Lois de probabilité

Catherine Labruère Chazal

Université de Bourgogne, Licence 2 Gestion

September 18, 2024

Terminologie

- Un *individu* est un élément d'un ensemble, appelé *population*
- Sur un individu dans une population, on peut regarder la valeur d'une *variable*.
- Si l'individu est pris au hasard, cette valeur est *aléatoire*.
- Exemples
  - Population : les salariés en France. Individu : un salarié. Variable : salaire.
  - Population : les entreprises en France en 2020. Individu : une entreprise. Variable : nombre de jours de congés maladie

- Une variable X est *discrète* si on peut dénombrer les valeurs qu'elle prend.
- Ex : nombre d'enfants d'une famille en France
- Une variable X est *continue* sinon.
- Ex : durée de vie d'un composant électronique (en secondes)

Lois de probabilité discrètes

Définitions

Soit X une variable aléatoire (v.a.) discrète prenant les valeurs $\{x_1, x_2, ..., x_n, ...\}$

- ☞ La *loi de probabilité (ou densité)* de X est définie par la donnée des nombres :

$$\mathbb{P}(X = x_1), \mathbb{P}(X = x_2,), \ldots, \mathbb{P}(X = x_n), ...$$

  dont la somme fait 1.

- ☞ L'exemple le plus naturel de loi discrète est la densité uniforme : toutes les valeurs ont la même probabilité.
- ☞ Exemple : on lance un dé à 6 faces, X est le numéro obtenu.
- ☞ Pour une loi continue, la densité sera définie différemment (par une fonction).

- ☞ *L'espérance de X* est

$$\mathbb{E}(X) = \sum_{i=1}^{n} x_i \times \mathbb{P}(X = x_i)$$

- ☞ L'espérance d'une variable aléatoire est une valeur théorique qui donne la valeur prise en moyenne par cette variable.

- *La variance* de X est

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

- *L'écart-type* de X est

$$\mathbb{S}(X) = \sqrt{\mathbb{V}(X)}.$$

- $\mathbb{V}(X)$ et $\mathbb{S}(X)$ indiquent si X a des valeurs très ou peu dispersées autour de $\mathbb{E}(X)$.

Lois de probabilité discrètes

Loi de Bernoulli

- Une *expérience aléatoire de Bernoulli* possède deux résultats possibles appelés succès et échec. Si p est la probabilité de succès, la probabilité d'échec est $1 - p$.
- Une v.a. X suit *une loi de Bernoulli de paramètre p* si elle vaut 1 quand il y a succès et 0 sinon.
- Notation : $X \sim \mathcal{B}(p)$.
- La loi de X est donnée par :

$$\mathbb{P}(X = 1) = p \; , \; \mathbb{P}(X = 0) = 1 - p$$

Lois de probabilité discrètes

Loi binomiale

- Soit n un entier. Un *processus de n expériences de Bernoulli* est la répétition de n expériences aléatoires de Bernoulli identiques et indépendantes.
- On peut représenter les résultats d'un processus de Bernoulli par un arbre.
- Soit k un entier compris entre 0 et n. Le *coefficient binomial $\binom{n}{k}$* est le nombre de chemins de l'arbre réalisant k succès.
- On a la propriété :

$$\binom{n}{k} = \binom{n}{n-k}$$

- Une v.a. X suit *une loi binomiale de paramètres* (n, p) si elle donne le nombre de succès lors d'un processus de n expériences de Bernoulli $\mathcal{B}(p)$.
- Notation : $X \sim \mathcal{B}(n, p)$
- Les valeurs possibles pour X sont $k = 0, 1, ..., n$.
- et la loi de X est donnée par :

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Densité de la loi binomiale

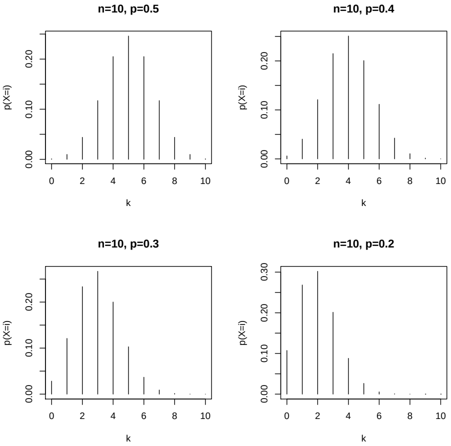


Figure: densité de la loi binomiale

Espérance et variance des lois de Bernoulli et binomiale

- ☞ Si $X \sim \mathcal{B}(p)$, $\mathbb{E}(X) = p$ et $\mathbb{V}(X) = p(1-p)$
- ☞ Si $X \sim \mathcal{B}(n, p)$, $\mathbb{E}(X) = np$ et $\mathbb{V}(X) = np(1-p)$

Exercice 1
Une entreprise de commerce électronique observe que 5% des commandes sont annulées par les clients. On prélève un échantillon de 10 commandes pour une analyse de performance.

1. Déterminer le nombre moyen de commandes annulées dans un échantillon de 10 commandes..
2. Calculez la probabilité de trouver au maximum 2 commandes annulées dans cet échantillon de 10 commandes.
3. Calculez la probabilité de trouver au minimum 2 commandes annulées dans cet échantillon de 10 commandes.

- Une v.a. X suit <u>*une loi géométrique de paramètres p*</u> si elle donne le nombre d'épreuves de Bernoulli $\mathcal{B}(p)$ à réaliser avant d'obtenir un succès.
- Notation : $X \sim \mathcal{G}(p)$.
- Les valeurs possibles pour X sont infinies : $k = 1, 2, ...$
- La loi de probabilité de X est donnée par

$$\mathbb{P}(X = k) = (1 - p)^{k-1} \times p$$

Exercice 2

- ☞ On joue à un jeu dans lequel on lance un dé plusieurs fois de suite et on s'arrête dès que l'on obtient un 6.
- ☞ Quelle est la probabilité que l'on s'arrête au bout de 5 lancers?

Lois de probabilité discrètes

Loi de Poisson

- La loi de Poisson sert par exemple à modéliser le nombre d'événements se produisant au cours d'une durée fixée.
  Ex : nombre d'accidents par an
- Elle se définit à partir du nombre moyen m d'évènements au cours de cette durée :
- On dit que X suit une <u>*loi de Poisson*</u> de paramètre m si X prend les valeurs $k = 0, 1, 2, ...$ et a pour loi

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

  avec $k\,! = 1 \times 2 \times 3 \times ... \times k$ et $0\,! = 1$.
- *Notation:* $X \sim \mathcal{P}(m)$.
- $$\mathbb{E}(X) = m, \quad \mathbb{V}(X) = m.$$

Lois de probabilité discrètes

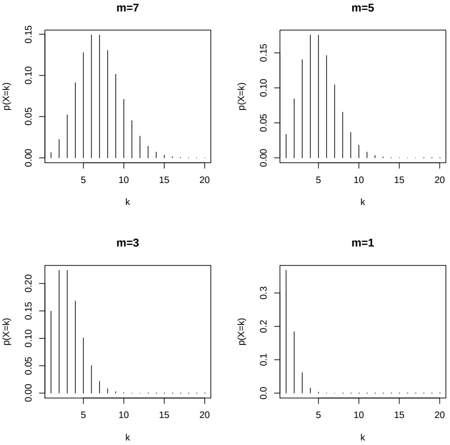


Figure: densité de la loi de Poisson

Une propriété importante :

☞ Si $X_1, \ldots, X_n$ sont des v.a. indépendantes de loi $\mathcal{P}(m)$ alors :

$$X_1 + ... + X_n \text{ suit la loi } \mathcal{P}(nm).$$

Exercice 3

Notons X le nombre d'accidents subis par le personnel d'une usine dans une journée de travail. On suppose que X suit une loi de Poisson. On sait que le nombre d'accidents moyen par jour dans ce type d'usine est 0.8.

☞ Calculer la probabilité qu'il y ait un accident au cours d'une journée, puis au moins deux accidents.

☞ Soit Y la variable aléatoire comptant le nombre d'accidents sur une semaine (5 jours). Donner la loi de Y.

☞ Calculer la probabilité qu'il y ait au moins 3 accidents sur une semaine.

Soit X un v.a. continue

- ☞ Sa loi est définie par une *fonction densité* f : fonction positive dont l'intégrale sur tout le domaine vaut 1.
- ☞ La probabilité que X prenne une valeur donnée est nulle $\mathbb{P}(X = x) = 0$.
- ☞ On calcule la probabilité d'un intervalle avec

$$\mathbb{P}(a \leq X \leq b) = \mathbb{P}(a < X < b) = \int_a^b f(t)dt.$$

C'est l'aire sous le graphe de f comprise entre a et b.

Lois de probabilité continues

Densité, fonction de répartition

- ☞ On appelle *fonction de répartition* la fonction F définie par :

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X < x) = \int_{-\infty}^{x} f(t)dt.$$

- ☞ C'est l'aire sous la courbe entre $-\infty$ et x.
- ☞ Elle permet de calculer des probabilités, par exemple

$$\mathbb{P}(a \leq X \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(X < a) = F(b) - F(a)$$

Lois de probabilité continues

Loi exponentielle

- Une v. a. X suit une *loi exponentielle de paramètre $\lambda > 0$* si elle prend des valeurs positives et a pour densité

$$f(x) = \lambda e^{-\lambda x}$$

- Notation : $X \sim \mathcal{E}(\lambda)$
- Cette loi modélise par exemple la durée de vie d'un composant électronique.

Lois de probabilité continues

Loi exponentielle

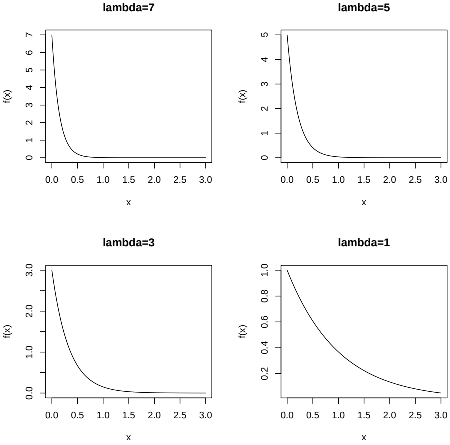


Figure: Densité de la loi exponentielle

- ☞ $-e^{-\lambda x}$ est une primitive de $f(x)$
- ☞ Donc la fonction de répartition est :

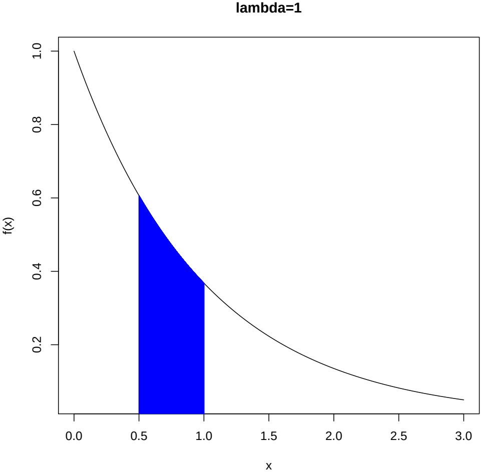
$$F(x) = -e^{-\lambda x} - (-e^{-\lambda \times 0}) = 1 - e^{-\lambda x}$$

- ☞ $\mathbb{P}(a \leq X \leq b) = F(b) - F(a) = e^{-\lambda a} - e^{-\lambda b}$
- ☞ $\mathbb{P}(X \geq a) = 1 - \mathbb{P}(X < a) = 1 - F(a) = e^{-\lambda a}$

Lois de probabilité continues

Loi exponentielle

$\mathbb{P}(a \leq X \leq b)$ est l'aire sous le graphe de la densité entre a et b.



Espérance, variance

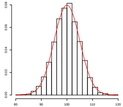
- ☞ $\mathbb{E}(X) = \frac{1}{\lambda}$
- ☞ $\mathbb{V}(X) = \frac{1}{\lambda^2}$.
- ☞ $\frac{1}{\lambda}$ est donc la valeur moyenne de X et également l'écart-type des valeurs de X.

Exercice 4

- ☞ Soit T donnant la durée de vie (en jours) d'un certain type de composant électronique. Supposons que $T \sim \mathcal{E}(0.0002)$
- ☞ Donner la densité et la fonction de répartition de T.
- ☞ Déterminer la probabilité pour qu'un composant pris au hasard ait une durée de vie supérieure à 2 000 jours.
- ☞ Déterminer la durée de vie moyenne des composants.
- ☞ Déterminer la valeur t_0 pour laquelle $\mathbb{P}(T < t_0) = 0.5$ (appelée demi-vie d'un composant).

Loi normale
Définition

☞ Depuis Gauss et Pascal, on sait démontrer que pour les grandes populations, les histogrammes représentant les mesures de pesées, de tailles, etc ... sont des courbes de Gauss :



définies par la fonction densité :

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

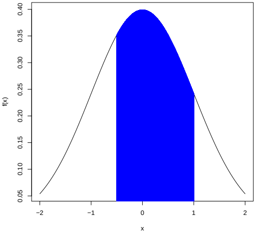
Loi normale
Définition

- ☞ Définition : une v.a. X suit *une loi normale* de moyenne m et d'écart type σ si sa densité est la fonction f ci-dessus.
- ☞ Notation: $X \sim \mathcal{N}(m, \sigma^2)$

Loi normale

Définition

On ne sait pas calculer de primitives de la densité de la loi normale. Les calculs de probabilité se font donc avec une table ou une calculatrice (intégration numérique).



Loi normale

Espérance, variance

Pour $X \sim \mathcal{N}(m, \sigma^2)$

☞ $\mathbb{E}(X) = m$

☞ $\mathbb{V}(X) = \sigma^2$.

Loi normale

Utilisation de la table

L'utilisation est basée sur la propriété *de changement de variable* :

- ☞ Si la variable X suit une loi normale $\mathcal{N}(m, \sigma^2)$ alors la variable $Z = \frac{X-m}{\sigma}$ suit une loi normale $\mathcal{N}(0,1)$.
- ☞ La loi $\mathcal{N}(0,1)$ est appelée loi normale centrée ($m = 0$), réduite ($\sigma = 1$).

Loi normale

Utilisation de la table

Soit $X \sim \mathcal{N}(m, \sigma^2)$

☞ Pour calculer , $\mathbb{P}(a \leq X \leq b)$, on écrit

$$\mathbb{P}(a \leq X \leq b) = \mathbb{P}\Big(\frac{a-m}{\sigma} \leq \frac{X-m}{\sigma} \leq \frac{b-m}{\sigma}\Big)$$

$$= \mathbb{P}\Big(\frac{a-m}{\sigma} \leq Z \leq \frac{b-m}{\sigma}\Big).$$

☞ Puis on utilise la table.

Calculs avec la loi Z normale centrée réduite

- La table donne la valeur approchée de

$$\mathbb{P}(0 \leq Z \leq z) \quad \text{où} \quad Z \sim \mathcal{N}(0;1).$$

  que l'on notera $\varphi(z)$.
- Ex : $\mathbb{P}(0 \leq Z \leq 1.23) = 0.3907$
- Remarque : comme l'intégrale sous la courbe fait 1

$$\mathbb{P}(Z \geq 0) = \mathbb{P}(Z \leq 0) = 0.5$$

Calculs avec la loi Z normale centrée réduite

Les autres calculs (faire des dessins!)

- ☞ $\mathbb{P}(-2.51 \leq Z \leq 0) = \varphi(2.51) = 0.494$
- ☞ $\mathbb{P}(1.33 \leq Z \leq 2.51) = \varphi(2.51) - \varphi(1.33) = 0.0858$
- ☞ $\mathbb{P}(-3.12 \leq Z \leq -2.51) = \varphi(3.12) - \varphi(2.51) = 0.0051$
- ☞ $\mathbb{P}(-2.51 \leq Z \leq 1.33) = \varphi(2.51) + \varphi(1.33) = 0.9022$
- ☞ $\mathbb{P}(Z \geq 1.33) = 0.5 - \varphi(1.33) = 0.0918$
- ☞ $\mathbb{P}(Z \geq -2.51) = 0.5 + \varphi(2.51) = 0.994$
- ☞ $\mathbb{P}(Z \leq 2.51) = 0.5 + \varphi(2.51) = 0.994$
- ☞ $\mathbb{P}(Z \leq -3.12) = 0.5 - \varphi(3.12) = 0.0009$

Calcul avec une loi normale non centrée réduite

Soit X une v.a. qui suit une loi $X \sim \mathcal{N}(100, 25)$.

☞ **Calculer** $\mathbb{P}(X \leq 105)$:

$$\mathbb{P}(X \leq 105) = \mathbb{P}\left(\frac{X - 100}{5} \leq \frac{105 - 100}{5}\right)$$

On pose $Z = \frac{X-100}{5}$. On a $Z \sim \mathcal{N}(0, 1)$.

$$= \mathbb{P}(Z \leq 1) = 0.5 + \varphi(1) = 0.5 + 0.3413 = 0.8413.$$

Lecture indirecte de la table

Lecture indirecte = on connaît $\varphi(z)$ et on veut z.

- ☞ Si $\varphi(z)$ n'est pas dans la table, on prend la valeur la plus proche.
- ☞ Exemple : chercher z tel que $\varphi(z) = 0.472$. Réponse : $z = 1.91$.

Trouver un quantile

Soit p une probabilité donnée. Le nombre z tel que $\mathbb{P}(Z \leq z) = p$ est appelé *quantile de probabilité p*. Comment trouver un quantile?

- ☞ **En faisant un dessin**, on se rend compte que deux cas se présentent selon les valeurs de p :
- ☞ Si $p \geq 0.5$, alors $z > 0$ et $\varphi(z) = p - 0.5$.
- ☞ Si $p \leq 0.5$, alors $z < 0$ et $\varphi(|z|) = 0.5 - p$.

Exemple

- **Trouver a tel que $\mathbb{P}(X \leq a) = 0.05$ avec $X \sim \mathcal{N}(100, 25)$.**
- Or
$$\mathbb{P}(X \leq a) = \mathbb{P}\left(Z \leq \frac{a - 100}{5}\right).$$
- Posons $z = \frac{a-100}{5}$. On cherche z tel que $\mathbb{P}(Z \leq z) = 0.05$.
- z est forcément négatif et $\varphi(|z|) = 0.5 - 0.05 = 0.45$.
- La valeur 0.45 n'est pas dans la table, on remarque que c'est exactement le milieu des deux valeurs $0.4495 = \varphi(1.64)$ et $0.4505 = \varphi(1.65)$.
- Donc $z = -1.645$ et donc $a = 5z + 100 = 91.775$.

Loi normale

Approximation de la loi binomiale par la loi normale

Soit X une v.a. suivant une loi binomiale $\mathcal{B}(n, p)$.

- ☞ <u>*Théorème de Moivre-Laplace*</u> Si $n > 30, np > 5, n(1-p) > 5$, alors X peut être approximée par la loi normale d'espérance np et de variance $np(1-p)$.
- ☞ Notation: $X \approx Y$ où $Y \sim \mathcal{N}(np, np(1-p))$
- ☞ Conséquences : Si $n > 30, np > 5, n(1-p) > 5$
- ☞ $\mathbb{P}(X = b) \approx \mathbb{P}(b - 0.5 \leq Y \leq b + 0.5)$
- ☞ $\mathbb{P}(X \leq b) \approx \mathbb{P}(Y \leq b + 0.5)$
- ☞ $\mathbb{P}(X \geq a) \approx \mathbb{P}(Y \geq a - 0.5)$
- ☞ $\mathbb{P}(a \leq X \leq b) \approx \mathbb{P}(a - 0.5 \leq Y \leq b + 0.5)$

Exemple

Pour $X \sim \mathcal{B}(40; 0.4)$. Calculer $\mathbb{P}(16 \leq X \leq 20)$

- ☞ Calcul exact : $\mathbb{P}(16 \leq X \leq 20) = 0.4853$.
- ☞ En approximant avec la loi normale,

$$\mathbb{P}(16 \leq X \leq 20) \approx \mathbb{P}(15.5 \leq Y \leq 20.5) \text{ où } Y \sim \mathcal{N}(16; 9.6).$$

$$\text{Or } \mathbb{P}(15.5 \leq Y \leq 20.5) = \mathbb{P}(-0.16 \leq Z \leq 1.45) \text{ où } Z \sim \mathcal{N}(0; 1).$$

$$\text{Donc } \mathbb{P}(16 \leq X \leq 20) \approx \varphi(1.45) + \varphi(0.16) = 0.4901.$$

Loi normale

Exercice 5

☞ Dans une entreprise, chaque personne téléphone en moyenne 10 minutes par heure à l'extérieur. Il y a 300 employés et un standard qui filtre les appels vers l'extérieur. L'entreprise souhaite s'abonner à un nombre limité de lignes extérieures. Quel est le nombre minimal de lignes qu'elle doit prévoir pour que le risque d'encombrement soit inférieur à 5%?

Loi normale

Somme de lois normales indépendantes

☞ Si $X_1, ..., X_n$ sont des v.a. indépendantes de loi $\mathcal{N}(m;\sigma^2)$ alors :

$$X_1 + \cdots + X_n \sim \mathcal{N}(nm; n\sigma^2).$$

Loi normale

Exercice 6

Vous êtes responsable de la sécurité de votre entreprise. Vous avez commandé un ascenseur sur lequel est inscrit : CHARGE MAX=800kg. En relevant le poids de votre personnel, vous avez constaté que le poids moyen est de 70kg avec un écart type de 6kg. En supposant que le poids d'une personne suit une loi normale, calculer le risque de surcharge si 12 personnes montent en même temps?