Contiguity and Asymptotics

John Duchi

Stats 300b – Winter Quarter 2021

Outline

- Absolute continuity and contiguity
- Le Cam's Lemmas
- Distances between distributions

Reading:

- van der Vaart, *Asymptotic Statistics* Ch. 6
- Lehmann & Romano, *Testing Statistical Hypothesis* Ch. 12.3

Recapitulation and motivation

- in asymptotic testing, we assumed locally "uniform" convergence guarantee

$$\sqrt{n}\frac{T_n - \mu(\theta_n)}{\sigma(\theta_n)} \xrightarrow{d} \mathcal{N}(0,1) \quad \text{where } \theta_n \to 0$$

- gave power of a test rejecting $\theta = 0$ for large $T_n - \mu(0)$
- slope $\mu'(0)/\sigma(0)$ governed (relative) efficiencies

goal:

- understand how distributions get "close" to one another
- perform power and level calculations simultaneously

The Portmanteau Lemma

Lemma

The following are all equivalent definitions of convergence in distribution $X_n \xrightarrow{d} X$*:*

(i) $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$ *for all bounded continuous* f

(ii) $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$ *for all bounded Lipschitz* f

(iii) $\liminf_n \mathbb{E}[f(X_n)] \geq \mathbb{E}[f(X)]$ *for all nonnegative continuous* f

(iv) $\liminf_n P(X_n \in G) \geq P(X \in G)$ *for all open sets* G

(v) $\liminf_n P(X_n \in C) \leq P(X \in C)$ *for all closed sets* C

Absolute Continuity

Definition

A distribution Q is *absolutely continuous* with respect to P, $Q \ll P$, if $P(A) = 0$ implies $Q(A) = 0$

consequence: if $Q \ll P$, Radon-Nikodym density $g = \frac{dQ}{dP}$ exists

- can compute means $Qf = Pfg$
- likelihood ratio g and P characterize Q

Changing measures with likelihood ratios

Observation

Let M be joint measure of the pair $(X, L) = (X, \frac{dQ}{dP})$ under P, where $Q \ll P$. Then

- $L \geq 0$ *and* $\mathbb{E}_M[L] = 1$
- $Q(X \in B) = \mathbb{E}_P[1\{X \in B\} L] = \int_{B \times \mathbb{R}_+} r dM(x, r)$
- $\mathbb{E}_Q[f(X)L] = \mathbb{E}_P[f(X)] = \int_{\mathcal{X} \times \mathbb{R}_+} f(x) r dM(x, r)$

said differently, knowing

$$\mathcal{L}(X \mid P) = \text{law of } X \text{ under } P$$
$$\mathcal{L}(L \mid P) = \text{law of } \frac{dQ}{dP} \text{ under } P$$

means we know $\mathcal{L}(X \mid Q)$

An asymptotic version of absolute continuity

idea: transfer "power" calculations under null P_0 to (local) alternatives $Q_n \to P_0$ in some way

Definition

A sequence Q_n is *contiguous* with respect to P_n, written $Q_n \triangleleft P_n$, if

$$P_n(A_n) \to 0 \quad \text{implies} \quad Q_n(A_n) \to 0.$$

They are *mutually contiguous*, $P_n \triangleleft\triangleright Q_n$, if $P_n \triangleleft Q_n$ and $Q_n \triangleleft P_n$

Asymptotics of the likelihood ratio

- let μ_n dominate P_n, Q_n and $p_n = \frac{dP_n}{d\mu_n}$ and $q_n = \frac{dQ_n}{d\mu_n}$
- define likelihood ratio $L_n = \frac{q_n}{p_n}$
- likelihood ratio is *tight* under P_n

Lemma (Le Cam's First Lemma)

The limits of L_n determine contiguity: the following are equivalent.

(1) $Q_n \triangleleft P_n$

(2) *If* $L_n^{-1} \xrightarrow{d}_{Q_n} U$ *along a subsequence, then* $\mathbb{P}(U > 0) = 1$

(3) *If* $L_n \xrightarrow{d}_{P_n} L$*, then* $\mathbb{E}[L] = 1$

(4) *If* $T_n \xrightarrow{p}_{P_n} 0$*, then* $T_n \xrightarrow{p}_{Q_n} 0$

Asymptotic log normality

An important special case: asymptotic normality when

$$\log \frac{dP_n}{dQ_n} \xrightarrow[Q_n]{d} \mathcal{N}(\mu, \sigma^2)$$

- certainly $Q_n \triangleleft P_n$ as $U = \exp(\mathcal{N}(\mu, \sigma^2)) > 0$

Lemma

$Q_n \triangleleft\triangleright P_n$ if and only if $\mu = -\frac{1}{2}\sigma^2$

Asymptotic log normality for smooth likelihoods

- log likelihood $\ell_{\theta_0} := \log p_{\theta_0}$, $\theta_0 \in \mathbb{R}^d$
- assume sufficiently smooth around θ_0, and let $h \in \mathbb{R}^d$

Lemma
For $X_1^n \overset{\text{iid}}{\sim} P_{\theta_0}$,

$$\log \frac{dP_{\theta_0 + h/\sqrt{n}}}{dP_{\theta_0}}(X_1, \ldots, X_n) = h^T \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \dot{\ell}_{\theta_0}(X_i) \right) - \frac{1}{2} h^T I_{\theta_0} h + o_P(1)$$

Asymptotically changing measure

Theorem (Le Cam; Theorem 6.6 in van der Vaart)

Let P_n, Q_n be distributions on a $X_n \in \mathcal{X}$ and $L_n = \frac{dQ_n}{dP_n}$. If $Q_n \triangleleft P_n$ and

$$(X_n, L_n) \xrightarrow[P_n]{d} (X, L)$$

where (X, L) has joint measure M on $\mathcal{X} \times \mathbb{R}_+$. Then

$$X_n \xrightarrow[Q_n]{d} Z \quad \textit{where} \quad \mathbb{P}(Z \in B) = \mathbb{E}_M[1\{X \in B\} L]$$

i.e. $\mathbb{P}(Z \in B) = \int_{B \times \mathbb{R}_+} r \, dM(x, r)$

Changing measures with asymptotic normality

Lemma (Le Cam's Third Lemma)

Assume

$$\left(X_n, \log \frac{dQ_n}{dP_n}\right) \xrightarrow[P_n]{d} \mathcal{N}\left(\begin{bmatrix} \mu \\ -\frac{1}{2}\sigma^2 \end{bmatrix}, \begin{bmatrix} \Sigma & \tau \\ \tau^T & \sigma^2 \end{bmatrix}\right).$$

Then $X_n \xrightarrow[Q_n]{d} \mathcal{N}(\mu + \tau, \Sigma)$.

Asymptotically linear statistics and smooth likelihoods

Assume typical likelihood expansions:

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_{\theta_0}(X_i) + o_{P_{\theta_0}}(1)$$

$$\log \frac{dP_{\theta_0 + h/\sqrt{n}}}{dP_{\theta_0}}(X_1, \ldots, X_n) = \frac{h^T}{\sqrt{n}} \sum_{i=1}^{n} \dot{\ell}_{\theta_0}(X_i) - \frac{1}{2} h^T I_{\theta_0} h + o_{P_{\theta_0}}(1)$$

where $P_{\theta_0}\psi_{\theta_0} = 0$, $\text{Cov}(\psi_{\theta_0}) = \Sigma$

Corollary

Under the above conditions,

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) \underset{P_{\theta_0+h/\sqrt{n}}}{\overset{d}{\longrightarrow}} \mathcal{N}\left(\text{Cov}(\psi_{\theta_0}, h^T \dot{\ell}_{\theta_0}), \Sigma\right).$$