Notation & Useful Facts

- **Uncentered 2nd Moments:** Matrices A, B: $\Sigma_{AB} \overset{\text{def}}{=} \mathbb{E}[AB]$
- **Sample 2nd Moments:** Matrices A, B: $S_{AB} \overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^n a_i b_i$
- **Sample average:** $\bar{a} \overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^n a_i$
- **BLP:** $\mathbb{E}[Y|X = x] = x'\beta$, where $\beta = \arg\min_b \mathbb{E}[(Y - X'b)^2] = \Sigma_{XX}^{-1}\Sigma_{XY}$.
- **Model:** $y_i = \mathbb{E}[y_i|x_i] + y_i - \mathbb{E}[y_i|x_i] = x_i'\beta + \epsilon_i$. Stack this: $Y = X\beta + \epsilon$, where Y vector n, X matrix $n \times k$, ϵ vector n.

- $\boldsymbol{P_X} = X(X'X)^{-1}X'$: $P_X' = P_X, P_X^2 = P_X$, rank/trace k, eigenvalues 0 or 1. $P_X X = X$.
- $\boldsymbol{M_X} = I_n - P_X$: $M_X' = M_X, M_X^2 = M_X$, rank/trace $n-k$, eigenvalues 0 or 1. $M_X X = 0 \implies M_X P_X = 0$.
- **Residual:** $\hat{\epsilon} = Y - X\hat{\beta} = M_X Y = M_X \epsilon = \epsilon + x(\beta - \hat{\beta})$

- **Fit** $\hat{Y} = P_X Y$
- $\boldsymbol{SSR} = \hat{\epsilon}'\hat{\epsilon} = Y'Y - Y'P_X Y$, $\boldsymbol{ESS} = \hat{Y}'\hat{Y} = Y'P_X Y$, $\boldsymbol{TSS} = Y'Y = SSR + ESS$
- **R-squared:** $R^2 = \frac{ESS}{TSS} = 1 - \frac{SSR}{TSS}$, $\bar{R}^2 = 1 - \frac{n}{n-k}(1 - R^2)$

- **FWL**: For $Y = X_1\beta_1 + X_2\beta_2 + \epsilon$, $\hat{\beta}_i = (X_i' M_{X_j} X_i)^{-1} X_i' M_{X_j} Y$, for $i \neq j \in \{1, 2\}$.

- $\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'(X\beta + \epsilon) = \beta + (X'X)^{-1}X'\epsilon$, useful for unbiased and cons.
- **Moment function:** $g_i = x_i\epsilon_i = x_i(y_i - x_i'\beta)$, $\mathbb{E}[g_i] = 0$ at BLP β.
- **Approx. Dist:** If $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}(0, V)$, then $\hat{\theta} \overset{a}{\sim} \mathcal{N}\left(\theta, \frac{1}{n}V\right)$.
- **Lagrangian:** For $\max_\theta f(\theta)$ s.t. $g(\theta) = 0$, $\mathcal{L}(\theta, \lambda) = f(\theta) + \lambda g(\theta)$.

$\mathcal{N}$ & Related Distributions

- **MVN** $X \in \mathbb{R}^p$ MVN $\iff$ $a'X$ is univariate normal $\forall\ 0 \neq a \in \mathbb{R}^p$. Write $X \sim \mathcal{N}_p(\mu, \Sigma)$. $f_X(x) = (2\pi)^{-p/2}|\Sigma|^{-1/2}\exp\left(-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right)$ if $\text{rank}(\Sigma) = p$. Then $\mathbb{E}[a'X] = a'\mu$, $\text{Var}[a'X] = a'\Sigma a$, $M_X(a) = \mathbb{E}[e^{a'X}] = \exp(a'\mu + \frac{1}{2}a'\Sigma a)$.

- **Linear map (A):** $X \sim \mathcal{N}_p(\mu, \Sigma)$, $Y = \eta + BX$ (B is $k \times p$) $\Rightarrow Y \sim \mathcal{N}_k(\eta + B\mu,\ B\Sigma B')$.
- **Density (B):** if $\text{rank}(\Sigma) = p$, $f_X(x) = (2\pi)^{-p/2}|\Sigma|^{-1/2}\exp\left(-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right)$.
- **Independent $\mathcal{N}$s (C):** $X_1 \sim \mathcal{N}_p(\mu_1, \Sigma_1)$, $X_2 \sim \mathcal{N}_q(\mu_2, \Sigma_2)$, $X_1 \perp\!\!\!\perp X_2 \Rightarrow (X_1', X_2')' \sim \mathcal{N}_{p+q}\left(\binom{\mu_1}{\mu_2}, \begin{pmatrix}\Sigma_1 & 0 \\ 0 & \Sigma_2\end{pmatrix}\right)$.
- **Sums (F):** $X_1 \perp\!\!\!\perp X_2$, $X_i \sim \mathcal{N}_p(\mu_i, \Sigma_i)$ $\Rightarrow X_1 + X_2 \sim \mathcal{N}_p(\mu_1 + \mu_2, \Sigma_1 + \Sigma_2)$.
- **Partition:** $X = (X_1', X_2')'$, $\mu = (\mu_1', \mu_2')'$, $\Sigma = \begin{pmatrix}\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}$.
  - **Marginal (G):** $X_1 \sim \mathcal{N}(\mu_1, \Sigma_{11})$.
  - **Cond. (D):** $(X_1|X_2 = x_2) \sim \mathcal{N}(\mu_1 + A,\ B)$ w/ $A = \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$, $B = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.
  - **Indep (H):** $\Sigma_{12} = 0 \iff X_1 \perp\!\!\!\perp X_2$.
- **Indep. of lin. comb. (I):** $X \sim \mathcal{N}_p(\mu, \Sigma)$, $B \in \mathbb{R}^{p \times k}$, $C \in \mathbb{R}^{p \times m}$: $B'X \perp\!\!\!\perp C'X \iff B'\Sigma C = 0$.
- **Converse construction (E):** if $X_2 \sim \mathcal{N}(\mu_2, \Sigma_{22})$ and $X_1|X_2 = x_2 \sim \mathcal{N}(A + Bx_2, \Omega)$ (constant A, B, Ω), then $(X_1', X_2')'$ is MVN with $\boldsymbol{\mu} = (A + B\mu_2,\ \mu_2)$ and
$$\text{Var}\begin{pmatrix}X_1 \\ X_2\end{pmatrix} = \begin{pmatrix}B\Sigma_{22}B' + \Omega & B\Sigma_{22} \\ \Sigma_{22}B' & \Sigma_{22}\end{pmatrix}.$$
- **Quadratic form:** $Y'AY$; WLOG take $A = A'$.
- **Orthonormal projection:** if $Z \sim \mathcal{N}(0, I_n)$ and $P \in \mathbb{R}^{n \times m}$ with $P'P = I_m$, then $P'Z \sim \mathcal{N}(0, I_m)$ and $Z'PP'Z \sim \chi^2_m$.
- **Mahalanobis (J):** if $X \sim \mathcal{N}_p(\mu, \Sigma)$, $\text{ra}(\Sigma) = p$, then $(X-\mu)'\Sigma^{-1}(X-\mu) \sim \chi^2_p$.
- **Idempotent (K):** if $Z \sim \mathcal{N}(0, I_p)$ and $M^2 = M$ with $\text{ra}(M) = k$, then $Z'MZ \sim \chi^2_k$.
- **Independence (L):** $X = PZ$, $Q = Z'AZ$, and $PA = 0 \Rightarrow X \perp\!\!\!\perp Q$.
- **Independence (M):** $Q_1 = Z'A_1Z$, $Q_2 = Z'A_2Z$, and $A_1A_2 = 0 \Rightarrow Q_1 \perp\!\!\!\perp Q_2$.

Normal OLS Finite-Sample

- **Assumptions:** (1) i.i.d. sample, (2) $y|x \sim \mathcal{N}(x'\beta, \sigma^2)$, (3) X full rank a.s., (4) $f_X(x) \perp\!\!\!\perp \beta, \sigma^2$. These imply homoskedasticity.

- $\boldsymbol{\mathcal{L}}$: $f(Y, X) = f(Y|X)f(X) \propto (\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}(Y - X\beta)'(Y - X\beta)\right)$
- $\hat{\boldsymbol{\beta}}_{\boldsymbol{MLE}} = \arg\max_\beta \mathcal{L}(\beta, \sigma^2|Y, X) = (X'X)^{-1}X'Y = S_{XX}^{-1}S_{XY} \xrightarrow{p} \Sigma_{XX}^{-1}\Sigma_{XY}$
- $\hat{\boldsymbol{\sigma}}^2_{\boldsymbol{MLE}}$: $\hat{\sigma}^2 = \frac{1}{n}\hat{\epsilon}'\hat{\epsilon} = S_{\hat{\epsilon}\hat{\epsilon}} = \frac{SSR}{n}$
- $\text{Var}[\hat{\beta}|X] = \frac{\sigma^2}{(X'X)} \overset{\text{def}}{=} V_{\hat{\beta}}, \hat{\beta} \sim \mathcal{N}(\beta, V_{\hat{\beta}})$
- $\hat{\sigma}^2$ biased. Unbiased: $s^2 = \frac{\hat{\epsilon}'\hat{\epsilon}}{n-k}$. $\text{Var}[\hat{\sigma}^2] = \text{Var}\left[\frac{\sigma^2}{n}\left(\frac{SSR}{\sigma^2}\right)\right] = \frac{2\sigma^4}{n^2}(n-k) \Leftarrow \frac{SSR}{\sigma^2} = \left(\frac{1}{\sigma}\epsilon\right)' M_X \left(\frac{1}{\sigma}\epsilon\right) \sim \chi^2_{n-k} \Rightarrow \text{Var}[s^2] = \frac{2\sigma^4}{n-k}$.
- $\hat{\beta} \perp\!\!\!\perp \hat{\sigma}^2$, since $(X'X)^{-1}X'M_X = 0$. This allows for t-tests and F-tests.
- **F-Stat.** $A = \frac{(\hat{\beta}-\beta)'(\hat{\beta}-\beta)}{(\sigma^2(X'X)^{-1})} \sim \chi^2_k$ and $B = \frac{SSR}{\sigma^2} \sim \chi^2_{n-k} \implies F = \frac{A/k}{B/(n-k)} \sim F_{k,n-k}$ and $F = \frac{(\hat{\beta}-\beta)'(X'X)(\hat{\beta}-\beta)}{ks^2} \sim F_{k,n-k}$ and $F = \frac{\hat{\beta}'(X'X)\hat{\beta}(n-k)}{SSRk} = \frac{ESS}{SSR}\frac{(n-k)}{k} \sim F_{k,n-k}$
- **t-Stat.** $t = \frac{\hat{\beta}_j - \beta_j}{SE(\hat{\beta}_j)} \sim t_{n-k}$, where $SE(\hat{\beta}_j)$ is the sqrt of j-th diagonal element of $s^2(X'X)^{-1}$.
- $\mathcal{I}(\beta, \sigma^2) \implies \hat{\beta}_{MLE}$ efficient. $\mathbb{E}[\sigma^2] \neq \sigma^2 \implies \hat{\sigma}^2_{MLE}$ not efficient. s^2 unbiased w/ $\text{Var}[s^2] = \frac{2}{n-k}\sigma^2$, asymptotically efficient.
- **B:** $f(\beta, \sigma^2|Y, X) \propto \mathcal{L}(\beta, \sigma^2|Y, X)\pi(\beta, \sigma^2)$. Joint posteriors can be derived numerically by drawing σ^2 prior, then drawing β conditional on σ^2 and repeating many times.
- **w/o normality:** (1) i.i.d. sample, (2) $\mathbb{E}[y|x] = x'\beta$, (3) X full rank a.s., (4) $\text{Var}[y|x] = \sigma^2$.
- $\mathbb{E}[\hat{\beta}|X] = \beta$, $\text{Var}[\hat{\beta}|X] = \sigma^2(X'X)^{-1}$, $\mathbb{E}[\hat{\sigma}^2] = \frac{n-k}{n}\sigma^2$, $\mathbb{E}[s^2] = \sigma^2$.
- **Gauss-Markov:** *With* $\mathbb{E}[y|x] = x'\beta$ *and* $\text{Var}[y|x] = \sigma^2$, *OLS is the BLUE of* β. Proof: Any LUE $\tilde{\beta} = CY$ with $CX = I_k$. $D := C - (X'X)^{-1}X'$, then $DX = 0$. Then, $\text{Var}\left[\tilde{\beta}|X\right] - \text{Var}\left[\hat{\beta}|X\right] = \sigma^2(CC' - (X'X)^{-1}) = \sigma^2(DD' + DX(X'X)^{-1} + (X'X)^{-1}(XD)') = \sigma^2 DD' \geq 0$, Since DD' is PSD for any D.
- **GLS:** If $\text{Var}[\epsilon|X] = \sigma^2\Omega$, then $\hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y$. Prove by $\Omega = HH'$, transform model by H^{-1}, then apply OLS $\implies$ efficiency, unbiasedness, cons. etc.

OLS Asymptotic Properties

- **Assumptions:** (1) i.i.d. sample, (2) $\beta = \Sigma_X X^{-1}\Sigma_{XY}$ is BLP and Σ_{XX} is non-singular, (3) fourth moments exist for (y, x). These imply $\mathbb{E}[x\epsilon] = \mathbb{E}[g] = 0$ and existance of $\Sigma_{gg} = \text{Var}[g] = \mathbb{E}[\epsilon^2 xx']$.
- **Properties::** $\hat{\beta} - \beta = S_{XX}^{-1}\bar{g} \implies \hat{\beta} \xrightarrow{p} \beta$ by WLLN. $\sqrt{n}\left(\hat{\beta} - \beta\right) = S_{XX}^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^n g_i \xrightarrow{d} \mathcal{N}(0, V)$ with $V = \Sigma_{XX}^{-1}\Sigma_{gg}\Sigma_{XX}^{-1}$ by Lindeberg-Feller CLT and Slutsky's theorem. Hence, $\hat{\beta} \overset{a}{\sim} \mathcal{N}\left(\beta, \frac{1}{n}V\right)$.
- **Wald:** $\left(\hat{\beta} - \beta\right)'\left(\frac{1}{n}V\right)^{-1}\left(\hat{\beta} - \beta\right) \xrightarrow{d} \chi^2_k$. $\delta \overset{\text{def}}{=} R\beta$, then $n(R\hat{\beta} - \delta)'(RVR')^{-1}(R\hat{\beta} - \delta) \xrightarrow{d} \chi^2_m$ and $\frac{(\hat{\beta}_j - \beta_j)}{\sqrt{n^{-1}(V)_{jj}}} \xrightarrow{d} \mathcal{N}(0, 1)$, holds for cons. $\hat{V}$
- $\hat{\boldsymbol{V}}$: Under homoskedasticity $\hat{\Sigma}_{gg} = \hat{\sigma}^2 S_{XX}$, then $\hat{V} = \hat{\sigma}^2(X'X)^{-1} \xrightarrow{p} V$, by WLLN.
- **Huber-White $\hat{\boldsymbol{V}}$:** $\hat{\Sigma}_{gg} = \frac{1}{n}\sum_{i=1}^n \hat{\epsilon}_i^2 x_i x_i'$, then $\hat{V} = S_{XX}^{-1}\hat{\Sigma}_{gg}S_{XX}^{-1} \xrightarrow{p} V$, by WLLN.

Misc OLS/Causality

- **Nonlinear** functions can be handled by adding quadratic terms to the regresion equation.
- **OVB:** True $\hat{\beta}_1 \xrightarrow{p} \beta_1 + \Sigma_{X_1X_1}^{-1}\Sigma_{X_1X_2}\beta_2$. No OVB if $\beta_2 = 0$ (irrelevant variable) or $\Sigma_{X_1X_2} = 0$ (exogenous variable).
- **Multicol.:** When X nearly singular $\text{Var}\left[\hat{\beta}\right]$ increases drastically.
- **Cluster SE** $\hat{\Sigma}_{gg} = \frac{1}{C}\sum_{c=1}^C \hat{g}_c\hat{g}_c'$ w/ $\hat{g}_c = \frac{1}{n}\sum_{i \in c} x_i\hat{\epsilon}_i \implies \hat{V} = S_{XX}^{-1}\hat{\Sigma}_{gg}S_{XX}^{-1} \xrightarrow{p} V$.

- **ATE(X):** $X \perp\!\!\!\perp (Y(1), Y(0)) \implies \mathbb{E}[T] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)]$. Using OLS: $\mathbb{E}[Y|X=x] = \mathbb{E}[Y(0)] + x\mathbb{E}[T] + \epsilon$, where $X \perp\!\!\!\perp (Y(1), Y(0)) \iff \mathbb{E}[\epsilon|X] = 0$. When x is continous, interpret $\mathbb{E}[T]$ as marginal effect. When $X \perp\!\!\!\perp (Y(1), Y(0))|Z$, incl. Z in model (if $\frac{\partial \mathbb{E}[T]}{\partial Z} = 0$) and incl. interac. (if $\frac{\partial \mathbb{E}[T]}{\partial Z} \neq 0$), derive coef. via FWL. $\text{Var}[T] = \text{Var}[Y(1) - Y(0)] = \text{Var}[Y(1)] + \text{Var}[Y(0)] - 2\,\text{Cov}[Y(1), Y(0)]$.

GMM/IV

- **GMM** relies on $\mathbb{E}[g_i(\theta)] = 0$, where $g_i(\theta)$ is vector function of data and param. Given data, can solve for θ by minimizing $\hat{\theta} = \arg\min_\theta \bar{g}(\theta)' W \bar{g}(\theta)$, w/ W PSD weights and $\bar{g}(\theta) = \frac{1}{n}\sum_{i=1}^n g_i(\theta)$. Sometimes useful to incl. scaling factor in objective fn.

- **IV** $y_i = z_i'\delta + \epsilon_i$, $z_i = x_i'\beta + v_i$, w/ z_i $\ell \times 1$ & x_i $k \times 1$. **(A1)** (y_i, z_i, x_i) i.i.d., **(A2)** $\mathbb{E}[x\epsilon] = \mathbb{E}[g] = 0$, **(A3)** rel. moments exist and $\mathbb{E}[g'g] = \Sigma_g g$, **(A4)** Σ_{XZ} rank ℓ.

$\ell = k$

- $\hat{\boldsymbol{\delta}}_{\boldsymbol{IV}} = S_{XZ}^{-1} S_{XY} = \delta + S_{XZ}^{-1}\bar{g}$.
- **Cons.** $\bar{g} \xrightarrow{p} 0 \implies \hat{\delta}_{IV} \xrightarrow{p} \delta$
- **AN** $\sqrt{n}(\hat{\delta}_{IV} - \delta) = S_{XZ}^{-1}\sqrt{n}\bar{g} \xrightarrow{d} \mathcal{N}(0, V_{\hat{\delta}})$
- $\boldsymbol{V}_{\hat{\boldsymbol{\delta}}} = \Sigma_{XZ}^{-1}\Sigma_{gg}\Sigma_{XZ}^{-1}$. $\hat{\Sigma}_{gg} = \frac{1}{n}\sum_{i=1}^n \hat{\epsilon}_i^2 x_i x_i' \xrightarrow{p} \Sigma_{gg}$, then $\hat{V}_{\hat{\delta}} = S_{XZ}^{-1}\hat{\Sigma}_{gg}S_{XZ}^{-1} \xrightarrow{p} V_{\hat{\delta}}$.

$\ell > k$

- Derrived by $\arg\min_d J(d) = n\bar{g}(d)' W \bar{g}(d)$
- $\hat{\boldsymbol{\delta}}_{\boldsymbol{IV}} = (S_{ZX} W S_{XZ})^{-1} S_{XZ} W S_{XY} = \delta + (S_{ZX} X S_{XZ})^{-1} S_{XZ} W \bar{g}$.
- **Cons.** $\bar{g} \xrightarrow{p} 0 \Rightarrow \hat{\delta} \xrightarrow{p} \delta$.
- **AN** $\sqrt{n}(\hat{\delta} - \delta) \xrightarrow{d} \mathcal{N}(0, V_{\hat{\delta}})$,
- $\boldsymbol{V}_{\hat{\boldsymbol{\delta}}} = (\Sigma_{ZX} W \Sigma_{XZ})^{-1}(\Sigma_{XZ} W \Sigma_{gg} W \Sigma_{ZX})$ $(\Sigma_{ZX} W \Sigma_{XZ})^{-1}$, $\hat{\Sigma}_{gg} = \frac{1}{n}\sum_{i=1}^n \hat{\epsilon}_i^2 x_i x_i'$, $\hat{V}_{\hat{\delta}} = (S_{ZX} W S_{XZ})^{-1}(S_{XZ} W \hat{\Sigma}_{gg} W S_{ZX})$ $(S_{ZX} W S_{XZ})^{-1} \xrightarrow{p} V_{\hat{\delta}}$.
- $W_{opt} = \Sigma_{gg}^{-1}$, $\hat{V}_{opt} = (S_{ZX}\hat{\Sigma}_{gg}^{-1} S_{XZ})^{-1}$.
- **J-Test**: $J(\hat{\delta}) = n\bar{g}(\hat{\delta})'\hat{W}\bar{g}(\hat{\delta}) \xrightarrow{d} \chi^2_{\ell-k}$, w/ $\hat{W} \xrightarrow{p} W$. Test assumption $\mathbb{E}[g_i] = 0$.

Misc IV

- **2SLS:** Under homosk. $\text{Var}[g] = \sigma^2\Sigma_{XX} \Rightarrow \hat{\delta}_{2SLS} = (Z'P_X Z)^{-1} Z' P_X Y$, derived via opt GMM or 2SLS. Control fn approach: add first stage residuals to 2nd stage regression.
- **Weak instruments:** First stage $F < 10$ or $t < \sqrt{10} = 3.16 \Rightarrow$ weak instrueumnt. Local-to-zero asymptotics: $\hat{\pi}_1 - \delta_0\hat{\pi}_2 \sim \mathcal{N}\left(\pi_1 - \delta_0\pi_2, n^{-1}(\sigma_1^2 + \delta_0^2\sigma_2^2 - 2\delta_0\sigma_{12})\right) \implies \tilde{\xi} = \frac{n(\hat{\pi}_1 - \delta_0\hat{\pi}_2)^2}{\sigma_1^2 + \delta_0^2\sigma_2^2 - 2\delta_0\sigma_{12}}$, reject for large $\tilde{\xi}$.
- **AR reformulation:** $y - z\delta_0 = z(\delta - \delta_0) + \epsilon = x\gamma + e$, test $H_0 : \delta = \delta_0$ via OLS of $y - z\delta_0$ on x, CIs formed by inverting test.
- **LATE:** With model $y = z_i\delta_i + \epsilon_i$, $z_i = x_i\pi_i + v_i$, $\hat{\delta}_{IV} \xrightarrow{p} \mathbb{E}[w_i\delta_i]$ w/ $w_i = \frac{\pi_i}{\mathbb{E}[\pi_i]}$. Never-takers ($\pi_i = 0$) and always-takers ($\pi_i = 0$) get weight 0, compliers and defiers get $|w_i| > 0$. Defiers ruled out by monotonicity assumption.

Time Series

- **Autocovariance** $\lambda_{t,k} = \text{Cov}[Y_t, Y_{t-k}]$, **Autocorrelation** $\rho_{t,k} = \frac{\lambda_{t,k}}{\lambda_{t,t}}$.
- **Covariance stationarity** $\iff \mathbb{E}[Y_t] = \mu_t = \mu$ and $\lambda_{t,k} = \text{Cov}[Y_t, Y_{t-k}] = \lambda_k \forall\ t$. Implies: $\lambda_k = \lambda_{-k}, \forall t, \rho_{t,k} = \rho_k = \rho_{-k}$.
- Y_t is **white noise** if $\mu = 0$ and $\lambda_k = 0 \forall\ |k| > 0$. Y_t is **Martingale** if $\mathbb{E}[Y_t|\Omega_{t-1}] = Y_{t-1}$, **Martingale difference** (md) if $\mathbb{E}[Y_t|\Omega_{t-1}] = 0$, where Ω_{t-1} is information set at $t-1$. Y_t is **random walk** if $Y_t - Y_{t-1}$ is white noise.
- **AR(1)** $Y_t = \phi Y_{t-1} + \epsilon_t$, where $\epsilon_t \sim iid(0, \sigma^2)$. $\lambda_{t,0} = \phi^{2t}\sigma_0^2 + \frac{1-\phi^2}{1-\phi^2}\sigma^2 \forall |\phi| \neq 1$, $\lambda_{t,0} = \sigma_0^2 + t\sigma^2$ for $|\phi| = 1$, $\mu_t = \phi^t\mu_0$. AR(1) is time invariant if $|\phi| < 1$ and $\mu_0 = 0$, $\sigma_0^2 = \frac{\sigma^2}{1-\phi^2}$. These imply $\lambda_{t,k} = \frac{\phi^k}{1-\phi^2}\sigma^2$ and $\rho_k = \phi^k$.
- **VAR(1)** $Y_t = \Phi Y_{t-1} + \epsilon_t$, where Y_t is $m \times 1$, Φ is $m \times m$, $\epsilon_t \sim iid(0, \Sigma)$. VAR(1) is time invariant if all eigenvalues of Φ have abs. value < 1 and $\mu_0 = 0$, Λ_0 solves $\Lambda_0 = \Phi\Lambda_0\Phi' + \Sigma$ and $\text{vec}\,\Lambda_0 =$
- **AR(p)** analyzed as VAR(1), where $Z_t = [Y_t, Y_{t-1}, \ldots, Y_{t-p+1}]'$, Φ is $p \times p$ w/ first row $[\phi_1, \phi_2, \ldots, \phi_p]$ and subdiagonal of ones, and $e_t = [\epsilon_t, 0, \ldots, 0]' \Rightarrow Z_t = \Phi Z_{t-1} + e_t$. Stationarity see above. $\mathbb{E}[Z_t] = \frac{c}{1-\sum_{i=1}^p \phi_i}$ if $Y_t = c + \sum_{i=1}^p \phi_i Y_{t-i} + \epsilon_t$.
- **Yule-Walker Equations:** For cov. stationary AR(p), $\lambda_k = \sum_{i=1}^p \phi_i\lambda_{k-i} + \sigma^2$ for $k = 0$, and $\lambda_k = \sum_{i=1}^p \phi_i\lambda_{k-i}$ for $k \geq 1$.
- **L** is lag operator, $LY_t = Y_{t-1}$, $L^kY_t = Y_{t-k}$. AR(p): $\phi(L)Y_t = \epsilon_t$, where $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p$.
- **MA(q)** $Y_t = \epsilon_t - \theta_1\epsilon_{t-1} - \ldots - \theta_q\epsilon_{t-q}$, where $\epsilon_t \sim iid(0, \sigma^2)$. In lag operator: $Y_t = \theta(L)\epsilon_t$, where $\theta(L) = 1 - \theta_1 L - \ldots - \theta_q L^q$. $\mathbb{E}[Y_t] = 0, \text{Var}[[]\, Y_t] = \sigma^2(1 + \sum_{i=1}^q \theta_i^2)$, $\lambda_k = \sigma^2(-\theta_k + \sum_{i=1}^{q-k}\theta_i\theta_{i+k})$ for $0 \leq k \leq q$, $\lambda_k = 0$ for $k > q \Rightarrow$ cov. stationarity. Invertible if roots of $\theta(L) = 0$ lie outside unit circle.
- **ARMA(p,q)** $\phi(L)Y_t = \theta(L)\epsilon_t$. Cov. stationarity if roots of $\phi(L) = 0$ lie outside unit circle. Invertible if roots of $\theta(L) = 0$ lie outside unit circle. **ARIMA(p,q,d)** $X_t \sim$ ARMA(p,q) then $X_t = (1-L)^d Y_t$ and $Y_t \sim$ ARIMA(p,q,d).
- **Wold's Theorem:** A covariance stationary process Y_t can be written as $Y_t = \theta(L)\epsilon_t$ for some $\theta(L)$ with roots outside unit circle and $\epsilon_t \sim iid(0, \sigma^2)$.
- **Ergodicity** $\approx$ elements are asymptotically independent. **Ergodic Theorem** If $\{z_t\}$ is ergodic and stationary with $\mathbb{E}[z_t] = \mu$, then $\frac{1}{T}\sum_{t=1}^T z_t \xrightarrow{p} \mu$. If z_t ergodic, then $x_t = f(z_t)$ is ergodic for any measurable f.
- **MDS CLT** If mds $\{z_t\}$ is ergodic and stationary with $\mathbb{E}[z_t] = \mu$ and $\text{Var}[[]\, z_t] = \sigma^2$, then $\frac{1}{\sqrt{T}}\sum_{t=1}^T (z_t - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$.
- **Linear Model w/ serial cor.** Assumptions (1) $Y = X\beta + \varepsilon$ (2) $\{y_t, x_t\}$ stationary and ergodic (3) $\mathbb{E}[\varepsilon_t x_t] = \mathbb{E}[g_t] = 0$ (4) Moments exist with Σ_{XX} non-singular (5) $\{g_t\}$ is mds with $\mathbb{E}[g_t g_t'] = \Sigma_{gg}$.
- **Properties** under (1)-(5), $\hat{\beta} \xrightarrow{p} \beta$ and $\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{d} \mathcal{N}(0, V)$ with $V = \Sigma_{XX}^{-1}\Sigma_{gg}\Sigma_{XX}^{-1}$. Suppose $\hat{\Sigma}_{gg} \xrightarrow{p} \Sigma_{gg}$, then $\hat{V} = S_{XX}^{-1}\hat{\Sigma}_{gg}S_{XX}^{-1} \xrightarrow{p} V$.
- **Testing** $t_j. = \left(\sqrt{T}(\hat{\beta}_j - \beta_j)\right) / \sqrt{(\hat{V}_{\hat{\beta}})_{jj}} \xrightarrow{d} \mathcal{N}(0,1)$, $\xi_W = T(R\hat{\beta} - r)'(RVR')^{-1}(R\hat{\beta} - r) \xrightarrow{d} \chi^2_m$, where R is $m \times k$ for $H_0 : R\beta = r$. $\frac{\xi_W}{m} \xrightarrow{d} F_{m,\infty}$
- Assume $\mathbb{E}[(x_{t,i}x_{t,j})^2] < \infty\ \forall i,j$ + (1)-(5), then $S_{\hat{g}\hat{g}} \xrightarrow{p} \Sigma_{gg}$, where $\hat{g}_t = \hat{\varepsilon}_t x_t = (y_t - \hat{\beta}' x_t)x_t$.