Module 3: Probability and Statistics

Week 9 Tutorial

Probability

Key goals for the class

1. What are some **useful mathematical rules** governing probabilities?

2. How do we analyse probability distributions of **discrete random variables**?

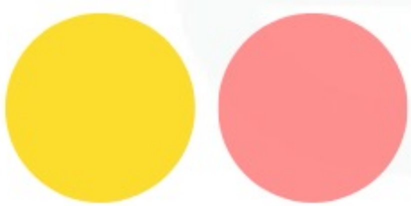
3. How do we analyse probability distributions of **continuous random variables**?

Basic rules of probability

Probability Addition Formula

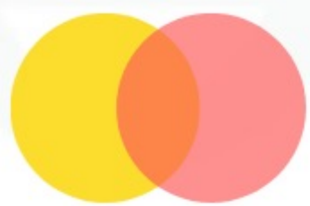
When the events are Mutually Exclusive

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



When the events are not mutually exclusive

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



Conditional Probability Formula

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

- $P(A \mid B)$: Probability of *A* given *B*
- $P(A \cap B)$: Probability of *A* and *B*
- $P(B)$: Probability of *B*

∪ means **union** (logical "or")

∩ means **intersection** (logical "and")

Mutually exclusive events cannot occur at the same time $P(A \cap B) = 0$

Independent events do not influence each other $P(A \cap B) = P(A)\, P(B)$

Mutually exclusive events are not independent, and vice versa!

Tutorial question

Try Q1 and Q2 on the tutorial sheet (**probability**).

- 1. A box contains 100 items, 2 of which are faulty. If 2 items are selected at random from the box, without replacement, calculate the probabilities that 0, 1or 2 of the selected items are faulty.

- 2. The probabilities of the events B, $A \cup B$ and $A \cap B$ are $P(B) = 0.3$, $P(A \cup B) = 0.9$ and $P(A \cap B) = 0.2$. Determine $P(A)$, $P(B|A)$ and $P(A|B)$. Conclude if the events A and B can be regarded as independent. Are the events A and B mutually exclusive?

Discrete random variables

A **discrete random variable** can only take specific values such as integers (e.g. rolling a dice: $x_i = \{1,2,3,4,5,6\}$)

Probabilities are a set of discrete values (e.g. $P(x_i) = \left\{\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right\}$)

Normalisation: $\sum_i P(x_i) = 1$

Population mean/expectation:

$$E(x) = \mu = \sum_i x_i P(x_i)$$

Population variance:

$$\sigma^2 = \sum_i (x_i - \mu)^2 P(x_i) = E(x^2) - \mu^2$$

This represents the results of an infinite number of experiments, as opposed to the *sample mean and variance*!

Tutorial question

Try Q3 on the tutorial sheet about (**discrete variables**).

- 3. A discrete random variable X takes the values $X = \{3, 4, -1, 2\}$ with probabilities $P(X = 3) = 1/5$, $P(X = 4) = 1/6$ and $P(X = 2) = 1/3$.

  (a) Determine the probability $P(X = -1)$.

  (b) The expected value $E(X)$ of a discrete random variable $X = \{X_1, X_2, X_3, \ldots\}$ is given by

  $$E(X) = \sum_i X_i P(X_i).$$

  Determine $E(X)$.

  (c) The variance $\mathrm{Var}(X)$ of a discrete random variable $X = \{X_1, X_2, X_3, \ldots\}$ is given by

  $$\mathrm{Var}(X) = \sum_i (X_i - E(X))^2 P(X_i)$$

  Prove that $\mathrm{Var}(X) = \sum_i (X_i - E(X))^2 P(X_i) = \sum_i X_i^2 P(X_i) - (E(X))^2$.

  (d) Determine the value of $\mathrm{Var}(X)$.

Continuous random variables

A **continuous random variable** can take on a smooth range of values with probabilities that are a function (e.g. $P(x) = \frac{1}{b-a}$ for $a < x < b$)

Normalisation: $\int_{-\infty}^{\infty} P(x)\, dx = 1$

Probability in a range: $\int_{x_1}^{x_2} P(x)\, dx$

Quartiles: $\int_{-\infty}^{Q_1} P(x)\, dx = \int_{Q_1}^{Q_2} P(x)\, dx = \cdots = 0.25$

Population mean/expectation:

$$E(x) = \mu = \int_{-\infty}^{\infty} x\, P(x)\, dx$$

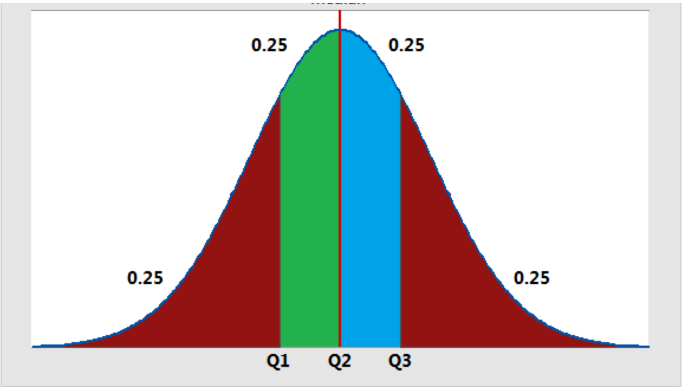
Population variance:

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2\, P(x)\, dx = E(x^2) - \mu^2$$

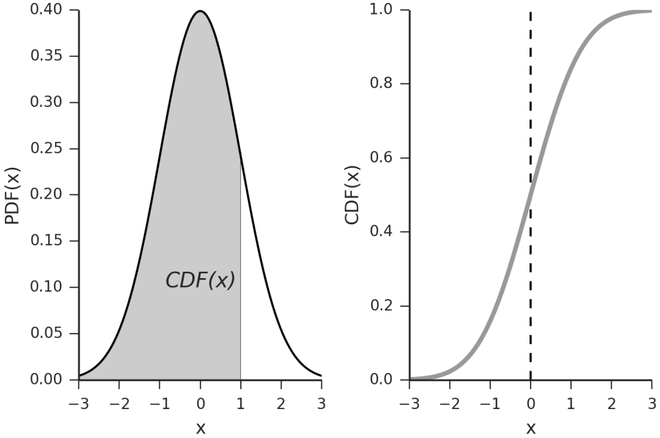
Cumulative distribution function:

$$\mathrm{cdf}(x) = \int_{-\infty}^{x} P(x')\, dx'$$

Quartiles



Cumulative distribution function



Tutorial question

Try Q4 on the tutorial sheet (**continuous variables**).

- 4. The pdf of a continuos random variable $X > 1$ is $f(x) = C/x^4$.

  (a) Determine the constant C

  (b) Determine the cdf function $F(x)$.

  (c) The expected value $E(X)$ is defined as

  $$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx.$$

  Determine the numerical value of $E(X)$.

  (d) Determine the numerical value of $\mathrm{Var}(X)$.

  (e) Determine the probability of $X \in [3, 5]$.

  (f) Determine the probability that $X = 3.5$

  (g) Determine the three qurtiles Q_1, Q_3, $Q_2 = \texttt{Md}$.

That’s all for today!