Econometrics – Exam 1

Summation Rules

$$\sum_{i=1}^{n} c = nc$$

$$\sum_{i=1}^{n} cx_i = c\sum_{i=1}^{n} x_i$$

$$\sum_{i=1}^{n} x_i - y_i = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i$$

$$\sum_{i=1}^{n}(ax_i - by_i) = a\sum_{i=1}^{n} x_i - b\sum_{i=1}^{n} y_i$$

$$\sum_{i=1}^{n}(x_i - y_i)^2 = \sum_{i=1}^{n} x_i^2 - 2\sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} y_i^2$$

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n(\bar{x})^2$$

$$\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i(y_i - \bar{y}) = \sum_{i=1}^{n} y_i(x_i - \bar{x}) = \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}$$

Descriptive Statistics

mean $= \boldsymbol{\bar{x}} = \frac{x_1+x_2+\cdots+x_n}{n} = \frac{\sum x_i}{n} = \frac{1}{n}\sum x_i$

expected value $= \boldsymbol{E(X)} = p_1x_1 + p_2x_2 + \cdots p_kx_k = \sum p_ix_i$

E(X+Y) = E(X) + E(Y)

E(aX+bY) = aE(X) + bE(Y)

standard deviation $= \boldsymbol{\sigma} = \sqrt{\sum(x_i - \bar{x})^2} = \sqrt{\sigma^2}$

variance $= \boldsymbol{\sigma^2} = E[(X-\mu)^2] = E(X^2) - \mu^2$

var(X+Y) = var(X) + var (Y) + 2*cov*(X,Y)

var(aX+bY) = a^2var(X) + b^2var(Y) + 2ab*cov(X,Y)

*** if X,Y are independent: take out covariance terms

cov(X,Y) $= \boldsymbol{\sigma_{XY}} = E[(X-\mu_X)(Y-\mu_Y)] = E(XY) - \mu_X\mu_Y$

corr(X,Y) $= \boldsymbol{\rho_{XY}} = \frac{cov(X,Y)}{sd(X)\,sd(Y)}$

proportional change $= \frac{\Delta x}{x_0} = \frac{x_1 - x_0}{x_0}$

percentage change $= \%\Delta x = 100\frac{\Delta x}{x_0}$

Estimators

unbiased estimator: E(X') = X

biased estimator: E(X') ≠ X

consistent: MSE → 0 as n → ∞

$\hat{\beta}_j \to \beta_j$ as $n \to \infty$

efficiency of V relative to W $= \frac{var(W)}{var(V)}$

MSE(V) = E(V – θ)^2 = var(V) + [E(V) – θ]^2

Regression

$$\boldsymbol{\hat{\beta}_1} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{cov(x_i, y_i)}{var(x_i)} = \frac{\sigma_{XY}}{\sigma_X^2} = \rho_{XY}\frac{\sigma_Y}{\sigma_X}$$

$$\boldsymbol{\hat{\beta}_0} = \bar{y} - \hat{\beta}_1\bar{x}$$

fitted value $= \boldsymbol{\hat{y}_i} = \widehat{\beta_0} + \widehat{\beta_1}x_i$

residual $= \boldsymbol{\hat{u}_i} = y_i - \hat{y}_i = y_i - \widehat{\beta_0} - \widehat{\beta_1}x_i$

$\sum \hat{u}_i = 0$

$\sum x_i\hat{u}_i = 0$

$\sum \hat{y}_i\hat{u}_i = 0$

SSR $= \sum \hat{u}_i^{\,2} = \sum(y_i - \hat{y}_i)^2 = \sum(y_i - \widehat{\beta_0} - \widehat{\beta_1}x_i)^2$

SST $= \sum(y_i - \bar{y})^2$

SSE $= \sum(\hat{y}_i - \bar{y})^2$

SST = SSE + SSR

$$\boldsymbol{R^2} = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}$$

Unbiased estimators: $E(\widehat{\beta_j}) = \beta_j$ for all j = 0, 1, 2, …, k

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum(x_i - \bar{x})^2} = \frac{\sigma^2}{SST_x}$$

$$sd(\hat{\beta}_1) = \frac{\sigma}{\sqrt{SST_x}}$$

$$Var(\hat{\beta}_0) = \frac{\sigma^2(n^{-1}\sum x_i^2)}{\sum(x_i - \bar{x})^2} = \frac{\sigma^2\ 1/n\ \sum x_i^2}{SST_x}$$

$$sd(\hat{\beta}_0) = \sqrt{var(\hat{\beta}_0)}$$

$Var(\hat{\beta}_j) = \frac{\sigma^2}{SST_j(1-R_j^2)}$ for all j (MLR)

df = *n – k – 1*

$$Var(\hat{u}) = \hat{\sigma}^2 = \frac{\sum \hat{u}_i^{\,2}}{n-k-1} = \frac{SSR}{df}$$

$$se(\hat{\beta}_j) = \frac{\hat{\sigma}}{\sqrt{SST_j(1-R_j^2)}}$$

t-Test

$$\frac{\widehat{\beta_j} - \beta_j}{sd(\widehat{\beta_j})} \sim Normal(0,1)$$

$$\boldsymbol{t_{df}} = \frac{\widehat{\beta_j} - \beta_j}{se(\widehat{\beta_j})} \sim t_{n-k-1}$$

($\sigma \to \hat{\sigma}$ = standard normal → t-distribution)

$t_{\hat{\beta}_j} = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)}$, for $H_0: \beta_j = 0$

$t = \frac{(\hat{\beta}_j - \alpha_j)}{se(\hat{\beta}_j)}$, for $H_0: \beta_j = \alpha_j$

$t = \frac{(\hat{\beta}_1 - \hat{\beta}_2)}{se(\hat{\beta}_1 - \hat{\beta}_2)}$, for $H_0: \beta_1 = \beta_2$

$$se(\hat{\boldsymbol{\beta}}_1 - \hat{\boldsymbol{\beta}}_2) = \sqrt{\left[se(\hat{\beta}_1)\right]^2 + \left[se(\hat{\beta}_2)\right]^2 - 2cov(\hat{\beta}_1, \hat{\beta}_2)}$$

p-value = $P(|T| > |t_{df}|) = 2P(T > t)$

CI: $\hat{\beta}_j \pm c \cdot se(\hat{\beta}_j)$

Common Critical Values

c	90%	95%	99%
One-Sided	1.645	1.96	2.576
Two-Sided	1.282	1.645	2.326

F-test

$H_0: \beta_i = \beta_j = \cdots = 0$

$\boldsymbol{F} = \frac{(SSR_r - SSR_{ur})}{SSR_{ur}} \cdot \frac{(n-k-1)_{ur}}{q}$, with $q = df_r - df_{ur}$

$\boldsymbol{F} = \frac{(R_{ur}^2 - R_r^2)}{1 - R_{ur}^2} \cdot \frac{(n-k-1)_{ur}}{q}$, with $q = df_r - df_{ur}$

More on Coefficients

z-score $= Z = \frac{\bar{Y}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$

beta coefficient: $\hat{b}_j = \frac{\hat{\sigma}_x}{\hat{\sigma}_y} \hat{\beta}_j$

standardized variable: $z_{x_n} = \frac{x_n - \bar{x}}{\hat{\sigma}_n}$, $z_y = \frac{y_i - \bar{y}}{\hat{\sigma}_y}$

standardized equation: $z_y = b_1 z_{x_1} + b_2 z_{x_2} + \cdots + b_k + \frac{\hat{u}_i}{\hat{\sigma_u}}$

approximate effect (log model): $\%\Delta y = 100\hat{\beta}_j$

exact effect (log model): $\%\Delta y = 100 \cdot (e^{\hat{\beta}_j} - 1)$

turning point of quadratic function: $x^* = -\frac{\hat{\beta}_1}{2\hat{\beta}_2}$

PE = partial effect of x_j on y (quadratic) $= \frac{\partial y}{\partial x_j} = \hat{\beta}_j + \hat{\beta}_k x_j$

APE = average partial effect (quadratic) $= \frac{\partial y}{\partial x_j} = \hat{\beta}_j + \hat{\beta}_k \bar{x}$

Goodness of Fit

adjusted $R^2 = \bar{\boldsymbol{R}}^2 = 1 - \frac{SSR/(n-k-1)}{SST/(n-1)} = 1 - \frac{\hat{\sigma}^2}{SST/(n-1)}$

adjusted $R^2 = \bar{\boldsymbol{R}}^2 = 1 - (1 - R^2)\frac{n-1}{n-k-1}$

Dummy Variables

Chow statistic $= \frac{SSR_p - (SSR_1 + SSR_2)}{SSR_1 + SSR_2} \cdot \frac{n - 2(k+1)}{k+1}$

SSR_1: SSR from group 1
SSR_2: SSR from group 2
SSR_p: pooled SSR

Other

Percentage change $= \frac{new\ value - old\ value}{old\ value} \cdot 100$

Variance $= \sigma^2 = \frac{\sum(x_i - \bar{x})^2}{n} = \frac{\sum x_i^2}{n} - \mu^2$

SD $= \sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}} = \sqrt{\frac{1}{n}\sum(x_i - \bar{x})^2} = \frac{\sum(x_i - \bar{x})}{\sqrt{n}}$

$\hat{\sigma}^2 = \frac{SSR}{n-2}$, Simple Regression

$\hat{\sigma}^2 = \frac{SSR}{n-k-1}$, Multiple Regression

$var(\bar{x}) = \frac{\sigma^2}{n}$