

Biostatistics-Lecture 9

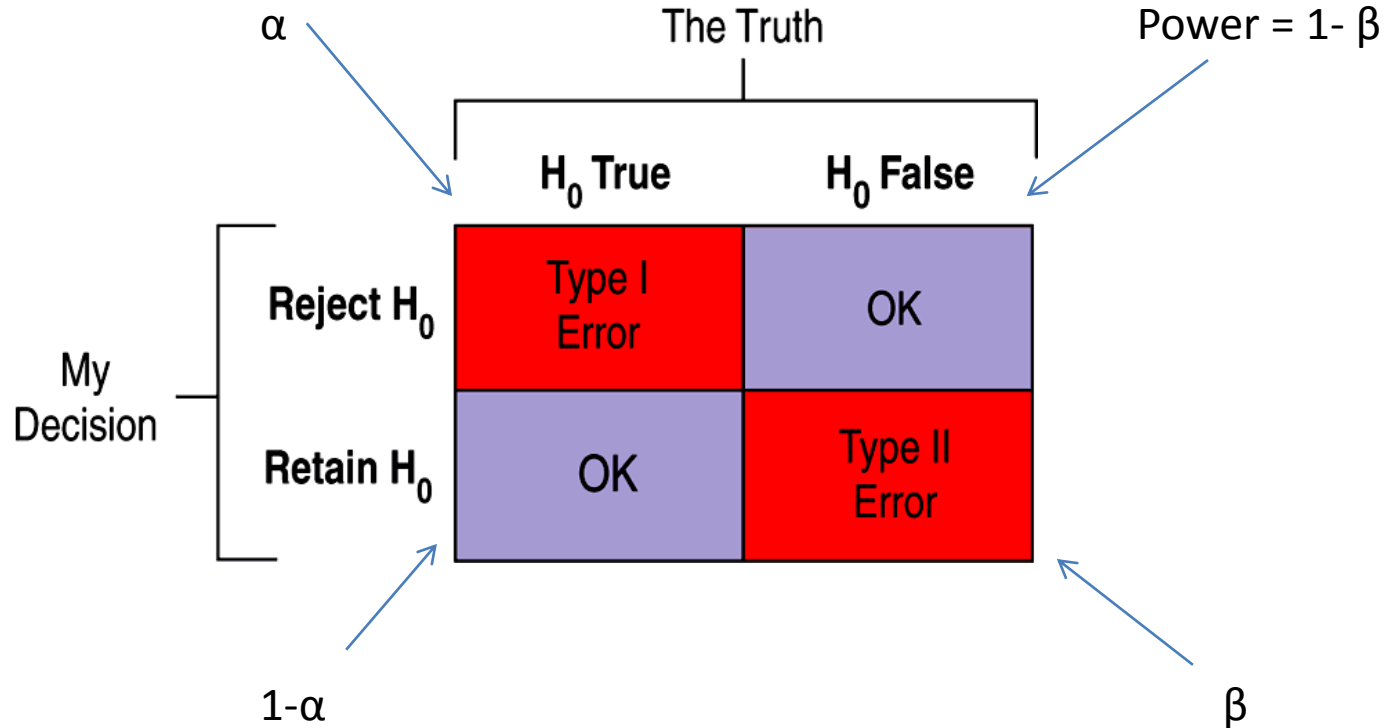
Multiple testing

Ruibin Xi

Peking University

School of Mathematical Sciences

Errors in hypothesis testing



Why multiple testing



Why multiple testing

- In general, perform m hypothesis testing, what is the probability of at least 1 false positive?

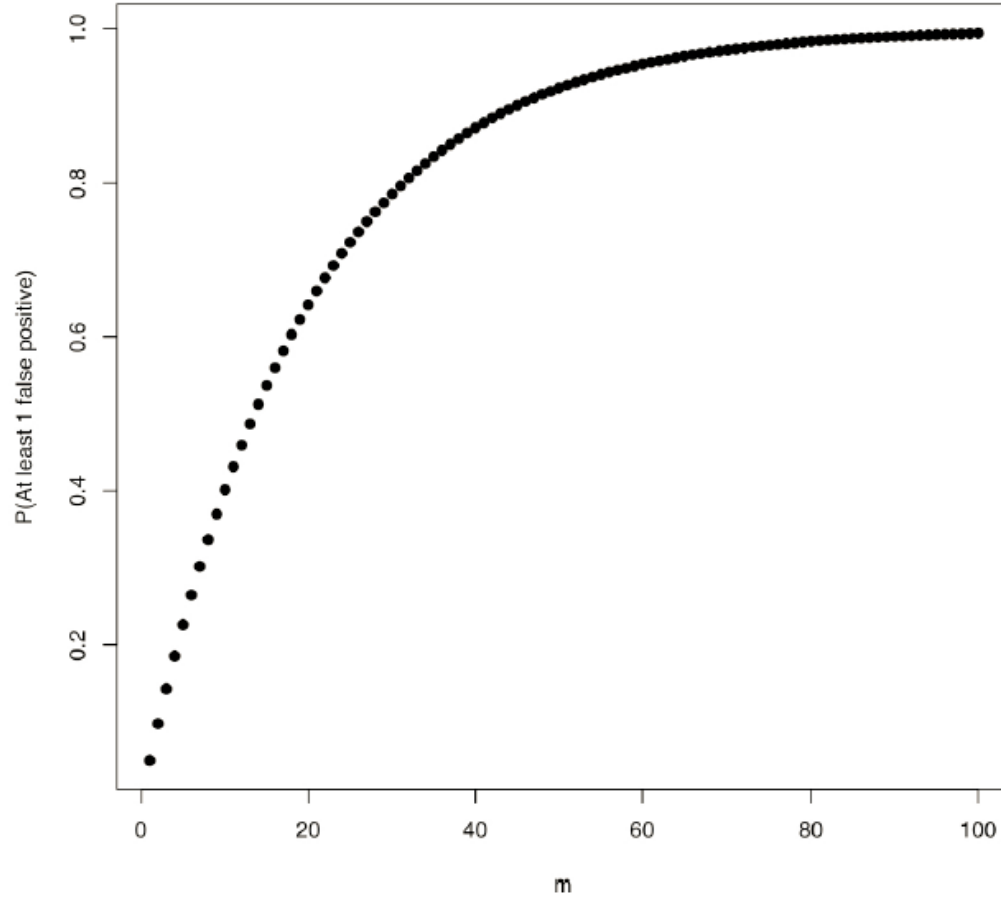
$$P(\text{Making an error}) = \alpha$$

$$P(\text{Not making an error}) = 1 - \alpha$$

$$P(\text{Not making an error in } m \text{ tests}) = (1 - \alpha)^m$$

$$P(\text{Making at least 1 error in } m \text{ tests}) = 1 - (1 - \alpha)^m$$

Probability of at least 1 false positive



Counting Errors

Assume we are testing H^1, H^2, \dots, H^m

$m_0 = \#$ of true hypotheses $R = \#$ of rejected hypotheses

	Null True	Alternative True	Total
Not Called Significant	U	T	$m - R$
Called Significant	V	S	R
	m_0	$m - m_0$	m

$V = \#$ Type I errors [false positives]

Measures to Control Type I Errors

- **Per comparison error rate** (PCER): the expected value of the number of Type I errors over the number of hypotheses,

$$\text{PCER} = E(V)/m$$

- **Per-family error rate** (PFER): the expected number of Type I errors,

$$\text{PFE} = E(V).$$

- **Family-wise error rate**: the probability of at least one type I error

$$\text{FEWR} = P(V \geq 1)$$

- **False discovery rate** (FDR) is the expected proportion of Type I errors among the rejected hypotheses

$$\text{FDR} = E(V/R \mid R > 0)P(R > 0)$$

- **Positive false discovery rate** (pFDR): the rate that discoveries are false

$$\text{pFDR} = E(V/R \mid R > 0)$$

Measures to Control Type I Errors

- **Per comparison error rate** (PCER): the expected value of the number of Type I errors over the number of hypotheses,

$$\text{PCER} = E(V)/m$$

- **Per-family error rate** (PFER): the expected number of Type I errors,

$$\text{PFE} = E(V).$$

- **Family-wise error rate**: the probability of at least one type I error

$$\text{FEWR} = P(V \geq 1)$$

- **False discovery rate** (FDR) is the expected proportion of Type I errors among the rejected hypotheses

$$\text{FDR} = E(V/R \mid R > 0)P(R > 0)$$

- **Positive false discovery rate** (pFDR): the rate that discoveries are false

$$\text{pFDR} = E(V/R \mid R > 0)$$

Methods based on FWER

- Many procedures have been developed to control the Family Wise Error Rate (the probability of at least one type I error):

$$P(V \geq 1)$$

- Two general types of FWER corrections:
 1. **Single step**: equivalent adjustments made to each p-value
 2. **Sequential**: adaptive adjustment made to each p-value

Single Step Approach: Bonferroni

- Very simple method for ensuring that the overall Type I error rate of α is maintained when performing m independent hypothesis tests
- Rejects any hypothesis with p-value $\leq \alpha/m$:

$$\tilde{p}_j = \min[mp_j, 1]$$

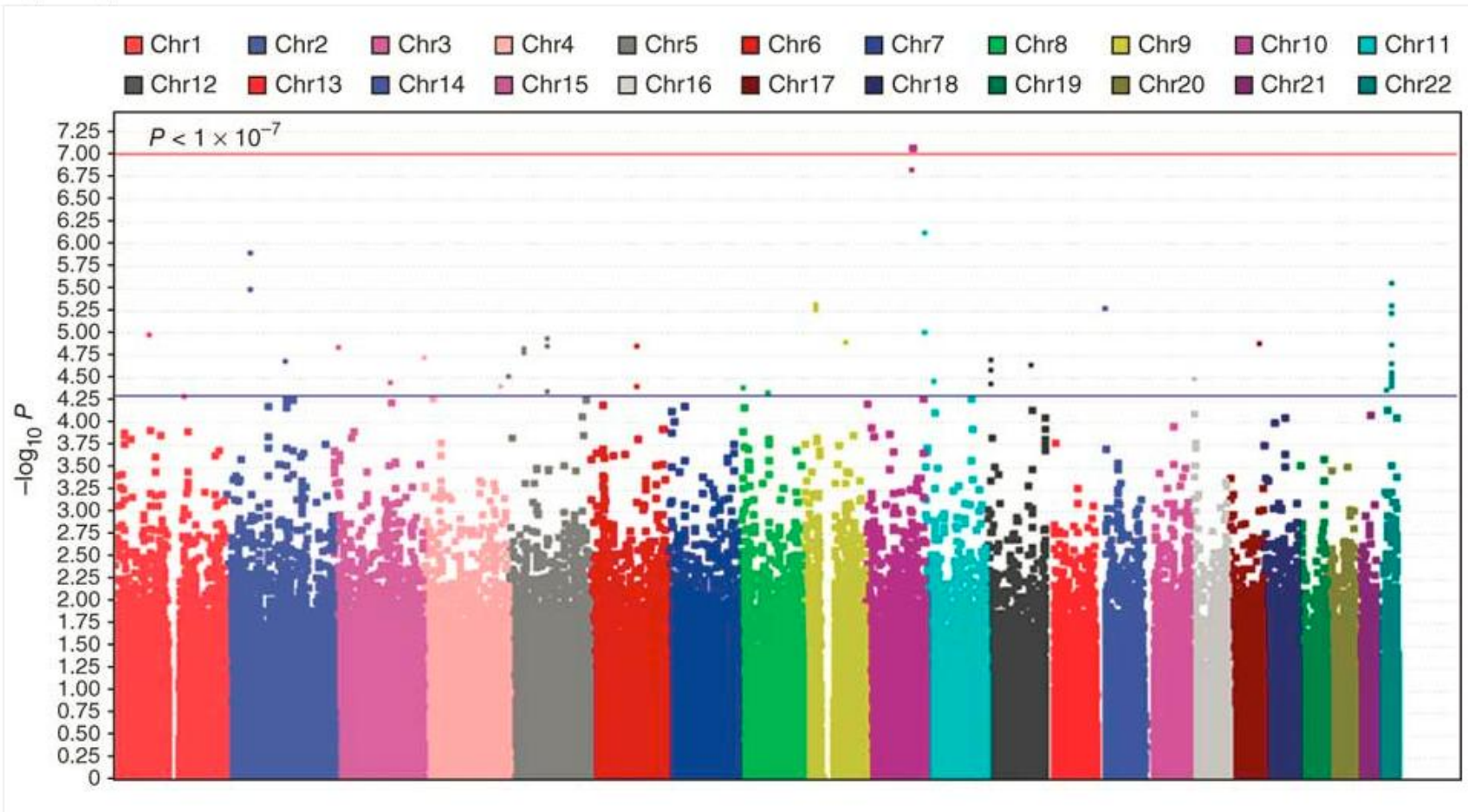
- For example, if we want to have an experiment wide Type I error rate of 0.05 when we perform 10,000 hypothesis tests, we'd need a p-value of $0.05/10000 = 5 \times 10^{-6}$ to declare significance

Rational behind Bonferroni's correction

$$\begin{aligned} P(V \geq 1 | H^1, \dots, H^m) &\leq P\left(\bigcup_{j=1}^m \{p_j \leq \alpha/m\}\right) \\ &\leq \sum_{j=1}^m P(p_j \leq \alpha/m) \\ &= \sum_{j=1}^m \alpha/m \\ &= m\alpha/m = \alpha \end{aligned}$$

Too conservative!

A common application of Benferroni's method



FWER: sequential adjustment

- Holm's method
 - Order the p-values $p_{(1)} \leq \dots \leq p_{(m)}$
 - Reject $H_{(i)}$ if $p_{(j)} \leq \alpha / (m - j + 1), j = 1, \dots, i$
or reject $H_{(i)}$ if $p_{(i)} \leq \alpha / (m - i + 1)$
and all preceding $H_{(j)}$ are rejected
 - Adjusted p-values
$$q_{(i)} = \min\{1, \max[(m - i + 1)p_{(i)}, q_{(i-1)}]\}$$
 - Controls the FWER

FWER: sequential adjustment

- Holm's method can be restated as the following procedure
 - Start testing $H_{(1)}$
Stops the procedure if $p_{(1)} > \alpha/m$
Otherwise reject $H_{(1)}$
 - At the k th step, if $H_{(1)}, \dots, H_{(k-1)}$ are rejected
Stops the procedure if $p_{(k)} > \alpha/(m - k + 1)$
Otherwise reject $H_{(k)}$ and continue the process

FWER: sequential adjustment

- Hochberg's method

- Reject $H_{(i)}$ if there is a $j = i, \dots, m$

$$p_{(j)} \leq \alpha / (m - j + 1)$$

- Adjusted p-values

$$q_{(i)} = \min\{1, \min[(m - i + 1)p_{(i)}, q_{(i+1)}]\}$$

- More powerful than Holm's method

FWER: sequential adjustment

- Hochberg's method can be restated as
 - Start testing $H_{(m)}$
 - If $p_{(m)} \leq \alpha$, reject all $H_{(1)}, \dots, H_{(m)}$ and stops
 - Otherwise, $H_{(m)}$ is retained, the procedure continues with a smaller significance level $\alpha/2$
 - If $p_{(m-1)} \leq \alpha/2$, reject all $\bar{H}_{(1)}, \dots, H_{(m-1)}$
 -

FWER: sequential adjustment

- Hommel's method
 - Step 1. If $p_{(m)} > \alpha$, retain $H_{(m)}$ and go to the next step. Otherwise reject all hypotheses and stop.
 - Steps $i = 2, \dots, m - 1$. If $p_{(m-j+1)} > (i - j + 1)\alpha/i$ for $j = 1, \dots, i$, retain $H_{(m-i+1)}$ and go to the next step. Otherwise reject all remaining hypotheses and stop.
 - Step m . If $p_{(m-j+1)} > (i - j + 1)\alpha/i$ for $j = 1, \dots, m$, retain $H_{(1)}$; otherwise reject it.

FWER is too stringent

- FWER is appropriate when you want to guard against ANY false positives
- However, in many cases (particularly in genomics) we can tolerate a certain number of false positives
- In these cases, the more relevant quantity to control is the false discovery rate (FDR)

False discovery rate

	Null True	Alternative True	Total
Not Called Significant	U	T	$m - R$
Called Significant	V	S	R
	m_0	$m - m_0$	m

V = # Type I errors [false positives]

- False discovery rate (FDR) is designed to control the proportion of false positives **among the set of rejected hypotheses** (R)

FDR and FPR

	Null True	Alternative True	Total
Not Called Significant	<i>U</i>	<i>T</i>	<i>m - R</i>
Called Significant	<i>V</i>	<i>S</i>	<i>R</i>
	<i>m₀</i>	<i>m - m₀</i>	<i>m</i>

$$FDR = \frac{V}{R}$$

$$FPR = \frac{V}{m_0}$$

The Benjamini–Hochberg procedure

- Benjamini and Hochberg (1995) proposed
- Find the largest k such that $p_{(k)} \leq \frac{k}{m}\alpha$
- Reject all hypothesis $H_{(i)}$ $i = 1, \dots, k$
- The adjusted p-values

$$p_{(j)}^* = \min_{k=j, \dots, m} \left\{ \min\left(\frac{m}{k} p_{(k)}, 1\right) \right\}$$

The Benjamini–Yekutieli procedure

- Benjamini and Yekutieli (2001) proposed
 - Find the largest k such that $p_{(k)} \leq \frac{k}{mc(m)}\alpha$
 - Reject all hypothesis $H_{(i)}$ $i = 1, \dots, k$
- For Independent and positive correlated tests

$$c(m) = 1$$

- For general case

$$c(m) = \sum_{i=1}^m \frac{1}{i}$$

- The adjusted p-values

$$p_{(j)}^* = \min_{k=j, \dots, m} \left\{ \min \left(\frac{mc(m)}{k} p_{(k)}, 1 \right) \right\}$$