FOUNDATIONS

Sigma-field. $\mathcal{F} \subseteq 2^{\Omega}$: $\Omega \in \mathcal{F}$; $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$; $A_i \in \mathcal{F} \Rightarrow \bigcup_{i\geq 1} A_i \in \mathcal{F}$.
Measure. $\mu : \mathcal{F} \to [0,\infty]$; $\mu(\varnothing) = 0$; pairwise disjoint A_i: $\mu(\bigcup_i A_i) = \sum_i \mu(A_i)$; $\mu(\Omega) = 1$: probability measure.
R.V. $X : (\Omega, \mathcal{F}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$; X measurable iff $X^{-1}(B) \in \mathcal{F}\ \forall B \in \mathcal{B}(\mathbb{R}^d)$. Real X: $\{X \leq x\} \in \mathcal{F}\ \forall x \in \mathbb{R}$.
MCT / DCT. For measurable f_n: MCT: $0 \leq f_n \uparrow f$ μ-a.e. $\Rightarrow \int f_n\, d\mu \uparrow \int f\, d\mu$.
DCT: $f_n \to f$ a.e., $|f_n| \leq h$, $h \geq 0$ integrable $\Rightarrow \int |f_n - f|\, d\mu \to 0$ and $\int f_n\, d\mu \to \int f\, d\mu$.
Change of variables. $S_Y = g(S_X)$. Scalar: g strict mono. C^1, $g' \neq 0$: $f_Y(y) = f_X(g^{-1}(y))|(g^{-1})'(y)| = f_X(g^{-1}(y))/|g'(g^{-1}(y))|$. Vector $d \to d$: $g : S_X \to S_Y$ C^1 diffeom., $J_g = \partial g/\partial x'$, $\det J_g \neq 0$; $f_Y(y) = f_X(g^{-1}(y))|\det J_{g^{-1}}(y)| = f_X(g^{-1}(y))/|\det J_g(g^{-1}(y))|$ on S_Y; 0 o.w.

COMMON LAWS

Pois(λ). $\lambda > 0$, $k \in \mathbb{N}_0$: $p_k = e^{-\lambda}\lambda^k/k!$; $\mathbb{E}X = \mathrm{Var}\,X = \lambda$; $M_X(t) = e^{\lambda(e^t-1)}$.
$N(\mu, \sigma^2)$. $\sigma > 0$; $f(x) = e^{-(x-\mu)^2/(2\sigma^2)}/(\sigma\sqrt{2\pi})$; $M_X(t) = e^{\mu t + \sigma^2 t^2/2}$.
Exp(λ). $\lambda > 0$: $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$; $F(x) = 1 - e^{-\lambda x}$;
$(\mathbb{E}X, \mathrm{Var}\,X) = (\lambda^{-1}, \lambda^{-2})$; $M_X(t) = \lambda/(\lambda - t)$, $t < \lambda$.
Gamma(r, λ). Shape $r > 0$, rate $\lambda > 0$: $f(x) = \lambda^r x^{r-1} e^{-\lambda x}/\Gamma(r)$, $x > 0$;
$(\mathbb{E}X, \mathrm{Var}\,X) = (r/\lambda, r/\lambda^2)$; $M_X(t) = (\lambda/(\lambda - t))^r$, $t < \lambda$.
χ^2_k. $k \in \mathbb{N}$; $\chi^2_k = \mathrm{Gamma}(k/2, 1/2)$; $(\mathbb{E}X, \mathrm{Var}\,X) = (k, 2k)$; $M_X(t) = (1-2t)^{-k/2}$, $t < 1/2$.

MGF / CF / CUMULANTS

MGF / CF. $M_X(t) = \mathbb{E}e^{tX}$; $M_X^{(r)}(0) = \mathbb{E}X^r$ if finite near 0; $\varphi_X(t) = \mathbb{E}e^{itX}$ always exists. Equality of MGFs on a common nbhd. of 0, or CFs on $\mathbb{R}$, implies equal dist.
Cumulants. M_X finite on a nbhd. of 0; $K_X = \log M_X$, $\kappa_j = K_X^{(j)}(0)$. $\kappa_1 = \mu$, $\kappa_2 = \sigma^2$, $\kappa_3 = \mathbb{E}(X-\mu)^3$, $\kappa_4 = \mathbb{E}(X-\mu)^4 - 3\sigma^4$.
$X \perp Y \Rightarrow \kappa_j(X+Y) = \kappa_j(X) + \kappa_j(Y)$.
Taylor at 0. $e^x = \sum_{j=0}^{\infty} x^j/j!$; $\log(1+x) = \sum_{j=1}^{\infty} (-1)^{j+1} x^j/j$, $|x| < 1$.

INEQUALITIES

L^p norm. $p \geq 1$: $\|X\|_p := (\mathbb{E}|X|^p)^{1/p}$; $p, q > 1$, $p^{-1} + q^{-1} = 1$.
Jensen. X, $g(X) \in L^1$; g convex: $g(\mathbb{E}X) \leq \mathbb{E}g(X)$; concave: reverse.
Hölder / CBS. $\mathbb{E}|XY| \leq \|X\|_p \|Y\|_q$; $|\mathbb{E}XY| \leq (\mathbb{E}X^2)^{1/2}(\mathbb{E}Y^2)^{1/2}$.
Minkowski / Lyapunov. $p \geq 1$: $\|X+Y\|_p \leq \|X\|_p + \|Y\|_p$; $X \in L^s$, $0 < r \leq s$: $\|X\|_r \leq \|X\|_s$.
c_r. $\mathbb{E}|X+Y|^r \leq c_r(\mathbb{E}|X|^r + \mathbb{E}|Y|^r)$; $c_r = 1$ for $0 < r \leq 1$, $c_r = 2^{r-1}$ for $r \geq 1$.
Marcinkiewicz–Zygmund (iid). $X_i \overset{iid}{\sim} X$, $\mathbb{E}X = 0$, $\mathbb{E}|X|^p < \infty$: $S_n = \sum_{i=1}^n X_i$,
$\mathbb{E}|S_n|^p \leq B_p \mathbb{E}|X|^p\, \{n,\ 1 \leq p \leq 2;\ n^{p/2},\ p \geq 2\}$.
Markov / Chebyshev. $r, a > 0$: $\mathbb{P}(|X| \geq a) \leq \mathbb{E}|X|^r / a^r$;
$\mathbb{P}(|X - \mathbb{E}X| \geq a) \leq \mathrm{Var}(X)/a^2$.
Hoeffding. X_i indep., $a_i \leq X_i \leq b_i$ a.s., $V_b = \sum_i (b_i - a_i)^2$:
$\mathbb{P}(|\sum_i (X_i - \mathbb{E}X_i)| \geq z) \leq 2e^{-2z^2/V_b}$.
Bernstein. X_i indep., $\mathbb{E}X_i = 0$, $|X_i| \leq M$, $v = \sum_i \mathrm{Var}\,X_i$:
$\mathbb{P}(|\sum_i X_i| \geq z) \leq 2e^{-z^2/[2(v + Mz/3)]}$.

CONDITIONAL MOMENTS / NORMAL

Tower / LIE. $Y \in L^1$, $\mathcal{G} \subseteq \mathcal{H} \subseteq \mathcal{F}$: $\mathbb{E}[\mathbb{E}(Y \mid \mathcal{H}) \mid \mathcal{G}] = \mathbb{E}(Y \mid \mathcal{G})$ a.s.; $\mathbb{E}\mathbb{E}(Y \mid \mathcal{G}) = \mathbb{E}Y$.
Conditional variance. $Y \in L^2$: $\mathrm{Var}(Y \mid X) = \mathbb{E}(Y^2 \mid X) - [\mathbb{E}(Y \mid X)]^2$; $\mathrm{Var}\,Y = \mathbb{E}[\mathrm{Var}(Y \mid X)] + \mathrm{Var}[\mathbb{E}(Y \mid X)]$.
Multivariate normal. $X \sim N_d(\mu, \Sigma)$, $\Sigma > 0$:
$f(x) = (2\pi)^{-d/2}(\det\Sigma)^{-1/2} e^{-(x-\mu)'\Sigma^{-1}(x-\mu)/2}$; $AX + b \sim N(A\mu + b, A\Sigma A')$.
Block normal. $X = (X_1', X_2')'$, $\mu = (\mu_1', \mu_2')'$, $\Sigma = (\Sigma_{ij})$, $\Sigma_{22} > 0$:
$X_1 \mid X_2 = x_2 \sim N(\mu_{1.2}, \Sigma_{1.2})$; $\mu_{1.2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$;
$\Sigma_{1.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. $X_1 \perp X_2$ iff $\Sigma_{12} = 0$.
Truncated normal. $Z \sim N(0,1)$: $\mathbb{E}(Z \mid Z > c) = \phi(c)/[1 - \Phi(c)]$;
$\mathbb{E}(Z \mid Z < c) = -\phi(c)/\Phi(c)$. For $Y \sim N(\mu, \sigma^2)$ standardize $a = (c - \mu)/\sigma$.
Dependence chain. $X, Y \in L^2$: $X \perp Y \Rightarrow \mathbb{E}(Y \mid X) = \mathbb{E}Y \Rightarrow \mathrm{Cov}(X, Y) = 0$; reverse false.

CONVERGENCE / STOCHASTIC ORDER

Strength. $a.s.$ or L^r $(r > 0) \Rightarrow p \Rightarrow d$; $X_n \to c$ iff $X_n \overset{p}{\to} c$.
Orders / products. $a_n, b_n > 0$ deterministic; $X_n = O_p(a_n)$ iff X_n/a_n is bounded in probability; $X_n = o_p(a_n)$ iff $X_n/a_n \overset{p}{\to} 0$. $o_p(a) \subset O_p(a)$; $O_p(a)O_p(b) = O_p(ab)$; $O_p(a) o_p(b) = o_p(ab)$; $o_p(a)o_p(b) = o_p(ab)$.

LLN / CLT

SLLN. X_i iid, $\mathbb{E}|X_1| < \infty$: $\bar{X}_n \overset{a.s.}{\to} \mu = \mathbb{E}X_1$.
CLT. $X_i \in \mathbb{R}$ iid, $\mathbb{E}X_1 = \mu$, $\mathbb{E}\|X_1\|^2 < \infty$, $\Sigma = \mathrm{Var}(X_1)$: $\sqrt{n}(\bar{X}_n - \mu) \overset{d}{\to} N(0, \Sigma)$.
Cramér–Wold. $Y_n \overset{d}{\to} Y$ iff $c'Y_n \overset{d}{\to} c'Y\ \forall c \in \mathbb{R}^k$.

ESTIMATION / INFORMATION

Bias–variance. $b = \mathbb{E}_\theta\hat{\theta} - \theta$; $\mathrm{MSE}_\theta(\hat{\theta}) := \mathbb{E}_\theta\|\hat{\theta} - \theta\|^2 = \mathrm{tr}\,\mathrm{Var}_\theta(\hat{\theta}) + b'b$.
MLE invariance. $\tau = g(\theta)$; $L^*(\tau) = \sup_{\theta : g(\theta) = \tau} L(\theta)$;
$\hat{\theta} \in \arg\max L \Rightarrow g(\hat{\theta}) \in \arg\max_\tau L^*(\tau)$.
Fisher / IME. $s_\theta = \partial_\theta \log f_\theta(Y)$, $H_\theta = \partial^2_{\theta\theta'} \log f_\theta(Y)$. Under common-support / differentiation-under-integral regularity: $\mathbb{E}_\theta s_\theta = 0$, $I_1(\theta) = \mathbb{E}_\theta(s_\theta s_\theta') = -\mathbb{E}_\theta H_\theta$; iid sample: $I_n = nI_1$.

MATRIX CALCULUS

$Z \sim N_k(0, I_k)$, $A = A' : Z'AZ \overset{d}{=} \sum_{j=1}^k \lambda_j(A) U_j$, $U_j \overset{iid}{\sim} \chi^2_1$; $P = P' = P^2$: $Z'PZ \sim \chi^2_{\mathrm{rank}(P)}$.
Vector derivatives. $\nabla_x(a'x) = a$; $\nabla_x(x'Ax) = (A + A')x$ $(= 2Ax$ if $A = A')$; $\nabla_b\|y - Xb\|^2 = -2X'(y - Xb)$; $\nabla^2_b = 2X'X$.
Gaussian integral. $a > 0$: $\int_{-\infty}^{\infty} e^{-ax^2}\, dx = \sqrt{\pi/a}$.

CLASSICAL TESTING

Power / size / level. $K(\theta) = \mathbb{P}_\theta(X \in \mathcal{R})$; size $= \sup_{\theta \in \Theta_0} K(\theta)$; level α iff size $\leq \alpha$.
p-value. Nested level-α regions: $p(x) = \inf\{\alpha : x \in \mathcal{R}_\alpha\}$.
Neyman–Pearson. $H_0 : f = f_0$ vs. $H_1 : f = f_1$; choose (c, γ) so $\phi^* = \mathbf{1}\{f_1 > cf_0\} + \gamma\mathbf{1}\{f_1 = cf_0\}$ has $\mathbb{E}_0\phi^* = \alpha$. Then $\mathbb{E}_1\phi \leq \mathbb{E}_1\phi^*$ for every $0 \leq \phi \leq 1$ with $\mathbb{E}_0\phi \leq \alpha$; every MP level-$\alpha$ test has this form P_0-a.s.
MLR. On common support, $f_{\theta_1}(x)/f_{\theta_0}(x)$ is nondecreasing in $T(x)\ \forall \theta_1 > \theta_0$.
Karlin–Rubin. MLR implies upper-tail power is nondecreasing: $\theta_2 > \theta_1 \Rightarrow \mathbb{P}_{\theta_2}(T > c) \geq \mathbb{P}_{\theta_1}(T > c)$. Calibrate rejection at θ_0 to probability α (randomize at $T = c$ if needed): reject for large T is level-α UMP for $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$.
When no UMP exists. If NP-optimal regions vary with alternative direction/nuisance, state a comparison criterion: require power $\geq \alpha$ at every alternative; require null rejection probability $= \alpha$ for every nuisance value; require invariance to irrelevant transformations; or maximize weighted-average, minimum, or local power. Weak-IV example: π nuisance. Under iid homoskedastic normal fixed-IV, use the two-sided invariant-similar class and condition on $\boldsymbol{Q}_T$ (a statistic absorbing first-stage nuisance π): choose $\kappa_\alpha(Q_T)$ so $\mathbb{P}_{\beta_0,\pi}\{LR > \kappa_\alpha(Q_T) \mid Q_T\} = \alpha$; iterated expectations gives size $\alpha\ \forall \pi$. Under these conditions, CLR power is uniformly close to the two-sided invariant-similar power envelope.

IDENTIFICATION / HECKIT

Identified set. $\Theta_I(P) = \{\theta : P_\theta = P\}$. Point ID: $\Theta_I(P) = \{\theta_0\}$. Partial ID: $|\Theta_I(P)| > 1$ (a proper subset is informative; $\Theta_I(P) = \Theta$ is no ID). Weak ID is a sequence/strength notion: the identification Jacobian/signal approaches rank deficiency so total information fails to diverge and root-n theory is nonuniform (e.g. $\Pi_n = C/\sqrt{n}$), although each finite-n model may be point-ID.
Mixed IV. $y = x'\beta + z_1'\gamma + e$, $z = (z_1', z_2')'$, $\mathbb{E}(z\varepsilon) = 0$; $x = \Pi_1 z_1 + \Pi_2 z_2 + \eta$, $\mathbb{E}(z\eta') = 0$, $\mathbb{E}(zz') > 0$. For $w = (x', z_1')'$, $M = \mathbb{E}[zw'] = \mathbb{E}[zz']\begin{pmatrix}\Pi_1' & I \\ \Pi_2' & 0\end{pmatrix}$:
$(\beta', \gamma')'$ is point-ID iff rank $M = d_x + d_{z_1}$ iff rank $\Pi_2 = d_x$.
Heckit model. $Y^* = X'\theta + \varepsilon$, $Y = DY^*$, $D = \mathbf{1}\{s + \eta \geq 0\}$, $s = X'\pi_1 + Z'\pi_2$; $(\varepsilon, \eta)' \mid X, Z \sim N(0, \Sigma)$, $\sigma_\eta = 1$, $\mathrm{Cov}(\varepsilon, \eta) = \sigma_{\varepsilon\eta}$: $\mathbb{P}(D = 1 \mid X, Z) = \Phi(s)$.
Heckit mean / ID. $D = 1 \iff \eta \geq -s$ and $\mathbb{E}(\varepsilon \mid \eta) = \sigma_{\varepsilon\eta}\eta$ imply $\mathbb{E}(Y \mid X, Z, D = 1) = X'\theta + \sigma_{\varepsilon\eta}\lambda(s)$, $\lambda(s) = \phi(s)/\Phi(s)$. For $R = (X', Z')'$, $\mathbb{E}(RR') > 0$ identifies π by probit; $\mathbb{E}[D\tilde{R}\tilde{R}'] > 0$, $\tilde{R} = (X', \lambda(s))'$, identifies $(\theta, \sigma_{\varepsilon\eta})$.
Heckit two-step. Probit D on (X, Z): $\hat{\lambda}_i = \phi(\hat{s}_i)/\Phi(\hat{s}_i)$. On $D = 1$, regress Y on $(X, \hat{\lambda})$. Need full rank; credible ID uses excluded Z.

SMOOTH EXTREMUM

EE / ID / UCON. $\hat{\theta}_n \in \Theta$, $Q_n(\hat{\theta}_n) \leq \inf_\theta Q_n(\theta) + o_p(1)$;
$\inf_{\|\theta - \theta_0\| \geq \varepsilon} Q(\theta) > Q(\theta_0)\ \forall \varepsilon > 0$; $\sup_\theta |Q_n(\theta) - Q(\theta)| \overset{p}{\to} 0$; together $\hat{\theta}_n \overset{p}{\to} \theta_0$.
Sample ULLN. $Q_n(\theta) = n^{-1}\sum_i m(W_i, \theta)$. W_i iid; Θ compact; $m(w, \cdot)$ continuous $\forall w$; $\mathbb{E}[\sup_{\theta\in\Theta}|m(W_i, \theta)|] < \infty \Longrightarrow$ UCON.
CF. $\theta_0 \in \mathrm{int}\,\Theta$; $B_n(\theta) := \partial^2_{\theta\theta'} Q_n(\theta)$; $\sqrt{n}\,\partial_\theta Q_n(\theta_0) \overset{d}{\to} N(0, \Omega_0)$;
$\sup_{\|\theta - \theta_0\| \leq \delta_n} \|B_n(\theta) - B_0\| \overset{p}{\to} 0\ \forall \delta_n \downarrow 0$, B_0 nonsingular;
$\partial_\theta Q_n(\hat{\theta}_n) = o_p(n^{-1/2})$:
$\sqrt{n}(\hat{\theta}_n - \theta_0) = -B_0^{-1}\sqrt{n}\,\partial_\theta Q_n(\theta_0) + o_p(1) \overset{d}{\to} N(0, B_0^{-1}\Omega_0 B_0^{-1})$.
NLS. $u_i(\theta) = Y_i - g(X_i, \theta)$, $D_i(\theta) = \partial_\theta g(X_i, \theta)$, $H_i(\theta) = \partial^2_{\theta\theta'} g(X_i, \theta)$;
$Q_n = (2n)^{-1}\sum_i u_i^2$, $\partial_\theta Q_n = -n^{-1}\sum_i u_i D_i$, $B_n(\theta) = n^{-1}\sum_i (D_i D_i' - u_i H_i)$. At θ_0: $U_i = u_i(\theta_0)$, $D_{i0} = D_i(\theta_0)$, $H_{i0} = H_i(\theta_0)$; $B_0 = \mathbb{E}(D_{i0}D_{i0}' - U_i H_{i0})$, $\Omega_0 = \mathbb{E}(U_i^2 D_{i0}D_{i0}')$; $\mathbb{E}(U_i \mid X_i) = 0 \Rightarrow B_0 = \mathbb{E}(D_{i0}D_{i0}')$; homoskedastic: $\Omega_0 = \sigma^2 B_0$.
GMM (fixed W_n). $\bar{g}_n(\theta) = n^{-1}\sum_i g_i(\theta)$, $G_n(\theta) = \partial_{\theta'}\bar{g}_n(\theta)$; $Q_n = \frac{1}{2}\bar{g}_n' W_n \bar{g}_n$, $\partial_\theta Q_n = G_n' W_n \bar{g}_n$, $B_n(\theta) = G_n' W_n G_n + \sum_{\ell=1}^k [W_n\bar{g}_n]_\ell \partial^2_{\theta\theta'} g_{n\ell}(\theta)$.
$W_n \overset{p}{\to} W = W' > 0$, $\mathbb{E}g_i(\theta_0) = 0$; $G = \mathbb{E}[\partial_{\theta'} g_i(\theta_0)]$, $\mathrm{rank}\,G = d$, $\Omega_g = \mathbb{E}[g_i(\theta_0) g_i(\theta_0)']$: $B_0 = G'WG$, $\Omega_0 = G'W\Omega_g WG$;
$W = \Omega_g^{-1} \Rightarrow B_0 = \Omega_0$, $\mathrm{Avar} = B_0^{-1}$.

SECOND-ORDER GMM RATE

Setup. $\theta = (\theta_{-p}', \theta_p)'$, $\Delta_- = \theta_{-p} - \theta_{-p,0}$, $\Delta_p = \hat{\theta}_p - \theta_{p0}$; $\hat{\theta} \overset{p}{\to} \theta_0$ and $\|\bar{g}_n(\hat{\theta})\|^2 \leq \|\bar{g}_n(\theta_0)\|^2 + o_p(n^{-1})$. $\Gamma = \partial_{\theta'_{-p}} g(\theta_0)$ full column rank;
$\partial_{\theta_p} g(\theta_0) = 0$, $B = \partial^2_{\theta_p} g(\theta_0)$. $P_\Gamma = \Gamma(\Gamma'\Gamma)^{-1}\Gamma'$, $M_\Gamma = I - P_\Gamma$, $B'M_\Gamma B > 0$.
Derivative averages. Let $d_{pn} = \partial_{\theta_p}\bar{g}_n(\theta_0) = O_p(n^{-1/2})$. Define $\hat{\Gamma}$, $\hat{B}$ as line-segment averages of $\partial_{\theta'_{-p}}\bar{g}_n$ and $\partial^2_{\theta_p}\bar{g}_n$; consistency/local ULLN give $\hat{\Gamma} \overset{p}{\to} \Gamma$, $\hat{B} \overset{p}{\to} B$.
Two expansions. $\bar{g}_n(\hat{\theta}) = \bar{g}_n(\theta_{-p,0}, \hat{\theta}_p) + \hat{\Gamma}\Delta_-$;
$\bar{g}_n(\theta_{-p,0}, \hat{\theta}_p) = \bar{g}_n(\theta_0) + d_{pn}\Delta_p + \frac{1}{2}\hat{B}\Delta_p^2$. Since $\Delta_p = o_p(1)$, $d_{pn}\Delta_p = o_p(n^{-1/2})$; hence $\bar{g}_n(\hat{\theta}) = \bar{g}_n(\theta_0) + \hat{\Gamma}\Delta_- + \frac{1}{2}\hat{B}\Delta_p^2 + o_p(n^{-1/2})$.
Solve / project. With $\hat{\Gamma}^+ = (\hat{\Gamma}'\hat{\Gamma})^{-1}\hat{\Gamma}'$,
$\Delta_- = \hat{\Gamma}^+[\bar{g}_n(\hat{\theta}) - \bar{g}_n(\theta_0) - \frac{1}{2}\hat{B}\Delta_p^2] + o_p(n^{-1/2})$. Substitution gives $\bar{g}_n(\hat{\theta}) = \bar{g}_n(\theta_0) + P_{\hat{\Gamma}}[\bar{g}_n(\hat{\theta}) - \bar{g}_n(\theta_0)] + \frac{1}{2}M_{\hat{\Gamma}}\hat{B}\Delta_p^2 + o_p(n^{-1/2})$.
Criterion order. $\|\bar{g}_n(\hat{\theta})\|^2 \leq \|\bar{g}_n(\theta_0)\|^2 + o_p(n^{-1})$ and $\bar{g}_n(\theta_0) = O_p(n^{-1/2})$ imply $\bar{g}_n(\hat{\theta})$ and their difference are $O_p(n^{-1/2})$. Squaring and using $P_{\hat{\Gamma}}M_{\hat{\Gamma}} = 0$ yields $O_p(1)\sqrt{n}\Delta_p^2 + \{B'M_\Gamma B/4 + o_p(1)\}n\Delta_p^4 = O_p(1)$.
Rates. Set $x_n = \sqrt{n}\Delta_p^2$. The last equation is $c_n x_n^2 + A_n x_n = R_n$, where $c_n \overset{p}{\to} B'M_\Gamma B/4 > 0$ and $A_n, R_n = O_p(1)$. Thus $x_n = O_p(1)$ and $\Delta_p = O_p(n^{-1/4})$. Then $\Delta_p^2 = O_p(n^{-1/2})$ in the solved equation gives $\Delta_- = O_p(n^{-1/2})$.

SMOOTH RESTRICTION TESTS

Common objects. $H_0 : h(\theta_0) = 0$, r; $H = \partial h(\theta_0)/\partial\theta'$, $\mathrm{rank}\,H = r$;
$V_\theta = B_0^{-1}\Omega_0 B_0^{-1}$; $V_h = HV_\theta H' = HB_0^{-1}\Omega_0 B_0^{-1}H'$.
Wald. $W_n = n\,h(\hat{\theta}_n)'\{\hat{H}_n\hat{V}_{\theta,n}\hat{H}_n'\}^{-1}h(\hat{\theta}_n) \overset{d}{\to} \chi^2_r$.
Restricted objects. $\hat{\theta}_n \in \arg\min_{\theta : h(\theta) = 0} Q_n(\theta)$; $\hat{B}_n := B_n(\hat{\theta}_n)$;
$\hat{V}_{h,n} = \hat{H}_n\hat{B}_n^{-1}\hat{\Omega}_n\hat{B}_n^{-1}\hat{H}_n'$.
LM. $LM_n = n\,\{\partial_\theta Q_n(\hat{\theta}_n)\}'\hat{B}_n^{-1}\hat{H}_n'\hat{V}_{h,n}^{-1}\hat{H}_n\hat{B}_n^{-1}\partial_\theta Q_n(\hat{\theta}_n) \overset{d}{\to} \chi^2_r$.
QLR. $\hat{\Omega}_0 = cB_0$, $\hat{c}_n \overset{p}{\to} c$:
$QLR_n = 2n\{Q_n(\hat{\theta}_n) - Q_n(\hat{\theta}_n)\}/\hat{c}_n = LM_n + o_p(1) \overset{d}{\to} \chi^2_r$.
Local direction. $\lambda \in \mathbb{R}^d$ fixed; $\theta_n = \theta_0 + \lambda/\sqrt{n}$;
$\delta(\lambda) = \lambda'H'V_h^{-1}H\lambda = \|V_h^{-1/2}H\lambda\|^2$; $W_n, LM_n, QLR_n \overset{d}{\to} \chi^2_r(\delta(\lambda))$.

NONSMOOTH / QUANTILE REGRESSION

EE3. $\hat{\theta}_n \in \Theta$, $\hat{\theta}_n \overset{p}{\to} \theta_0$; $Q_n(\theta) = \|\bar{g}_n(\theta)\|$; $Q_n(\hat{\theta}_n) \leq \inf_\theta Q_n(\theta) + o_p(n^{-1/2})$.
CF-NS. $\theta_0 \in \mathrm{int}\,\Theta$, $g(\theta_0) = 0$; g differentiable at θ_0, $\Gamma = \partial_\theta g(\theta_0)$ full column rank; $\sqrt{n}\,\bar{g}_n(\theta_0) \overset{d}{\to} N(0, V_0)$; $\sup_{\|\theta - \theta_0\| < \delta_n} \sqrt{n}\|[\bar{g}_n - g](\theta) - [\bar{g}_n - g](\theta_0)\| \overset{p}{\to} 0\ \forall \delta_n \downarrow 0$.
Linear representation. $A_\Gamma = (\Gamma'\Gamma)^{-1}\Gamma'$; $\sqrt{n}(\hat{\theta}_n - \theta_0) = -A_\Gamma\sqrt{n}\,\bar{g}_n(\theta_0) + o_p(1)$.

BOOTSTRAP CANCELLATION / SUBSAMPLING

Mean-test setup. X_i iid, $\mathbb{E}X_i = \mu$, $\mathrm{Var}\,X_i = 1$; $T_n = \sqrt{n}\bar{X}_n$; reject upper-tail $H_0 : \mu = 0$ vs. $H_1 : \mu > 0$ when $T_n > c_V$; for two-sided testing use absolute recentered roots.
Uncentered bootstrap cancels. Conditionally, $Z_n^* := \sqrt{n}(\bar{X}_n^* - \bar{X}_n) \overset{d^*}{\to} N(0, 1)$, but the uncentered bootstrap statistic is $T_n + Z_n^*$. Thus ideal $c_{1-\alpha}^* = T_n + z_{1-\alpha} + o_p(1)$ and rejection cancels to $0 > z_{1-\alpha} + o_p(1)$: as $B \to \infty$, NRP and fixed-alt. power tend to 0 if $\alpha < 1/2$ and 1 if $\alpha > 1/2$. For fixed B, with $r = \lceil (1-\alpha)B \rceil$, both limits are $\mathbb{P}\{Z_{(r)} < 0\} = \mathbb{P}\{\mathrm{Bin}(B, 1/2) \geq r\}$, so the test is inconsistent.
Recentered repair. Bootstrap Z_n^* itself and let $B \to \infty$: $cv \overset{p}{\to} z_{1-\alpha}$, NRP $\to \alpha$, fixed-$\mu > 0$ power $\to 1$.
Subsampling works. $s \to \infty$, $s/n \to 0$ (e.g. $s = \sqrt{n}$); $\sqrt{s}(\bar{X}_S - \bar{X}_n) \overset{d^*}{\to} N(0, 1)$ conditionally, $c_{1-\alpha}^{sub} = \sqrt{s}\,\bar{X}_n + z_{1-\alpha} + o_p(1)$. Under H_0, $\sqrt{s}\bar{X}_n = \sqrt{s/n}\,T_n = o_p(1)$, so NRP $\to \alpha$; for fixed $\mu > 0$, $T_n - c_{1-\alpha}^{sub} = (\sqrt{n} - \sqrt{s})\mu + O_p(1) \to \infty$, so power $\to 1$. Use absolute subsample roots for a two-sided test.

UNIFORM SIZE / TEST INVERSION

Null rejection. $H_0 : \theta = \theta_0$, nuisance $\gamma \in \Gamma$; $RP_n(\gamma) = \mathbb{P}_{\theta_0,\gamma}(\text{reject } H_0)$.
Pointwise NRP. $\forall$ fixed γ: $\lim\sup_{n\to\infty} RP_n(\gamma) \leq \alpha$; fix γ where $n \to \infty$.
Finite-sample size. $S_{2n} = \sup_{\gamma\in\Gamma} RP_n(\gamma)$; correct iff $S_{2n} \leq \alpha$ for every n.
Asymptotic size. $\mathrm{AsySz}(\theta_0) = \lim\sup_{n\to\infty}\sup_{\gamma\in\Gamma} RP_n(\gamma) = \lim\sup_n S_{2n}$; correct iff $\leq \alpha$. Finite-sample correct $\Rightarrow$ asymptotic correct $\Rightarrow$ pointwise.
Test inversion. $CS_n = \{\theta : T_n(\theta) \leq c_{1-\alpha}(\theta)\}$.
AsyCS. $\mathrm{AsyCS} := \lim\inf_n \inf_{\gamma\in\Gamma} \mathbb{P}_{\theta,\gamma}\{\theta \in CS_n\} \geq 1 - \alpha$.

MOMENT-INEQUALITY SIZE CORRECTION

Moment / type-h limit. $m = (m_1, \ldots, m_k)'$, $k = p + v$, $\bar{m}_n = n^{-1}\sum_i m(W_i, \theta)$. $S(m, \Sigma)$ is continuous, noninc reasing in the first p coordinates, and $S(m, \Sigma) = S(\Delta m, \Delta\Sigma\Delta)$ for all pd diagonal Δ. $\mathbb{E}_F m_j(\theta_0) \geq 0\ (j \leq p)$, $\mathbb{E}_F m_j(\theta_0) = 0\ (j > p)$; $\gamma_{1j} = \mathbb{E}_F m_j/\sigma_{F,j}$, $\sqrt{n}\gamma_{1n} \to h_1 \in [0,\infty]^p$, $h_2 = \mathrm{Corr}_F(m)$.
$T_n \overset{d}{\to} J_h := S\{h_1^{1/2}Z + (h_1, 0_v), h_2\}$; $c_h(1-\tau)$ is its $(1-\tau)$ quantile.
Local nuisance / CR. $\hat{\Sigma} = n^{-1}\sum_i (m_i - \bar{m}_n)(m_i - \bar{m}_n)'$, $\hat{D} = \mathrm{Diag}(\hat{\Sigma})$, $\hat{\eta} = \sqrt{n}\,\hat{D}_{11}^{-1/2}\bar{m}_{1n}$; $\hat{R} = \hat{D}^{-1/2}\hat{\Sigma}\hat{D}^{-1/2}$; $\hat{\eta} - h_1 \overset{d}{\to} N(0, h_{2,11})$, so h_1 is not consistently estimable; $\hat{h}_2 \overset{p}{\to} h_2$ by WLLN/CMT. $k_\beta = z_{1-\beta/p}$;
$CR_{h_1}(1-\beta) = \{\hat{h}_1 \in [0,\infty]^p : \hat{h}_{1j} \geq \max(0, \hat{\eta}_j - k_\beta)\ \forall j \leq p\}$, so $\liminf \mathbb{P}\{h_1 \in CR\} \geq 1 - \beta$.
Why size is only $\alpha + \beta$. $cv(1-\alpha) = \sup_{\hat{h}_1 \in CR} c_{(\hat{h}_1, \hat{h}_2)}(1-\alpha)$. On $A_n = \{h_1 \in CR\}$, $cv \geq c_{(h_1,\hat{h}_2)}(1-\alpha) = c_{(h_1,h_2)}(1-\alpha) + o_p(1)$, so for $R = \{T_n > cv\}$, $\mathbb{P}(R) \leq \mathbb{P}(R, h_1 \in CR) + \mathbb{P}(h_1 \notin CR) \leq \alpha + \beta + o_p(1)$. Thus only AsySz $\leq \alpha + \beta$; the test need not be conservative relative to α.
Bonferroni repair. $0 < \beta < \alpha$; set $\delta = \alpha - \beta$ and use $cv(1-\delta)$, i.e. the $1 - \alpha + \beta$ quantile. Then $\lim\sup \mathbb{P}(R) \leq (\alpha - \beta) + \beta = \alpha$. Power vs. global worst-case critical values has no uniform ordering: the CR shrinks the nuisance set, but the higher quantile raises the critical value.

POST-SELECTION MAX-MEAN CI

Intervals / CLT. $\theta_n = \max\{\mu_{Xn}, \mu_{Yn}\}$; $C_{Jn} = [J_n \pm \hat{\sigma}_{Jn} z/\sqrt{n}]$, $J = X, Y$, $z = z_{1-\alpha/2}$; choose C_{Xn} iff $\bar{X}_n \geq \bar{Y}_n$. Uniform moments give $\left(\frac{\sqrt{n}(\bar{X}_n - \mu_{Xn})}{\hat{\sigma}_{Xn}}, \frac{\sqrt{n}(\bar{Y}_n - \mu_{Yn})}{\hat{\sigma}_{Yn}}\right) \overset{d}{\to} (Z_X, Z_Y)$, standard bivariate normal with correlation ρ.
Fixed F. If $\mu_X > \mu_Y$, WLLN gives $\mathbb{P}(\bar{X}_n \geq \bar{Y}_n) \to 1$, hence coverage $\to \mathbb{P}(|Z_X| \leq z) = 1 - \alpha$; symmetric if $\mu_Y > \mu_X$. If $\mu_X = \mu_Y$, with $A_0 = \{\sigma_X Z_X \geq \sigma_Y Z_Y\}$, $\mathbb{P}(\theta \in C_n) \to \mathbb{P}(|Z_X| \leq z, A_0) + \mathbb{P}(|Z_Y| \leq z, A_0^c)$, not automatically $1 - \alpha$.
Drift / selection. $\sqrt{n}(\mu_{Xn} - \mu_{Yn}) \to h_1$, $(\sigma_{Xn}, \sigma_{Yn}, \rho_n) \to (s_X, s_Y, \rho)$. Then the selection event converges to $A_h = \{s_X Z_X + h_1 \geq s_Y Z_Y\}$. If $h_1 \geq 0$, target $= \mu_{Xn}$ and the wrong-Y root is $Z_Y - h_1/s_Y$; if $h_1 < 0$, target $= \mu_{Yn}$ and the wrong-X root is $Z_X + h_1/s_X$.
Local coverage. $C(h) = \begin{cases}\mathbb{P}(|Z_X| \leq z, A_h) + \mathbb{P}(|Z_Y - h_1/s_Y| \leq z, A_h^c), & h_1 \geq 0, \\ \mathbb{