Expectation: The population mean. For a discrete random variable $X$

$$\mu = \mathbb{E}[X] = \sum_x x \cdot P(x),$$

Linearity of expectations. For a collection of random variables $X_1, ..., X_n$

$$\mathbb{E} \left[ \sum_i X_i \right] = \sum_i \mathbb{E}[X_i].$$

Expectation of a function $f(s)$ given a random variable $X$ is

$$\mathbb{E}[f(x)] = \sum_x f(x) \cdot P(x),$$

and the distribution of $f(x)$ is

$$P(f(X)) = P(f(X) = y) = \sum_{x : g(x) = y} P(x).$$

Markov’s inequality is an example of a law of large numbers, for a positive random variable $X \geq 0$

$$P(X \geq a) \leq \frac{\mathbb{E}(X)}{a}, \quad \forall a > 0.$$ 

Tail sum formula, for a positive random variable $X \geq 0$

$$\mathbb{E}(X) = \sum_{i=1}^{\infty} P(X > i).$$

For independent random variables $X$ and $Y$

$$\mathbb{E}[XY] = \mathbb{E}[X] \mathbb{E}[Y].$$

Moments: The $k$-th moment of a discrete distribution is

$$M_k = \mathbb{E}(X^k) = \sum_x x^k \cdot P(x).$$

Variance: The variance of a random variable measures its spread. For a discrete random variable $X$

$$\sigma^2 = \text{Var}(X) = \sum_x (x - \mu)^2 \cdot P(x).$$

The standard deviation is $\sigma = \sqrt{\text{Var}(X)}$. Shifting and scaling a r.v. $X$ changes the variance by

$$\text{Var}(aX + b) = a^2 \cdot \text{Var}(X).$$
If $X$ and $Y$ are independent then
\[ \text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y). \]

Given a r.v. $X$, the random variable $Z = \frac{X - \mu}{\sigma}$ is standardized which means $\mathbb{E}(Z) = 0$ and $\text{Var}(Z) = 1$.

**Sums of random variables:** Given $n$ random variables $X_1, \ldots, X_n$ that are independent and identically distributed (iid) the sum
\[ S_n = X_1 + \cdots + X_n, \]
has mean $n\mu$ and variance $n\sigma^2$. If $n$ is big the the following standardized random variable
\[ Z_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}, \]
is normally distributed with mean 0 and variance 1. This is an example of a central limit theorem. Also
\[ \Pr\left(a \leq \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \leq b\right) \approx \Phi(b) - \Phi(a), \]
where $\Phi(u)$ is the cumulative distribution function of the standard normal distribution.