

Combinatorics

We want to choose k elements among n elements.

	Repetition not allowed	Repetition allowed
Order does not matter	$\frac{n!}{(n-k)!k!}$ <p>(n choices, then $n-1$ choices, etc. and we stop at $n-k$, we also remove $n!$ because order does not matter and $n!$ is the number of permutations)</p>	$\binom{n-1+r}{n-1}$ <p>It's stars and bars method. We want r stars + n stars for each box. Then, we transform $n-1$ of these stars into bars to separate the r stars into boxes.</p>
Order matters	$n!$ <p>(n choices, then $n-1$ choices, etc.)</p>	n^k <p>Cartesian product.</p>

Probabilities

Solving a probability problem

- list possible outcomes, define the probability space
- sometimes we keep a general Ω and different \mathcal{F} depending on the point of view (colorblind/not colorblind, etc.)

Terminology

- Ω is the **sample space**, containing all possible outcomes ω .
- \mathcal{F} is an **event space** (there are multiple event spaces!). It is a set of the subsets of Ω . The powerset of Ω includes all \mathcal{F} . $|\mathcal{F}| = 2^{|\Omega|}$ only if Ω is finite.

\mathcal{F} is also called a sigma-algebra.

Example for a fair die:

- **Sample space:** $\{1, 2, 3, 4, 5, 6\}$
- **Events:** $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{1, 2\}, \{1, 3\}, \dots, \{1, 2, 3, 4, 5, 6\}$ note that this is only for one throw! $\{a, b\}$ is read "getting a or getting b ".
- **Event space:** all events

Axioms

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Generalization of the OR between events:

$$P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_2 \cap A_3) - P(A_1 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$$

$$\mathbb{P}(\cup_{i=1}^n A_i) = \sum_{r=1}^n (-1)^{r+1} \sum_{1 \leq i_1 < i_2 < \dots < i_r \leq n} (A_{i_1} \cap \dots \cap A_{i_r})$$

(les i_k doivent donc être différents)

Conditional probabilities:

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

Law of total probability

Let $\{B_i\}_{i=1}^{\infty}$ be pairwise disjoint events, and let $A \subset \cup_{i=1}^{\infty} B_i$ then:

$$P(A) = \sum_{i=1}^{\infty} P(A \cap B_i) = \sum_{i=1}^{\infty} \mathbb{P}(A|B_i)\mathbb{P}(B_i)$$

Bayes' Theorem:

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

but we can replace $P(B)$ by what we know, law of total probability:

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

Derangement

$$!n = \sum_{i=0}^n \frac{(-1)^i}{i!}, n \geq 0$$

Si on veut un élément non dérangé parmi n :

$$n \cdot \frac{1}{n} \cdot !(n-1)$$

Si on en veut r :

(comment on choisit les r bien rangés · le fait qu'ils soient bien rangés et que les restants soient dérangés)

$$\binom{n}{r} \cdot \frac{(n-r)!}{n!} \cdot !(n-r)$$

Independence

If A and B are independent then $\mathbb{P}(A|B) = \mathbb{P}(A)$.

They are independent iff $\mathbb{P}(A \cup B) = \mathbb{P}(A)\mathbb{P}(B)$.

If two events are disjoint they can not be independent unless their respective probabilities are 0.

Pairwise Independence

If you take any two events, they are independent.

Random variable (notation)

We will use Y to denote a random variable.

We will use y to denote a specific value that Y can take.

$Y : \Omega \rightarrow \mathbb{R}$ is a function.

$$D_Y = \{x \in \mathbb{R} : \exists \omega \in \Omega \text{ s.t. } X(\omega) = y\}$$

D_Y is called the support of X . If D_Y is countable, then Y is a discrete random variable.

$$\mathbb{P}(Y = y) \equiv p_Y(y)$$

p_Y is called the probability mass function.

A probability distribution is a table or a graph that provides $p(y) \forall y$.

For everything to hold, $\sum_y p(y) = 1$.

Expected Value

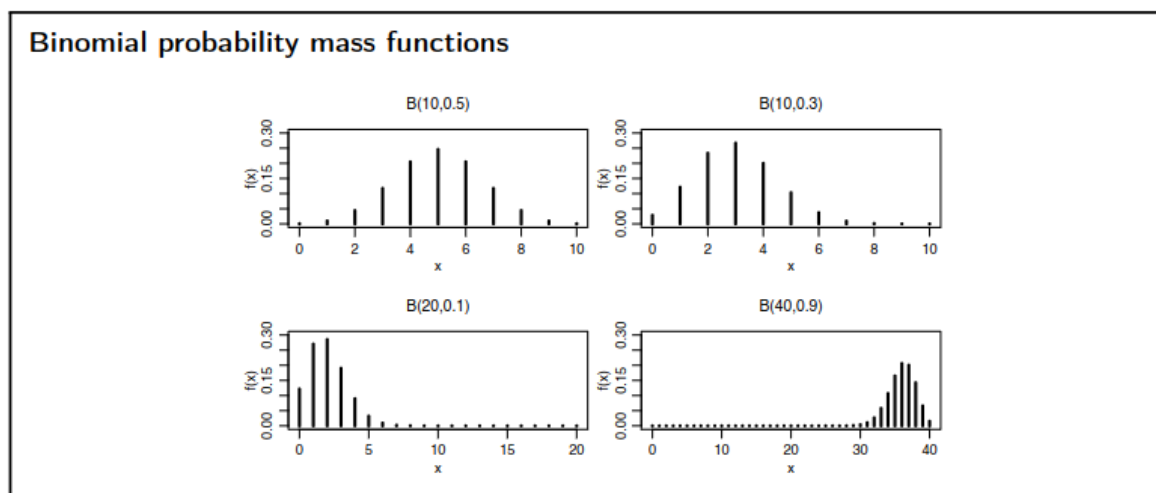
$$E(Y) = \sum_y yp(y)$$

Binomial Random Variable

A binomial random variable X has PMF:

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, x = 0, 1, \dots, n, n \in \mathbb{N}, 0 \leq p \leq 1$$

We write $X \sim B(n, p)$ and call n the denominator and p probability of success.



Geometric distribution

$$f_{X(x)} = p(1-p)^{x-1}$$

1 success and $x - 1$ failures. the probability to have exactly one success

Memorylessness :

$$P(X > n + m \mid X > m) = P(X > n)$$

Thanks to independence.

Negative binomial distribution

Loi binomiale négative

- **Définition** : La loi binomiale négative généralise la loi géométrique. Elle modélise le nombre d'essais nécessaires pour obtenir un certain nombre de succès r (au lieu d'un seul succès comme dans la loi géométrique), avec une probabilité de succès constante p à chaque essai.
- **Paramètres** : r (nombre de succès désirés) et p (probabilité de succès).
- **Exemple** : Si on lance une pièce de monnaie et qu'on souhaite savoir combien de lancers sont nécessaires pour obtenir 3 faces, on utilise la loi binomiale négative.
- **Formule** :

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

où k est le nombre total d'essais nécessaires pour obtenir r succès.

$k-1$ parce que le dernier doit être un succès.

Hypergeometric distribution

Définition

Soit une population de N éléments parmi lesquels K sont des succès (ou éléments d'intérêt) et $N - K$ sont des échecs (ou éléments sans intérêt). Si on effectue un échantillonnage de n éléments sans remise, la probabilité d'obtenir exactement k succès dans l'échantillon suit une distribution hypergéométrique.

La fonction de probabilité de la distribution hypergéométrique est donnée par la formule :

$$P(X = k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$$

où :

- N est la taille totale de la population,
- K est le nombre total de succès dans la population,
- n est le nombre d'éléments tirés de la population,
- k est le nombre de succès dans l'échantillon,
- $\binom{a}{b}$ représente le coefficient binomial, qui donne le nombre de façons de choisir b objets parmi a .

Discrete uniform distribution

Discrete uniform distribution

Definition 74. A **discrete uniform** random variable X has PMF

$$f_X(x) = \frac{1}{b - a + 1}, \quad x = a, a + 1, \dots, b, \quad a < b, \quad a, b \in \mathbb{Z}.$$

We write $U \sim \text{DU}(a, b)$.

This definition generalizes the outcome of a die throw, which corresponds to the $\text{DU}(1, 6)$ distribution.

Chaque évènement a la même probabilité de se produire.

Poisson random variable

Poisson distribution

Definition 75. A **Poisson** random variable X has the PMF

$$f_X(x) = \frac{\lambda^x}{x!} e^{-\lambda}, \quad x = 0, 1, \dots, \quad \lambda > 0.$$

We write $X \sim \text{Pois}(\lambda)$.

□ Since $\lambda^x/x! > 0$ for any $\lambda > 0$ and $x \in \{0, 1, \dots\}$, and

$$e^{-\lambda} = \frac{1}{e^\lambda} = \frac{1}{\sum_{x=0}^{\infty} \frac{\lambda^x}{x!}} > 0,$$

we see that $f_X(x) > 0$ and $\sum_{x=0}^{\infty} f_X(x) = 1$, so this is a probability distribution.

- The Poisson distribution appears everywhere in probability and statistics, often as a model for counts, or for a number of rare events.
- It also provides approximations to probabilities, for example for random permutations (Example 47, random hats) or the binomial distribution (later).

Le nombre d'événements qui se produisent dans un intervalle de temps ou d'espace fixe, lorsque ces événements se produisent avec une certaine moyenne (taux d'événements λ) et de manière indépendante.

Par exemple le nombre d'appels téléphoniques reçus par un centre d'appel en une heure.

Cumulative Distribution Function

Cumulative distribution function

Definition 76. The **cumulative distribution function (CDF)** of a random variable X is

$$F_X(x) = P(X \leq x), \quad x \in \mathbb{R}.$$

If X is discrete, we can write

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

which is a step function with jumps at the points of the support D_X of $f_X(x)$.

When there is no risk of confusion, we write $F \equiv F_X$.

Example 77. Give the support and the probability mass and cumulative distribution functions of a Bernoulli random variable.

Example 78. Give the cumulative distribution function of a geometric random variable.

La probabilité que X prenne une valeur inférieure ou égale à x .

Moments

$$E(X^r) = \sum_x (x_i)^r f_X(x_i)$$

Variance

On pourrait étudier la différence entre X et la moyenne pour voir à quel point X “varie” (une fonction constante aurait 0) : $E|X - E(X)|$.

Cependant on met un carré pour éviter les problèmes avec l'utilisation de la valeur absolue.