Stats 300C: Theory of Statistics Spring 2018

Lecture 4 — April 9, 2018

Prof. Emmanuel Candes

Scribe: Jeha Yang; Edited by: Nima Hamidi, Qian Zhao

1 Outline

Agenda: Global testing

1. Simes Test
2. Tests Based on Empirical CDF's:
   (a) Kolmogorov-Smirnov Test
   (b) Anderson-Darling Test
   (c) Tukey's Second-Level Significance Testing

Start: Sparse Mixtures

2 Simes Test [Simes 1987, Eklund]:

As usual, we have n hypotheses $H_{0,i}$ and p values p_i for each. Under $H_{0,i}$, $p_i \sim U(0,1)$. We are interested in testing the global null $H_0 = \bigcap_i H_{0,i}$.

As before we start with n p-values and order them

$$p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(n)}$$

The Simes statistic is

$$T_n = \min_i \left\{ p_{(i)} \frac{n}{i} \right\}$$

n/i is an adjustment factor.

Theorem 1. *Under H_0 and independence of the p_i,*

$$T_n \sim U(0,1).$$

Thus the Simes test rejects H_0 if $T_n \leq \alpha$.

Equivalent Formulation: Reject H_0 if

$$\exists i : \ \ p_{(i)} \leq \alpha \frac{i}{n}$$

In fact we don't really need independence for this test to have level α. We merely need a sort of positive dependence that we will define later in the course [Sarkar 1988].

The Simes procedure is strictly less conservative than Bonferroni, which rejects when $p_{(1)} \leq \frac{\alpha}{n}$.

Example: $n = 2$: Simes rejects if $p_{(1)} \leq \alpha/2$ *or* $p_{(2)} \leq \alpha$. Below we plot the rejection regions of each test:

We can easily check that in this case, the size of Bonferroni is $\alpha - \alpha^2/4$, while the size of Simes is α. Nevertheless, Simes still tends to look at lower p-values, since higher p-values are unlikely to be less or equal to $\alpha\frac{i}{n}$. Now let's prove the theorem above.

Proof. By induction: The theorem is true for $n = 1$ by inspection.

Assume that it is true for $n - 1$. Then $T_{n-1} \sim U(0, 1)$. We have

$$p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(n)}.$$

The density of $p_{(n)}$ is

$$f(t) = nt^{n-1}$$

for $t \in [0, 1]$. Then

$$\begin{aligned}\mathbb{P}(T_n \leq \alpha) &= \int_0^1 \mathbb{P}(T_n \leq \alpha | p_{(n)} = t) f(t)\, dt \\ &= \int_0^\alpha f(t)\, dt + \int_\alpha^1 \mathbb{P}(T_n \leq \alpha | p_{(n)} = t) f(t)\, dt.\end{aligned}$$

We first handle the second integral. Conditional on $p_{(n)} = t$, the other p values are independently uniform on $U(0, t)$, so if we divide them by t, we can apply the inductive hypothesis, once we observe:

$$\min_{1 \leq i \leq n-1} p_{(i)} \frac{n}{i} \leq \alpha \quad \Leftrightarrow \quad \min_{1 \leq i \leq n-1} \frac{p_{(i)}}{t} \cdot \frac{n-1}{i} \leq \frac{\alpha}{t} \cdot \frac{n-1}{n}$$

Therefore, $\mathbb{P}(T_n \leq \alpha | p_{(n)} = t) = \frac{\alpha}{t} \cdot \frac{n-1}{n}$ for $t \geq \alpha$.

Then

$$\begin{aligned}
\mathbb{P}(T_n \leq \alpha) &= \int_0^\alpha nt^{n-1}\,dt + \int_\alpha^1 \frac{\alpha}{t} \cdot \frac{n-1}{n} nt^{n-1}\,dt \\
&= \int_0^\alpha nt^{n-1}\,dt + \int_\alpha^1 \frac{\alpha}{t} \cdot \frac{n-1}{n} nt^{n-1}\,dt \\
&= \alpha^n + \alpha \int_\alpha^1 (n-1)t^{n-2}\,dt \\
&= \alpha^n + \alpha[1 - \alpha^{n-1}] \\
&= \alpha.
\end{aligned}$$

$\square$

Summary: The Simes procedure is powerful for a single strong effect, but has moderate power for many mild effects.

3 Tests Based on Empirical CDF's

Define *empirical* CDF of $p_1, \cdots, p_n$ as

$$\hat{F}_n(t) = \frac{1}{n} \#\{i : \, p_i \leq t\}.$$

Under the global null H_0 we have that

$$\mathbb{E}(\hat{F}_n(t)) = t.$$

Moreover, if we assume that p_i's are independent, $n\hat{F}_n(t)$ is a binomial random variable with parameter t.

Now, the idea is that, under the global null, $\hat{F}_n(t)$ should be around t. Hence, we measure the distance between what we observe and what we expect and reject if the difference is large.

3.1 Kolmogorov-Smirnov Test

Define the following test statistics

$$KS = \sup_t |\hat{F}_n(t) - t|,$$

and reject if KS exceeds certain threshold. The threshold can be computed using simulations or asymptotic calculations. We are usually interested in cases when $\sup_t(\hat{F}_n(t) - t)$ is large.

3.2 Anderson-Darling Test

Let weight function $\omega(t)$ be a non-negative function. Then, define the following statistic

$$A^2 = n \int_0^1 \left(\hat{F}_n(t) - t\right)^2 \omega(t)\, dt.$$

When $\omega(t) = 1$, the above statistic is called the *Cramer-von Mises* statistic. Anderson and Darling (1954) suggested the weight function $\omega(t) = \left[t(1-t)\right]^{-1}$ as to standardize $(\hat{F}_n(t) - t)$ which is binomial for independent p_i. This puts more weight on small/large p values than the Cramer-von Mises statistic. Thus, our statistics in this special case is

$$A^2 = n \int_0^1 \frac{\left(\hat{F}_n(t) - t\right)^2}{t(1-t)}\, dt.$$

A useful relation that holds about this statistic is the following

$$A^2 = -n - \sum_{i=1}^{n} \frac{2i-1}{n} \left[\log(p_{(i)}) + \log(1 - p_{(n+1-i)})\right].$$

Anderson-Darling gives more weight to p-values in the bulk than Fisher's statistic.

3.3 Tukey's Second-Level Significance Testing (1976)

Second-Level Significance Testing: [Tukey 1976] Define the **Higher Criticism Statistic**

$$HC_n^* = \max_{0 \le t \le \alpha_0} \frac{\hat{F}_n(t) - t}{\sqrt{t(1-t)/n}}$$

The difference between this test and Anderson-Darling statistic is that this uses a maximum value rather than a (squared) average.

Define statistics $HC_n(t)$ as

$$HC_n(t) = \frac{\hat{F}_n(t) - t}{\sqrt{t(1-t)/n}} = \frac{\#\{\text{significance of level t}\} - nt}{\sqrt{nt(1-t)}},$$

then HC_n^* scans across significance levels for departure from H_0. Hence, a large value of HC_n^* indicates significance of an overall body of tests.

4 Sparse Mixtures

Original Model: We have independent statistics X_i distributed as

$$\begin{aligned} H_{0,i}&: \; X_i \sim N(0,1) \\ H_{1,i}&: \; X_i \sim N(\mu_i, 1), \quad \mu_i > 0 \end{aligned}$$

Here we consider a framework in which we are interested in possibilities within H_1 with a small fraction of non-null hypotheses. Rather than directly saying that there are some amount of nonzero means under H_1, we assume that our samples follow a mixture of $N(0,1)$ and $N(\mu, 1)$ with μ fixed, resulting in the following:

Simple Model with Equal Means:

$$\begin{aligned} H_0 &: \ X_i \overset{\text{i.i.d.}}{\sim} N(0,1) \\ H_1 &: \ X_i \overset{\text{i.i.d.}}{\sim} (1-\varepsilon)N(0,1) + \varepsilon N(\mu,1) \end{aligned}$$

Put another way, there are about $n\varepsilon$ non-nulls under H_1.

The likelihood ratio for this model is then

$$L = \prod_{i=1}^{n} \left[(1-\varepsilon) + \varepsilon e^{\mu X_i - \mu^2/2} \right].$$

Asymptotic Analysis [Ingster 99, Jin 03]: To carry out asymptotic analysis, we must specify the dependence scheme of ε and μ on n. Ingster (99) and Jin (03) considered

$$\begin{aligned} \varepsilon_n &= n^{-\beta} & \quad \frac{1}{2} < \beta < 1 \\ \mu_n &= \sqrt{2r \log n} & \quad 0 < r < 1 \end{aligned}$$

This automatically incorporates the settings that are explored in the past few lectures: the needle in a haystack problem corresponds to $\beta = 1$ and $r = 1$; the small distributed effects case corresponds to $\beta = 1/2$. So we are actually studying situations in between those two cases.

Threshold Effect: Ingster and Jin found that there is a threshold curve for r of the form

$$\rho^*(\beta) = \begin{cases} \beta - 1/2 & \frac{1}{2} < \beta \leq \frac{3}{4} \\ (1 - \sqrt{1-\beta})^2 & \frac{3}{4} \leq \beta \leq 1 \end{cases}$$

such that

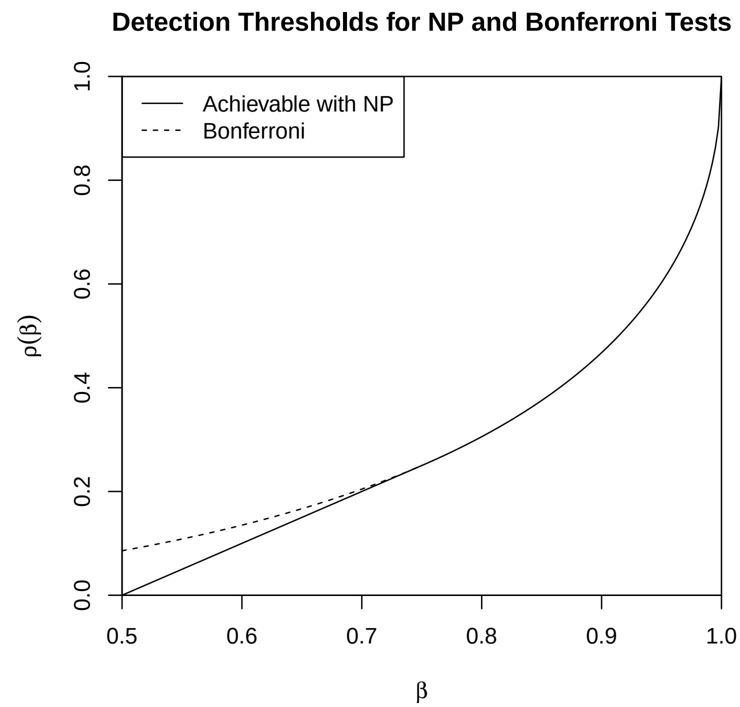
1. If $r > \rho^*(\beta)$ we can adjust the NP test to achieve

$$\mathbb{P}_0(\text{Type I Error}) + \mathbb{P}_1(\text{Type II Error}) \to 0$$

2. If $r < \rho^*(\beta)$ then for *any* test

$$\liminf_n \quad \mathbb{P}_0(\text{Type I Error}) + \mathbb{P}_1(\text{Type II Error}) \geq 1$$

Higher Criticism statistic: Donoho and Jin proved for $r > \rho^*(\beta)$ in the sparse mixture setting, the higher criticism statistic with a proper threshold has full power asymptotically. This is interesting because the HC statistic does not need knowledge of ϵ and/or μ.


Detection Thresholds for NP and Bonferroni Tests
Achievable with NP
Bonferroni
ρ(β)
β
0.0
0.2
0.4
0.6
0.8
1.0
0.5
0.6
0.7
0.8
0.9
1.0