ST104a Statistics 1

Examination Formula Sheet

Expected value of a discrete random variable:

$$\mu = \mathrm{E}(X) = \sum_{i=1}^{N} p_i \, x_i$$

Standard deviation of a discrete random variable:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_{i=1}^{N} p_i (x_i - \mu)^2}$$

The transformation formula:

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

Finding Z for the sampling distribution of the sample mean:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

Finding Z for the sampling distribution of the sample proportion:

$$Z = \frac{P - \pi}{\sqrt{\pi(1-\pi)/n}}$$

Confidence interval endpoints for a single mean (σ known):

$$\bar{x} \pm z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}$$

Confidence interval endpoints for a single mean (σ unknown):

$$\bar{x} \pm t_{\alpha/2,\, n-1} \times \frac{s}{\sqrt{n}}$$

Confidence interval endpoints for a single proportion:

$$p \pm z_{\alpha/2} \times \sqrt{\frac{p\,(1-p)}{n}}$$

Sample size determination for a mean:

$$n \geq \frac{(z_{\alpha/2})^2 \, \sigma^2}{e^2}$$

Sample size determination for a proportion:

$$n \geq \frac{(z_{\alpha/2})^2 \, p\,(1-p)}{e^2}$$

z test of hypothesis for a single mean (σ known):

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

t test of hypothesis for a single mean (σ unknown):

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

z test of hypothesis for a single proportion:

$$Z \cong \frac{P - \pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}$$

z test for the difference between two means (variances known):

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

t test for the difference between two means (variances unknown):

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{S_p^2\,(1/n_1 + 1/n_2)}}$$

Confidence interval endpoints for the difference between two means:

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

Pooled variance estimator:

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}$$

t test for the difference in means in paired samples:

$$T = \frac{\bar{X}_d - \mu_d}{S_d/\sqrt{n}}$$

Confidence interval endpoints for the difference in means in paired samples:

$$\bar{x}_d \pm t_{\alpha/2,\, n-1} \times \frac{s_d}{\sqrt{n}}$$

z test for the difference between two proportions:

$$Z = \frac{P_1 - P_2 - (\pi_1 - \pi_2)}{\sqrt{P\,(1-P)\,(1/n_1 + 1/n_2)}}$$

Pooled proportion estimator:

$$P = \frac{R_1 + R_2}{n_1 + n_2}$$

Confidence interval endpoints for the difference between two proportions:

$$p_1 - p_2 \pm z_{\alpha/2} \times \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

χ^2 statistic for test of association:

$$\sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Sample correlation coefficient:

$$r = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sqrt{\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)\left(\sum_{i=1}^{n} y_i^2 - n\bar{y}^2\right)}}$$

Spearman rank correlation:

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}$$

Simple linear regression line estimates:

$$b = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

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