UNIVERSITY OF STIRLING
Computing Science & Mathematics

MATU9D2 : PRACTICAL STATISTICS : FORMULA SHEET

1. One Sample Summary Statistics

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

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

2. Two Sample Summary Statistics

$$\bar{x}_1 = \frac{\sum x_{1i}}{n_1} \qquad s_1^2 = \frac{\sum x_{1i}^2 - \frac{\left(\sum x_{1i}\right)^2}{n_1}}{n_1 - 1}$$

$$\bar{x}_2 = \frac{\sum x_{2i}}{n_2} \qquad s_2^2 = \frac{\sum x_{2i}^2 - \frac{\left(\sum x_{2i}\right)^2}{n_2}}{n_2 - 1}$$

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

3. TO COME

4. Relationships

The Chi-squared test has test statistic

$$X^2 = \sum_{all\ cells} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

Using a significance level of 0.05, this is compared with χ^2(df; 0.05), where df= (r-1)(c-1).

TO COME

5. Probability

A random variable X following a Binomial probability model with n trials and probability of success θ has probability distribution

$$\Pr[X = x] = \frac{n!}{x!(n-x)!}\theta^x (1-\theta)^{n-x}, \qquad x = 0, \ldots\ldots n.$$

and mean and variance given by

$$E(X) = n\theta, \qquad Var(X) = n\theta(1 - \theta).$$

If X follows a Normal probability model with mean μ and variance σ^2, then $Z = \dfrac{(X-\mu)}{\sigma}$ follows a N(0,1) distribution.

6. One Sample Confidence Intervals

An interval estimate for a population *proportion*, having approximate confidence 0.95, is

$$\hat{\theta} \quad \pm \quad 1.96 \times \sqrt{\frac{\hat{\theta}\left(1-\hat{\theta}\right)}{n}}$$

An interval estimate for a population *mean*, having confidence 0.95, is

$$\overline{x} \quad \pm \quad t(n-1;0.025) \times \frac{s}{\sqrt{n}}$$

An interval estimate for a population *variance*, having confidence 0.95, is

$$\left(\frac{(n-1)s^2}{\chi^2(n-1;0.025)}, \frac{(n-1)s^2}{\chi^2(n-1;0.975)}\right).$$

7. Two Sample Tests

The test statistics for the test of the Null Hypothesis $H_0 : \mu_1 = \mu_2$ against the alternative $H_1 : \mu_1 \neq \mu_2$ is

$$t \quad = \quad \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

For a test having significance level 0.05, we compare this with $\pm$ t($n_1 + n_2 - 2$; 0.025).

The test statistic for the test of the Null Hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$ against the alternative $H_1 : \sigma_1^2 \neq \sigma_2^2$ is

$$\frac{s_1^2}{s_2^2} \quad \left(\frac{larger}{smaller}\right)$$

For a test having significance level 0.05, we compare this with

$$F(n_1 - 1, n_2 - 1; 0.025).$$

8. TO COME

9. TO COME