DSO 699: Statistics Theory
Special Topics in Data Sciences an Operations

Week 2

Adel Javanmard

USC Marshall
Department of Data Sciences and Operations

September 1, 2023

Last class

We talked about:

- **Inference thinking:** going from data to the underlying phenomena behind the data
- **Hypothesis testing:** Concepts of
  - ✓ null hypothesis, alternative hypothesis
  - ✓ test statistic
  - ✓ p-value
  - ✓ ROC curves, AUC as measure of performance
- Multiple hypotheses testing
- Challenges of multiple testing

Outline for today

1. Global testing: introduction and methods

2. Optimality of Bonferroni's Method against Sparse Alternatives

Multiple testing: an example

	Healthy (n_0)	Prostate cancer (n_1)
Expression level of gene i	$Y_{ij}^{(0)}, 1 \leq j \leq n_0$	$Y_{ij}^{(1)}, 1 \leq j \leq n_1$

- 20K genes in the human body
- hundreds of healthy patients
- 50-100 patients with prostate cancer ($n_1 \ll n_0$)

$$H_{0,i} : \mathbb{E}[\,Y_{ij}^{(0)}] = \mathbb{E}[\,Y_{ij}^{(1)}]\,,$$

or

$$H_{0,i} :\ Y_{ij}^{(0)} \stackrel{d}{=}\ Y_{ij}^{(1)}\,.$$

Multiple testing: an example

	Healthy (n_0)	Prostate cancer (n_1)
Expression level of gene i	$Y_{ij}^{(0)}, 1 \le j \le n_0$	$Y_{ij}^{(1)}, 1 \le j \le n_1$

- 20K genes in the human body
- hundreds of healthy patients
- 50-100 patients with prostate cancer ($n_1 \ll n_0$)

$$H_{0,i} : \mathbb{E}[\,Y_{ij}^{(0)}] = \mathbb{E}[\,Y_{ij}^{(1)}]\,,$$

or

$$H_{0,i} :\ Y_{ij}^{(0)} \overset{d}{=}\ Y_{ij}^{(1)}\,.$$

In this lecture we focus on the simplest task in multiple testing, which is global testing.

Global testing: introduction and methods

Global testing

We would like to test the global null

$$H_0 = \cap_{i=1}^{n} H_{0,i} \,,$$

which holds if and only if all the individual nulls are true.
Suppose that for each $H_{0,i}$ we already have a p-value, $p_i \sim U[0,1]$.

Global testing

We would like to test the global null

$$H_0 = \cap_{i=1}^{n} H_{0,i} \, ,$$

which holds if and only if all the individual nulls are true.
Suppose that for each $H_{0,i}$ we already have a p-value, $p_i \sim U[0,1]$.

How to combine the individual p-values to test the global null H_0?

- Bonferroni's method
- Fisher's combination test

Bonferroni's method

Procedure. Given a desired level α, reject the global null whenever

$$\min_{1 \leq i \leq n} p_i \leq \alpha / n\,.$$

Size of the test?

Bonferroni's method

Procedure. Given a desired level α, reject the global null whenever

$$\min_{1 \le i \le n} p_i \le \alpha/n\,.$$

Size of the test?

$$\begin{aligned}\mathbb{P}_{H_0}(\text{type I error}) &= \mathbb{P}_{H_0}\big(\min_{1\le i\le n} p_i \le \alpha/n\big)\\ &= \mathbb{P}_{H_0}\big(\cup_{i=1}^{n}\mathbb{I}(p_i \le \alpha/n)\big)\\ &\le \sum_{i=1}^{n}\mathbb{P}_{H_0}(p_i \le \alpha/n) = \sum_{i=1}^{n}\alpha/n = \alpha\,.\end{aligned}$$

Bonferroni's method

Procedure. Given a desired level α, reject the global null whenever

$$\min_{1 \le i \le n} p_i \le \alpha / n \,.$$

Size of the test?

$$\begin{aligned}
\mathbb{P}_{H_0}(\text{type I error}) &= \mathbb{P}_{H_0}\big(\min_{1 \le i \le n} p_i \le \alpha/n\big) \\
&= \mathbb{P}_{H_0}\big(\cup_{i=1}^{n} \mathbb{I}(p_i \le \alpha/n)\big) \\
&\le \sum_{i=1}^{n} \mathbb{P}_{H_0}(p_i \le \alpha/n) = \sum_{i=1}^{n} \alpha/n = \alpha \,.
\end{aligned}$$

Pros. Controls the size regardless of the dependence structure of p-values (union bounding)

Is Bonferroni's test conservative?

Recall that conservative means size(test) $\ll \alpha$ (target level)

A misconception is that Bonferroni's method is conservative!

- *Why a misconception?*

Is Bonferroni's test conservative?

Recall that conservative means size(test) $\ll \alpha$ (target level)

A misconception is that Bonferroni's method is conservative!

- *Why a misconception?*

Suppose that $H_{0,i}$ are independent. Then,

$$\begin{aligned}
\mathbb{P}_{H_0}(\text{type I error}) &= 1 - \mathbb{P}_{H_0}\big(\min_{1\le i\le n} p_i > \alpha/n\big) \\
&= 1 - \mathbb{P}_{H_0}(\cap_{i=1}^n \mathbb{I}(p_i > \alpha/n)) \\
&= 1 - (1-\alpha/n)^n \\
&\overset{n\to\infty}{\longrightarrow} 1 - e^{-\alpha} \approx \alpha
\end{aligned}$$

Is Bonferroni's test conservative?

Recall that conservative means size(test) $\ll \alpha$ (target level)

A misconception is that Bonferroni's method is conservative!

- *Why a misconception?*

Suppose that $H_{0,i}$ are independent. Then,

$$
\begin{aligned}
\mathbb{P}_{H_0}(\text{type I error}) &= 1 - \mathbb{P}_{H_0}\big(\min_{1\le i\le n} p_i > \alpha/n\big) \\
&= 1 - \mathbb{P}_{H_0}\big(\cap_{i=1}^{n}\mathbb{I}(p_i > \alpha/n)\big) \\
&= 1 - (1-\alpha/n)^n \\
&\overset{n\to\infty}{\longrightarrow} 1 - e^{-\alpha} \approx \alpha
\end{aligned}
$$

For small α and large n, size can get close to α.
(e.g., for $\alpha = 0.05$, we have $1 - e^{-\alpha} = 0.0488$)

Fisher's combination test

Another method for global testing!

Procedure. Given the individual p-values p_i,

- computes the statistics

$$T = -2 \sum_{i=1}^{n} \log p_i$$

- Reject the global null when $T > \chi^2_{2n}(1-\alpha)$.

Fisher's combination test

Size if Fisher's combination test:

Proposition

Suppose p_i are independent. Then under the global null H_0 we have $T \sim \chi^2_{2n}$ and so the size of Fisher's test is controlled at level α.

Fisher's combination test

Size if Fisher's combination test:

Proposition

Suppose p_i are independent. Then under the global null H_0 we have $T \sim \chi^2_{2n}$ and so the size of Fisher's test is controlled at level α.

Proof:

- $p_i \sim U[0,1] \implies -\log p_i \sim \mathrm{Exp}(1)$.
- $X \sim \mathrm{Exp}(1) \implies 2X \sim \chi^2_2$.
- Sum of independent $\chi^2_{n_i}$ is also a $\chi^2_{\sum n_i}$.
- Therefore, $T \sim \chi^2_{2n}$.

size is controlled only for independent p-values!

Fisher's combination test (case of dependent p-values)

When the p-values are no dependent, the null distribution of T is more complicated.

Idea of scaled χ^2:

- Approximating T using a scaled χ^2-distribution, $c\chi^2(n')$, with n' degrees of freedom.
- The mean and variance of this scaled χ^2 are:

$$\mathbb{E}[c\chi^2(n')] = cn' \,,$$
$$\text{Var}[c\chi^2(n')] = 2c^2 n' \,.$$

- Matching the first two moments with data:

$$c = \frac{\text{Var}(T)}{2\mathbb{E}[T]} \qquad n' = \frac{2(\mathbb{E}[T])^2}{\text{Var}(T)} \,.$$

[Brown's method, Kost's method]

Fisher's combination test (case of dependent p-values)

Harmonic mean p-value (HMP) technique:

- For a set of weights $w_1, \ldots, w_n$ the weighted harmonic mean is given by

$$\mathring{p} = \left(\frac{\sum_{i=1}^{n} w_i p_i^{-1}}{\sum_{i=1}^{n} w_i} \right)^{-1}$$

- Reject the global null if $\mathring{p} \leq \alpha \sum_{i=1}^{n} w_i$.

Fisher's combination test (case of dependent p-values)

Harmonic mean p-value (HMP) technique:

- For a set of weights $w_1, \ldots, w_n$ the weighted harmonic mean is given by

$$\mathring{p} = \left(\frac{\sum_{i=1}^{n} w_i p_i^{-1}}{\sum_{i=1}^{n} w_i} \right)^{-1}$$

- Reject the global null if $\mathring{p} \le \alpha \sum_{i=1}^{n} w_i$.

The procedure is valid for any choice of weights, but the power will be affected by this choice.

Often choose $w_i = 1/n$.

[The harmonic mean p-value for combining dependent tests Daniel J. Wilsona,PNAS 2019]

Application in meta-analysis

- You can (and often is the case) apply Fisher's method to a collection of independent test statistics, usually from separate studies having the same null hypothesis.
- In this case the test helps to combine evidence from different studies/ environments against null hypothesis.
- Sometimes it makes sense to consider the possibility of "heterogeneity," in which the null hypothesis holds in some studies but not in others, or where different alternative hypotheses may hold in different studies.

Comparison between Bonferroni and Fisher's methods

Bonferroni's method: Reject the global null if

$$\min_{1 \le i \le n} p_i \le \alpha / n \,.$$

Fisher's method: Construct the statistic

$$T = -2 \sum_{i=1}^{n} \log p_i \,.$$

Comparison between Bonferroni and Fisher's methods

Bonferroni's method: Reject the global null if

$$\min_{1\le i\le n} p_i \le \alpha/n\,.$$

Fisher's method: Construct the statistic

$$T = -2\sum_{i=1}^{n}\log p_i\,.$$

- Bonferroni's method works better for detecting *a few large changes* in the individual tests
- Fisher's test works better for detecting *many subtle* changes.

How to formalize this intuition?

Optimality of Bonferroni's Method against Sparse Alternatives

Gaussian sequence model

Consider an independent Gaussian sequence model:

$$X_i \sim \mathrm{N}(\mu_i, 1), \quad i = 1, \ldots n\,,$$

where we are interested in the n hypotheses

$$H_{0,i} : \mu_i = 0\,.$$

Gaussian sequence model

Consider an independent Gaussian sequence model:

$$X_i \sim \mathrm{N}(\mu_i, 1), \quad i = 1, \ldots n\,,$$

where we are interested in the n hypotheses

$$H_{0,i} : \mu_i = 0\,.$$

Bonferroni's method rejects the global null if:

- (one sided: $H_a : \mu_i > 0$)

$$p_i = 1 - \Phi(X_i) \Longrightarrow \max_{1 \le i \le n} X_i > \Phi^{-1}(1 - \tfrac{\alpha}{n}) := z(\tfrac{\alpha}{n})$$

Gaussian sequence model

Consider an independent Gaussian sequence model:

$$X_i \sim \mathrm{N}(\mu_i, 1), \quad i = 1, \dots n\,,$$

where we are interested in the n hypotheses

$$H_{0,i} : \mu_i = 0\,.$$

Bonferroni's method rejects the global null if:

- (one sided: $H_a : \mu_i > 0$)

$$p_i = 1 - \Phi(X_i) \Longrightarrow \max_{1 \le i \le n} X_i > \Phi^{-1}(1 - \tfrac{\alpha}{n}) := z(\tfrac{\alpha}{n})$$

- (two sided: $H_a : \mu_i \neq 0$)

$$p_i = 2(1 - \Phi(|X_i|)) \Longrightarrow \max_{1 \le i \le n} |X_i| > \Phi^{-1}(1 - \tfrac{\alpha}{2n}) := z(\tfrac{\alpha}{2n})$$

Magnitude of Bonferroni's threshold

Why does the Bonferroni's threshold makes sense?

Magnitude of Bonferroni's threshold

Why does the Bonferroni's threshold makes sense?

- Of course we already gave a union bound argument!

Magnitude of Bonferroni's threshold

Why does the Bonferroni's threshold makes sense?

- Of course we already gave a union bound argument!
- We know that under H_0, $X_i \sim N(0,1)$, i.i.d, and so

$$\frac{\max_{1\le i\le n} X_i}{\sqrt{2\log n}} \to 1 \quad \text{(in probability)}$$

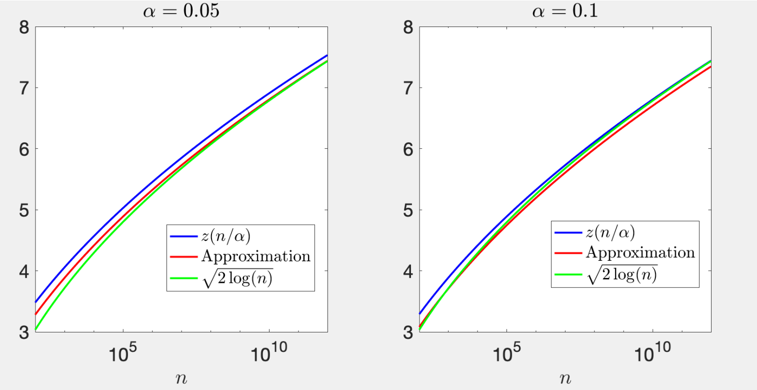
- Interestingly, one can show

$$z\left(\frac{\alpha}{n}\right) = \sqrt{2\log n}(1+o(1))\,.$$

More specifically,

$$z\left(\frac{\alpha}{n}\right) \approx \sqrt{B\left(1-\frac{\log B}{B}\right)} \qquad B = 2\log(n/\alpha) - \log(2\pi)\,.$$

Approximating z-values



Asymptotic Power of Bonferroni's method

Needle in a haystack problem:

The alternative is that *one* $\mu_i = \mu > 0$ but we don't know for which i.

Asymptotic Power of Bonferroni's method

Needle in a haystack problem:

The alternative is that *one* $\mu_i = \mu > 0$ but we don't know for which i.

A sharp detection threshold for Bonferroni's method:

- *Asymptotic full power above threshold:* Suppose $\mu > (1+\varepsilon)\sqrt{2\log n}$. Then,

$$\begin{aligned}\mathbb{P}_{H_1}\left(\max X_i > z(\alpha/n)\right) &\geq \mathbb{P}\left(X_1 > z(\alpha/n)\right) \\ &= \mathbb{P}\left(z_1 > z(\alpha/n) - \mu\right) \to 1\end{aligned}$$

  because $z(\alpha/n) - \mu \to -\infty$ as $n \to \infty$.

Asymptotic Power of Bonferroni's method (con't)

Needle in a haystack problem:

A sharp detection threshold for Bonferroni's method:

- *Asymptotic powerlessness below threshold:* Suppose $\mu < (1-\varepsilon)\sqrt{2\log n}$. Then,

$$\begin{aligned}
&\mathbb{P}_{H_1}\left(\max X_i > z(\alpha/n)\right) \\
&\le \mathbb{P}\left(X_1 > z(\alpha/n) - \mu\right) + \mathbb{P}\left(\max_{2\le i\le n} z_i > z(\alpha/n)\right) \\
&\le \mathbb{P}\left(X_1 > z(\alpha/n) - \mu\right) + \mathbb{P}\left(\max_{1\le i\le n} z_i > z(\alpha/n)\right) \\
&= 0 + 1 - (1-\alpha/n)^n \to 1 - e^{-\alpha} \approx \alpha\,.
\end{aligned}$$

In this case, the test is as bad as random guessing.

Asymptotic Power of Bonferroni's method (con't)

Needle in a haystack problem:

A sharp detection threshold for Bonferroni's method:

- *Asymptotic powerlessness below threshold:* Suppose $\mu < (1-\varepsilon)\sqrt{2\log n}$. Then,

$$\begin{aligned}
&\mathbb{P}_{H_1}\left(\max X_i > z(\alpha/n)\right) \\
&\leq \mathbb{P}\left(X_1 > z(\alpha/n) - \mu\right) + \mathbb{P}\left(\max_{2\leq i\leq n} z_i > z(\alpha/n)\right) \\
&\leq \mathbb{P}\left(X_1 > z(\alpha/n) - \mu\right) + \mathbb{P}\left(\max_{1\leq i\leq n} z_i > z(\alpha/n)\right) \\
&= 0 + 1 - (1-\alpha/n)^n \to 1 - e^{-\alpha} \approx \alpha\,.
\end{aligned}$$

  In this case, the test is as bad as random guessing.

- *Optimality?* We next show that there is no test with useful power in this case.

Optimality of Bonferroni's method

Claim: The decision threshold $\mu = \sqrt{2 \log n}$ cannot be improved using any test of the global null against the "needle in a haystack" alternative.

Optimality of Bonferroni's method

Claim: The decision threshold $\mu = \sqrt{2 \log n}$ cannot be improved using any test of the global null against the “needle in a haystack” alternative.

“Bayesian” Decision Problem:

- A trick to reduce composite alternative to a simple alternative
- Consider

$$H_0^n : \ \mu_i = 0 \text{ for all } i$$
$$H_1^n : \ \{\mu_i\}_{i=1}^n \sim \pi \,,$$

  where π selects an index i^* uniformly at random and sets $\mu_{i^*} = (1-\varepsilon)\sqrt{2 \log n}$ with all other $\mu_i = 0$.
- Since null and alternative hypotheses are simple, we can apply the Neyman-Pearson Lemma!

Neyman-Pearson Lemma

(*Simple hypothesis:* if the hypothesis uniquely specifies the distribution of the population from which the sample is taken)

Lemma

Consider testing simple hypothesis $H_0 : \mathbb{P} = \mathbb{P}_0$ (or $\theta = \theta_0$) against simple alternative $H_a : \mathbb{P} = \mathbb{P}_1$ (or $\theta = \theta_1$). Then the likelihood ratio test (LRT) is the most powerful test and takes the following form. Given sampled data $\boldsymbol{X}$

$$R(\boldsymbol{X}) = \begin{cases} 1 & \text{if } \frac{\mathbb{P}_1(\boldsymbol{X})}{\mathbb{P}_0(\boldsymbol{X})} > \eta\,, \\ 0 & \text{if } \frac{\mathbb{P}_1(\boldsymbol{X})}{\mathbb{P}_0(\boldsymbol{X})} < \eta\,, \end{cases}$$

where threshold $\eta > 0$ is chosen so that $\mathbb{P}_0(R(\boldsymbol{X}) = 1) = \alpha$, for the prefixed significance level α.

LRT for “needle in a haystack” model

The null and alternative distributions are given by (for $\mu_n = (1-\varepsilon)\sqrt{2\log n}$)

$$\mathbb{P}_0^n(\boldsymbol{X}) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}X_j^2}$$

$$\mathbb{P}_1^n(\boldsymbol{X}) = \frac{1}{n}\sum_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(X_i-\mu_n)^2} \prod_{j:j\neq i} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}X_j^2} \right)$$

Therefore, the likelihood ratio is given by

$$L_n = \frac{\mathbb{P}_1^n}{\mathbb{P}_0^n} = \frac{1}{n}\sum_{i=1}^{n} e^{X_i\mu_n - \frac{1}{2}\mu_n^2} \quad and \quad \mathbb{E}_{H_0}[L_n] = 1\,.$$

LRT for “needle in a haystack” model (con’t)

To prove LRT (and so any other test) is powerless in this regime, we show:

Lemma

If $\mu_n = (1-\varepsilon)\sqrt{2\log n}$, then $L_n \xrightarrow{p} 1$.

LRT for "needle in a haystack" model (con't)

To prove LRT (and so any other test) is powerless in this regime, we show:

Lemma

If $\mu_n = (1-\varepsilon)\sqrt{2\log n}$, then $L_n \xrightarrow{p} 1$.

First impulse:

$$L_n = \frac{1}{n}\sum_{i=1}^{n} e^{X_i\mu_n - \frac{1}{2}\mu_n^2}.$$

Under H_0, this is is sum of i.i.d terms. How about WLLN?

Let's bound the variance

Let $Z_i := e^{X_i \mu_n - \frac{1}{2}\mu_n^2}$. Then,

E^2 = 1>0, so inequality holds

$$\begin{aligned}
\mathrm{Var}_0(L) = \frac{1}{n}\mathrm{Var}_0(Z_1) &\le \frac{1}{n}\mathbb{E}_0[Z_1^2] \\
&= \frac{1}{n}\int_{-\infty}^{\infty} e^{-\mu_n^2} e^{2\mu_n x}\phi(x)\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}n}\int_{-\infty}^{\infty} e^{-\mu_n^2} e^{2\mu_n x - \frac{1}{2}x^2}\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}n} e^{\mu_n^2}\int_{-\infty}^{\infty} e^{-(x-2\mu_n)^2/2}\mathrm{d}x = \frac{1}{n}e^{\mu_n^2}
\end{aligned}$$

just definition

Let's bound the variance

Let $Z_i := e^{X_i \mu_n - \frac{1}{2}\mu_n^2}$. Then,

$$\begin{aligned}
\mathrm{Var}_0(L) = \frac{1}{n}\mathrm{Var}_0(Z_1) &\le \frac{1}{n}\mathbb{E}_0[Z_1^2] \\
&= \frac{1}{n}\int_{-\infty}^{\infty} e^{-\mu_n^2} e^{2\mu_n x}\phi(x)\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}n}\int_{-\infty}^{\infty} e^{-\mu_n^2} e^{2\mu_n x - \frac{1}{2}x^2}\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}n} e^{\mu_n^2}\int_{-\infty}^{\infty} e^{-(x-2\mu_n)^2/2}\mathrm{d}x = \frac{1}{n}e^{\mu_n^2}
\end{aligned}$$

But,

$$\frac{1}{n}e^{\mu_n^2} = n^{2(1-\varepsilon)^2 - 1} \to \infty$$

if $\varepsilon < 1 - \frac{1}{\sqrt{2}}$:(

Let's bound the varaince

Truncation trick!

Recall that $L = \frac{1}{n}\sum_{i=1}^{n} Z_i$ with $Z_i = e^{X_i\mu_n - \frac{1}{2}\mu_n^2}$.

Let $T_n = \sqrt{2\log n}$ and write

$$\tilde{L} = \frac{1}{n}\sum_{i=1}^{n} Z_i \mathbb{I}(X_i \le T_n)$$

We have

$$\mathbb{P}(L \neq \tilde{L}) \le \mathbb{P}(\max X_i \ge T_n) \to 0\,.$$

So we can work with $\tilde{L}$ instead of L.

Truncation trick!

$$\begin{aligned}\mathbb{E}_0[\tilde{L}] = \mathbb{E}_0[Z_1 \mathbb{I}(X_1 \le T_n)] &= \int_{-\infty}^{T_n} e^{\mu x - \mu^2/2} \phi(x) \mathrm{d}x \\ &= \int_{-\infty}^{T_n} \frac{1}{\sqrt{2\pi}} e^{-(x-\mu_n)^2/2} \mathrm{d}x \\ &= \Phi(T_n - \mu_n) = \Phi(\varepsilon\sqrt{2\log n}) \to 1\end{aligned}$$

Truncation trick!

asymptotically same mean

$$
\begin{aligned}
\mathbb{E}_0[\tilde{L}] = \mathbb{E}_0[Z_1 \mathbb{I}(X_1 \le T_n)] &= \int_{-\infty}^{T_n} e^{\mu x - \mu^2/2} \phi(x) \mathrm{d}x \\
&= \int_{-\infty}^{T_n} \frac{1}{\sqrt{2\pi}} e^{-(x-\mu_n)^2/2} \mathrm{d}x \\
&= \Phi(T_n - \mu_n) = \Phi(\varepsilon \sqrt{2 \log n}) \to 1
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}_0(\tilde{L}) = \frac{1}{n} \mathrm{Var}_0(Z_1 \mathbb{I}(X_1 \le T_n)) &\le \frac{1}{n} \mathbb{E}_0[Z_1^2 \mathbb{I}(X_1 \le T_n)] \\
&= \frac{1}{n} \int_{-\infty}^{T_n} e^{-\mu_n^2} e^{2\mu_n x} \phi(x) \mathrm{d}x \\
&= \frac{1}{n} e^{\mu_n^2} \Phi(T_n - 2\mu_n)
\end{aligned}
$$

Truncation trick!

$$\mathbb{E}_0[\tilde{L}] = \mathbb{E}_0[Z_1 \mathbb{I}(X_1 \le T_n)] = \int_{-\infty}^{T_n} e^{\mu x - \mu^2/2} \phi(x) \mathrm{d}x$$
$$= \int_{-\infty}^{T_n} \frac{1}{\sqrt{2\pi}} e^{-(x-\mu_n)^2/2} \mathrm{d}x$$
$$= \Phi(T_n - \mu_n) = \Phi(\varepsilon\sqrt{2\log n}) \to 1$$

$$\mathrm{Var}_0(\tilde{L}) = \frac{1}{n} \mathrm{Var}_0(Z_1 \mathbb{I}(X_1 \le T_n)) \le \frac{1}{n} \mathbb{E}_0[Z_1^2 \mathbb{I}(X_1 \le T_n)]$$
$$= \frac{1}{n} \int_{-\infty}^{T_n} e^{-\mu_n^2} e^{2\mu_n x} \phi(x) \mathrm{d}x$$
$$= \frac{1}{n} e^{\mu_n^2} \Phi(T_n - 2\mu_n)$$
$$\le \frac{1}{n} e^{\mu_n^2} \to 0$$

damping

if $\varepsilon > 1 - \frac{1}{\sqrt{2}}$

Let's bound the variance

Let's focus on $\varepsilon < 1 - \frac{1}{\sqrt{2}}$. So

$$2\mu_n - T_n = (2(1-\varepsilon) - 1)\sqrt{2\log n} = (1-2\varepsilon)\sqrt{2\log n} > 0\,.$$

Using inequality $\Phi(-t) \leq \frac{\phi(t)}{t}$, for $t > 0$ ....)

asymptotically cancels the rest

$$\begin{aligned}
\mathrm{Var}_0(\tilde{L}) &\leq \frac{1}{n} e^{\mu_n^2} \Phi(T_n - 2\mu_n) \\
&\leq \frac{1}{n} e^{\mu_n^2} \frac{\phi(2\mu_n - T_n)}{2\mu_n - T_n} = \frac{1}{n} e^{(1-\varepsilon)^2 T_n^2} \frac{1}{\sqrt{2\pi}} \frac{e^{-(1-2\varepsilon)^2 T_n^2/2}}{(1-2\varepsilon) T_n} \\
&= \frac{1}{\sqrt{2\pi} n} \frac{e^{(1-2\varepsilon^2) T_n^2/2}}{(1-2\varepsilon) T_n} = \frac{1}{\sqrt{2\pi}} \frac{n^{-2\varepsilon^2}}{(1-2\varepsilon)\sqrt{2\log n}} \to 0\,.
\end{aligned}$$

exclude many large values from var, since they are squared

Are we done?

Almost!

Are we done?

Almost!

Proposition

Set the threshold $\eta_n(\alpha)$ so that $\mathbb{P}_0(L_n \geq \eta_n(\alpha)) = \alpha$. Then,

$$\lim_{n \to \infty} \mathbb{P}(\text{Type II error}) = 1 - \alpha\,.$$

(so LRT has asymptotic power α.)

Are we done?

Almost!

Proposition

Set the threshold $\eta_n(\alpha)$ so that $\mathbb{P}_0(L_n \geq \eta_n(\alpha)) = \alpha$. Then,

$$\lim_{n\to\infty} \mathbb{P}(\text{Type II error}) = 1 - \alpha\,.$$

(so LRT has asymptotic power α.)

Since $L_n \xrightarrow{p} 1$, the likelihood ration of data under null and under alternative converges together, and there is no way to distinguish them.
In other words, the optimal test (Neyman-Pearson) is as good as random guessing.

Putting things together ....

Corollary

For **Bayesian decision problem**, if $\mu_n = (1 - \varepsilon)\sqrt{2 \log n}$, then the *optimal test* has

$$\mathbb{P}_{H_0}(\text{Type I error}) + \mathbb{P}_{H_1}(\text{Type II error}) \to 1\,.$$

Putting things together ....

Corollary

For **Bayesian decision problem**, if $\mu_n = (1-\varepsilon)\sqrt{2\log n}$, then the *optimal test* has

implies it's random guess

$$\mathbb{P}_{H_0}(\text{Type I error}) + \mathbb{P}_{H_1}(\text{Type II error}) \to 1\,.$$

How about composite alternative we started with?

Note that for any test,

$$\mathbb{P}_{H_0}(\text{Type I error}) + \sup_{H_1} \mathbb{P}_{H_1}(\text{Type II error}) \geq \mathbb{P}_{H_0}(\text{Type I error}) + \mathbb{E}_{H_1\sim\pi}\mathbb{P}_{H_1}(\text{Type II error})$$

Therefore,

there exists at least one hypo that ...

$$\liminf_{n\to\infty}\left(\mathbb{P}_{H_0}(\text{Type I error}) + \sup_{H_1} \mathbb{P}_{H_1}(\text{Type II error})\right) \geq 1\,.$$

Recap

We talked about

- Global testing as the simplest task in multiple testing
- We discussed two methods for global testing: Bonferroni's correction and Fisher's combination test
- We argued that Bonferroni's method works better for detecting *a few large changes* in the individual tests, while Fisher's test works better for detecting *many subtle* changes.
- We showed that Bonferroni's method has a sharp detection boundary in the "needle in a haystack" model (with detection level $\mu_n = \sqrt{2\log n}$)
- Any other test is also powerless below this threshold!
(optimality of Bonferrnoi's method against sparse alternatives)

FIGHT ON!