PROBABILITIES

1 Probability space

Proposition (Properties of a probability function).

Let $\left(\Omega, \mathcal{P}(\Omega), P\right)$ *be a probability space.*

1. $P(\emptyset) = 0$.
2. *If* $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$, *the values* $P(\{\omega_1\})$, $P(\{\omega_2\})$, ...,$P(\{\omega_n\})$ *enable one to define the function* P.
3. $\forall (A, B) \in \left[\mathcal{P}(\Omega)\right]^2$, $P(A) = P(A \setminus B) + P(A \cap B)$.
4. $\forall (A, B) \in \left[\mathcal{P}(\Omega)\right]^2$, $P(A \cup B) + P(A \cap B) = P(A) + P(B)$.

General comment: you know these properties well, you have seen them before. The purpose here is to see that they can be proved using the basic axioms of a probability function.

Proof:
From the definition of a probability function, we know that:

1. $P(\Omega) = 1$
2. For all $(A, B) \in \left[\mathcal{P}(\Omega)\right]^2$ such that $A \cap B = \emptyset$,

$$P(A \sqcup B) = P(A) + P(B)$$

Now, let us prove the four properties.

1. For any set $A \in \mathcal{P}(\Omega)$, we have $A \cap \emptyset = \emptyset$. Thus, by using the second axiom satisfied by the function P, we get:

$$P(A \sqcup \emptyset) = P(A) + P(\emptyset)$$

However, $A \sqcup \emptyset = A$. Hence,

$$P(A) = P(A) + P(\emptyset)$$

which implies: $P(\emptyset) = 0$.

2. We have to show that for all $A \in \mathcal{P}(\Omega)$, the value of $P(A)$ can be determined. Let $A \in \mathcal{P}(\Omega)$. Since Ω is a finite set, $A \subset \Omega$ is a finite set too. We can hence enumerate its elements and write A with the form

$$A = \{x_1, x_2, \cdots, x_p\} \qquad \text{with} \qquad \text{Card}(A) = p$$

Thus, A can be expressed as the disjoint union:

$$A = \{x_1\} \sqcup \{x_2\} \sqcup \cdots \sqcup \{x_p\}$$

By using the second axiom satisfied by the function P, we get:

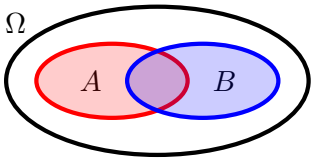
$$P(A) = P(\{x_1\}) + P(\{x_2\}) + \cdots + P(\{x_p\})$$

3. Let $A \in \mathcal{P}(\Omega)$. A can be written as the disjoint union

$$A = (A \setminus B) \sqcup (A \cap B)$$

Thus,

$$P(A) = P(A \setminus B) + P(A \cap B)$$



4. Let $(A, B) \in \left[\mathcal{P}(\Omega)\right]^2$. The set $A \cup B$ can be written as the disjoint union

$$A \cup B = (A \setminus B) \sqcup (A \cap B) \sqcup (B \setminus A)$$

We hence get:

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

Furthermore, using the property 3,

$$\begin{cases} P(A) &= P(A \setminus B) + P(A \cap B) \\ P(B) &= P(B \setminus A) + P(A \cap B) \end{cases}$$

By summing these two relations, we get:

$$P(A) + P(B) = \underbrace{P(A \setminus B) + P(A \cap B) + P(B \setminus A)}_{P(A \cup B)} + P(A \cap B)$$

Thus, $P(A \cup B) + P(A \cap B) = P(A) + P(B)$.

2 Expectation and variance

Theorem (**Expectation of** $aX+b$)**.** *Let X be a finite random variable and $(a, b) \in \mathbb{R}^2$. Then* $\mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b$.

Proof:

Let X be a finite random variable and $(a, b) \in \mathbb{R}^2$. Let us enumerate the possible values of X: $X(\Omega) = \{x_1, x_2, \cdots, x_n\}$. Now, consider the random variable $Y = aX + b$.

This random variable has the form $Y = f(X)$ where f is the function $x \longmapsto ax + b$. By using a proposition stated (and proven) in a MiMo's video, we know that

$$\mathrm{E}(Y) = \sum_{i=1}^{n} P(X{=}x_i) f(x_i) = \sum_{i=1}^{n} P(X{=}x_i)(a\,x_i + b\,)$$

But for all $i \in [\![1, n]\!]$, $\quad P(X{=}x_i)(a\,x_i + b) = a\,P(X{=}x_i)\,x_i + b\,P(X{=}x_i)$. Hence,

$$\mathrm{E}(Y) = \sum_{i=1}^{n} \Big(a\,P(X{=}x_i)\,x_i + b\,P(X{=}x_i)\Big) = \sum_{i=1}^{n} \Big(a\,P(X{=}x_i)\,x_i\Big) + \sum_{i=1}^{n} \Big(b\,P(X{=}x_i)\Big)$$

In the latter expression, a can be factorized in the first sum and b can be factorized in the second one. We get:

$$\mathrm{E}(Y) = a \sum_{i=1}^{n} P(X{=}x_i)\,x_i \quad + \quad b \sum_{i=1}^{n} P(X{=}x_i)$$

But in this expression, the first sum equals $\mathrm{E}(X)$ and the second sum equals 1. This leads to:

$$\mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b$$

Theorem (**Properties of the variance**). *Let X be a finite random variable and $(a, b) \in \mathbb{R}^2$. Then:*

1. $\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - \left(\mathrm{E}(X)\right)^2$,
2. $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$

Proof:

Let X be a finite random variable and $(a, b) \in \mathbb{R}^2$. Let us prove these two properties.

1. We know that, by definition of the variance, $\mathrm{Var}(X) = \mathrm{E}\left(\left(X - \mathrm{E}(X)\right)^2\right)$.

   But $\left(X - \mathrm{E}(X)\right)^2 = X^2 - 2\,\mathrm{E}(X)X + \left(\mathrm{E}(X)\right)^2$.

   Thus, $\mathrm{Var}(X) = \mathrm{E}\left(X^2 - 2\,\mathrm{E}(X)X + \left(\mathrm{E}(X)\right)^2\right)$.

   We can now use a property of the expectation, which has been stated in the lecture but is not proven in this document: if X and Y are two finite random variables, then

   $$\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$$

   Using this property in the expression above of $\mathrm{Var}(X)$, we get

   $$\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}\left(2\,\mathrm{E}(X)\,X\right) + \mathrm{E}\left(\left(\mathrm{E}(X)\right)^2\right)$$

   In the latter expression, note that:

   (a) $2\,\mathrm{E}(X)\,X$ is a random variable of the form $aX + b$ with $a = 2\,\mathrm{E}(X)$ and $b = 0$. Indeed, $\mathrm{E}(X)$ is a constant value. It depends on the **distribution** of X, but it does not depend on the **value** of X. If you randomly pick X several times, X will take different values but $\mathrm{E}(X)$ will remain constant.

   Thus, $\mathrm{E}\left(2\,\mathrm{E}(X)\,X\right) = 2\,\mathrm{E}(X)\;\mathrm{E}(X) = 2\left(\mathrm{E}(X)\right)^2$.

   (b) $\mathrm{E}\left(\left(\mathrm{E}(X)\right)^2\right)$ is the expectation of the variable $\left(\mathrm{E}(X)\right)^2$ which is a constant variable. Its expectation is hence the value of this constant, that is:

   $$\mathrm{E}\left(\left(\mathrm{E}(X)\right)^2\right) = \left(\mathrm{E}(X)\right)^2$$

   Using these two remarks, we get:

   $$\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - 2\left(\mathrm{E}(X)\right)^2 + \left(\mathrm{E}(X)\right)^2 = \mathrm{E}\left(X^2\right) - \left(\mathrm{E}(X)\right)^2$$

2. To prove property 2, let us write the definition of the variance of $aX + b$:

   $$\mathrm{Var}(aX + b) = \mathrm{E}\left(\left((aX + b) - \mathrm{E}(aX + b)\right)^2\right)$$

   Since $\mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b$, we have

   $$(aX + b) - \mathrm{E}(aX + b) = aX + b - a\,\mathrm{E}(X) - b = a\left(X - \mathrm{E}(X)\right)$$

   Thus,

   $$\mathrm{Var}(aX + b) = \mathrm{E}\left(a^2\left(X - \mathrm{E}(X)\right)^2\right) = a^2\;\mathrm{E}\left(\left(X - \mathrm{E}(X)\right)^2\right)$$

   Now, note that the expectation which comes in the latter expression is the variance of X.

   Hence, $\mathrm{Var}(aX + b) = a^2\;\mathrm{Var}(X)$.