- **Sample mean :** $\bar{X} = \frac{\sum X}{n}$
- **Sample variance:**

$$S^2 = \frac{\sum (X - \bar{X})^2}{n-1} = \frac{\sum X^2 - \frac{1}{n}(\sum X^2)}{n-1}$$

- **Locating percentiles:** P_α

$$R_\alpha = \frac{\alpha}{100}(n+1) = i.d$$

$$P_\alpha = X_{(i)} + d(X_{(i+1)} - X_{(i)})$$

- **Coefficient of Variation:** $CV = s / \bar{X}$
- **Coefficient of skewness:** $CS = \frac{3(\bar{X} - \tilde{X})}{s}$
- **Probability**

$$P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B)$$

$$P(A') = 1 - P(A)$$

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

- **Total Rule of Probability:**

$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i)$$

$$P(A | B_r) = \frac{P(A|B_r)P(B_r)}{P(A)}$$

- **Bayes' Rule:** $\frac{P(A| B_r)P(B_r)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}$
- **Expected values**

$$E(X) = \sum xf(x) or \int_{-\infty}^{\infty} xf(x)dx$$

$$\sigma^2 = \sum [x-\mu]^2 f(x) or \int_{-\infty}^{\infty} [x-\mu]^2 f(x)dx$$

$$\sigma^2 = E(x^2) - [E(x)]^2$$

- **Binomial distribution**

$$b(x;n,p) = \binom{n}{x} p^x q^{n-x}, x = 0,1,2,....n$$

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}, \ \mu = np, q = 1-p, \sigma^2 = npq$$

- **Poisson distribution**

$$p(x;\lambda) = \Pr[X = x] = \frac{e^{-\lambda}\lambda^x}{x!}, X = 0,1,2...$$

$$E(X) = Var(X) = \lambda$$

- **Geometric Distribution**

$$g(x;p) = pq^{x-1}, \mu = \frac{1}{p}, \sigma^2 = \frac{q}{p^2}$$

- **Hypergeometric Distribution**

$$h(x,n,k,N) = P(X = x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

$$\mu = E(X) = \frac{nK}{N}, \sigma^2 = \frac{nK}{N}(1-\frac{K}{N})(\frac{N-n}{N-1})$$

- **The Exponential Distribution**

$$f(x) = \frac{1}{\beta} e^{-x/\beta}, x > 0$$

$\mu = \beta$ and $\sigma^2 = \beta^2$

- **Sampling distribution**

$$\mu_{\bar{X}} = \mu, \ \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2,$$

$$\sigma^2_{\bar{X}_1 - \bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

- **Student t-distribution**

$$t = \frac{\bar{X} - \mu}{S / \sqrt{n}}, df = n-1$$

- **Confidence Interval Estimation single Sample; C.I. for** μ

  when σ known: $\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

- **When** σ unknown:
  - **a. n large:** $\bar{x} \pm z_{\alpha/2} \frac{S}{\sqrt{n}}$
  - **b. n small:** $\bar{x} \pm t_{\alpha/2} \frac{S}{\sqrt{n}}$

- **Error =** $e = \frac{z_{\alpha/2}\sigma}{\sqrt{n}}$ **,** $e = \frac{z_{\alpha/2}S}{\sqrt{n}}$ **or**

  $e = \frac{t_{\alpha/2}S}{\sqrt{n}}$

- **Required sample size** $n \geq \left(\frac{Z_{\alpha/2}\sigma}{e}\right)^2$
- **Single sample C.I. for P**

  a. $\hat{P} \pm z_{\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$

  b. Error = $z_{\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$

  c. Required sample size n is

  $n \geq \frac{Z_{\alpha/2}^{2}\hat{P}(1-\hat{P})}{e^2}$

- **Two-Sample C.I. for** $\mu_1 - \mu_2$

  1) If σ_1 and σ_2 **are known**

  $(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

  2) If σ_1 and σ_2 **are unknown; large samples**

  $(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

  3) Small samples and variances are equal

  $(\bar{x}_1 - \bar{x}_2) + \pm t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$, where

  $s_P^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$

  $df = n_1 + n_2 - 2$

- **Two-Sample C.I for** $P_1 - P_2$

  $(\hat{P}_1 - \hat{P}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{P}_1\hat{Q}_1}{n_1} + \frac{\hat{P}_2\hat{Q}_2}{n_2}}$, where

  $\hat{P}_i = \frac{X_i}{n_i}, \hat{q}_i = 1 - \hat{P}_i$

- **Testing Hypotheses**

1) Single Sample: Testing about μ

a. When σ **known**

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

b. Large sample

$$Z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

c. When σ **unknown and n small**

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}, df = n - 1$$

2) Single sample: testing about P

$$z = \frac{X - np_0}{\sqrt{np_0q_0}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0q_0}{n}}}$$

3) Two samples: Testing about $\mu_1 - \mu_2$

a. If σ_1 and σ_2 **are known**

$$z = \frac{\bar{x}_1 - \bar{x}_2 - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

b. If σ_1 and σ_2 **are unknown**

Large samples test:

$$z = \frac{\bar{x}_1 - \bar{x}_2 - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

c. Small samples and variances equal

$$t = \frac{\bar{x}_1 - \bar{x}_2 - d_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

4) Two samples: Testing about: $P_1 - P_2$

Large samples under $H_0 : P_1 - P_2 = 0$

$$z = \frac{\hat{P}_1 - \hat{P}_2 - 0}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n2}\right)}},$$

where $\hat{p} = \frac{n_1 \hat{P}_1 + n_2 \hat{P}_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2}$, and

$\hat{q} = 1 - \hat{p}$

- **Simple linear Regression**

1) Estimated regression model:

$\hat{y} = a + bx$, where:

$$b = \frac{\sum xy - \sum x \sum y / n}{\sum x^2 - (\sum x)^2 / n} = \frac{S_{xy}}{S_{xx}},$$

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

2) Sum of squares:

a. $S_{xx} = \sum (x - \bar{x})^2 = \sum x^2 - (\sum x)^2 / n$

b. $S_{yy} = \sum (y - \bar{y})^2 = \sum y^2 - (\sum y)^2 / n$ =**SST**

c. $S_{xy} = \sum xy - (\sum x \sum y) / n$

d. $SSE = \sum (y_i - \hat{y}_i)^2 = S_{yy} - bS_{xy}$

e. $S^2 = MSE = \frac{SSE}{n-2}$

3) Inference about the regression coefficients

a. **C.I. for** $\beta : b \pm t_{\alpha/2} \frac{S}{\sqrt{S_{XX}}}, df = n - 2$

b. **C.I. for** $\alpha : \alpha \pm \frac{t_{\alpha/2} S \sqrt{\sum x^2}}{\sqrt{nS_{XX}}}$

c. **Testing about** $\beta : t = \frac{b - \beta_0}{S / \sqrt{S_{XX}}}$

d. **Testing about** $\alpha : t = \frac{a - \alpha_0}{S \sqrt{\frac{x^2}{nS_{XX}}}}$

4) C.I. for the mean response $\mu_{y|x_0}$

$$\hat{y}_0 \pm t_{\alpha/2} S \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{XX}}}$$

5) P.I. for a single response y_0 is:

$$\hat{y}_0 \pm t_{\alpha/2} S \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{XX}}}$$

6) Coefficient of determination

$R^2 = 1 - \frac{SSE}{SST}$

7) Correlation coefficient

$r = b\sqrt{\frac{S_{XX}}{S_{YY}}} = \frac{S_{XY}}{\sqrt{(S_{XX})(S_{YY})}}, r^2 = R^2$

8) Testing about the population correlation coefficient:

$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}, df = n - 2$