

# Прикладной статистический анализ данных

## Множественная проверка гипотез

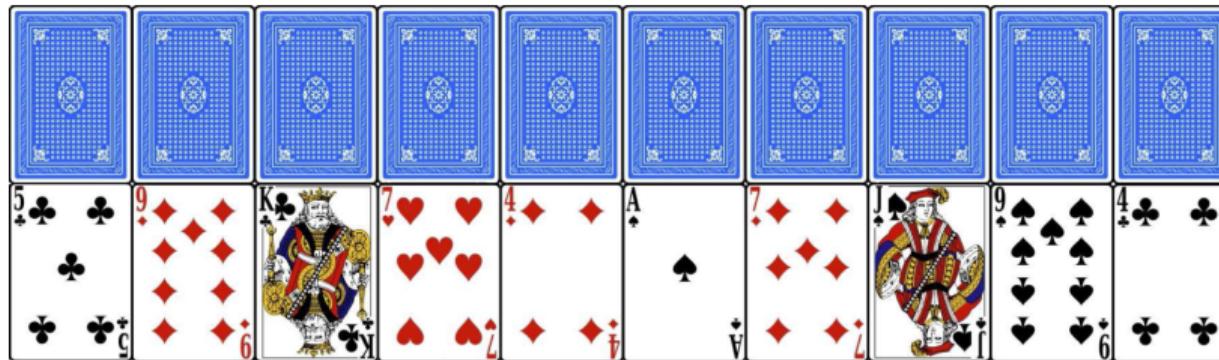
Олег Бахтеев  
psad@phystech.edu

2022

## Поиск экстрасенсов

(Rhine, 1950): исследования возможности экстрасенсорного восприятия. Первый этап — поиск экстрасенсов.

Испытуемому предлагается угадать цвет 10 карт.



$H_0$ : испытуемый выбирает ответ наугад.

$H_1$ : испытуемый может предсказывать цвета карт.

Статистика  $t$  — число карт, цвета которых угаданы.

$$P(t \geq 9 | H_0) = 11 \cdot \frac{1}{2}^{10} = 0.0107421875,$$

т. е. при  $t = 9$  получаем достигаемый уровень значимости  $p \approx 0.01$  — можно отклонять  $H_0$ .

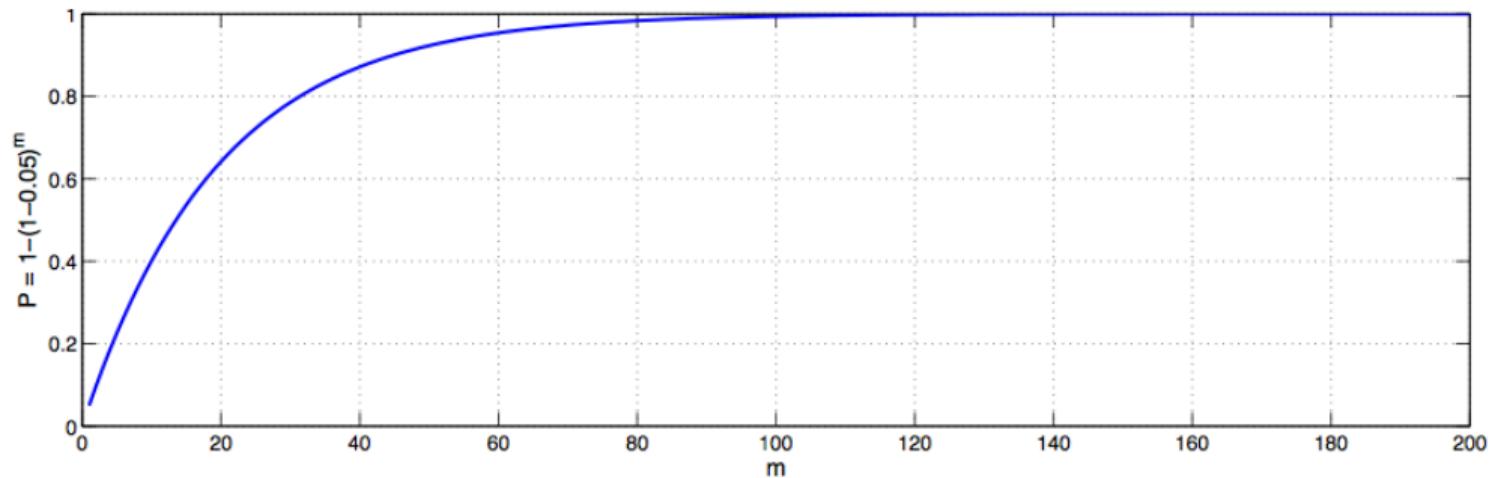
## Поиск экстрасенсов

Процедуру отбора прошли 1000 человек.

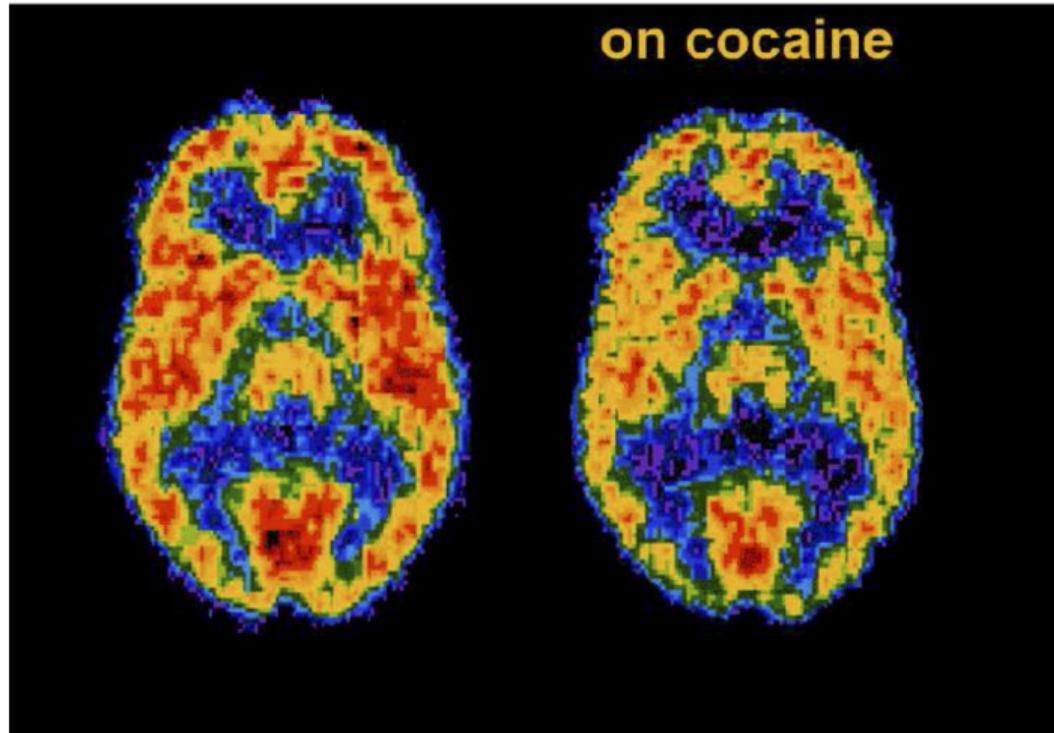
Девять из них угадали цвета 9 из 10 карт, двое — цвета всех 10 карт. Ни один в последующих экспериментах не подтвердил своих способностей.

Вероятность того, что из 1000 человек хотя бы один случайно угадает цвета 9 или 10 из 10 карт:

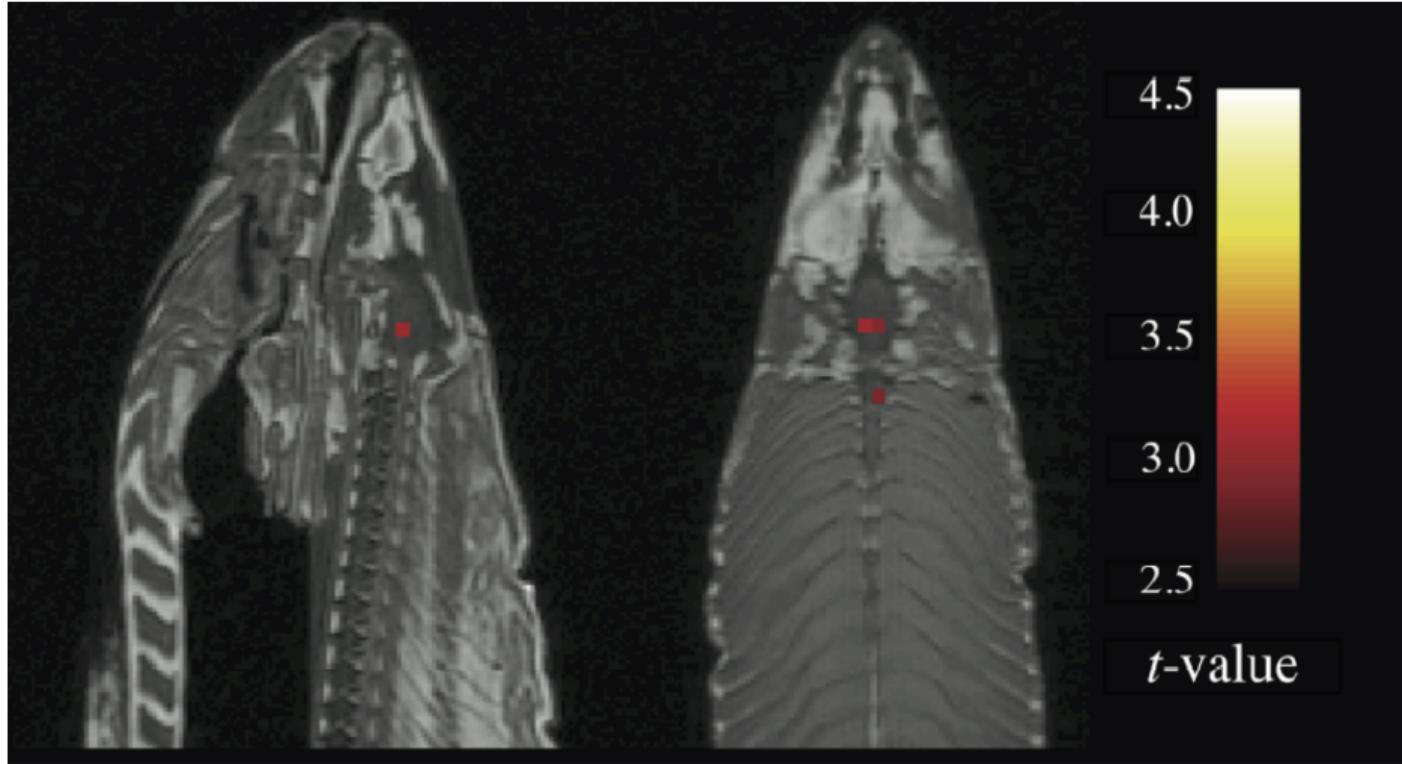
$$1 - \left(1 - 11 \cdot \frac{1}{2}^{10}\right)^{1000} \approx 0.9999796.$$



## Случайные поля



# Лосось (бывший)



## Математическая формулировка

выборка:  $X^n = (X_1, \dots, X_n)$ ,  $X \sim P \in \Omega$

нулевая гипотеза:  $H_0: P \in \omega$ ,  $\omega \in \Omega$

альтернатива:  $H_1: P \notin \omega$

статистика:  $T(X^n)$ ,  $T(X^n) \sim F(x)$  при  $P \in \omega$

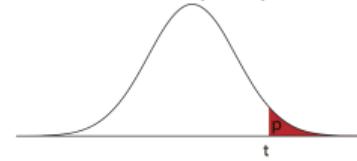
$T(X^n) \not\sim F(x)$  при  $P \notin \omega$



реализация выборки:  $x^n = (x_1, \dots, x_n)$

реализация статистики:  $t = T(x^n)$

достигаемый уровень значимости:  $p(x^n)$  — вероятность при  $H_0$  получить  $T(X^n) = t$  или ещё более экстремальное



$$p(x^n) = P(T \geq t | H_0)$$

Гипотеза отвергается при  $p(x^n) \leq \alpha$ ,  $\alpha$  — уровень значимости

## Правило проверки гипотезы



## Несимметричность задачи проверки гипотезы

	$H_0$ верна	$H_0$ неверна
$H_0$ принимается	$H_0$ верно принята	Ошибка второго рода
$H_0$ отвергается	Ошибка первого рода	$H_0$ верно отвергнута

Вероятность ошибки первого рода жёстко ограничивается малой величиной:

$$p = P(T(X^n) \geq t | H_0) = P(p \leq \alpha | H_0) \leq \alpha.$$

Вероятность ошибки второго рода минимизируется путём выбора достаточно мощного критерия.

## Математическая постановка

- данные:  $\mathbf{X} = \{X_1^{n_1}, \dots, X_m^{n_m}\}, X_i \sim P_i \in \Omega_i$   
нулевые гипотезы:  $H_i: P_i \in \omega_i, \omega_i \in \Omega_i$   
альтернативы:  $H'_i: P_i \notin \omega_i$   
статистики:  $T_i = T(X_i^{n_i})$  проверяет  $H_i$  против  $H'_i$   
реализации статистик:  $t_i = T(x_i^{n_i})$   
достигаемые уровни значимости:  $p_i = p(x_i^{n_i}), i = 1, \dots, m$

$$\mathbf{M} = \{1, 2, \dots, m\}$$

$\mathbf{M}_0 = \mathbf{M}_0(P) = \{i: H_i \text{ верна}\}$  — индексы верных гипотез,  $|\mathbf{M}_0| = m_0$

$\mathbf{R} = \mathbf{R}(P, \alpha) = \{i: H_i \text{ отвергнута}\}$  — индексы отвергаемых гипотез,  $|\mathbf{R}| = R$

$V = |\mathbf{M}_0 \cap \mathbf{R}|$  — число ошибок первого рода

	Число верных $H_i$	Число неверных $H_i$	Всего
Число принятых $H_i$	$U$	$T$	$m - R$
Число отвергнутых $H_i$	$V$	$S$	$R$
Всего	$m_0$	$m - m_0$	$m$

# Многомерные обобщения ошибки первого рода

Групповая вероятность ошибки первого рода (familywise error rate):

$$\text{FWER} = P(V > 0).$$

Контроль над групповой вероятностью ошибки на уровне  $\alpha$  означает

$$\text{FWER} = P(V > 0) \leq \alpha \quad \forall P.$$

Как этого добиться?

Параметры  $\alpha_1, \dots, \alpha_m$  — уровни значимости, на которых необходимо проверять гипотезы  $H_1, \dots, H_m$ ; задача — выбрать их так, чтобы обеспечить  $\text{FWER} \leq \alpha$ .

# Поправка Бонферрони

Метод Бонферрони:

$$\alpha_1 = \dots = \alpha_m = \alpha/m.$$

## Теорема

Если гипотезы  $H_i$ ,  $i = 1, \dots, m$ , отвергаются при  $p_i \leq \alpha/m$ , то FWER  $\leq \alpha$ .

## Доказательство.

$$\begin{aligned} \text{FWER} &= P(V > 0) = P\left(\bigcup_{i=1}^{m_0} \{p_i \leq \alpha/m\}\right) \leq \sum_{i=1}^{m_0} P(p_i \leq \alpha/m) \leq \\ &\leq \sum_{i=1}^{m_0} \alpha/m = \frac{m_0}{m} \alpha \leq \alpha. \end{aligned}$$

□

Альтернативный вид — переход к модифицированным достижимым уровням значимости:

$$\tilde{p}_i = \min(1, mp_i).$$

## Поправка Бонферрони

При увеличении  $m$  в результате применения поправки Бонферрони мощность статистической процедуры резко уменьшается — шансы отклонить неверные гипотезы падают.

Пример: критерий Стьюдента для независимых выборок,

$$X_1^n, X_1 \sim N(\mu_1, 1), \quad X_2^n, X_2 \sim N(\mu_2, 1), \quad \mu_1 - \mu_2 = 1,$$

$$H_0: \mathbb{E}X_1 = \mathbb{E}X_2, \quad H_1: \mathbb{E}X_1 \neq \mathbb{E}X_2.$$

$m$	$n$	Мощность
1	23	0.9
10	23	0.67
100	23	0.37
1000	23	0.16
1000	62	0.9

Если проверяется одновременно 1000000 гипотез, при размере выборок  $n = 10$  мощность 0.9 достигается при расстоянии между средними выборок в пять стандартных отклонений.

## Модельный эксперимент

$$n = 20, \ m = 200, \ m_0 = 150;$$

$$X_i^n, X_i \sim N(0, 1), \ i = 1, \dots, m_0;$$

$$X_i^n, X_i \sim N(1, 1), \ i = m_0 + 1, \dots, m;$$

$$H_i: \mathbb{E}X_i = 0, \ H'_i: \mathbb{E}X_i \neq 0.$$

Для проверки используем одновыборочный критерий Стьюдента.

Без поправок:

	Верных $H_i$	Неверных $H_i$	Всего
Принятых $H_i$	142	0	142
Отвергнутых $H_i$	8	50	58
Всего	150	50	200

Бонферрони:

	Верных $H_i$	Неверных $H_i$	Всего
Принятых $H_i$	150	27	177
Отвергнутых $H_i$	0	23	23
Всего	150	50	200

## Нисходящие методы множественной проверки гипотез

Составим вариационный ряд достижимых уровней значимости:

$$p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(m)},$$

$H_{(1)}, H_{(2)}, \dots, H_{(m)}$  — соответствующие гипотезы.

- ① Если  $p_{(1)} \geq \alpha_1$ , принять все нулевые гипотезы  $H_{(1)}, H_{(2)}, \dots, H_{(m)}$  и остановиться; иначе отвергнуть  $H_{(1)}$  и продолжить.
- ② Если  $p_{(2)} \geq \alpha_2$ , принять все нулевые гипотезы  $H_{(2)}, H_{(3)}, \dots, H_{(m)}$  и остановиться; иначе отвергнуть  $H_{(2)}$  и продолжить.
- ③ ...

Каждый достижимый уровень значимости  $p_{(i)}$  сравнивается со своим уровнем значимости  $\alpha_i$ .

## Метод Холма

Метод Холма — нисходящая процедура со следующими уровнями значимости:

$$\alpha_1 = \frac{\alpha}{m}, \alpha_2 = \frac{\alpha}{m-1}, \dots, \alpha_i = \frac{\alpha}{m-i+1}, \dots, \alpha_m = \alpha.$$

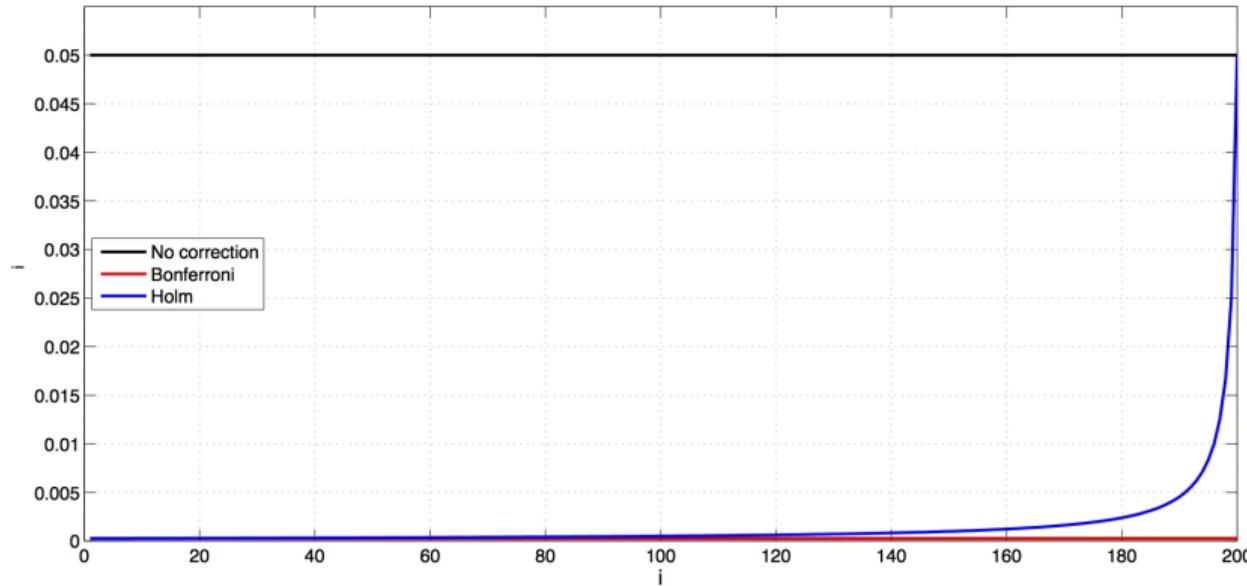
Метод обеспечивает контроль над FWER на уровне  $\alpha$  при любых  $p_i$  и  $T_i$ .

Модифицированные достижимые уровни значимости:

$$\tilde{p}_{(i)} = \min \left( 1, \max \left( (m - i + 1) p_{(i)}, \tilde{p}_{(i-1)} \right) \right).$$

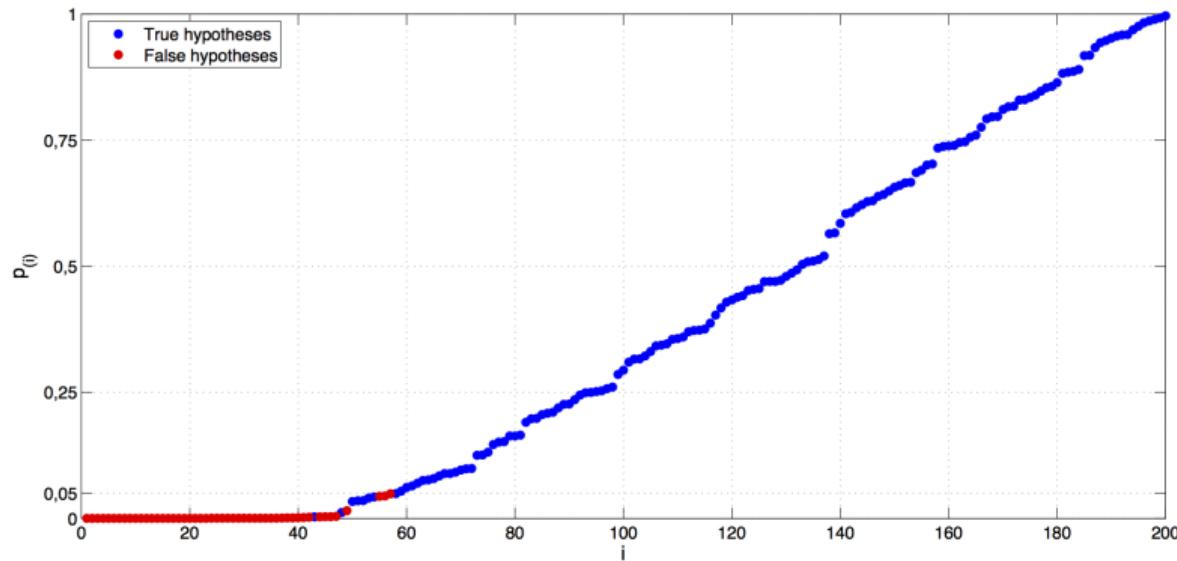
## Метод Холма

Метод Холма равномерно мощнее поправки Бонферрони, поскольку все его уровни значимости  $\alpha_i$  не меньше:



# Модельный эксперимент

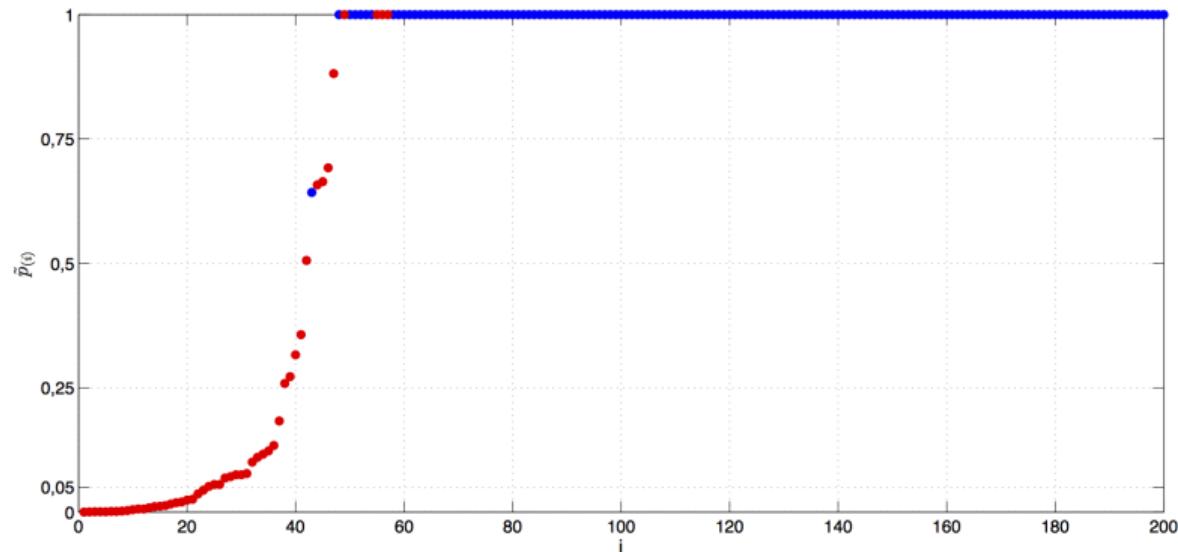
Отсортированные достижаемые уровни значимости:



	Верных $H_i$	Неверных $H_i$	Всего
Принятых $H_i$	142	0	142
Отвергнутых $H_i$	8	50	58
Всего	150	50	200

## Модельный эксперимент

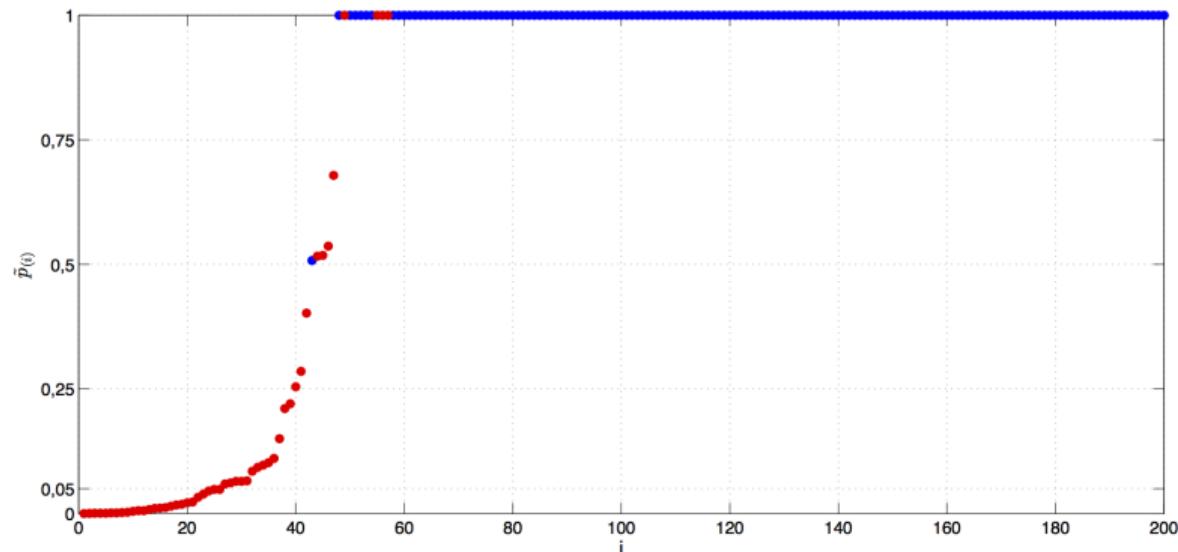
Модифицированные достигаемые уровни значимости, метод Бонферрони:



	Верных $H_i$	Неверных $H_i$	Всего
Принятых $H_i$	150	27	177
Отвергнутых $H_i$	0	23	23
Всего	150	50	200

# Модельный эксперимент

Модифицированные достижимые уровни значимости, метод Холма:



	Верных $H_i$	Неверных $H_i$	Всего
Принятых $H_i$	150	24	174
Отвергнутых $H_i$	0	26	26
Всего	150	50	200

## Идеи для дальнейших улучшений

- Дополнительно оценить  $m_0$ .
- Сделать дополнительные предположения:
  - ▶ о характере зависимости между статистиками;
  - ▶ о совместном распределении статистик.
- Учесть зависимость между статистиками с помощью перестановочных методов.

## Предварительное оценивание $m_0$

Многие методы контролируют FWER на уровне  $\alpha_T = \frac{m_0}{m} \alpha \Rightarrow$  можно оценить  $m_0$  и выбрать  $\alpha$  так, чтобы  $\alpha_T$  было равно желаемой величине.

Метод Шведера-Спьётволла:

$$\hat{m}_0(\lambda) = \frac{1}{1-\lambda} \left( 1 + \sum_{i=1}^m [p_i > \lambda] \right), \quad \lambda \in [0, 1).$$

Имеет положительное смещение, а также большую дисперсию, особенно при сильно коррелированных  $p$ , поэтому FWER не контролируется.

# Одношаговый метод Шидака

Метод Шидака:

$$\alpha_1 = \alpha_2 = \dots = \alpha_m = 1 - (1 - \alpha)^{\frac{1}{m}}.$$

Идея:

$$\text{FWER} \leq \alpha \rightarrow P(V > 0) \leq \alpha \rightarrow 1 - P(V = 0) = 1 - (1 - \alpha)^m \rightarrow \alpha = 1 - (1 - \alpha)^m.$$

Метод обеспечивает контроль над FWER на уровне  $\alpha$  при условии, что статистики  $T_i$  **независимы** или выполняется следующее свойство:

$$P(T_1 \leq t_1, \dots, T_m \leq t_m) \geq \prod_{i=1}^m P(T_i \leq t_i) \quad \forall t_1, \dots, t_m$$

(**positive lower orthant dependence**).

Модифицированные достижимые уровни значимости:

$$\tilde{p}_i = 1 - (1 - p_i)^m.$$

## Нисходящая модификация

Нисходящий метод Шидака (метод Шидака-Холма) — нисходящая процедура со следующими уровнями значимости:

$$\alpha_1 = 1 - (1 - \alpha)^{\frac{1}{m}}, \dots, \alpha_i = 1 - (1 - \alpha)^{\frac{1}{m-i+1}}, \dots, \alpha_m = \alpha.$$

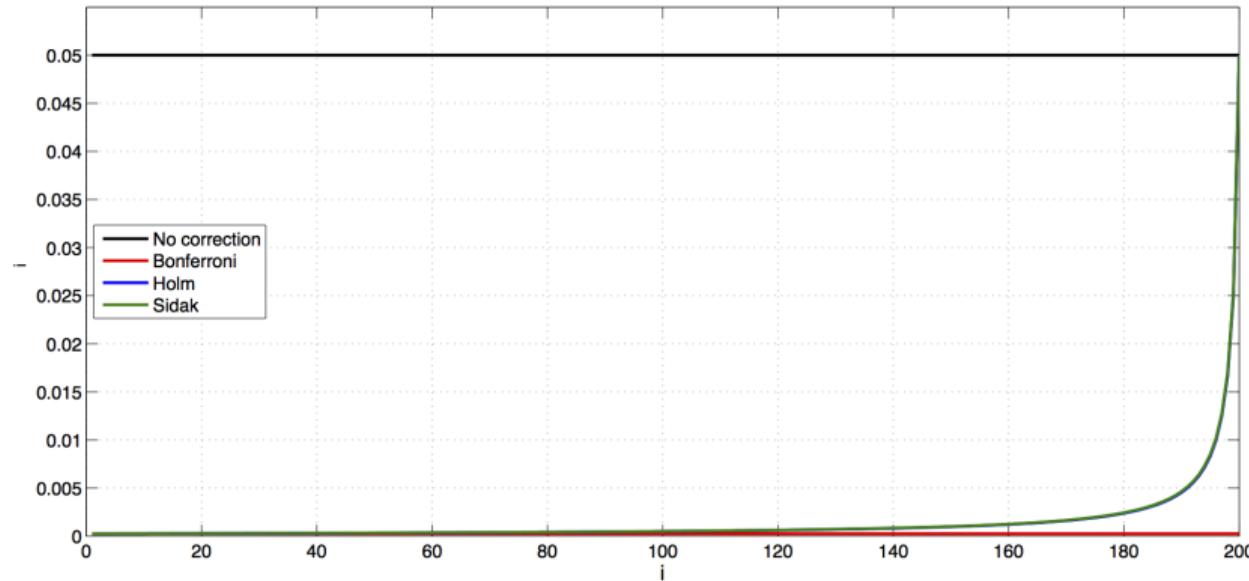
Метод обеспечивает контроль над FWER на уровне  $\alpha$  при условии, что статистики  $T_i$  **независимы**.

Модифицированные достижимые уровни значимости:

$$\tilde{p}_{(i)} = \max \left( 1 - (1 - p_{(i)})^{(m-i+1)}, \tilde{p}_{(i-1)} \right).$$

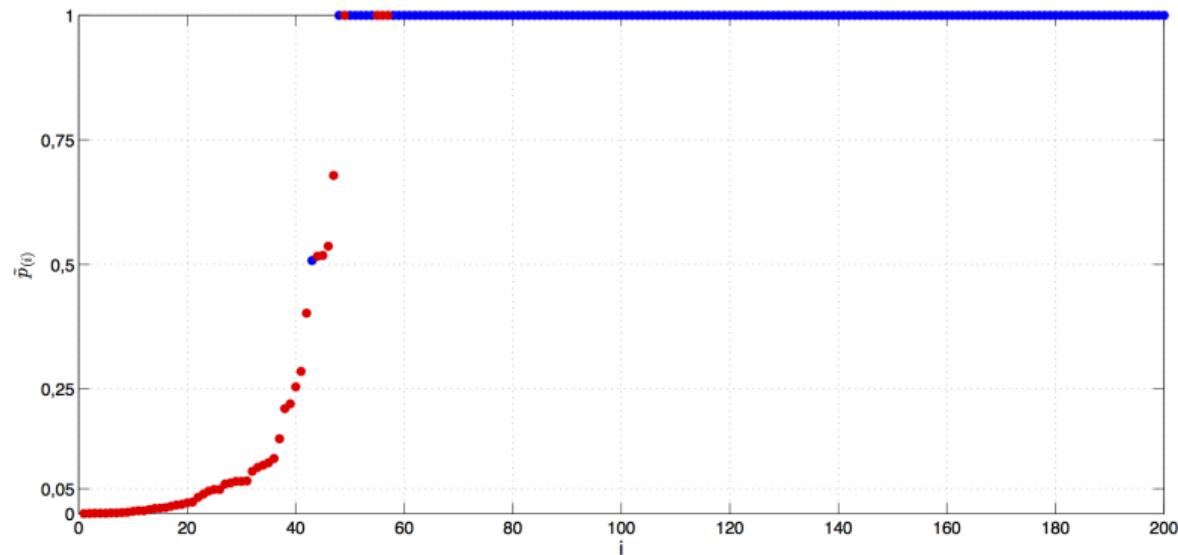
## Нынешняя модификация

На практике при достаточно больших  $m$  не слишком отличается от метода Холма:



# Модельный эксперимент

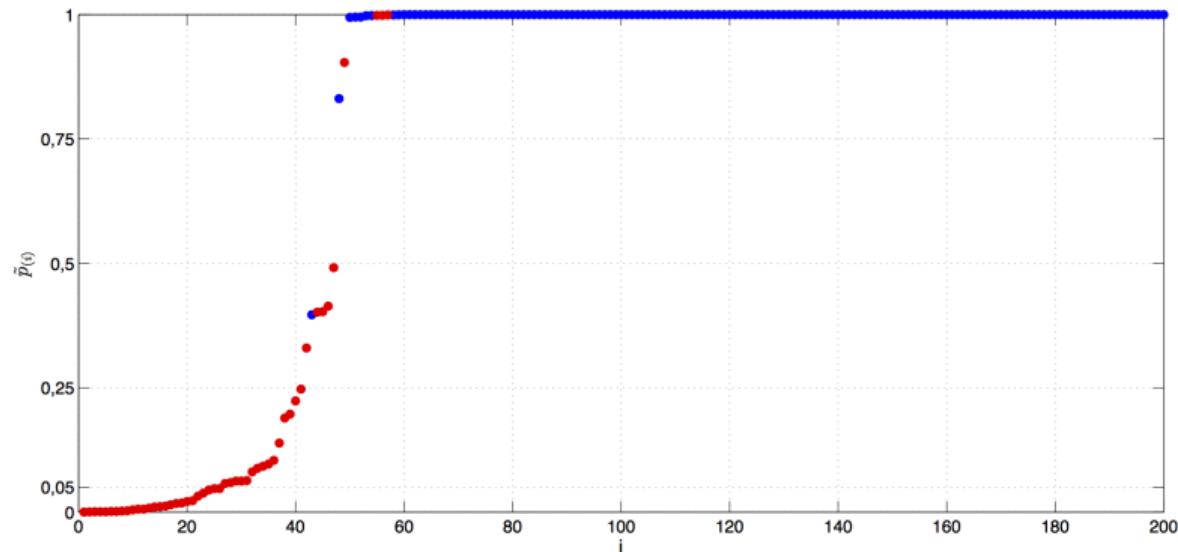
Модифицированные достижимые уровни значимости, метод Холма:



	Верных $H_i$	Неверных $H_i$	Всего
Принятых $H_i$	150	24	174
Отвергнутых $H_i$	0	26	26
Всего	150	50	200

## Модельный эксперимент

Модифицированные достигаемые уровни значимости, исходящий метод Шидака:



	Верных $H_i$	Неверных $H_i$	Всего
Принятых $H_i$	150	24	174
Отвергнутых $H_i$	0	26	26
Всего	150	50	200

## Зависимость между статистиками

- Не учитывая характер зависимости между статистиками, нельзя построить контролирующую FWER процедуру мощнее, чем метод Холма.
- Если статистики независимы, нельзя построить контролирующую FWER процедуру мощнее, чем метод Шидака-Холма.
- Чем сильнее связь между статистиками, тем меньше нужно модифицировать уровни значимости.

Для построения мощной процедуры множественной проверки гипотез необходимо учесть структуру зависимости статистик.

## Параметрические методы

Если совместное нулевое распределение статистик  $T_1, \dots, T_m$  известно, константы  $\alpha_i$  могут быть так, что контроль над FWER будет точным ( $\text{FWER} = \alpha$ ).

**Примеры:**

- метод HSD Тьюки для попарных сравнений нормально распределённых выборок друг с другом;
- критерий Даннета для сравнения средних  $m$  нормально распределённых выборок со средним контрольной выборки.

## Перестановочные методы, пример

Group	$E_1$	$E_2$	$E_3$
Trt	1	1	1
Trt	1	0	0
Trt	1	0	0
Trt	1	0	0
Ctrl	0	0	0
Ctrl	0	0	0
Ctrl	0	0	0

Источник: Bretz et al.

$p$ -value для разных признаков:

0.02, 0.57, 0.57.

Итог: берем  $\min(0.02, 0.57, 0.57) = 0.02$

## Перестановочные методы

Неявно учесть зависимости между статистиками можно при помощи перестановочных методов.  
Подробнее: Bretz, раздел 5.1.

Методы обеспечивают контроль над FWER на уровне  $\alpha$  при условии выполнения свойства **subset pivotality**:

$$P\left(\left.\bigcap_{i \in M^*} \{T_i \geq t^*\}\right| \bigcap_{i \in M^*} H_i\right) = P\left(\left.\bigcap_{i \in M^*} \{T_i \geq t^*\}\right| \bigcap_{i \in M} H_i\right) \quad \forall M^* \in M$$

(нулевое распределение любого подмножества статистик  $T_i$  не зависит от того, верны или неверны соответствующие оставшимся статистикам гипотезы).

Примеры задач:

- проверка гипотез о средних коррелированных нормальных выборок;
- проверка гипотез о линейных комбинациях средних нормальных выборок;
- попарные сравнения средних в нормальных выборках.

## Многомерные обобщения ошибки первого рода

Ожидаемая доля ложных отклонений гипотез (false discovery rate):

$$\text{FDR} = \mathbb{E} \left( \frac{V}{\max(R, 1)} \right).$$

Контроль над ожидаемой долей ложных отклонений на уровне  $\alpha$  означает

$$\text{FDR} = \mathbb{E} \left( \frac{V}{\max(R, 1)} \right) \leq \alpha \quad \forall P.$$

Для любой процедуры множественной проверки гипотез  $\text{FDR} \leq \text{FWER}$ .

## Восходящие методы множественной проверки гипотез

Составим вариационный ряд достижимых уровней значимости:

$$p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(m)},$$

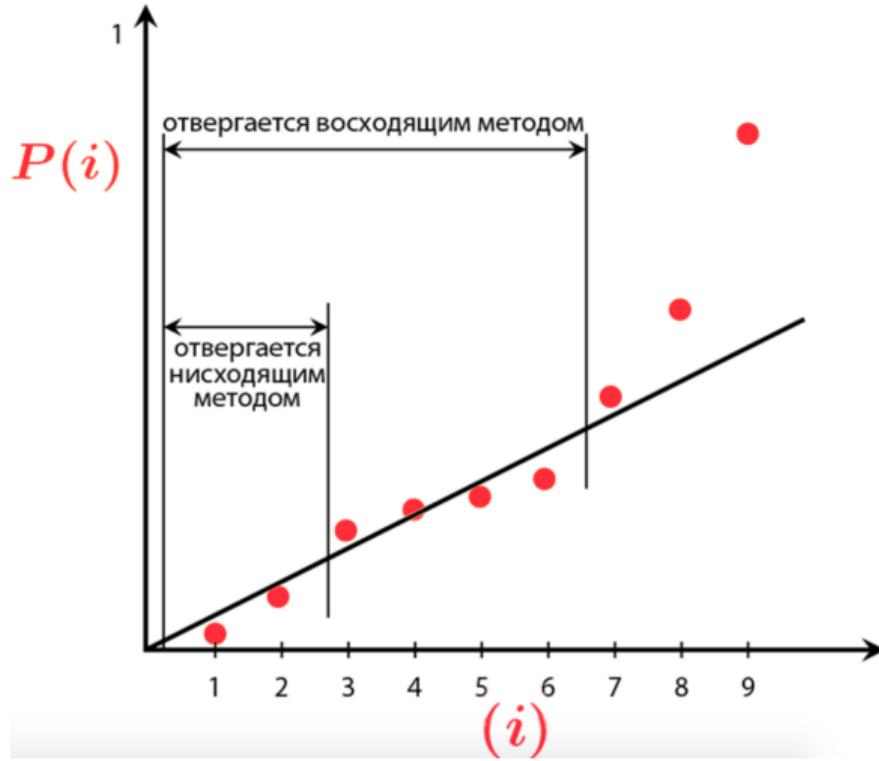
$H_{(1)}, H_{(2)}, \dots, H_{(m)}$  — соответствующие гипотезы.

- ① Если  $p_{(m)} \leq \alpha_m$ , отвергнуть все нулевые гипотезы  $H_{(1)}, H_{(2)}, \dots, H_{(m)}$  и остановиться; иначе принять  $H_{(m)}$  и продолжить.
- ② Если  $p_{(m-1)} \leq \alpha_{m-1}$ , отвергнуть все нулевые гипотезы  $H_{(1)}, H_{(2)}, \dots, H_{(m-1)}$  и остановиться; иначе принять  $H_{(m-1)}$  и продолжить.
- ③ ...

Каждый достижимый уровень значимости  $p_{(i)}$  сравнивается со своим уровнем значимости  $\alpha_i$ .

## Восходящие методы множественной проверки гипотез

Восходящая процедура всегда отвергает не меньше гипотез, чем нисходящая с теми же уровнями значимости:



## Метод Бенджамини-Хохберга

Метод Бенджамини-Хохберга — восходящая процедура со следующими уровнями значимости:

$$\alpha_1 = \frac{\alpha}{m}, \dots, \alpha_i = \frac{\alpha i}{m}, \dots, \alpha_m = \alpha.$$

Метод обеспечивает контроль над FDR на уровне  $\alpha$  при условии, что статистики  $T_i$  **независимы** или выполняется следующее свойство:

$$P(X \in D | T_i = x) \text{ неубывает по } x \quad \forall i \in M_0,$$

где  $D$  — произвольное возрастающее множество, то есть, такое, что из  $x \in D$  и  $y \geq x$  следует  $y \in D$ .  
**(PRDS on  $T_i, i \in M_0$ )** (positive regression dependency on each one from a subset)).

Модифицированные достигаемые уровни значимости:

$$\tilde{p}_{(i)} = \min \left( 1, \frac{mp_{(i)}}{i}, \tilde{p}_{(i+1)} \right).$$

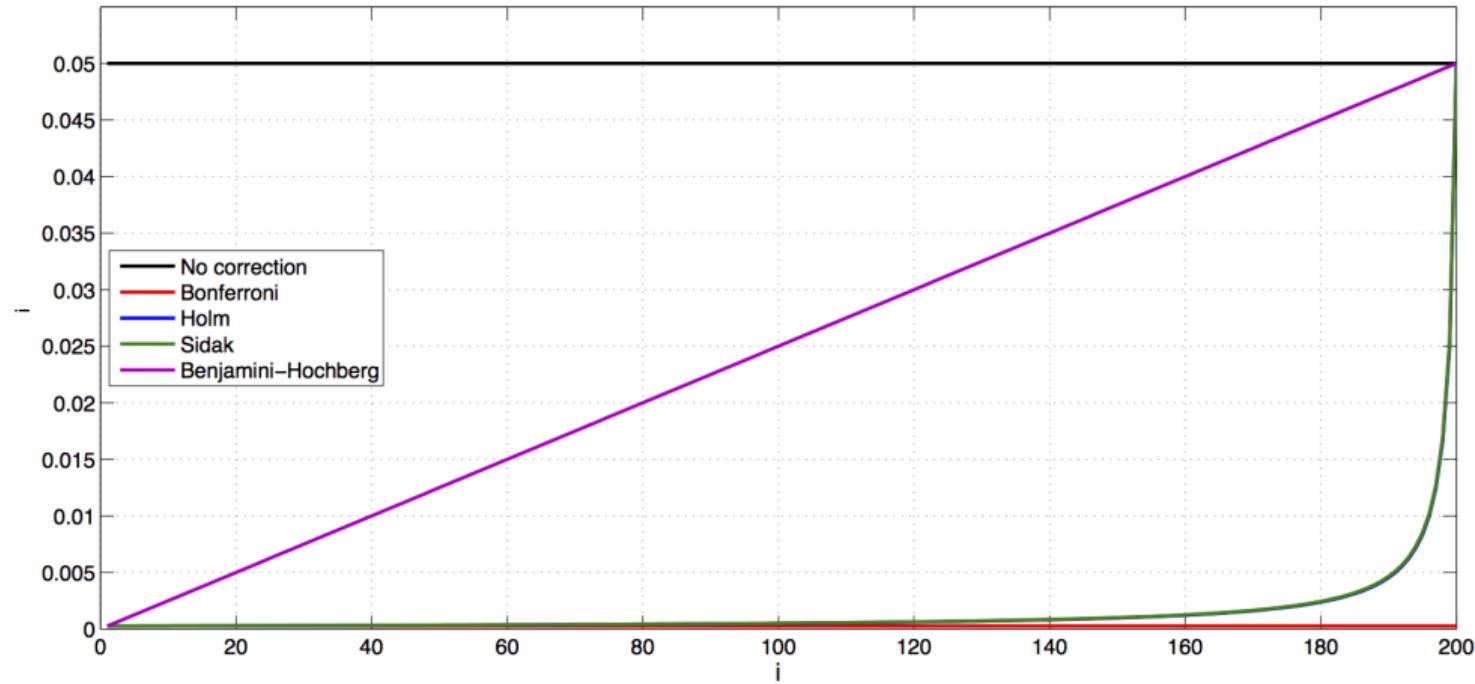
## Метод Бенджамини-Хохберга

PRDS выполняется, например, для многомерного нормального распределения с нулевыми средними и неотрицательными корреляциями элементов из  $M_0$ , а также для некоторых его производных.

Примеры задач:

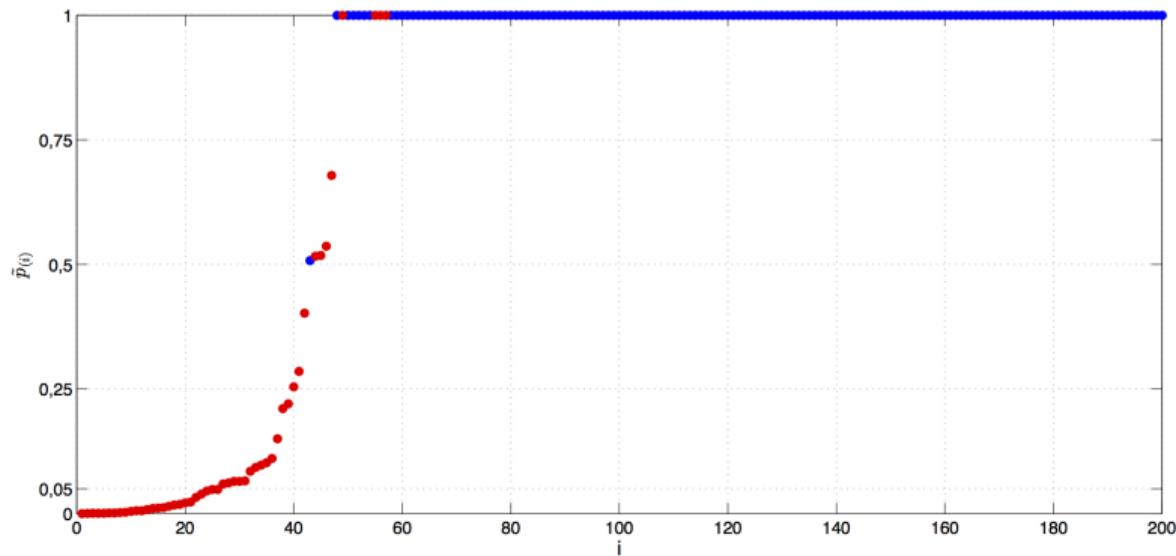
- анализ непересекающихся подгрупп при сравнении двух выборок критерием Стьюдента с общей оценкой дисперсии;
- сравнение одной нормальной выборки с многими при использовании общей оценки дисперсии;

## Метод Бенджамини–Хохберга



## Модельный эксперимент

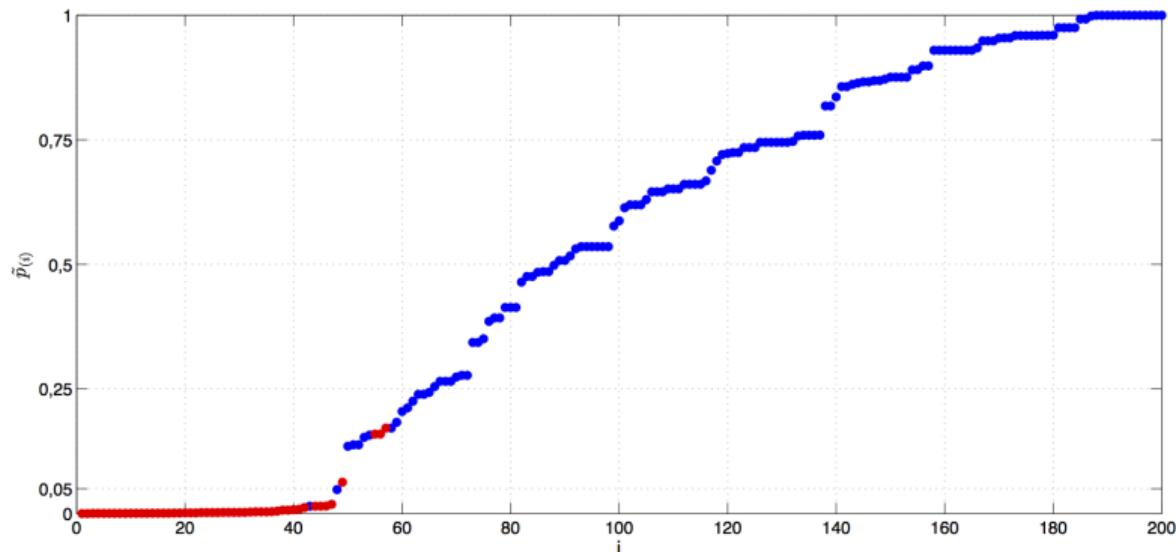
Модифицированные достижимые уровни значимости, метод Холма:



	Верных $H_i$	Неверных $H_i$	Всего
Принятых $H_i$	150	24	174
Отвергнутых $H_i$	0	26	26
Всего	150	50	200

## Модельный эксперимент

Модифицированные достигаемые уровни значимости, метод Бенджамина-Хохберга:



	Верных $H_i$	Неверных $H_i$	Всего
Принятых $H_i$	148	4	152
Отвергнутых $H_i$	2	46	48
Всего	150	50	200

## Метод Бенджамини-Иекутиели

Метод Бенджамини-Иекутиели — восходящая процедура со следующими уровнями значимости:

$$\alpha_1 = \frac{\alpha}{m \sum_{j=1}^m \frac{1}{j}}, \dots, \alpha_i = \frac{\alpha i}{m \sum_{j=1}^m \frac{1}{j}}, \dots, \alpha_m = \frac{\alpha}{\sum_{j=1}^m \frac{1}{j}}.$$

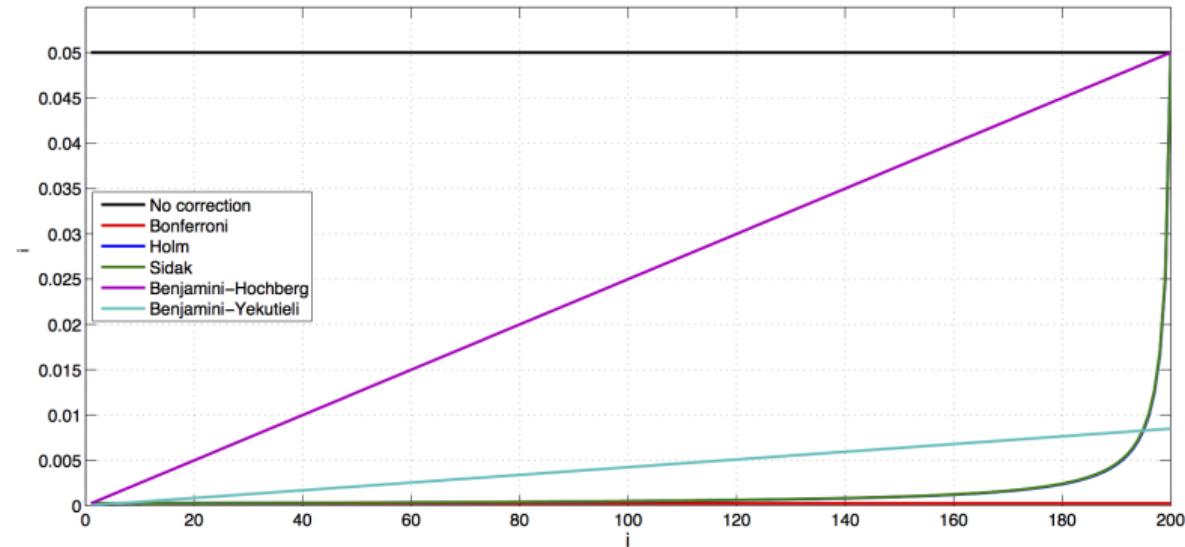
Метод обеспечивает контроль над FDR на уровне  $\frac{m_0}{m}\alpha \leq \alpha$  при любых  $p_i$  и  $T_i$ . При отсутствии информации о зависимости между статистиками метод неулучшаем.

Модифицированные достигаемые уровни значимости:

$$\tilde{\alpha}_{(i)} = \min \left( 1, \frac{mp_{(i)} \sum_{j=1}^m \frac{1}{j}}{i}, \tilde{\alpha}_{(i+1)} \right).$$

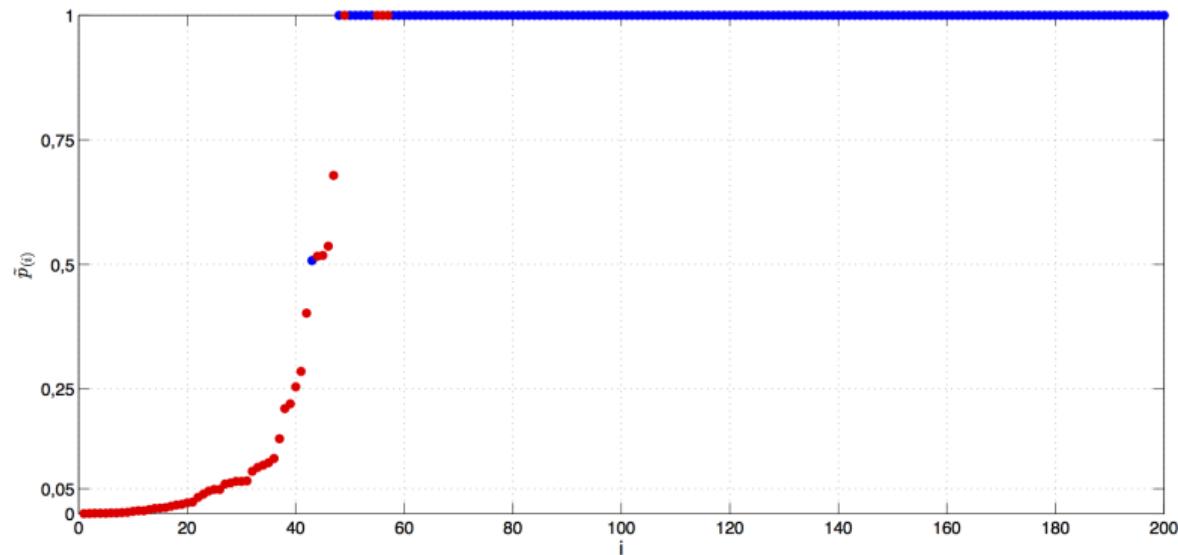
Если доля неверных гипотез мала, метод Бенджамини-Иекутиели отвергает меньше гипотез, чем метод Холма.

## Метод Бенджамини-Иекутиели



## Модельный эксперимент

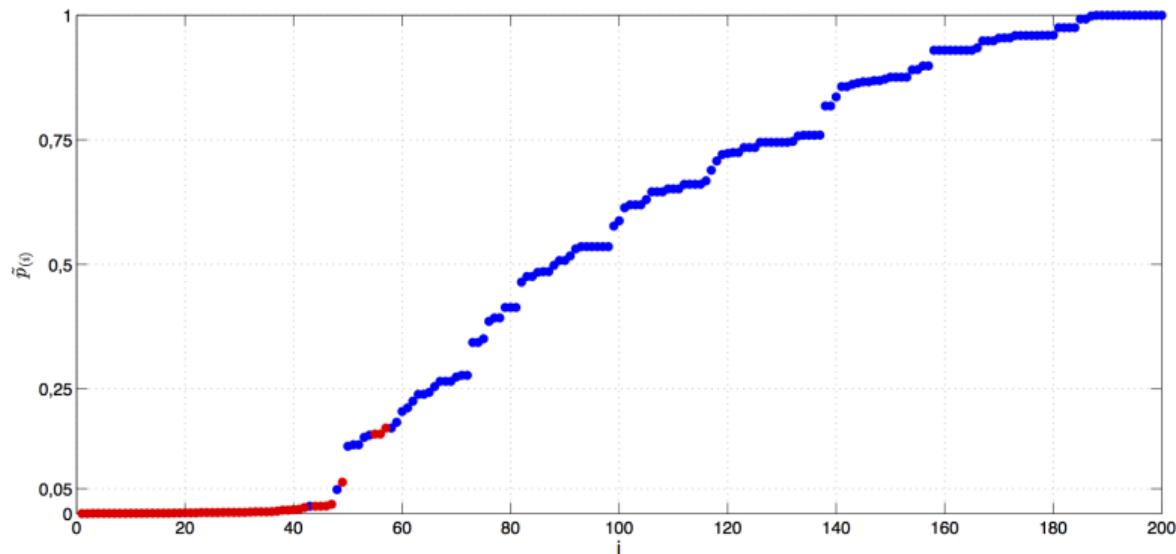
Модифицированные достижимые уровни значимости, метод Холма:



	Верных $H_i$	Неверных $H_i$	Всего
Принятых $H_i$	150	24	174
Отвергнутых $H_i$	0	26	26
Всего	150	50	200

## Модельный эксперимент

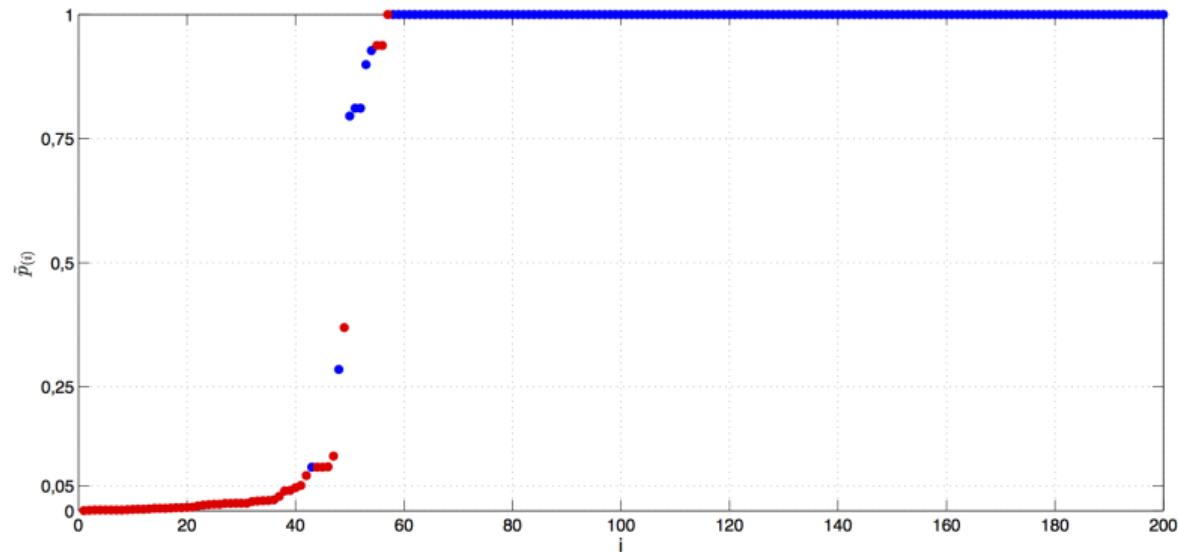
Модифицированные достигаемые уровни значимости, исходящий метод Бенджамина-Хохберга:



	Верных $H_i$	Неверных $H_i$	Всего
Принятых $H_i$	148	4	152
Отвергнутых $H_i$	2	46	48
Всего	150	50	200

## Модельный эксперимент

Модифицированные достигаемые уровни значимости, исходящий метод Бенджамина-Иекутиели:



	Верных $H_i$	Неверных $H_i$	Всего
Принятых $H_i$	150	10	160
Отвергнутых $H_i$	0	40	40
Всего	150	50	200

# Мутации

	Контроль (100)	Больные (100)	$p$
Мутация	1 из 100	8 из 100	0.0349
Фамилия начинается с гласной	36 из 100	40 из 100	0.6622

Бонферрони, Холм:  $p_1$  сравнивается с  $\frac{0.05}{2} = 0.025$

Шидак:  $p_1$  сравнивается с  $1 - (1 - 0.05)^{\frac{1}{2}} \approx 0.02532$

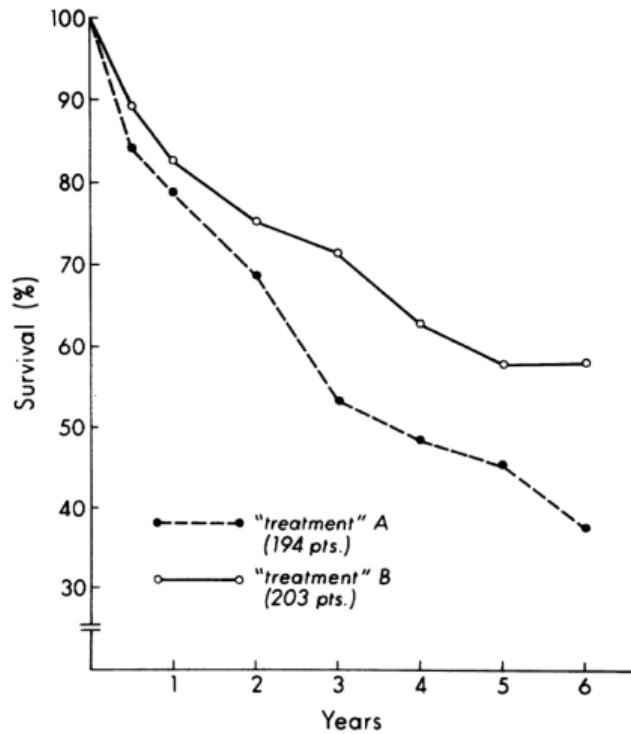
## Подгруппы

(Lee et al., 1980): 1073 пациента с ишемической болезнью сердца были искусственно разделены на две случайные группы, лечение в двух группах проходило одинаково. Исследовалась выживаемость пациентов.

Важными факторами, влияющими на выживаемость, являются число поражённых артерий (одна, две, три) и тип сокращений левого желудочка (нормальный, аномальный).

Для одной из шести подгрупп по этим уровням фактора были обнаружены значимые различия в выживаемости пациентов первого и второго типов.

## Подгруппы



## Подгруппы

(Ishitani, Lin, 2008): анализировалась связь потребления кофеина, кофе и чая и риска возникновения рака груди, всего около 50 подгрупп. Показано, что:

- употребление четырёх и более чашек кофе в день связано с увеличением риска злокачественного рака груди ( $p = 0.08$ );
- потребление кофеина связано с увеличением риска возникновения эстроген- и прогестерон-независимых опухолей и опухолей больше 2 см ( $p = 0.02$  и  $p = 0.02$ );
- потребление кофе без кофеина связано со снижением риска рака груди у женщин в постменопаузе, принимающих гормоны ( $p = 0.02$ ).

См. также:

- (Гален, II в. н.э.): "Все больные, принявшие это средство, вскоре выздоровели, за исключением тех, кому оно не помогло — они умерли. Отсюда очевидно, что это средство помогает во всех случаях, кроме безнадежных."
- <http://youtu.be/QysrgLXMPwA>
- <http://xkcd.com/882/>
- <http://wmbriggs.com/blog/?p=9308>

# Cherry-picking

Метод cherry-picking:

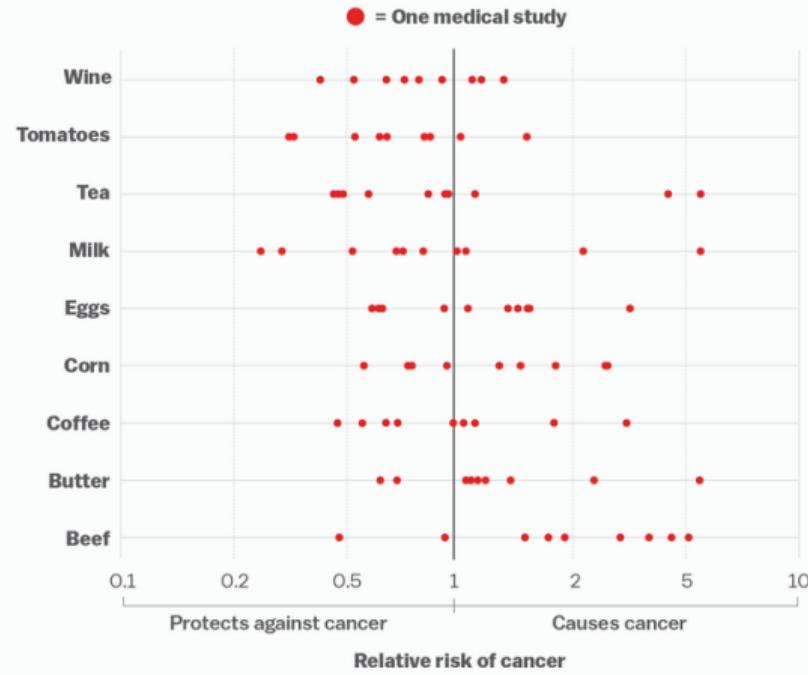
- выбираем произвольное множество  $\mathbf{R}$  гипотез, которые мы хотим отвергнуть;
- оцениваем сверху долю ложных отклонений:

$$FDP = \frac{V}{R}.$$

Оценка справедлива сразу для всех  $2^m - 1$  возможных множеств  $\mathbf{R}$ .

# Большая часть исследований неверна

Everything we eat both causes and prevents cancer



SOURCE: Schoenfeld and Ioannidis, *American Journal of Clinical Nutrition*

Vox

Ioannidis, Why most published research findings are false, 2005.

## Литература

- попроще — Bretz, посложнее — Dickhaus, хороший краткий обзор — Goeman, 2014;
- перестановочные методы (permutation methods) — Westfall, 2008, и другие работы этого автора;
- cherry-picking — Goeman, 2011;
- случайные поля — Nichols, 2003; [filion.ucl.ac.uk/spm/](http://filion.ucl.ac.uk/spm/), [coursera.org/course/fmri](https://coursera.org/course/fmri).

Bretz F., Hothorn T., Westfall P. *Multiple Comparisons Using R*, 2010.

Dickhaus T. *Simultaneous Statistical Inference With Applications in the Life Sciences*, 2014.

Goeman J.J., Solari A. (2011). *Multiple testing for exploratory research*. Statistical Science, 26(4), 584–597.

Goeman J.J., Solari A. (2014). *Multiple hypothesis testing in genomics*. Statistics in Medicine, 33(11), 1946–1978.

Ioannidis J.P.A. (2005). *Why most published research findings are false*. PLoS Medicine, 2(8), e124.

## Литература

- Ishitani K., Lin J. (2008). *Caffeine consumption and the risk of breast cancer in a large prospective cohort of women.* Archives of Internal Medicine, 168(18), 2022–2031.
- Lee K.L., McNeer J.F., Starmer C.F., Harris P.J., Rosati R.A. (1980). *Clinical judgment and statistics. Lessons from a simulated randomized trial in coronary artery disease.* Circulation, 61(3), 508–515.
- Nichols T.E., Hayasaka, S. (2003). *Controlling the familywise error rate in functional neuroimaging: a comparative review.* Statistical Methods in Medical Research, 12(5), 419–446.
- Westfall P., Troendle J. (2008). *Multiple testing with minimal assumptions.* Biometrical Journal, 50(5), 745–755.

# Проверочная работа

**Формат:** несколько типовых задач из занятий 1–4.

**Длительность:** 30 минут.

**Когда:** 1 марта после семинара.