Rule	**Description / Formula**
Sample Space (Ω or S)	The set of all possible outcomes of an experiment.
Probability of an Event	The likelihood of an event occurring. For event A: $$P(A) = \frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}}$$
Complementary Rule	The probability of the complement of an event A is: $$P(A^c) = 1 - P(A)$$
Addition Rule	For two mutually exclusive events A and B: $$P(A \cup B) = P(A) + P(B)$$ For two non-mutually exclusive events: $$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
Multiplication Rule	For two independent events A and B: $$P(A \cap B) = P(A) \cdot P(B)$$ For two dependent events: $$P(A \cap B) = P(A) \cdot P(B\|A)$$
Conditional Probability	The probability of event A given that event B has occurred: $$P(A\|B) = \frac{P(A \cap B)}{P(B)}$$
Bayes' Theorem	Relates the conditional and marginal probabilities of random events: $$P(A\|B) = \frac{P(B\|A) \cdot P(A)}{P(B)}$$
Total Probability Rule	If $B_1, B_2, \ldots, B_n$ are mutually exclusive and exhaustive events: $$P(A) = \sum_{i=1}^{n} P(A\|B_i) \cdot P(B_i)$$
Law of Total Probability	If $B_1, B_2, \ldots, B_n$ form a partition of the sample space S: $$P(A) = \sum_{i=1}^{n} P(A \cap B_i)$$

Table 1: Basic Rules of Probability

Concept	**Definition / Formula**
Mean (Expected Value)	For a random variable X with a probability mass function $P(X = x)$ or probability density function $f(x)$: $$\mathbb{E}[X] = \begin{cases} \sum_x xP(X = x) & \text{(Discrete)} \\ \int_{-\infty}^{\infty} xf(x)\,dx & \text{(Continuous)} \end{cases}$$
Variance	For a random variable X with mean μ: $$\text{Var}(X) = \mathbb{E}[(X-\mu)^2] = \begin{cases} \sum_x (x-\mu)^2 P(X = x) & \text{(Discrete)} \\ \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx & \text{(Continuous)} \end{cases}$$
Standard Deviation	$$\sigma = \sqrt{\text{Var}(X)}$$
Covariance	For two random variables X and Y with means μ_X and μ_Y: $$\text{Cov}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)]$$
Correlation	For two random variables X and Y: $$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$
Cumulative Distribution Function (CDF)	For a random variable X: $$F(x) = P(X \leqslant x) = \begin{cases} \sum_{t \leqslant x} P(X = t) & \text{(Discrete)} \\ \int_{-\infty}^{x} f(t)\,dt & \text{(Continuous)} \end{cases}$$
Moment	The n-th moment of a random variable X about the origin: $$\mathbb{E}[X^n]$$
Moment Generating Function (MGF)	For a random variable X: $$M_X(t) = \mathbb{E}[e^{tX}]$$
Law of Large Numbers	The average of the results obtained from a large number of trials should be close to the expected value, and will tend to become closer as more trials are performed.
Central Limit Theorem	The distribution of the sum (or average) of a large number of independent, identically distributed random variables approaches a normal distribution, regardless of the shape of the original distribution.

Table 2: Formulas and Definitions of Basic Concepts in Probability and Statistics

Distribution	**PDF/PMF**	**Mean**	**Variance**
Bernoulli	$P(X = x) = p^x(1-p)^{1-x}, \quad x \in \{0, 1\}$	p	$p(1-p)$
Binomial	$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$	np	$np(1-p)$
Poisson	$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \ldots$	λ	λ
Uniform (Continuous)	$f(x) = \frac{1}{b-a}, \quad a \leqslant x \leqslant b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2
Exponential	$f(x) = \lambda e^{-\lambda x}, \quad x \geqslant 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Geometric	$P(X = k) = (1-p)^{k-1} p, \quad k = 1, 2, \ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Multinomial	$P(X_1 = k_1, \ldots, X_k = k_k) = \frac{n!}{k_1! \cdots k_k!} p_1^{k_1} \cdots p_k^{k_k}$	np_i	$np_i(1-p_i)$
Gamma	$f(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)}, \quad x \geqslant 0$	$\frac{k}{\lambda}$	$\frac{k}{\lambda^2}$
Beta	$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, \quad 0 \leqslant x \leqslant 1$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

Table 3: Common Discrete and Continuous Distributions

Hypothesis Test	**Purpose**	**Test Statistic**
One-Sample t-test	Determine if the sample mean is significantly different from a known or hypothesized population mean.	$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$
Two-Sample t-test	Compare the means of two independent samples	$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$ where s_p is the pooled standard deviation.
Paired t-test	Compare the means of two related groups.	$$t = \frac{\bar{d}}{s_d/\sqrt{n}}$$ where $\bar{d}$ is the mean of the differences and s_d is the standard deviation of the differences.
Chi-Square Test for Independence	checks whether two variables are likely related or not.	$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$ where O_i and E_i are observed and expected frequencies, respectively.
Chi-Square Goodness of Fit Test	determine whether a variable likely comes from a specified distribution or not	$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$ where O_i and E_i are observed and expected frequencies, respectively.
ANOVA (Analysis of Variance)	compare the variances of two samples or the ratio of variances among multiple samples.	$$F = \frac{\text{MS}_{\text{between}}}{\text{MS}_{\text{within}}}$$ where MS is the mean square.
Mann-Whitney U Test	test under the assumption of continuous responses, with the alternative hypothesis that one distribution is stochastically greater than the other..	$$U = \min(U_1, U_2)$$ where U_1 and U_2 are the test statistics for each group.
Wilcoxon Signed-Rank Test	Compare two related samples, matched samples, or repeated measurements on a single sample to assess whether their population mean ranks differ.	$$W = \sum \text{ranked differences}$$
Kruskal-Wallis Test	Compare two or **more** independent groups without the normal assumption.	$$H = \frac{12}{N(N+1)} \sum \frac{R_i^2}{n_i} - 3(N+1)$$ where R_i is the sum of ranks in the i-th group and n_i is the number of observations in the i-th group.
Fisher's Exact Test	Test to assess the association between two categorical variables, especially useful for small sample sizes.	Calculated using the hypergeometric distribution.

Table 4: Common Hypothesis Tests