Probabilités et Statistiques :
Exemples et Applications

Table des matières

2 Lois de probabilité discrètes

L'ensemble du texte est un recueil de texte de :

- **Aide-mémoire Statistique et probabilités pour l'ingénieur, Dunod, 2014.**

Les lois discrètes sont utilisées pour modéliser les résultats des jeux de hasard, les sondages d'opinion, les phénomènes biologiques, les processus aléatoires (files d'attente, évolution de l'état de matériels) ... Les plus utilisées sont la loi uniforme, la loi binomiale et les lois dérivées, la loi de Poisson.

<u>Exemple :</u>

Soit X la variable aléatoire prenant trois valeurs 0, 1, 2 avec les probabilités :

$$\Pr(X = 0) = 1/2 \qquad \Pr(X = 1) = 1/3 \qquad \Pr(X = 2) = 1/6$$

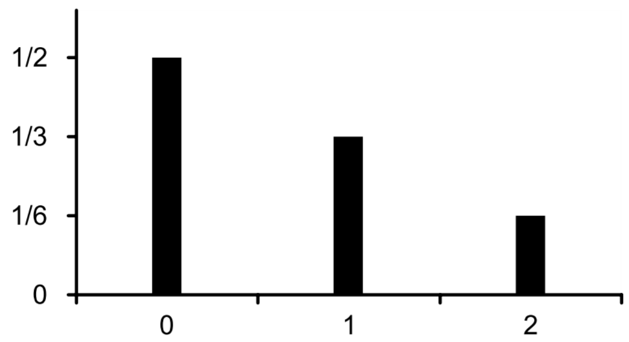
$(1/2 + 1/3 + 1/6 = 1)$

Espérance mathématique :

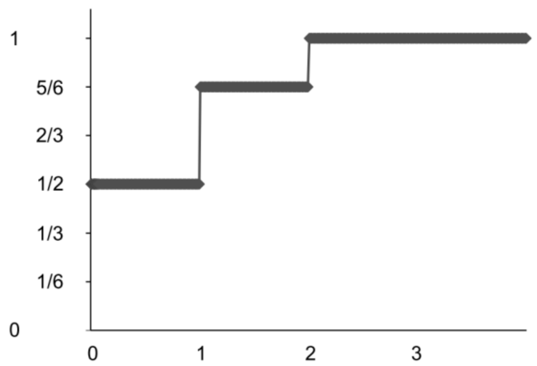
$$E(X) = 0 \times 1/2 + 1 \times 1/3 + 2 \times 1/6 = 2/3$$

Variance :

$$\mathrm{Var}(X) = (\sigma_x)^2 = E(X^2) - [E(X)]^2 = 1 \times 1/3 + 4 \times 1/6 - (2/3)^2 = 5/9$$



Histogramme de la loi donnée dans l'exemple



Fonction de répartition de la loi donnée dans l'exemple

2.1 Loi uniforme

2.1.1 Domaines d'application

La loi de probabilité uniforme intervient dans de nombreux domaines comme les jeux de pile ou face ou les jeux de dés (avec une pièce ou un dé parfaitement équilibré(e)), les jeux de cartes, les loteries, les sondages...

2.2 Loi binomiale ou loi des tirages avec remise

2.2.1 Domaines d'application

- La loi binomiale décrit des phénomènes ne pouvant prendre que deux états s'excluant mutuellement, succès ou échec dans un jeu, bonne pièce ou pièce défectueuse dans une fabrication, lot acceptable ou lot refusé, défaillance ou fonctionnement d'un matériel...
- Elle est utilisée dans le domaine technique pour déterminer la probabilité de défaillance à la sollicitation de matériels, en contrôle qualité, mais elle ne peut s'appliquer rigoureusement que si les expériences sont non exhaustives, c'est la loi du tirage avec remise.
- Les événements considérés doivent être indépendants et la probabilité de réalisation d'un événement doit être constante.

2.2.2 Exemples

Exemple

On veut réaliser une étude clinique sur des malades se présentant à une consultation hospitalière. Pour cette étude, seuls les malades répondant à un ensemble de critères C sont retenus. Des statistiques antérieures ont montré que 20 % des consultants présentent les critères C.

10 malades viennent consulter le premier jour.

Soit X la variable aléatoire « nombre de malades retenus » c'est-à-dire répondant à l'ensemble des critères C. La variable X suit *la loi binomiale* $B(10\,;0{,}20)$.

La probabilité qu'aucun malade ne soit recruté ce jour est égale à :

$$\Pr(X = 0) = \mathrm{C}_{10}^{0}\,(0{,}20)^0\,(0{,}80)^{10} = 0{,}107$$

La probabilité pour qu'il y ait au moins un malade recruté est égale à :

$$\Pr(X \geqslant 1) = 1 - \Pr(X = 0) = 1 - 0{,}107 = 0{,}803$$

EXAMPLE | Digital Channel

The chance that a bit transmitted through a digital transmission channel is received in error is 0.1. Also, assume that the transmission trials are independent. Let X = the number of bits in error in the next four bits transmitted. Determine $P(X = 2)$.

Let the letter E denote a bit in error, and let the letter O denote that the bit is okay, that is, received without error. We can represent the outcomes of this experiment as a list of four letters that indicate the bits that are in error and those that are okay. For example, the outcome $OEOE$ indicates that the second and fourth bits are in error and the other two bits are okay. The corresponding values for x are

Outcome	x	Outcome	x
$OOOO$	0	$EOOO$	1
$OOOE$	1	$EOOE$	2
$OOEO$	1	$EOEO$	2
$OOEE$	2	$EOEE$	3
$OEOO$	1	$EEOO$	2
$OEOE$	2	$EEOE$	3
$OEEO$	2	$EEEO$	3
$OEEE$	3	$EEEE$	4

The event that $X = 2$ consists of the six outcomes:

$$\{EEOO, EOEO, EOOE, OEEO, OEOE, OOEE\}$$

Using the assumption that the trials are independent, the probability of $\{EEOO\}$ is

$$P(EEOO) = P(E)P(E)P(O)P(O) = (0.1)^2(0.9)^2 = 0.0081$$

Also, any one of the six mutually exclusive outcomes for which $X = 2$ has the same probability of occurring. Therefore,

$$P(X = 2) = 6(0.0081) = 0.0486$$

In general, $P(X = x)$ = (number of outcomes that result in x errors) $\times (0.1)^x(0.9)^{4-x}$.

To complete a general probability formula, only an expression for the number of outcomes that contain x errors is needed. An outcome that contains x errors can be constructed by partitioning the four trials (letters) in the outcome into two groups. One group is of size x and contains the errors, and the other group is of size $n - x$ and consists of the trials that are okay. The number of ways of partitioning four trials into two groups, one of which is of size x, is $\binom{4}{x} = \frac{4!}{x!(4-x)!}$. Therefore, in this example,

$$P(X = x) = \binom{4}{x}(0.1)^x(0.9)^{4-x}$$

2.3 Loi géométrique

2.3.1 Domaines d'application

La loi géométrique est la loi de la variable Y « loi du nombre d'essais nécessaires » pour qu'un événement de probabilité p apparaisse pour la première fois, les hypothèses étant les mêmes que pour la loi binomiale, en particulier, la probabilité p reste constante au cours des essais

2.3.2 Exemples

Exemple

Un certain matériel a une probabilité $p = 0{,}02$ constante de défaillance à chaque mise en service. On procède à l'expérience suivante, l'appareil est mis en marche, arrêté, remis en marche, arrêté, jusqu'à ce qu'il tombe en panne. Le nombre d'essais nécessaires pour obtenir la panne est une variable aléatoire suivant la loi géométrique de paramètre p. La probabilité que ce matériel tombe en panne (pour la première fois) au dixième essai est égale à :

$$\Pr(Y = 10) = (0{,}02)(1 - 0{,}02)^9 = 0{,}0167$$

2.4 Loi multinomiale

2.4.1 Définitions

La loi multinomiale est une généralisation de la loi binomiale.

Une population P est composée d'individus appartenant à k types différents, dans des proportions $p_1, p_2 \ldots p_k$ telles que $\sum_{i=1}^{k} p_i = 1$.

On tire un échantillon de n individus, de façon équiprobable et indépendante et on s'intéresse à la composition de l'échantillon.

Soit X_i la variable aléatoire représentant le nombre d'individus de type i dans l'échantillon. Par définition, le vecteur $\underline{X} = (X_1, \ldots, X_k)$ est un vecteur aléatoire suivant une *loi multinomiale de paramètres* $(n\,; p_1, \ldots, p_k)$, notée $M(n\,; p_1, \ldots, p_k)$.

$$\begin{aligned}
\Pr[X = (x_1, \ldots, x_k)] &= \mathrm{C}_n^{x_1}\, \mathrm{C}_{n-x_1}^{x_2} \ldots \mathrm{C}_{n-(x_1+\cdots+x_{k-1})}^{x_k}\, p_1^{x_1} \ldots p_k^{x_k} \\
&= \frac{n!}{x_1! \ldots x_k!}\, p_1^{x_1} \ldots p_k^{x_k}
\end{aligned}$$

$$\sum_{i=1}^{k} p_i = 1 \qquad \sum_{i=1}^{k} x_i = n$$

Le coefficient $\frac{n!}{x_1! \ldots x_k!}$ est le nombre de partitions d'un échantillon de taille n en sous-populations d'effectifs x_i (voir annexe 1).

Le nombre d'individus du type i dans l'échantillon suit donc la loi binomiale $B(n\,; p_i)$. D'où :

$$E(X_i) = np_i \qquad \mathrm{Var}(X_i) = np_i(1 - p_i)$$

Les variables X_i et X_k ne sont pas indépendantes.

Le couple (X_i, X_k) suit une loi multinomiale de dimension 3. En effet, un élément tiré est :

- soit du type i (probabilité p_i),
- soit du type k (probabilité p_k),
- soit de n'importe quel autre type (probabilité $1 - p_i - p_k$).

En partant de ces propriétés, on démontre que :

$$E\,(X_i X_k) = n\,(n-1)\;p_i\,p_k$$

$$\operatorname{Cov}(X_i,\ X_k) = E\,(X_i X_k) - E\,(X_i)\;E\,(X_k) = -n\,p_i\,p_k$$

Les variables X_i et X_k ne peuvent donc pas être indépendantes.

2.4.2 Domaines d'application

Soit X une variable aléatoire continue de densité $f(x)$. On suppose que l'espace D_X des valeurs prises par cette variable est partagé en k classes distinctes C_i d'extrémités e_{i-1} et e_i, par exemple tranches d'âges, de revenus, d'impôts...

On considère un échantillon $(X_1, \ldots, X_n)$ de n observations de cette variable et on cherche le nombre de points N_i de l'échantillon dans la classe C_i.

2.4.3 Exemples

Exemple

Un produit d'éclairage de l'entreprise M peut présenter des défectuosités regroupées en trois catégories : défectuosité critique, défectuosité majeure, défectuosité mineure.

Un contrôle final est effectué une semaine après la sortie du produit pour vérifier si certaines défectuosités se seraient développées au cours de cette période. Le résultat du contrôle est le suivant :

- 80 % du produit ne présente aucune défectuosité (ensemble E_1),
- 10 % du produit présente des défectuosités mineures (ensemble E_2),
- 6 % du produit présente des défectuosités majeures (ensemble E_3),
- 4 % du produit présente des défectuosités critiques (ensemble E_4).

Un échantillon de taille $n = 20$ est prélevé au hasard dans un grand lot et vérifié selon les critères précédents.

Soit X_i le nombre d'unités appartenant au sous-ensemble E_i dans l'échantillon contrôlé. L'ensemble (X_1, X_2, X_3, X_4) suit une loi multinomiale qui a pour paramètres les pourcentages donnés par le contrôle. D'où la probabilité :

$$\Pr(X_1 = x_1,\ X_2 = x_2,\ X_3 = x_3,\ X_4 = x_4)$$

$$= C_{20}^{x_1} C_{20-x_1}^{x_2} C_{20-(x_1+x_2)}^{x_3} C_{20-(x_1+x_2+x_3)}^{x_4} (0{,}80)^{x_1} (0{,}10)^{x_2} (0{,}06)^{x_3} (0{,}04)^{x_4}$$

$$= \frac{20!}{x_1!\,x_2!\,x_3!\,x_4!} (0{,}80)^{x_1} (0{,}10)^{x_2} (0{,}06)^{x_3} (0{,}04)^{x_4}$$

Espérance mathématique des variables X_i :

$$E(X_1) = 0{,}80 \times 20 = 16 \qquad E(X_2) = 0{,}10 \times 20 = 2$$

$$E(X_3) = 0{,}06 \times 20 = 1{,}2 \qquad E(X_2) = 0{,}04 \times 20 = 0{,}8$$

On peut calculer différentes probabilités :

$\Pr(X_1 = 10,\ X_2 = 6,\ X_3 = 3,\ X_4 = 1) = 0{,}0001439$

$\Pr(X_1 = 20) = 0{,}0115292$

$\Pr(X_1 = 15,\ X_2 = 5) = 0{,}0054549$

2.5 Loi de poisson

2.5.1 Domaines d'application

La loi de Poisson est la loi discrète d'une variable aléatoire représentant un nombre d'événements. Elle est utilisée pour décrire :

- la réalisation d'événements peu probables, dans une succession d'épreuves très nombreuses, au moins 50,
- le nombre d'accidents dans un atelier, le nombre de défauts sur un appareil,

Elle a des applications dans le domaine des files d'attente. Elle est la loi limite de la loi binomiale quand n tend vers l'infini et p tend vers zéro, le produit *np* restant fini.

2.5.2 Exemples

Exemple

Selon les données recueillies depuis plusieurs années, le nombre de pannes hebdomadaires du système informatique d'une entreprise suit une loi de Poisson de paramètre $\lambda = 0{,}05$.

Soit X la variable aléatoire « nombre de pannes hebdomadaires » :

$$\Pr(X = k) = \frac{e^{-0{,}05}\,(0{,}05)^{k}}{k!}$$

La probabilité que le système tombe en panne une fois au cours d'une semaine quelconque ($k = 1$) est égale à 0,04756.

La probabilité qu'il fonctionne sans panne ($k = 0$) est égale à 0,95122.

On considère une année (50 semaines) de fonctionnement de ce système. Le nombre de pannes Y obéit à une loi de Poisson de paramètre $\mu = 0{,}05 \times 50 = 2{,}5$.

$$\Pr(Y = k) = \frac{e^{-2{,}5}\,(2{,}5)^{k}}{k!}$$

La probabilité d'observer 2 pannes au cours de l'année ($k = 2$) est égale à 0,2565 et la probabilité d'en observer 4 est égale à 0,1336.

3 Lois de probabilité continues

L'ensemble du texte est un recueil de texte de :

- **Aide-mémoire Statistique et probabilités pour l'ingénieur, Dunod, 2014.**

3.1 Loi exponentielle

3.1.1 Domaines d'application

– La distribution exponentielle est associée aux processus de Poisson. Un tel processus génère des événements dont les temps d'occurrence sont indépendants et distribués suivant une loi exponentielle (chapitre 9).

– La loi exponentielle est utilisée en *fiabilité* (chapitre 18), le paramètre λ représente le taux de défaillance alors que son inverse $\theta = 1/\lambda$ est le temps moyen de bon fonctionnement MTBF (*Mean Time Between Failure*). Avec le paramètre θ, la densité de probabilité s'écrit :

$$f(x) = \frac{1}{\theta}\ \mathrm{e}^{-\frac{x}{\theta}}$$

et les moments sont égaux à :

$$E(X) = \theta$$

$$\mathrm{Var}(X) = \theta^{2}$$

– La loi exponentielle s'applique bien aux matériels électroniques, c'est-à-dire aux matériels fonctionnant pratiquement sans usure, aux matériels subissant des défaillances brutales ou à des systèmes complexes dont les composants ont des lois de fiabilité différentes. Elle permet de décrire la période de fonctionnement durant laquelle le taux de défaillance est constant ou presque constant.

3.1.2 Exemples

Exemple

On suppose que le temps, en heures, nécessaire pour réparer une machine est une variable aléatoire suivant une loi exponentielle de paramètre $\lambda = 0{,}5$.
La densité de probabilité est $f(t) = 0{,}5\,e^{-0{,}5\,t}$ et la fonction de répartition $F(t) = 1 - e^{-0{,}5\,t}$.
La probabilité pour que le temps de réparation dépasse 2 heures est :

$$\Pr(T > 2) = 1 - \Pr(T < 2) = 1 - F(2) = e^{-1} = 0{,}368$$

Sachant que la réparation a déjà dépassé 9 heures, quelle est la probabilité qu'elle prenne au moins 10 heures ?
La loi exponentielle étant une loi sans « mémoire », on obtient :

$$\Pr(T > 10/T > 9) = \Pr(T > 10 - 9 = 1) = e^{-0{,}5} = 0{,}606$$

3.2 Loi normale

Exemple Utilisation de la table de la loi normale

Soit X une variable suivant la loi normale $N(3\,;2)$, donc de moyenne 3 et d'écart-type 2. On veut calculer les probabilités suivantes : $\Pr(X < 4)$, $\Pr(X < -1)$, $\Pr(X > 1)$ ou les nombres a_i tels que $\Pr(X < a_1) = 0{,}75$, $\Pr(X > a_2) = 0{,}85$.
On utilise la variable centrée réduite U associée à la variable X :

$$U = \frac{X - 3}{2} \quad \text{et} \quad X = 2\,U + 3$$

$$\Pr(X < 4) = \Pr(2\,U + 3 < 4) = \Pr(U < 0{,}50) = 0{,}6915$$

$$\Pr(X < -1) = \Pr(2\,U + 3 < -1) = \Pr(U < -2) = 0{,}0228$$

$$\Pr(X > 1) = \Pr(2\,U + 3 > 1) = \Pr(U > -1) = 0{,}8413$$

$$\Pr(X < a_1) = \Pr(2\,U + 3 < a_1) = \Pr\left(U < \frac{a_1 - 3}{2}\right) = 0{,}75$$

$$\Pr(U < 0{,}6745) = 0{,}75 \quad \text{D'où : } a_1 = 4{,}35$$

$$\Pr(X > a_2) = \Pr(2\,U + 3 > a_2) = \Pr\left(U > \frac{a_2 - 3}{2}\right) = 0{,}85$$

$$\Pr(U < -1{,}0364) = 0{,}15 \quad \Pr(U > -1{,}0364) = 0{,}85$$

$$\text{D'où : } a_2 = -1{,}0364 \times 2 + 3 = 0{,}9272$$

EXAMPLE 4.10 | Normal Distribution Calculations

The following calculations are shown pictorially in Figure 4.13. In practice, a probability is often rounded to one or two significant digits.

1. $P(Z > 1.26) = 1 - P(Z \leq 1.26) = 1 - 0.89616 = 0.10384$.
2. $P(Z < -0.86) = 0.19490$.
3. $P(Z > -1.37) = P(Z < 1.37) = 0.91465$.
4. $P(-1.25 < Z < 0.37)$. This probability can be found from the difference of two areas, $P(Z < 0.37) - P(Z < -1.25)$. Now,

$$P(Z < 0.37) = 0.64431$$

and

$$P(Z < -1.25) = 0.10565$$

Therefore,

$$\begin{aligned} P(-1.25 < Z < 0.37) &= 0.64431 - 0.10565 \\ &= 0.53866 \end{aligned}$$

5. $P(Z \leq -4.6)$ cannot be found exactly from Appendix Table III. However, the last entry in the table can be used to find that $P(Z \leq -3.99) = 0.00003$. Because $P(Z \leq -4.6) < P(Z \leq -3.99)$, $P(Z \leq -4.6)$ is nearly zero.
6. Find the value z such that $P(Z > z) = 0.05$. This probability expression can be written as $P(Z \leq z) = 0.95$. Now Table III is used in reverse. We search through the probabilities to find the value that corresponds to 0.95. The solution is illustrated in Figure 4.13. We do not find 0.95 exactly; the nearest value is 0.95053, corresponding to $z = 1.65$.
7. Find the value of z such that $P(-z < Z < z) = 0.99$. Because of the symmetry of the normal distribution, if the area of the shaded region in Figure 4.13(7) is to equal 0.99, the area in each tail of the distribution must equal 0.005. Therefore, the value for z corresponds to a probability of 0.995 in Table III. The nearest probability in Table III is 0.99506 when $z = 2.58$.

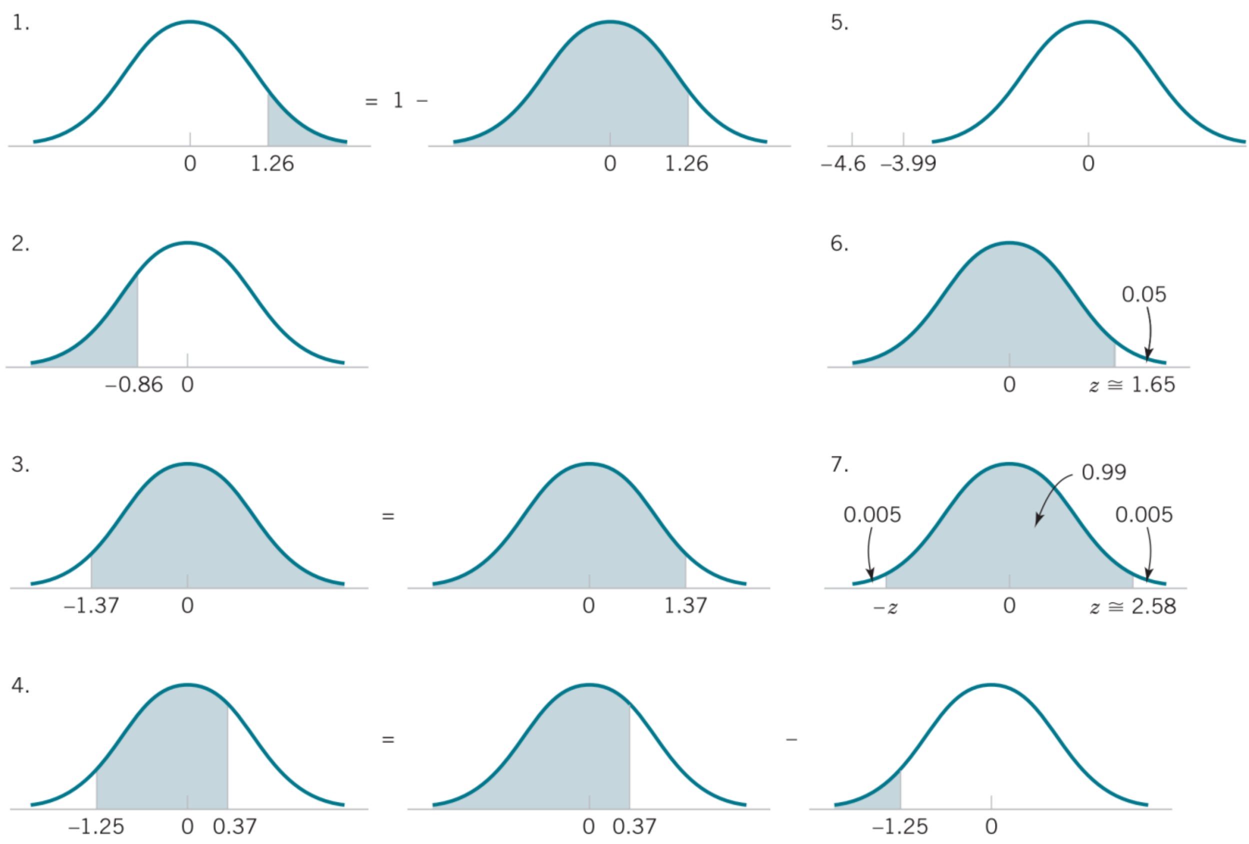


FIGURE 4.13

Graphical displays for standard normal distributions.

EXAMPLE 4.11 | Normally Distributed Current

Suppose that the current measurements in a strip of wire are assumed to follow a normal distribution with a mean of 10 milliamperes and a variance of 4 (milliamperes)^2. What is the probability that a measurement exceeds 13 milliamperes?

Let X denote the current in milliamperes. The requested probability can be represented as $P(X > 13)$. Let $Z = (X - 10)/2$. The relationship between the several values of X and the transformed values of Z are shown in Figure 4.14. We note that $X > 13$ corresponds to $Z > 1.5$. Therefore, from Appendix Table III,

$$P(X > 13) = P(Z > 1.5) = 1 - P(Z \leq 1.5)$$
$$= 1 - 0.93319 = 0.06681$$

Rather than using Figure 4.14, the probability can be found from the inequality $X > 13$. That is,

$$P(X > 13) = P\left(\frac{X - 10}{2} > \frac{13 - 10}{2}\right)$$
$$= P(Z > 1.5) = 0.06681$$

Practical Interpretation: Probabilities for any normal random variable can be computed with a simple transform to a standard normal random variable.

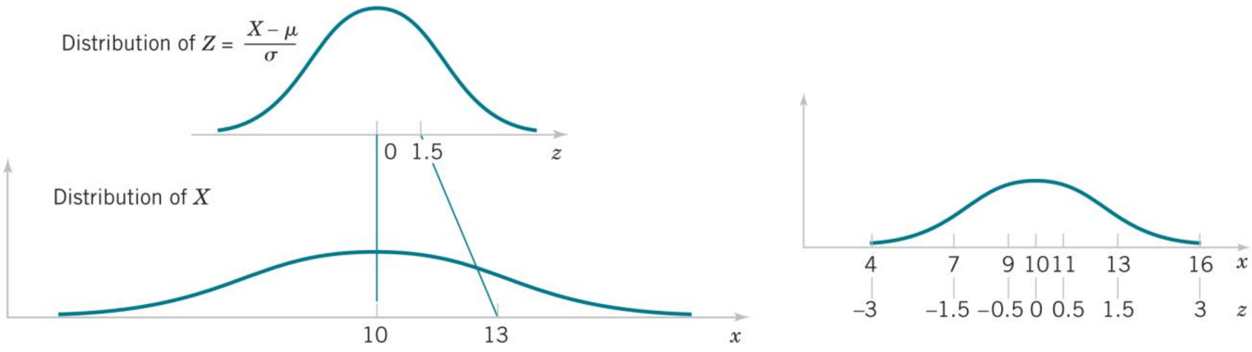


FIGURE 4.14

Standardizing a normal random variable.

EXAMPLE 4.12 | Normally Distributed Current

Continuing Example 4.11, what is the probability that a current measurement is between 9 and 11 milliamperes? From Figure 4.14, or by proceeding algebraically, we have

$$\begin{aligned} P(9 < X < 11) &= P\left(\frac{9-10}{2} < \frac{X-10}{2} < \frac{11-10}{2}\right) \\ &= P(-0.5 < Z < 0.5) \\ &= P(Z < 0.5) - P(Z < -0.5) \\ &= 0.69146 - 0.30854 = 0.38292 \end{aligned}$$

Determine the value for which the probability that a current measurement is less than this value is 0.98. The requested value is shown graphically in Figure 4.15. We need the value of x such that $P(X < x) = 0.98$. By standardizing, this probability expression can be written as

$$\begin{aligned} P(X < x) &= P\left(\frac{X-10}{2} < \frac{x-10}{2}\right) \\ &= P\left(Z < \frac{x-10}{2}\right) = 0.98 \end{aligned}$$

Appendix Table III is used to find the z-value such that $P(Z < z) = 0.98$. The nearest probability from Table III results in

$$P(Z < 2.06) = 0.980301$$

Therefore, $(x - 10)/2 = 2.06$, and the standardizing transformation is used in reverse to solve for x. The result is

$$x = 2(2.06) + 10 = 14.1 \text{ mA}$$

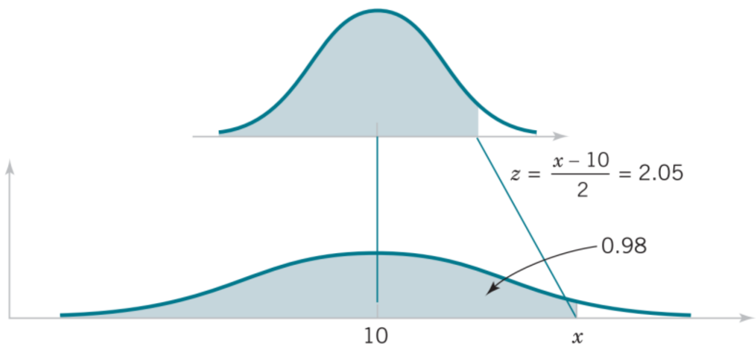


FIGURE 4.15

Determining the value of x to meet a specified probability.

3.3 Limit approximations

3.3.1 Normal Approximation to the Binomial and Poisson Distributions

EXAMPLE 4.13

Assume that in a digital communication channel, the number of bits received in error can be modeled by a binomial random variable, and assume that the probability that a bit is received in error is 1×10^{-5}. If 16 million bits are transmitted, what is the probability that 150 or fewer errors occur?

Let the random variable X denote the number of errors. Then X is a binomial random variable and

$$P(X \le 150) = \sum_{x=0}^{150} \binom{16{,}000{,}000}{x} (10^{-5})^x (1 - 10^{-5})^{16{,}000{,}000-x}$$

Practical Interpretation: Clearly, this probability is difficult to compute. Fortunately, the normal distribution can be used to provide an excellent approximation in this example.

Normal Approximation to the Binomial Distribution

If X is a binomial random variable with parameters n and p,

$$Z = \frac{X - np}{\sqrt{np(1-p)}} \tag{4.12}$$

is approximately a standard normal random variable. To approximate a binomial probability with a normal distribution, a **continuity correction** is applied as follows:

$$P(X \le x) = P(X \le x + 0.5) \approx P\left(Z \le \frac{x + 0.5 - np}{\sqrt{np(1-p)}}\right)$$

and

$$P(x \le X) = P(x - 0.5 \le X) \approx P\left(\frac{x - 0.5 - np}{\sqrt{np(1-p)}} \le Z\right)$$

The approximation is good for $np > 5$ and $n(1-p) > 5$.

EXAMPLE 4.14

The digital communication problem in Example 4.13 is solved as follows:

$$\begin{aligned} P(X \le 150) &= P(X \le 150.5) \\ &= P\left(\frac{X - 160}{\sqrt{160(1 - 10^{-5})}} \le \frac{150.5 - 160}{\sqrt{160(1 - 10^{-5})}}\right) \\ &\approx P(Z \le -0.75) = 0.227 \end{aligned}$$

Because $np = (16 \times 10^6)(1 \times 10^{-5}) = 160$ and $n(1-p)$ is much larger, the approximation is expected to work well in this case.

Practical Interpretation: Binomial probabilities that are difficult to compute exactly can be approximated with easy-to-compute probabilities based on the normal distribution.

EXAMPLE 4.15 | Normal Approximation to Binomial

Again consider the transmission of bits in Example 4.14. To judge how well the normal approximation works, assume that only $n = 50$ bits are to be transmitted and that the probability of an error is $p = 0.1$. The exact probability that two or fewer errors occur is

$$P(X \le 2) = \binom{50}{0} 0.9^{50} + \binom{50}{1} 0.1(0.9^{49}) + \binom{50}{2} 0.1^2(0.9^{48}) = 0.112$$

Based on the normal approximation,

$$P(X \le 2) = P\left(\frac{X - 5}{\sqrt{50(0.1)(0.9)}} \le \frac{2.5 - 5}{\sqrt{50(0.1)(0.9)}}\right) \approx P(Z < -1.18) = 0.119$$

We can also approximate $P(X = 5)$ as

$$P(5 \le X \le 5) = P(4.5 \le X \le 5.5) \approx P\left(\frac{4.5 - 5}{2.12} \le Z \le \frac{5.5 - 5}{2.12}\right) = P(-0.24 \le Z \le 0.24) = 0.19$$

and this compares well with the exact answer of 0.1849.

Practical Interpretation: Even for a sample as small as 50 bits, the normal approximation is reasonable, when $p = 0.1$.

Normal Approximation to the Poisson Distribution

If X is a Poisson random variable with $E(X) = \lambda$ and $V(X) = \lambda$,

$$Z = \frac{X - \lambda}{\sqrt{\lambda}} \qquad (4.13)$$

is approximately a standard normal random variable. The same continuity correction used for the binomial distribution can also be applied. The approximation is good for

$$\lambda > 5$$

EXAMPLE 4.16 | Normal Approximation to Poisson

Assume that the number of asbestos particles in a squared meter of dust on a surface follows a Poisson distribution with a mean of 1000. If a squared meter of dust is analyzed, what is the probability that 950 or fewer particles are found?

This probability can be expressed exactly as

$$P(X \le 950) = \sum_{x=0}^{950} \frac{e^{-1000} 1000^x}{x!}$$

The computational difficulty is clear. The probability can be approximated as

$$P(X \le 950) = P(X \le 950.5) \approx P\left(Z \le \frac{950.5 - 1000}{\sqrt{1000}}\right) = P(Z \le -1.57) = 0.058$$

Practical Interpretation: Poisson probabilities that are difficult to compute exactly can be approximated with easy-to-compute probabilities based on the normal distribution.

3.3.2 Exponential Distribution

Exponential Distribution

The random variable X that equals the distance between successive events from a Poisson process with mean number of events $\lambda > 0$ per unit interval is an **exponential random variable** with parameter λ. The probability density function of X is

$$f(x) = \lambda e^{-\lambda x} \quad \text{for} \quad 0 \le x < \infty \tag{4.14}$$

Mean and Variance

If the random variable X has an exponential distribution with parameter λ,

$$\mu = E(X) = \frac{1}{\lambda} \quad \text{and} \quad \sigma^2 = V(X) = \frac{1}{\lambda^2} \tag{4.15}$$

EXAMPLE | Computer Usage

In a large corporate computer network, user log-ons to the system can be modeled as a Poisson process with a mean of 25 log-ons per hour. What is the probability that there are no log-ons in an interval of 6 minutes?

Let X denote the time in hours from the start of the interval until the first log-on. Then X has an exponential distribution with $\lambda = 25$ log-ons per hour. We are interested in the probability that X exceeds 6 minutes. Because λ is given in log-ons per hour, we express all time units in hours. That is, 6 minutes = 0.1 hour. Therefore,

$$P(X > 0.1) = \int_{0.1}^{\infty} 25e^{-25x}dx = e^{-25(0.1)} = 0.082$$

The cumulative distribution function can also be used to obtain the same result as follows:

$$P(X > 0.1) = 1 - F(0.1) = e^{-25(0.1)}$$

An identical answer is obtained by expressing the mean number of log-ons as 0.417 log-ons per minute and computing the probability that the time until the next log-on exceeds 6 minutes. Try it.

What is the probability that the time until the next log-on is between 2 and 3 minutes? Upon converting all units to hours,

$$P(0.033 < X < 0.05) = \int_{0.033}^{0.05} 25e^{-25x}dx = -e^{-25x}\Big|_{0.033}^{0.05} = 0.152$$

An alternative solution is

$$P(0.033 < X < 0.05) = F(0.05) - F(0.033) = 0.152$$

Determine the interval of time such that the probability that no log-on occurs in the interval is 0.90. The question asks for the length of time x such that $P(X > x) = 0.90$. Now,

$$P(X > x) = e^{-25x} = 0.90$$

Take the (natural) log of both sides to obtain $-25x = \ln(0.90) = -0.1054$. Therefore,

$$x = 0.00421 \text{ hour} = 0.25 \text{ minute}$$

Furthermore, the mean time until the next log-on is

$$\mu = 1/25 = 0.04 \text{ hour} = 2.4 \text{ minutes}$$

The standard deviation of the time until the next log-on is

$$\sigma = 1/25 \text{ hours} = 2.4 \text{ minutes}$$

Practical Interpretation: Organizations make wide use of probabilities for exponential random variables to evaluate resources and staffing levels to meet customer service needs.

4 Introduction à l'apprentissage automatique

Références :

- https://eduscol.education.fr/sti/sites/eduscol.education.fr.sti/files/ressources/pedagogiques/14512/14512-introduction-lapprentissage-automatique-ensps.pdf

Le domaine de l'intelligence artificielle a pour objectif de parvenir à simuler l'intelligence humaine et en particulier l'apprentissage de nombreuses tâches. Deux méthodes sont alors possibles pour apprendre :

- L'apprentissage par cœur consiste à mémoriser explicitement tous les exemples possibles afin de pouvoir les restituer ;
- L'apprentissage par généralisation a pour objectif d'extraire des règles implicites à partir d'une quantité d'exemples afin de les réappliquer à de nouvelles situations jamais rencontrées. L'apprentissage par cœur est relativement aisé pour une machine à condition de disposer des exemples. En revanche, l'apprentissage par généralisation est difficile car il demande d'extraire des règles qui ne sont pas explicitement mentionnées dans les exemples. Ce défi constitue le cœur l'apprentissage automatique.

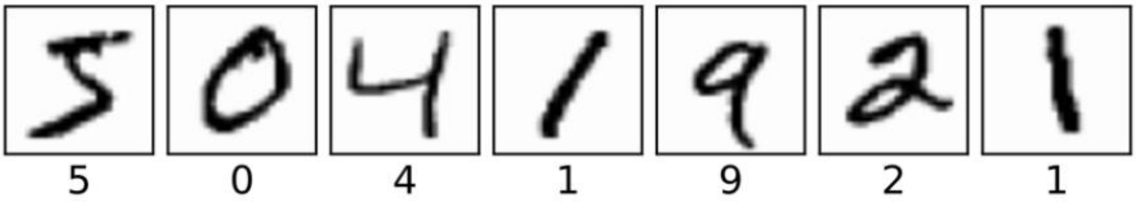
Comme expliqué précédemment, l'apprentissage automatique est un champ de l'IA porté sur l'analyse statistiques de données d'apprentissage. Historiquement, cette branche est définie comme le développement de machines capables d'apprendre sans avoir été explicitement programmées à apprendre une tâche.

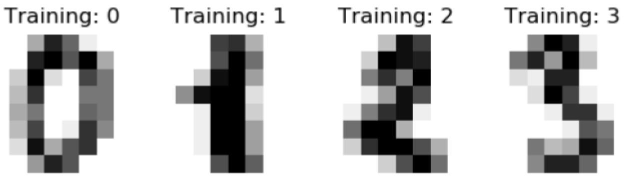
4.1 Exemple 1

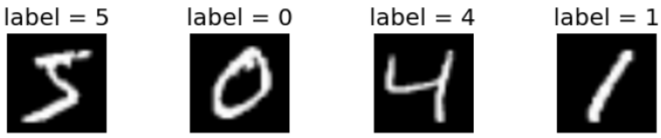
Figure : Exemple de chiffres manuscrits
Extrait de la base de données MNIST, couramment utilisée en apprentissage automatique

Etapes d'apprentissage automatique (supervisé) :

1. **Classification manuelle** des objets de la base d'apprentissage :



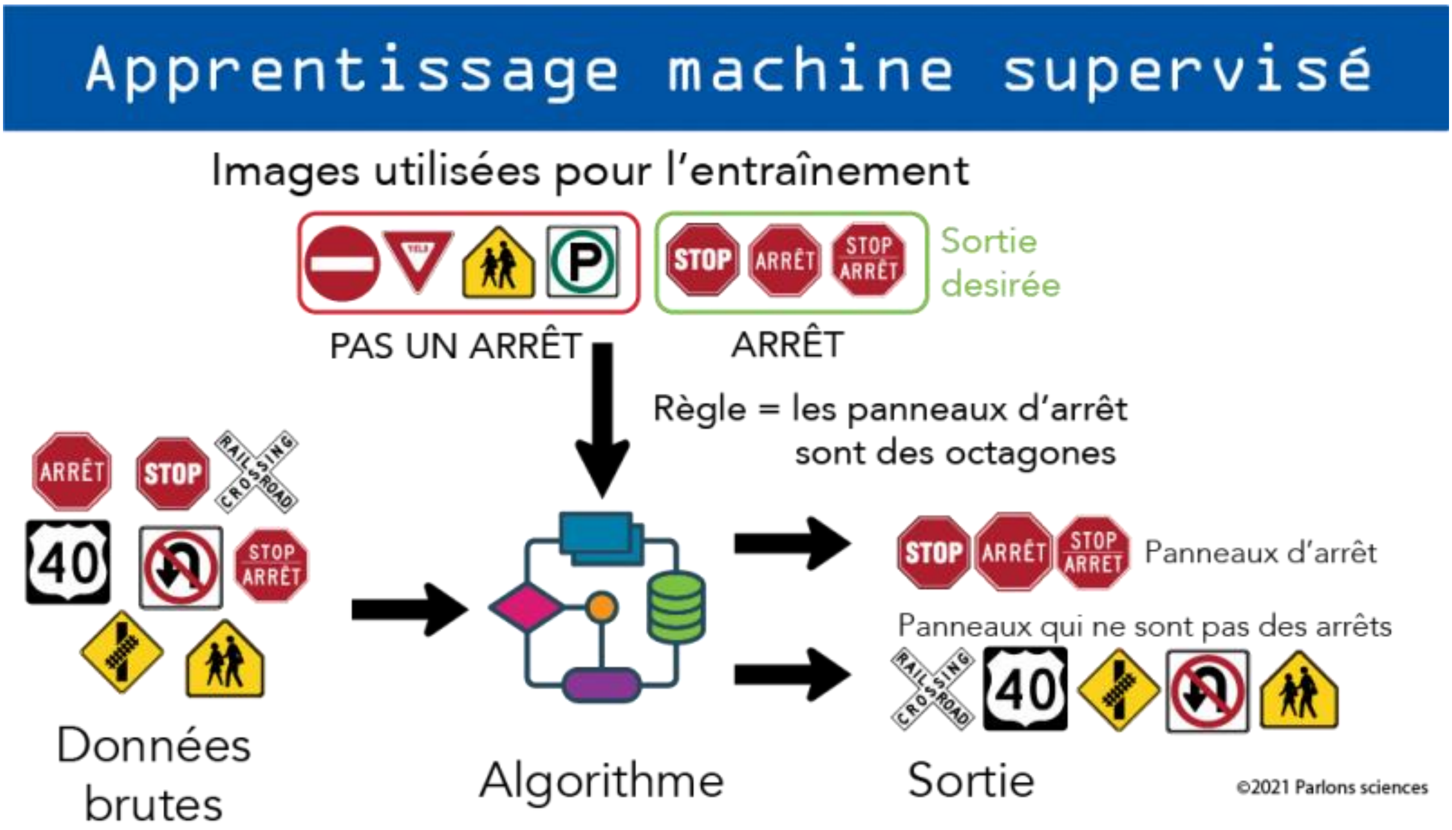




…………………………
…………………………

2. Application d'un **algorithme d'apprentissage** des classes et leurs objets
3. Application grandeur réelle pour **reconnaitre de nouveaux objets** :

4.2 Exemple 2


Apprentissage machine supervisé
Images utilisées pour l'entraînement
Sortie desirée
PAS UN ARRÊT
ARRÊT
Règle = les panneaux d'arrêt sont des octagones
Données brutes
Algorithme
Panneaux d'arrêt
Panneaux qui ne sont pas des arrêts
Sortie
©2021 Parlons sciences