36-226 Exam 3 Reference Sheet

Sunday, April 30, 2013

General Hypothesis Testing

$\alpha = P(\text{Type I Error})$ is the probability of rejecting H_0 when H_0 is true.
$\beta = P(\text{Type II Error})$ is the probability of accepting H_0 when H_a is true.

Z-Test: $H_0 : \theta = \theta_0$ $\quad Z = \frac{\hat{\theta}-\theta_0}{\sigma} \approx \frac{\hat{\theta}-\theta_0}{s/\sqrt{n}}$ $\quad RR = \begin{cases} Z > z_\alpha & \text{if } H_a : \theta > \theta_0 \\ Z < z_\alpha & \text{if } H_a : \theta < \theta_0 \\ |Z| > z_{\alpha/2} & \text{if } H_a : \theta \neq \theta_0 \end{cases}$

T-Tests are the same for $n - 1 \leq 30$ df.

For proportions: $H_0 : p = p_0$, $Z = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$

For two populations: $H_0 : \mu_1 - \mu_2 = D_0$,

$$Z = \frac{(\overline{Y}_1 - \overline{Y}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \qquad T = \frac{(\overline{Y}_1 - \overline{Y}_2) - D_0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad \text{where } S_p = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}}$$

Testing hypotheses about variances (with independent samples from normal distributions):
$H_0 : \sigma^2 = \sigma_0^2$

$$RR = \begin{cases} \chi^2 > \chi^2_\alpha & \text{if } H_a : \theta > \theta_0 \\ \chi^2 < \chi^2_{1-\alpha} & \text{if } H_a : \theta < \theta_0 \\ \chi^2 > \chi^2_{\alpha/2} OR \chi^2 < \chi^2_{1-\alpha/2} & \text{if } H_a : \theta \neq \theta_0 \end{cases},$$

where $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$ and χ^2_α is chosen so that, for $\nu = n - 1$ df, $P(\chi^2 > \chi^2_\alpha) = \alpha$.

Test of the hypothesis $\sigma_1^2 = \sigma_2^2$ (with independent samples from normal distributions):
$H_0 : \sigma_1^2 = \sigma_2^2 \quad H_a : \sigma_1^2 > \sigma_2^2$

$$RR = \{F > F_\alpha\},$$

where $F = \frac{S_1^2}{S_2^2}$, F_α chosen with $P(F > F_\alpha) = \alpha$ if S_1 has $\nu_1 = n_1 - 1$ df and S_2 has $\nu_2 = n_2 - 1$ df.

Neyman-Pearson Lemma

The most powerful test $H_0 : \theta = \theta_0$ versus $H_a : \theta = \theta_a$ at a given α has a rejection region determined by

$$\frac{L(\theta_0)}{L(\theta_a)} < k,$$

where k depends on α.

Linear Regression

$$S_{xx} = \sum_{i=1}^{n}(x_i - \overline{x})^2 = \left(\sum_{i=1}^{n} x_i^2\right) - n\overline{x}^2 \qquad S_{xy} = \sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = \left(\sum_{i=1}^{n} x_i y_i\right) - n\overline{x}\,\overline{y}$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = r\sqrt{\frac{S_{yy}}{S_{xx}}} \qquad \hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \cdot \overline{x}$$

$$SSE = S_{yy} - \hat{\beta}_1 \cdot S_{xy} \qquad S = \sqrt{\frac{SSE}{n-2}}$$

If each $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$, then

$$\hat{\beta}_0 \sim \mathcal{N}\left(\beta_0, \frac{\sigma^2 \sum_{i=1}^{n} x_i^2}{nS_{xx}}\right) \qquad \hat{\beta}_1 \sim \mathcal{N}\left(\beta_1, \frac{\sigma^2}{S_{xx}}\right) \qquad \mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_1) = \frac{-\overline{x}\sigma^2}{S_{xx}}$$

$$\frac{(n-2)S^2}{\sigma^2} \sim \chi^2 \quad \text{with } n-2 \text{ df} \qquad S^2 \perp \hat{\beta}_0, \hat{\beta}_1$$

$$H_0 : \beta_i = \beta_{i0} \qquad RR = \begin{cases} T > t_\alpha & \text{if } H_a : \beta_i > \beta_{i0} \\ T < t_\alpha & \text{if } H_a : \beta_i < \beta_{i0} \\ |T| > t_{\alpha/2} & \text{if } H_a : \beta_i \neq \beta_{i0} \end{cases} \qquad (t_\alpha \text{ has } n-2 \text{ df})$$

$$T = \frac{\hat{\beta}_i - \beta_{i0}}{S\sqrt{c_{ii}}} \qquad \text{where } c_{00} = \frac{\sum_{i=1}^{n} x_i^2}{nS_{xx}} \qquad c_{11} = \frac{1}{S_{xx}}$$

$$100(1-\alpha)\% \text{ CI for } \beta_i: \qquad \hat{\beta}_i \pm t_{\alpha/2} S\sqrt{c_{ii}}$$

$$100(1-\alpha)\% \text{ CI for } E(Y) = \beta_0 + \beta_1 x^*: \qquad \hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{\alpha/2} S\sqrt{\frac{1}{n} + \frac{(x^* - \overline{x})^2}{S_{xx}}}$$

$$100(1-\alpha)\% \text{ CI for } Y \text{ when } x = x^*: \qquad \hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{\alpha/2} S\sqrt{1 + \frac{1}{n} + \frac{(x^* - \overline{x})^2}{S_{xx}}}$$

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \hat{\beta}_1\sqrt{\frac{S_{xx}}{S_{yy}}} \qquad \rho = \beta_1 \frac{\sigma_X}{\sigma_Y} \qquad T = \frac{\hat{\beta}_1}{S/\sqrt{S_{xx}}} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t_{n-2}$$

$$r^2 = 1 - \frac{SSE}{S_{yy}}$$