The Rules of Summation

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n$$

$$\sum_{i=1}^{n} a = na$$

$$\sum_{i=1}^{n} ax_i = a \sum_{i=1}^{n} x_i$$

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

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

$$\sum_{i=1}^{n} (a + bx_i) = na + b \sum_{i=1}^{n} x_i$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

$$\sum_{i=1}^{n} (x_i - \bar{x}) = 0$$

$$\begin{aligned}\sum_{i=1}^{2}\sum_{j=1}^{3} f(x_i, y_j) &= \sum_{i=1}^{2} [f(x_i, y_1) + f(x_i, y_2) + f(x_i, y_3)] \\ &= f(x_1, y_1) + f(x_1, y_2) + f(x_1, y_3) \\ &\quad + f(x_2, y_1) + f(x_2, y_2) + f(x_2, y_3)\end{aligned}$$

Expected Values & Variances

$$\begin{aligned}E(X) &= x_1 f(x_1) + x_2 f(x_2) + \cdots + x_n f(x_n) \\ &= \sum_{i=1}^{n} x_i f(x_i) = \sum_x x f(x)\end{aligned}$$

$$E[g(X)] = \sum_x g(x) f(x)$$

$$\begin{aligned}E[g_1(X) + g_2(X)] &= \sum_x [g_1(x) + g_2(x)] f(x) \\ &= \sum_x g_1(x) f(x) + \sum_x g_2(x) f(x) \\ &= E[g_1(X)] + E[g_2(X)]\end{aligned}$$

$E(c) = c$
$E(cX) = cE(X)$
$E(a + cX) = a + cE(X)$
$\text{var}(X) = \sigma^2 = E[X - E(X)]^2 = E(X^2) - [E(X)]^2$
$\text{var}(a + cX) = E[(a + cX) - E(a + cX)]^2 = c^2\text{var}(X)$

Marginal and Conditional Distributions

$f(x) = \sum_y f(x, y)$ for each value X can take

$f(y) = \sum_x f(x, y)$ for each value Y can take

$$f(x|y) = P[X = x|Y = y] = \frac{f(x, y)}{f(y)}$$

If X and Y are independent random variables, then $f(x,y) = f(x)f(y)$ for each and every pair of values x and y. The converse is also true.

If X and Y are independent random variables, then the conditional probability density function of X given that

$$Y = y \text{ is } f(x|y) = \frac{f(x, y)}{f(y)} = \frac{f(x)f(y)}{f(y)} = f(x)$$

for each and every pair of values x and y. The converse is also true.

Expectations, Variances & Covariances

$$\begin{aligned}\text{cov}(X, Y) &= E[(X - E[X])(Y - E[Y])] \\ &= \sum_x \sum_y [x - E(X)][y - E(Y)] f(x, y)\end{aligned}$$

$$\rho = \frac{\text{cov}(X,Y)}{\sqrt{\text{var}(X)\text{var}(Y)}}$$

$E(c_1X + c_2Y) = c_1E(X) + c_2E(Y)$
$E(X + Y) = E(X) + E(Y)$

$$\text{var}(aX + bY + cZ) = a^2\text{var}(X) + b^2\text{var}(Y) + c^2\text{var}(Z) + 2ab\text{cov}(X,Y) + 2ac\text{cov}(X,Z) + 2bc\text{cov}(Y,Z)$$

If X, Y, and Z are independent, or uncorrelated, random variables, then the covariance terms are zero and:

$$\text{var}(aX + bY + cZ) = a^2\text{var}(X) + b^2\text{var}(Y) + c^2\text{var}(Z)$$

Normal Probabilities

If $X \sim N(\mu, \sigma^2)$, then $Z = \dfrac{X - \mu}{\sigma} \sim N(0, 1)$

If $X \sim N(\mu, \sigma^2)$ and a is a constant, then

$$P(X \geq a) = P\left(Z \geq \frac{a - \mu}{\sigma}\right)$$

If $X \sim N(\mu, \sigma^2)$ and a and b are constants, then

$$P(a \leq X \leq b) = P\left(\frac{a - \mu}{\sigma} \leq Z \leq \frac{b - \mu}{\sigma}\right)$$

Assumptions of the Simple Linear Regression Model

- SR1 The value of y, for each value of x, is $y = \beta_1 + \beta_2 x + e$
- SR2 The average value of the random error e is $E(e) = 0$ since we assume that $E(y) = \beta_1 + \beta_2 x$
- SR3 The variance of the random error e is $\text{var}(e) = \sigma^2 = \text{var}(y)$
- SR4 The covariance between any pair of random errors, e_i and e_j is $\text{cov}(e_i, e_j) = \text{cov}(y_i, y_j) = 0$
- SR5 The variable x is not random and must take at least two different values.
- SR6 (*optional*) The values of e are *normally distributed* about their mean $e \sim N(0, \sigma^2)$

Least Squares Estimation

If b_1 and b_2 are the least squares estimates, then

$$\hat{y}_i = b_1 + b_2 x_i$$
$$\hat{e}_i = y_i - \hat{y}_i = y_i - b_1 - b_2 x_i$$

The Normal Equations

$$Nb_1 + \Sigma x_i b_2 = \Sigma y_i$$
$$\Sigma x_i b_1 + \Sigma x_i^2 b_2 = \Sigma x_i y_i$$

Least Squares Estimators

$$b_2 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2}$$

$$b_1 = \bar{y} - b_2\bar{x}$$

Elasticity

$$\eta = \frac{\text{percentage change in } y}{\text{percentage change in } x} = \frac{\Delta y/y}{\Delta x/x} = \frac{\Delta y}{\Delta x} \cdot \frac{x}{y}$$

$$\eta = \frac{\Delta E(y)/E(y)}{\Delta x/x} = \frac{\Delta E(y)}{\Delta x} \cdot \frac{x}{E(y)} = \beta_2 \cdot \frac{x}{E(y)}$$

Least Squares Expressions Useful for Theory

$b_2 = \beta_2 + \Sigma w_i e_i$

$$w_i = \frac{x_i - \bar{x}}{\Sigma(x_i - \bar{x})^2}$$

$\Sigma w_i = 0, \quad \Sigma w_i x_i = 1, \quad \Sigma w_i^2 = 1/\Sigma(x_i - \bar{x})^2$

Properties of the Least Squares Estimators

$$\text{var}(b_1) = \sigma^2 \left[\frac{\Sigma x_i^2}{N\Sigma(x_i - \bar{x})^2} \right] \quad \text{var}(b_2) = \frac{\sigma^2}{\Sigma(x_i - \bar{x})^2}$$

$$\text{cov}(b_1, b_2) = \sigma^2 \left[\frac{-\bar{x}}{\Sigma(x_i - \bar{x})^2} \right]$$

Gauss-Markov Theorem: Under the assumptions SR1–SR5 of the linear regression model the estimators b_1 and b_2 have the *smallest variance of all linear and unbiased estimators* of β_1 and β_2. They are the Best Linear Unbiased Estimators (BLUE) of β_1 and β_2.

If we make the normality assumption, assumption SR6, about the error term, then the least squares estimators are normally distributed.

$$b_1 \sim N\left(\beta_1, \frac{\sigma^2 \Sigma x_i^2}{N\Sigma(x_i - \bar{x})^2}\right), b_2 \sim N\left(\beta_2, \frac{\sigma^2}{\Sigma(x_i - \bar{x})^2}\right)$$

Estimated Error Variance

$$\hat{\sigma}^2 = \frac{\Sigma \hat{e}_i^2}{N - 2}$$

Estimator Standard Errors

$$\text{se}(b_1) = \sqrt{\widehat{\text{var}(b_1)}}, \quad \text{se}(b_2) = \sqrt{\widehat{\text{var}(b_2)}}$$

***t*-distribution**

If assumptions SR1–SR6 of the simple linear regression model hold, then

$$t = \frac{b_k - \beta_k}{\text{se}(b_k)} \sim t_{(N-2)}, \quad k = 1, 2$$

Interval Estimates

$P[b_2 - t_c \text{se}(b_2) \le \beta_2 \le b_2 + t_c \text{se}(b_2)] = 1 - \alpha$

Hypothesis Testing

Components of Hypothesis Tests

1. A *null* hypothesis, H_0
2. An *alternative* hypothesis, H_1
3. A test *statistic*
4. A *rejection* region
5. A conclusion

If the null hypothesis $H_0: \beta_2 = c$ is ***true***, then

$$t = \frac{b_2 - c}{\text{se}(b_2)} \sim t_{(N-2)}$$

Rejection rule for a two-tail test: If the value of the test statistic falls in the rejection region, either tail of the t-distribution, then we reject the null hypothesis and accept the alternative.

Type I error: The null hypothesis is *true* and we decide to *reject* it.

Type II error: The null hypothesis is *false* and we decide *not* to reject it.

***p*-value rejection rule:** When the p-value of a hypothesis test is *smaller* than the chosen value of α, then the test procedure leads to *rejection* of the null hypothesis.

Prediction

$y_0 = \beta_1 + \beta_2 x_0 + e_0, \; \hat{y}_0 = b_1 + b_2 x_0, \; f = \hat{y}_0 - y_0$

$$\widehat{\text{var}(f)} = \hat{\sigma}^2 \left[1 + \frac{1}{N} + \frac{(x_0 - \bar{x})^2}{\Sigma(x_i - \bar{x})^2} \right], \text{se}(f) = \sqrt{\widehat{\text{var}(f)}}$$

A $(1 - \alpha) \times 100\%$ confidence interval, or prediction interval, for y_0

$$\hat{y}_0 \pm t_c \text{se}(f)$$

Goodness of Fit

$\Sigma(y_i - \bar{y})^2 = \Sigma(\hat{y}_i - \bar{y})^2 + \Sigma \hat{e}_i^2$

$SST = SSR + SSE$

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = (\text{corr}(y, \hat{y}))^2$$

Log-Linear Model

$\ln(y) = \beta_1 + \beta_2 x + e, \; \widehat{\ln(y)} = b_1 + b_2 x$

$100 \times \beta_2 \approx \%$ change in y given a one-unit change in x.

$\hat{y}_n = \exp(b_1 + b_2 x)$

$\hat{y}_c = \exp(b_1 + b_2 x)\exp(\hat{\sigma}^2/2)$

Prediction interval:

$$\exp\left[\widehat{\ln(y)} - t_c \text{se}(f)\right], \; \exp\left[\widehat{\ln(y)} + t_c \text{se}(f)\right]$$

Generalized goodness-of-fit measure $R_g^2 = (\text{corr}(y, \hat{y}_n))^2$

Assumptions of the Multiple Regression Model

- MR1 $y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} + e_i$
- MR2 $E(y_i) = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} \Leftrightarrow E(e_i) = 0.$
- MR3 $\text{var}(y_i) = \text{var}(e_i) = \sigma^2$
- MR4 $\text{cov}(y_i, y_j) = \text{cov}(e_i, e_j) = 0$
- MR5 The values of x_{ik} are not random and are not exact linear functions of the other explanatory variables.
- MR6 $y_i \sim N[(\beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK}), \sigma^2] \Leftrightarrow e_i \sim N(0, \sigma^2)$

Least Squares Estimates in MR Model

Least squares estimates $b_1, b_2, \ldots, b_K$ minimize
$S(\beta_1, \beta_2, \ldots, \beta_K) = \Sigma(y_i - \beta_1 - \beta_2 x_{i2} - \cdots - \beta_K x_{iK})^2$

Estimated Error Variance and Estimator Standard Errors

$$\hat{\sigma}^2 = \frac{\Sigma \hat{e}_i^2}{N - K} \quad \text{se}(b_k) = \sqrt{\widehat{\text{var}(b_k)}}$$

Hypothesis Tests and Interval Estimates for Single Parameters

Use t-distribution $\quad t = \dfrac{b_k - \beta_k}{\text{se}(b_k)} \sim t_{(N-K)}$

t-test for More than One Parameter

$$H_0 : \beta_2 + c\beta_3 = a$$

When H_0 is true $\quad t = \dfrac{b_2 + cb_3 - a}{\text{se}(b_2 + cb_3)} \sim t_{(N-K)}$

$$\text{se}(b_2 + cb_3) = \sqrt{\widehat{\text{var}(b_2)} + c^2\widehat{\text{var}(b_3)} + 2c \times \widehat{\text{cov}(b_2, b_3)}}$$

Joint F-tests

To test J joint hypotheses,

$$F = \frac{(SSE_R - SSE_U)/J}{SSE_U/(N-K)}$$

To test the overall significance of the model the null and alternative hypotheses and F statistic are

$$H_0 : \beta_2 = 0,\ \beta_3 = 0, \ldots, \beta_K = 0$$

H_1 : *at least one* of the β_k is nonzero

$$F = \frac{(SST - SSE)/(K-1)}{SSE/(N-K)}$$

RESET: A Specification Test

$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + e_i \qquad \hat{y}_i = b_1 + b_2 x_{i2} + b_3 x_{i3}$

$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \gamma_1 \hat{y}_i^2 + e_i, \quad H_0 : \gamma_1 = 0$

$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \gamma_1 \hat{y}_i^2 + \gamma_2 \hat{y}_i^3 + e_i, \quad H_0 : \gamma_1 = \gamma_2 = 0$

Model Selection

$\text{AIC} = \ln(SSE/N) + 2K/N$

$\text{SC} = \ln(SSE/N) + K\ln(N)/N$

Collinearity and Omitted Variables

$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + e_i$

$$\text{var}(b_2) = \frac{\sigma^2}{(1 - r_{23}^2)\sum(x_{i2} - \bar{x}_2)^2}$$

When x_3 is omitted, $\text{bias}(b_2^*) = E(b_2^*) - \beta_2 = \beta_3 \dfrac{\widehat{\text{cov}(x_2, x_3)}}{\widehat{\text{var}(x_2)}}$

Heteroskedasticity

$\text{var}(y_i) = \text{var}(e_i) = \sigma_i^2$

General variance function

$\sigma_i^2 = \exp(\alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS})$

Breusch-Pagan and White Tests for H_0: $\alpha_2 = \alpha_3 = \cdots = \alpha_S = 0$

When H_0 is true $\quad \chi^2 = N \times R^2 \sim \chi^2_{(S-1)}$

Goldfeld-Quandt test for $H_0 : \sigma_M^2 = \sigma_R^2$ versus $H_1 : \sigma_M^2 \neq \sigma_R^2$

When H_0 is true $\quad F = \hat{\sigma}_M^2/\hat{\sigma}_R^2 \sim F_{(N_M - K_M, N_R - K_R)}$

Transformed model for $\text{var}(e_i) = \sigma_i^2 = \sigma^2 x_i$

$y_i/\sqrt{x_i} = \beta_1(1/\sqrt{x_i}) + \beta_2(x_i/\sqrt{x_i}) + e_i/\sqrt{x_i}$

Estimating the variance function

$\ln(\hat{e}_i^2) = \ln(\sigma_i^2) + v_i = \alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS} + v_i$

Grouped data

$$\text{var}(e_i) = \sigma_i^2 = \begin{cases} \sigma_M^2 & i = 1, 2, \ldots, N_M \\ \sigma_R^2 & i = 1, 2, \ldots, N_R \end{cases}$$

Transformed model for feasible generalized least squares

$$y_i \Big/ \sqrt{\hat{\sigma}_i} = \beta_1\left(1 \Big/ \sqrt{\hat{\sigma}_i}\right) + \beta_2\left(x_i \Big/ \sqrt{\hat{\sigma}_i}\right) + e_i \Big/ \sqrt{\hat{\sigma}_i}$$

Regression with Stationary Time Series Variables

Finite distributed lag model

$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_q x_{t-q} + v_t$

Correlogram

$r_k = \sum(y_t - \bar{y})(y_{t-k} - \bar{y}) / \sum(y_t - \bar{y})^2$

For $H_0 : \rho_k = 0, \quad z = \sqrt{T} r_k \sim N(0, 1)$

LM test

$y_t = \beta_1 + \beta_2 x_t + \rho\hat{e}_{t-1} + \hat{v}_t \quad$ Test $H_0 : \rho = 0$ with t-test

$\hat{e}_t = \gamma_1 + \gamma_2 x_t + \rho\hat{e}_{t-1} + \hat{v}_t \quad$ Test using $LM = T \times R^2$

AR(1) error $\quad y_t = \beta_1 + \beta_2 x_t + e_t \quad e_t = \rho e_{t-1} + v_t$

Nonlinear least squares estimation

$$y_t = \beta_1(1 - \rho) + \beta_2 x_t + \rho y_{t-1} - \beta_2 \rho x_{t-1} + v_t$$

ARDL(p, q) model

$$y_t = \delta + \delta_0 x_t + \delta_1 x_{t-1} + \cdots + \delta_q x_{t-q} + \theta_1 y_{t-1} + \cdots + \theta_p y_{t-p} + v_t$$

AR(p) forecasting model

$$y_t = \delta + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \cdots + \theta_p y_{t-p} + v_t$$

Exponential smoothing $\quad \hat{y}_t = \alpha y_{t-1} + (1 - \alpha)\hat{y}_{t-1}$

Multiplier analysis

$$\delta_0 + \delta_1 L + \delta_2 L^2 + \cdots + \delta_q L^q = (1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_p L^p) \times (\beta_0 + \beta_1 L + \beta_2 L^2 + \cdots)$$

Unit Roots and Cointegration

Unit Root Test for Stationarity: Null hypothesis:

$H_0 : \gamma = 0$

Dickey-Fuller Test 1 (no constant and no trend):

$\Delta y_t = \gamma y_{t-1} + v_t$

Dickey-Fuller Test 2 (with constant but no trend):

$\Delta y_t = \alpha + \gamma y_{t-1} + v_t$

Dickey-Fuller Test 3 (with constant and with trend):

$\Delta y_t = \alpha + \gamma y_{t-1} + \lambda t + v_t$

Augmented Dickey-Fuller Tests:

$$\Delta y_t = \alpha + \gamma y_{t-1} + \sum_{s=1}^{m} a_s \Delta y_{t-s} + v_t$$

Test for cointegration

$\Delta\hat{e}_t = \gamma\hat{e}_{t-1} + v_t$

Random walk: $\quad y_t = y_{t-1} + v_t$

Random walk with drift: $\quad y_t = \alpha + y_{t-1} + v_t$

Random walk model with drift and time trend:

$y_t = \alpha + \delta t + y_{t-1} + v_t$

Panel Data

Pooled least squares regression

$y_{it} = \beta_1 + \beta_2 x_{2it} + \beta_3 x_{3it} + e_{it}$

Cluster robust standard errors $\quad \text{cov}(e_{it}, e_{is}) = \psi_{ts}$

Fixed effects model

$y_{it} = \beta_{1i} + \beta_2 x_{2it} + \beta_3 x_{3it} + e_{it} \qquad \beta_{1i}$ not random

$y_{it} - \bar{y}_i = \beta_2(x_{2it} - \bar{x}_{2i}) + \beta_3(x_{3it} - \bar{x}_{3i}) + (e_{it} - \bar{e}_i)$

Random effects model

$y_{it} = \beta_{1i} + \beta_2 x_{2it} + \beta_3 x_{3it} + e_{it} \qquad \beta_{it} = \bar{\beta}_1 + u_i$ random

$y_{it} - \alpha\bar{y}_i = \bar{\beta}_1(1 - \alpha) + \beta_2(x_{2it} - \alpha\bar{x}_{2i}) + \beta_3(x_{3it} - \alpha\bar{x}_{3i}) + v_{it}^*$

$$\alpha = 1 - \sigma_e \Big/ \sqrt{T\sigma_u^2 + \sigma_e^2}$$

Hausman test

$$t = (b_{FE,k} - b_{RE,k}) \Big/ \left[\widehat{\text{var}(b_{FE,k})} - \widehat{\text{var}(b_{RE,k})}\right]^{1/2}$$