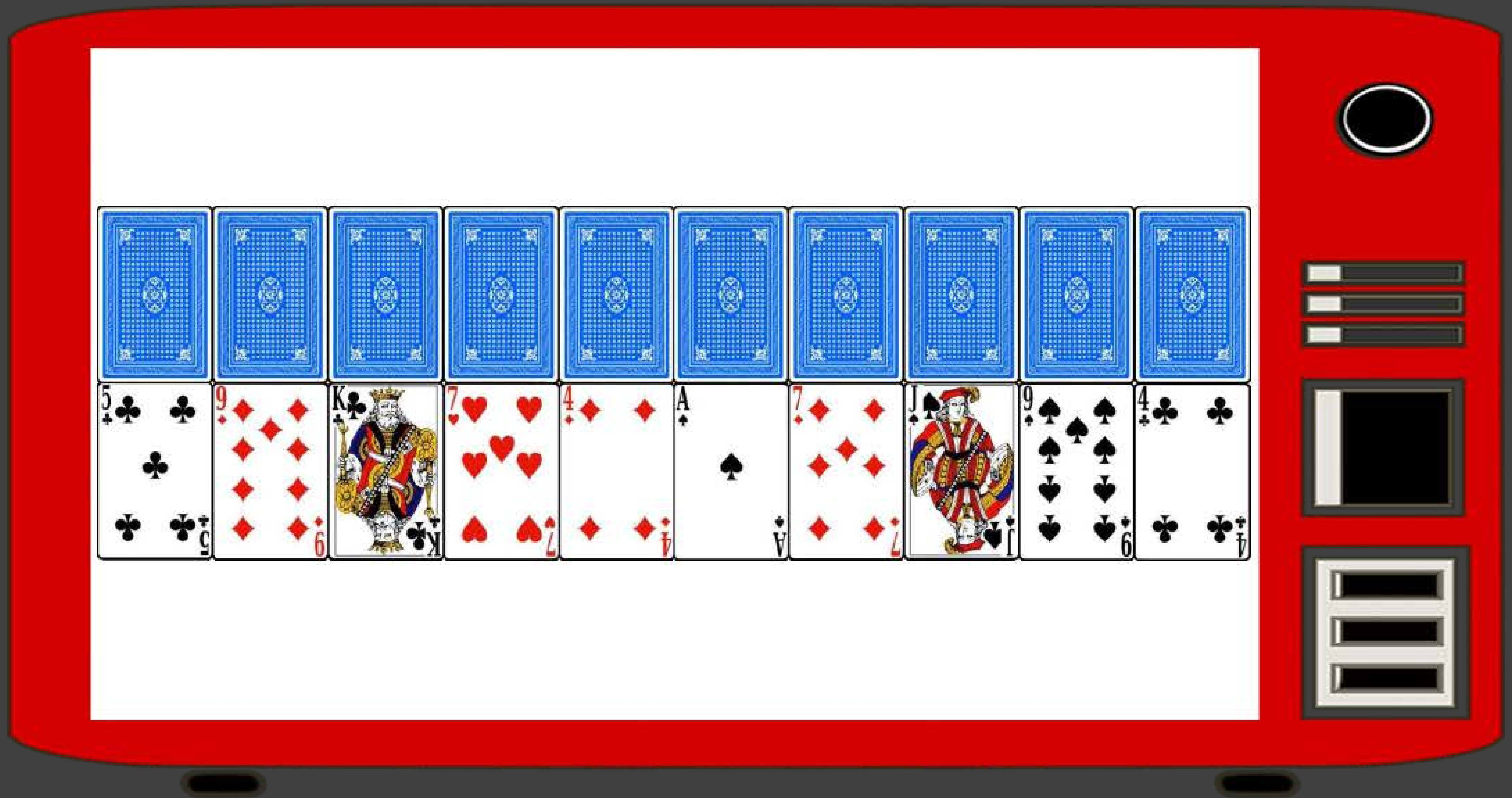
В ЧЁМ ПРОБЛЕМА?

ПОИСК ЭКСТРАСЕНСОВ

- Джозеф Райн, 1950: исследования возможности экстрасенсорного восприятия. Первый этап — поиск эстрасенсов

- Испытуемому предлагается угадать цвет 10 карт



Поиск экстрасенсов

- H_0: испытуемый выбирает ответ наугад
- H_1: испытуемый может предсказывать цвета карт
- Статистика t — число карт, цвета которых угаданы

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

ПОИСК ЭКСТРАСЕНСОВ

- 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

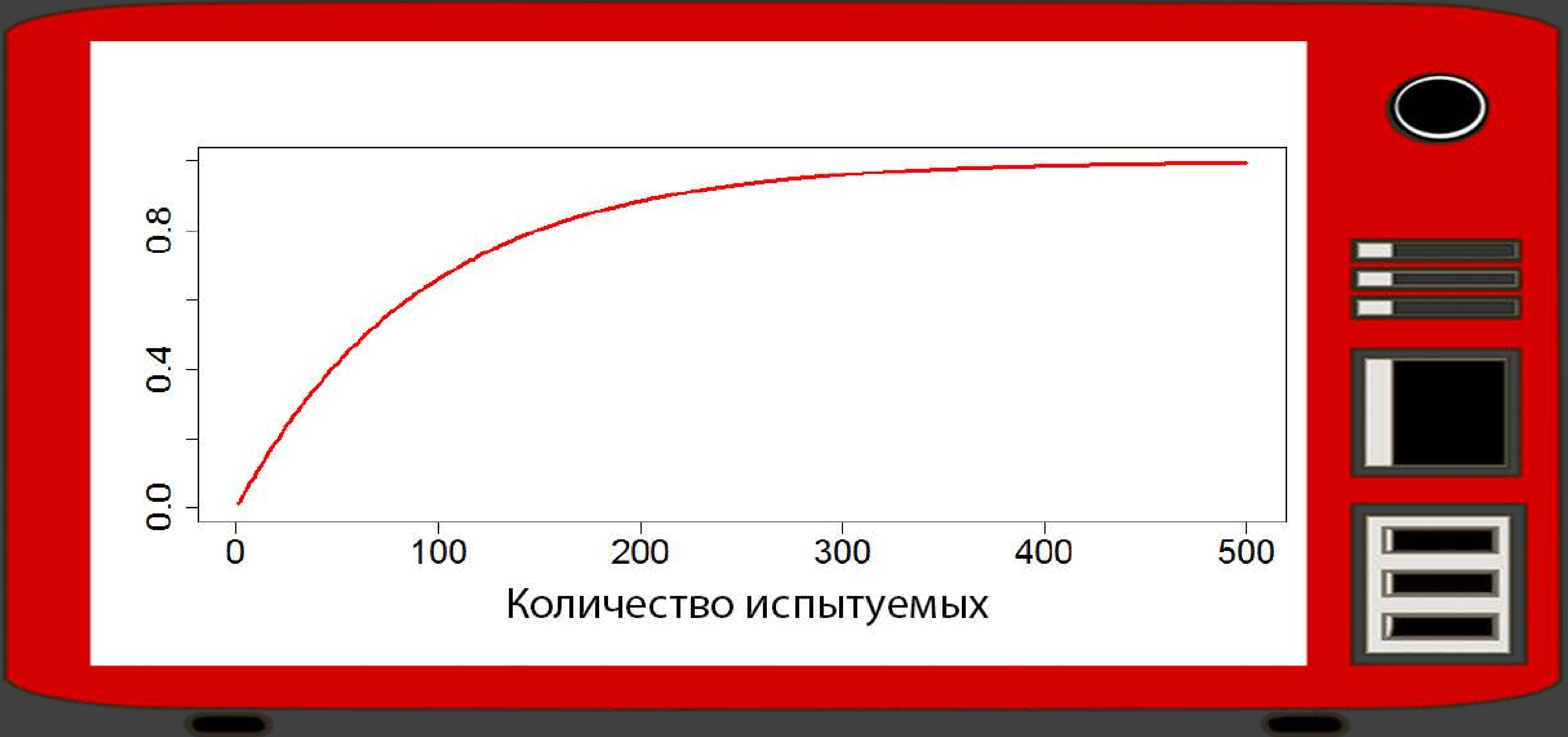
ПОИСК ЭКСТРАСЕНСОВ

- 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

Поиск экстрасенсов

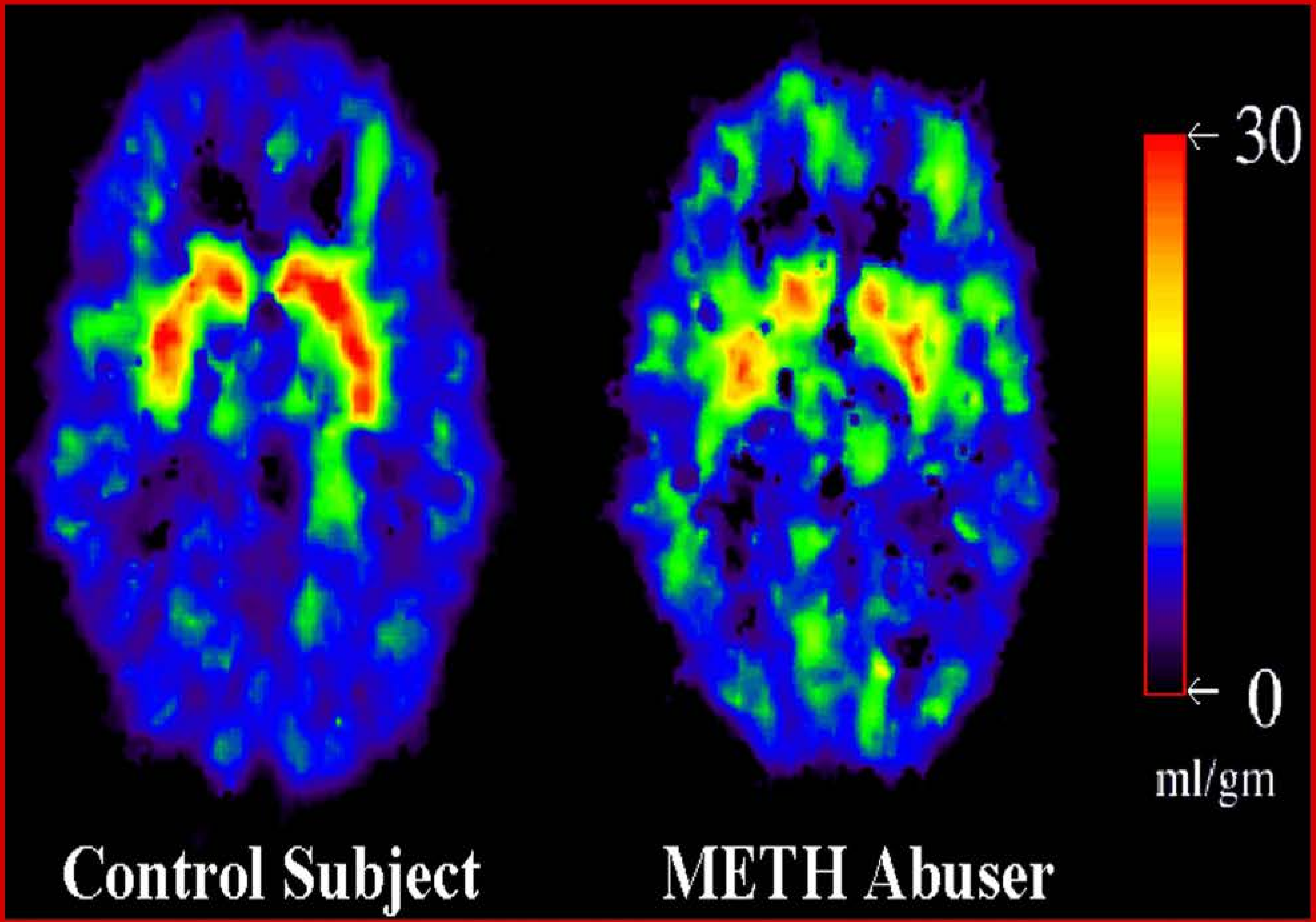
- Вероятность того, что из 1000 человек хотя бы один случайно угадает цвета 9 или 10 из 10 карт:

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

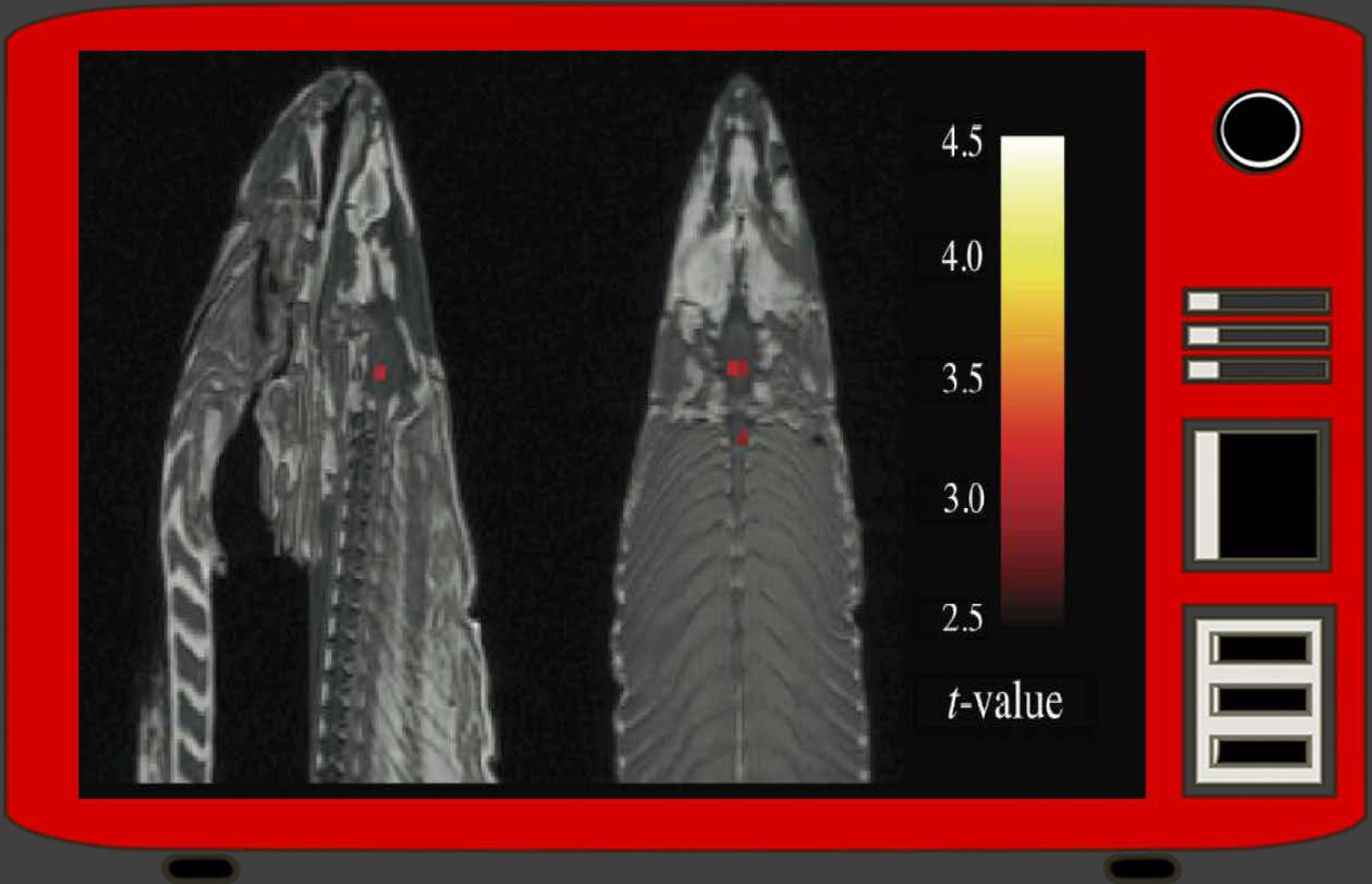


0.0
0.4
0.8
0
100
200
300
400
500
Количество испытуемых

НЕЙРОНАУКА


30
0
ml/gm
Control Subject
METH Abuser

МЁРТВЫЙ ЛОСОСЬ


4.5
4.0
3.5
3.0
2.5
t-value

ДАЛЕЕ В ПРОГРАММЕ

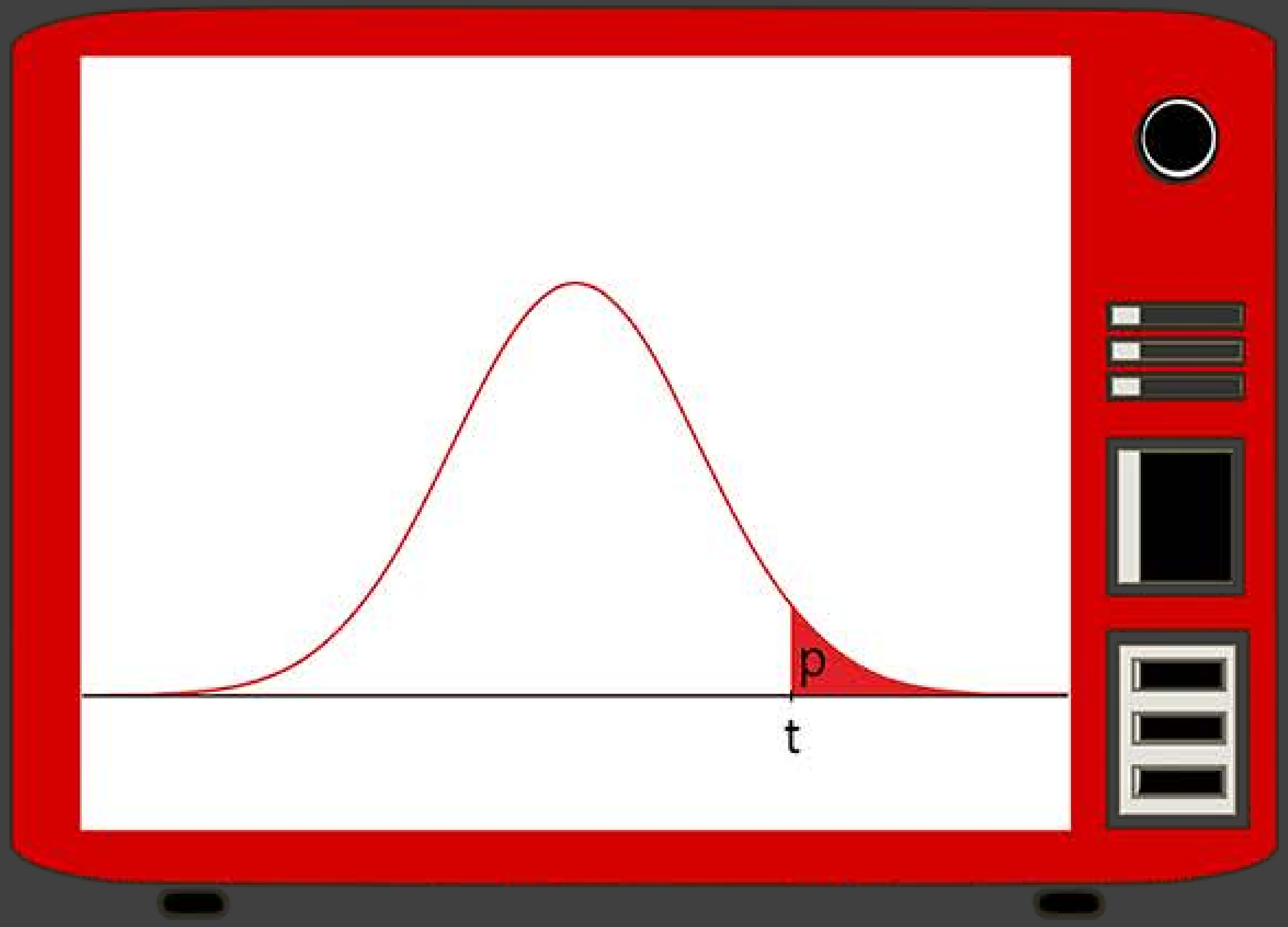
- Решаем проблему

ПОСТАНОВКА

Проверка гипотез

выборка:	$X^n = (X_1, \ldots, X_n)\,,\ X \sim \mathrm{P}$
нулевая гипотеза:	$H_0:\ \mathrm{P} \in \omega$
альтернатива:	$H_1:\ \mathrm{P} \notin \omega$
статистика:	$T\left(X^n\right)$
нулевое распределение:	$F\left(x\right)$

- Достигаемый уровень значимости: $p = \mathrm{P}(T \geq t\,|H_0\,)$.
- При $p \leq \alpha$ гипотеза отвергается.


p
t

ПРОВЕРКА ГИПОТЕЗ

	H_0 верна	H_0 неверна
H_0 принимается	H_0 верно принята	Ошибка II
H_0 отвергается	Ошибка I рода	H_0 верно отвергнута

$$\mathrm{P}(\text{ошибка I рода}) = \mathrm{P}(T \geq t \,|\, H_0) \leq \alpha.$$

Множественная проверка

данные: $\mathbf{X} = \{X_1^{n_1}, \dots, X_m^{n_m}\}, \quad X_i \sim \mathrm{P}_i$

нулевые гипотезы: $H_i\colon\ \mathrm{P}_i \in \omega_i$

альтернативы: $H_i'\colon\ \mathrm{P}_i \notin \omega_i$

статистики: $T_i = T\left(X_i^{n_i}\right)$

- Достигаемые уровни значимости:

$$p_i\,,\ i = 1, \dots, m.$$

МНОЖЕСТВЕННАЯ ПРОВЕРКА

- $\mathrm{M} = \{1, 2, \ldots, m\}$
- $\mathrm{M}_0 = \{i\colon H_i \text{ верна}\}$, $|\mathrm{M}_0| = m_0$
- $\mathrm{R} = \{i\colon H_i \text{ отвергнута}\}$, $|\mathrm{R}| = R$
- $V = |\mathrm{M}_0 \cap \mathrm{R}|$ — число ошибок первого рода.

	# верных H_i	# неверных H_i	Σ
# принятых H_i	U	T	$m - R$
# отвергнутых H_i	V	S	R
Σ	m_0	$m - m_0$	m

Множественная проверка

- $V = |\mathrm{M}_0 \cap \mathrm{R}|$— число ошибок первого рода.
- Надо что-то делать с V.

	# верных H_i	# неверных H_i	Σ
# принятых H_i	U	T	$m - R$
# отвергнутых H_i	V	S	R
Σ	m_0	$m - m_0$	m

ДАЛЕЕ В ПРОГРАММЕ

- Что-то делаем с V

FWER. ПОПРАВКА БОНФЕРРОНИ

FWER

	# верных H_i	# неверных H_i	Σ
# принятых H_i	U	T	$m - R$
# отвергнутых H_i	V	S	R
Σ	m_0	$m - m_0$	m

FWER

- Групповая вероятность ошибки первого рода (familywise error rate):
$$\mathrm{FWER} = P\left(V > 0\right).$$

- Контроль над $\mathbf{FWER}$ на уровне α:
$$\mathrm{FWER} = P\left(V > 0\right) \leq \alpha\,.$$

- Как этого добиться?

FWER

- Контроль над $\mathbf{FWER}$ на уровне α:
$$\mathbf{FWER} = P\left(V > 0\right) \leq \alpha\,.$$

- Как этого добиться?

- $\alpha_1, \ldots, \alpha_m$ — уровни значимости для $H_1, \ldots, H_m$.
  Задача — выбрать их так, чтобы обеспечить
  $\mathbf{FWER} \leq \alpha$.

Поправка Бонферрони

- Метод Бонферрони:

$$\alpha_1 = \ldots = \alpha_m = \frac{\alpha}{m}.$$

- Альтернативный вид — переход к модифицированным достигаемым уровням значимости:

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

- H_i отвергается при $\tilde{p}_i \le \alpha$.

Поправка Бонферрони

- Теорема
  Если $H_i, \; i = 1, \ldots, m$, отвергаются при $p_i \leq \alpha/m$, то $\mathbf{FWER} \leq \alpha$.

Поправка Бонферрони

- Доказательство.

$$\mathbf{FWER} = \mathbf{P}(V > 0) = \mathbf{P}\left(\bigcup_{i=1}^{m_0}\left\{p_i \le \frac{\alpha}{m}\right\}\right) \le$$

$$\le \sum_{i=1}^{m_0} \mathbf{P}\left(p_i \le \frac{\alpha}{m}\right) \le$$

$$\le \sum_{i=1}^{m_0} \frac{\alpha}{m} = \frac{m_0}{m}\alpha \le \alpha\,.$$

Модельный эксперимент

- 50 выборок из $N(1,1)$ и 150 — из $N(0,1)$, $n = 20$.
  $H_i\colon \mathbb{E}X_i = 0$, $H_i'\colon \mathbb{E}X_i \neq 0$, критерий Стьюдента.

- Без поправок:

	# верных H_i	# неверных H_i	Σ
# принятых H_i	142	0	142
# отвергнутых H_i	8	50	58
Σ	150	50	200

Модельный эксперимент

- Без поправок:

	# верных H_i	# неверных H_i	Σ
# принятых H_i	142	0	142
# отвергнутых H_i	8	50	58
Σ	150	50	200

- Бонферрони:

	# верных H_i	# неверных H_i	Σ
# принятых H_i	150	27	177
# отвергнутых H_i	0	23	23
Σ	150	50	200

Резюме

- $\mathbf{FWER}$
- Поправка Бонферрони — просто применять, работает всегда, но мощность низкая
- Далее: метод Холма для контроля $\mathbf{FWER}$

FWER. МЕТОД ХОЛМА

FWER

$$\mathbf{FWER} = P\left(V > 0\right) \leq \alpha$$

$\alpha_1, \ldots, \alpha_m$ — уровни значимости для $H_1, \ldots, H_m$.

- Задача — выбрать их так, чтобы обеспечить $\mathbf{FWER} \leq \alpha$.

- Метод Бонферрони:

$$\alpha_1 = \ldots = \alpha_m = \frac{\alpha}{m}$$

FWER

- Задача — выбрать их так, чтобы обеспечить $\mathbf{FWER} \leq \alpha$.
- Метод Бонферрони:

$$\alpha_1 = \ldots = \alpha_m = \frac{\alpha}{m}$$

- Можно сделать лучше, если α_i брать разными.

НИСХОДЯЩИЕ МЕТОДЫ

- Вариационный ряд достигаемых уровней значимости:

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

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

Нисходящие методы

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

НИСХОДЯЩИЕ МЕТОДЫ

- Процедура называется нисходящей, потому что гипотезы проверяются по убыванию значимости.

Метод Холма

- Метод Холма — нисходящая процедура со следующими уровнями значимости:

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

- Метод обеспечивает безусловный контроль над $\mathbf{FWER}$ на уровне α.

Метод Холма

- Модифицированные достигаемые уровни значимости:

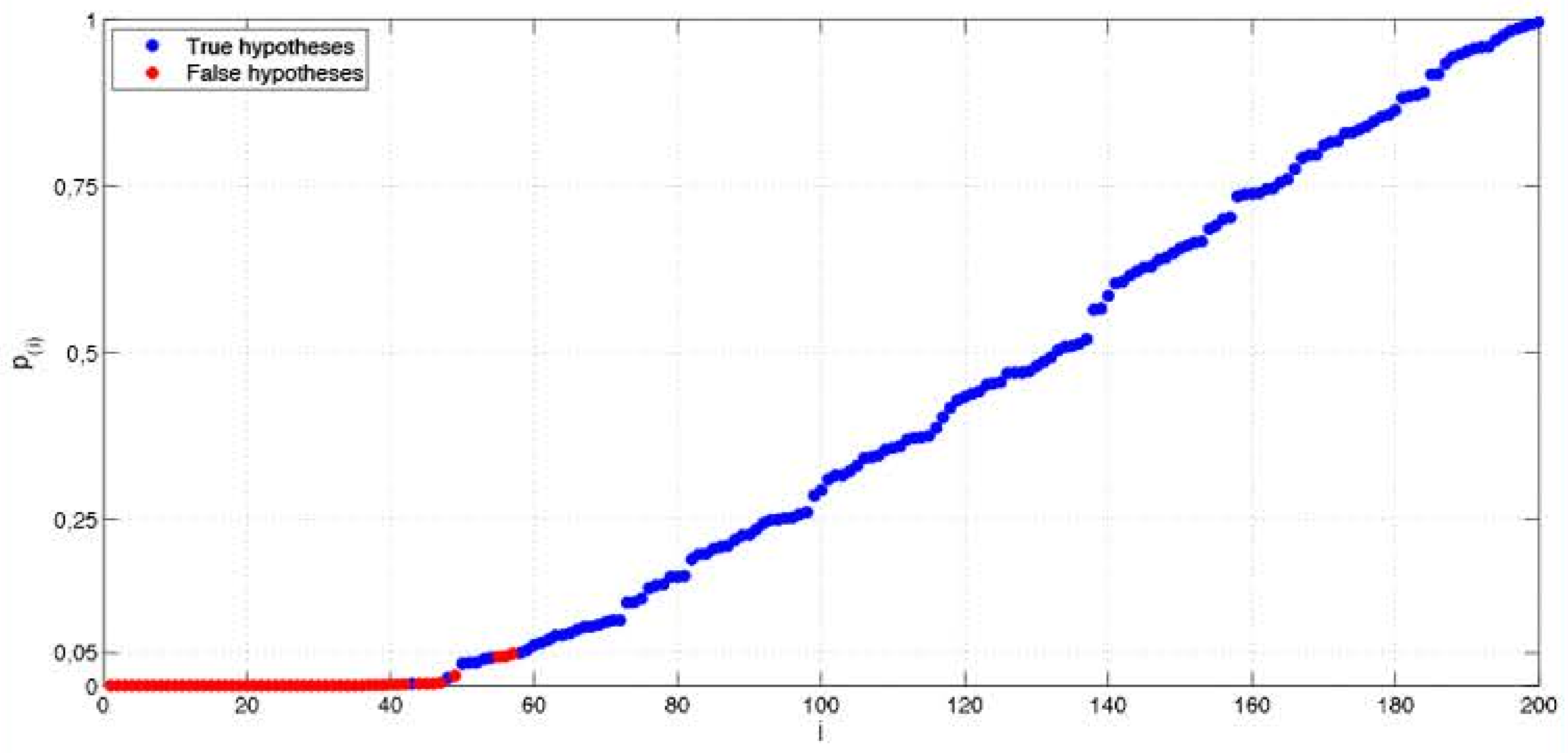
$$\tilde{p}_{(i)} = \min\left(1, \max\left((m-i+1)\,p_{(i)}, \tilde{p}_{(i-1)}\right)\right)$$

Метод Холма

- Метод всегда мощнее поправки Бонферрони, потому что его α_i всегда не меньше.

Модельный эксперимент

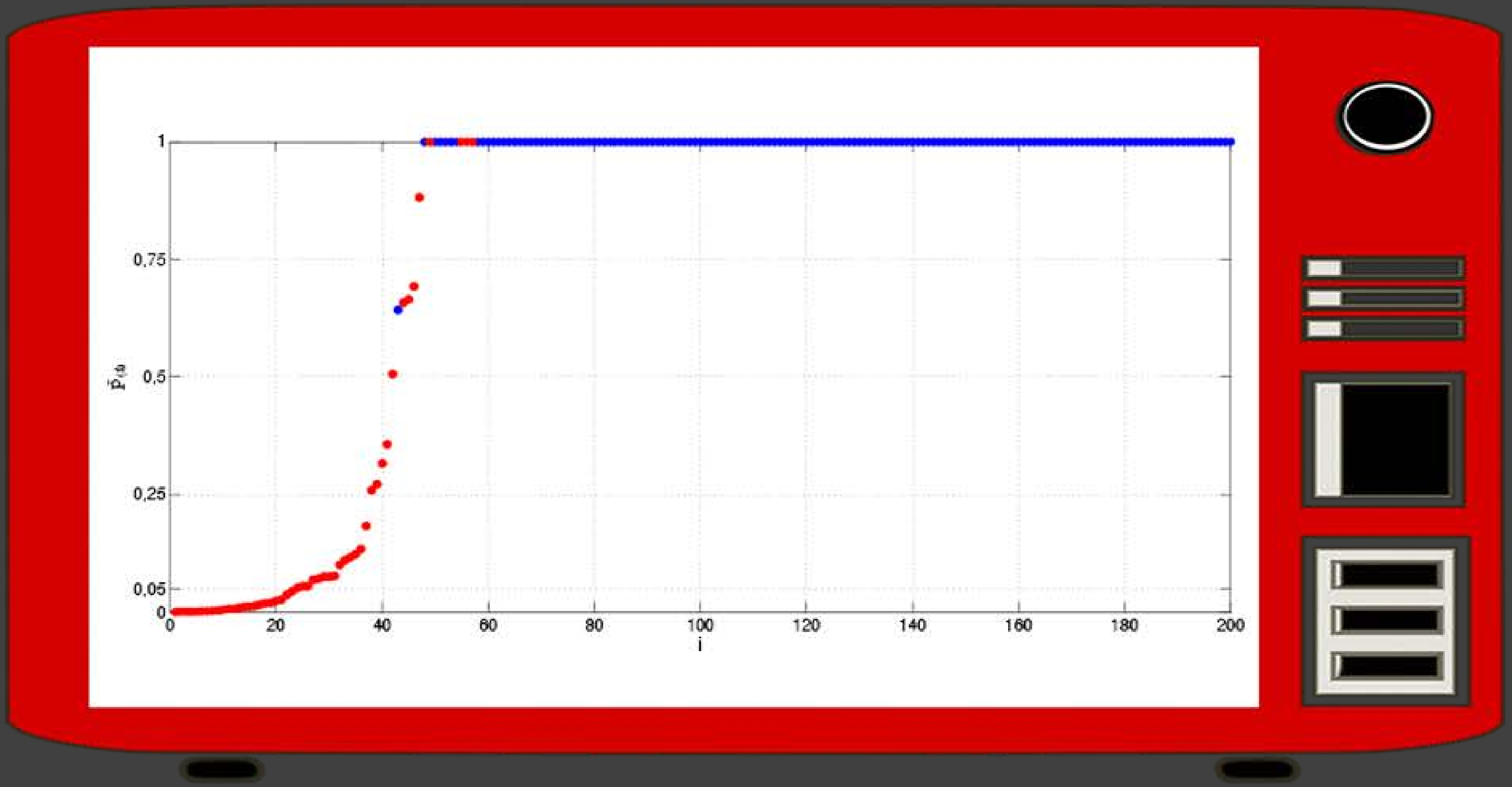
- Отсортированные достигаемые уровни значимости:


True hypotheses
False hypotheses
$p_{(i)}$
1
0,75
0,5
0,25
0,05
0
0
20
40
60
80
100
120
140
160
180
200
i

Модельный эксперимент

- $\tilde{p}_{(i)}$, метод Бонферрони:

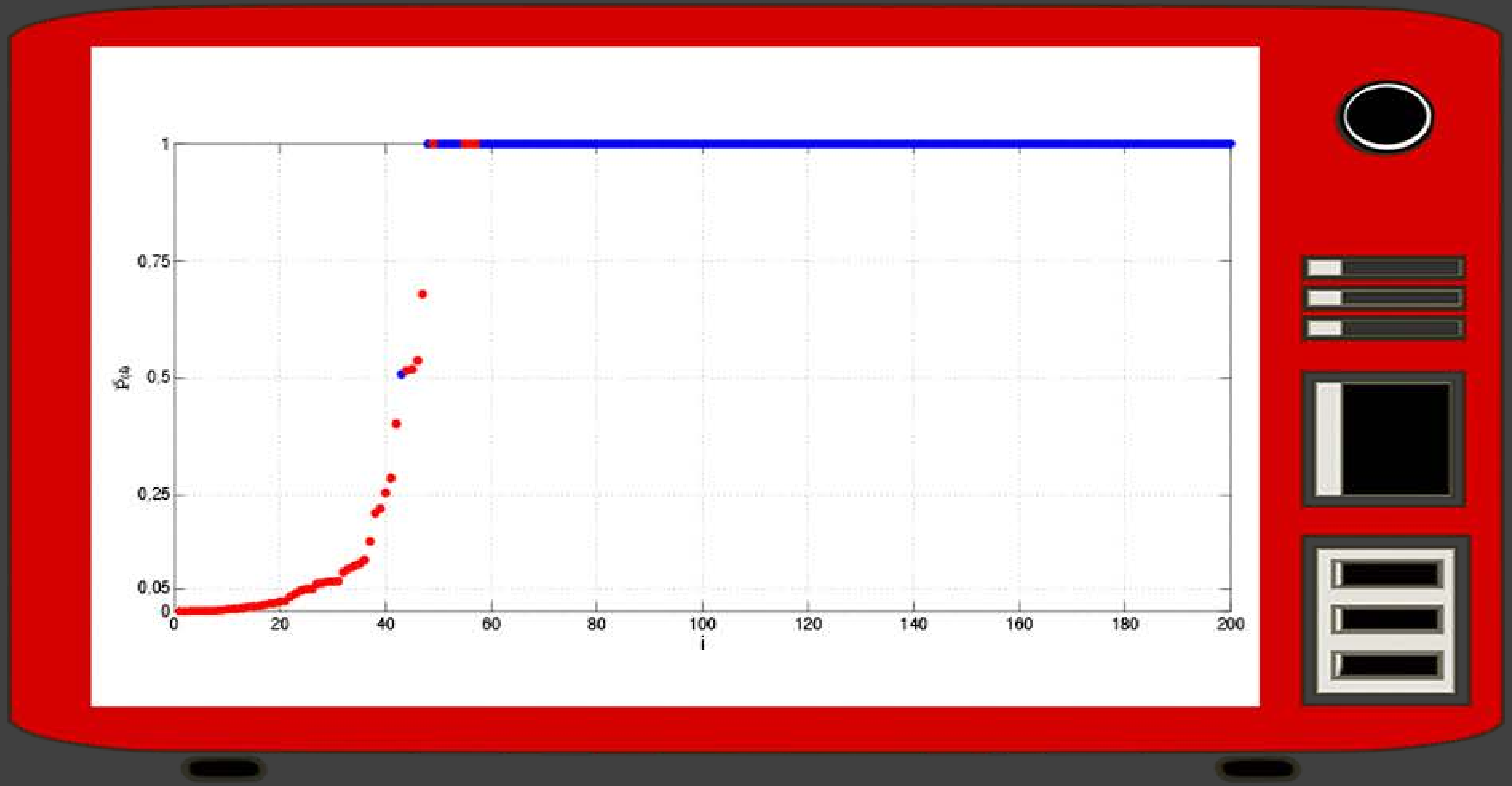
	# верных H_i	# неверных H_i	Σ
# принятых H_i	150	27	177
# отвергнутых H_i	0	23	23
Σ	150	50	200


1
0,75
0,5
0,25
0,05
0
0
20
40
60
80
100
120
140
160
180
200
i

Модельный эксперимент

- $\tilde{p}_{(i)}$, метод Холма:

	# верных H_i	# неверных H_i	Σ
# принятых H_i	150	24	174
# отвергнутых H_i	0	26	26
Σ	150	50	200


1
0,75
0,5
0,25
0,05
0
0
20
40
60
80
100
120
140
160
180
200
i

Модельный эксперимент

- Бонферрони:

	# верных H_i	# неверных H_i	Σ
# принятых H_i	150	27	177
# отвергнутых H_i	0	23	23
Σ	150	50	200

- Холм:

	# верных H_i	# неверных H_i	Σ
# принятых H_i	150	24	174
# отвергнутых H_i	0	26	26
Σ	150	50	200

РЕЗЮМЕ

- Метод Холма — немного сложнее, но всегда лучше поправки Бонферрони, работает всегда

- Далее: **FDR**

FDR. МЕТОД БЕНДЖАМИНИ-ХОХБЕРГА

FDR

	# верных H_i	# неверных H_i	Σ
# принятых H_i	U	T	$m-R$
# отвергнутых H_i	V	S	R
Σ	m_0	$m-m_0$	m

- Групповая вероятность ошибки первого рода (familywise error rate):

$$\mathrm{FWER} = P\left(V > 0\right)$$

FDR

- Групповая вероятность ошибки первого рода (familywise error rate):

$$\mathbf{FWER} = P\left(V > 0\right)$$

- Ожидаемая доля ложных отклонений (false discovery rate):

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

FDR

- Для любой процедуры множественной проверки гипотез $\mathbf{FDR} \le \mathbf{FWER}$.

Восходящие методы

- Вариационный ряд достигаемых уровней значимости:

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

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

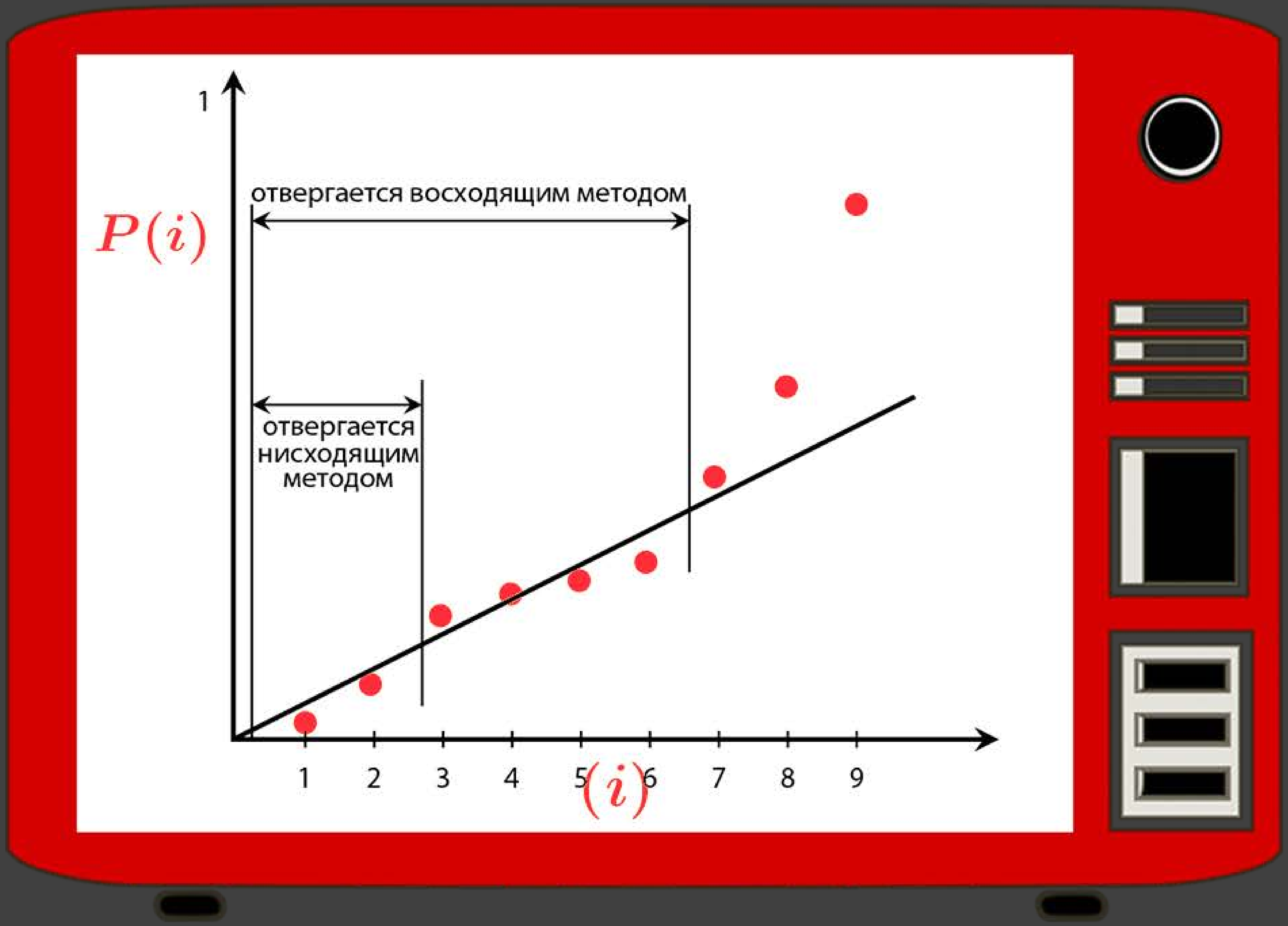
- Если $p_{(m)} \leq \alpha_m$, отвергнуть все нулевые гипотезы $H_{(1)}, H_{(2)}, \dots, H_{(m)}$ и остановиться; иначе принять $H_{(m)}$ и продолжить.

Восходящие методы

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

Восходящие методы

- Восходящая процедура всегда отвергает не меньше гипотез, чем нисходящая с теми же α_i:


1
отвергается восходящим методом
$P(i)$
отвергается
нисходящим
методом
1 2 3 4 5 6 7 8 9
(i)

Метод Бенджамини-Хохберга

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

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

- Модифицированные достигаемые уровни значимости:

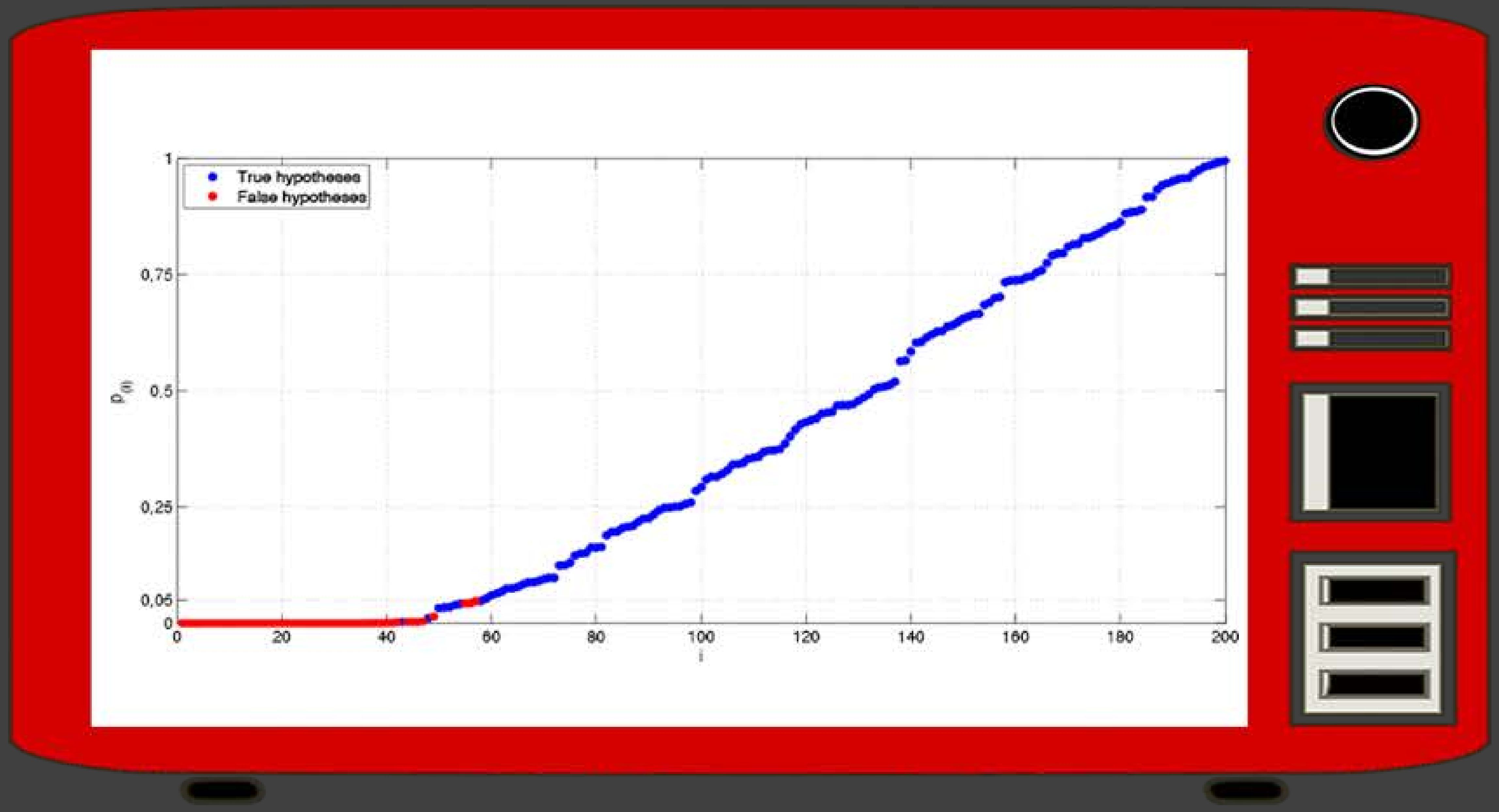
$$\tilde{p}_{(i)} = \min\left(1, \frac{m p_{(i)}}{i}, \tilde{p}_{(i+1)}\right)$$

Метод Бенджамини-Хохберга

- Метод обеспечивает контроль над FDR на уровне α при условии, что проверяющие гипотезы статистики независимы.

Модельный эксперимент

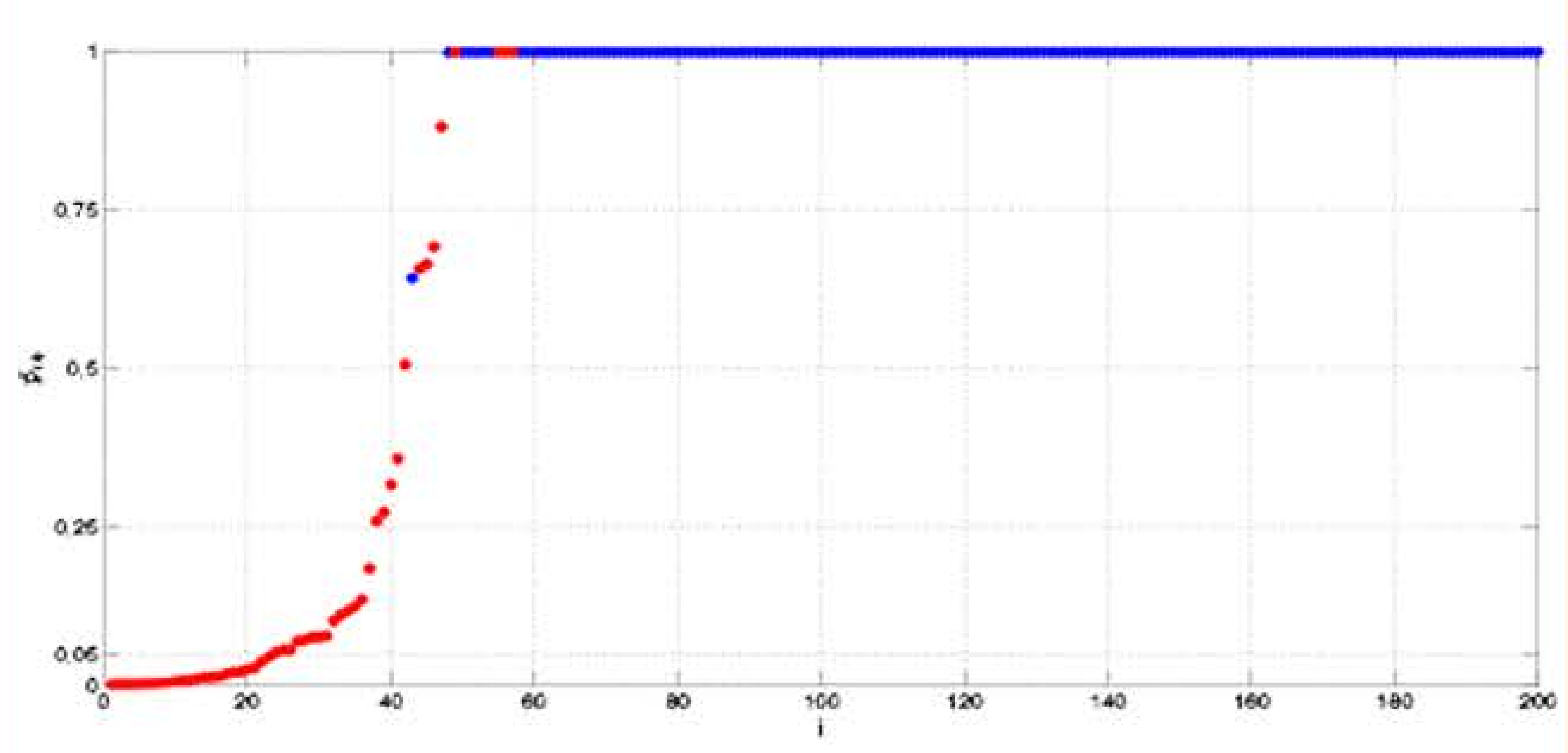
- Отсортированные достигаемые уровни значимости:


True hypotheses
False hypotheses
1
0,75
0,5
0,25
0,05
0
0
20
40
60
80
100
120
140
160
180
200
i
$p_{(i)}$

Модельный эксперимент

- $\tilde{p}_{(i)}$, метод Холма:

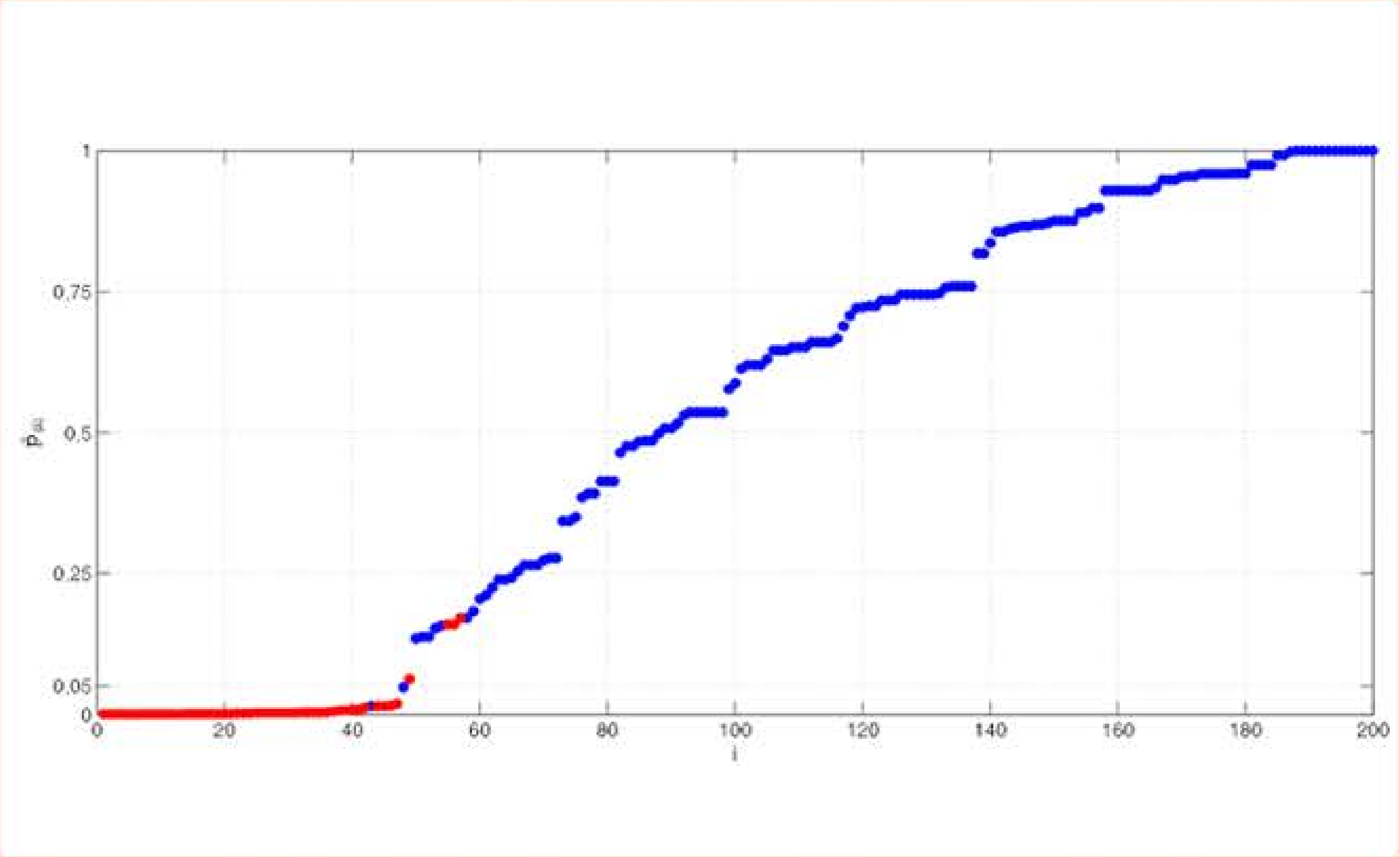
	# верных H_i	# неверных H_i	Σ
# принятых H_i	150	24	174
# отвергнутых H_i	0	26	26
Σ	150	50	200


1
0.75
0.5
0.25
0.05
0
0
20
40
60
80
100
120
140
160
180
200
i

Модельный эксперимент

- $\tilde{p}_{(i)}$, метод Бенджамини-Хохберга:

	# верных H_i	# неверных H_i	Σ
# принятых H_i	148	4	152
# отвергнутых H_i	2	46	48
Σ	150	50	200



1
0,75
0,5
0,25
0,05
0
0
20
40
60
80
100
120
140
160
180
200
i

Модельный эксперимент

- Холм:

	# верных H_i	# неверных H_i	Σ
# принятых H_i	150	24	174
# отвергнутых H_i	0	26	26
Σ	150	50	200

- Бенджамини-Хохберг:

	# верных H_i	# неверных H_i	Σ
# принятых H_i	148	4	152
# отвергнутых H_i	2	46	48
Σ	150	50	200

Резюме

- Контролируя $\mathbf{FDR}$ вместо $\mathbf{FWER}$, мы позволяем больше ошибок I рода и отвергаем больше гипотез
- Метод Бенджамини-Хохберга работает не всегда, но используется повсеместно
- Далее: анализ подгрупп

АНАЛИЗ ПОДГРУПП

ПОДГРУППЫ

- 1073 пациента с ишемической болезнью сердца были разделены на две группы по типу лечения. Исследовалась выживаемость пациентов.

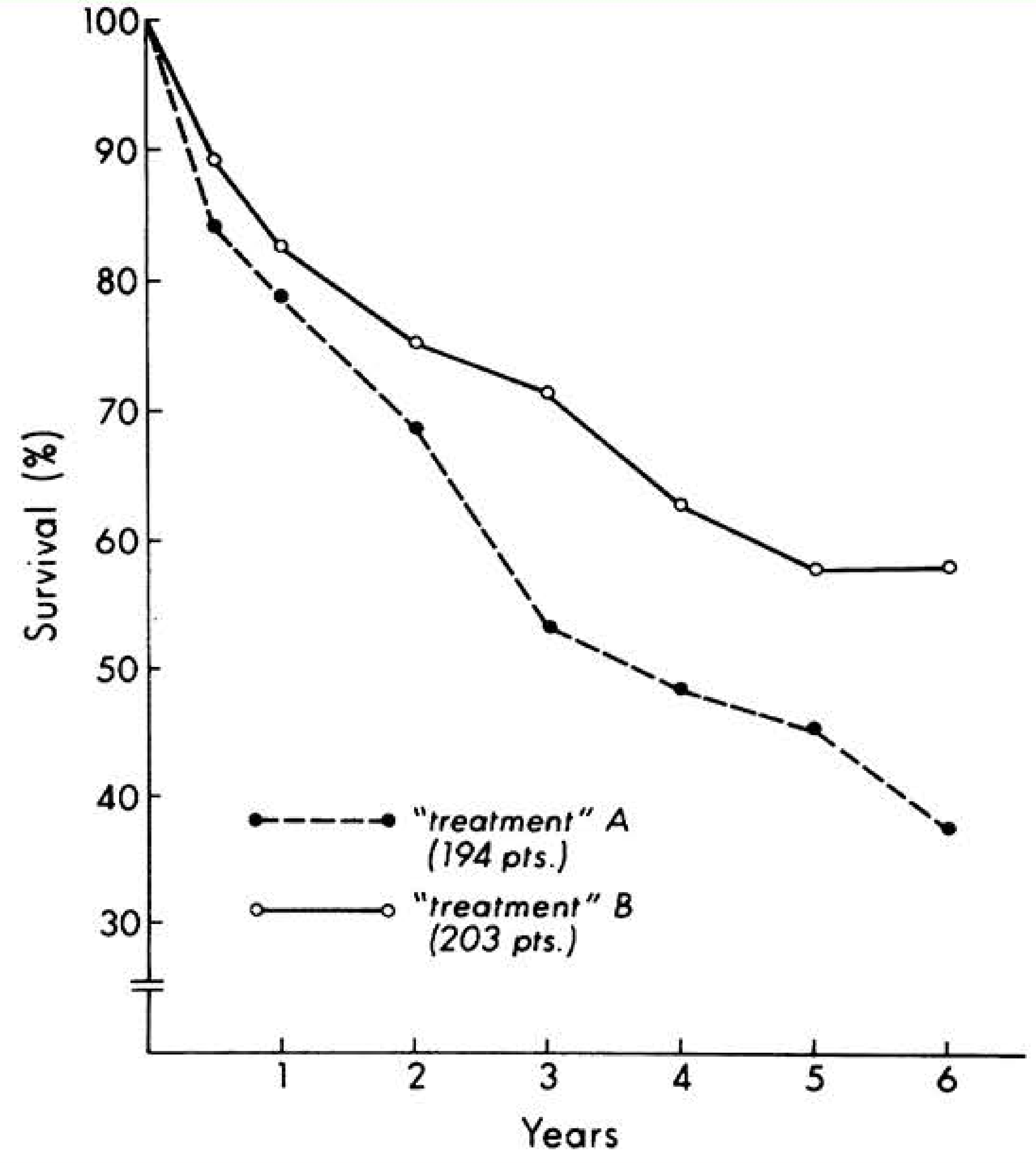
- Важные факторы, влияющие на выживаемость, — число поражённых артерий (1, 2, 3) и тип сокращений левого желудочка (нормальный, абнормальный).

ПОДГРУППЫ

- Важные факторы, влияющие на выживаемость, — число поражённых артерий (1, 2, 3) и тип сокращений левого желудочка (нормальный, абнормальный).

- Для одной из шести подгрупп были обнаружены значимыеразличия в выживаемости пациентов первого и второго типов.

ПОДГРУППЫ



Survival (%)
100
90
80
70
60
50
40
30
1
2
3
4
5
6
Years
"treatment" A
(194 pts.)
"treatment" B
(203 pts.)

Подгруппы

- Исследовалась связь потребления кофеина и рака груди, всего около 50 подгрупп. Показано, что:
  - употребление ≥ 4 чашек кофе в день связано с увеличением риска злокачественного рака груди ($p = 0.08$)

ПОДГРУППЫ

- Исследовалась связь потребления кофеина и рака груди, всего около 50 подгрупп. Показано, что:

  - потребление кофеина связано с увеличением риска возникновения эстроген- и прогестерон-независимых опухолей и опухолей больше 2 см ($p = 0.02$ и $p = 0.02$)

Подгруппы

- Исследовалась связь потребления кофеина и рака груди, всего около 50 подгрупп. Показано, что:

  - потребление кофе без кофеина связано со снижением риска рака груди у женщин в постменопаузе, принимающих гормоны ($p = 0.02$)

ПОДГРУППЫ

- (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.”

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, Бенджамини-Хохберг:

p_1 сравнивается с $\dfrac{0.05}{2} = 0.025$

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

- Лучшее решение проблемы множественной проверки гипотез — проверять меньше гипотез.

РЕЗЮМЕ

- При анализе подгрупп нужно делать поправку на множественную проверку
- Чем меньше гипотез, тем меньше боли