PROBABILITY A4, Lesson 3.
Random Variables

- *Random Variable:* an $X : \Omega \to \mathbb{R}$ stochastic function that is measurable with respect to $\mathcal{A}$. It means that
$$\{\omega \,|\, X(\omega) \in B\} = A \in \mathcal{A} \qquad \forall B \in \mathcal{B},$$
where $\mathcal{B}$ denotes the set of Borel-sets of $\mathbb{R}$. *Distribution of* X: the collection of the probabilities $\mathbb{P}(A)$'s of the above A's. Of course, we need not give all of them.

- *Special types of random variables:*
  1. **Discrete probability distributions**: X takes on values $x_1, x_2, \dots$ $\quad \mathbb{P}(X = x_i) = p_i$, $i = 1, 2, \dots$ $\quad (\sum_i p_i = 1)$. The collection of p_i's is called *probability mass function (p.m.f.)* of X. The *mode* of X: the value(s) taken on with the largest probability.
  2. **Absolutely continuous probability distributions:** The range of X is not countable and for any $x \in \mathbb{R}$: $\mathbb{P}(X = x) = 0$. However, there is an $f : \mathbb{R} \to \mathbb{R}$ nonnegative, integrable function such that
  $$\int_{-\infty}^{\infty} f(x)\,dx = 1 \quad \text{and} \quad \int_B f(x)\,dx = \mathbb{P}(X \in B), \quad \forall B \in \mathcal{B}.$$
  f is called *probability density function (p.d.f.)* of X.
  Cumulative distribution function (c.d.f.) of X: $F : \mathbb{R} \to \mathbb{R}$ such that
  $$F(x) = \mathbb{P}(X < x) = \int_{-\infty}^{x} f(t)\,dt, \quad x \in \mathbb{R}.$$
  F is continuous, increasing, $\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to\infty} F(x) = 1$. F is almost everywhere differentiable (at the points of continuity of f) and for such x's: $F'(x) = f(x)$.
  $$\mathbb{P}(a < X < b) = F(b) - F(a) = \int_a^b f(x)\,dx \qquad (a < b).$$
  (For discrete distributions the above F is a stepwise constant, increasing, left-continuous function, $\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to\infty} F(x) = 1$.)

- **Expectation** of X (center of gravity of the mass distribution):
  1. $\mathbb{E}(X) = \sum_i x_i p_i$ $\quad$ (if it is absolutely convergent).
  2. $\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x)\,dx$ $\quad$ (if it is absolutely convergent).

  For $X \geq 0$: 1. $\mathbb{E}(X) = \sum_{i=0}^{\infty} \mathbb{P}(X > i)$ if $X \in \mathbb{N}$, 2. $\mathbb{E}(X) = \int_0^{\infty} (1 - F(x))\,dx$.

- **Variance** of X (inertia with resp. to the center of gravity of the mass distribution):
$$V(X) = \mathbb{D}^2(X) = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}(X^2) - \mathbb{E}^2(X), \quad \text{provided } \mathbb{E}(X^2) < \infty.$$
Standard deviation of X: $\mathbb{D}(X) = \sqrt{V(X)} \geq 0$ and $=0$ if and only if $\mathbb{P}(X = cst.) = 1$.

- k-th *Moment* of X: $M_k(X) = \mathbb{E}(X^k)$, k-th *Central Moment* of X: $M_k^c(X) = \mathbb{E}(X - \mathbb{E}X)^k$ (if exists, then also exists for $1 \leq s < k$). $\mathbb{E}(X) = M_1(X)$, $V(X) = M_2^c(X) = M_2(X) - [M_1(X)]^2$.

- *Steiner's Theorem*: $\mathbb{E}(X - c)^2 = \mathbb{E}(X - \mathbb{E}X)^2 + (\mathbb{E}X - c)^2 \geq \mathbb{D}^2 X$, min. if $c = \mathbb{E}X$.

- p**-quantile** value or $100p$-*percentile* of X is x_p if $F(x_p) = p$. **Median**: 0.5-quantile value.