Formula Sheet

Expected Values, Variances, Correlation

$$E(c) = c$$

$$E(cx) = cE(x)$$

$$E(a + cx) = a + cE(x)$$

$$E(x + y) = E(x) + E(y)$$

$$E(c_1 x + c_2 y) = c_1 E(x) + c_2 E(y)$$

$$var(x) = \sigma^2 = E(x - E(x))^2$$

$$std(x) = \sigma = \sqrt{E(x - E(x))^2}$$

$$var(a + cx) = c^2 var(x)$$

$$cov(x, y) = E[(x - E(x))(y - E(y))]$$

$$corr(x, y) = \rho = \frac{cov(x,y)}{\sqrt{var(x)var(y)}}$$

$$P(y = y_1 | x = x_1) = \frac{P(x=x_1, y=y_1)}{p(X=x_1)}$$

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

$$var(\bar{Y}) = \frac{\sigma_Y^2}{n}$$

$$std(\bar{Y}) = \frac{\sigma_Y}{\sqrt{n}}$$

$$s_y^2 = \frac{1}{n-1} \sum_{i=1}^{N} (y_i - \bar{y})^2$$

$$s_y = \sqrt{\frac{1}{n-1} \sum_{i=1}^{N} (y_i - \bar{y})^2}$$

$$SE(\bar{y}) = \frac{s_y}{\sqrt{n}}$$

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

$$r_{xy} = \frac{s_{xy}}{s_x s_y}$$

Logarithms

$x = \ln(e^x)$

$\frac{d\ln(x)}{dx} = \frac{1}{x}$

$\ln(1/x) = -\ln(x)$

$\ln(ax) = \ln(a) + \ln(x)$

$\ln(x/a) = \ln(x) - \ln(a)$

$\ln(x^a) = a\ln(x)$

$\ln(x+\Delta x) - \ln(x) \approx \frac{\Delta x}{x}$ (approximately equal for small Δx)

Calculus

x^* that maximizes (minimizes) a strictly concave (convex) function, $f(x)$, solves $\frac{df(x)}{dx} = 0$

OLS Estimator

$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{s_{XY}}{s_X^2}$

$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$

$\sigma^2_{\hat{\beta}_1} = \frac{1}{n}\frac{var((X_i - \mu_X)u_i))}{(var(X_i))^2}$

$\sigma^2_{\hat{\beta}_0} = \frac{1}{n}\frac{var(H_i u_i)}{(E(H_i^2))^2}$; where $H_i = 1 - (\frac{\mu_X}{E(X_i^2)})X_i$

$\hat{\beta}_1 \rightarrow \beta_1 + \rho_{Xu}\frac{\sigma_u}{\sigma_X}$

Hypothesis Testing

Difference in means from different populations

$H_0 : \mu_w - \mu_m = d_0; \quad vs. \;\; H_1 : \mu_w - \mu_m \neq d_0$

$SE(\bar{Y}_w - \bar{Y}_m) = \sqrt{s_w^2/n_w + s_m^2/n_m}$

$t^{act} = \frac{(\bar{Y}_w - \bar{Y}_m) - d_0}{SE(\bar{Y}_w - \bar{Y}_m)}$

Linear Regression

$t^{act} = \frac{\hat{\beta}_1 - \beta_{1,0}}{SE(\hat{\beta}_1)}$

$H_0 : \beta_1 = \beta_{1,0}$ vs. $H_1 : \beta_1 \neq \beta_{1,0}$, p-value $= 2\Phi(-|t^{act}|)$

$H_0 : \beta_1 = \beta_{1,0}$ vs. $H_1 : \beta_1 < \beta_{1,0}$, p-value $= \Phi(t^{act})$

$H_0 : \beta_1 = \beta_{1,0}$ vs. $H_1 : \beta_1 > \beta_{1,0}$, p-value $= 1 - \Phi(t^{act})$

t^α is the critical value for a two-sided test with α significance level

$\alpha = 2\Phi(-|t^\alpha|)$

$(1-\alpha)$ CI: $[\hat{\beta}_1 - t^\alpha SE(\hat{\beta}_1), \hat{\beta}_1 + t^\alpha SE(\hat{\beta}_1)]$

For testing means, replace β with μ_X and $\hat{\beta}$ with $\bar{X}$

Joint hypotheses

$H_0 : \beta_j = \beta_{j,0}, \; \beta_m = \beta_{m,0}, \ldots$ for a total of q restrictions

H_1 : one or more of the q restrictions under H_0 does not hold

the F-statistic is distributed $F_{q,n-k-1}$

p-value $= \Pr[F_{q,n-k-1} > F^{act}] = 1 - G(F^{act}; q, n-k-1)$

$$F = \frac{1}{2}\left(\frac{(t_1^{act})^2 + (t_2^{act})^2 - 2\hat{\rho}_{t_1^{act},t_2^{act}} t_1^{act} t_2^{act}}{1 - \hat{\rho}_{t_1^{act},t_2^{act}}}\right) \text{ if } q = 2$$

$$F^{act} = \frac{(SSR_{restricted} - SSR_{unrestricted})/q}{SSR_{unrestricted}/(n-k-1)} = \frac{(R^2_{unrestricted} - R^2_{restricted})/q}{(1 - R^2_{unrestricted})/(n-k-1)}$$

Goodness of Fit

$$SSR = \sum_{i=1}^{n} \hat{u}_i^2$$
$$ESS = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2$$
$$TSS = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$
$$R^2 = \frac{ESS}{TSS} = 1 - \frac{SSR}{TSS}$$
$$SER = s_{\hat{u}} = \sqrt{s_{\hat{u}}^2},\ s_{\hat{u}}^2 = \frac{SSR}{n-k-1}$$
$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1}\frac{SSR}{TSS} = 1 - \frac{s_{\hat{u}}^2}{s_Y^2}$$

Nonlinear and Time Series Regression

$$E[Y|X_1, X_2, \ldots, X_k] = f(X_1, X_2, \ldots, X_k)$$

$$\Delta\hat{Y} = \hat{f}(X_1 + \Delta X_1, X_2, \ldots, X_k) - \hat{f}(X_1, X_2, \ldots, X_k)$$

$$SE(\Delta\hat{Y}) = \frac{|\Delta\hat{Y}|}{\sqrt{F}}$$

$$(1-\alpha)\ \text{CI: } [\Delta\hat{Y} - t^{\alpha} SE(\Delta\hat{Y}), \Delta\hat{Y} + t^{\alpha} SE(\Delta\hat{Y})]$$

$$\text{RMSFE} = \sqrt{E[(Y_{T+1} - \hat{Y}_{T+1|T})^2]}$$

$$SE(Y_{T+1} - \hat{Y}_{T+1|T}) = \widehat{RMSFE} = \sqrt{var(\hat{u}_t)} = SER$$

$$(1-\alpha)\ \text{CI: } [\hat{Y}_{T+1|T} - t^{\alpha} \times SE(Y_{T+1} - \hat{Y}_{T+1|T}), \hat{Y}_{T+1|T} + t^{\alpha} \times SE(Y_{T+1} - \hat{Y}_{T+1|T})]$$

$$\text{BIC}(K) = \ln\left[\frac{SSR(K)}{T}\right] + K\frac{\ln(T)}{T}$$

$$\text{AIC}(K) = \ln\left[\frac{SSR(K)}{T}\right] + K\frac{2}{T}$$