Introduction à la théorie des probabilités et à la statistique

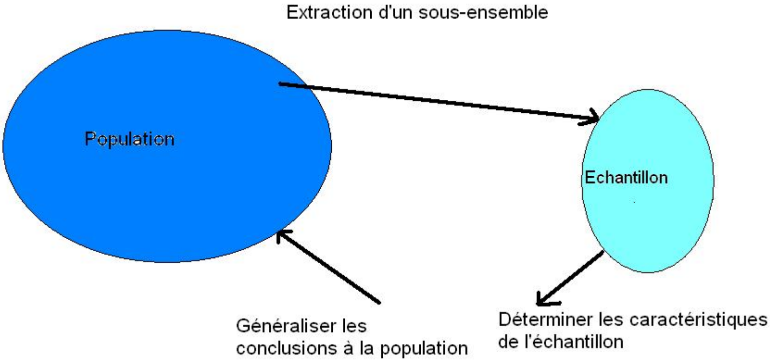
Ségolen Geffray

IUT Carquefou

Année 2008-2009

segolen.geffray@univ-nantes.fr

La démarche


Extraction d'un sous-ensemble
Population
Echantillon
Généraliser les
conclusions à la population
Déterminer les caractéristiques
de l'échantillon

Distinction Proba/Stat

- **La Théorie des probabilités** :
  - permet de modéliser des phénomènes aléatoires et d'y effectuer des calculs théoriques
  - concerne les **populations** : on ne peut donc pas faire de mesures.
- **La Statistique** :
  - concerne les **échantillons**, le monde réel, la pratique,
  - on fait des mesures (observations) sur des individus,
  - repose sur la modélisation probabiliste des observations.

Un exemple concret

- Un fabricant d'ampoules souhaite vérifier la qualité des ampoules électriques produites dans sa chaine de montage.
- Pour cela, il propose donc d'évaluer la durée moyenne de bon fonctionnement d'une ampoule.
- Comment faire ? on ne peut pas tester toutes les ampoules produites par la chaine de montage !
- On tire un échantillon au hasard
- On réalise l'expérience, on effectue des mesures, on calcule la durée moyenne de bon fonctionnement des ampoules de l'échantillon
- On approxime la durée moyenne de bon fonctionnement des ampoules de la population entière par la durée moyenne de bon fonctionnement des ampoules de l'échantillon

Un exemple concret (suite)

- Comment savoir si ce qu'on vient de faire est licite ? Quelle est la qualité de l'approximation ? Il faut étudier la théorie des probabilités et la statistique !
- Si on tire un autre échantillon, il y a de fortes chances que l'on n'obtienne pas les mêmes résultats.
- Ces fluctuations (ou erreurs d'échantillonnage) sont dues à la variabilité. Cela signifie que des objets semblables en apparence peuvent présenter des différences lorsqu'on effectue des mesures.

Les différents aspects de la Statistique

Observer ne suffit pas !

- **Statistique descriptive** :
  - Résumer les mesures sur un échantillon (moyenne, variance,...)
  - Représenter les mesures (histogramme, distribution)
- **Statistique inférentielle** :
  - Généraliser les propriétés d'un échantillon à une population en prenant en compte les fluctuations d'échantillonnage
  - il faut modéliser les observations (par des variables aléatoires) : on fait appel à la théorie des probabilités
- **Tests d'hypothèses** :
  - Contrôler la validité d'un modèle
  - Comparer un échantillon à une référence
- **Statistique décisionnelle** :
  - Savoir prendre une décision alors que les résultats sont exprimés en termes de probabilités (i.e. de pourcentage de chances, de risques)

1$^{\text{ère}}$ partie

Notions de théorie des probabilités

Expérience aléatoire, évènements

- Une expérience est **aléatoire** si on ne peut pas prévoir à l'avance son résultat, et si, répétée dans des conditions identiques, elle peut donner lieu à des résultats différents. Lorsqu'on effectue une expérience, les valeurs obtenues s'appellent des **réalisations** ou des observations.
- **Univers** (noté Ω) : ensemble de tous les résultats possibles d'une expérience. Il peut être :
  - fini, par ex $\{x_1, ..., x_k\}$
  - infini dénombrable : on peut indicer, numéroter ses éléments jusqu'à l'infini, par ex $\{x_1, , x_2, ..., x_n, ...\}$
  - infini non-dénombrable : ceci signifie qu'il n'est pas possible de décrire l'ensemble sous la forme d'une liste numérotée $\{x_1, x_2, ..., x_k, ...\}$, par ex l'intervalle $[0, 1]$ est un ensemble infini non-dénombrable.
- **Evènement élémentaire** : un des éléments de Ω lorsqu'on peut les énumérer.
- **Evènement** : sous-ensemble de Ω.

Exercice : univers finis

1. Soit l'expérience "une personne lance un dé cubique à 6 faces et note la valeur obtenue sur la face supérieure de ce dé". Quel est l'univers associé à cette expérience ? Quels sont les évènements élémentaires ?
2. Soit l'expérience "une personne lance simultanément deux dés cubiques à 6 faces et note la valeur obtenue sur la face supérieure de chaque dé". Quel est l'univers associé à cette expérience ? Quels sont les évènements élémentaires ?
3. Soit l'expérience "une personne lance simultanément deux dés cubiques à 6 faces et note la somme des valeurs obtenues sur la face supérieure de chaque dé". Quel est l'univers associé à cette expérience ? Quels sont les évènements élémentaires ?

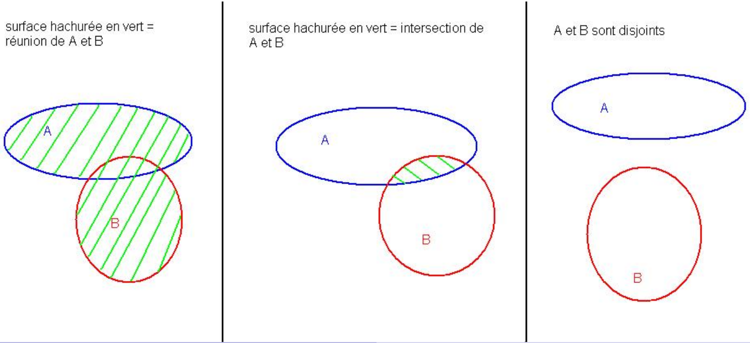
Exercice : univers infinis

1. Soit l'expérience “Alex compte le nombre de véhicules se présentant au péage de l'autoroute en une journée”. Quel est l'univers associé à cette expérience ?
2. Soit l'expérience “Mr Martin note, comme chaque midi, la température extérieure”. Mr Martin habite à Paris où la température à 12h peut varier de -10°C à 43°C. Quel est l'univers associé à cette expérience ?
3. Soit l'expérience “Mr Jean note, comme chaque lundi, la durée de son vol Paris-Berlin”. Le vol entre Paris et Berlin dure 1h45, peut avoir jusqu'à 15 minutes d'avance si le vent est favorable et jusqu'à 3h de retard en cas de problème. Quel est l'univers associé à cette expérience ?
4. Soit l'expérience “le technicien mesure et pèse une tige tirée de la production”. La tige est usinée de sorte qu'elle pèse entre 12g et 25g et mesure entre 8.5cm et 11.5cm. Quel est l'univers associé à cette expérience ?

Rappel : opérations sur les ensembles

Soient A et B deux évènements d'un ensemble fondamental Ω

- $\{A \text{ ou } B\} = A \cup B =$ réunion de A et B
- $\{A \text{ et } B\} = A \cap B =$ intersection de A et B
- complémentaire de A dans $\Omega = \overline{A} = \Omega - A$
- $\emptyset =$ évènement impossible
- $\Omega =$ évènement certain
- A et B sont incompatibles ou disjoints lorsque $A \cap B = \emptyset$



Exercice : évènements

Soit l'expérience "une personne lance un dé cubique à 6 faces et note la valeur obtenue sur la face supérieure de ce dé".

1. Comment décrire l'évènement A="obtenir une valeur inférieure à 4"?
2. Comment décrire l'évènement B="obtenir une valeur paire"?
3. Comment décrire l'évènement C="obtenir une valeur inférieure à 4 ou une valeur paire"?
4. Comment décrire l'évènement D="obtenir une valeur inférieure à 4 et une valeur paire"?
5. Comment décrire l'évènement E="obtenir une valeur supérieure ou égale à 8"?
6. Comment décrire l'évènement F="obtenir une valeur inférieure ou égale à 6"?

Exercice : évènements (suite)

Soit l'expérience “une personne lance simultanément deux dés cubiques à 6 faces et note la valeur obtenue sur la face supérieure de chaque dé”.

1. Comment décrire l'évènement A=“obtenir au moins un 6”?
2. Comment décrire l'évènement B=“obtenir une somme des deux valeurs supérieure ou égale à 10”?
3. Comment décrire l'évènement C=“obtenir au moins un 6 et obtenir une somme des deux valeurs supérieure ou égale à 10”?
4. Comment décrire l'évènement D=“obtenir un produit des deux valeurs supérieur à 100”?

Règle de calcul des probabilités

- Une probabilité est une **fonction** notée $\mathbb{P}$ qui attribue à tout évènement A une valeur $\mathbb{P}(A)$ désignant la probabilité que A se réalise.
- Une probabilité possède les propriétés suivantes :
  - $0 \leq \mathbb{P}(A) \leq 1$ pour tout évènement A
  - $\mathbb{P}(\Omega) = 1$
  - $\mathbb{P}(\emptyset) = 0$
  - $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$
  - $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$
  - en général, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
  - mais si A et B sont disjoints, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$
  - en général, on a $\mathbb{P}\left(\cup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i)$
  - mais si les A_i sont 2 à 2 disjoints, $\mathbb{P}\left(\cup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)$

Probabilités sur les univers finis

- Soit Ω est un univers **fini**, on peut alors l'écrire sous la forme $\Omega = \{\omega_1, \ldots, \omega_n\}$. On note card($\Omega$) le nombre d'éléments de Ω qui représente le nombre de cas possibles à l'issue de l'expérience aléatoire.
- Lorsque Ω est un univers **fini**, il est parfois approprié de supposer que la probabilité associée à chaque évènement élémentaire est identique i.e. $\mathbb{P}[\omega_i] = 1/\text{card}(\Omega)$ pour $i = 1, \ldots, n$. On dit alors qu'il y a **équiprobabilité**.
- L'hypothèse d'équiprobabilité implique que la probabilité d'un évènement A s'obtient en calculant le rapport du nombre de cas possibles correspondant à l'évènement A (noté card(A)) sur le nombre de cas possibles, soit :

$$\mathbb{P}[A] = \frac{\text{card}(A)}{\text{card}(\Omega)}$$

Exemple d'équiprobabilité sur un univers fini

Soit l'expérience "une personne lance un dé cubique à 6 faces et note la valeur obtenue sur la face supérieure de ce dé". Soit, pour $i = 1, 2, ..., 6$, l'évènement ω_i="obtenir la valeur i". Effectuer l'hypothèse d'équiprobabilité revient à supposer que $\mathbb{P}[\omega_i] = 1/6$ pour $i = 1, 2, ..., 6$, ce qui correspond au cas où le dé est bien équilibré.

- Considérons l'évènement B="obtenir une valeur paire". Que vaut $\mathbb{P}[B]$?
- Considérons C="obtenir une valeur inférieure à 4 ou une valeur paire". Que vaut $\mathbb{P}[C]$?

Attention, ceci ne tient pas si le dé est truqué de sorte que $\mathbb{P}[\omega_1] = 1/24$, $\mathbb{P}[\omega_i] = 1/6$ pour $i = 2, ..., 5$ et $\mathbb{P}[\omega_6] = 7/24$.

- Que valent $\mathbb{P}[B]$ et $\mathbb{P}[C]$ dans ce cas ?

Probabilités conditionnelles : introduction

- Considérons une expérience réalisée sur une certaine population et un évènement A qui a une probabilité $\mathbb{P}[A]$ de se réaliser,
  par ex : $A =$ présence d'une maladie M.
- Que devient $\mathbb{P}[A]$ si on se restreint à une sous-population ?
  par ex : sous-population $=$ les individus présentant un signe S.
- On introduit un évènement B conditionnant, qui définit la sous-population,
  par ex : $B =$ présenter le signe S.
- $\mathbb{P}[B]$ ne doit pas être nul.

Probabilités conditionnelles : définition

- La probabilité que l'évènement A se réalise sachant que l'évènement B a eu lieu (=probabilité de A parmi la sous-population caractérisée par B) est définie par :

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}.$$

- De même, la probabilité que l'évènement B se réalise sachant que l'évènement A a eu lieu est définie par :

$$\mathbb{P}[B|A] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[A]}.$$

- Ne PAS confondre $\mathbb{P}[A|B]=$ probabilité que A se réalise sachant qu'on a observé B avec $\mathbb{P}[A \cap B]$=probabilité que A et B se réalisent simultanément ! !

Exemples de probabilités conditionnelles

Considérons les chiffres suivants valables pour la France :

- $\mathbb{P}[\text{être VHC+}] = 600000/60000000 = 1\%$
- $\mathbb{P}[\text{être VHC+ sachant que âge} = 15 \text{ ans}] =$ faible, certainement $< 10^{-4}$
- $\mathbb{P}[\text{être VHC+ sachant qu'il y a toxicomanie IV depuis} > 5 \text{ ans}]$ $=$ forte, certainement $> 50\%$
- $\mathbb{P}[\text{être VHC+ sachant que la personne est asthmatique}] = 1\%$

Probabilités conditionnelles : règles de calcul

- Soit B un évènement **fixé**. La fonction $A \to \mathbb{P}[A|B]$ est une vraie probabilité i.e. les règles de calcul avec les probabilités conditionnelles sont les mêmes qu'avec les probabilités classiques.
- $0 \leq \mathbb{P}(A|B) \leq 1$ pour tout évènement A
- $\mathbb{P}(\Omega|B) = 1$
- $\mathbb{P}(\emptyset|B) = 0$
- $\mathbb{P}(\overline{A}|B) = 1 - \mathbb{P}(A|B)$
- $A_1 \subset A_2 \Rightarrow \mathbb{P}(A_1|B) \leq \mathbb{P}(A_2|B)$
- en général,
  $\mathbb{P}(A_1 \cup A_2|B) = \mathbb{P}(A_1|B) + \mathbb{P}(A_2|B) - \mathbb{P}(A_1 \cap A_2|B)$
- mais si A_1 et A_2 sont disjoints,
  $\mathbb{P}(A_1 \cup A_2|B) = \mathbb{P}(A_1|B) + \mathbb{P}(A_2|B)$
- en général, on a $\mathbb{P}\left(\cup_{i=1}^{n} A_i \middle| B\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i|B)$
- mais si les A_i sont 2 à 2 disjoints,
  $\mathbb{P}\left(\cup_{i=1}^{n} A_i|B\right) = \sum_{i=1}^{n} \mathbb{P}(A_i|B)$

Indépendance de 2 évènements : introduction

- définition sans formule : soient A et B deux évènements. Si, lorsqu'on reçoit l'information que B s'est produit, cela ne modifie pas la probabilité de A, on dit que A et B sont indépendants. Autrement dit, des événements indépendants n'apportent pas d'information l'un sur l'autre.
- ex : l'asthme et le VHC sont indépendants : l'un n'aide pas au diagnostic de l'autre. Notons que $\mathbb{P}[\text{être VHC}+] = 1\%$ et que $\mathbb{P}[\text{être VHC}+ \text{ sachant que la personne est asthmatique}] = 1\%$.

Indépendance de 2 évènements : définition formelle

- $1^{\text{ère}}$ définition avec formule : A et B sont indépendants si

$$\mathbb{P}[A|B] = \mathbb{P}[A] \qquad \text{ou/et} \qquad \mathbb{P}[B|A] = \mathbb{P}[B].$$

- ex : La fréquence du VHC est 1% : $\mathbb{P}[\text{VHC}+] = 0.01$. La fréquence du VHC chez les asthmatiques est également de 1% : $\mathbb{P}[\text{VHC}+|\text{asthmatique}] = 0.01$.
- $2^{\text{ème}}$ définition avec formule : A et B sont indépendants si

$$\mathbb{P}[A \cap B] = \mathbb{P}[A] \times \mathbb{P}[B].$$

Cette formule est symétrique en A et B, on en déduit que si $\mathbb{P}[A|B] = \mathbb{P}[A]$, alors on a aussi $\mathbb{P}[B|A] = \mathbb{P}[B]$.
ex : $\mathbb{P}[\text{asthmatique}|\text{VHC}+] = \mathbb{P}[\text{asthmatique}]$ puisque asthme et VHC sont indépendants.

Indépendance et incompatibilité

- Ne pas confondre des événements incompatibles et des évènements indépendants...
- ex : considérons A="l'enfant à naître est un garçon" et B="l'enfant à naître est une fille". Les évènements A et B sont incompatibles. Mais ils ne sont pas indépendants ! ! ! En effet,

$$\mathbb{P}[A \cap B] = 0 \neq \mathbb{P}[A] \times \mathbb{P}[B] = 0.5 \times 0.5 = 0.25$$

Système complet

- Une famille d'évènements $\{A_1, A_2, \ldots, A_n\}$ forme un système complet d'évènements si
  - $A_1 \neq \emptyset$, $A_2 \neq \emptyset$,...,$A_n \neq \emptyset$
  - $A_1 \cap A_2 = \emptyset$, $A_1 \cap A_3 = \emptyset$,...,$A_{n-1} \cap A_n = \emptyset$ (i.e. les évènements du système complet sont deux à deux disjoints)
  - $\Omega = A_1 \cup A_2 \cup \ldots \cup A_n$
- Par exemple, soit l'univers $\Omega = \{1, 2, 3, 4, 5, 6\}$. On peut définir un système complet d'évènements $\{A_1, A_2, A_3\}$ avec, par exemple, $A_1 = \{1, 2\}$, $A_2 = \{3, 5\}$ et $A_3 = \{4, 6\}$. On a bien $A_1 \neq \emptyset$, $A_2 \neq \emptyset$, $A_3 \neq \emptyset$, $A_1 \cap A_2 = \emptyset$, $A_1 \cap A_3 = \emptyset$, $A_2 \cap A_3 = \emptyset$ et $\Omega = A_1 \cup A_2 \cup A_3$.
- Lorsque A est un évènement de probabilité non nulle, $\{A, \overline{A}\}$ forme un système complet d'évènements.
- Pour tout évènement B, on a
  $\mathbb{P}[B] = \sum_{i=1}^{n} \mathbb{P}[B \cap A_i] = \sum_{i=1}^{n} \mathbb{P}[B|A_i]\mathbb{P}[A_i]$
  et en particulier
  $\mathbb{P}[B] = \mathbb{P}[B \cap A] + \mathbb{P}[B \cap \overline{A}] = \mathbb{P}[B|A]\mathbb{P}[A] + \mathbb{P}[B|\overline{A}]\mathbb{P}[\overline{A}]$

Quelques formules utiles

- **Formule de Bayes** : lorsque A et B sont deux évènements de probabilité non-nulle, on peut écrire :

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$$

- **Formule des probabilités totales** : Soit $\{A_1, \ldots, A_n\}$ un système complet d'évènements. Pour $k = 1, \ldots, n$, on peut écrire :

$$\mathbb{P}(A_k|B) = \frac{\mathbb{P}(B|A_k)\mathbb{P}(A_k)}{\sum_{j=1}^{n} \mathbb{P}(B|A_j)\mathbb{P}(A_j)}$$

ce qui dans le cas d'un système complet d'évènements donné par $\{A, \overline{A}\}$ revient à :

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|\overline{A})\mathbb{P}(\overline{A})}$$

Exercice : tubes d'aluminium

Un échantillon de 100 tubes d'aluminium est prélevé dans la production de l'usine et chaque tube est classé en fonction de sa longueur (L) et de sa qualité de surface (QS). Chacune de ces caractéristiques peut être considérée comme "conforme" ou non "conforme". Les résultats suivants sont obtenus :

	L conforme	L non-conforme
QS conforme	75	7
QS non-conforme	10	8

Soit A l'évènement "le tube a une qualité de surface conforme" et soit B l'évènement "le tube est de longueur conforme".

- Déterminer les probabilités suivantes : $\mathbb{P}(A)$, $\mathbb{P}(B)$, $\mathbb{P}(\overline{A})$, $\mathbb{P}(A \cap B)$, $\mathbb{P}(A \cup B)$, $\mathbb{P}(\overline{A} \cup B)$, $\mathbb{P}(A|B)$ et $\mathbb{P}(B|A)$.

Exercice : chasse au canard

Trois amis vont à la chasse. Le premier est bon tireur et a la probabilité $2/3$ d'atteindre sa cible. Les deux autres sont moins performants et ont la probabilité $1/6$ d'atteindre leur cible. Un canard s'envole : les trois amis tirent dessus.

- Quelles sont les chances de survie du canard ?

Exercice : grippe et symptômes

Durant l'hiver, la probabilité pour qu'une personne ait la grippe est estimée à 30%. Le diagnostique clinique est posé lorsque la personne présente les symptômes suivants : courbatures, fièvre subite, signes respiratoires. Durant l'hiver, la probabilité pour qu'une personne présente ces symptômes est estimée à 40%. On sait aussi qu'une personne ayant la grippe a 80 chances sur 100 d'avoir ces symptômes.

- Quelle est la probabilité d'avoir la grippe et de présenter les symptômes décrits ci-dessus ?
- Quelle est la probabilité d'avoir la grippe sachant qu'on présente les symptômes ci-dessus ?

Exercice : tirage de 2 pièces parmi une production

L'ingénieur d'usine de l'entreprise Product a noté, en se basant sur une évaluation de plusieurs années, que 2% de la production de l'usine est défectueuse. L'ingénieur tire successivement deux pièces dans la production. On suppose que le fait d'observer une pièce défectueuse n'influe pas sur la qualité de l'autre pièce.

1. Quelle est la probabilité que les 2 pièces tirées par l'ingénieur soient défectueuses ?
2. Quelle est la probabilité que l'une des deux pièces soit bonne et l'autre défectueuse ?

Exercice : bris d'un dispositif électronique et arrêt d'une chaine d'empaquatage

L'ingénieur d'usine de l'entreprise Electropak a noté, en se basant sur une évaluation de plusieurs années, qu'un dispositif électronique installé sur une chaine d'empaquetage a une probabilité de 20% de tomber en panne. Lorsque ce dispositif tombe en panne, la probabilité d'être obligé d'arrêter complètement la chaine d'empaquetage (à cause d'un bris trop important) est de 50%.

- Quelle est la probabilité d'observer que le dispositif tombe en panne et que la chaine d'empaquetage soit complètement arrêtée ?

Exercice : kermesse

Pour une kermesse d'école, un stand propose le jeu suivant. Le joueur tire une carte dans un jeu comportant 32 cartes. S'il obtient une figure (i.e. un valet, une dame ou un roi), il tire un billet dans la corbeille "Super Chance" qui contient 20 billets gagnants et 30 billets perdants. Si le joueur n'obtient pas de figure, il tire un billet dans la corbeille "Petite Chance" qui contient 10 billets gagnants et 40 billets perdants.

- Quelle est la probabilité pour un joueur de tirer un billet gagnant ?

Exercice : orthographe anglophone

Les anglais et les américans orthographient le mot “rigueur” respectivement “rigour” et “rigor”. Un homme ayant pris une chambre dans un hôtel parisien a écrit ce mot sur un bout de papier. Une lettre est prise au hasard dans ce mot et c’est une voyelle. Or, 40% des anglophones de l’hôtel sont des Anglais et 60% des américains.

- Quelle est la probabilité que l’auteur du mot soit anglais ?

Exercice : alcootest

Une entreprise commercialisant un alcootest décide d'en vérifier la fiabilité. Les chiffres sont les suivants :

- 25% des personnes contrôlées par la police sont effectivement en état d'ébriété
- 95 fois sur 100, l'alcootest s'est révélé positif alors que la personne était réellement en état d'ébriété
- 1 fois sur 100, l'alcootest s'est révélé positif alors que la personne n'était pas en état d'ébriété

1. Quelle est la probabilité que l'alcootest donne une indication correcte ?
2. Quelle est la probabilité qu'une personne soit réellement en état d'ébriété lorsque l'alcootest est positif ?

Exercice : tirage de boules

Une urne étiquetée A contient 3 boules blanches, 2 boules noires et 5 boules rouges. Une urne étiquetée B contient 5 boules blanches, 3 boules noires et 2 boules rouges. On effectue l'hypothèse d'équiprobabilité à la fois en ce qui concerne le tirage des urnes et le tirage des boules.

1. On tire une boule au hasard dans une des urnes elle-même tirée au hasard. Quelle est la probabilité que la boule tirée soit rouge ?
2. On tire une boule au hasard dans une des urnes elle-même tirée au hasard. La boule tirée étant rouge, quelle est la probabilité que celle-ci provienne de l'urne A ?

Exercice : puces électroniques contaminées

Lors d'un procédé de fabrication de semi-conducteurs, les puces qui ont été soumises à une contamination due à la présence de poussières tombent en panne avec une probabilité 0.1 tandis que les puces qui n'ont pas été soumises à une contamination tombent en panne avec une probabilité 0.005. Pendant une séquence particulière de production, la probabilité que les puces soient soumises à une contamination est 0.2.

1. Calculer la probabilité qu'une de ces puces tombent en panne.
2. Sachant qu'une puce est tombée en panne, calculer la probabilité que celle-ci ait été contaminée lors de la production.

Exercice : microprocesseurs défectueux

On suppose que trois types de microprocesseurs utilisés dans la fabrication d'ordinateurs se partagent le marché à raison de 25% pour le type X, 35% pour le type Y et 40% pour le type Z. Les pourcentages de défauts de fabrication sont : 5% pour les microprocesseurs de type X, 4% pour ceux de type Y et 2% pour ceux de type Z. Dans un lot constitué de microprocesseurs dans les proportions indiquées pour les types X, Y et Z, on prélève un microprocesseur.

1. Quelle est la probabilité qu'il soit défectueux ?
2. Sachant que le microprocesseur présente un défaut de fabrication, quelle est la probabilité qu'il soit de type X ?

Définition d'une variable aléatoire

- Une **variable aléatoire** X est le procédé qui relie l'expérience aléatoire à un nombre. On note D_X l'ensemble des valeurs que X peut prendre après réalisation de l'expérience : D_X s'appelle le domaine de définition de X.
- A chaque fois que l'on reproduit l'expérience, on obtient une réalisation de X que l'on note x : x est un nombre alors que X est une fonction ! ! !
- Soit l'expérience "tirer une pièce parmi une production" et soit X la variable aléatoire représentant la longueur de la pièce tirée. L'ingénieur d'usine effectue une 1ère fois cette expérience, il obtient la réalisation $x_1 = 10.2$cm. Il recommence une 2ème fois l'expérience et obtient la réalisation $x_2 = 9.9$cm, etc...
- Soit l'expérience "jeter un dé" et soit X la variable aléatoire représentant la valeur inscrite sur la face supérieure. Un joueur effectue une 1ère fois cette expérience, il obtient la réalisation $x_1 = 4$. Il recommence une 2ème fois l'expérience et obtient la réalisation $x_2 = 3$, etc...

Les différents types de variables

On distingue :

- les variables aléatoires **continues** : toute valeur d'un intervalle de $\mathbb{R}$ est acceptable, ex : taille, poids, volume, temps écoulé...
- les variables aléatoires **discrètes** : elles prennent un nombre dénombrable (fini ou infini) de valeurs, par ex : nombre de pièces défectueuses dans la production journalière d'une usine, nombre de clients arrivant à un guichet en une journée, variable aléatoire binaire codant pour “succès” ou “échec”...

Exemples de variables aléatoires et d'évènements associés

- A l'usine, on dispose d'un lot de 30 pièces prélevées dans la production sur lesquelles on effectue un contrôle de qualité à l'issue duquel on déclare les pièces conformes ou non-conformes. Soit X la variable aléatoire qui compte le nombre de pièces non-conformes.
  - L'ensemble des valeurs possibles pour X est $D_X = \{0, 1, ..., 30\}$.
  - L'évènement "2 pièces sont non-conformes" se note $\{X = 2\}$.
  - $\{X = 50\} = \emptyset$
  - $\{8.5 \leq X \leq 10.5\}$ code pour l'évènement "9 ou 10 pièces sont non-conformes".
- On s'intéresse au poids des pièces qui peut varier de 10g à 15g. Soit X la variable aléatoire représentant le poids (en g) d'une pièce.
  - L'ensemble des valeurs acceptables pour X est $D_X = [10, 15]$.
  - $\{X = 12\} =$ le poids d'une pièce est de 12g.
  - $\{X = 50\} = \emptyset$
  - $\{8.5 \leq X \leq 10.5\} =$ le poids d'une pièce est compris entre 8.5g et 10.5g.

Distribution de probabilité et fonction de répartition

- La **loi de probabilité** aussi appellée **distribution de probabilité** et notée f_X d'une variable aléatoire X a pour but de décrire quelles sont les valeurs possibles prises par la variable et avec quelle probabilité ces différentes valeurs sont prises.
- Une variable aléatoire est entièrement caractérisée par sa distribution de probabilité f_X ou de manière équivalente par sa **fonction de répartition** notée F_X.
- La **théorie des probabilités** vise à évaluer le comportement des variables aléatoires (espérance, variance, probabilités de dépassement d'un seuil, comportement de sommes,...) étant donné la distribution de probabilité F_X.
- La **statistique** fournit des méthodes pour résoudre le problème inverse dit d'inférence statistique : caractériser F_X au vu des observations des variables.

Fonction de répartition d'une variable aléatoire

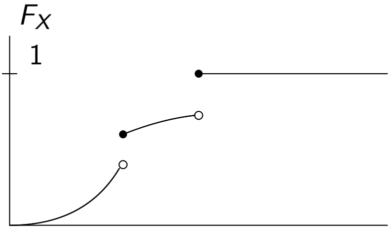
- Soit X une variable aléatoire et soit x un nombre.
- Considérons l'évènement $\{X \leq x\}$ = ensemble des résultats d'expérience dont le codage est inférieur ou égal à x.
- $\mathbb{P}[X \leq x]$ est un nombre qui dépend de la valeur de x
- On définit F_X la fonction de répartition de X par

$$F_X(x) = \mathbb{P}[X \leq x].$$

- Pour tout x, on a $0 \leq F_X(x) \leq 1$ avec $\lim_{x \to +\infty} F_X(x) = 1$ et $\lim_{x \to -\infty} F_X(x) = 0$.
- On note $F_X(x^-) = \mathbb{P}[X < x] = \lim_{y \to x, y < x} F_X(y)$.
- $\mathbb{P}[X > x] = 1 - \mathbb{P}[X \leq x]$ i.e. $\mathbb{P}[X > x] = 1 - F_X(x)$
- $\mathbb{P}[X \leq x] - \mathbb{P}[X < x] = \mathbb{P}[X = x]$ i.e. $F_X(x) - F_X(x^-) = \mathbb{P}[X = x]$
- Pour $a < b$, on a :
  - $F_X(b) - F_X(a) = \mathbb{P}[a < X \leq b]$
  - $F_X(b) - F_X(a^-) = \mathbb{P}[a \leq X \leq b]$
  - $F_X(b^-) - F_X(a) = \mathbb{P}[a < X < b]$
  - $F_X(b^-) - F_X(a^-) = \mathbb{P}[a \leq X < b]$

Quelques propriétés de la fonction de répartition

- F_X est croissante ($x \leq y \Rightarrow F_X(x) \leq F_X(y)$).
- F_X est continue à droite ce qui signifie que F_X est continue sauf éventuellement en un nombre dénombrable de points isolés $(a_i)_{i=1,..}$ en lesquels $F_X(a_i) \neq F_X(a_i^-)$.
- Si F_X est discontinue en un point b alors on a $\mathbb{P}[X = b] = F_X(b) - F_X(b^-) > 0$.
- Si F_X est continue en un point a alors on a $F_X(a) = F_X(a^-)$ et $\mathbb{P}[X = a] = 0$.



Fonction de répartition d'une variable discrète

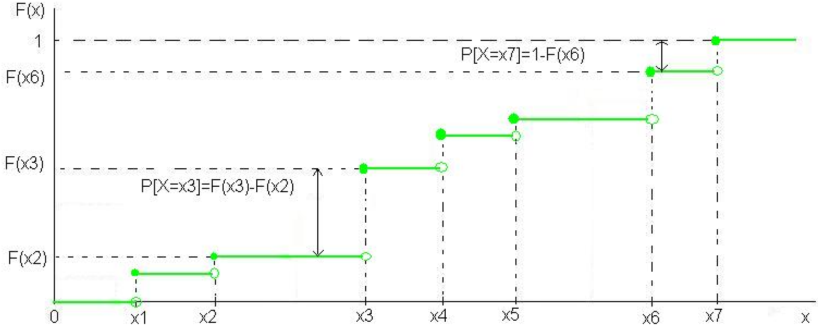
- La fonction de répartition d'une variable discrète au point x correspond à l'accumulation des probabilités des valeurs inférieures ou égales à x :

$$F_X(x) = \sum_{x_i \leq x, x_i \in D_X} \mathbb{P}[X = x_i]$$

- Ainsi, F_X est une fonction en escalier, continue à droite.
- Pour $a < b$, on a :

$$\begin{aligned}
F_X(b) - F_X(a) &= \sum_{a < x_i \leq b, x_i \in D_X} \mathbb{P}[X = x_i] \\
F_X(b) - F_X(a^-) &= \sum_{a \leq x_i \leq b, x_i \in D_X} \mathbb{P}[X = x_i] \\
F_X(b^-) - F_X(a) &= \sum_{a < x_i < b, x_i \in D_X} \mathbb{P}[X = x_i] \\
F_X(b^-) - F_X(a^-) &= \sum_{a \leq x_i < b, x_i \in D_X} \mathbb{P}[X = x_i]
\end{aligned}$$

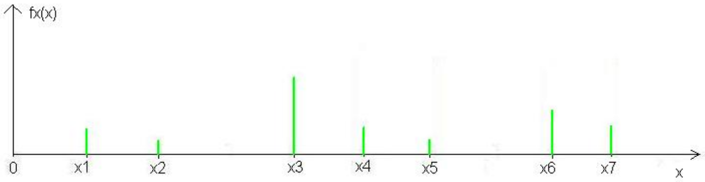
Allure de la fonction de répartition d'une variable discrète



Distribution de probabilité d'une variable discrète

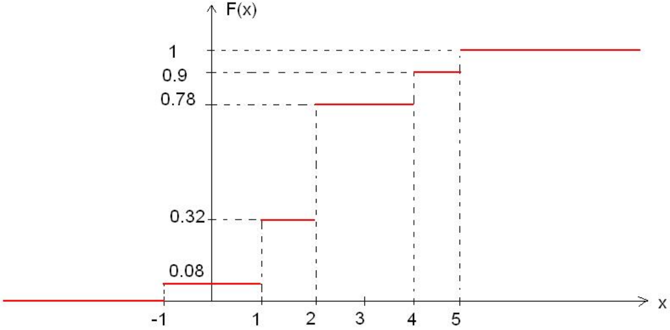
- La distribution de probabilité d'une variable aléatoire discrète est la donnée des nombres $p_i = \mathbb{P}[X = x_i]$ pour chacune des valeurs possibles x_i pour X. Ces nombres sont compris entre 0 et 1 et leur somme vaut 1. Cela revient à définir une fonction f_X telle que

$$f_X(x) = \begin{cases} \mathbb{P}[X = x] & \text{si } x \in D_X \\ 0 & \text{si } x \notin D_X \end{cases}$$



Exercice : variable discrète (1)

Soit X une variable de fonction de répartition F tracée ci-dessous. Quel est l'ensemble des valeurs prises par X ? Quelle est la loi de probabilité de X ? Calculer $\mathbb{P}[X \leq -2]$, $\mathbb{P}[X > 6]$, $\mathbb{P}[X \leq 3]$, $\mathbb{P}[X \leq 4]$, $\mathbb{P}[X < 2]$ et $\mathbb{P}[1 \leq X \leq 4]$. Déterminer les a vérifiant $\mathbb{P}[X \leq a] \geq 1/4$ puis $\mathbb{P}[X \geq a] \geq 1/4$.



Exercice : variable discrète (2)

Une personne lance simultanément 2 dés. Soit X la variable aléatoire représentant la somme des valeurs obtenues sur la face supérieure de chacun des dés.

1. Quel est le domaine de définition D_X de X ?
2. Déterminer puis tracer la distribution de probabilité f_X de X.
3. Déterminer puis tracer la fonction de répartition F_X de X.
4. Calculer les probabilités suivantes : $\mathbb{P}[X \leq 7]$, $\mathbb{P}[X \geq 5]$ et $\mathbb{P}[7 \leq X \leq 9]$.
5. Déterminer les a vérifiant $\mathbb{P}[X \leq a] \leq 1/6$ puis $\mathbb{P}[X \geq a] \leq 1/6$.

Exercice : variable discrète (3)

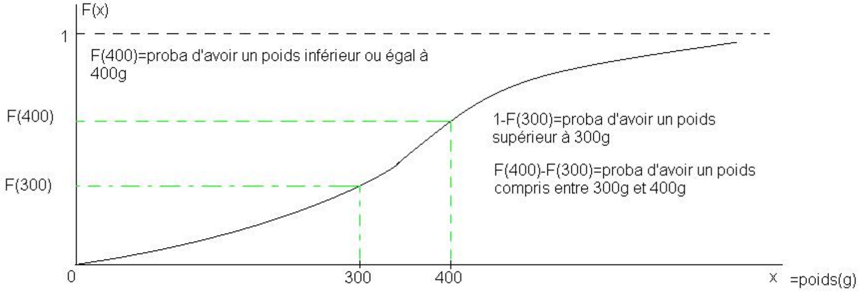
Soit X une variable aléatoire de domaine de définition $D_X = \{0, 1, 2, 3, 4\}$ et dont la distribution de probabilité est donnée pour $x \in D_X$ par

$$f_X(x) = \frac{2x+1}{25}$$

1. Déterminer puis tracer la fonction de répartition F_X de X.
2. Calculer les probabilités suivantes : $\mathbb{P}[X = 4]$, $\mathbb{P}[X \geq -2]$, $\mathbb{P}[X \leq 1]$ et $\mathbb{P}[2 \leq X < 4]$.
3. Déterminer les a vérifiant $\mathbb{P}[X \leq a] \leq 1/5$ puis $\mathbb{P}[X \geq a] \leq 1/5$.

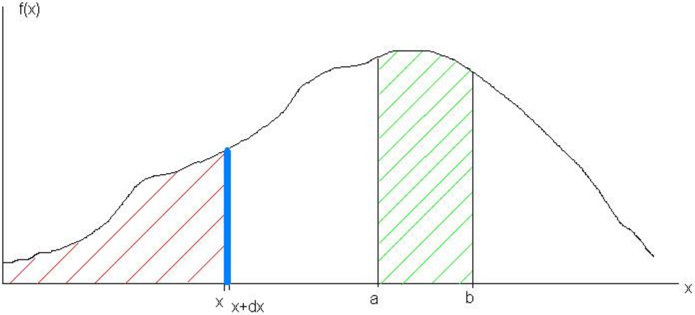
Fonction de répartition d'une variable continue

- La fonction de répartition d'une variable aléatoire continue est une fonction continue, dérivable presque-partout.
- En tout point x, on a $\mathbb{P}[X = x] = 0$ et $\mathbb{P}[X < x] = \mathbb{P}[X \leq x]$.
- Pour tout $a < b$, on a :
  $\mathbb{P}[a \leq X \leq b] = \mathbb{P}[a \leq X < b] = \mathbb{P}[a < X \leq b] = \mathbb{P}[a < X < b]$



Rappels sur le calcul des intégrales

- $f(x)dx$=surface bleue
- $\int_{-\infty}^{x} f(t)dt = F(x)$=surface en rouge : F=primitive de f
- $\int_a^b f(x)dx$=$F(b) - F(a)$=surface verte $(=\int_a^b f(t)dt)$
- $f(x) = \frac{dF(x)}{dx}$=dérivée de F en x= pente de F en x



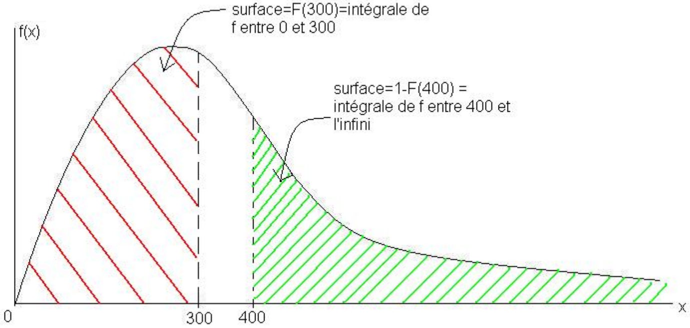
Distribution de probabilité d'une variable aléatoire continue

- La distribution de probabilité d'une variable aléatoire continue (aussi appelée **densité**) est donnée par :

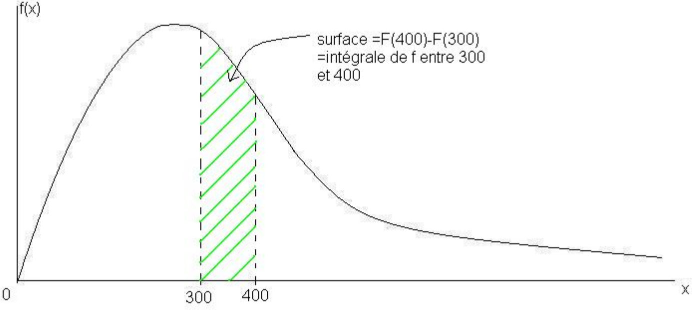
$$f_X(x) = \begin{cases} \frac{dF_X(x)}{dx} & \text{pour } x \in D_X \text{ tel que } F_X \text{ est dérivable} \\ 0 & \text{pour } x \notin D_X \text{ et lorsque } F_X \text{ n'est pas dérivable} \end{cases}$$

- $f_X(x)dx = \mathbb{P}[x \leq X \leq x + dx] \approx \mathbb{P}[X = x]$
- $F_X(x) = \int_{-\infty}^{x} f_X(u)du$
- f_X est positive (car $F_X \nearrow$) et $\int_{-\infty}^{\infty} f_X(u)du = 1$
- $\mathbb{P}[a < X \leq b] = F_X(b) - F_X(a) = \int_a^b f_X(u)du$

Densité d'une variable aléatoire continue (suite)



Densité d'une variable aléatoire continue (suite 2)



Exercice : variable continue (1)

Soit une variable aléatoire X définie sur $[-1, 1]$ régie par la densité suivante : $f(x) = 3(1 - x^2)/4$.

1. Quel est le domaine de définition D_X de X ?
2. Tracer la distribution de probabilité f_X de X.
3. Déterminer puis tracer la fonction de répartition F_X de X.
4. Calculer les probabilités suivantes : $\mathbb{P}[X = 0.5]$, $\mathbb{P}[X \leq 0.5]$, $\mathbb{P}[X > 0.2]$ et $\mathbb{P}[0.2 \leq X < 0.5]$.

Exercice : variable continue (2)

Soit une variable aléatoire X régie par la densité suivante :

$$f(x) = \begin{cases} \frac{x}{k} \quad \text{pour } 0 \leq x \leq 10 \\ \frac{20-x}{k} \quad \text{pour } 10 < x \leq 20 \\ 0 \quad \text{ailleurs} \end{cases}$$

1. Déterminer la valeur de la constante k pour s'assurer que la fonction f soit bien une densité de probabilité.
2. Tracer la densité de probabilité obtenue.
3. Déterminer la fonction de répartition.
4. Déterminer $\mathbb{P}[X \leq 0]$, $\mathbb{P}[1 \leq X \leq 2]$, $\mathbb{P}[15 \leq X \leq 18]$, $\mathbb{P}[5 \leq X \leq 15]$, $\mathbb{P}[X \geq 19]$, $\mathbb{P}[X \geq 9]$ et $\mathbb{P}[X \geq 29]$.
5. Déterminer a tel que $\mathbb{P}[X \leq a]$=1/4 puis $\mathbb{P}[X \leq a]$=3/4.

Variables aléatoires indépendantes

- Deux variable X_1 et X_2 sont **indépendantes** lorsque le fait de connaître la valeur obtenue par X_1 n'apporte aucune information sur la valeur qui sera prise par X_2 et réciproquement.
- ex : X_1="poids d'une souris" et X_2="couleur du pelage" sont indépendantes alors que X_1 et Y_1="taille d'une souris" ne le sont vraisemblablement pas.
- Caractérisation de l'indépendance pour un couple de variables **discrètes** : X et Y sont indépendants lorsque pour tout couple de valeurs (x_i, y_j) pris par (X, Y), on a

$$\mathbb{P}[X = x_i, Y = y_j] = \mathbb{P}[X = x_i]\mathbb{P}[Y = y_j]$$

- Caractérisation de l'indépendance pour un couple de variables **continues** : X et Y sont indépendants lorsqu'on a

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) \qquad \text{pour tout } (x, y)$$

  où $f_{X,Y}$ représente la densité jointe du couple (X, Y), f_X représente la densité de X et f_Y représente la densité de Y.

Caractéristiques de position et de dispersion

- Caractérisques de **position** :
  - l'**espérance** est un nombre qui représente la valeur moyenne prise par X.
  - la **médiane** (lorqu'elle est définie) est la valeur telle que X a autant de chance de se réaliser au-dessus qu'en dessous.
  - un **mode** correspond à une valeur ayant une probabilité maximale de se réaliser (il existe des distributions n'ayant aucun mode).
  - moyenne=médiane=mode pour les distributions symétriques i.e. telles que les valeurs prises par X sont également réparties autour d'une valeur centrale.
- Caractéristiques de **dispersion** :
  - La **variance** exprime à quel point les valeurs prises par X sont dispersées autour de la moyenne. Une grande variance indique une grande dispersion. A l'inverse, une variance nulle révèle que X est en fait non-aléatoire.
  - L'**écart-type** fournit la même information.
  - Les **quantiles** (lorqu'ils sont définis) permettent de fournir l'intervalle dans lequel X se réalise avec 95% de chances par ex.

Espérance mathématique

- Soit X une variable aléatoire. On note $\mathbb{E}[X]$ l'**espérance** de X. C'est un nombre qui représente la valeur moyenne prise par X.
- Si X est discrète, on calcule $\mathbb{E}[X]$ par la formule :

$$\mathbb{E}[X] = \sum_{x_i \in D_X} x_i \mathbb{P}[X = x_i].$$

- Si X est continue, on calcule $\mathbb{E}[X]$ par la formule :

$$\mathbb{E}[X] = \int x f_X(x) dx.$$

- On dit qu'une variable est **centrée** lorsque $\mathbb{E}[X] = 0$.
- On a toujours $\mathbb{E}[aX] = a\mathbb{E}[X]$, $\mathbb{E}[a + X] = a + \mathbb{E}[X]$ et $\mathbb{E}[X_1 + X_2] = \mathbb{E}[X_1] + \mathbb{E}[X_2]$.
- On calcule $\mathbb{E}[X^2]$ par la formule :
  - $\mathbb{E}[X^2] = \sum_{x_i \in D_X} x_i^2 \mathbb{P}[X = x_i]$ si X est discrète,
  - $\mathbb{E}[X^2] = \int x^2 f_X(x) dx$ si X est continue.

Variance mathématique, écart-type mathématique

- La **variance** d'une variable X est le nombre **positif** défini par :

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

- Si X est **discrète**, on calcule $\mathrm{Var}(X)$ par l'une des 2 formules :

$$\mathrm{Var}(X) = \sum (x_i - \mathbb{E}[X])^2 \mathbb{P}[X = x_i] = \sum x_i^2 \mathbb{P}[X = x_i] - \mathbb{E}[X]^2.$$

- Si X est **continue**, on calcule $\mathrm{Var}(X)$ par l'une des 2 formules :

$$\mathrm{Var}(X) = \int (x - \mathbb{E}[X])^2 f(x) dx = \int x^2 f(x) dx - \mathbb{E}[X]^2.$$

- Une variable aléatoire est dite **réduite** lorsque $\mathrm{Var}(X) = 1$.
- On a toujours $\mathrm{Var}(X + a) = \mathrm{Var}(X)$ et $\mathrm{Var}(aX) = a^2 \mathrm{Var}(X)$ mais la relation $\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2)$ n'est vraie que lorsque X_1 et X_2 sont indépendantes ! ! !
- L'écart-type est défini par : $\sigma(X) = \sqrt{\mathrm{Var}(X)}$.

Exercice : espérance, variance et écart-type

Reprendre les 3 exercices précédents portant sur les variables discrètes et les 2 exercices précédents portant les variables continues. Dans chaque cas, calculer l'espérance, la variance et l'écart-type de la variable aléatoire considérée.

Exercice : nombre d'accidents par jour

Le responsable du comité de sécurité de l'entreprise Micom a évalué le nombre moyen d'accidents de travail par jour à 1.6 avec un écart-type de 1.265 accident/jour. Soit X la variable représentant le nombre d'accidents par jour. Pour maintenir un service d'urgence, l'entreprise subit des frais fixes de 200 euros par jour ainsi que des frais variables de 50 euros par accidents. Notons Y la variable représentant les frais encourus par jour.

1. Exprimer Y en fonction de X.
2. Quels sont en moyenne les frais encourus par jour ?
3. Quel est l'écart-type des frais ?

Médiane, quantiles

- La **médiane** (si elle existe) de la variable X est la valeur u qui vérifie la relation $F_X(u) = 1/2$. Par conséquent, la médiane d'une variable aléatoire X est la valeur u telle que $\mathbb{P}[X \leq u] = \mathbb{P}[X > u] = 1/2$.
- Le **quantile** (ou **fractile**) d'ordre $1 - \alpha$ (si il existe) de la variable X est la valeur u qui vérifie la relation $F_X(u) = 1 - \alpha$. Le nombre α représente le risque que X dépasse la valeur u.
- L'intervalle interpercentile (si il existe) de la variable X est la donnée des valeurs u et v qui vérifient la relation $\mathbb{P}[u \leq X \leq v] = 1 - \alpha$. Le nombre α représente le risque que X sorte de l'intervalle $[u, v]$.
- L'intervalle interquartile (si il existe) de la variable X est la donnée des valeurs u et v qui vérifient la relation $\mathbb{P}[u \leq X \leq v] = 0.5$.

Mode

- On dit que u est un mode de la variable aléatoire X si $f_X(u)$ est un maximum (local ou global).
  - Si X est une variable continue, u doit donc vérifier $f'_X(u) = 0$ et $f''_X(u) < 0$.
  - Si X est une variable discrète définie sur $\{x_1, x_2, ...\}$, alors $u = \max\{f_X(x_1), f_X(x_2), ...\}$.
- Les distributions de probabilité qui n'ont qu'un seul mode sont dites unimodales et celles qui possèdent plusieurs modes sont dites multimodales.
- Il existe des distributions de probabilité qui n'ont aucun mode.

Exercice : mode, médiane, quantile

- Déterminer le mode de la variable X des exercices : variable discrète (1), (2) et (3).
- Déterminer le mode et la médiane de la variable X des exercices : variable continue (1) et (2).
- Déterminer l'intervalle interquartile de la variable X de l'exercice : variable continue (2).

Caractéristiques de forme : skewness, kurtosis

- Le **coefficient d'asymétrie** (skewness) ν_1 est défini par :

$$\nu_1 = \frac{\mathbb{E}[(X - \mathbb{E}[X])^3]}{\mathrm{Var}(X)^{3/2}}$$

- $\nu_1 = 0$ pour les distributions symétriques, $\nu_1 > 0$ lorsque les valeurs prises par X sont très étalées sur la droite, $\nu_1 < 0$ lorsque les valeurs prises par X sont très étalées sur la gauche.
- Le **coefficient d'aplatissement** (kurtosis) ν_2 est défini par :

$$\nu_2 = \frac{\mathbb{E}[(X - \mathbb{E}[X])^4]}{\mathrm{Var}(X)^2}$$

- La valeur de référence est $\nu_2 = 3$, c'est celle pour la loi gaussienne standard (cf plus tard). Pour $\nu_2 > 3$, la courbe est aigue. Pour $\nu_2 < 3$, la courbe est aplatie. La valeur de ν_2 est intéressante seulement pour des distributions peu asymétriques.

Distributions usuelles continues

- Des familles de lois de proba usuelles continues sont
  - lois normales
  - lois exponentielles
- Ces lois sont paramétrées. Cela signifie que la famille de lois donne la forme générale mais qu'à l'intérieur de ces familles chaque loi dépend de un ou plusieurs nombres appelés **paramètres**.

Loi normale

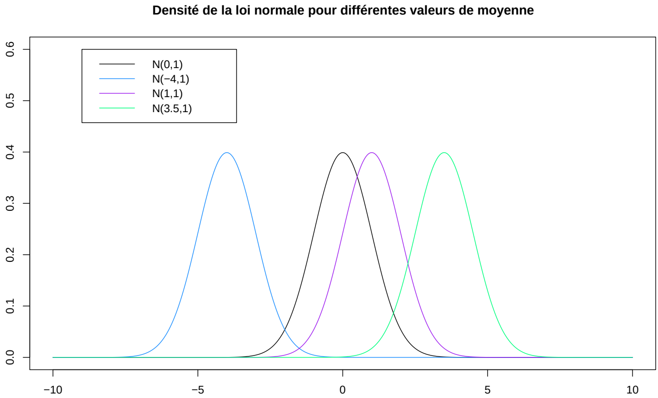
- On dit qu'une variable X suit une loi normale de paramètres m et σ^2, notée $\mathcal{N}(m, \sigma^2)$, lorsqu'elle prend ses valeurs dans $\mathbb{R}$ avec la densité suivante pour $x \in \mathbb{R}$:

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

- La densité est symétrique par rapport à la droite verticale d'absisse $x = m$.
- Une variable X de loi normale $\mathcal{N}(m, \sigma^2)$ représente une variable qui oscille de façon symétrique autour de sa moyenne.
- $\mathbb{E}[X] = m =$ médiane $=$ mode et $\text{var}(X) = \sigma^2$.
- Si $X_1, ..., X_n$ est une suite de variables indépendantes de loi normales telle que $X_i \simeq \mathcal{N}(m_i, \sigma_i^2)$, alors $\sum_{i=1}^n X_i$ suit une loi $\mathcal{N}(m_1 + ...m_n, \sigma_1^2 + ... + \sigma_n^2)$.

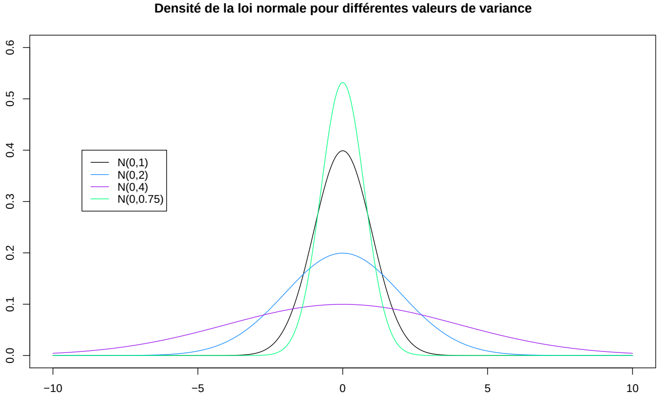
Loi normale : influence de la moyenne

L'allure de la courbe se conserve si on change de moyenne. Il s'agit d'un simple décalage.



Loi normale : influence de la variance

La coube s'aplatit lorsque la variance augmente, elle se resserre si la variance diminue, le maximum s'ajuste pour que la surface vaille 1, le maximum peut dépasser 1.



Loi normale et transformation

- D'une manière générale, si X suit une loi $\mathcal{N}(m, \sigma^2)$, alors $aX + b$ suit une loi $\mathcal{N}(am + b, a^2\sigma^2)$.
- On ne peut déterminer la fonction de répartition de X de loi $\mathcal{N}(m, \sigma^2)$ que par approximations numériques (par ordinateur).
- Un cas important est la loi $\mathcal{N}(0, 1)$ appellée **loi gaussienne standard** ou **loi normale centrée réduite** qui est tabulée.
- Pour se ramener à la loi $\mathcal{N}(0, 1)$ à partir d'une variable X de loi $\mathcal{N}(m, \sigma^2)$, on utilise la variable $Y = \frac{X-m}{\sigma}$: Y suit une loi $\mathcal{N}(0, 1)$ et s'appelle la variable centrée réduite associée à X.

Loi normale et tabulation

- On cherche la valeur de $\mathbb{P}[a \leq X \leq b]$ pour X de loi $\mathcal{N}(m, \sigma^2)$.
- Notons Φ la fonction de répartition associée à la variable Y de loi $\mathcal{N}(0, 1)$.
- Or, seule Φ est tabulée et en plus seulement pour $x > 0$.
- Pour $x < 0$, on utilise la formule $\Phi(x) + \Phi(-x) = 1$.
- Pour $x > 0$, on dispose aussi de la relation $\mathbb{P}[-x \leq Y \leq x] = 2\Phi(x) - 1$.
- Pour la variable X de loi $\mathcal{N}(m, \sigma^2)$, on utilise $Y = (X - m)/\sigma$,la variable centrée réduite associée à X et on obtient :

$$\begin{aligned} \mathbb{P}[a \leq X \leq b] &= \mathbb{P}\left[\frac{a-m}{\sigma} \leq \frac{X-m}{\sigma} \leq \frac{b-m}{\sigma}\right] \\ &= \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right). \end{aligned}$$

Lecture de la table $\mathcal{N}(0,1)$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7793	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8906	0.8925	0.8943	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9986	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Exercice : loi gaussienne standard

Soit X une variable aléatoire de loi gaussienne standard.

1. Déterminer les probabilités suivantes :

$\mathbb{P}[X = 1.2]$	$\mathbb{P}[X \leq 2]$	$\mathbb{P}[0 \leq X \leq 0.5]$
$\mathbb{P}[1.1 \leq X \leq 3.2]$	$\mathbb{P}[X > 0.8]$	$\mathbb{P}[X \leq -1]$
$\mathbb{P}[X > -0.23]$	$\mathbb{P}[0 \leq X \leq 0.83]$	$\mathbb{P}[-2.2 \leq X \leq 2.2]$
$\mathbb{P}[-1 \leq X \leq 0]$	$\mathbb{P}[-2 < X < -0.8]$	$\mathbb{P}[-1.5 \leq X < 0]$
$\mathbb{P}[-1.98 \leq X \leq 0.49]$	$\mathbb{P}[-0.25 \leq X \leq 1.5]$	$\mathbb{P}[X \geq 1.25]$

2. Déterminer x satisfaisant les égalités suivantes :

$\mathbb{P}[X \leq x] = 0.6255$	$\mathbb{P}[X \geq x] = 0.9971$
$\mathbb{P}[X \leq x] = 0.2119$	$\mathbb{P}[X \geq x] = 0.1314$
$\mathbb{P}[0 \leq X \leq x] = 0.4750$	$\mathbb{P}[-x \leq X \leq 0] = 0.2291$
$\mathbb{P}[-x \leq X \leq x] = 0.2052$	$\mathbb{P}[-1 \leq X \leq x] = 0.1785$

Exercice : aliment pour chat

L'entreprise Granulex distribue un aliment pour chat dans un contenant métallique dont le poids après remplissage est en moyenne 340g. Le processus de remplissage est donc calibré pour obtenir cette valeur moyenne. En fait, des études ont montré que le poids est distribué normalement avec un écart-type de 6g.

1. Quelle est la probabilité pour qu'un contenant choisi au hasard de la production ait un poids entre 334g et 346g ?
2. Quelle est la probabilité pour qu'un contenant choisi au hasard de la production ait un poids qui diffère de la moyenne par moins de 2g ?
3. Sur une production de 1000 contenants, combien auront un poids inférieur à 330g ?
4. A quel niveau moyen doit-on fixer le remplissage pour assurer que seulement 1 contenant sur 100 aura un poids inférieur à 340g ?
5. A quel niveau moyen doit-on fixer le remplissage pour assurer que seulement 5% des contenants auront un poids supérieur à 348g ?

Exercice : demande de logiciels

D'après la responsable de la mise sur le marché de l'entreprise Sicom, la demande annuelle pour leur logiciel suit une loi normale. Elle précise également qu'il y a une probabilité 0.195 pour que la demande soit inférieure à 1500 unités et une probabilité 0.025 d'être supérieure à 2910 unités.

1. Déterminer la demande annuelle moyenne ainsi que sa variance.
2. En admettant que la demande est équitablement répartie au cours de l'année et que la demande mensuelle suit une loi normale, déterminer la demande mensuelle moyenne et sa variance.
3. Les coûts fixes mensuels de production sont de 21000 euros alors que le coût unitaire est de 300 euros. Le prix de vente d'un logiciel est de 500 euros. Quelle est la probabilité que le seuil de rentabilité mensuel soit atteint ?

Exercice : système de sécurité

Un système de sécurité opère avec deux composants électroniques. La conception du système est telle que le second composant entre en fonction lorsque le premier devient défaillant. Lorsque le second composant devient défaillant, le système de sécurité devient inopérant. La durée de vie de chaque composant est distribué normalement avec une moyenne de 500h et une variance de $2450h^2$. On admet que les durées de vie de chaque composant sont indépendantes.

1. Quelle est la loi de la durée de vie du système de sécurité ? En préciser la moyenne et la variance.
2. Avec quelle probabilité peut-on affirmer que le système devrait fonctionner au moins 850h ?
3. Quelle est la durée de vie minimum du système dans 95% des cas ?

Lois exponentielles

- On dit qu'une variable X suit une loi exponentielle de paramètre λ, notée $\mathcal{E}(\lambda)$, lorsqu'elle prend ses valeurs dans $\mathbb{R}^+ = [0, +\infty[$ avec la densité et la fonction de répartition suivantes pour $x \in \mathbb{R}$:

$$f(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

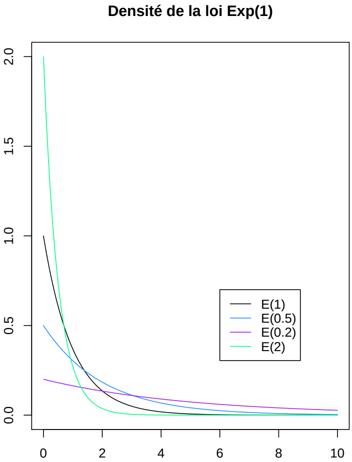
$$F(x) = \begin{cases} 1 - \exp(-\lambda x) & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

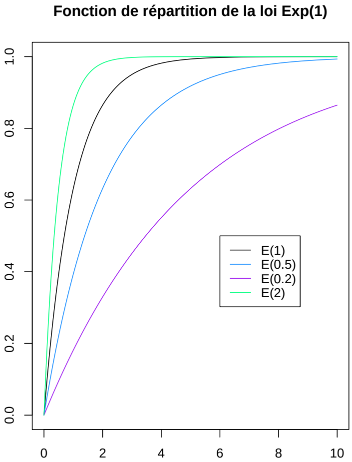
- $\mathbb{E}[X] = \frac{1}{\lambda}$ et $\mathrm{var}(X) = \frac{1}{\lambda^2}$.
- La loi exponentielle est une loi sans mémoire :

$$\mathbb{P}[T > t + s | T > s] = \mathbb{P}[T > t].$$

Elle ne convient donc pas pour modéliser des objets soumis à une usure non-négligable.

Lois exponentielles (suite)





Exercice : durée de vie de pneus

Une étude réalisée sur un grand nombre de pneus de la marque Michel montre que leur durée de vie (en km) est une variable aléatoire X de loi exponentielle de paramètre $\lambda = 0.0007$.

1. Déterminer $\mathbb{P}[X \leq 1000]$, $\mathbb{P}[X > 1000]$ et $\mathbb{P}[1000 \leq X \leq 2000]$
2. Déterminer x tel que $\mathbb{P}[X \leq x] = 0.05$.
3. Déterminer la médiane de X.

Exercice : désintégration d'un atome

Une entreprise du nucléaire s'intéresse à la durée de vie de certains types d'atomes. Cette durée de vie est donnée par le temps qui s'écoule entre l'instant initial $t_0 = 0$ et le moment où le noyau de l'atome se désintègre. On sait que la probabilité pour qu'au temps $t \geq 0$ le noyau ne se soit pas encore désintégré vaut $s(t) = \exp(-\lambda t)$ pour un $\lambda > 0$ qui varie avec le type d'atome. Notons X la variable aléatoire qui représente la durée de vie d'un atome.

1. Quelle est la fonction de répartition de X en fonction de λ ?
2. Quelles sont la densité, l'espérance et la variance de X ?
3. On observe un atome jusqu'à un instant t. Sachant qu'à cet instant t l'atome n'est pas désintégré, quelle est la probabilité qu'il ne se désintègre pas avant l'instant $t + s$ pour un $s > 0$?.

Distributions usuelles discrètes

- Les familles de lois de proba usuelles discrètes sont
  - lois de Bernoulli
  - lois binômiales
  - lois hypergéométriques
  - lois de Poisson
- Ces lois sont paramétrées. Cela signifie que la famille de lois donne la forme générale mais qu'à l'intérieur de ces familles chaque loi dépend de un ou plusieurs nombres appelés **paramètres**.

Lois de Bernoulli

- On dit qu'une variable X suit une loi de Bernoulli de paramètre p, notée $\mathcal{B}(p)$, lorsqu'elle prend les valeurs 0 avec probabilité $(1-p)$ et 1 avec probabilité p : $\mathbb{P}[X=1]=p$ et $\mathbb{P}[X=0]=1-p$.
- $\mathbb{E}[X]=p$ et $\mathrm{var}(X)=p(1-p)$.
- Définir une variable de Bernoulli revient à coder par 1 la réalisation d'un évènement et par 0 sa non-réalisation, autrement dit, $X=1$ si l'évènement est réalisé et $X=0$ sinon.
- La loi de Bernoulli est utilisée pour coder des caractéristiques à 2 modalités selon un schéma succès/échec, par ex : défectueux/non-défectueux, conforme/non-conforme, ce qui donne : $X=1$ si la pièce est conforme et $X=0$ sinon. Un autre exemple est $X=1$ si X est inférieur à un seuil et $X=0$ si X dépasse ce seuil.

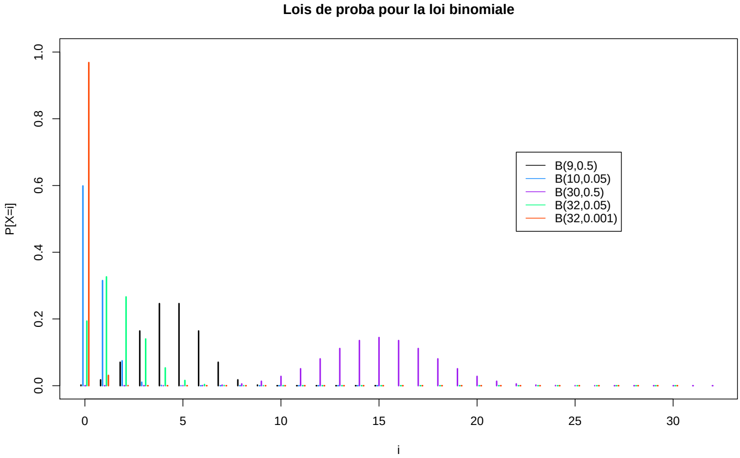
Lois binômiales

- On dit qu'une variable X suit une loi binômiale de paramètres n et p, notée $\mathcal{B}(n, p)$, lorsqu'elle prend ses valeurs parmi $\{0, ..., n\}$ avec les probabilités suivantes pour $k \in \{0, ..., n\}$:

$$\mathbb{P}[X = k] = C_n^k p^k (1 - p)^{n-k}.$$

- $\mathbb{E}[X] = np$ et $\mathrm{var}(X) = np(1 - p)$.
- X représente le nombre de fois où un évènement se réalise en n répétitions indépendantes de l'expérience. Par ex, lorsque qu'on prélève au hasard n factures de façon indépendante, X représente le nombre de factures erronées. La loi binômiale sert aussi à compter le nombre de succès lors d'un tirage avec remise de n éléments.
- Si $Y_1, ..., Y_n$ est une suite de variable indépendantes de loi de Bernoulli de même paramètre p, $\mathcal{B}(p)$, alors $\sum_{i=1}^n Y_i$ suit une loi $\mathcal{B}(n, p)$.
- Si X_1 et X_2 sont deux variables indépendantes de lois respectives $\mathcal{B}(n_1, p)$ et $\mathcal{B}(n_2, p)$ (lois de même paramètre p !), alors $X_1 + X_2$ suit une loi $\mathcal{B}(n_1 + n_2, p)$.

Lois binômiales



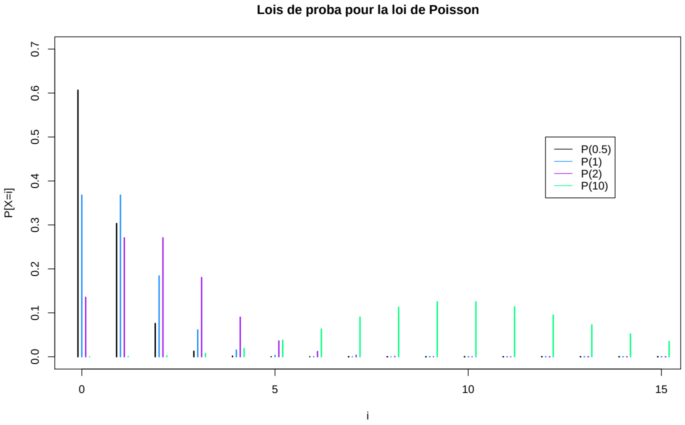
Lois de Poisson

- On dit qu'une variable X suit une loi de Poisson de paramètre λ, $\mathcal{P}(\lambda)$, lorsqu'elle prend ses valeurs dans $\mathbb{N} = \{0, 1, ..., n, ...\}$ (infinité de valeurs possibles) avec les probabilités suivantes pour $k \in \mathbb{N}$:

$$\mathbb{P}[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}.$$

- $\mathbb{E}[X] = \lambda$ et $\mathrm{var}(X) = \lambda$.
- X représente le résultat d'un comptage effectué sur une durée fixée, par ex : X compte le nombre d'arrivées de clients à un guichet de banque en une semaine, le nombre de véhicules passant par un péage donné en une journée, le nombre de pannes en un an...
- Si $X_1, ..., X_n$ est une suite de variables indépendantes de loi de Poisson telle que $X_i \sim \mathcal{P}(\lambda_i)$, alors $\sum_{i=1}^{n} X_i$ suit une loi $\mathcal{P}(\lambda_1 + ...\lambda_n)$.

Lois de Poisson (suite)



Approximations de lois

- Pour $n = 5, 10, 15, 20, 30, 40, 50$ et $p = 0.01, 0.02, 0.03, 0.04, 0.05$,0.06,0.07,0.08,0.09,0.1,0.2,0.3,0.4,0.5, la loi binômiale est tabulée.
- Pour n et p satisfaisant $np > 5$ et $n(1-p) > 5$, on approxime la loi binômiale $\mathcal{B}(n,p)$ par la loi normale $\mathcal{N}(np, np(1-p))$.
- Pour n grand ($n \geq 30$) et p petit ($p < 0.1$) et $np \leq 5$, on approxime la loi binômiale $\mathcal{B}(n,p)$ par la loi de Poisson $\mathcal{P}(np)$.
- Pour n grand ($n \geq 30$) et p petit ($p < 0.1$) et $np \geq 10$, on approxime la loi binômiale $\mathcal{B}(n,p)$ par la loi normale $\mathcal{N}(np, np)$.
- Pour λ grand ($\lambda \geq 10$), on approxime la loi de Poisson $\mathcal{P}(\lambda)$ par la loi normale $\mathcal{N}(\lambda, \lambda)$.

Lois hypergéométriques

- Une variable X suit une loi hypergéométrique de paramètres N, n et p, notée $\mathcal{H}(N, n, p)$, lorsqu'elle prend ses valeurs parmi $\{\max(0, n - N(1-p)), \dots, \min(Np, n)\}$ avec les probabilités suivantes pour $k \in \{\max(0, n - N(1-p)), \dots, \min(Np, n)\}$:

$$\mathbb{P}[X = k] = \frac{C_{Np}^{k} C_{N(1-p)}^{n-k}}{C_N^n}.$$

- $\mathbb{E}[X] = np$ et $\mathrm{var}(X) = np(1-p)\frac{N-n}{N-1}$.
- X représente le nombre de fois où un évènement se réalise en n tirages sans remise parmi N éléments, par ex : X représente le nombre de boulons non-conformes obtenus au cours d'un tirage sans remise de n boulons parmi N.
- En pratique, on n'hésite pas à substituer la loi $\mathcal{B}(n, p)$ à la loi $\mathcal{H}(N, n, p)$ dès que $n/N \leq 0.1$ (en fait, lorsque $N \to \infty$, les tirages avec remise ou sans remise sont pratiquement équivalents).

Exercice : contamination VHC

Dans un laboratoire d'analyses médicales, 12 personnes se présentent pour effectuer un test de dépistage du VHC. Sachant que 1% de la population française est porteuse du VHC, quelle est la probabilité qu'aucune des 12 personnes ne soit contaminée ? que la moitié des 12 personnes soit contaminée ?

Exercice : contrôle de qualité

Un client commande à son fournisseur un lot de pièces dont la qualité est spécifiée par contrat : le taux de pièces défectueuses doit être inférieur ou égal à 4%. Avant de livrer, le fournisseur effectue un contrôle sur 50 pièces et compte le nombre de pièces défectueuses. Quelle proportion de pièces defectueuses dans sa production ne doit-il pas dépasser pour livrer la marchandise avec un risque de 5% qu'elle soit refusée avec ce plan de contrôle ?

Exercice : comportement du consommateur

D'après une étude sur le comportement du consommateur, il apparaît que 3 consommateurs sur 6 sont influencés par la marque de commerce lors de l'achat d'un bien. La directrice du marketing d'un grand magasin interroge 20 consommateurs au hasard afin de connaître leur réaction sur ce sujet.

1. Identifier la variable concernée et donner les valeurs possibles de cette variable
2. Donner l'expression générale de la loi de probabilité régissant le comportement du consommateur sur ce sujet.
3. Parmi l'échantillon de taille 20, combien de consommateurs, en moyenne, se déclareront influencés par la marque ?
4. Parmi l'échantillon de taille 20, quelle est la probabilité que moins de 10 consommateurs se déclarent influencés par la marque ?
5. Parmi l'échantillon de taille 20, quel est le nombre de consommateurs le plus probable qui se déclareront influencés par la marque ?

Exercice : loi binômiale

Soit X_1 une variable de loi $\mathcal{B}(4, 0.1)$ et soit X_2 une autre variable de loi $\mathcal{B}(6, 0.1)$. On suppose que X_1 et X_2 sont indépendantes.

1. Quelle est la loi de probabilité qui régit la somme $X_1 + X_2$?
2. Quelles sont la moyenne et la variance de la somme $X_1 + X_2$?
3. Déterminer $\mathbb{P}[X_1 + X_2 = 2]$.
4. Déterminer $\mathbb{P}[X_1 + X_2 \geq 8]$.
5. Déterminer $\mathbb{P}[X_2 = 4 | X_1 = 3]$.

Exercice : location de voitures de luxe

Une société de location de voitures possède entre autres trois voitures de luxe qui peuvent être louées à la journée. Le nombre de demandes reçues par la société pour ce genre de voitures est distribué suivant une loi de Poisson avec une moyenne de 1.8 voitures par jour.

1. Quelle est l'expression de la loi de probabilité régissant la demande et quelles sont les valeurs possibles de la variable ?
2. Déterminer la proportion de jours où aucune demande n'est faite pour ce genre de voitures.
3. Calculer la proportion de jours pour lesquels les demandes de location ne peuvent être entièrement satisfaites.

Exercice : vente de frigidaires

Les ventes journalières de frigidaires suivent une loi de Poisson de moyenne 8.8 unités par jour.

1. Quelle est l'expression de la loi de probabilité régissant les ventes journalières et quelles sont les valeurs possibles de la variable aléatoire associée ?
2. Quelle est la variance de la variable ?
3. Quelle est la probabilité de n'observer aucune vente de ce bien sur un jour donné ?
4. Quelle est la proportion de jours pour lesquels les ventes journalières sont inférieures à 5 unités ?
5. Quel est le nombre le plus probable d'unités vendues en un jour donné ?
6. Déterminer le nombre de jours où les ventes journalières ont été exactement de 10 unités et ceci pour une période de 250 jours ouvrables ?

Exercice : nombre d'accidents de travail par jour

Le responsable du comité de sécurité de l'entreprise Nicom a effectué une compilation du nombre d'accidents du travail qui se sont produits depuis deux ans dans l'usine. Ceci a permis d'établir que le nombre moyen d'accidents par jour est 1.6. On admet que le nombre d'accidents du travail par jour suit une loi de Poisson.

1. Quelle est l'expression permettant de calculer la probabilité d'observer x accidents par jour ?
2. Quel est l'écart-type de la variable concernée ?
3. Quelle est la probabilité d'observer plus de 2 accidents par jour ?
4. Calculer la probabilité d'observer un nombre d'accidents compris dans l'intervalle $[\mathbb{E}[X] - \sigma(X), \mathbb{E}[X] + \sigma(X)]$?

Exercice : nombre d'accidents variable

Une assurance s'intéresse au nombre d'accidents touchant un individu au cours d'une année donnée. On admet que ce nombre suit une loi de Poisson. On suppose que l'espérance de cette loi de Poisson varie en fonction des personnes et qu'elle vaut 2 pour 60% des personnes et 3 pour les 40% restants.

1. Quelle est la probabilité qu'au cours d'une année une personne n'ait aucun accident ? Qu'elle en ait trois ?
2. Sachant qu'une personne n'a pas eu d'accident au cours de l'année précédente, quelle est la probabilité qu'elle ait trois accidents dans l'année en cours ?

Exercice : défaillance chez Electropak

Chez Electropak, l'appareil servant à l'étiquetage de contenants est sujet à deux types de pannes, soit une défaillance électronique, soit une défaillance mécanique. Les deux sources de panne sont indépendantes. Selon l'ingénieur d'usine de l'entreprise, le nombre de pannes attribuables à une défaillance électronique au cours d'un mois d'opération est distribué selon une loi de Poisson de paramètre $\lambda_1 = 1.4$ tandis que le nombre de pannes attribuables à une défaillance électronique au cours d'un mois d'opération est distribué selon une loi de Poisson de paramètre $\lambda_2 = 2$.

1. Notons X la variable correspondant au nombre de pannes par mois attribuables à une défaillance électronique et Y celle correspondant au nombre de pannes par mois attribuables à une défaillance mécanique. Donner, dans chaque cas, l'expression de la loi de probabilité correspondante.
2. Quelle est la probabilité pour qu'au cours d'un mois d'opération, il y ait une seule panne de l'appareil d'étiquetage ?

Exercice : défaillance chez Electropak (suite)

1. Quelle est la probabilité pour qu'au cours d'un mois d'opération, l'appareil présente deux pannes, l'une attribuable à une défaillance électronique, l'autre attribuable à une défaillance mécanique ?
2. Quelle est l'expression de la loi de probabilité du nombre total de pannes $W = X + Y$ au cours d'un mois d'opération ?
3. Quelles sont l'expérance et la variance de W ?
4. Quelle est la probabilité pour qu'au cours d'un mois d'opération, l'appareil présente moins de deux pannes ?
5. Quelle est la probabilité pour qu'au cours d'un mois d'opération, l'appareil présente au moins deux pannes ?

Exercice : vérification de comptes-clients

Les vérifications effectuées au cours des dix dernières années révèlent que sur les 50 comptes-clients de l'entreprise Simex, en moyenne 3 sont inexacts. Le vérificateur de l'entreprise prélève, au hasard des 50 comptes, 6 comptes pour en vérifier l'exactitude.

1. Notons X la variable correspondant au nombre de comptes inexacts dans l'échantillon de taille 6 provenant des 50 comptes-clients. Donner l'expression de la loi de probabilité de X.
2. Quelle est la variance du nombre de comptes inexacts pour un échantillon de taille 6 ?
3. Quelle est la probabilité que le vérificateur ne trouve aucun compte inexact dans l'échantillon de taille 6 ?
4. Quelle est la probabilité que le vérificateur trouve plus de 1 compte inexact dans l'échantillon de taille 6 ?
5. Peut-on obtenir plus de 3 comptes inexacts dans l'échantillon de taille 6 ?

Exercice : fusibles défectueux

Une fabrication de fusibles comporte habituellement 2% de défaut.
Quelle est la probabilité d'observer 4 fusibles défectueux

1. dans un lot de taille 20 ?
2. dans un lot de taille 50 ?
3. dans un lot de taille 100 ?
4. pour chaque taille de lot mentionnée précédemment, combien de fusibles défectueux peut-on s'attendre à observer en moyenne ?