EC 203 Applied Econometrics: Formula sheet

Hypothesis Testing:

Test statistics for tests on means

		σ^2 Known	σ^2 Not known
	One population		
Large sample	$H_0: \mu = \mu_0$	$z = \dfrac{(\bar{x} - \mu_0)}{\sigma / \sqrt{n}}$	$z = \dfrac{(\bar{x} - \mu_0)}{s_X / \sqrt{n}}$
	Two populations		
Large sample	$H_0: \mu_1 - \mu_2 = \delta$	$z = \dfrac{(\bar{x}_1 - \bar{x}_2 - \delta)}{\sqrt{[\sigma_1^2 / n_1 + \sigma_2^2 / n_2]}}$	$z = \dfrac{(\bar{x}_1 - \bar{x}_2 - \delta)}{\sqrt{[s_1^2 / n_1 + s_2^2 / n_2]}}$
	One population		
Small sample	$H_0: \mu = \mu_0$	$z = \dfrac{(\bar{x} - \mu_0)}{\sigma / \sqrt{n}}$	$t = \dfrac{(\bar{x} - \mu_0)}{s_X / \sqrt{n}}$
	Two populations		
Small sample	$H_0: \mu_1 - \mu_2 = \delta$	$z = \dfrac{(\bar{x}_1 - \bar{x}_2 - \delta)}{\sqrt{[\sigma_1^2 / n_1 + \sigma_2^2 / n_2]}}$	$t = \dfrac{(\bar{x}_1 - \bar{x}_2 - \delta)}{s_p \sqrt{[1/n_1 + 1/n_2]}}$

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

Small sample	$H_0: \mu_1 - \mu_2 = \delta$	$z = \dfrac{(\bar{x}_1 - \bar{x}_2 - \delta)}{\sqrt{[\sigma_1^2 / n_1 + \sigma_2^2 / n_2]}}$	$t = \dfrac{(\bar{X}_1 - \bar{X}_2 - \delta)}{\sqrt{[s_1^2 / n_1 + s_2^2 / n_2]}} \sim t_{dof}$

Where $dof = \dfrac{[s_1^2 / n_1 + s_2^2 / n_2]^2}{\sqrt{[(s_1^2 / n_1)^2 / (n_1 - 1) + (s_2^2 / n_2)^2 / (n_2 - 1)]}}$

Tests on proportions

One population $\quad H_0: \pi = \pi_0 \quad z = \dfrac{p - \pi_0}{\sqrt{\{\pi_0 (1 - \pi_0)/n\}}}$

Two populations $\quad H_0: \pi_1 - \pi_2 = 0 \quad z = \dfrac{p_1 - p_2}{\sqrt{\hat{\pi}(1 - \hat{\pi})(1/n_1 + 1/n_2)}}$ where $\hat{\pi} = \dfrac{p_1 n_1 + p_2 n_2}{n_1 + n_2}$

Tests on variances

One population $\quad H_0: \sigma^2 = \sigma_o^2 \quad u = \dfrac{(n-1)s_x^2}{\sigma_0^2} \sim \chi_{n-1}^2$

Two populations $\quad H_0: \sigma_1^2 = \sigma_2^2 \quad F = s_1^2 / s_2^2 \sim F_{(n_1 - 1, n_2 - 1)}$

Two variable linear regression model: $\quad Y_i = \alpha + \beta X_i + u_i$

Least squares estimates (OLS): $$b = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \qquad a = \bar{Y} - b\bar{X}$$

Estimation of the error variance: $$s^2 = \frac{RSS}{n-2} = \frac{\sum_{i=1}^{n} e_i^{\,2}}{n-2}$$

Tests on regression slope coefficient: $H_0 : \beta = \beta_0$, $t = \frac{(b - \beta_0)}{se(b)} \sim t_{n-2}$, where

$$se(b) = \frac{s}{\sqrt{\sum_{i-1}^{n}(X_i - \bar{X})^2}}$$

Multiple Regression Model: $$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki} + u_i$$

Estimation of the error variance: $$s^2 = \frac{RSS}{n-k-1} = \frac{\sum_{i=1}^{n} e_i^{\,2}}{n-k-1}$$

R-squared: $R^2 = \frac{ESS}{TSS}$ $\quad \bar{R}^2 = 1 - \frac{RSS/(n-k-1)}{TSS/(n-1)}$

Tests on regression coefficients:

(i) Single coefficient is equal to some hypothesised value:

$H_0 : \beta_i = \beta_{i0}$, $i = 1,2,...,k$, $\qquad t = \frac{(b_i - \beta_{i0})}{se(b_i)} \sim t_{n-k-1}$

(ii) All slope coefficients are equal to zero:

$H_0 : \beta_1 = \beta_2 = ... = \beta_k = 0$, $\quad F = \frac{R^2/k}{(1-R^2)/(n-k-1)} \sim F_{k,n-k-1}$

(iii) A sub-set (h) of coefficients are equal to zero:

$H_0 : \beta_1 = \beta_2 = ... = \beta_q = 0$, $\quad F = \frac{(RSS_R - RSS_{UR})/q}{(RSS_{UR})/(n-k-1)} \sim F_{q,n-k}{}^{-1}$

(iv) The coefficients for a first sub-population are equal to those from the second sub-population:

$H_0 : \alpha^1 = \alpha^2, \beta_1^1 = \beta_1^2, ..., \beta_k^1 = \beta_k^2$ $\quad F = \frac{(RSS_R - (RSS_1 + RSS_2)/(k+1)}{(RSS_1 + RSS_2)/(n - 2(k+1))} \sim F_{k+1,n-2(k+1)}$

Diagnostics

HETEROSKEDASTICITY: $e_t^2 = \delta_0 + \delta_1 Z_{1t} + ... + \delta_p Z_{pt} + \eta_t$

$H_0 : \delta_1 = ... = \delta_p = 0 \quad H_1 :$ any $\delta_i \neq 0$

X²: $p = 2k \qquad Z_{1t} = X_{1t} ..., Z_{kt} = X_{kt}, Z_{k+1,t} = X_{1t}^2, ..., Z_{2k,t} = X_{kt}^2$