1 Asymptotic Comparisons of Tests

Consider the following generic version of a testing problem. One observes data $X_i, i = 1, \ldots, n$ i.i.d. with distribution $P \in \mathbf{P} = \{P_\theta : \theta \in \Theta\}$ and wishes to test the null hypothesis $H_0 : \theta \in \Theta_0$ versus the alternative $H_1 : \theta \in \Theta_1$. A test is simply a function $\phi_n = \phi_n(X_1, \ldots, X_n)$ that returns the probability of rejecting the null hypothesis after observing $X_1, \ldots, X_n$. For example, ϕ_n might be the indicator function of a certain test statistic $T_n = T_n(X_1, \ldots, X_n)$ being greater than some critical value $c_n(1-\alpha)$. The test is said to be (pointwise) asymptotically of level α if for each $\theta \in \Theta_0$

$$\limsup_{n\to\infty} E_\theta[\phi_n] \leq \alpha \ .$$

In the class so far, one has encountered many such tests: Wald tests, quasi-likelihood ratio tests, and Lagrange multiplier tests. Suppose one is given two different tests of the same null hypothesis, $\phi_{1,n}$ and $\phi_{2,n}$, and both tests are (pointwise) asymptotically of level α. How can one choose between these two competing tests of the same null hypothesis? We will now explore the answer to this question in the context of a specific example.

2 A Symmetric Location Model

Suppose P_θ is the distribution with density $f(x-\theta)$ on the real line (w.r.t. Lebesgue measure). Suppose further that f is symmetric about 0 and that it's median, 0, is unique. Because f is symmetric about 0, $f(x-\theta)$ is symmetric about θ. We also have that $E_\theta[X] = \theta$ and $\text{Med}_\theta[X] = \theta$. Finally, suppose that the variance of P_0 is finite; that is, $\sigma_0^2 = \int x^2 f(x)dx < \infty$.

Notice that we could take f to be the density of a normal distribution and satisfy all of our assumptions. But many other choices of f satisfy these assumptions. For example, we could take f to be the uniform density on $[-1, 1]$, the logistic density, or the Laplace density.

Suppose $\Theta_0 = \{0\}$ and $\Theta_1 = \{\theta \in \mathbf{R} : \theta > 0\}$; i.e., we wish to test the null hypothesis $H_0 : \theta = 0$ versus the alternative $H_1 : \theta > 0$. How could we test this null hypothesis?

One such test is of course based on the familiar t-statistic:

$$\frac{\sqrt{n}\bar{X}_n}{\hat{\sigma}_n} \ .$$

Under the assumptions above, it is easy to show that

$$\frac{\sqrt{n}\bar{X}_n}{\hat{\sigma}_n} \xrightarrow{d} N(0,1)$$

under P_0. Thus, we may take

$$\phi_{1,n} = I\{\frac{\sqrt{n}\bar{X}_n}{\hat{\sigma}_n} > z_{1-\alpha}\}$$

where $z_{1-\alpha}$ is the $1-\alpha$ quantile of the standard normal distribution. Obviously, this test is asymptotically of level α (because $z_{1-\alpha}$ is a continuity point of the standard normal distribution).

A second test is based off of the following observation. Since f has median 0 under the null hypothesis, the number of positive and negative observations should be roughly equal (at least asymptotically). This suggests a test based on the test statistic:

$$\frac{1}{n}\sum_{1\leq i\leq n} I\{X_i > 0\} \ .$$

How does this statistic behave under the null hypothesis? We can compute that

$$E_0[I\{X_i < 0\}] = \Pr_0\{X_i > 0\} = 1 - F(0) = \frac{1}{2} \ ,$$

and thus

$$V_0[I\{X_i < 0\}] = F(0)(1-F(0)) = \frac{1}{4} \ .$$

Thus, by the central limit theorem for i.i.d. observations, we have that

$$\frac{1}{\sqrt{n}}\sum_{1\leq i\leq n}(I\{X_i > 0\} - \frac{1}{2}) \xrightarrow{d} N(0,\frac{1}{4}) \ .$$

So, we could take

$$\phi_{2,n} = I\{\frac{1}{\sqrt{n}}\sum_{1\leq i\leq n}(I\{X_i > 0\} - \frac{1}{2}) > \frac{1}{2}z_{1-\alpha}\} \ .$$

This test is known as the sign test. Obviously, this test is also asymptotically of level α.

3 A Naive Approach

It is natural to base comparisons of two different tests on their power functions. The power function of a test is the function $\pi_n(\theta) = E_\theta[\phi_n]$; i.e., it is the probability of rejecting the null hypothesis as a function of the unknown parameter θ. In this problem it will be difficult to compare the finite-sample power functions of the two tests, but we may try to do so in an asymptotic sense. To this end, let's compute the power functions of each of the above two tests at a fixed $\theta > 0$.

Let's start with the t-test. The key trick is to realize that

$$\begin{aligned}
\pi_{1,n}(\theta) &= P_\theta\{\frac{\sqrt{n}\bar{X}_n}{\hat{\sigma}_n} > z_{1-\alpha}\} \\
&= P_0\{\frac{\sqrt{n}\bar{Y}_n + \sqrt{n}\theta}{\hat{\sigma}_n} > z_{1-\alpha}\} \\
&= P_0\{\frac{\sqrt{n}\bar{Y}_n}{\hat{\sigma}_n} > z_{1-\alpha} - \frac{\sqrt{n}\theta}{\hat{\sigma}_n}\} ,
\end{aligned}$$

where $Y_i = X_i - \theta$ is distributed according to P_0. Importantly, we have done this in the denominator, too, using the fact that

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{1\le i\le n}(X_i - \bar{X}_n)^2 = \frac{1}{n}\sum_{1\le i\le n}(Y_i - \bar{Y}_n)^2 .$$

Since $\frac{\sqrt{n}\bar{Y}_n}{\hat{\sigma}_n}$ converges in distribution to a standard normal under P_0 and $z_{1-\alpha} - \frac{\sqrt{n}\theta}{\hat{\sigma}_n}$ diverges in probability to $-\infty$ under P_0, it follows that

$$\pi_{1,n}(\theta) \to 1$$

for every $\theta > 0$.

Now let's consider the sign test. Begin by considering the behavior of

$$\frac{1}{n}\sum_{1\le i\le n} I\{X_i > 0\}$$

under P_θ. Using the same trick as above, it is easy to compute that

$$\begin{aligned}
E_\theta[I\{X_i > 0\}] &= P_\theta\{X_i > 0\} \\
&= P_0\{Y_i > -\theta\} \\
&= 1 - F(-\theta) ,
\end{aligned}$$

which implies that

$$V_\theta[I\{X_i > 0\}] = F(-\theta)(1 - F(-\theta)) \ .$$

Thus, by the central limit theorem for i.i.d. observations, we have that

$$S_n(\theta) = \frac{1}{n} \sum_{1 \le i \le n} (I\{X_i > 0\} - (1 - F(-\theta)))$$

converges in distribution to $N(0, F(-\theta)(1 - F(-\theta)))$. We can now see that

$$\begin{aligned} \pi_{2,n}(\theta) &= P_\theta\{\frac{1}{\sqrt{n}} \sum_{1 \le i \le n} (I\{X_i > 0\} - \frac{1}{2}) > \frac{1}{2} z_{1-\alpha}\} \\ &= P_\theta\{S_n(\theta) > \frac{1}{2} z_{1-\alpha} - \sqrt{n}(\frac{1}{2} - F(-\theta))\} \ . \end{aligned}$$

Because f is symmetric about 0, $F(-\theta) < \frac{1}{2}$. We can now conclude as before that

$$\pi_{2,n}(\theta) \to 1$$

for every $\theta > 0$.

So, we see that a pointwise comparison of power functions of the two tests is completely uninformative. Both tests have power tending to 1 against any fixed alternative $\theta > 0$. In general, tests that have power tending to 1 against any fixed $\theta \in \Theta_1$ are said to be consistent. Any reasonable test will be consistent, so consistency is too weak of a requirement to be of use when trying to choose among different tests.

4 Local Asymptotic Power

Here, as always, there are an innumerable number of ways of embedding our situation with a sample of size n in a sequence of hypothetical situations with sample sizes larger than n. When choosing among these different asymptotic frameworks, it is important to keep in mind that what we are really interested in is the finite-sample behavior of the power function; that is, the behavior of the power function for our sample of size n. In the preceding section, we have shown that the power tends to 1 at any fixed $\theta > 0$

as n tends to infinity. Of course, in our sample of size n we know that the power is not 1 uniformly for $\theta > 0$. It may be very close to 1 for θ "far" from 0, but for θ "close" to 0 we would expect the finite-sample power function to be < 1. Of course, what we mean by "far" and "close" will change with our sample size n. Our asymptotic framework should reflect this fact. The above framework in which the alternative $\theta > 0$ is fixed does not. This suggests that we should consider the behavior of the power function evaluated at a sequence of alternatives θ_n, where θ_n tends to 0 at some rate. One can think of this as providing a locally asymptotic approximation to the power function.

It turns out that if θ_n tends to 0 slowly enough, then the power function will still tend to 1 as n tends to infinity. This follows from the following useful fact: If for every $\epsilon > 0$, $E_n(\epsilon) \to 1$, then there exists a sequence ϵ_n tending to 0 slowly enough so that $E_n(\epsilon_n) \to 1$. I won't prove this fact, but it isn't too hard to do it yourself. You can also find a proof in David Pollard's *A User's Guide to Measure-Theoretic Probability.*

Likewise, if θ_n tends to 0 quickly enough, then for asymptotic purposes it's as if $\theta_n = 0$. For any such sequence, the power function tends to α as n tends to infinity in each of the above two examples.

There is a delicate rate in between the two extremes above such that if θ_n tends to 0 at this rate, then the power will tend to a limit in $(\alpha, 1)$. This rate may be different in different problems, but in problems such as this one in which the distribution depends on θ in a "smooth" way it must be that $\theta_n = O(\frac{1}{\sqrt{n}})$. So, we will consider sequences $\theta_n = \frac{h}{\sqrt{n}}$, where $h \in \mathbf{R}$.

Let's again consider the t-test first. The calculation will be very similar to the one in the preceding section for the t-test. An important distinction is that now we must consider a triangular array of random variables because the distribution of the data is changing with each n. For each n, let $X_{i,n}, i = 1, \ldots, n$ be an i.i.d. sequence of random variables with distribution P_{θ_n}. The trick, as before, will be to write the power in terms of $Y_{i,n} = X_{i,n} - \theta_n$, which is distributed according to P_0. We can now see that

$$\begin{aligned}\pi_{1,n}(\theta_n) &= P_{\theta_n}\{\frac{\sqrt{n}\bar{X}_{n,n}}{\hat{\sigma}_{n,n}} > z_{1-\alpha}\} \\ &= P_0\{\frac{\sqrt{n}\bar{Y}_{n,n} + \sqrt{n}\theta_n}{\hat{\sigma}_{n,n}} > z_{1-\alpha}\} \\ &= P_0\{\frac{\sqrt{n}\bar{Y}_{n,n}}{\hat{\sigma}_{n,n}} > z_{1-\alpha} - \frac{h}{\hat{\sigma}_{n,n}}\} \ .\end{aligned}$$

Since the distribution of $Y_{i,n}$ is no longer changing with n, our analysis from before applies and we see that

$$\frac{\sqrt{n}\bar{Y}_{n,n}}{\hat{\sigma}_{n,n}} \stackrel{d}{\to} N(0,1)$$

under P_0. Since $\hat{\sigma}_{n,n}$ converges in probability under P_0 to σ_0^2, we have that

$$\pi_{1,n}(\theta_n) \to 1 - \Phi(z_{1-\alpha} - \frac{h}{\sigma_0}) \ .$$

This limit is called the local asymptotic power function of the t-test. Notice that it depends on the so-called local paramter h.

A remark on interpretation is warranted here. We are really only interested in the power of the test at a single $\theta > 0$, not a sequence θ_n. So, how should we use the above approximation in practice? Given a sample of size n and a $\theta > 0$, we can solve for the corresponding value of h by equating θ and θ_n. By dong so, we find that $h = \sqrt{n}\theta$. Plugging this value of h into the above expression, we get our approximation to the power of the test at θ.

Now let's consider the sign test. Begin as before by considering the behavior of

$$\frac{1}{n}\sum_{1 \leq i \leq n} I\{X_{i,n} > 0\}$$

under P_{θ_n}. Our earlier analysis shows that

$$E_{\theta_n}[I\{X_{i,n} > 0\}] = 1 - F(-\theta_n)$$

and

$$V_{\theta_n}[I\{X_{i,n} > 0\}] = F(-\theta_n)(1 - F(-\theta_n)) \ .$$

We'd like to assert that

$$S_n(\theta_n) = \frac{1}{\sqrt{n}} \sum_{1 \le i \le n} (I\{X_{i,n} > 0\} - (1 - F(-\theta_n)))$$

converges in distribution under P_{θ_n} to a normal distribution. To do this, we will need a central limit theorem for a triangular array. The most general such theorem is the Lindeberg-Feller central limit theorem. Here's a special case of it:

Theorem 4.1 For each n, let $Z_{n,i}, i = 1, \ldots, n$ be i.i.d. with distribution P_n. Suppose $E_n[Z_{n,i}] = 0$ and $V_n[Z_{n,i}] = \sigma_n^2 < \infty$. If for each $\epsilon > 0$

$$\lim_{n \to \infty} \frac{1}{\sigma_n^2} E_n[Z_{n,i}^2 I\{|Z_{n,i}| > \epsilon \sqrt{n} \sigma_n\}] = 0 \ ,$$

then

$$S_n = \sqrt{n} \bar{Z}_{n,n} / \sigma_n \xrightarrow{d} N(0,1)$$

under P_n.

For the general version of the Lindeberg-Feller central limit theorem see, for example, Theorem 11.2.5 of Romano and Lehmann (2005). For a proof see Theorem 27.2 of Billingsley (1995).

So let's apply the theorem with

$$\begin{aligned} Z_{n,i} &= I\{X_{i,n} > 0\} - (1 - F(-\theta_n)) \\ \sigma_n^2 &= F(-\theta_n)(1 - F(-\theta_n)) \ . \end{aligned}$$

For any fixed h, σ_n is also bounded away from 0 because $F(0) = \frac{1}{2}$, F is continuous by assumption (it's the integral of f), and $\theta_n \approx 0$ for large n. We also have that σ_n is bounded from above because F is bounded. Finally, we have that $Z_{n,i}$ is bounded because I and F are both bounded. Therefore, the condition required in the theorem holds trivially in this case. Since $\sigma_n^2 \to F(0)(1 - F(0)) = \frac{1}{4}$, we have that

$$S_n(\theta_n) \xrightarrow{d} N(0, \frac{1}{4})$$

under P_{θ_n}.

We can now finish our analysis for the sign test. We have that

$$\begin{aligned}\pi_{2,n}(\theta_n) &= P_{\theta_n}\{\frac{1}{\sqrt{n}}\sum_{1\le i\le n}(I\{X_i>0\}-\frac{1}{2})>\frac{1}{2}z_{1-\alpha}\} \\ &= P_{\theta_n}\{S_n(\theta_n)>\frac{1}{2}z_{1-\alpha}-\sqrt{n}(\frac{1}{2}-F(-\theta_n))\}\ .\end{aligned}$$

Since F is differentiable by assumption (with derivative equal to f), we see that

$$\sqrt{n}(\frac{1}{2}-F(-\theta_n))=\sqrt{n}(F(0)-F(-\theta_n))\approx\sqrt{n}\theta_n f(0)=hf(0)\ ,$$

assuming f is continuous at 0. Together with the result about the asymptotic normality of $S_n(\theta_n)$ above, we find that

$$\pi_{2,n}(\theta_n)\to 1-\Phi(z_{1-\alpha}-2hf(0))\ .$$

We are now (finally) in a position to compare these two tests based on their local asymptotic power functions. It is easy to see that if $2f(0)>\frac{1}{\sigma_0}$, then the sign test will be preferred to the t-test in a local asymptotic power sense; otherwise, the t-test will be preferred to the sign test.

If f is the normal density, then we know that the t-test should be uniformly most powerful for testing the null hypothesis. Reassuringly, if we plug in the standard normal density for f, we find that the above analysis bears this out. Likewise, if f is the density of a logistic or a uniform distribution, then the t-test is preferred to the sign test.

If, on the other hand, we consider distributions with "fatter" tails, we find that the situation is reversed. For example, if we take f to be the density of a Laplace distribution, the above analysis implies that the sign test is preferred to the t-test in a local asymptotic power sense. In fact, we can make the ratio of $2f(0)$ to $\frac{1}{\sigma_0}$ arbitrarily large by considering densities f with more and more mass in the tails. Thus, the moral of this story is that if the underlying distribution is symmetric, then, the t-test, while preferred for many distributions, is not as robust as the sign test to "fat" tails (and can in fact be arbitrarily worse than the sign test!).

The square of the ratio of $2f(0)$ to $\frac{1}{\sigma_0}$ is sometimes referred to as the asymptotic relative efficiency of the sign test w.r.t. the t-test. Asymptotic relative efficiency is defined analogously for other pairs of tests.