*6. Bahadur's rescue of efficiency

The representation of the likelihood ratio in Theorems <14> and <18> provides the framework for a method to exclude the superefficiency phenomenon described in Section 1.4. In fact, the form of the underlying statistical models (independent, identically distributed observations from a smoothly parametrized family of densities,...) becomes almost irrelevant once we have the limit behavior

<30> $$\frac{d\mathbb{Q}_n}{d\mathbb{P}_n} \rightsquigarrow \exp\left(tZ - \tfrac{1}{2}t^2\sigma^{-2}\right) \qquad \text{with } Z \sim N(0,\sigma^{-2}) \text{ under } \mathbb{P},$$

for a constant $t > 0$. For the cases considered in Section 2, the constant σ^{-2} was the information function evaluated at the value θ_0 that defined $\mathbb{P}_n$. The only other vestige of the underlying parameter is an assumption about the asymptotic behavior of some estimator T_n. Specifically, suppose there is a number θ_0 for which

<31> $$\sqrt{n}\,(T_n - \theta_0) \rightsquigarrow N(0,\tau^2) \qquad \text{under } \mathbb{P}_n.$$

With a very mild assumption—weaker than the assumption that $\sqrt{n}(T_n - \theta_n)$ has a limiting $N(0,\tau^2)$ distribution under $\mathbb{Q}_n$—on the behavior of T_n under $\mathbb{Q}_n$, Bahadur (1964) was able to rule out the possibility that $\tau^2 < \sigma^2$, the inequality corresponding to superefficiency of T_n at θ_0.

<32> **Theorem.** *Suppose <30> and <31> hold, and*

$$\liminf_n \mathbb{Q}_n\{\sqrt{n}\,(T_n - \theta_n) < 0\} \le \tfrac{1}{2} \qquad \text{where } \theta_n := \theta_0 + t/\sqrt{n}.$$

Then $\tau^2 \ge \sigma^2$.

The proof will follow as a simple consequence of the following Lemma, which captures the essence of Bahadur's main argument.

<33> **Lemma.** *Suppose $\mathbb{P}_n$ and $\mathbb{Q}_n$ are probability measures with $\mathbb{Q}_n$ contiguous to $\mathbb{P}_n$. Suppose $d\mathbb{Q}_n/d\mathbb{P}_n$, as random variables on $(\mathcal{X}_n, \mathcal{A}_n, \mathbb{P}_n)$, converge in distribution to a random variable L on $(\mathcal{X}, \mathcal{A}, \mathbb{P})$. Then for each sequence of measurable functions ψ_n with $0 \le \psi_n \le 1$, and each positive constant C,*

$$\liminf_n \left(\mathbb{P}_n\psi_n + C\mathbb{Q}_n\bar{\psi}_n\right) \ge \|\mathbb{P} \wedge (C\mathbb{Q})\|_1,$$

where $\mathbb{Q}$ is the probability measure on $(\mathcal{X}, \mathcal{A})$ defined by $d\mathbb{Q}/d\mathbb{P} = L$.

Proof. Write L_n for the density of the part of $\mathbb{Q}_n$ that is absolutely contnuous with respect to $\mathbb{P}_n$. We are assumimg that $L_n \rightsquigarrow L$. Thus

$$\mathbb{P}_n\psi_n + C\mathbb{Q}_n\bar{\psi}_n \ge \inf_{0\le\psi\le 1} \mathbb{P}_n\left(\psi + CL_n\bar{\psi}\right) = \mathbb{P}_n\left(\{CL_n \le 1\} + CL_n\{CL_n > 1\}\right).$$

That is, the infimum is achieved when $\psi := \{CL_n \le 1\}$. Rewrite the last expectation as $\mathbb{P}_n\left(1 \wedge (CL_n)\right)$. The map $x \mapsto 1 \wedge (Cx)$ is bounded and continuous on $\mathbb{R}^+$. The lower bound converges to $\mathbb{P}\left(1 \wedge (CL)\right) = \|\mathbb{P} \wedge (C\mathbb{Q})\|_1$, as asserted. □

Proof of Theorem <32>. Identify the limit distribution for L_n with the distribution of the density $d\mathbb{Q}/d\mathbb{P}$, where $\mathbb{P} := N(0,\sigma^2)$ and $\mathbb{Q} := N(t,\sigma^2)$. Invoke the Lemma with $\psi_n := \{\sqrt{n}(T_n - \theta_n) \ge 0\}$ and $C := \exp\left(-\sigma^{-2}t^2/2\right)$. The lim inf of $\mathbb{P}_n\psi_n + C\mathbb{Q}_n\bar{\psi}_n$ is less than $\mathbb{P}\{N(-t,\tau^2) \ge 0\} + \tfrac{1}{2}C$. To calculate the norm of

$\mathbb{P}\wedge(C\mathbb{Q})$, note that the $N(0,\sigma^2)$ density is smaller than C times the $N(t,\sigma^2)$ density at those points x of the real line for which

$$-\tfrac{1}{2}x^2\sigma^{-2} \le -\tfrac{1}{2}\sigma^{-2}t^2 - \tfrac{1}{2}\sigma^{-2}(x-t)^2,$$

that is, when $x \ge t$. Thus

$$\|\mathbb{P}\wedge(C\mathbb{Q})\| = \mathbb{P}[t,\infty) + C\mathbb{Q}(-\infty,t] = \bar{\Phi}(t/\sigma) + \tfrac{1}{2}C.$$

□ In order that $\bar{\Phi}(t/\tau) + \frac{1}{2}C \ge \bar{\Phi}(t/\sigma) + \frac{1}{2}C$, we must have $\tau \ge \sigma$.

Extra note

The argument in Section 6 is an extension of the method of Bahadur (1964). He noted that there is an easy generalization to the case where the parameter is vector valued. Bahadur imposed classical regularity conditions to produce the required approximation for the likelihood ratio.

Extra problems

[8] Show that the affinity between two finite Borel measures λ and μ on a metric space $\mathcal{X}$ equals the infimum of $\lambda g + \mu\bar{g}$ taken over all continuous functions g for which $0 \le g \le 1$. Hint: Use the fact that the bounded continuous functions are dense in $\mathcal{L}^1(\lambda+\mu)$. Also, if $0 \le f \le 1$ show that $|f - g_0| \le |f - g|$ where $g_0 = 1 \wedge g^+$.

[9] Let $\mathbb{P}_n$ and $\mathbb{Q}_n$ be as in Lemma <33>. Suppose $\{Y_n\}$ is sequence of random vectors for which $Y_n \rightsquigarrow \lambda$ under $\mathbb{P}_n$ and $Y_n \rightsquigarrow \mu$ under $\mathbb{Q}_n$, where λ and μ are probability measures on $\mathbb{R}^k$. For each positive constant C, show that

$$\|\lambda \wedge (C\mu)\|_1 \ge \|P \wedge (CQ)\|_1$$

Deduce that $\|\lambda - \mu\|_1 \le \|P - Q\|_1$. Hint: Invoke the Lemma with $\psi_n := g(Y_n)$, with g continuous and $0 \le g \le 1$, then appeal to Problem [8].

[10] Show that $\|N(t_1,\sigma^2) - N(t_2,\sigma^2)\|_1 = 2\mathbb{P}\{|N(0,1)| \le |t_1 - t_2|/\sigma\}$.

References

Bahadur, R. R. (1964), 'On Fisher's bound for asymptotic variances', *Annals of Mathematical Statistics* **35**, 1545–1552.

Hájek, J. & Šidák, Z. (1967), *Theory of Rank Tests*, Academic Press. Also published by Academia, the Publishing House of the Czechoslavak 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. & Yang, G. L. (2000), *Asymptotics in Statistics: Some Basic Concepts*, 2nd edn, Springer-Verlag.