Множественное тестирование

ФББ

27 марта 2013 г.

Задача: найти гены с разной «активностью» в больных и здоровых людях

Активность генов в здоровых людях

Активность генов в больных людях

Ген	H1	H2	H3	H4	H5	H6	D1	D2	D3	D4	D5	D6	P-val
G1	7.15	6.70	7.28	8.32	7.78	6.86	6.41	7.22	7.83	7.29	5.84	6.19	0.53
G2	…	…	…	…	…	…	…	…	…	…	…	…	…
G3	…	…	…	…	…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…	…	…	…	…	…

Как отобрать гены на основе полученных значений p-value так, чтобы получить **«разумное»** количество перепредсказаний (ошибок I рода) и недопредсказаний (ошибок II рода)?

Задача: найти гены с разной «активностью» в больных и здоровых людях

- Всего генов **10’000**.
- Мы использовали порог на p-value **0.05** и нашли **500** генов со значимо разной активностью.
- Это круто? Можно писать статью в *Nature*?

Допустим, на самом деле среди всех генов не было ни одного с разной активностью. Сколько мы ожидаем перепредсказаний (FP)?

10’000*0.05 = **500**

Ошибки разных родов

	На самом деле H_0	На самом деле H_a
Мы приняли H_0	TN – true negatives (правильно приняли)	FN – false negatives (ошибки II рода)
Мы отвергли H_0	FP – false positives (ошибки I рода)	TP – true positives (правильно отвергли)

Positives – это «открытие», «сигнал», что-то необычное, альтернативная гипотеза

Negatives – фон, нулевая гипотеза

Ученых часто больше волнуют ошибки I рода (перепредсказание).

Типы ошибок (error rates) при множественном тестировании

	На самом деле H_0	На самом деле H_a
Мы приняли H_0	TN – true negatives (правильно приняли)	FN – false negatives (ошибка II рода)
Мы отвергли H_0	FP – false positives (ошибка I рода)	TP – true positives (правильно отвергли)

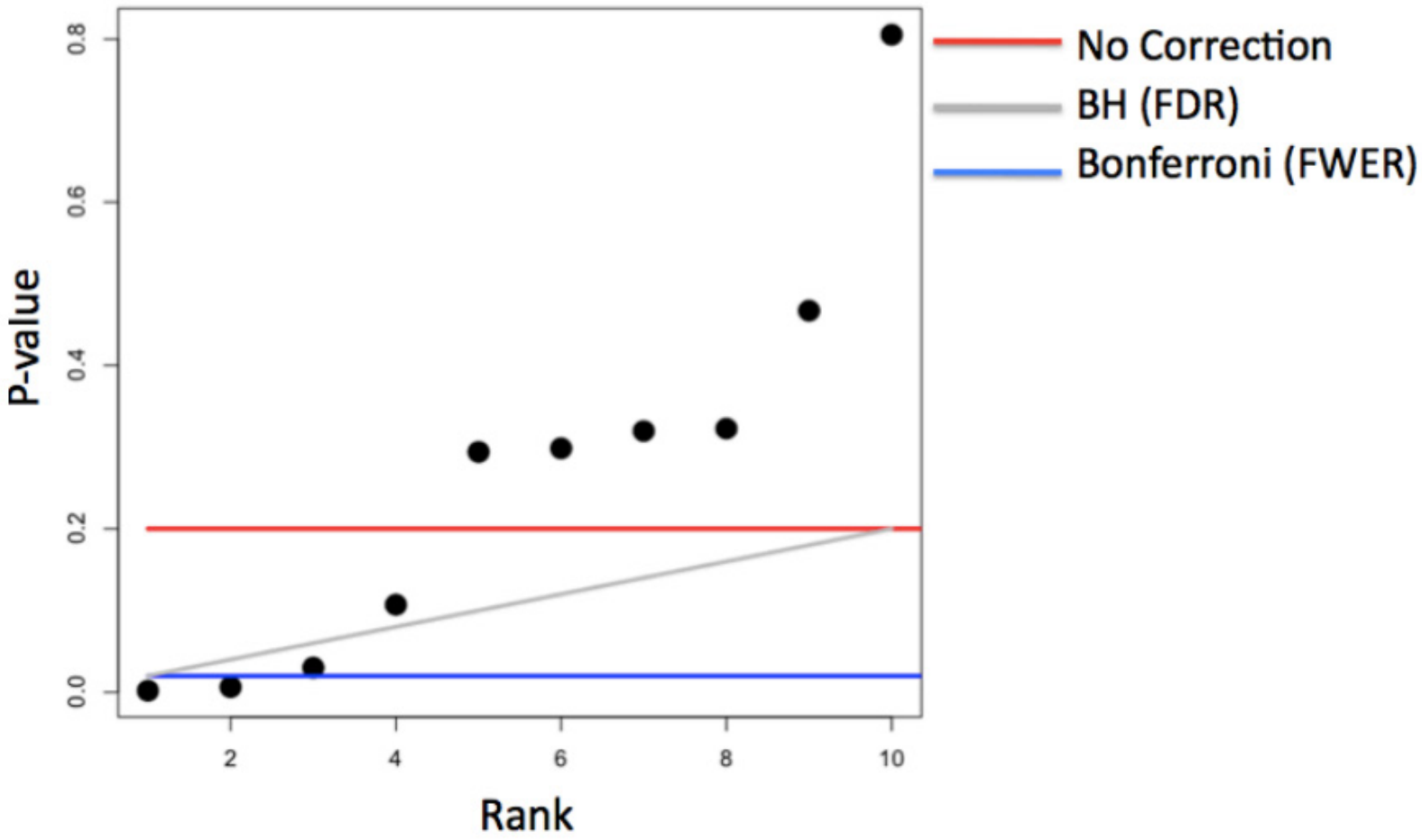
Для достижения желаемого уровня ошибок I рода можно контролировать разные типы ошибок (error rates):

- False Positive Rate : $p = E\left(\frac{FP}{FP+TN}\right)$
- Family wise error rate (FWER): $FWER = P(FP \geq 1)$
- False Discovery Rate (FDR): $FDR = E\left(\frac{FP}{FP+TP}\right)$

Множественное тестирование

Тип контролируемой ошибки (error rate)	**Метод коррекции**	**Коррекция порога на p-value, чтобы error < α (при m тестах)**	**Коррекция p-value (*p*), чтобы error < α (при m тестах)**
False positive rate: $p = E\left(\frac{FP}{FP+TN}\right)$	Без коррекции	нет	нет
Family wise error rate (FWER): *FWER = P(FP ≥ 1)*	Поправка Бонферрони (Bonferroni)	Считаем значимыми $p-value(i) < \frac{\alpha}{m}$	Сравниваем с α $p_{FWER} = \max(m \times p, 1)$ для каждого *p* (p-value)
False Discovery Rate: $FDR = E\left(\frac{FP}{FP+TP}\right)$	Поправка Benjamini-Hochberg	Считаем значимыми $p-value(i) < \alpha \times \frac{i}{m}$ (***i*** – ранк p-value)	Сравниваем с α $p(i)_{BH} < \frac{p(i) \times m}{i}$ (***i*** – ранк p-value)

Пример для 10 p-value



Пример: в выборке нет TP

```
> set.seed(12345)
> pValues <- rep(NA,1000)
> for(i in 1:1000)
+ {
+   x <- rnorm(20)
+   pValues[i] <- t.test(x)$p.value
+ }

# Control false positive rate
> sum(pValues < 0.05) # ≈ 0.05*1000
[1] 51

# Control FWER
> sum(p.adjust(pValues, method="bonferroni") < 0.05)
[1] 0

# Control FDR
> sum(p.adjust(pValues, method="BH") < 0.05)
[1] 0
```

Пример: в выборке 50% TP

```
# Генирируем 500 выборок с mean = 0 и
# 500 выборок с mean = 1.5 -> применяем t-test
# тестирования среднего 0


> set.seed(12345)
> pValues <- rep(NA,1000)
> for(i in 1:500){pValues[i] <- t.test(rnorm(20))$p.value}
> for(i in 501:1000){pValues[i] <- t.test(rnorm(20,
+ mean=1.5))$p.value}
# сохраняем правильные ответы
> trueStatus <- rep(c("zero", "not zero"), each=500)
```

Пример: в выборке 50% TP

```
> trueStatus <- rep(c("zero", "not zero"), each=500)
# Control false positive rate
> table(pValues < 0.05, trueStatus)
       trueStatus
        not zero zero
  FALSE        0  478
  TRUE       500   22
# Control FWER
> table(p.adjust(pValues, method="bonferroni")<0.05,
trueStatus)
       trueStatus
        not zero zero
  FALSE       59  500
  TRUE       441    0
# Control FDR
> table(p.adjust(pValues, method="BH")<0.05,trueStatus)
trueStatus
       not zero zero
  FALSE        0  487
  TRUE       500   13
```