Useful formulas for econometrics

This document contains various formulas useful in econometrics courses. This sheet can not replace a full formula collection but some of the better known formulas used in econometrics are listed.

1 Random variables

X, Y: random variables. a, b, c: constants

$$P(X \in A, Y \in B) = \int_A \int_B f(x,y)dydx$$

Expectation

$$E[X] = \int_{-\infty}^{\infty} xdF(x)$$

$$E[X] = \begin{cases} \int_{-\infty}^{\infty} xf(x)dx & \text{if } X \text{ is continuous} \\ \sum_x xp(x) & \text{if } X \text{ is discrete.} \end{cases}$$

$$E[h(x)] = \int_{-\infty}^{\infty} h(x)dF(x)$$

Variance

$$\mathrm{var}(X) = E\left[(X - E[X])^2\right]$$

$$\mathrm{var}(X+Y) = \mathrm{var}(X) + 2\mathrm{cov}(X,Y) + \mathrm{var}(Y)$$

Covariance

$$\mathrm{cov}(X,Y) = E\left[(X - E[X])(Y - E[Y])\right]$$

$$\mathrm{cov}(X+Y,Z) = \mathrm{cov}(X,Z) + \mathrm{cov}(Y,Z)$$

$$\mathrm{cov}(a,X) = 0$$

$$\mathrm{cov}(aX+bY,Z) = a\mathrm{cov}(X,Z) + b\mathrm{cov}(Y,Z)$$

If x is normally distributed with mean μ and variance σ, then

$$E[e^x] = e^{\mu + \frac{1}{2}\sigma^2}$$

X, Y: random variables. a, b: constants.

2 Basic statistics

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

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

$$S_x = \sqrt{S_x^2}$$

$$\widehat{\mathrm{cov}(x,y)} = \widehat{\sigma_{x,y}} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

$$\widehat{\mathrm{corr}(x,y)} = \frac{\widehat{\sigma_{x,y}}}{S_x S_y}$$

Notation: n: number of observations. x: data

3 Univariate linear regression model

$$y_i = a + bx_i + e_i$$

$$\hat{b} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\mathrm{cov}(x,y)}{\mathrm{var}(x)}$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$

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

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

$$RSS = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{a} + \hat{b}x_i - \bar{y})^2$$

$$R^2 = 1 - \frac{SSR}{TSS} = \frac{RSS}{TSS}$$

$$S^2 = \frac{1}{n}\sum_{i=1}^{n} e_i^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - (a + bx_i))^2$$

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

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

$$s_a = \sqrt{\frac{1}{n}\frac{\left(\sum_{i=1}^{n} e_i^2\right)\left(\sum_{i=1}^{n} x_i^2\right)}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

$$s_b = \sqrt{\frac{1}{n}\frac{\left(\sum_{i=1}^{n} e_i^2\right)}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

Finite sample correction: $n-1$ and $n-2$ instead of n.

Notation: n: number of observations. x, y: data

4 Matrix algebra

Null matrix

$$\mathbf{0} = \begin{bmatrix} 0 & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & 0 \end{bmatrix}$$

$$\mathbf{A} + \mathbf{0} = \mathbf{A}$$

Identity matrix

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \\ 0 & & & & 0 & 1 \end{bmatrix}$$

$$\mathbf{AI} = \mathbf{A}$$

Matrix Multiplication.
Let $\mathbf{A} = [a_{ij}]$ and $\mathbf{B} = [b_{ij}]$.

$$\mathbf{AB} = [c_{ij}] \quad c_{ij} = \sum_r a_{ir} b_{rj}$$

If $\mathbf{A}$ is $p \times q$ and $\mathbf{B}$ is $q \times t$, then $\mathbf{C} = \mathbf{AB}$ is $p \times q$.

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$

$$(\mathbf{A}+\mathbf{B})(\mathbf{C}+\mathbf{D}) = \mathbf{A}(\mathbf{C}+\mathbf{D}) + \mathbf{B}(\mathbf{C}+\mathbf{D})$$

Transpose: $\mathbf{A}'$: Exchange rows for columns in $\mathbf{A}$

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$$

Inverse

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

The *rank* $r(\mathbf{A})$ of a matrix $\mathbf{A}$ is the number of linearly independent rows (columns) in $\mathbf{A}$.

$$r(\mathbf{AB}) \leq \min(r(\mathbf{A}), r(\mathbf{B}))$$

A matrix $\mathbf{A}$ is *symmetric* if $\mathbf{A} = \mathbf{A}'$
A matrix $\mathbf{A}$ is *idempotent* if $\mathbf{AA} = \mathbf{A}$.

5 Kroenecker product

$$A = \begin{pmatrix} a_{11} & a_{21} & & a_{m1} \\ a_{12} & a_{22} & & a_{m2} \\ & & \ddots & \\ a_{1n} & & & a_{mn} \end{pmatrix}$$

$$B = \begin{pmatrix} b_{11} & b_{21} & & b_{p1} \\ b_{12} & b_{22} & & b_{p2} \\ & & \ddots & \\ b_{1q} & & & b_{pq} \end{pmatrix}$$

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{21}B & & a_{m1}B \\ a_{12}B & a_{22}B & & a_{m2}B \\ & & \ddots & \\ a_{1n}B & & & a_{mn}B \end{pmatrix}$$

6 Quadratic forms

$$\mathbf{x}'\mathbf{A}\mathbf{x}$$

A matrix $\mathbf{A}$ is *positive (semi) definite* if

$$\mathbf{x}'\mathbf{A}\mathbf{x} > \ (\geq)\ 0 \ \forall\ \mathbf{x} \neq \mathbf{0}$$

7 Eigenvalues

$$\lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}$$

solves

$$|\mathbf{A} - \lambda\mathbf{I}| = 0$$

8 Optimization

$$f(\mathbf{x}) = f(x_1, x_2, \cdots, x_n)$$

Gradient

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

9 Probability distributions

Binomial distriution

$$f(x; n, p) = \begin{pmatrix} n \\ x \end{pmatrix} p^x (1-p)^{n-x}$$

Normal distribution

$$n(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

Chi Square distribution with k degrees of freedom

$$f(x) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k-2}{2}} e^{-\frac{x}{2}}$$

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$$

10 OLS estimation

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$

$$\widehat{\mathbf{b}}_{\mathbf{n}}^{\mathbf{ols}} = (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{y}$$

$$S^2 = \frac{1}{n}(\mathbf{e}'\mathbf{e})$$

$$\text{var}(\hat{b}^{ols}) = S^2(\mathbf{X}'X)^{-1}$$

$$SSR = (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b})$$

$$TSS = (\mathbf{y} - \overline{\mathbf{y}})'(\mathbf{y} - \overline{\mathbf{y}})$$

$$R^2 = 1 - \frac{SSR}{TSS}$$

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

Under normality:

$$\widehat{\mathbf{b}}_n^{ols} \sim N\left(\mathbf{b}, S^2(X'X)^{-1}\right)$$

$$S^2 \sim \left(\sigma^2, \frac{2\sigma^4}{n-k}\right)$$

Notation: k - number of parameters, n number of data.

11 Testing in OLS settings

Wald principle

$$W = (\mathbf{R}\widehat{\mathbf{b}} - \mathbf{r})'(\mathbf{R}\mathbf{V}\mathbf{R}')^{-1}(\mathbf{R}\widehat{\mathbf{b}} - \mathbf{r})$$

Lagrange Multiplier principle

$$(\mathbf{R}\lambda)'(\sigma^2\mathbf{X}'\mathbf{X})^{-1}(\mathbf{R}\lambda)$$

$$(\mathbf{y} - \mathbf{X}\tilde{\mathbf{b}})'\mathbf{X}'(\sigma^2\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}(\mathbf{y} - \mathbf{X}\tilde{\mathbf{b}}))$$

Likelihood Ratio principle

$$LR = \frac{\frac{1}{r}\left(SSE\left(\tilde{b}\right) - SSE\left(\hat{b}\right)\right)}{\frac{1}{n-k}SSE\left(\hat{b}\right)}$$

Chow test

$$Chow = \frac{(RSS - (RSS_1 + RSS_2))/k}{(RSS_1 + RSS_2)/(T - 2k)}$$

Test statistics
$\sim F(r, n-k)$ in small samples
$\sim \chi^2(r)$ in large samples.

Notation: $\tilde{b}$: restricted estimate. $\hat{b}$: unrestricted estimate. r: number of restrictions. k number of parameters. n: number of observations.

12 WLS

$$\mathbf{e} \sim N(0, \mathbf{\Omega})$$

$$\hat{\mathbf{b}}^{wls} = \left(\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X}\right)^{-1}\left(X'\mathbf{\Omega}^{-1}\mathbf{y}\right)$$

13 Time series models (Box-Jenkins)

Autocovariance function, $s = 1, 2, \cdots$

$$\gamma_s = E\left[(y_t - E[y_t])(y_{t-s} - E[y_{t-s}])\right]$$

Autocorrelation function (ACF) $s = 0, 1, 2, \cdots$

$$\tau_s = \frac{\gamma_s}{\gamma_0}$$

A process u_t is white noise if it satisfies the following conditions

$$E[u_t] = \mu$$

$$\text{var}(u_t) = \sigma^2$$

$$\gamma_{t-r} = \begin{cases} \sigma^2 & \text{if } t = r \\ 0 & \text{otherwise} \end{cases}$$

A 95% confidence interval for τ_l:

$$\left[-1.96\frac{1}{\sqrt{T}}, +1.96\frac{1}{\sqrt{T}}\right]$$

Box-Pierce statistic

$$Q_m = T\sum_{k=1}^{m} \hat{\tau}_k{}^2$$

Ljung-Box statistic

$$Q_m^* = T(T+2)\sum_{k=1}^{m}\frac{\hat{\tau_k}^2}{T-k}$$

$AR(p)$

$$y_t = \mu + \sum_{i=1}^{p}\phi_i y_{t-i} + u_t$$

$MA(q)$

$$y_t = \mu + \sum_{i=1}^{q}\theta u_{t-i} + u_t$$

$ARMA(p,q)$

$$y_t = \mu + \sum_{i=1}^{p}\phi_i y_{t-i} + \sum_{i=1}^{q}\theta u_{t-i} + u_t$$

Conditions for covariance stationarity:

$$E[y_t] = \mu \quad \forall t$$

$$E[(y_t-\mu)^2] = \sigma^2 < \infty \quad \forall t$$

$$E\left[(y_{t_1}-\mu)(y_{t_2}-\mu)\right] = \gamma_{t_2-t_1} < \infty \quad \forall t_1, t_2$$

14 Forecasting

Root Mean Square Error

$$RMSE = \sqrt{\frac{1}{n^0}\sum_i (y_i - \hat{y}_i)^2}$$

Mean Absolute error

$$MAE = \frac{1}{n^0}\sum_i |y_i - \hat{y}_i|$$

Theil U Statistic

$$U = \sqrt{\frac{(1/n^0)\sum_i (y_i-\hat{y}_i)^2}{(1/n^0)\sum_i y_i^2}}$$

Notation: y_i data, $\hat{y}_i$ forecast, n^0: number of periods forecasted.

15 Stationarity, unit roots and cointegration

Trend stationary process

$$y_t = \alpha + \beta t + u_t$$

Nonstationary model, random walk with drift

$$y_t = \mu + y_{t-1} + u_t$$

Dickey-Fuller test

$$y_t = \beta_0 + \gamma t + \rho y_{t-1} + \varepsilon_t$$

Augmented Dickey Fuller test

$$\Delta y_t = \beta_0 + \gamma t + \rho y_{t-1} + \sum_{j=1}^{k}\beta_j \Delta y_{t-j} + \varepsilon_t$$

Notation: t: time, Δ: first difference, u_t and ε_t: noise terms.

16 ARCH/GARCH

$ARCH(p)$

$$\sigma_t^2 = \sigma_0 + \sum_{i=1}^{p}\alpha_i u_{t-i}^2$$

$GARCH(p,q)$

$$\sigma_t^2 = \sigma_0 + \sum_{i=1}^{p}\alpha_i u_{t-i}^2 + \sum_{j=1}^{p}\phi_j \sigma_{t-j}^2$$

17 Maximum Likelihood

18 GMM

Moment condition

$$E[h(\theta_0, w_t)] = 0$$

$$\mathcal{Y}_T = \{w_1, w_2, \ldots, w_T\}$$

$$g(\theta, \mathcal{Y}_T) = \frac{1}{T}\sum_{t=1}^{T} h(\theta, w_t)$$

Exactly identified

$$g(\hat{\theta}, \mathcal{Y}_T) = 0$$

Overidentified

$$\hat{\theta} = \arg\min_{\theta} J(\theta, \mathcal{Y}_T) = \arg\min_{\theta} g(\theta, \mathcal{Y}_T)' W_T g(\theta, \mathcal{Y}_T)$$

$$W_T = \widehat{S}^{-1}$$

$$\widehat{S} = \text{var}\left(\sum_t h(\theta, w_t)' h(\theta, w_t)\right)$$

$$J(\theta) = g_T(\theta, y) S^{-1} g_T(\theta, y)$$

Notation: θ: parameters, w_t: data.