# Statistical Methods Reference By Ouzidane Reda

## 1. Hypothesis Testing

- (a) t-tests
  - One-sample t-test:

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \quad (df = n - 1)$$

• Independent two-sample t-test:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$
 (Welch's df)

• Paired t-test:

$$t = \frac{\bar{D}}{s_D/\sqrt{n}} \quad (df = n - 1)$$

(b) p-value

$$p = P(T \ge |t| \mid H_0)$$
 (Reject  $H_0$  if  $p < \alpha$ )

## 2. Confidence Intervals

• For mean  $\mu$ :

$$CI = \bar{X} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

• For proportion p:

CI = 
$$\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

# 3. ANOVA (Analysis of Variance)

(a) One-Way ANOVA

$$F = \frac{MS_{\text{between}}}{MS_{\text{within}}} = \frac{SS_{\text{between}}/(k-1)}{SS_{\text{within}}/(N-k)}$$

• Total Sum of Squares:

$$SS_{\text{total}} = \sum (X_{ij} - \bar{X}_{\text{total}})^2$$

• Post-Hoc (Tukey's HSD):

$$HSD = q_{\alpha,k,df_{\text{within}}} \cdot \sqrt{\frac{MS_{\text{within}}}{n}}$$

# 4. Regression Analysis

(a) Linear Regression

$$Y = \beta_0 + \beta_1 X + \epsilon, \quad \beta_1 = \frac{\text{Cov}(X, Y)}{\text{Var}(X)}$$

$$R^2 = \frac{SS_{\text{regression}}}{SS_{\text{total}}}$$

(b) Logistic Regression

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X$$

#### 5. Non-Parametric Tests

• Mann-Whitney U:

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

• Kruskal-Wallis:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{N} \frac{R_i^2}{n_i} - 3(N+1)$$

#### 6. Effect Sizes

Metric	Formula	Interpretation
Cohen's $d$	$rac{ar{X}_1 - ar{X}_2}{s_{ ext{pooled}}} \ SS_{ ext{effect}}$	Small: 0.2, Medium: 0.5, Large: 0.8
$\eta^2$	$\frac{SS_{\text{effect}}}{SS_{\text{total}}}$	% variance explained

### 7. Decision Guide

Precision-crafted by Ouzidane Reda.

Scenario	Test
Compare 2 means	t-test
Compare $\geq 3$ means	ANOVA + post-hoc
Continuous outcome prediction	Linear regression
Binary outcome prediction	Logistic regression
Non-normal data	Mann-Whitney/Kruskal-Wallis