ECON 468 Cheat Sheet
Zachary Probst Page:1

1 Basics

Distributions

Normal: $\phi(x) = (2\pi)^{-1/2}\exp(-\frac{1}{2}x^2)$

Normal CDF: $\mathbf{\Phi}(\mathbf{x}) = \int_{-\infty}^{x}\phi(y)dy$

Moments

Continuous: $m_k(X) \equiv \int_{-\infty}^{\infty} x^k f(x)dx$

Central:
$\mu_k \equiv E(X - E(X)^k) = \int_{-\infty}^{\infty}(x-\mu)^k f(x)dx$

Discrete Central:
$\mu_k \equiv E(X - E(X)^k) = \sum_{i=1}^{m} p(x)(x_i - \mu)^k$

Multivariate

Joint Density Func: $f(x_1, x_2) = \frac{\partial^2 F(x_1,x_2)}{\partial x_1 \partial x_2}$

Independence:
$F(x_1, x_2) = F(x_1, \infty)F(\infty, x_2)$
or $f(x_1, x_2) = f(x_1)f(x_2)$

Marginal Density:
$f(x_1) \equiv F_1(x_1, \infty) = \int_{-\infty}^{\infty} f(x_1, x_2)dx_2$

Conditional Density: $f(x_1|x_2) = \frac{f(x_1,x_2)}{f(x_2)}$

Law Iterate Expec: $E(E(X_1|X_2)) = E(X_1)$

Any Deterministic Func h:
$E(X_1 h(X_2) \mid X_2) = h(X_2)E(X_1 \mid X_2)$

Matrix Algebra

Dot Product: $\mathbf{a}^\top\mathbf{b} = \sum_{i=1}^{n} a_i b_i$

Matrix Multiplication: $C_{ij} = \sum_{k=1}^{m} A_{ik}B_{ki}$

Transpose: $(\mathbf{AB})^\top = \mathbf{B}^\top\mathbf{A}^\top$

$\|\mathbf{x}\| = (\mathbf{x}^\top\mathbf{y})^{1/2} = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$

$\langle \mathbf{x}, \mathbf{y}\rangle = \|\mathbf{x}\|\|\mathbf{y}\|\cos(\theta)$

Cauchy-Schwartz: $\langle \mathbf{x}, \mathbf{y}\rangle \le \|\mathbf{x}\|\|\mathbf{y}\|$

Linear Indep: $\mathbf{X}^\top\mathbf{X}\mathbf{b} = \mathbf{0}$, $\mathbf{b} \ne \mathbf{0}$

Singular: $\exists \mathbf{x} \ne 0 : \mathbf{Ax} = \mathbf{0}$

2 Linear Regression

OLS

Estimator: $\hat\beta = (\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{y}$

$E(x_{ti}u_t) = 0 \implies \mathbf{X}^\top(\mathbf{y} - \mathbf{X}\beta) = \mathbf{0}$

SSR(β): $= \sum_{t=1}^{n}(y_i - \mathbf{X}_t\beta)^2$

$\mathbf{y}^\top\mathbf{y} = \hat\beta^\top\mathbf{X}^\top\mathbf{X}\hat\beta + (\mathbf{y} - \mathbf{X}\hat\beta)^\top(\mathbf{y} - \mathbf{X}\hat\beta)$

Projection: $\mathbf{P_X} = \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top$

$\mathbf{M_X} = \mathbf{I} - \mathbf{P_X} = \mathbf{I} - \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top$

$\mathbf{P_X}\mathbf{y} = \mathbf{X}((\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{y})$; $\mathbf{P_X}\mathbf{P_X} = \mathbf{P_X}$

$\mathbf{M_X}\mathbf{y} = \hat{\mathbf{u}}$; $\mathbf{M_X}\mathbf{X} = \mathbf{0}$; $\mathbf{M_X}\mathbf{M_X} = \mathbf{M_X}$

$\mathbf{P_X} + \mathbf{M_X} = \mathbf{I}$; $\mathbf{P_X}\mathbf{M_X} = \mathbf{0}$; $\mathbf{P_X}^\top = \mathbf{P_X}$

$\|\mathbf{y}\|^2 = \|\mathbf{P_X}\mathbf{y}\|^2 + \|\mathbf{M_X}\mathbf{y}\|^2$; $\|\mathbf{P_X}\mathbf{y}\| \le \|\mathbf{y}\|$

Centering: $\mathbf{M}_\iota\mathbf{x} = \mathbf{z} = \mathbf{x} - \bar{x}\iota$; $\iota^\top\mathbf{z} = 0$

$\mathbf{P_1} \equiv \mathbf{P_{X_1}}$; $\mathbf{P_1}\mathbf{P_X} = \mathbf{P_X}\mathbf{P_1} = \mathbf{P_1}$

FWL: β_2 from $\mathbf{y} = \mathbf{X_1}\beta_1 + \mathbf{X_2}\beta_2 + \mathbf{u}$ and $\mathbf{M_1}\mathbf{y} = \mathbf{M_1}\mathbf{X_2}\beta_2 + \text{res}$ are the same. (+ res)

Seasonal w const: $\mathbf{s}_i' = \mathbf{s}_i - \mathbf{s}_4$, $i = 1, 2, 3$.

Avg is const coeff. $\mathbf{M_S}\mathbf{y}$ is deseasonalized.

$\beta^{(t)} - \hat\beta = -1\backslash 1 - h_t(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X_t}^\top\hat{u_t}$ where h_t denotes the t[th] diagonal element of $\mathbf{P_X}$.

Bias

vector of (true) model params: θ

Bias: $E(\hat\theta) - \theta_0$, $E((\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{u}) = 0$

estimating eq: $g(\mathbf{y}, \theta) = 0$ unbiased iff $\forall \mu \in \mathbb{M}, E_\mu g(\mathbf{y}, \theta_\mu) = \mathbf{0}$ or $E(\mathbf{X}^\top\mathbf{u}) = \mathbf{0}$

X exogenous $\implies E(\mathbf{u} \mid \mathbf{X}) = \mathbf{0}$ and both $\hat\beta$ and estimating equations unbiased.

Make exog assumpt in cross-sec not time.

regressors predetermined: $E(\mathbf{X}^\top\mathbf{u}) = \mathbf{0}$

Stochastic Limits

Converg in prob: $\lim Pr(|Y_n - Y_\infty| > \epsilon) = 0 \Rightarrow \text{plim}\, Y_n = Y_\infty \implies$ converg dist

Converg dist: $\lim F_n(x) = F(x) \equiv Y_n \to F$

LLN: $\text{plim}\overline{Y_n} = \text{plim}\frac{1}{n}\sum_t^n Y_t = \mu_Y$, Y_t IID, $\overline{Y_n}$ sample mean of Y_t, μ_T pop mean.

LLN2: $\text{plim}\frac{1}{n}\sum_t^n Y_t = \lim\frac{1}{n}\sum_t^n E(Y_t)$

$\text{plim}\, Y_n Z_n = \text{plim}\, Y_n \text{plim}\, Z_n$ if converg

$\mathbf{X}^\top\mathbf{X}$ may not have plim so mult by $1\backslash n$.

consistent: $\text{plim}_\mu\hat\beta = \beta_\mu$, may be bias

$E(\mu_t \mid \mathbf{X}_t) = 0 \implies \hat\beta$ consistent.

Covariance and Precision Matrices

$\text{Cov}(b_i, b_j) \equiv E((b_i - E(b_i))(b_j - E(b_j)))$

if $i = j$, $\text{Cov}(b_i, b_j) = \text{Var}(b_i)$

$\text{Var}(\mathbf{b}) \equiv E((\mathbf{b} - E(\mathbf{b}))(\mathbf{b} - E(\mathbf{b}))^\top)$

when $E(\mathbf{b}) = \mathbf{0}$, $\text{Var}(\mathbf{b}) = E(\mathbf{bb}^\top)$ b_i, b_j indep: $\text{Cov}(b_i, b_j) = 0$, converse false

correlation: $\rho(b_i, b_j) \equiv \frac{\text{Cov}(b_i,b_j)}{(\text{Var}(b_i)\text{Var}(b_j))^{1\backslash 2}}$

$\text{Var}(\mathbf{b})$ positive semidefinite. cov and corr matrix positive definite most of the time.

positive definite: $\mathbf{x}^\top\mathbf{Ax} > 0$ for $\mathbf{x} \in k \times 1$.

$\mathbf{x}^\top\mathbf{Ax} = \sum_i^k\sum_j^k x_i x_j A_{ij}$. If $\ge 0 \Rightarrow$ semidef.

Any $\mathbf{B}^\top\mathbf{B}$ is pos semidef. If full col rank then pos def. pos def $\Rightarrow$ diag > 0 & non-singular. (pos def)$^{-1}$ $\exists$ & is pos def.

Precision mtrx: invers of cov mtrx of estmatr. $\exists$ & pos def iff cov mtrx pos def.

If μ IID w Var σ^2 and cov of any pair $= 0$: $\text{Var}(\mathbf{u}) = E(\mathbf{uu}^\top) = \sigma^2\mathbf{I}$. If false, $\mathbf{\Omega}$ = err cov mtrx. If diag of $\mathbf{\Omega}$ differ, heteroskedastic. Homoskedastic: all μ same Var. Autocorrelated: off-diag $\mathbf{\Omega} \ne 0$.

$\hat\beta$ unbiased & $\mathbf{\Omega} = \sigma^2\mathbf{I}$ so no hetero or autocorr, then $\text{Var}(\hat\beta) = \sigma^2(\mathbf{X}^\top\mathbf{X})^{-1}$.

Precision affected by n, σ^2, X.

Collinearity: precision for β_1 dep on $\mathbf{X_2}$.

Efficiency

$\tilde\beta$ more efficient than $\hat\beta$ iff $\text{Var}(\tilde\beta)^{-1} - \text{Var}(\hat\beta)^{-1}$ is nonzero pos semidef mtrx.

Gauss-Markov: If $E(\mathbf{u} \mid \mathbf{X}) = \mathbf{0}$ and $E(\mathbf{uu}^\top \mid \mathbf{X}) = \sigma^2\mathbf{I}$ then OLS $\hat\beta$ is BLUE (best linear unbiased estimator). Not necessary that μ normally distributed.

Residuals & Disturbances

$\hat\mu = \mathbf{M_X}\mathbf{u}$ (hat resid, μ dist).

If $E(\mathbf{u} \mid \mathbf{X}) = \mathbf{0} \Rightarrow E(\|\hat{\mathbf{u}}\|^2) \le E(\|\mathbf{u}\|^2)$

$\text{Var}(\hat{u_t}) < \sigma^2$; $\hat\sigma^2 \equiv \frac{1}{n}\sum_t^n \hat{u}_t^2$

$E(\hat\sigma^2) = \frac{n-k}{n}\sigma^2$

$E(\mathbf{u}^\top\mathbf{M_X}\mathbf{u}) = E(\text{SSR}(\hat\beta)) = (n-k)\sigma^2$

unbiased: $s^2 \equiv \frac{1}{n-k}\sum_t^n u_t^2$; s = std err.

unbias est of $\text{Var}(\hat\beta)$: $\widehat{\text{Var}}(\hat\beta) = s^2(\mathbf{X}^\top\mathbf{X})^{-1}$

s^2 unbiased and consistent.

$\text{MSE}(\tilde\beta) \equiv E((\tilde\beta - \beta_0)(\tilde\beta - \beta_0)^\top)$

I $\tilde\beta$ unbiased $\text{MSE}(\tilde\beta) = \text{Var}(\tilde\beta)$.

Measures of Goodness of Fit

$R_u^2 = \frac{\text{ESS}}{\text{TSS}} = \frac{\|\mathbf{P_X y}\|^2}{\|\mathbf{y}\|^2} = \cos^2\theta$, where θ angle between $\mathbf{y}$ and $\mathbf{P_X y}$. $0 \le R_u^2 \le 1$.

R_c^2: center all vars first. Invalid if $\iota \notin \mathcal{S}(\mathbf{X})$.

$R_c^2 = 1 - \sum_t^n \hat{u}_t^2 \backslash \sum_t^n (y_t - \bar{y})^2$.

Adj R^2: unbiased estimators. maybe < 0.

$\overline{R}^2 \equiv 1 - \frac{\frac{1}{n-k}\sum_t^n \hat{u_t}^2}{\frac{1}{n-1}\sum_t^n (y_t - \bar{y})^2} = 1 - \frac{(n-1)\mathbf{y}^\top\mathbf{M_X y}}{(n-k)\mathbf{y}^\top\mathbf{M}_\iota\mathbf{y}}$

$\overline{R}^2$ does not always $\uparrow$ in regressors.

Hypothesis Testing

If u_t normal, and σ known, test $\beta = \beta_0$ w

$z = \frac{\hat\beta - \beta_0}{(\text{Var}(\hat\beta))^{1\backslash 2}} = \frac{n^{1\backslash 2}}{\sigma}(\hat\beta - \beta_0)$, $z \sim N(0, 1)$

NCP: $\lambda = \frac{n^{1\backslash 2}}{\sigma}(\beta_1 - \beta_0)$, $\beta_1 \ne \beta_0$

Reject null if z large enough. 2-tail: $|z|$.

Type 1: reject true null, 2: accept false

left-tail $\Phi(-c_\alpha) = \alpha\backslash 2$, $c_\alpha = \Phi^{-1}(\alpha\backslash 2)$.

$\Phi^{-1}(.975) = 1.96$. Power: prob test rejects the null. Prob of Type $2 = 1 - P(\text{power})$.

Power $\uparrow$ with $(\beta_1 - \beta_0)\uparrow$ or $\sigma\downarrow$ or $n\uparrow$.

$p(z) = 2(1 - \Phi(|z|))$

$x \sim N(\mu, \sigma^2) \Rightarrow z = (x - \mu)\backslash\sigma$, $z \sim N(0, 1)$.

Lin comb of rand vars that are jointly multivariate normal must be $\sim N$. If $\mathbf{x}$ multivar norm with 0 cov, componenets of $\mathbf{x}$ are mutually indep.

χ^2: $y \equiv \|\mathbf{z}\|^2 = \mathbf{z}^\top\mathbf{z} = \sum_i^m z_i^2$., $y \sim \chi^2(m)$ with $\mathbf{z} \sim N(\mathbf{0}, \mathbf{I})$; $E(y) = m$. $\text{Var}(y) = 2m$.

$y_1 \sim \chi^2(m_1)$ & $y_2 \sim \chi^2(m_2)$ indep $\Rightarrow$ $y_1 + y_2 \sim \chi^2(m_1 + m_2)$

$m \times 1$ $\mathbf{x} \sim N(\mathbf{0}, \mathbf{\Omega})$, then $\mathbf{x}^\top\mathbf{\Omega}^{-1}\mathbf{x} \sim \chi^2(m)$

If $\mathbf{P}$ $n \times n$ w rank $r < n$ and $\mathbf{z} \sim N(\mathbf{0}, \mathbf{I})$ then $\mathbf{z}^\top\mathbf{P}\mathbf{z} \sim \chi^2(r)$.

$z \sim N(0, 1)$ & $y \sim \chi^2(m)$, z, y indep, then $t \equiv z \backslash (y \backslash m)^{1\backslash 2}$. Or $t \sim t(m)$. Only first $m - 1$ moments exist. Cauchy: $t(1)$.

$\text{Var}(t) = m\backslash(m-2)$. $t(m)$ tends to std norm.

y_1, y_2 indep rand var $\sim \chi^2(m_1)$ & $\chi^2(m_2)$, then $F \equiv \frac{y_1\backslash m_1}{y_2\backslash m_2}$. $F \sim F(m_1, m_2)$. As $m_2 \to \infty$, $F \sim 1 \backslash m_1$ times $\chi^2(m_1)$. $t \sim t(m_2) \Rightarrow t^2 \sim F(1, m_2)$.