Statistics Formula Sheet

1. Descriptive Statistics

Measures of Central Tendency

- **Mean**:

$$\bar{x} = \frac{\sum x_i}{n}$$

- **Median**: Middle value in ordered data
- **Mode**: Most frequent value

Measures of Dispersion

- **Variance (Sample)**:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

- **Standard Deviation**:

$$s = \sqrt{s^2}$$

- **Range**:

$$\text{Max} - \text{Min}$$

2. Probability Basics

Key Concepts

- **Classical Probability**:

$$P(A) = \frac{\text{Favorable outcomes}}{\text{Total outcomes}}$$

- **Addition Rule**:

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

- **Conditional Probability**:

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

Random Variables

- **Expected Value**:

$$E(X) = \sum x_i P(x_i)$$

- **Variance**:

$$Var(X) = E(X^2) - [E(X)]^2$$

3. Probability Distributions

Discrete Distributions

- **Binomial** ($X \sim B(n,p)$):

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

  - Mean: np, Variance: $np(1-p)$

Continuous Distributions

- **Normal Distribution** ($X \sim N(\mu, \sigma^2)$):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- **Standard Normal** ($Z \sim N(0,1)$):

$$Z = \frac{X - \mu}{\sigma}$$

4. Correlation & Regression (Intro)

Correlation

- **Pearson's r:**

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

  - Range: $-1 \leq r \leq 1$

Simple Linear Regression

- **Regression Line:**

$$\hat{y} = a + bx$$

  - **Slope (b):**

$$b = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

  - **Intercept (a):**

$$a = \bar{y} - b\bar{x}$$

5. Index Numbers

- **Price Index:**

$$\text{Price Index} = \frac{\sum p_t q_0}{\sum p_0 q_0} \times 100 \quad \text{(Laspeyres)}$$

- **Quantity Index:**

$$\text{Quantity Index} = \frac{\sum q_t p_0}{\sum q_0 p_0} \times 100$$