de trajet domicile-travail des résidents de New York est de 38,3 minutes.

**a)** Supposez que la loi exponentielle soit appropriée et donnez la fonction de densité de probabilité du temps de trajet domicile-travail d'un New-Yorkais.

**b)** Quelle est la probabilité que le temps de trajet d'un New-Yorkais soit compris entre 20 et 40 minutes ?

**c)** Quelle est la probabilité que le temps de trajet d'un New-Yorkais soit supérieur à une heure ?

**54.** Le temps (en minutes) entre les appels téléphoniques dans une agence d'assurance suit la loi exponentielle suivante :

$$f(x) = 0{,}50e^{-0{,}50x} \quad \text{pour } x \geq 0$$

**a)** Quel est le temps moyen entre les appels téléphoniques ?

**b)** Quelle est la probabilité d'avoir au plus 30 secondes de répit entre deux appels téléphoniques ?

**c)** Quelle est la probabilité d'avoir au plus une minute de répit entre deux appels téléphoniques ?

**d)** Quelle est la probabilité de ne pas avoir d'appel téléphonique pendant au moins 5 minutes ?

## PROBLÈME *Specialty Toys*

La société Specialty Toys vend de nombreux jouets pour enfants. Les dirigeants savent que la période avant les fêtes de fin d'année est la plus propice à l'introduction de nouveaux jouets, parce que beaucoup de familles mettent à profit ce moment pour rechercher de nouvelles idées de cadeaux de Noël. Lorsque la société Specialty découvre un nouveau jouet avec un fort potentiel de vente, elle choisit de le mettre sur le marché en octobre.

Pour avoir les jouets dans ses rayons en octobre, la société passe commande à ses fabricants en juin ou juillet chaque année. La demande de jouets pour enfants peut être très volatile. Si le nouveau jouet connaît un certain engouement, un sentiment de rareté sur le marché accroît souvent la demande et d'importants profits peuvent être réalisés. Cependant, l'introduction de nouveaux jouets peut également se solder par un échec, laissant la société avec des stocks importants sur les bras, qui devront être vendus à prix réduit. La plus importante décision à laquelle doit faire face la société est de définir le nombre d'unités qui seront produites pour satisfaire la demande potentielle. Si trop peu de jouets sont produits, la société perd des ventes ; si trop de jouets sont produits, les profits seront réduits à cause de la baisse de prix nécessaire pour écouler les stocks.

Pour la saison à venir, Specialty envisage de mettre sur le marché un nouveau produit appelé Weather Teddy. Cette nouvelle version d'un ours parlant est fabriquée par une société à Taïwan. Lorsqu'un enfant presse la main de la peluche, l'ours se met à parler. Un baromètre, placé à l'intérieur de la peluche, sélectionne l'une des cinq prévisions de temps possibles. Les prévisions vont de « Ce sera une très belle journée. Profitez-en ! » à « Je crains qu'il ne pleuve aujourd'hui. N'oubliez pas votre parapluie ! ». Les tests ont prouvé que, sans être parfaites, les prévisions étaient plutôt bonnes. Plusieurs responsables de la société ont déclaré que les prévisions de Weather Teddy étaient aussi bonnes que celles des prévisionnistes des chaînes de télévision locales.

Comme pour tout produit, Specialty doit décider combien d'unités fabriquer. Différentes suggestions ont été faites par les membres de l'équipe dirigeante : 15 000, 18 000, 24 000 ou 28 000 unités. L'écart entre ces propositions souligne les divergences d'opinion quant au potentiel de vente de ce produit. Les dirigeants font appel à vous pour analyser les probabilités que des unités restent invendues dans les différents cas de figure (15 000, 18 000, 24 000 ou 28 000 unités commandées), pour estimer le profit potentiel et pour faire une recommandation quant à la quantité à commander. Specialty souhaite vendre Weather Teddy 24 dollars, sachant que le coût de production unitaire est de 16 dollars. Si un stock d'invendus reste après les fêtes, Specialty vendra chaque unité 5 dollars. Après avoir revu l'historique des ventes de produits similaires, le prévisionniste en chef des ventes de Specialty prévoit une demande de 20 000 unités, avec une probabilité de 0,95 que la demande soit comprise entre 10 000 et 30 000 unités.

### *Rapport*

Préparez un rapport managérial qui répond aux questions suivantes et recommandez quelle quantité de Weather Teddy commander.

1. Utiliser les prévisions de ventes pour décrire une distribution de probabilité normale qui peut être utilisée pour estimer la distribution de la demande. Représenter la distribution et indiquer sa moyenne et son écart type.
2. Calculer la probabilité qu'il y ait des invendus pour chacune des quantités de commande suggérées par l'équipe des dirigeants.
3. Calculer le profit attendu pour chacune des quantités de commande suggérées par l'équipe des dirigeants, sous trois scénarios alternatifs : le pire cas avec 10 000 unités vendues ; le cas le plus vraisemblable avec 20 000 unités vendues ; le cas le plus optimiste avec 30 000 unités vendues.
4. L'un des dirigeants de Specialty pense que la quantité commandée a 70 % de chances de satisfaire la demande et seulement 30 % de chances d'entraîner la constitution de stocks d'invendus. Dans ce contexte, quelle quantité devrait être commandée ? Quel est le profit espéré sous les trois scénarios de vente ?
5. Fournissez votre propre recommandation quant à la quantité à commander et donnez le profit espéré pour chacun des trois scénarios. Justifiez votre recommandation.

## ANNEXE 6.1 LOIS DE PROBABILITÉ CONTINUES AVEC MINITAB

Étudions la procédure de calcul des probabilités continues avec Minitab, en nous référant au problème de la société Grear Tire, dans lequel le kilométrage des pneus est décrit par une loi normale de moyenne $\mu = 36\,500$ et d'écart type $\sigma = 5\,000$. Une des questions posées était : quelle est la probabilité que le kilométrage d'un pneu dépasse 40 000 km ?

Pour des lois continues, Minitab fournit une probabilité cumulée. En d'autres termes, Minitab fournit la probabilité qu'une variable aléatoire prenne une valeur inférieure ou égale à une certaine valeur prédéterminée. Dans le cadre du problème de la société Grear Tire, Minitab peut être utilisé pour déterminer la probabilité cumulée que le kilométrage du pneu soit inférieur ou égal à 40 000 km. Après avoir obtenu la probabilité cumulée de Minitab, on doit la soustraire à 1 pour trouver la probabilité que le kilométrage du pneu excède 40 000 km.

Avant d'utiliser Minitab pour calculer une probabilité, on doit entrer la valeur prédéterminée dans une colonne de la feuille de calcul. Pour répondre à la question du kilométrage des pneus Grear, on a entré la valeur prédéterminée de 40 000 dans la colonne C1 de la feuille de calcul Minitab. Les étapes de l'utilisation de Minitab pour calculer la probabilité cumulée d'une variable aléatoire normale prenant une valeur inférieure ou égale à 40 000, sont décrites ci-dessous.

**Étape 1.** Sélectionner le menu **Calc**
**Étape 2.** Sélectionner le menu **Probability Distributions**
**Étape 3.** Sélectionner l'option **Normal**

**Étape 4.** Lorsque la boîte de dialogue apparaît :
Sélectionner **Cumulative probability**
Entrer 36 500 dans la boîte **Mean**
Entrer 5 000 dans la boîte **Standard deviation**
Entrer C1 dans la boîte **Input column** (la cellule contient la valeur 40 000)
Cliquer sur **OK**

Minitab fournira une probabilité égale à 0,7580. Puisque nous nous intéressons à la probabilité que le kilométrage du pneu dépasse 40 000 km, la probabilité souhaitée est égale à 0,2420 ($1-0,7580=0,2420$).

Une seconde question posée dans le cadre du problème de la société Grear Tire était : quelle est la garantie de kilométrage que Grear devrait fixer pour s'assurer que la garantie ne s'applique pas à plus de 10 % des pneus ? Ici la probabilité est donnée et l'on veut trouver la valeur de la variable aléatoire qui y correspond. Minitab utilise une fonction de calcul inverse pour trouver la valeur de la variable aléatoire associée à une probabilité cumulée donnée. D'abord, nous devons entrer la probabilité cumulée dans une colonne de la feuille de calcul de Minitab (disons C1). Dans cet exemple, la probabilité cumulée est égale à 0,10. Ensuite, les trois premières étapes de la procédure Minitab sont les mêmes que celles décrites ci-dessus. À l'étape 4, on sélectionne **Inverse cumulative probability** au lieu de **Cumulative probability** et on exécute le reste de la procédure. Minitab fournit alors le chiffre de 30 092 km.

Minitab est capable de calculer des probabilités pour d'autres lois continues, dont la loi exponentielle. Pour calculer des probabilités exponentielles, il suffit de suivre la procédure décrite précédemment pour la loi normale et de sélectionner l'option **Exponential** à l'étape 3. L'étape 4 est la même, mis à part le fait qu'il est inutile de rentrer la valeur de l'écart type. Les résultats des probabilités cumulées et des probabilités cumulées inversées sont identiques à ceux décrits pour la loi normale.

## ANNEXE 6.2 LOIS DE PROBABILITÉ CONTINUES AVEC EXCEL

Excel a la capacité de calculer des probabilités pour plusieurs lois de probabilité continues, dont la loi normale. Dans cette annexe, nous décrirons comment utiliser Excel pour calculer les probabilités d'une distribution normale. Les procédures pour les autres lois continues sont similaires à celle que nous décrirons pour la loi normale.

Reprenons le problème de la société Grear Tire, dans lequel le kilométrage est décrit par une loi normale, de moyenne $\mu = 36\,500$ et d'écart type $\sigma = 5\,000$. Supposons que nous nous intéressions à la probabilité que le kilométrage d'un pneu dépasse 40 000 km.

La fonction NORM.DIST d'Excel fournit les probabilités cumulées d'une distribution normale. La forme générale de la fonction est NORM.DIST ($x$, $\mu$, $\sigma$, *cumulative*). Le qualificatif TRUE est choisi pour définir le quatrième élément (*cumulative*) si on souhaite obtenir la probabilité cumulée. Ainsi, pour calculer la probabilité cumulée que le kilométrage du pneu soit inférieur ou égal à 40 000 km, on entre la formule suivante dans une cellule d'une feuille de calcul Excel :

= NORM.DIST (40000, 36500, 5000, TRUE)

À ce moment-là, 0,7580 apparaîtra dans la cellule dans laquelle la formule a été entrée, indiquant que la probabilité que le kilométrage soit inférieur ou égal à 40 000 km, est égale à 0,7580. Par conséquent, la probabilité que le kilométrage du pneu excède 40 000 km est égale à 0,2420 ($1-0,7580=0,2420$).

La fonction NORM.INV d'Excel permet de trouver la valeur de la variable aléatoire correspondant à une probabilité cumulée donnée. Par exemple, supposons que nous cherchions la garantie de kilométrage que Grear devrait fixer pour s'assurer qu'elle ne s'applique pas à plus de 10 % des pneus. Pour cela, nous devons entrer la formule suivante dans une feuille de calcul Excel :

= NORM.INV (0.1, 36500, 5000)

À ce moment-là, 30 092 apparaîtra dans la cellule dans laquelle la formule a été entrée, indiquant que la probabilité que le pneu effectue au plus 30 092 km est égale à 0,10.

La fonction Excel pour calculer des probabilités exponentielles est EXPON.DIST. Cette fonction nécessite d'entrer trois facteurs : $x$, la valeur de la variable ; lambda égal à $1/\mu$ et TRUE si vous souhaitez calculer une probabilité cumulée. Par exemple, considérez une loi exponentielle de moyenne $\mu=15$. La probabilité qu'une variable exponentielle soit inférieure ou égale à 6 peut être calculée en utilisant la formule Excel suivante :

= EXPON.DIST (6, 1/15, TRUE).

Si vous avez besoin d'aide pour déterminer les bons arguments, vous pouvez utiliser la fonction Insert (cf. annexe E).