

Chapter 7

Contiguity

Special version with missing proofs to be filled in by Stat 618 class

7.1 Definition and stochastic order equivalences

In many asymptotic problems one needs to study estimators under various sequences of probability models. For example, in Chapter 1, we saw that the Hodges estimator θ_n^* behaves badly under a sequence of alternatives $\theta_n := \theta_0 + \delta/\sqrt{n}$. For a careful analysis we would have to consider behavior of $\theta_n^*(x_1, \dots, x_n)$ under the product measure $\mathbb{P}_{n, \theta_n} := P_{\theta_n}^n$ on \mathcal{X}^n . From the theory developed in Chapter 3, we already know a lot about the behavior of θ_n^* under the product measure $\mathbb{P}_{n, \theta_0} := P_{\theta_0}^n$. We could repeat the arguments with θ_n taking over the role played by θ_0 , following closely the steps used for the θ_0 analysis, to derive the asymptotics under the alternatives. There is, however, a more elegant approach, whereby the analysis is concentrated into a study of the density $dP_{\theta_n}/dP_{\theta_0}$. The underlying magic is called **contiguity**, a subtle (see the Notes in Section 5) invention of Le Cam (1960).

As you will learn in the next few Chapters, contiguity lies at the root of a number of well known asymptotic facts.

The contiguity idea is not restricted to independent sampling. It makes sense—and has interesting consequences—for any two sequences $\{\mathbb{P}_n\}$ and $\{\mathbb{Q}_n\}$ of probability measures. For each n , both \mathbb{P}_n and \mathbb{Q}_n should live on the same space $(\Omega_n, \mathcal{F}_n)$, but there need be no constraint on how the spaces change with n . For example, \mathbb{P}_n and \mathbb{Q}_n might be the joint distributions of random vectors with dimension k_n , corresponding to parametric models whose dimensions change with sample size.

<1> **Definition.** A sequence $\{\mathbb{Q}_n\}$ is said to be *contiguous* to $\{\mathbb{P}_n\}$ if, for each sequence of sets $\{F_n\}$, with $F_n \in \mathcal{F}_n$:

$$\text{if } \mathbb{P}_n F_n \rightarrow 0 \text{ then } \mathbb{Q}_n F_n \rightarrow 0.$$

Write $\{\mathbb{Q}_n\} \triangleleft \{\mathbb{P}_n\}$, or just $\mathbb{Q}_n \triangleleft \mathbb{P}_n$, to denote contiguity.

Rewriting the limiting requirements of the definition as explicit δ, ϵ inequalities, we get a more cumbersome (but more versatile) characterization.

<2> **Lemma.** *The contiguity $\mathbb{Q}_n \triangleleft \mathbb{P}_n$ is equivalent to the assertion: for each $\epsilon > 0$ there exists an n_0 and a $\delta > 0$, both depending on ϵ , such that, for each $n \geq n_0$,*

$$\sup\{\mathbb{Q}_n F : F \in \mathcal{F}_n \text{ and } \mathbb{P}_n F < \delta\} \leq \epsilon,$$

That is, if $F \in \mathcal{F}_n$ and $\mathbb{P}_n F < \delta$, for some $n \geq n_0$, then $\mathbb{Q}_n F \leq \epsilon$.

PROOF

□

<3> **Example.** Let \mathbb{P}_n denote the $N(\alpha_n, 1)$ distribution and \mathbb{Q}_n denote the $N(\beta_n, 1)$ distribution, both on the real line. Under what conditions on the sequences of constants $\{\alpha_n\}$ and $\{\beta_n\}$ do we have $\mathbb{Q}_n \triangleleft \mathbb{P}_n$?

If the sequence $\delta_n := \beta_n - \alpha_n$ is not bounded then contiguity fails. For example, suppose $\delta_n \rightarrow \infty$ along some subsequence \mathbb{N}_1 . Define $F_n := [\beta_n, \infty)$ if $n \in \mathbb{N}_1$ and $F_n := \emptyset$ otherwise. Then $\mathbb{P}_n F_n \rightarrow 0$ but $\mathbb{Q}_n F_n = 1/2$ along the subsequence.

We will soon have elegant ways to show that $\mathbb{Q}_n \triangleleft \mathbb{P}_n$ if δ_n is bounded in absolute value by some finite constant C . For the moment, brute force will suffice. Then

$$\begin{aligned} \mathbb{Q}_n F &= (2\pi)^{-1/2} \int \{x \in F\} \exp(-(x - \beta_n)^2/2) \, dx \\ &= (2\pi)^{-1/2} \int \{x \in F\} \exp(\delta_n(x - \alpha_n) - \delta_n^2/2 - (x - \alpha_n)^2/2) \, dx \\ &\leq \mathbb{P}_n^x(\{x \in F\} \exp(C|x - \alpha_n|)). \end{aligned}$$

If we split the last integrand according to whether $|x - \alpha_n| \leq M$ or not, for some constant M , then make the change of variable $z = x - \alpha_n$ in the second contribution, we get a bound for the expectation:

$$\exp(CM)\mathbb{P}_n F + (2\pi)^{-1/2} \int \{|z| > M\} \exp(C|z| - z^2/2) \, dz.$$

If M is large enough, the second contribution is smaller than $\epsilon/2$. The first contribution is also smaller than $\epsilon/2$ if $\mathbb{P}_n F < \epsilon \exp(-CM)/2$.

□

The more cumbersome characterization makes it easy to prove that contiguity is equivalent to preservation of the $o_p(1)$ or $O_p(1)$ properties for sequences of random variables. Because there are so many different probability measures involved, we need to be explicit about the measure involved in these assertions.

<4> **Definition.** For a sequence of random variables $\{Y_n\}$, with Y_n measurable with respect to \mathcal{F}_n , write

(i) $Y_n = o_p(1; \mathbb{P}_n)$ to mean: $\mathbb{P}_n\{|Y_n| > \eta\} \rightarrow 0$ as $n \rightarrow \infty$, for each fixed $\eta > 0$

(ii) $Y_n = O_p(1; \mathbb{P}_n)$ to mean: for each $\epsilon > 0$ there exists a finite constant M_ϵ such that $\limsup_n \mathbb{P}_n\{|Y_n| > M_\epsilon\} < \epsilon$.

<5> **Lemma.** The following three conditions are equivalent.

(i) $\mathbb{Q}_n \triangleleft \mathbb{P}_n$

(ii) For each sequence of random variables $\{Y_n\}$, with Y_n measurable with respect to \mathcal{F}_n : if $Y_n = o_p(1; \mathbb{P}_n)$ then $Y_n = o_p(1; \mathbb{Q}_n)$.

(iii) For each sequence of random variables $\{Y_n\}$, with Y_n measurable with respect to \mathcal{F}_n : if $Y_n = O_p(1; \mathbb{P}_n)$ then $Y_n = O_p(1; \mathbb{Q}_n)$.

PROOF

□

Remark. A sequence of real random variables $\{Y_n\}$ of order $O_p(1; \mathbb{P}_n)$ is sometimes said to be stochastically bounded (under $\{\mathbb{P}_n\}$), or uniformly tight. Such a sequence must have a subsequence that converges in distribution to a probability measure concentrated on \mathbb{R} . For real-valued random variables the proof is easy: a Cantor diagonalization argument applied to the sequence of distribution functions evaluated on a countable dense subset of \mathbb{R} . The analog for more general spaces is often called the Prohorov/Le Cam theorem (UGMTP §7.5).

7.2 Likelihood ratios

Contiguity has simple characterizations in terms of the behavior of the likelihood ratios, the densities dQ_n/dP_n .

It pays to be quite precise in the definition of a likelihood ratio, to avoid later ambiguities concerning singular parts. Suppose both P and Q are probability measures defined on the same space (Ω, \mathcal{F}) . There is a unique decomposition of Q into a sum $Q_a + Q_s$, where Q_a is absolutely continuous with respect to P and Q_s is singular with respect to P , that is, Q_s concentrates on a set N_P with zero P measure (UGMTP §3.2). Some authors write dQ/dP for the density of Q_a with respect to P . I will write L for the density dQ_a/dP , an \mathbb{R}^+ -valued measurable function, and call it the **likelihood ratio for Q with respect to P** . I will also refer to (L, N_P) as the **Lebesgue decomposition of Q with respect to P** . Thus, at least for nonnegative measurable functions f ,

$$<6> \quad Qf = Q_a f + Q_s f = P(fLN_P^c) + Q(fN_P)$$

Of course the N_P^c is irrelevant for the P contribution, but it sometimes helps to be reminded indirectly that the density applies only to the contribution from Q_a .

Remark. If both P and Q are absolutely continuous with respect to a measure λ , with densities p and q , then we can take

$$L := (q/p)\{p \neq 0\} \quad \text{and} \quad N_P := \{p = 0\}.$$

Redefinition of the likelihood ratio on the set N_P would have no effect on the equality $<6>$. Some authors purposefully define L to be $(q/p)\{p \neq 0\} + \infty\{p = 0\}$, which leads to some economy of notation. For example, with likelihood ratios $\{L_n\}$ for sequences $\{P_n\}$ and $\{Q_n\}$, a statement like $L_n = O_p(1; Q_n)$ would imply both

$$\frac{dQ_n}{dP_n} = O_p(1; Q_n) \quad \text{and} \quad Q_n N_{P_n} \rightarrow 0.$$

The set $N_{P_n} = \{L_n = \infty\}$ would get absorbed into the set $\{L_n > M\}$ for each finite constant M .

After some experimentation on live audiences, I have decided that the possibilities for confusion outweigh the notational disadvantages of the more explicit treatment of singular parts of the $\{Q_n\}$. I will always regard the likelihood ratio as a real-valued random variable.

Lemma $<5>$ shows that contiguity is a matter of inheritance of a stochastic order property: to verify contiguity we could check the $O_p(1; Q_n)$ property for all $O_p(1; P_n)$ sequences. The next characterization simplifies the task by allowing us to check the inheritance for just one particular case.

Notice that $1 = \mathbb{Q}_n \Omega_n \geq \mathbb{P}_n L_n$, so L_n is always of order $O_p(1; \mathbb{P}_n)$.

<7> **Lemma.** *If \mathbb{Q}_n has Lebesgue decomposition (L_n, N_n) with respect to \mathbb{P}_n then $\mathbb{Q}_n \triangleleft \mathbb{P}_n$ if and only if both $L_n = O_p(1; \mathbb{Q}_n)$ and $\mathbb{Q}_n(N_n) \rightarrow 0$.*

PROOF

□

Remark. If I had adopted the convention that $L_n = \infty$ on N_n , the proof would have been slightly shorter. The case where $\mathbb{Q}_n N_n = 1$, with $L_n \equiv 0$, shows that the condition $L_n = O_p(1; \mathbb{Q}_n)$ by itself would not suffice for contiguity.

<8> **Example.** For the \mathbb{P}_n and \mathbb{Q}_n from Example <3>,

$$L_n = \exp(\delta_n(x - \alpha_n) - \delta_n^2/2) \quad \text{where } \delta_n := \beta_n - \alpha_n.$$

Under \mathbb{Q}_n the random variable $x - \alpha_n$ has a $N(\delta_n, 1)$ distribution. If $\{\delta_n\}$ is bounded then $\delta_n(x - \alpha_n)$, and hence L_n , is of order $O_p(1; \mathbb{Q}_n)$.

□

The automatic $O_p(1; \mathbb{P}_n)$ property of $\{L_n\}$ implies existence of subsequences that converge in distribution. Suppose L , on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, represents the limit distribution along some such subsequence $\{L_n : n \in \mathbb{N}_1\}$.

Remark. Be careful: \mathbb{P} need not be a limit of the \mathbb{P}_n in any sense; the probability \mathbb{P} exists only to give L a distribution. The image of \mathbb{P}_n under L_n converges, along the subsequence, to the image of \mathbb{P} under L , that is, $L_n(\mathbb{P}_n) \rightsquigarrow L(\mathbb{P})$.

For each finite constant M ,

$$\mathbb{P}(L \wedge M) = \lim_{n \in \mathbb{N}_1} \mathbb{P}_n(L_n \wedge M) \leq \liminf_{n \in \mathbb{N}_1} \mathbb{P}_n L_n \leq 1.$$

Let M increase to infinity to deduce that $\mathbb{P}L \leq 1$. Equality here will translate into a $O_p(1; \mathbb{Q}_n)$ property of $\{L_n : n \in \mathbb{N}_1\}$; equality for all such subsequences will translate into contiguity.

<9> **Lemma.** *The contiguity $\mathbb{Q}_n \triangleleft \mathbb{P}_n$ is equivalent to the equality $\mathbb{P}L = 1$ for every L that is a limit in distribution of a subsequence of the likelihood ratios $\{L_n\}$ under $\{\mathbb{P}_n\}$.*

PROOF Problems [1] and [2] show (via subsequencing arguments) that there is no loss of generality in considering only the case where L_n itself converges in distribution to some random variable L on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

□

The last Lemma has an interesting interpretation, which lends support to the idea that contiguity is a form of asymptotic absolute continuity. For simplicity, suppose L_n converges in distribution under \mathbb{P}_n to an L on $(\Omega, \mathcal{F}, \mathbb{P})$. Contiguity requires $\mathbb{P}L = 1$, a condition that begs for interpretation of L as the density of another probability measure \mathbb{Q} with respect to \mathbb{P} . The limit assertion then becomes

$$<10> \quad L_n = \frac{d\mathbb{Q}_n}{d\mathbb{P}_n} \text{ (under } \{\mathbb{P}_n\} \text{) } \rightsquigarrow \frac{d\mathbb{Q}}{d\mathbb{P}} \text{ (under } \mathbb{P} \text{)},$$

with \mathbb{Q} a probability measure absolutely continuous with respect to \mathbb{P} .

<11> **Example.** Once again consider the \mathbb{P}_n and \mathbb{Q}_n from Example <3>, with likelihood ratio $L_n = \exp(\delta_n(x - \alpha_n) - \delta_n^2/2)$, where $\delta_n := \beta_n - \alpha_n$. The difference $x - \alpha_n$ has a $N(0, 1)$ distribution, and thus $\log L_n$ is distributed as $N(-\delta_n^2/2, \delta_n^2)$, under \mathbb{P}_n . For L_n to converge in \mathbb{P}_n -distribution we must have $\delta_n^2 \rightarrow \delta^2 < \infty$ (compare with Problem [3]). The limit distribution is that of $L := \exp(\delta x - \delta^2/2)$ under the $N(0, 1)$ distribution \mathbb{P} on the real line. By direct calculation, $\mathbb{P}L = 1$. (Compare with the moment generating function of the normal distribution.) The corresponding \mathbb{Q} is the $N(\delta, 1)$ distribution.

□

The form of the limit distribution in the previous Example is not coincidental.

<12> **Example.** In many classical situations, $\log L_n$ has a limiting normal distribution, or, more precisely, $L_n \rightsquigarrow \exp(X)$, with X defined on some $(\Omega, \mathcal{A}, \mathbb{P})$, with distribution $N(\mu, \sigma^2)$. For contiguity we must have $1 = \mathbb{P} \exp(X) = \exp(\mu + \frac{1}{2}\sigma^2)$. That is, $\mu = -\frac{1}{2}\sigma^2$ is equivalent to contiguity in this setting.

□

7.3 Contiguity for product measures

For the study of asymptotic behavior under sequences of alternatives, we often need to consider sequences of probability measures $\mathbb{Q}_n := P_{\theta_n}^n$ and $\mathbb{P}_n := P_{\theta_0}^n$, where θ_n is a sequence converging to θ_0 at a $1/\sqrt{n}$ rate. For simplicity suppose θ is a real parameter, and P_θ has a smooth density f_θ with respect to a dominating measure λ .

Classical approximation arguments (similar to those used to establish quadratic approximations in Chapter 3) can be used to establish contiguity, $\mathbb{Q}_n \triangleleft \mathbb{P}_n$, when the density is twice continuously differentiable. The arguments become a little subtle when the densities do not all have the same support. The difficulties are avoided when $\{f_\theta > 0\}$ does not change with θ . For this case, by restricting λ to the common support set, we may even suppose $f_\theta(x) > 0$ for all θ and x , which ensures that there are no log 0 problems when defining $\ell_\theta(x) := \log f_\theta(x)$.

Under the classical regularity conditions the logarithm of the likelihood ratio

$$L_n(\theta) = \prod_{i \leq n} \frac{f(x_i, \theta)}{f(x_i, \theta_0)}$$

has a local quadratic approximation in $1/\sqrt{n}$ neighborhoods of θ_0 . More formally, the approximation results from the usual pointwise Taylor expansion of the log density $\ell(x, \theta) = \log f(x, \theta)$. For example, in one dimension,

$$\begin{aligned} \log L_n(\theta_0 + t/\sqrt{n}) &= \sum_{i \leq n} (\ell(x_i, \theta_0 + t/\sqrt{n}) - \ell(x_i, \theta_0)) \\ &= \frac{t}{\sqrt{n}} \sum_{i \leq n} \dot{\ell}(x_i, \theta_0) + \frac{t^2}{2n} \sum_{i \leq n} \ddot{\ell}(x_i, \theta_0) + \dots \\ &\approx tZ_n - \frac{t^2}{2}\Gamma, \end{aligned}$$

where $\Gamma = -P_{\theta_0} \ddot{\ell}(x, \theta_0)$ and

$$Z_n = \sum_{i \leq n} \dot{\ell}(x_i, \theta_0)/\sqrt{n} \rightsquigarrow N(0, \text{var}_{\theta_0} \dot{\ell}(x, \theta_0)).$$

The limiting variance for Z_n and the coefficient Γ from the quadratic term both equal the information function evaluated at θ_0 .

The equality $-P_{\theta_0}^x \ddot{\ell}_{\theta_0}(x) = \text{var}_{\theta_0}(\dot{\ell}_{\theta_0}(x))$ is the classical dual representation for the information function at θ_0 . As Le Cam and Yang (2000, page 41) commented,

The equality ... is the classical one. One finds it for instance in the standard treatment of maximum likelihood estimation under Cramér's conditions. There it is derived from conditions of differentiability under the integral sign.

The classical equality is nothing more than contiguity in disguise.

A rigorous analysis becomes more complicated if the sets $\{f_\theta > 0\}$ are not all the same. We then need to impose a condition regarding the mass of the part of P_θ that is singular with respect to P_{θ_0} . See Chapter 11 for details.

7.4 Limit distributions under contiguous alternatives

Contiguity was advertized in Section 1 as a way to transfer either $o_p(\cdot)$ or $O_p(\cdot)$ assertions from $\{\mathbb{P}_n\}$ to $\{\mathbb{Q}_n\}$. It can also be used to transfer assertions of convergence in distribution for sequences of random vectors $\{Y_n\}$, if we control the joint behaviour of Y_n and the likelihood ratio. The idea behind the proof is straightforward if we ignore complications such as unbounded likelihoods: for bounded, uniformly continuous g ,

$$\mathbb{Q}_n g(Y_n) \stackrel{?}{=} \mathbb{P}_n L_n g(Y_n) \stackrel{?}{\rightarrow} \mathbb{P} L g(Y).$$

In a rigorous proof, contiguity controls the contributions from regions of large L_n , and from the singularity region \mathcal{N}_n , and then convergence in distribution of (L_n, Y_n) takes care of the convergence assertion. The limit expression becomes $\mathbb{Q}g(Y)$, where \mathbb{Q} is the probability measure defined to have density L with respect to \mathbb{P} . That is, the limit distribution of Y_n under \mathbb{Q}_n is given by Y , as a random vector on $(\Omega, \mathcal{F}, \mathbb{Q})$.

We will need the result only for random vectors Y_n , but the proof actually works for random elements more general spaces.

<13> **Lemma.** *Suppose (Y_n, L_n) converges in distribution under $\{\mathbb{P}_n\}$ to a limit represented by a pair (Y, L) on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with $\mathbb{P}L = 1$. Then $\{Y_n\}$ converges in distribution under $\{\mathbb{Q}_n\}$ to the limit represented by Y as a random element on the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, where \mathbb{Q} has*

density L with respect to \mathbb{P} . That is, $\mathbb{Q}_n g(Y_n) \rightarrow \mathbb{Q} g(Y) := \mathbb{P} L g(Y)$, at least for bounded, continuous g .

PROOF The condition $\mathbb{P} L = 1$ ensures that $\mathbb{Q}_n \triangleleft \mathbb{P}_n$. Fix $\epsilon > 0$ and let g be a bounded, continuous function. For convenience suppose $0 \leq g \leq 1$.

...

□

Remark. By the same argument (or just by substitution of (Y_n, L_n) for Y_n in the conclusion of the Lemma), the pair (Y, L) under \mathbb{Q} also represents the limit distribution for the pairs (Y_n, L_n) under $\{\mathbb{Q}_n\}$.

Convergence in distribution of (Y_n, L_n) is equivalent to convergence in distribution of $(Y_n, \log L_n)$. When the joint limit is normal, the assertion of the preceding Lemma takes a particularly simple form. The result is known as *Le Cam's Third Lemma*.

<14> **Example.** Suppose $(Y_n, L_n) \rightsquigarrow (Y, e^Z)$ under $\{\mathbb{P}_n\}$, where the pair (Y, Z) , defined on $(\Omega, \mathcal{F}, \mathbb{P})$, has a joint normal distribution. To ensure contiguity, the marginal Z distribution must be $N(-1/2\sigma^2, \sigma^2)$ for some $\sigma^2 > 0$. Let the marginal Y distribution be $N(\mu, V)$, and let γ denote the vector of covariances between Y and Z . Under \mathbb{P} the pair (Y, Z) has moment generating function ...

That is, under \mathbb{Q}_n the limiting variances and covariances stay the same, but the mean of Y is shifted to $\mu + \gamma$.

□

<15> **Example.** Chapter 3 derived the asymptotic behavior under $\mathbb{P}_n := P_{\theta_0}^n$ of an estimator $\hat{\theta}_n$ that minimizes $\sum_{i \leq n} g(x_i, \theta)$. For simplicity, reconsider

only the case where θ is a real-valued parameter. Under classical regularity conditions, the standardized estimator had a representation

$$Y_n := \sqrt{n} (\hat{\theta}_n - \theta_0) = n^{-1/2} h(x_i) + o_p(1; \mathbb{P}_n),$$

for some h in $\mathcal{L}^2(P_{\theta_0})$ with $P_{\theta_0} h = 0$. Under mild assumptions, this representation also gives the asymptotic distribution of $\sqrt{n} (\hat{\theta}_n - \theta_n) = Y_n - t$ under $\mathbb{Q}_n := P_{\theta_n}^n$ when $\theta_n = \theta_0 + t/\sqrt{n}$ for a fixed t .

Suppose the likelihood ratio for \mathbb{Q}_n with respect to \mathbb{P}_n has the representation

$$L_n = (1 + o_p(1; \mathbb{P}_n)) \exp(tZ_n - \frac{1}{2}t^2\Gamma),$$

with $Z_n = n^{-1/2} \sum_{i \leq n} \Delta(x_i)$ for a Δ with $P_{\theta_0} \Delta = 0$ and $P_{\theta_0} \Delta^2 = \Gamma$. Then

$$(Y_n, Z_n) = o_p(1; \mathbb{P}_n) + n^{-1/2} \sum_{i \leq n} (h(x_i), \Delta(x_i)),$$

which has a limiting bivariate normal distribution (Y, Z) with ...

That is, the limit distribution for $\sqrt{n} (\hat{\theta}_n - \theta_n)$ under \mathbb{Q}_n is the same as the limiting distribution of $\sqrt{n} (\hat{\theta}_n - \theta_0)$ under \mathbb{P}_n if $\text{cov}_{\mathbb{P}}(Y, Z) = t$. This equality is precisely the condition derived in Chapter 1 from the assumption that $P_{\theta} g(x, t)$ is minimized at $t = \theta$.

Estimators for which the limiting distribution of $\sqrt{n} (\hat{\theta}_n - \theta_n)$ under $P_{\theta_n}^n$ is the same for each $\theta_n = \theta_0 + t/\sqrt{n}$ are said to be **Hájek regular** at θ_0 . This regularity property will return in a later chapter as one of the assumptions for the Hájek-Le Cam Convolution Theorem.

□

Problems

- [1] Suppose $\{\mathbb{P}_n\}$ and $\{\mathbb{Q}_n\}$ are sequences of probability measures with the following property: for each subsequence $\mathbb{N}_1 \subseteq \mathbb{N}$ there exists a subsubsequence $\mathbb{N}_2 \subseteq \mathbb{N}_1$ for which $\{\mathbb{Q}_n : n \in \mathbb{N}_2\} \triangleleft \{\mathbb{P}_n : n \in \mathbb{N}_2\}$. Show that $\{\mathbb{Q}_n : n \in \mathbb{N}\} \triangleleft \{\mathbb{P}_n : n \in \mathbb{N}\}$. Hint: If contiguity fails, there is subsequence for which there are sets with $\mathbb{P}_n F_n \rightarrow 0$ but $\mathbb{Q}_n F_n > \epsilon$, for some $\epsilon > 0$.

- [2] Suppose $\{X_n\}$ is a sequence of random variables with the following property: for each subsequence $\mathbb{N}_1 \subseteq \mathbb{N}$ there exists a subsubsequence $\mathbb{N}_2 \subseteq \mathbb{N}_1$ for which $\{X_n : n \in \mathbb{N}_2\} = O_p(1)$. Show that $\{X_n : n \in \mathbb{N}\} = O_p(1)$.
- [3] Suppose $Z_n \rightsquigarrow N(0, I_k)$ and that $\alpha_n Z_n + \beta_n$ has a nondegenerate limit distribution, for a pair of deterministic sequences $\{\alpha_n\}$ and $\{\beta_n\}$. Show that both $|\alpha_n|$ and β_n must converge to finite limits.

7.5 Notes

Le Cam (1960) defined contiguity and derived its most important properties, in a few pages. The name “Le Cam’s Third Lemma” seems due to Hájek and Šidák (1967, Chapter VI). It was the third of the lemmas in their chapter describing contiguity. The numbering now should have little significance.

Lucien Le Cam himself felt that describing contiguity as a subtle invention was an exaggeration. In a private letter to me he wrote “Really, contiguity is a very trivial affair. I just gave it a name that pleased people.” Maybe the only subtlety lies in the recognition that something so trivial is worth noticing. To my chagrin, I ignored the concept for many years, because it seemed hardly worth bothering about. Moreover, I have found that I was not alone in my oversight. Maybe subtlety lies in the eye of the beholder.

References

- Hájek, J. and Z. Šidák (1967). *Theory of Rank Tests*. Academic Press. Also published by Academia, the Publishing House of the Czechoslovak Academy of Sciences, Prague.
- Le Cam, L. (1960). Locally asymptotically normal families of distributions. *University of California Publications in Statistics* 3, 37–98.
- Le Cam, L. and G. L. Yang (2000). *Asymptotics in Statistics: Some Basic Concepts* (2nd ed.). Springer-Verlag.