

Local Asymptotic Normality in Tests

Lecturer: Michael I. Jordan

Scribe: Blaine Nelson

0.1 Notation

experiment $(\mathcal{X}, \mathcal{A}, P_h : h \in H)$ test $\phi : \mathcal{X} \rightarrow [0, 1]$ power $\pi(h) = \mathbb{E}_h \phi(X)$ level $\alpha : \sup_{h \in H_0} \pi(h) \leq \alpha$

1 Local Asymptotic Normality in Testing

Proposition 1 ((see van der Vaart, 1998, Proposition 15.2, p. 217)). Suppose $X \sim N(h, \Sigma)$ and $c^T \Sigma c > 0$ for some known c . In testing the null hypothesis $H_0 : c^T h = 0$ versus the alternative hypothesis $H_1 : c^T h > 0$, the Neyman-Pearson lemma implies that the test that rejects H_0 when $c^T X > z_\alpha \sqrt{c^T \Sigma c}$ is UMP at level α ; i.e., for any other power function $\pi(h)$ such that $\pi(h) \leq \alpha$ for every h with $c^T h = 0$, then for every h with $c^T h > 0$,

$$\pi(h) < 1 - \Phi \left(z_\alpha - \frac{c^T h}{\sqrt{c^T \Sigma c}} \right).$$

Theorem 2 ((see van der Vaart, 1998, Theorem 15.4, p. 219)). Suppose the following:

- $\{P_{n,0}\}$ are locally asymptotically normal (LAN) at θ_0 with non-singular Fisher information I_{θ_0} .
- let $\psi(\theta)$ be differentiable at θ_0 , $\psi(\theta_0) = 0$ and $\dot{\psi}_{\theta_0}$ be non-zero.

Then for any sequence of level α tests for testing $H_0 : \psi(\theta) \leq 0$ versus $H_1 : \psi(\theta) > 0$ and for every h such that $\dot{\psi}_{\theta_0} h > 0$, we have that their corresponding power functions satisfy:

$$\limsup_{n \rightarrow \infty} \pi_n \left(\theta_0 + \frac{h}{\sqrt{n}} \right) \leq 1 - \Phi \left(z_\alpha - \frac{\dot{\psi}_{\theta_0} h}{\sqrt{\dot{\psi}_{\theta_0} I_{\theta_0}^{-1} \dot{\psi}_{\theta_0}^T}} \right) \quad (1)$$

Proof. Choose h such that $\dot{\psi}_{\theta_0} h > 0$. Take a subsequence along which the limsup is attained. Take a further subsequence along which $\pi_n \left(\theta_0 + \frac{h}{\sqrt{n}} \right)$ converges for all h to the limit $\pi(h)$.

The function $\pi(h)$ is a power function in the Gaussian limit experiment. For $\dot{\psi}_{\theta_0}h < 0$, we have

$$\psi\left(\theta_0 + \frac{h}{\sqrt{n}}\right) = \frac{1}{\sqrt{n}}\left(\dot{\psi}_{\theta_0}h + o(1)\right) < 0$$

eventually. This implies that

$$\pi(h) \leq \limsup_{n \rightarrow \infty} \pi_n\left(\theta_0 + \frac{h}{\sqrt{n}}\right) \leq \alpha$$

and the continuity of ψ implies that $\pi(h) \leq \alpha$ for all h such that $\dot{\psi}_{\theta_0}h \leq 0$. Thus, $\pi(h)$ is of level α for testing $H_0 : \dot{\psi}_{\theta_0}h \leq 0$ versus $H_1 : \dot{\psi}_{\theta_0}h > 0$. The power function for this test is bounded above by that of the UMP test given in Proposition 1 (replace e^T by $\dot{\psi}_{\theta_0}$). \square

Theorem 3 (Rao-score statistic (see van der Vaart, 1998, Theorem 15.5, p. 219)). *Let T_n be the statistic*

$$T_n = \frac{\dot{\psi}_{\theta_0} I_{\theta_0}^{-1} \Delta_{n, \theta_0}}{\sqrt{\dot{\psi}_{\theta_0} I_{\theta_0}^{-1} \dot{\psi}_{\theta_0}^T}} + o_{P_{\theta_0}}(1)$$

where $\Delta_{n, \theta_0} = \frac{1}{\sqrt{n}} \sum_i \dot{\ell}_{\theta_0}(X_i)$ (Note: $\text{var}[T_n] = 1$). Then the test that rejects H_0 for large values of T_n is asymptotically optimal; i.e., it attains the RHS of Eq. (1).

Proof. We have

$$\begin{aligned} \left(\Delta_{n, \theta_0}, \log \frac{dP_{n, \theta_0 + h/\sqrt{n}}}{dP_{n, \theta_0}}\right) &\stackrel{\theta_0}{\rightsquigarrow} \left(\Delta, h^T \Delta - \frac{1}{2} h^T I_{\theta_0} h\right) \\ &\sim N\left(\begin{pmatrix} 0 \\ -\frac{1}{2} h^T I_{\theta_0} h \end{pmatrix}, \begin{pmatrix} I_{\theta_0} & I_{\theta_0} h \\ h^T I_{\theta_0} & h^T I_{\theta_0} h \end{pmatrix}\right) \end{aligned}$$

using Slutsky and LAN. Now, Le Cam's 3rd Lemma implies

$$\Delta_{n, \theta_0} \stackrel{h}{\rightsquigarrow} N(I_{\theta_0} h, I_{\theta_0})$$

Hence,

$$T_n \stackrel{h}{\rightsquigarrow} N\left(\frac{\dot{\psi}_{\theta_0} h}{\sqrt{\dot{\psi}_{\theta_0} I_{\theta_0}^{-1} \dot{\psi}_{\theta_0}^T}}, 1\right)$$

and this has as its power function the RHS of Eq. (1). \square

2 Wald Tests

We'll show that efficient estimates take the form

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = I_{\theta_0}^{-1} \Delta_{n, \theta_0} + o_{P_{\theta_0}}(1).$$

This implies that the test that rejects $\psi(\theta) \leq 0$ if

$$\sqrt{n} \psi(\hat{\theta}_n) \geq z_{\alpha} \sqrt{\dot{\psi}_{\hat{\theta}_n} I_{\hat{\theta}_n}^{-1} \dot{\psi}_{\hat{\theta}_n}^T}$$

is asymptotically optimal (use the delta method and Rao-score result) for testing $H_0 : \psi(\theta) \leq 0$, at every point θ_0 on the boundary of H_0 .

Example 4 (One-sample Location Model). Suppose that X_1, \dots, X_n are drawn from a location model $f(x - \theta)$ where f is a symmetric density with finite information I_f . We want to test the hypothesis $H_0 : \theta = 0$ versus the alternative $H_1 : \theta > 0$. The optimal asymptotic power of this test is given by,

$$1 - \Phi\left(z_\alpha - h\sqrt{I_f}\right).$$

To achieve this power, we use the score test:

$$T_n = \frac{-1}{\sqrt{n}} \frac{1}{\sqrt{I_f}} \sum_i \frac{f'}{f}(X_i) + o_{P_{\theta_0}}(1).$$

When f is completely known, we just use the first term. Often, though, f has the form $f(x) = \frac{1}{\sigma} f_0\left(\frac{x}{\sigma}\right)$ where σ is an unknown scale parameter. In this case, we have

$$\frac{1}{\sqrt{I_f}} \frac{f'}{f}(x) = \frac{1}{\sqrt{I_f}} \frac{f'_0}{f_0}\left(\frac{x}{\sigma}\right).$$

Nonetheless, using a consistent estimate of σ , $\hat{\sigma}_n$, yields the RHS result in Eq. (1).

For example, consider the t -test for a standard Gaussian density f_0 :

$$\begin{aligned} \frac{f'_0}{f_0}(x) &= -x \\ I_{f_0} &= 1 \end{aligned}$$

Thus, we want $T_n = \frac{1}{\sqrt{n}} \sum_i \frac{X_i}{\sigma} + o_P(n^{-1/2})$, or rather $T_n = \sqrt{n} \frac{X_n}{\sigma} + o_P(n^{-1/2})$. The statistic,

$$t = \sqrt{n} \frac{X_n}{S_n}$$

satisfies this requirement.

3 Signed Rank Statistic

Consider the statistic of the form:

$$T_n = \frac{1}{\sqrt{n}} \sum_i a_{n, R_{n,i}^+} \text{sign}(X_i)$$

where $R_{n,i}^+$ is the rank of $|X_i|$. These scores can be generated through a function ϕ by

$$a_{n,i} = \mathbb{E}\phi(U_{n(i)})$$

where $U_{n(i)}$ are the order statistics from a uniform distribution. As we will see next time, rank statistics achieve the optimal asymptotic power.

References

van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press, Cambridge.