

НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ

НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ

» $X^n = (X_1, \dots, X_n)$, $X \sim F(x)$

» Равно ли среднее \bar{X} нулю?

Непараметрические критерии

Непараметрические критерии

$X^n = (X_1, \dots, X_n), X \sim F(x)$

Равно ли среднее X нулю?

Непараметрические критерии

$$X^n = (X_1, \dots, X_n), X \sim F(x)$$

Равно ли среднее X нулю?

Статистика T ; нулевое распределение — ?

Непараметрические критерии

$$X^n = (X_1, \dots, X_n), X \sim F(x)$$

Равно ли среднее X нулю?

Статистика T ; нулевое распределение — ?

Проблемы:

- ▶ распределение $F(x)$ может быть нестандартным
- ▶ ЦПТ работает не всегда

Непараметрические критерии

$$X^n = (X_1, \dots, X_n), X \sim F(x)$$

Равно ли среднее X нулю?

Статистика T ; нулевое распределение — ?

Решения:

- ▶ превратить выборку во что-то более понятное
- ▶ сделать какие-то предположения о $F(x)$

НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ

- » $X^n = (X_1, \dots, X_n)$, $X \sim F(x)$
- » Равно ли среднее \bar{X} нулю?
- » Статистика T ; нулевое распределение — ?

НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ

- » $X^n = (X_1, \dots, X_n)$, $X \sim F(x)$
- » Равно ли среднее \bar{X} нулю?
- » Статистика T ; нулевое распределение — ?
- » Проблемы:
 - ▶ Распределение $F(x)$ может быть нестандартным
 - ▶ ЦПТ работает не всегда

НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ

- » $X^n = (X_1, \dots, X_n)$, $X \sim F(x)$
- » Равно ли среднее \bar{X} нулю?
- » Статистика T ; нулевое распределение — ?

» Проблемы:

- ▶ Распределение $F(x)$ может быть нестандартным
- ▶ ЦПТ работает не всегда

НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ

- » $X^n = (X_1, \dots, X_n)$, $X \sim F(x)$
- » Равно ли среднее \bar{X} нулю?
- » Статистика T ; нулевое распределение — ?
- » Решения:
 - ▶ Превратить выборку во что-то более понятное
 - ▶ Сделать какие-то предположения о $F(x)$

» Решения:

- ▶ Превратить выборку во что-то более понятное
- ▶ Сделать какие-то предположения о $F(x)$

КРИТЕРИИ ЗНАКОВ

ВРЕМЯ РЕМОНТА

- » Время ремонта оборудования местных клиентов провайдера Verizon ($n = 23$):

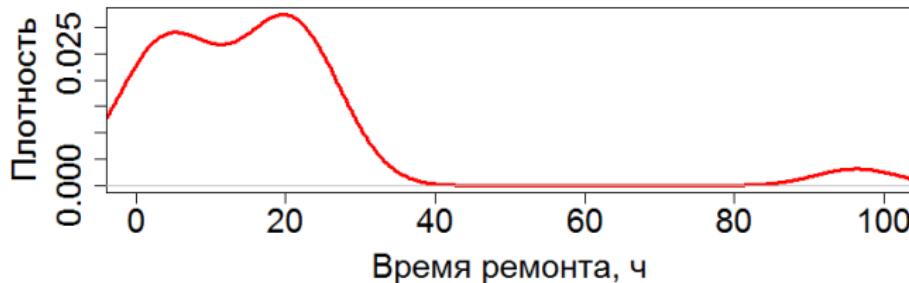
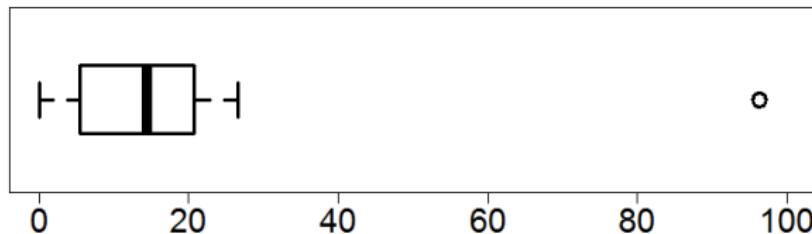
ВРЕМЯ РЕМОНТА

- › Время ремонта оборудования местных клиентов провайдера Verizon ($n = 23$):
- › Можно ли утверждать, что среднее время больше восьми часов?

Критерии знаков

Время ремонта

Время ремонта оборудования местных клиентов провайдера Verizon ($n = 23$):



Можно ли утверждать, что среднее время больше восьми часов?

Одновыборочный критерий знаков

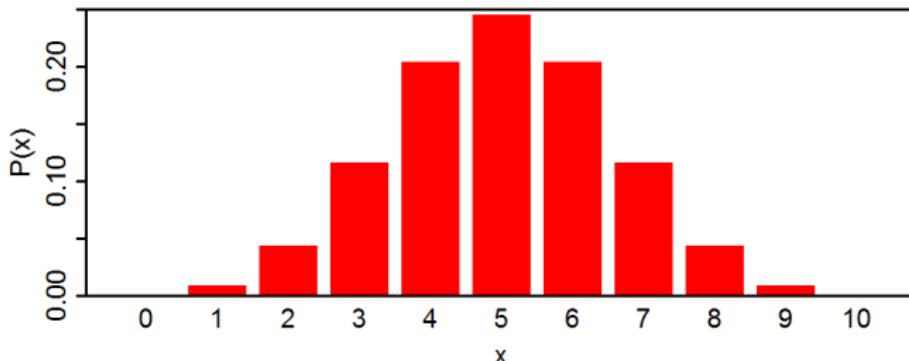
выборка: $X^n = (X_1, \dots, X_n)$, $X_i \neq m_0$;

нулевая гипотеза: $H_0: \text{med } X = m_0$;

альтернатива: $H_1: \text{med } X < \neq > m_0$;

статистика: $T(X^n) = \sum_{i=1}^n [X_i > m_0]$;

нулевое распределение: $T(X^n) \sim Bin(n, \frac{1}{2})$.



Время ремонта

H_0 : среднее время ремонта — 8 часов, $\text{med } X = 8$.

H_1 : ремонт в среднем длится дольше 8 часов, $\text{med } X > 8$.

Ремонт занял больше 8 часов в 15 случаях из 23.

Критерий знаков: $p = 0.105$, нельзя утверждать, что ремонт в среднем длится дольше 8 часов.

Цензурированная выборка

Выживаемость пациентов с лимфоцитарной лимфомой
(в неделях):

49, 58, 75, 110, 112, 132, 151, 276, 281, 362*

Исследование длилось 7 лет, поэтому для пациентов,
проживших дольше, выживаемость неизвестна.
Выборка цензурирована сверху.

Превышает ли среднее время дожития 200 недель?
Критерий знаков: $p = 0.9453$.

Качество классификаторов

	AUC _{C4.5}	AUC _{C4.5+m}
adult (sample)	0.763	0.768
breast cancer	0.599	0.591
breast cancer wisconsin	0.954	0.971
cmc	0.628	0.661
ionosphere	0.882	0.888
iris	0.936	0.931
liver disorders	0.661	0.668
lung cancer	0.583	0.583
lymphography	0.775	0.838
mushroom	1.000	1.000
primary tumor	0.940	0.962
rheum	0.619	0.666
voting	0.972	0.981
wine	0.957	0.978

Качество классификаторов

	AUC _{C4.5}	AUC _{C4.5+m}
adult (sample)	0.763	0.768
breast cancer	0.599	0.591
breast cancer wisconsin	0.954	0.971
cmc	0.628	0.661
ionosphere	0.882	0.888
iris	0.936	0.931
liver disorders	0.661	0.668
lung cancer	0.583	0.583
lymphography	0.775	0.838
mushroom	1.000	1.000
primary tumor	0.940	0.962
rheum	0.619	0.666
voting	0.972	0.981
wine	0.957	0.978

Двухвыборочный критерий знаков

выборки: $X_1^n = (X_{11}, \dots, X_{1n})$,

$X_2^n = (X_{21}, \dots, X_{2n})$,

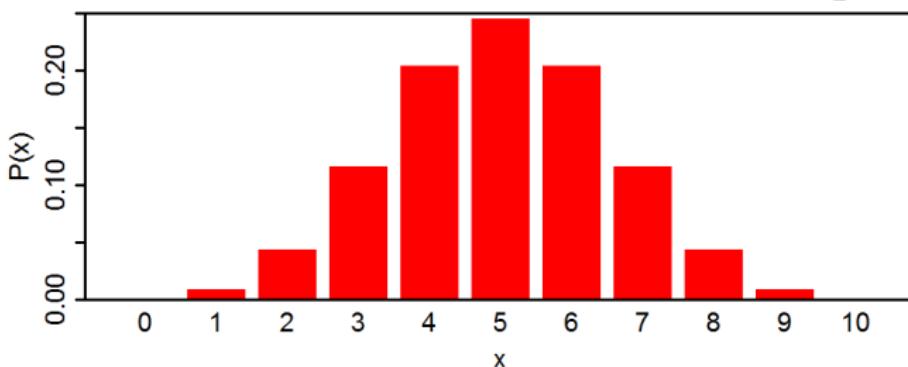
$X_{1i} \neq X_{2i}$, выборки связанные;

нулевая гипотеза: $H_0: P(X_1 > X_2) = \frac{1}{2}$;

альтернатива: $H_1: P(X_1 > X_2) < \neq > \frac{1}{2}$;

статистика: $T(X_1^n, X_2^n) = \sum_{i=1}^n [X_{1i} > X_{2i}]$;

нулевое распределение: $T(X_1^n, X_2^n) \sim Bin(n, \frac{1}{2})$.



Качество классификаторов

H_0 : у классификаторов одинаковое среднее качество,
 $\mathbf{P}(\text{AUC}_{C4.5+m} > \text{AUC}_{C4.5}) = \frac{1}{2}$.

H_1 : среднее качество модифицированного
классификатора выше, $\mathbf{P}(\text{AUC}_{C4.5+m} > \text{AUC}_{C4.5}) > \frac{1}{2}$.

Модифицированный алгоритм выигрывает на 10 датасетах из 14, ещё на 2 ничья.

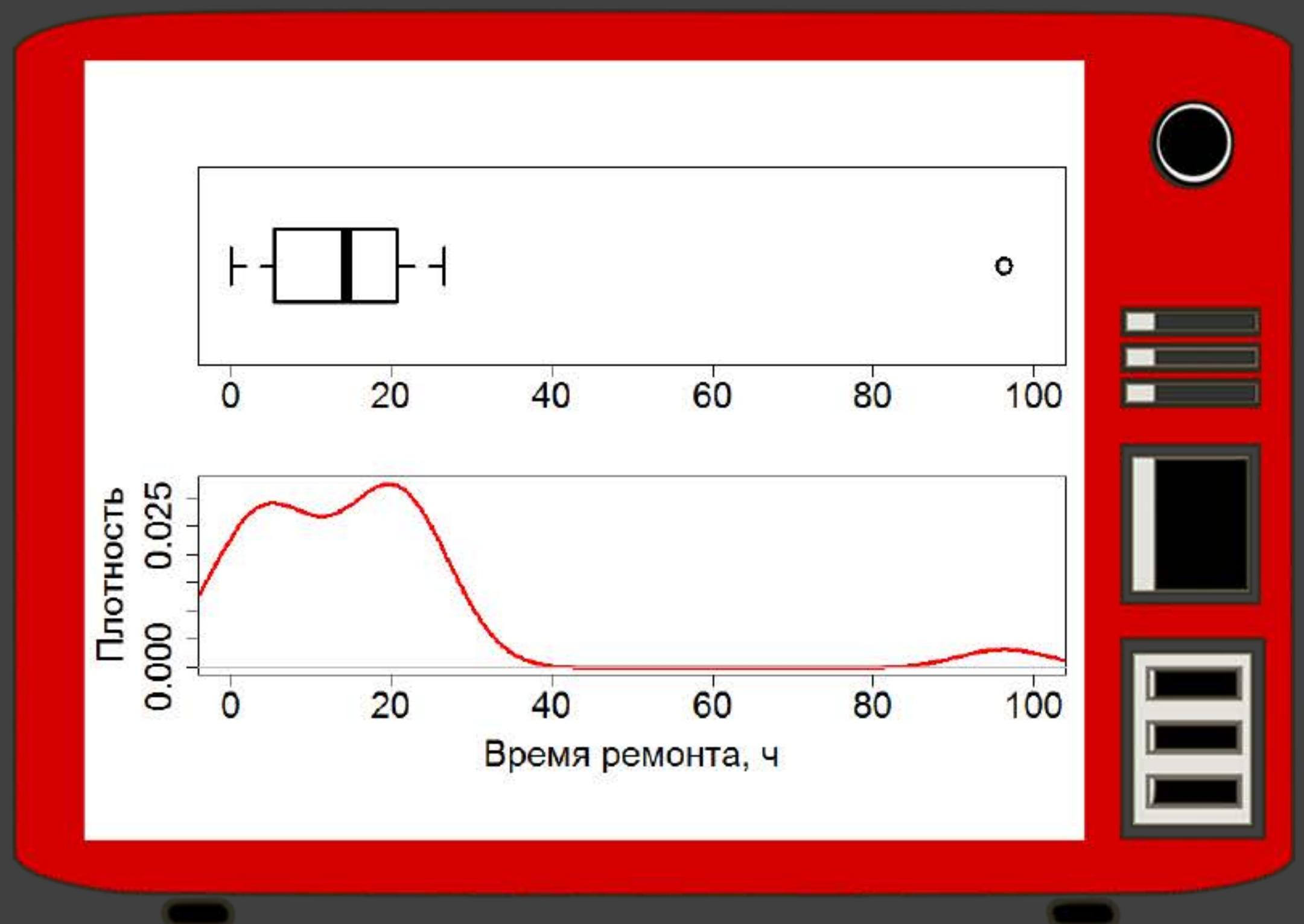
Критерий знаков: $p = 0.019$, модифицированный алгоритм лучше на 83% датасетов (95% нижний доверительный предел — 56.2%).

Резюме

- ▶ выборки из неизвестных распределений можно превращать в нули и единицы
- ▶ критерий знаков, одновыборочный и для связанных выборок

Далее в программе

- ▶ ранговые критерии



ОДНОВЫБОРОЧНЫЙ КРИТЕРИЙ ЗНАКОВ

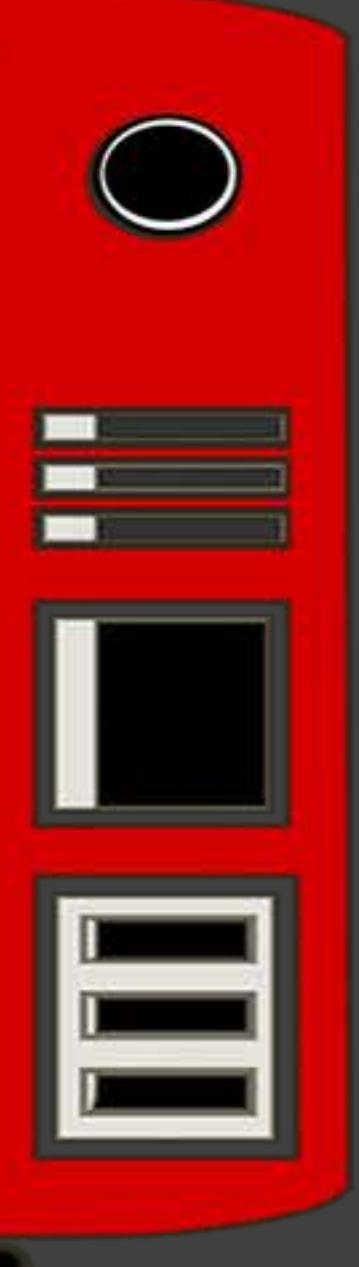
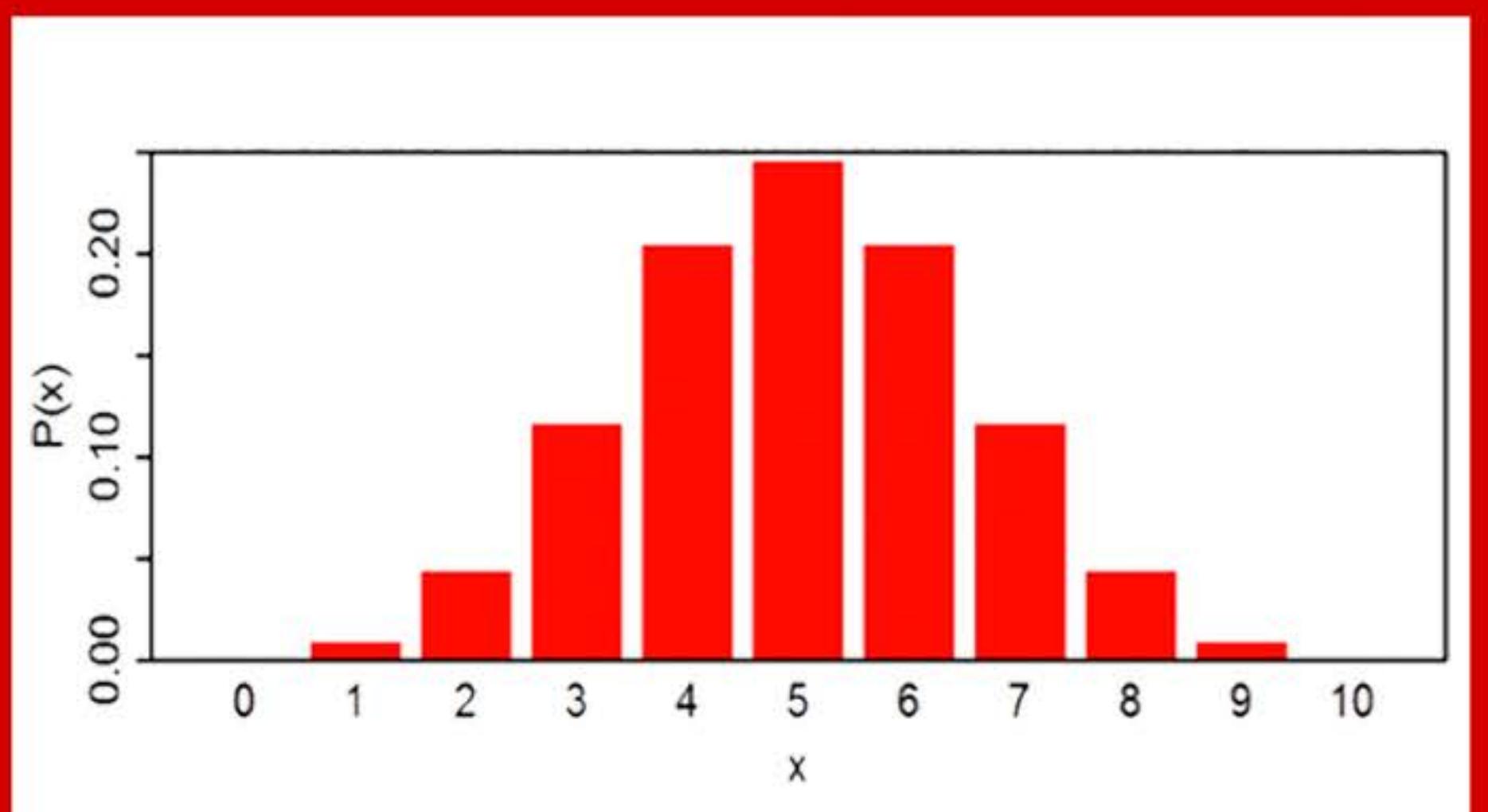
выборка: $X^n = (X_1, \dots, X_n), X_i \neq m_0$

нулевая гипотеза: $H_0: \text{med } X = m_0$

альтернатива: $H_1: \text{med } X <\neq> m_0$

статистика: $T(X^n) = \sum_{i=1}^n [X_i > m_0]$

нулевое распределение: $T(X^n) \sim Bin(n, \frac{1}{2})$



- » H_0 : среднее время ремонта — 8 часов,
 $\text{med } X = 8$
- » H_1 : медиана времени ремонта — 8 часов,
 $\text{med } X > 8$
- » Ремонт занял больше 8 часов в 15 случаях из 23

- » Ремонт занял больше 8 часов в 15 случаях из 23
- » Критерий знаков: $p = 0.105$, нельзя утверждать, что ремонт в среднем длится дольше 8 часов

ЦЕНЗУРИРОВАННАЯ ВЫБОРКА

- Выживаемость пациентов с лимфоцитарной лимфомой (в неделях):

49, 58, 75, 110, 112, 132, 151, 276, 281, 362*

- Исследование длилось 7 лет, поэтому для пациентов, проживших дольше, выживаемость неизвестна

ЦЕНЗУРИРОВАННАЯ ВЫБОРКА

- › Исследование длилось **7** лет, поэтому для пациентов, проживших дольше, выживаемость неизвестна
- › Превышает ли среднее время дожития **200** недель?
- › Критерий знаков: **$p = 0.9453$**

КАЧЕСТВО КЛАССИФИКАТОРОВ

	$AUC_{C4.5}$	$AUC_{C4.5+m}$
adult (sample)	0.763	0.768
breast cancer	0.599	0.591
breast cancer wisconsin	0.954	0.971
cmc	0.628	0.661
ionosphere	0.882	0.888
iris	0.936	0.931
liver disorders	0.661	0.668
lung cancer	0.583	0.583
lymphography	0.775	0.838
mushroom	1.000	1.000
primary tumor	0.940	0.962
rheum	0.619	0.666
voting	0.972	0.981
wine	0.957	0.978

КАЧЕСТВО КЛАССИФИКАТОРОВ

	$AUC_{C4.5}$	$AUC_{C4.5+m}$
adult (sample)	0.763	0.768
breast cancer	0.599	0.591
breast cancer wisconsin	0.954	0.971
cmc	0.628	0.661
ionosphere	0.882	0.888
iris	0.936	0.931
liver disorders	0.661	0.668
lung cancer	0.583	0.583
lymphography	0.775	0.838
mushroom	1.000	1.000
primary tumor	0.940	0.962
rheum	0.619	0.666
voting	0.972	0.981
wine	0.957	0.978

ДВУХВЫБОРОЧНЫЙ КРИТЕРИЙ ЗНАКОВ

выборки: $X_1^n = (X_{11}, \dots, X_{1n})$

$X_2^n = (X_{21}, \dots, X_{2n})$

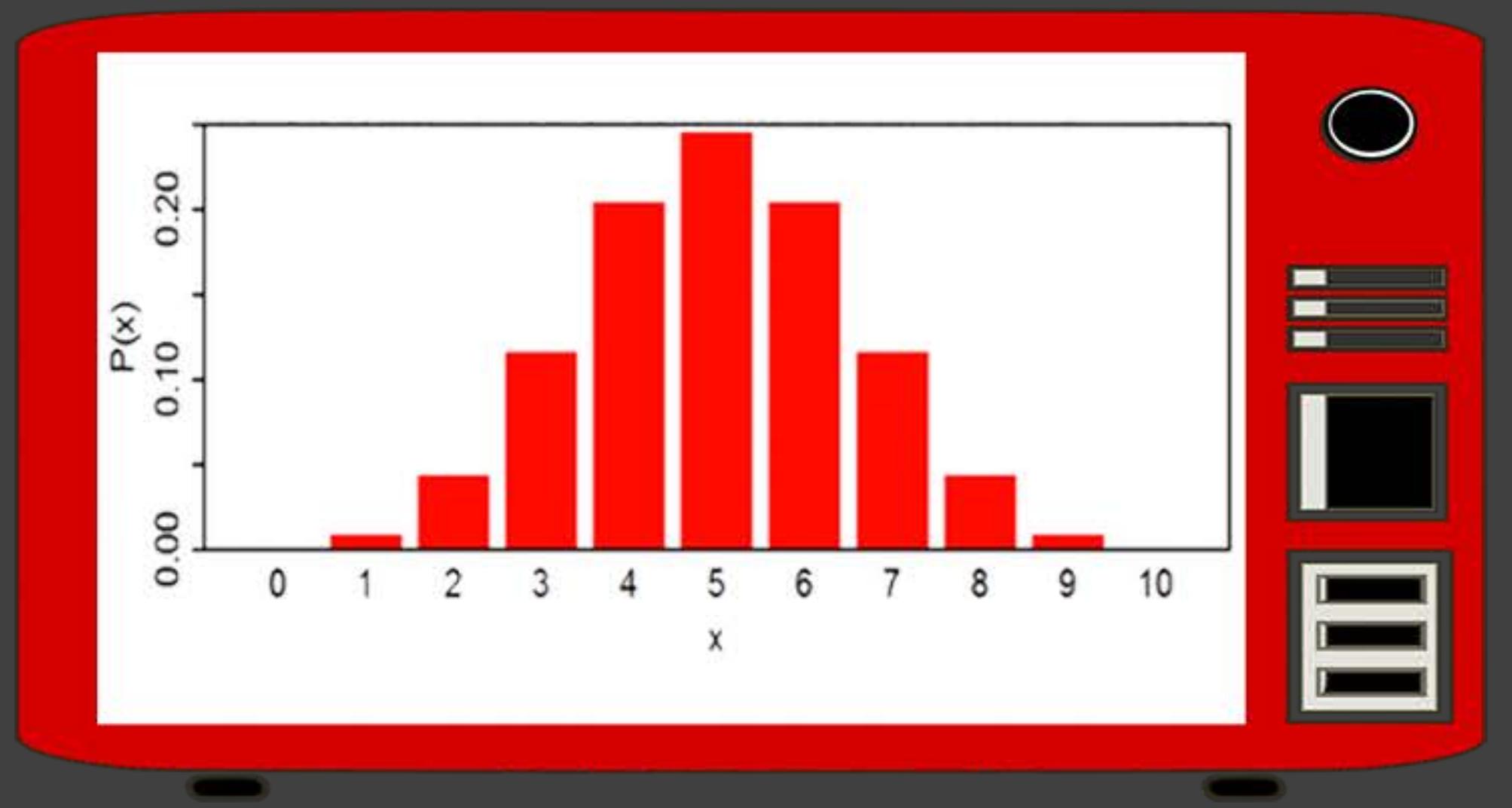
$X_{1i} \neq X_{2i}$, выборки связанные

нулевая гипотеза: $H_0: P(X_1 > X_2) = \frac{1}{2}$

альтернатива: $H_1: P(X_1 > X_2) < \neq > \frac{1}{2}$

статистика: $T(X_1^n, X_2^n) = \sum_{i=1}^n [X_{1i} > X_{2i}]$

нулевое распределение: $T(X_1^n, X_2^n) \sim Bin(n, \frac{1}{2})$



КАЧЕСТВО КЛАССИФИКАТОРОВ

- » H_0 : у классификаторов одинаковое среднее качество,

$$P(\text{AUC}_{C4.5+m} > \text{AUC}_{C4.5}) = \frac{1}{2}$$

- » H_1 : среднее качество модифицированного классификатора выше,

$$P(\text{AUC}_{C4.5+m} > \text{AUC}_{C4.5}) > \frac{1}{2}$$

- » Модифицированный алгоритм выигрывает на 10 датасетах из 14, ещё на 2 ничья

КАЧЕСТВО КЛАССИФИКАТОРОВ

- › Модифицированный алгоритм выигрывает на 10 датасетах из 14, ещё на 2 ничья
- › Критерий знаков: $p = 0.019$, модифицированный алгоритм лучше на 83% датасетов (95% нижний доверительный предел — 56.2%)

- › Выборки из неизвестных распределений можно превращать в нули и единицы
- › Критерий знаков, одновыборочный и для связанных выборок
- › Далее: ранговые критерии

РАНГОВЫЕ КРИТЕРИИ

ВАРИАЦИОННЫЙ РЯД

» $X_1, \dots, X_n \Rightarrow$

$$X_{(1)} \leq \dots < \underbrace{X_{(k_1)} = \dots = X_{(k_2)}}_{\text{связка размера } k_2 - k_1 + 1} < \dots \leq X_{(n)}$$

» Ранг наблюдения X_i :

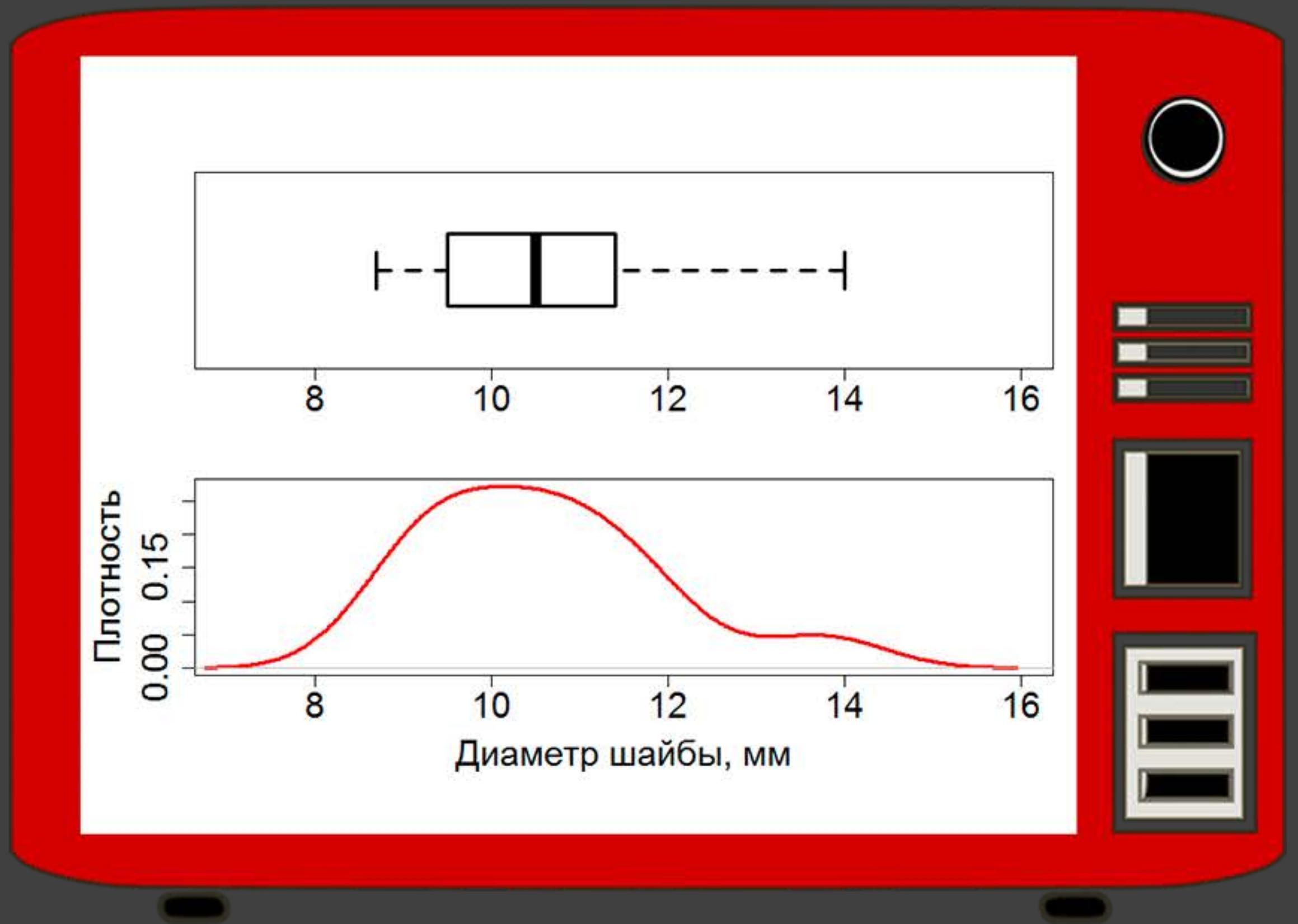
если X_i не в связке, то $\text{rank}(X_i) = r$: $X_i = X_{(r)}$,

если X_i в связке $X_{(k_1)}, \dots, X_{(k_2)}$, то

$$\text{rank}(X_i) = \frac{k_1 + k_2}{2}$$

ДИАМЕТР ШАЙБЫ

- » Диаметр шайб на производстве ($n = 24$):
- » Соответствуют ли шайбы стандартному размеру 10 мм?



КРИТЕРИЙ ЗНАКОВЫХ РАНГОВ

выборка: $X^n = (X_1, \dots, X_n), X_i \neq m_0$

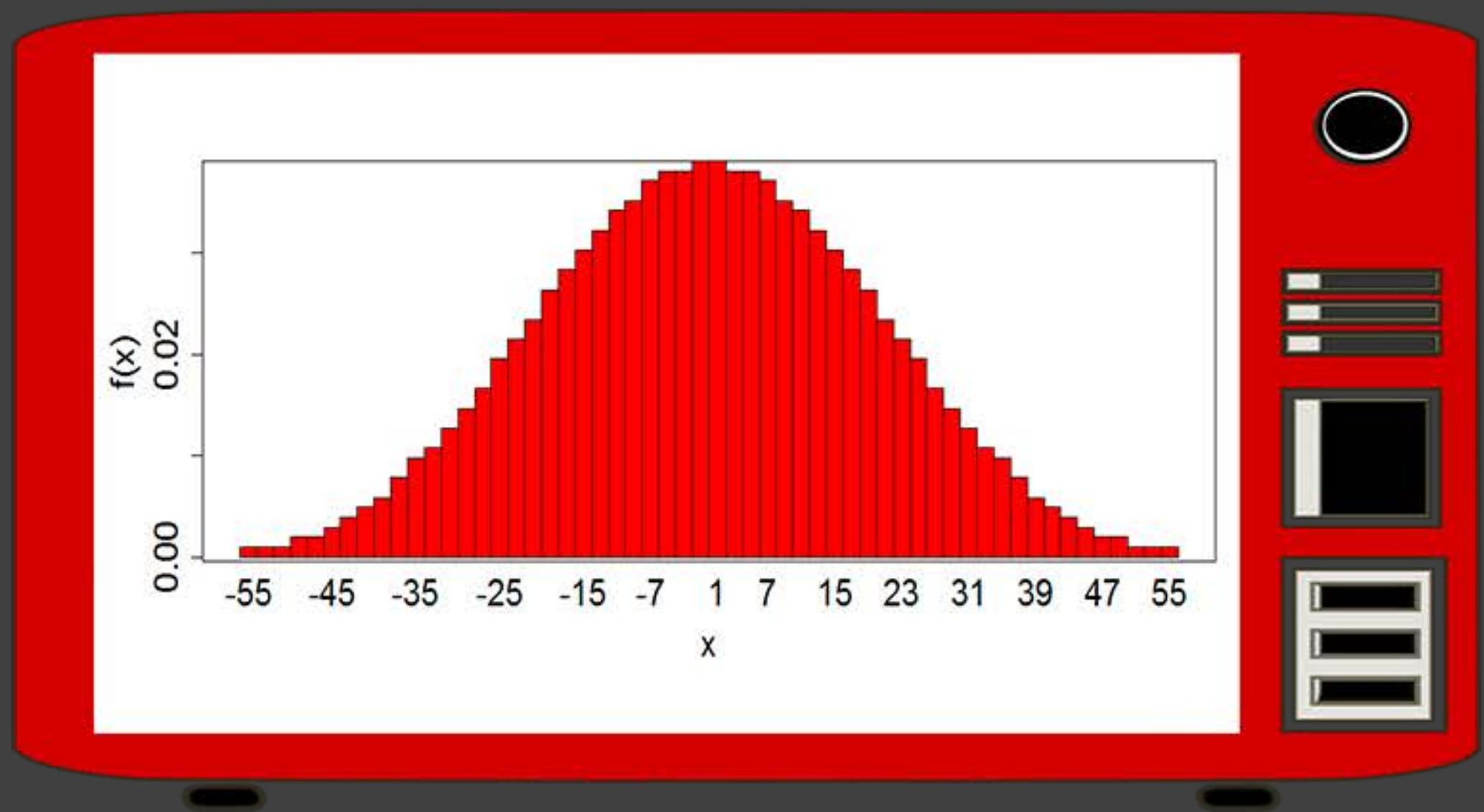
F_X симметрично относительно медианы

нулевая гипотеза: $H_0: \text{med } X = m_0$

альтернатива: $H_1: \text{med } X <\neq> m_0$

статистика: $W(X^n) = \sum_{i=1}^n \text{rank}(|X_i - m_0|) \cdot$
 $\cdot \text{sign}(X_i - m_0)$

нулевое распределение: табличное

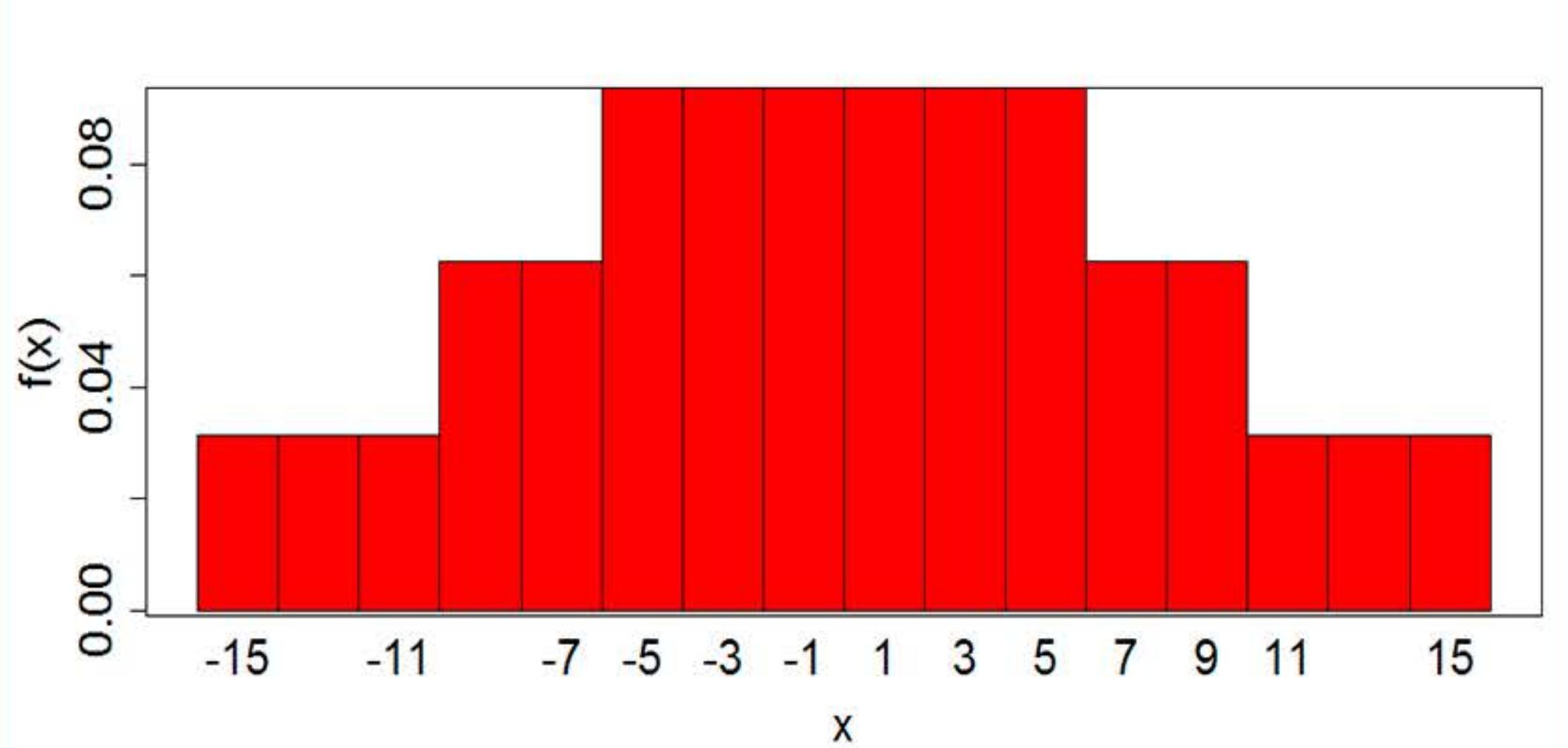


НУЛЕВОЕ РАСПРЕДЕЛЕНИЕ

1	2	3	4	5	W	Всего 2^n вариантов
-	-	-	-	-	-15	
+	-	-	-	-	-13	
-	+	-	-	-	-11	
...	
+	-	+	+	+	11	
-	+	+	+	+	13	
+	+	+	+	+	15	

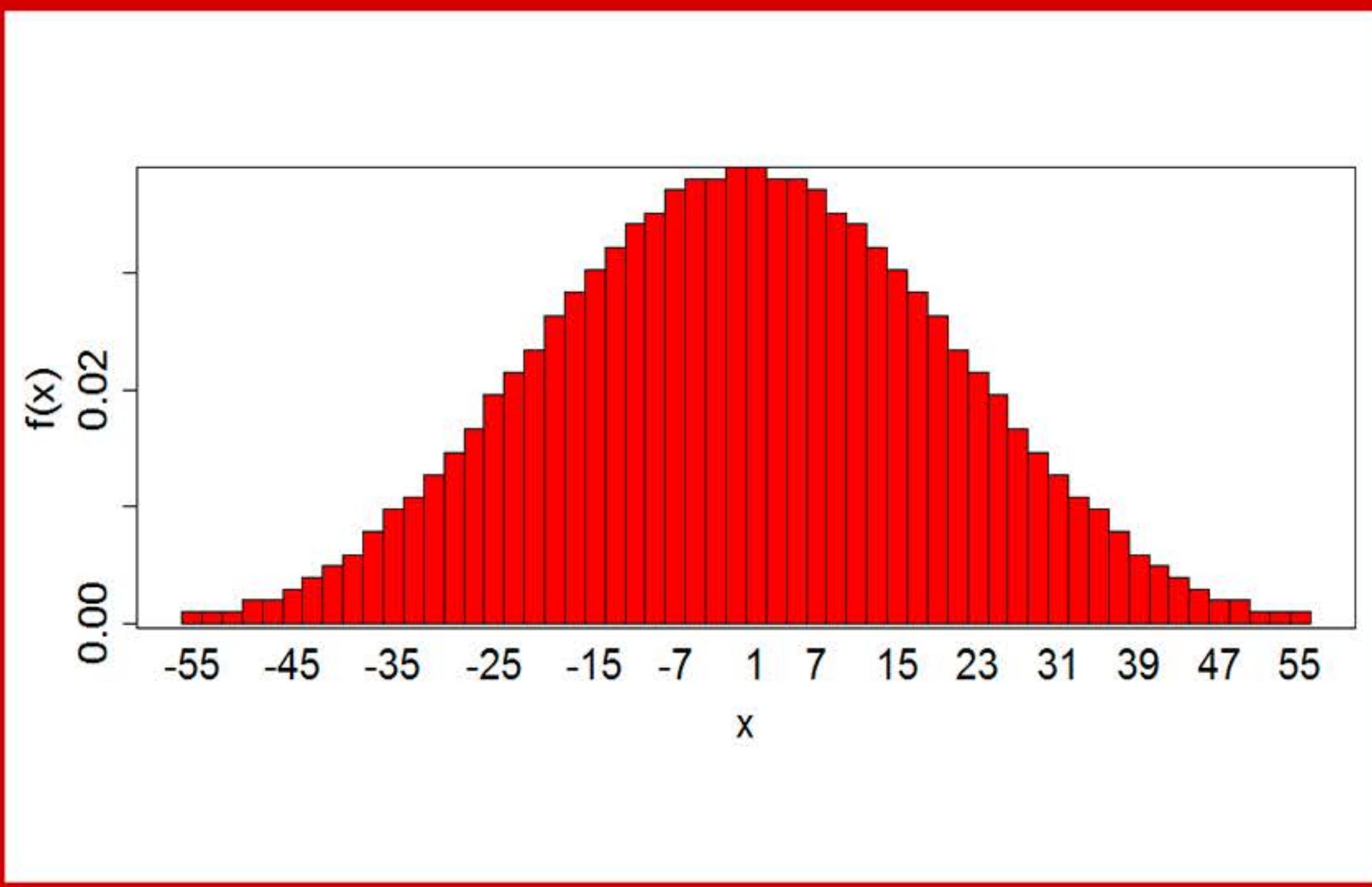
НУЛЕВОЕ РАСПРЕДЕЛЕНИЕ

» $n = 5$



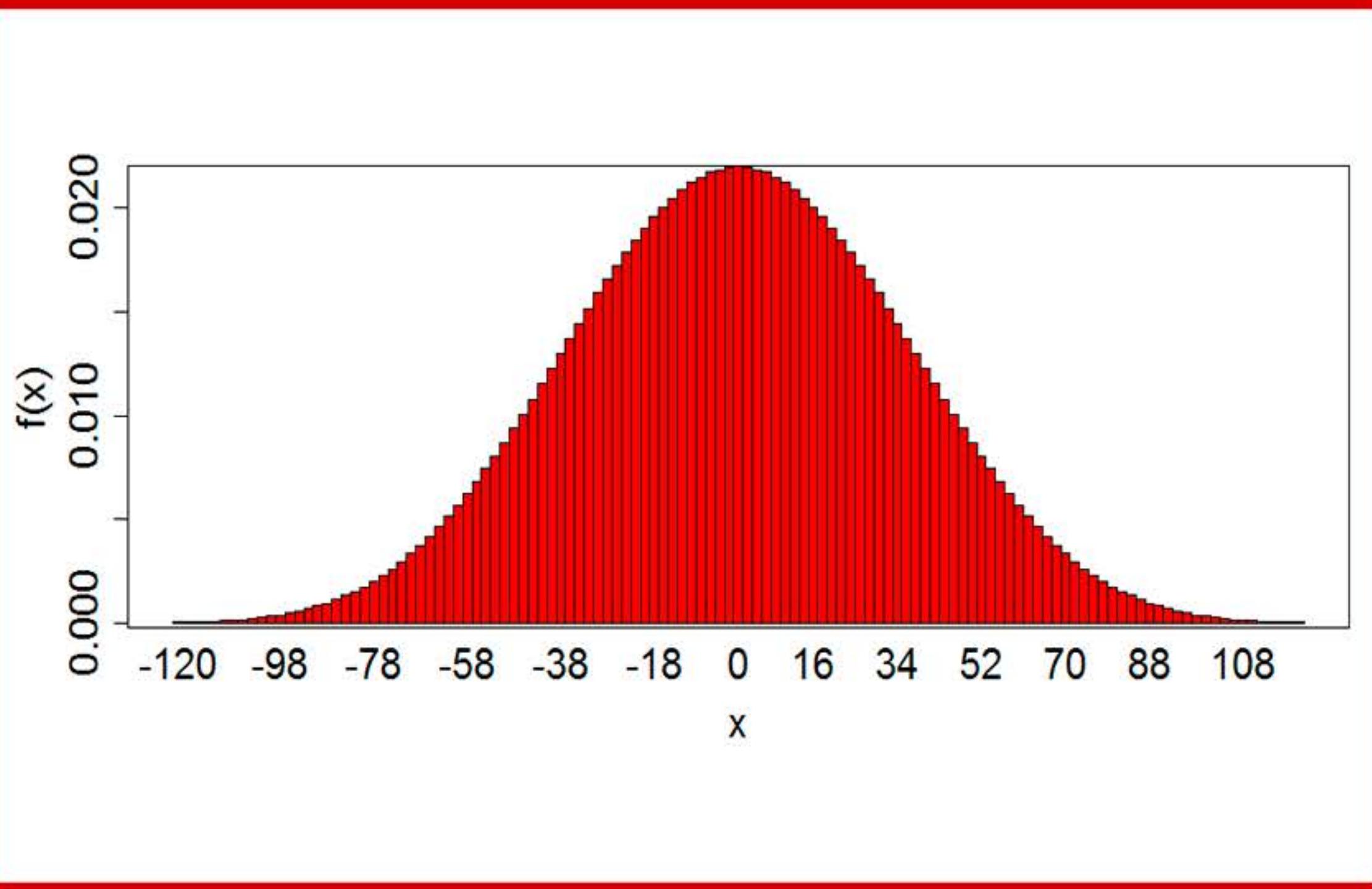
НУЛЕВОЕ РАСПРЕДЕЛЕНИЕ

» $n = 10$



НУЛЕВОЕ РАСПРЕДЕЛЕНИЕ

» $n = 15$



НУЛЕВОЕ РАСПРЕДЕЛЕНИЕ

» $n = 15$

Апроксимация для $n > 20$:

$$W \approx N \left(0, \frac{n(n+1)(2n+1)}{6} \right)$$

ДИАМЕТР ШАЙБЫ

- » H_0 : средний диаметр шайбы — 10 мм, $\text{med } X = 10$
- » H_1 : средний диаметр шайбы не соответствует стандарту, $\text{med } X \neq 10$
- » Критерий знаковых рангов: $p = 0.0673$, выборочная медиана диаметра — 10.5 мм (95% доверительный интервал — [9.95, 11.15] мм)

КРИТЕРИЙ ЗНАКОВЫХ РАНГОВ

выборки: $X_1^n = (X_{11}, \dots, X_{1n})$

$X_2^n = (X_{21}, \dots, X_{2n})$

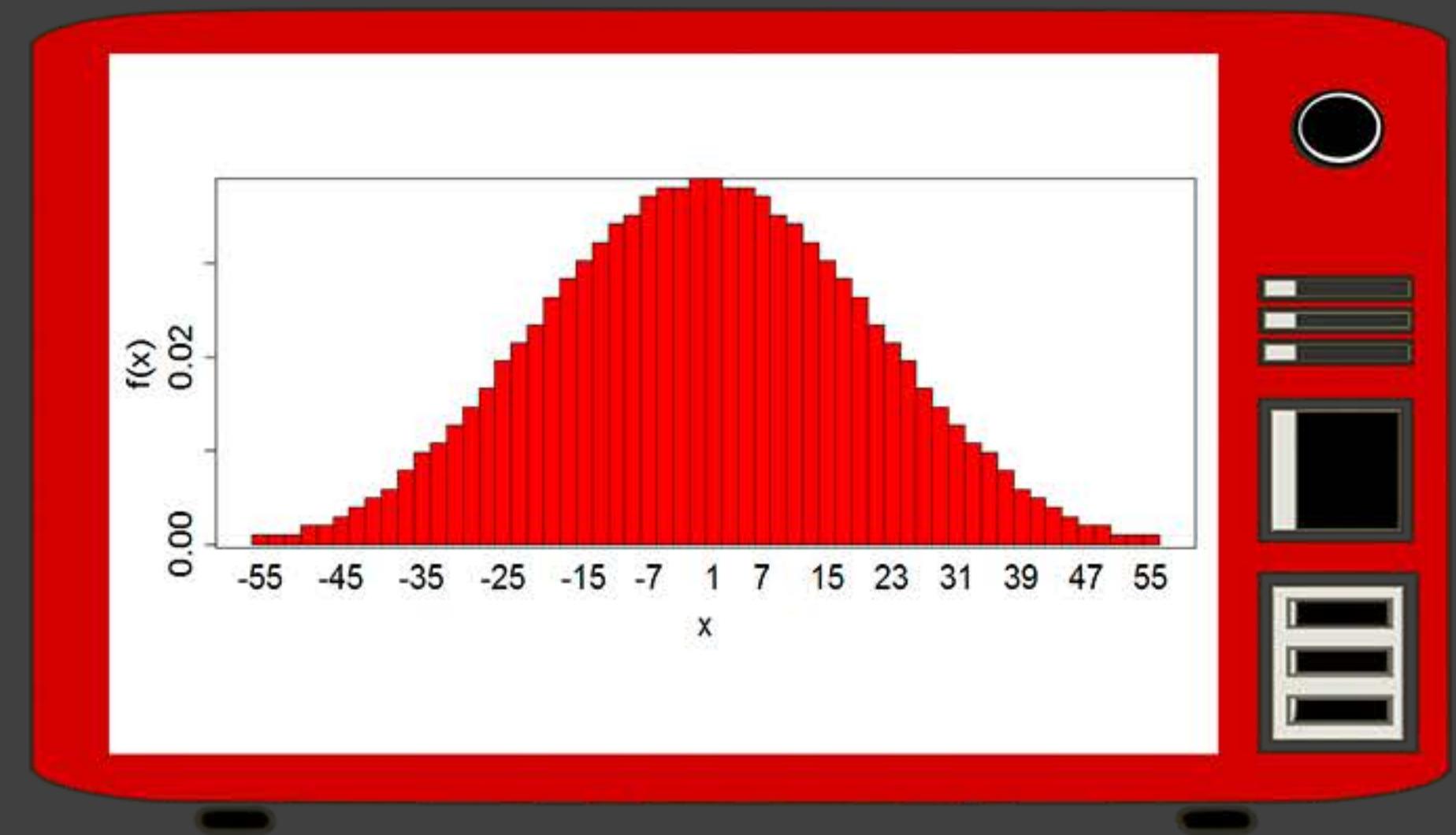
$X_{1i} \neq X_{2i}$, выборки связанные

нулевая гипотеза: $H_0: \text{med}(X_1 - X_2) = 0$

альтернатива: $H_1: \text{med}(X_1 - X_2) < \neq > 0$

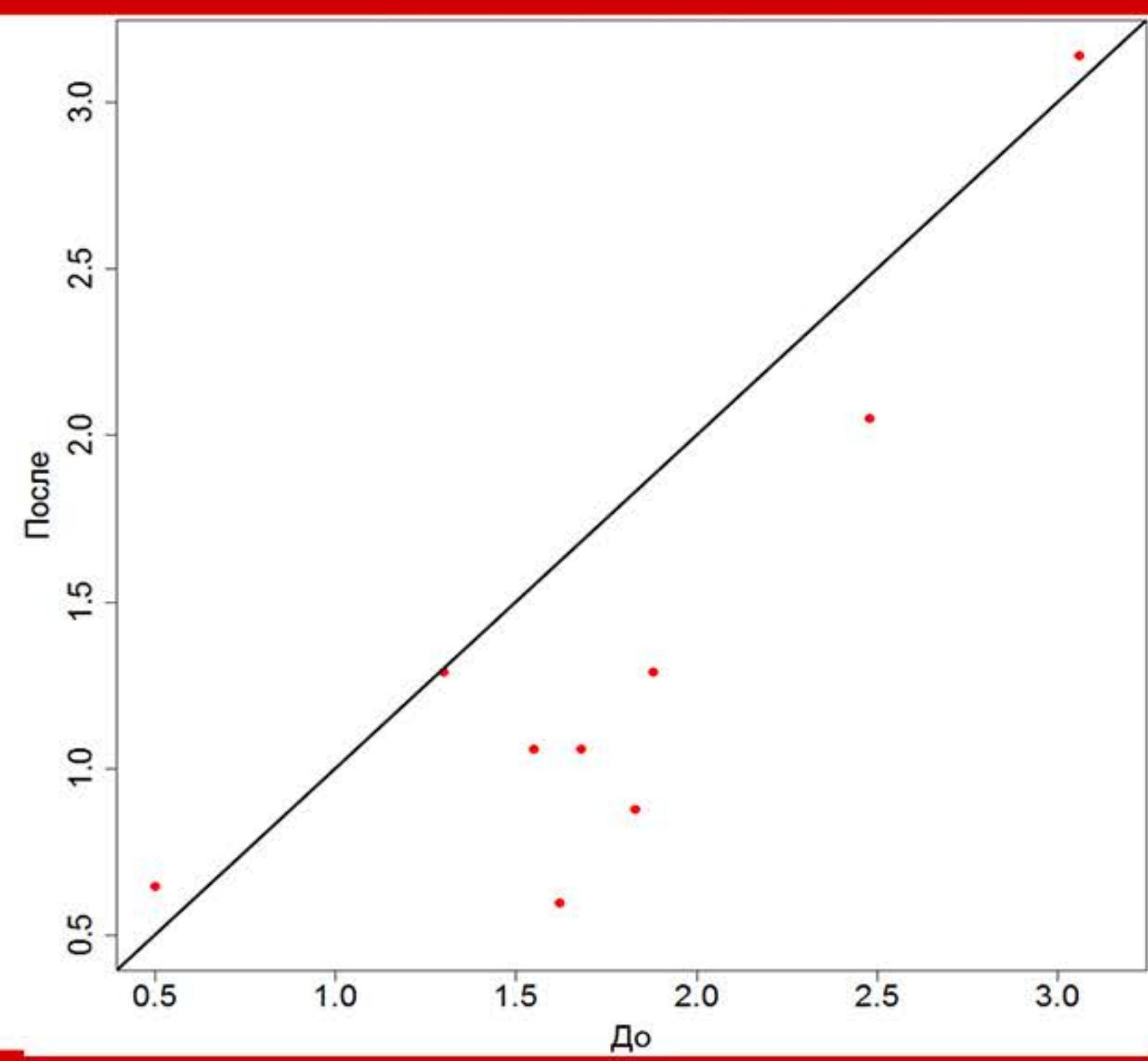
статистика: $W(X_1^n, X_2^n) = \sum_{i=1}^n \text{rank}(|X_{1i} - X_{2i}|) \cdot$
 $\cdot \text{sign}(X_{1i} - X_{2i})$

нулевое распределение: табличное



ЛЕЧЕНИЕ ДЕПРЕССИИ

- » Депрессивность 9 пациентов измерена по шкале Гамильтона до и после первого приёма транквилизатора. Подействовал ли транквилизатор?



ЛЕЧЕНИЕ ДЕПРЕССИИ

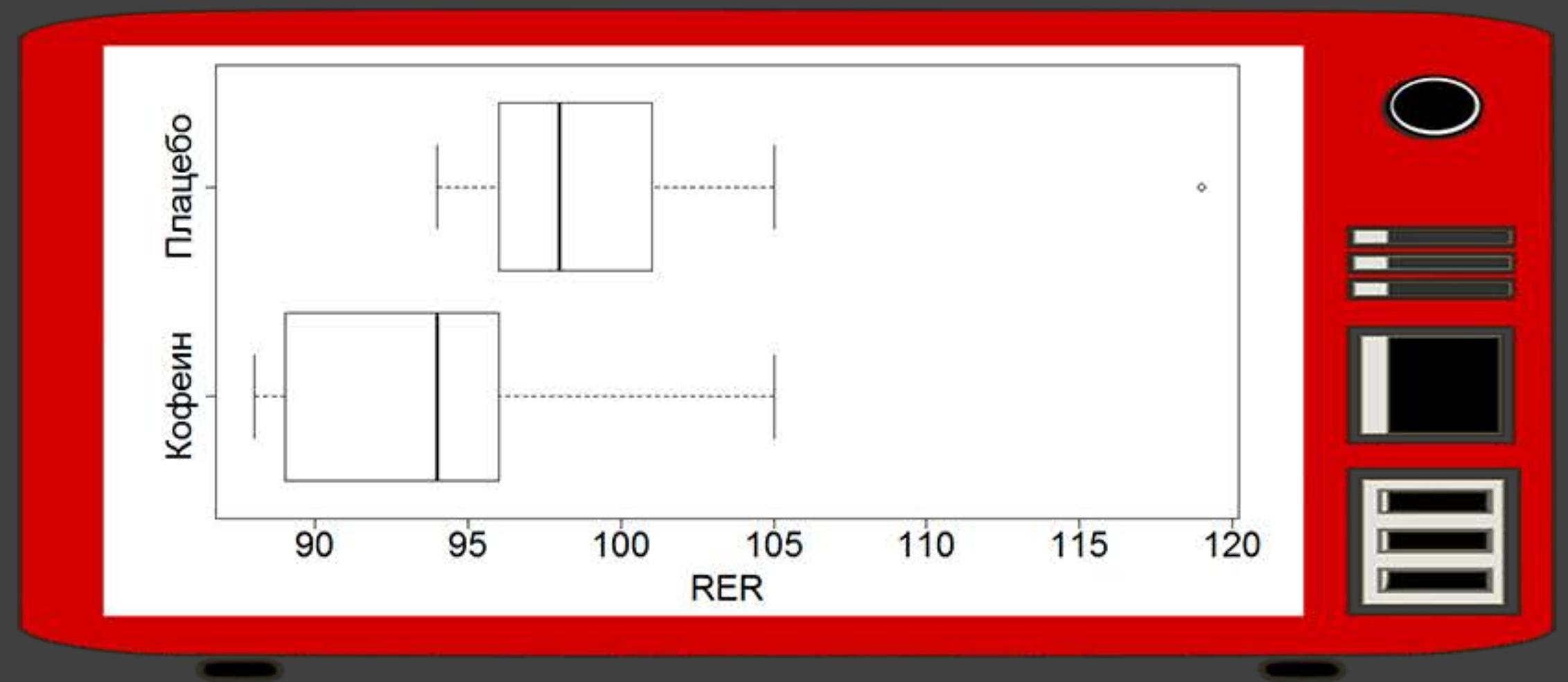
- » H_0 : депрессивность не изменилась,
 $\text{med}(X_2 - X_1) = 0$
- » H_1 : депрессивность снизилась, $\text{med}(X_2 - X_1) < 0$

ЛЕЧЕНИЕ ДЕПРЕССИИ

- » Критерий знаковых рангов: $p = 0.019$, медиана снижения — 0.49 пт (95% нижний доверительный предел — 0.175 пт)

КОФЕИН И РЕСПИРАТОРНЫЙ ОБМЕН

- › RER — соотношение числа молекул CO₂ и O₂ в выдыхаемом воздухе
- › В эксперименте измерялся респираторный обмен 18 испытуемых в процессе физических упражнений. За час до этого 9 из них получили таблетку кофеина, 9 — плацебо
- › Повлиял ли кофеин на значение RER?



КРИТЕРИЙ МАННА-УИТНИ

выборки: $X_1^{n_1} = (X_{11}, \dots, X_{1n_1})$

$X_2^{n_2} = (X_{21}, \dots, X_{2n_2})$

нулевая гипотеза: $H_0: F_{X_1}(x) = F_{X_2}(x)$

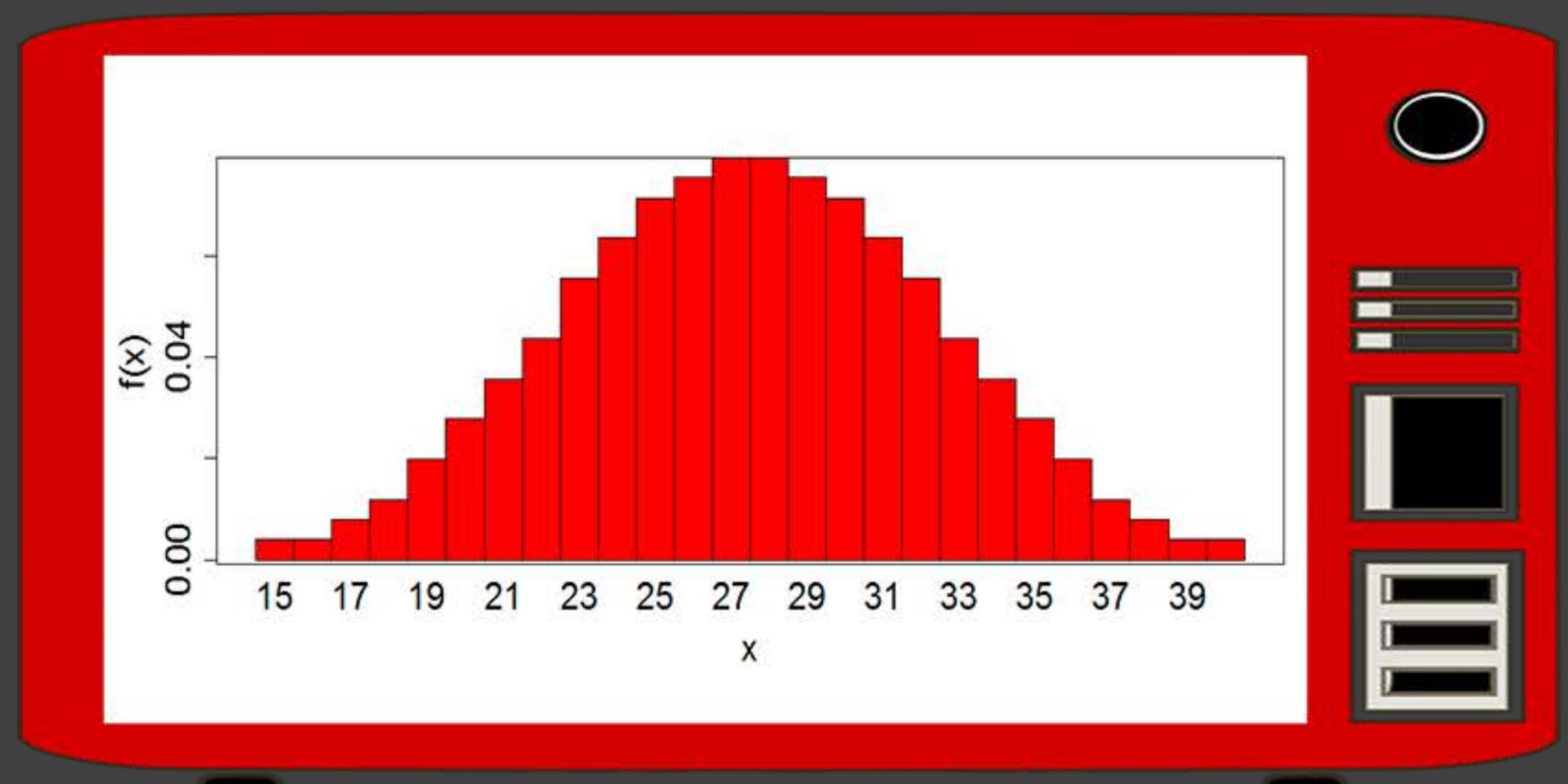
альтернатива: $H_1: F_{X_1}(x) = F_{X_2}(x + \Delta), \Delta < \neq 0$

статистика: $X_{(1)} \leq \dots \leq X_{(n_1+n_2)}$ — вариационный ряд

объединённой выборки $X = X_1^{n_1} \cup X_2^{n_2}$

$$R_1(X_1^{n_1}, X_2^{n_2}) = \sum_{i=1}^{n_1} \text{rank}(X_{1i})$$

нулевое распределение: табличное

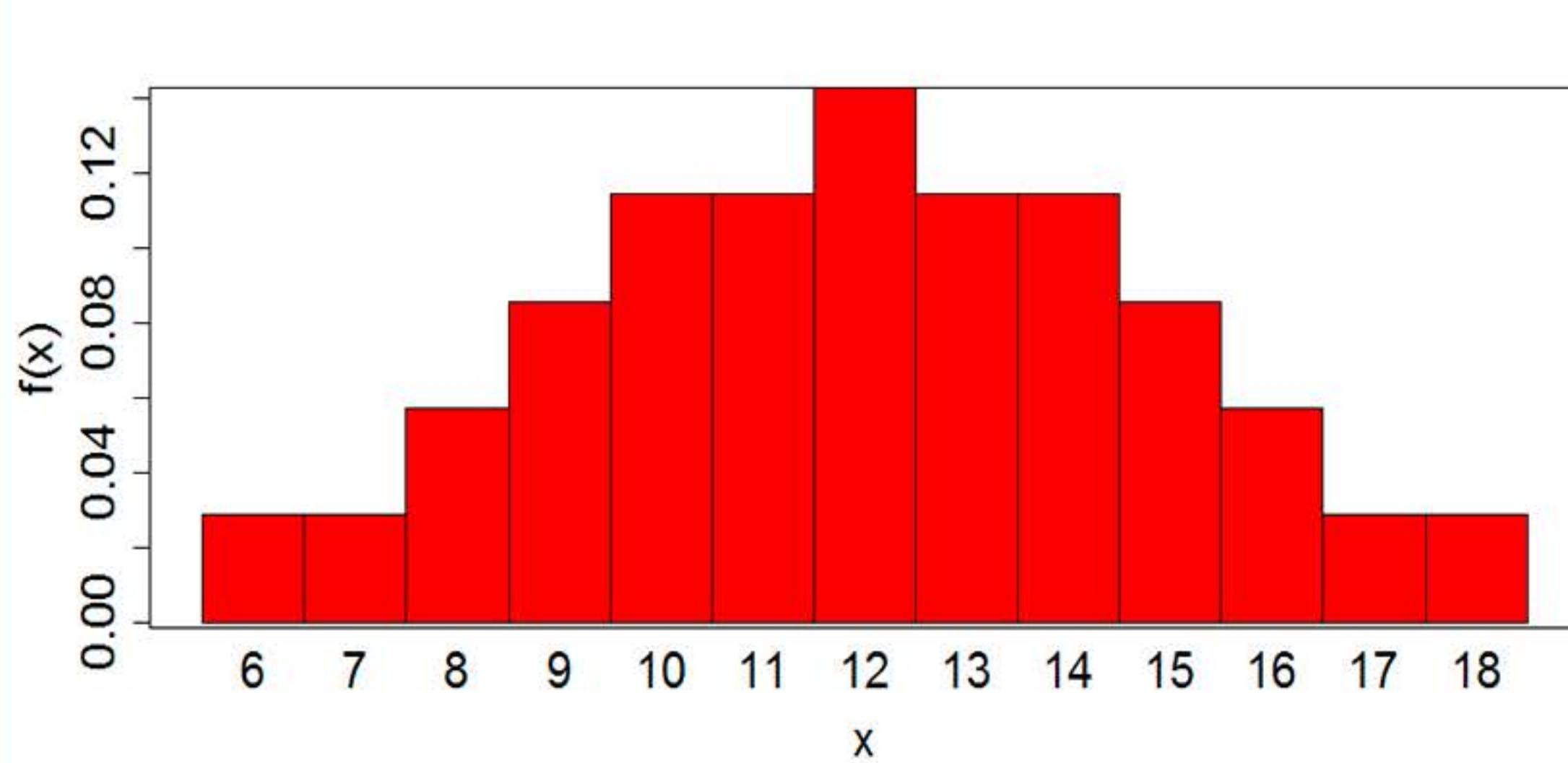


НУЛЕВОЕ РАСПРЕДЕЛЕНИЕ

X_1	X_2	R_1	
{1,2,3}	{4,5,6,7}	6	Всего $C_{n_1+n_2}^{n_1}$ вариантов
{1,2,4}	{3,5,6,7}	7	
{1,2,5}	{3,4,6,7}	8	
...	
{3,5,7}	{1,2,4,6}	15	
{3,6,7}	{1,2,4,5}	16	
{4,5,6}	{1,2,3,7}	15	
{4,5,7}	{1,2,3,6}	16	

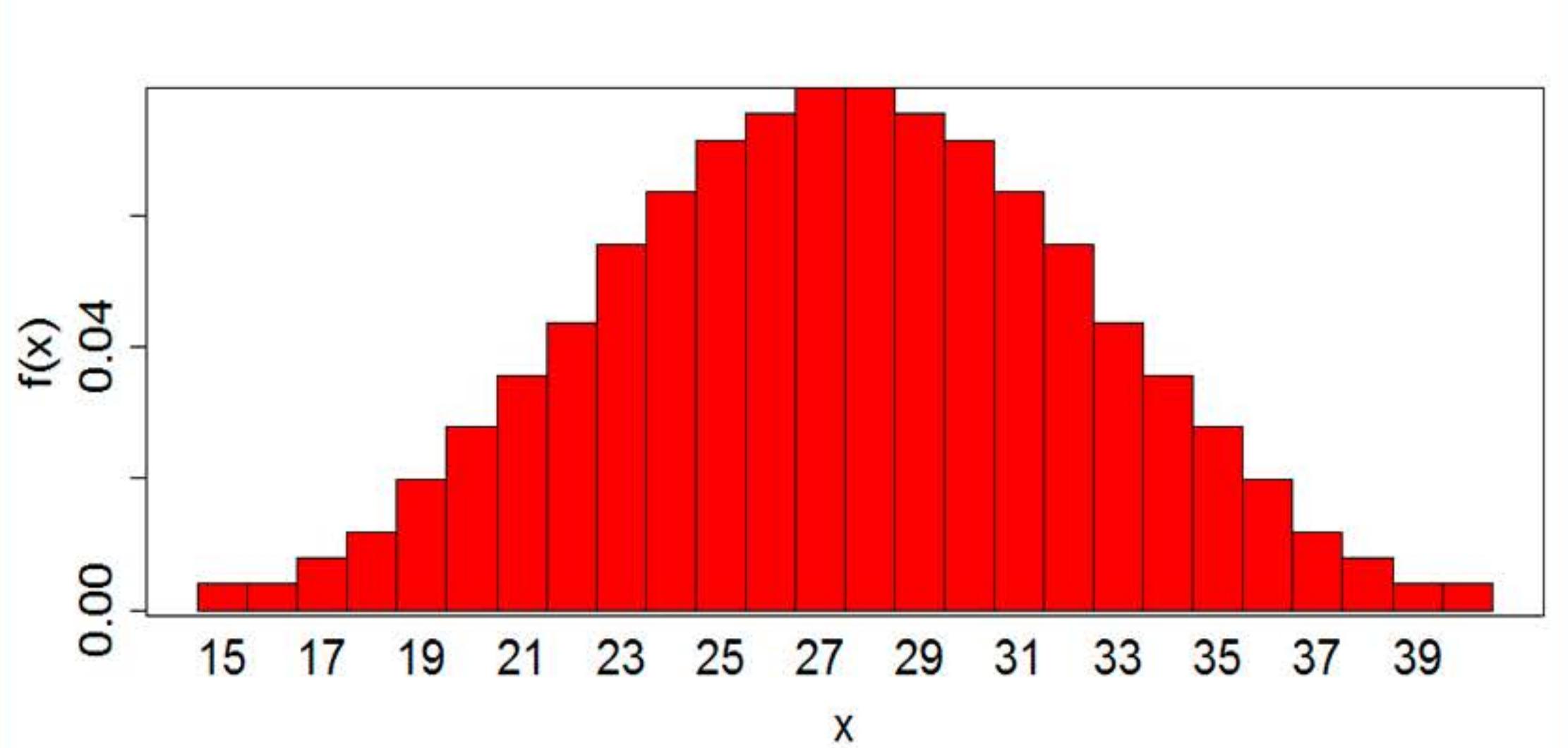
НУЛЕВОЕ РАСПРЕДЕЛЕНИЕ

» $n_1 = 3, \quad n_2 = 4$



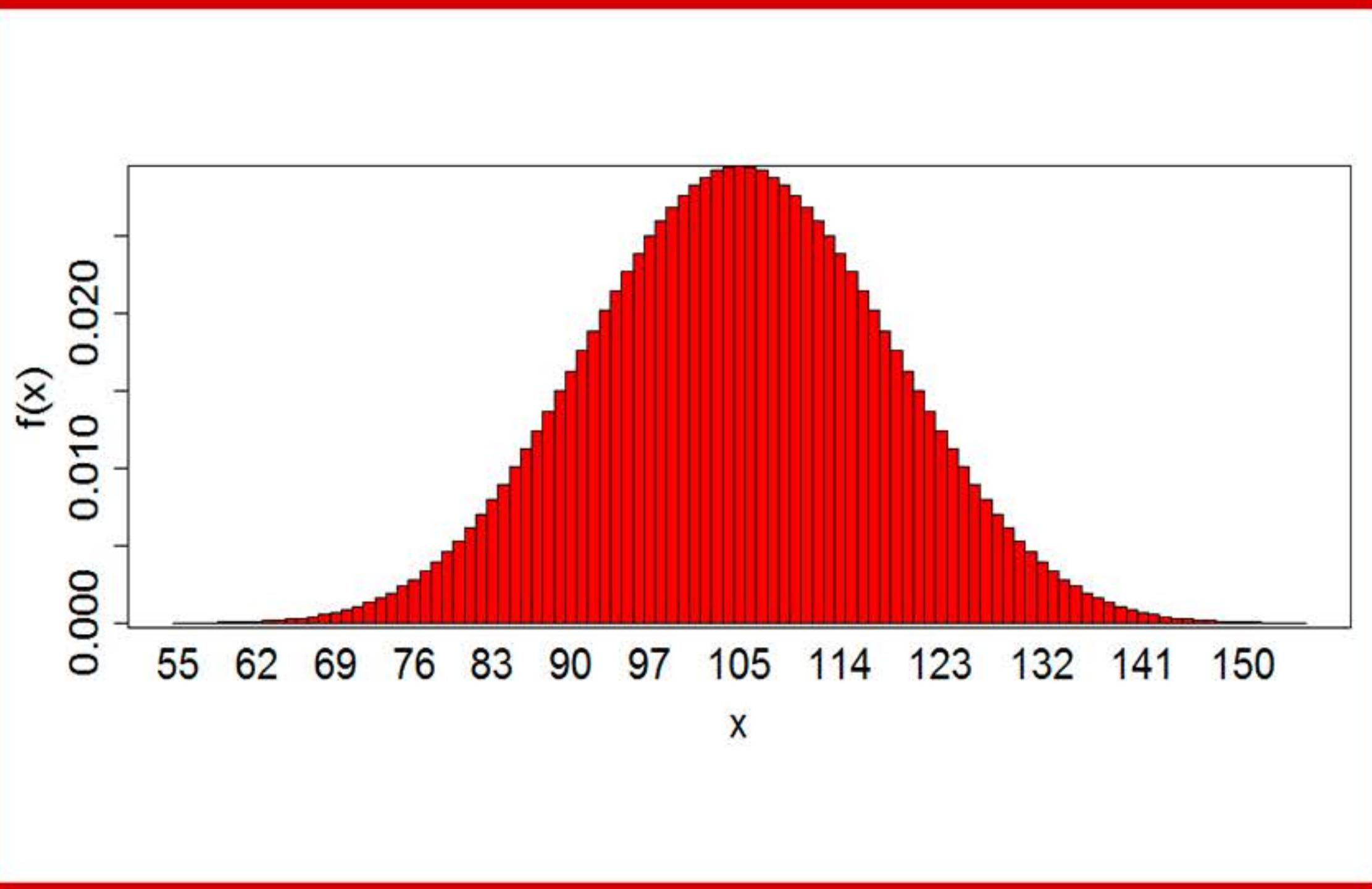
НУЛЕВОЕ РАСПРЕДЕЛЕНИЕ

» $n_1 = n_2 = 5$



НУЛЕВОЕ РАСПРЕДЕЛЕНИЕ

» $n_1 = n_2 = 10$



НУЛЕВОЕ РАСПРЕДЕЛЕНИЕ

» $n_1 = n_2 = 10$

Аппроксимация для $n_1, n_2 > 10$:

$$R_1 \sim N\left(\frac{n_1(n_1 + n_2 + 1)}{2}, \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}\right)$$

КОФЕИН И РЕСПИРАТОРНЫЙ ОБМЕН

- » H_0 : среднее значение показателя респираторного обмена не отличается в двух группах
- » H_1 : среднее значение показателя респираторного обмена отличается в двух группах
- » Критерий Манна-Уитни: $p = 0.0521$, сдвиг между средними — 6 пунктов, (95% доверительный интервал — [-0.00005, 12] пт)

- › Выборки из неизвестных распределений можно превращать в ранги, но для построения критериев нужны дополнительные предположения
- › Критерии знаковых рангов, одновыборочный и для связанных выборок
- › Критерий Манна-Уитни для независимых выборок

- › Выборки из неизвестных распределений можно превращать в ранги, но для построения критериев нужны дополнительные предположения
- › Критерии знаковых рангов, одновыборочный и для связанных выборок
- › Критерий Манна-Уитни для независимых выборок

- › Далее: перестановочные критерии

ПЕРЕСТАНОВОЧНЫЕ КРИТЕРИИ

ПЕРЕСТАНОВОЧНЫЕ КРИТЕРИИ

- » Ранговые критерии:
 - ▶ Выборки \Rightarrow ранги
 - ▶ Дополнительное предположение
 - ▶ Перестановки \Rightarrow нулевое распределение статистики

- » Что если пропустить первый пункт?

ОДНОВЫБОРОЧНЫЙ КРИТЕРИЙ

выборка: $X^n = (X_1, \dots, X_n)$

$F(X)$ симметрично относительно
матожидания

нулевая гипотеза: $H_0: \mathbb{E}X = m_0$

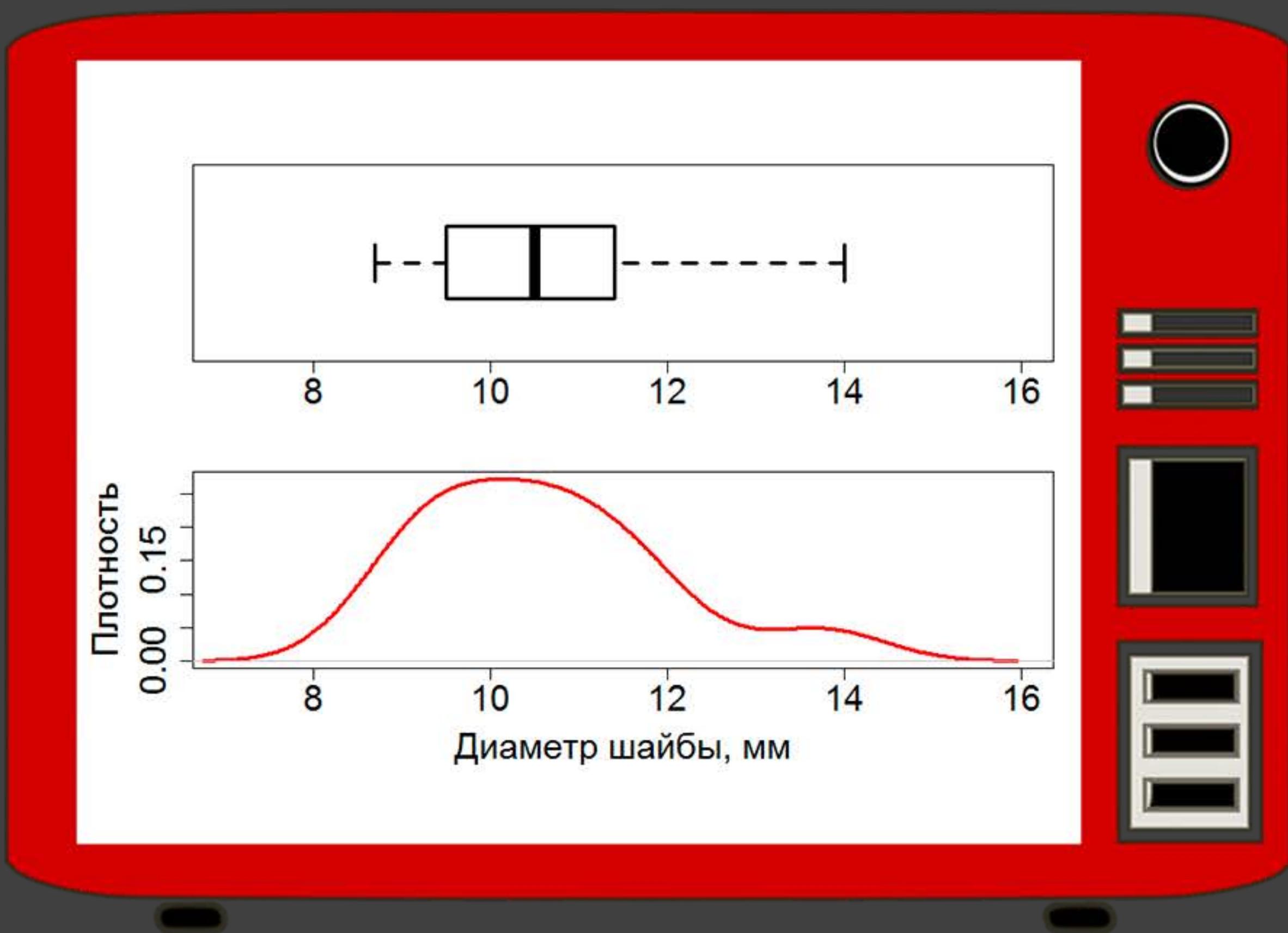
альтернатива: $H_1: \mathbb{E}X <\neq> m_0$

статистика: $T(X^n) = \sum_{i=1}^n (X_i - m_0)$

нулевое распределение: порождается перебором 2^n знаков
перед слагаемыми $X_i - m_0$

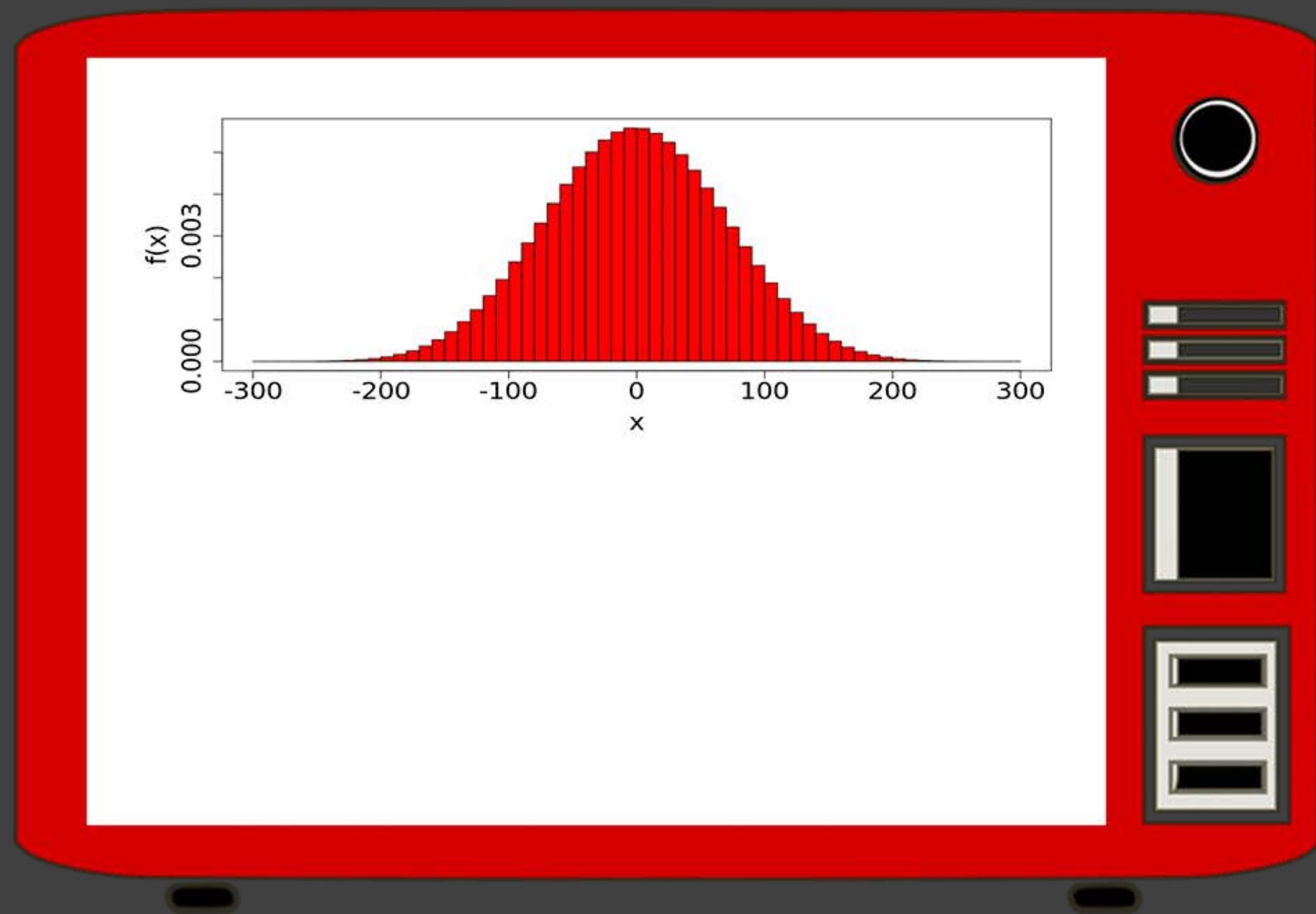
ДИАМЕТР ШАЙБЫ

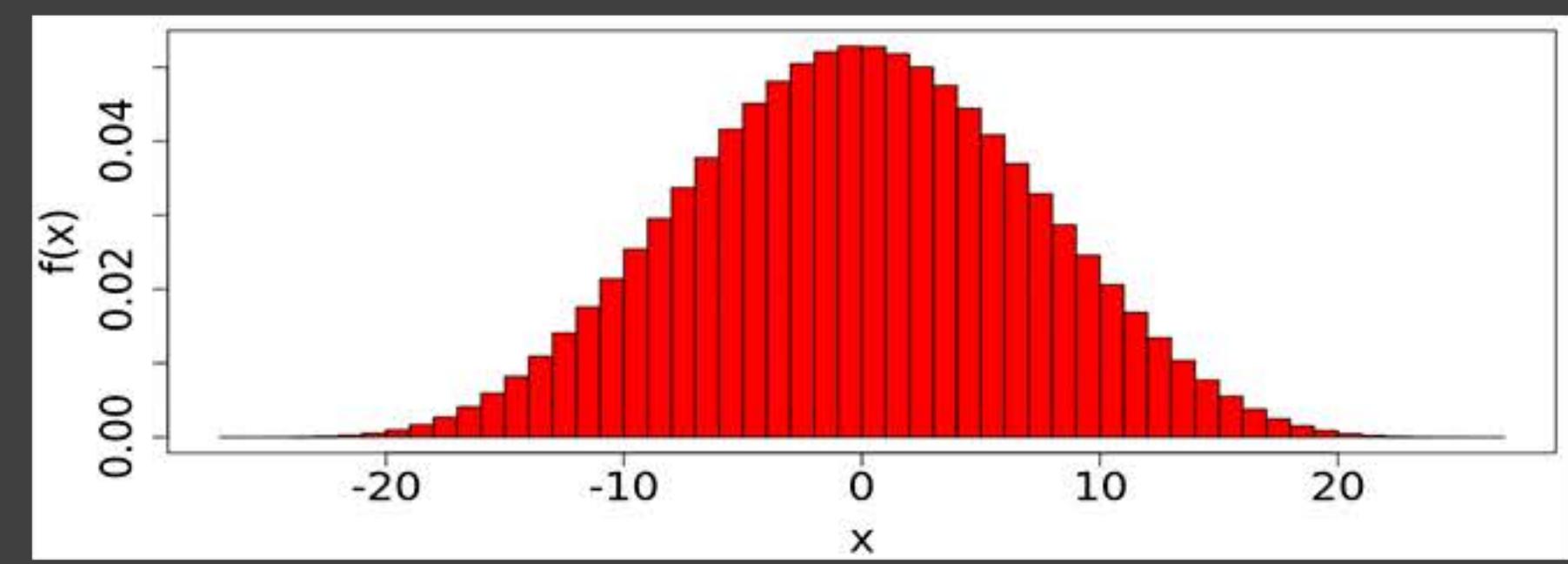
- » H_0 : средний диаметр шайбы — 10 мм, $\mathbb{E}X = 10$
- » H_1 : средний диаметр шайбы не соответствует стандарту, $\mathbb{E}X \neq 10$



ДИАМЕТР ШАЙБЫ

- » Критерий знаковых рангов: $p = 0.0673$
- » Перестановочный критерий: $T = 14.6, p = 0.1026$





ДЛЯ СВЯЗАННЫХ ВЫБОРОК

выборки: $X_1^n = (X_{11}, \dots, X_{1n})$

$X_2^n = (X_{21}, \dots, X_{2n})$

выборки связанные

нулевая гипотеза: $H_0: \mathbb{E}(X_1 - X_2) = 0$

альтернатива: $H_1: \mathbb{E}(X_1 - X_2) < \neq > 0$

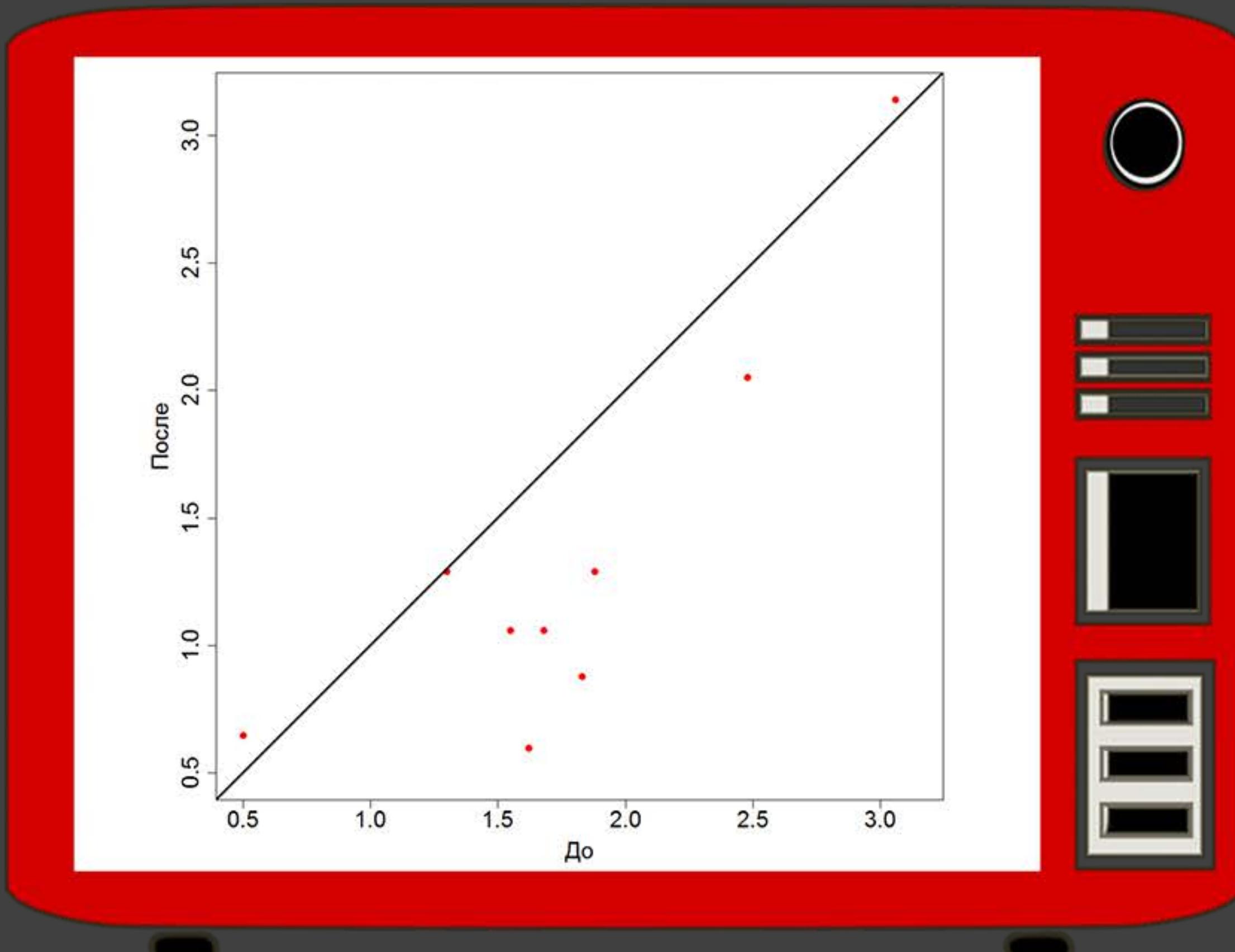
статистика: $D^n = (X_{1i} - X_{2i})$

$$T(X_1^n, X_2^n) = T(D^n) = \sum_{i=1}^n D_i$$

нулевое распределение: порождается перебором 2^n знаков перед слагаемыми D_i

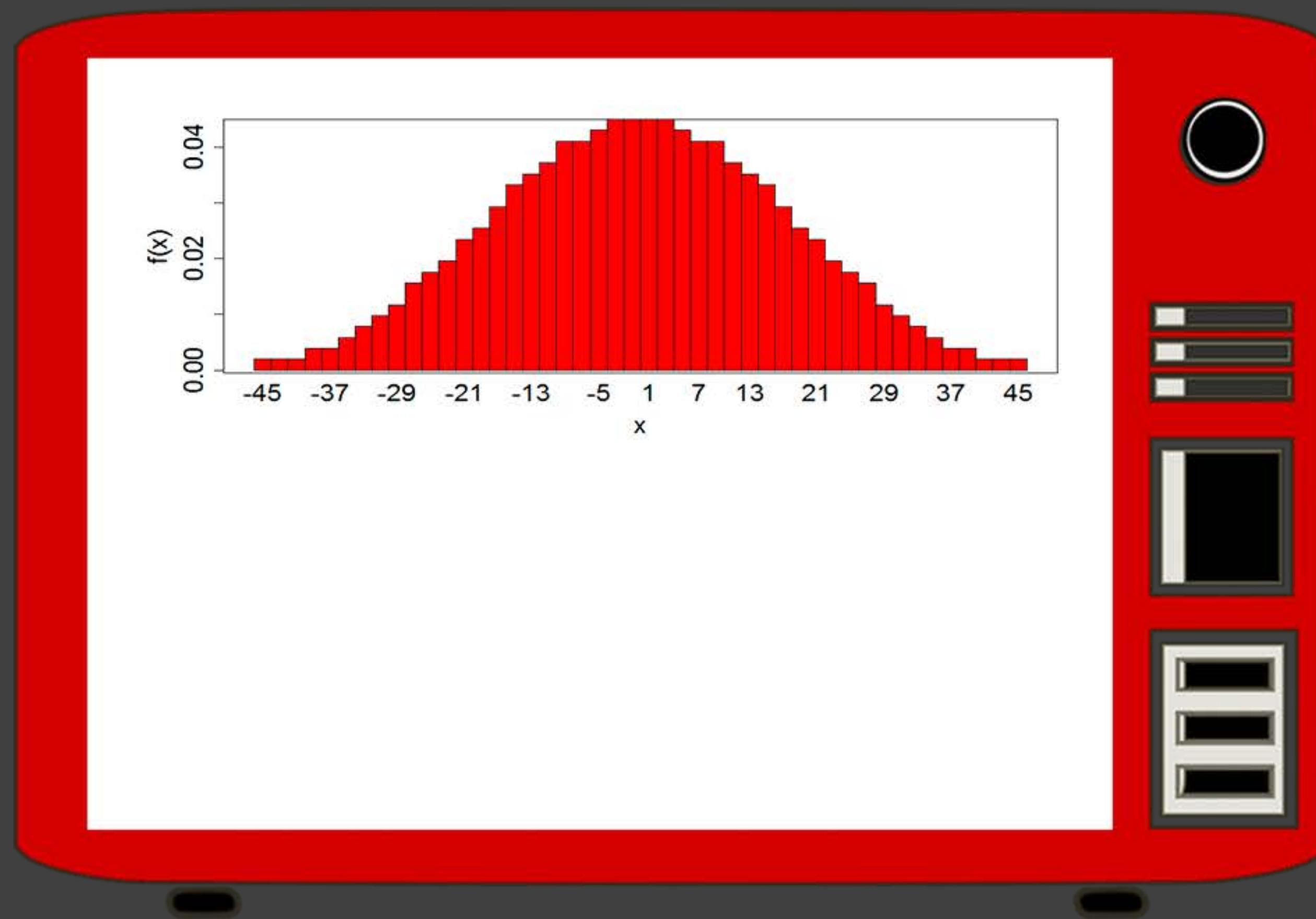
ЛЕЧЕНИЕ ДЕПРЕССИИ

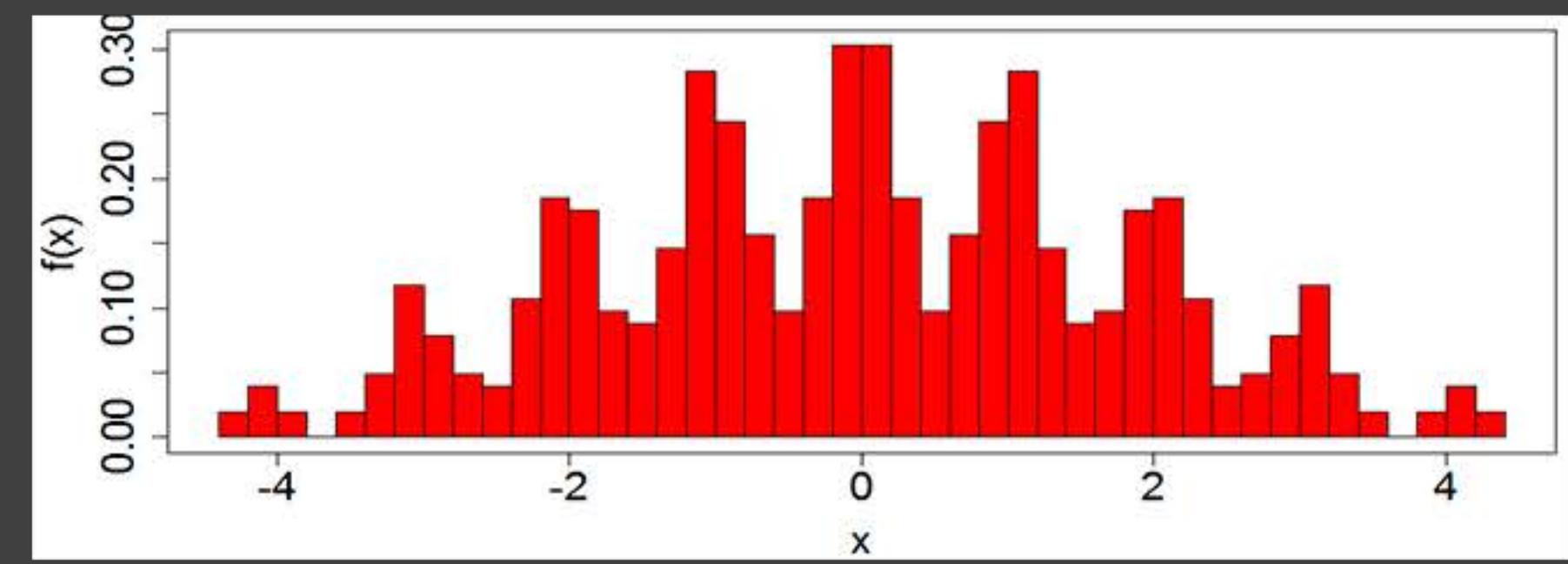
- » H_0 : депрессивность не изменилась,
 $\mathbb{E}(X_1 - X_2) = 0$
- » H_1 : депрессивность снизилась, $\mathbb{E}(X_1 - X_2) > 0$



ЛЕЧЕНИЕ ДЕПРЕССИИ

- » Критерий знаковых рангов: $p = 0.019$
- » Перестановочный критерий:
 $T = 3.887, p = 0.0137$





ДЛЯ НЕЗАВИСИМЫХ ВЫБОРОК

выборки: $X_1^{n_1} = (X_{11}, \dots, X_{1n_1})$

$X_2^{n_2} = (X_{21}, \dots, X_{2n_2})$

нулевая гипотеза: $H_0: F_{X_1}(x) = F_{X_2}(x)$

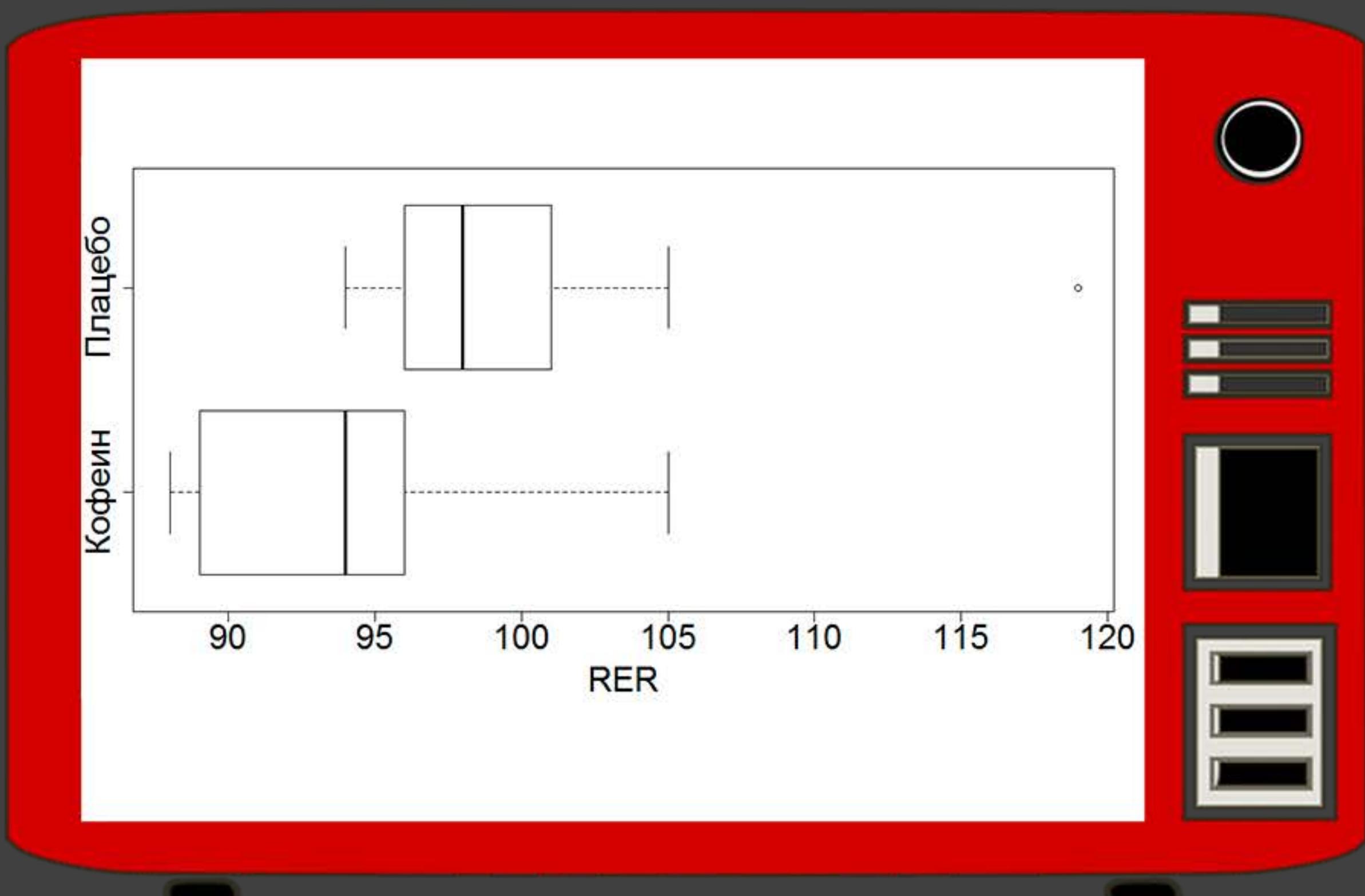
альтернатива: $H_1: F_{X_1}(x) = F_{X_2}(x + \Delta), \Delta <\neq> 0$

статистика: $T(X_1^{n_1}, X_2^{n_2}) = \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1i} - \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2i}$

нулевое распределение: порождается перебором $C_{n_1+n_2}^{n_1}$ размещений объединённой выборки.

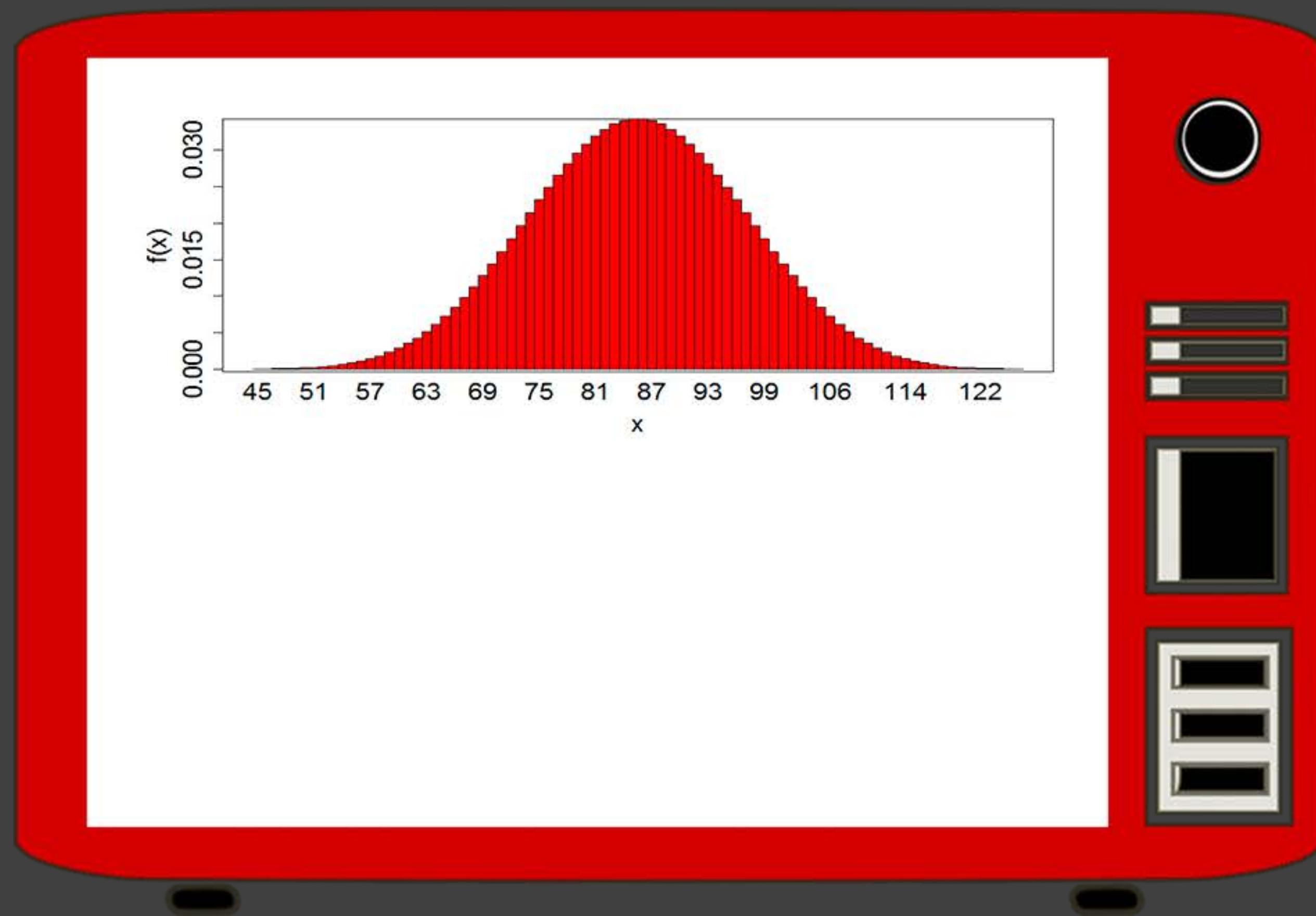
КОФЕИН И РЕСПИРАТОРНЫЙ ОБМЕН

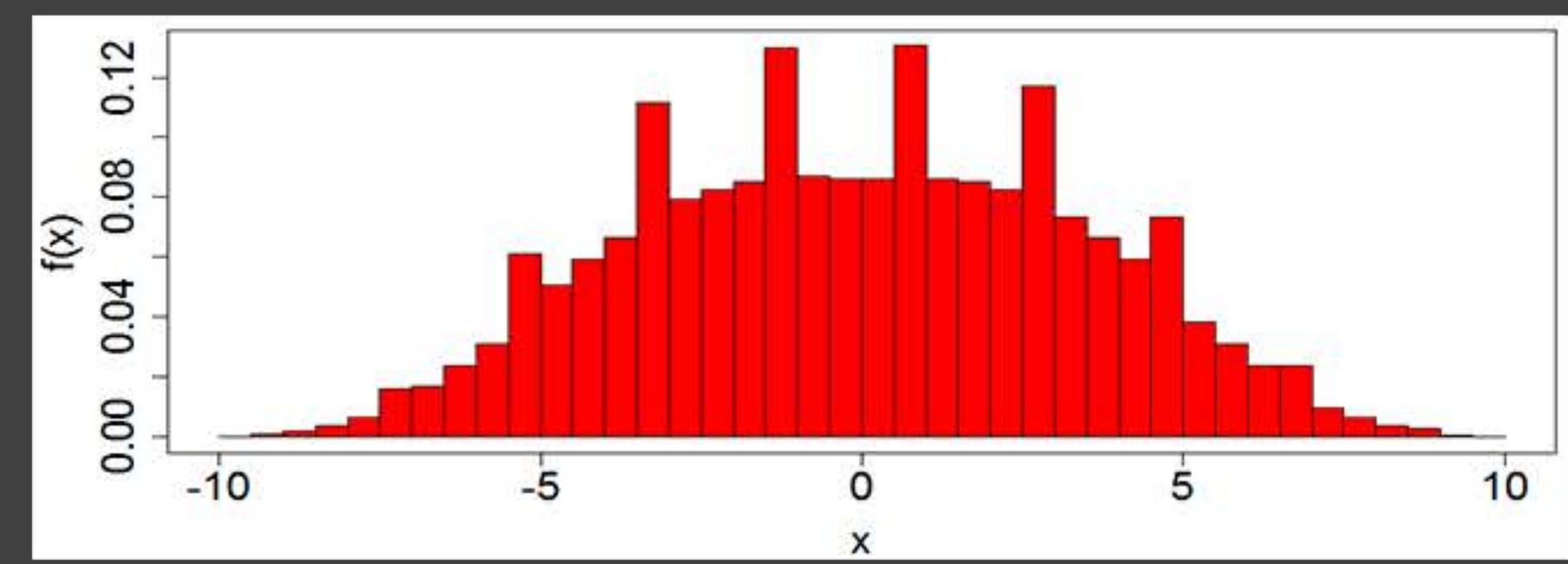
- › H_0 : среднее значение показателя респираторного обмена не отличается в двух группах
- › H_1 : среднее значение показателя респираторного обмена отличается в двух группах



КОФЕИН И РЕСПИРАТОРНЫЙ ОБМЕН

- » Критерий Манна-Уитни: $p = 0.0521$
- » Перестановочный критерий:
 $T = 6.33, p = 0.0578$





ОСОБЕННОСТИ ПЕРЕСТАНОВОЧНЫХ КРИТЕРИЕВ

- » Статистику можно выбрать разными способами
- » В некоторых случаях разные статистики приведут к одному и тому же достигаемому уровню значимости:

$$X^n, H_0: \mathbb{E}X = 0, H_1: \mathbb{E}X \neq 0$$

$$T_1(X^n) = \sum_{i=1}^n X_i \sim T_2(X^n) = \bar{X}$$

ОСОБЕННОСТИ ПЕРЕСТАНОВОЧНЫХ КРИТЕРИЕВ

- » Статистику можно выбрать разными способами
- » В других случаях достигаемый уровень значимости будет зависеть от выбора статистики:

$$T_2(X^n) = \bar{X} \quad \propto \quad T_3(X^n) = \frac{\bar{X}}{S/\sqrt{n}}$$

ОСОБЕННОСТИ ПЕРЕСТАНОВОЧНЫХ КРИТЕРИЕВ

- » Если множество всех перестановок G слишком велико, для оценки нулевого распределения T достаточно взять случайное подмножество G' . При этом стандартное отклонение достигаемого уровня значимости будет равно примерно

$$\sqrt{\frac{p(1-p)}{|G'|}}$$

РЕЗЮМЕ



- › В тех же предположениях, что используют ранговые критерии, можно учитывать больше информации с помощью перестановочных

- › Далее: связь с бутстрепом

ПЕРЕСТАНОВКИ И БУТСТРЕП

- » Перестановочные критерии:
 - ▶ Выборки, статистика
 - ▶ Дополнительное предположение
 - ▶ Перестановки \Rightarrow нулевое распределение статистики
- » Бутстреповые доверительные интервалы:
 - ▶ Выборки, статистика, оценивающая параметр
 - ▶ Бутстреп-псевдовыборки \Rightarrow приближённое распределение статистики

ПЕРЕСТАНОВКИ И БУТСТРЕП

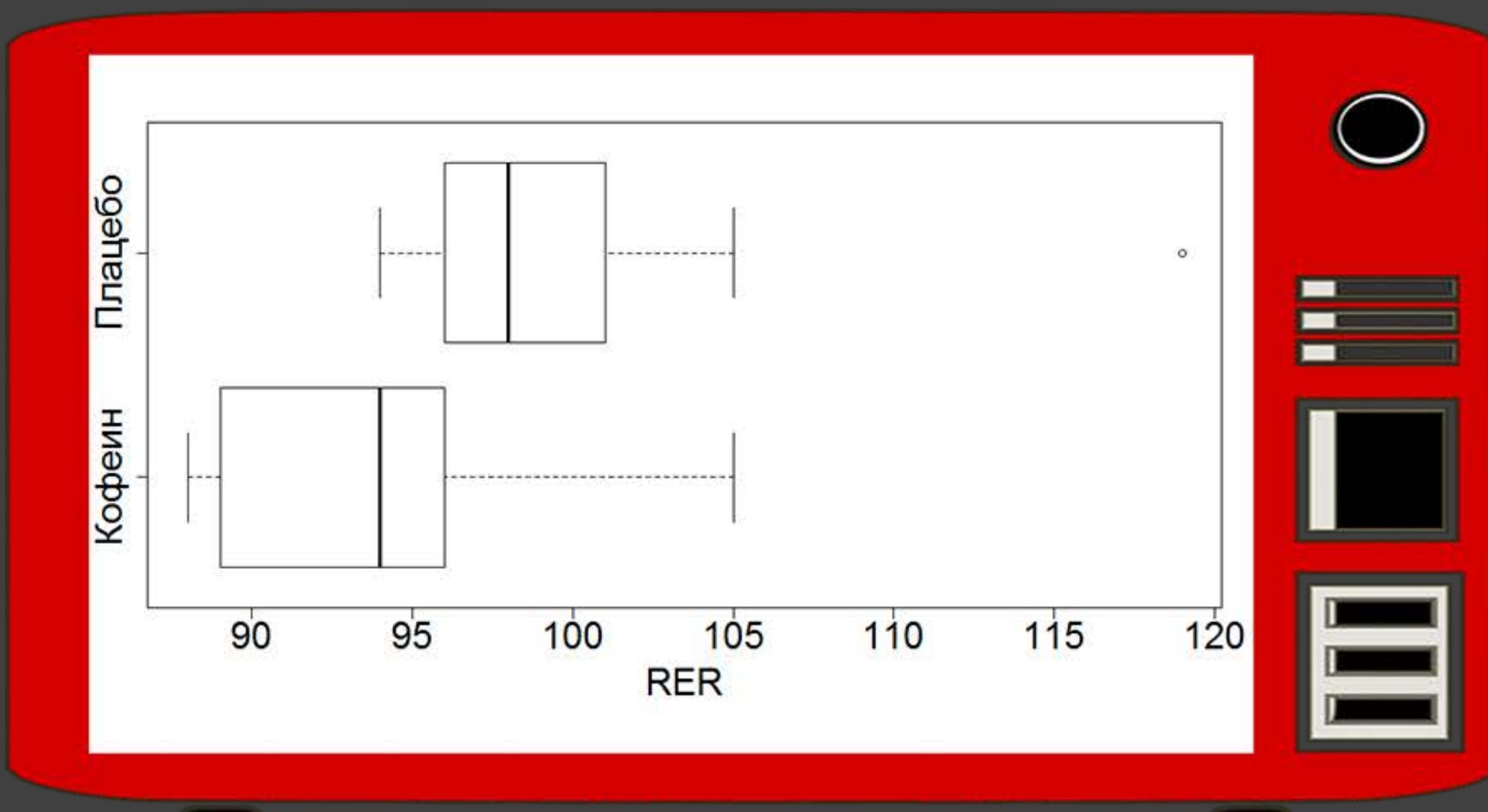
- Перестановочные критерии:
 - ▶ Выборки, статистика
 - ▶ Дополнительное предположение
 - ▶ Перестановки \Rightarrow нулевое распределение статистики

- » Бутстреповые доверительные интервалы:
 - ▶ Выборки, статистика, оценивающая параметр
 - ▶ Бутстреп-псевдовыборки \Rightarrow приближённое распределение статистики

КОФЕИН И РЕСПИРАТОРНЫЙ ОБМЕН

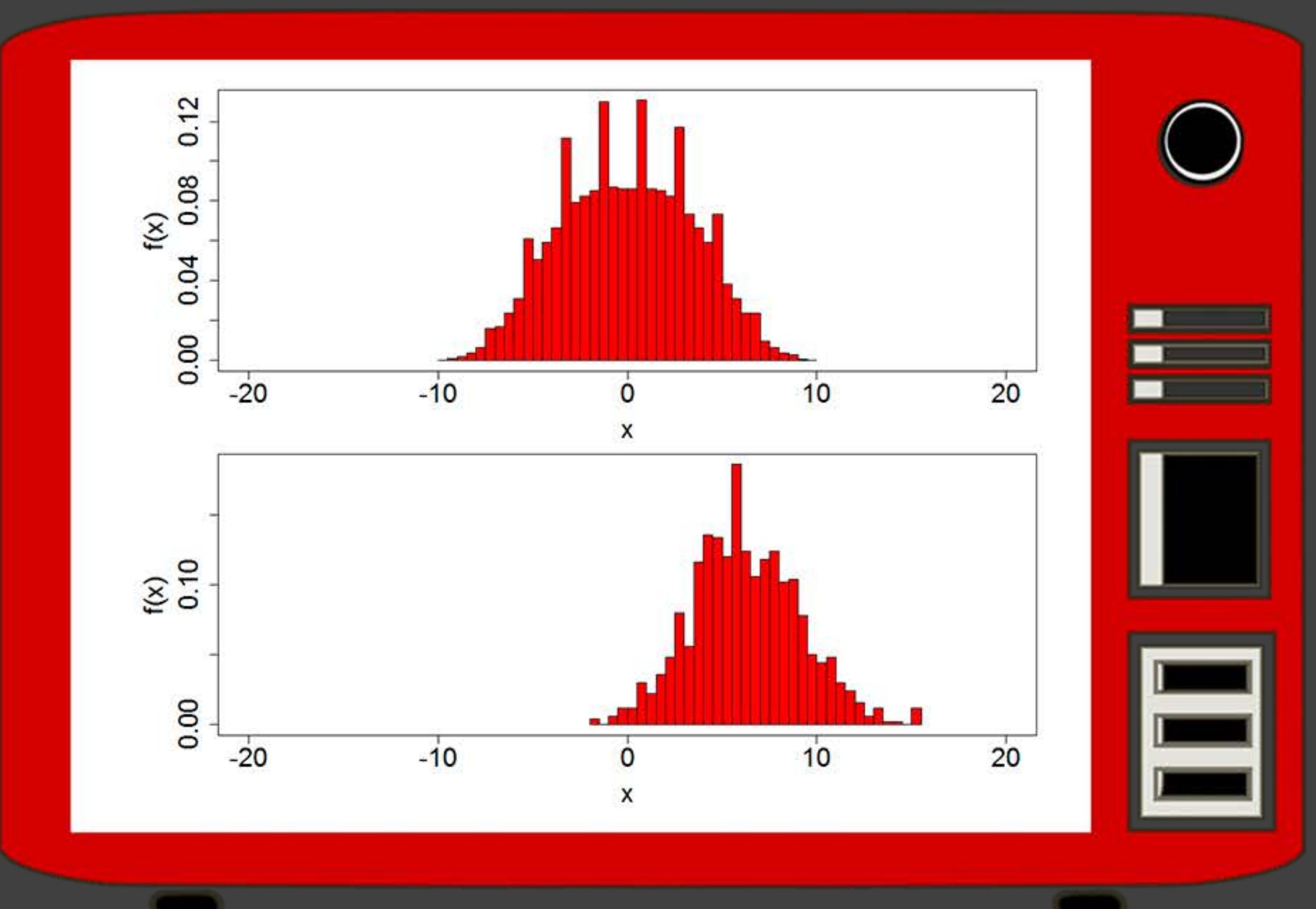
- » H_0 : среднее значение показателя респираторного обмена не отличается в двух группах
- » H_1 : под воздействием кофеина среднее значение показателя респираторного обмена снижается

$$\bar{X}_{1n} - \bar{X}_{2n} = 6.33$$



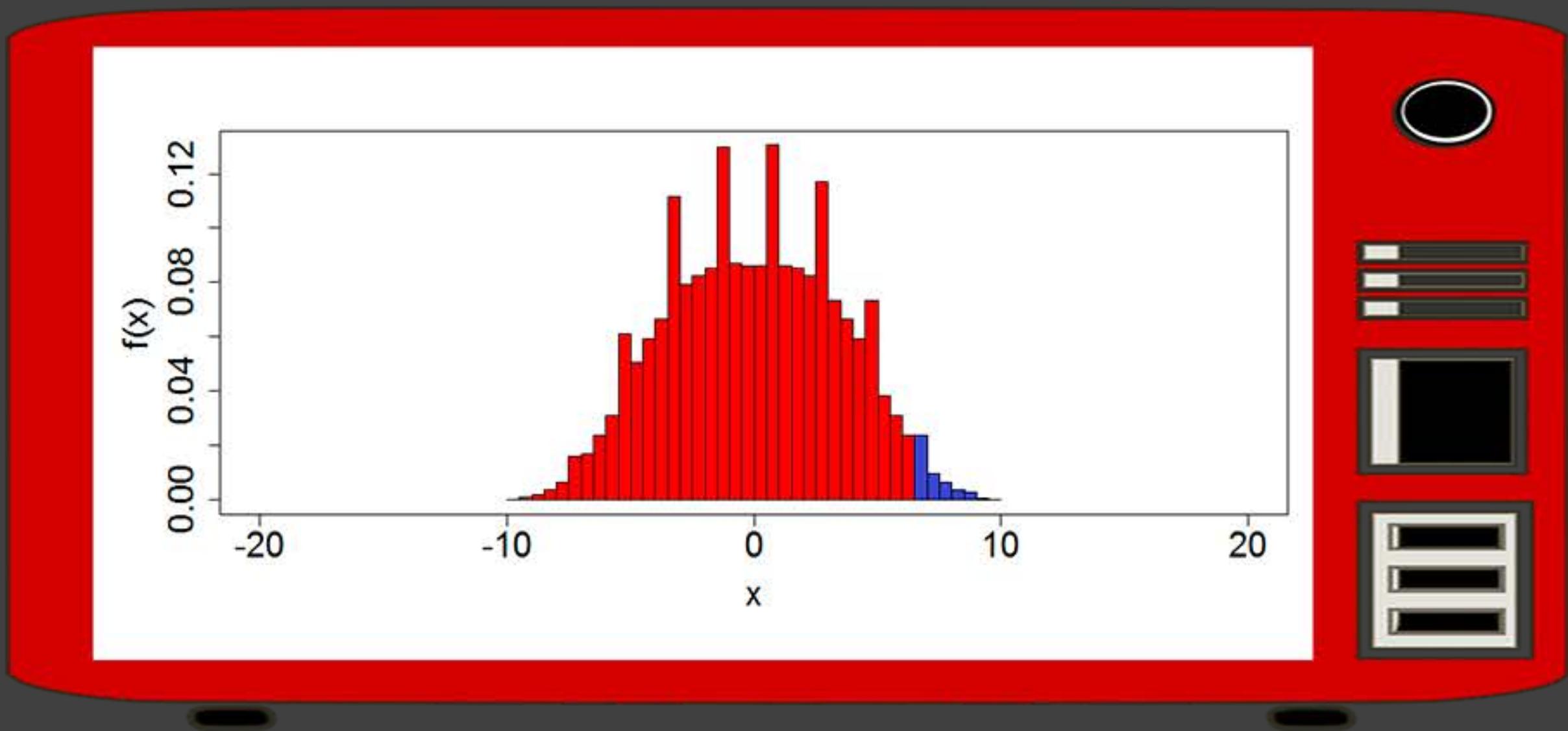
КОФЕИН И РЕСПИРАТОРНЫЙ ОБМЕН

- › Нулевое распределение перестановочного критерия со статистикой $\bar{X}_{1n} - \bar{X}_{2n}$
- › Бутстреп-распределение статистики $\bar{X}_{1n} - \bar{X}_{2n}$



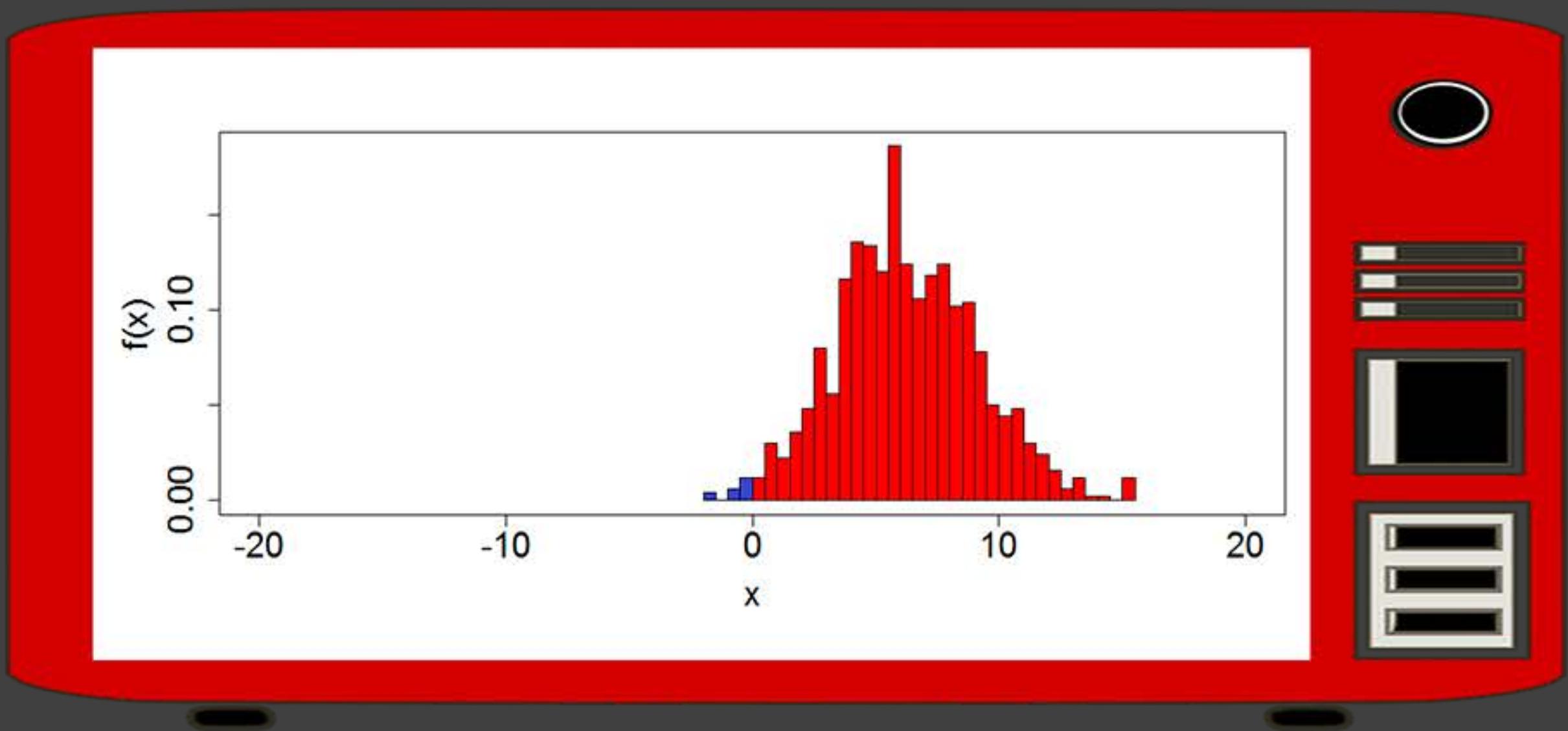
КОФЕИН И РЕСПИРАТОРНЫЙ ОБМЕН

- › Нулевое распределение перестановочного критерия со статистикой $\bar{X}_{1n} - \bar{X}_{2n}$
- › Доля перестановок, на которых среднее больше либо равно 6.33 — 0.0289
- › Это точный достигаемый уровень значимости перестановочного критерия



КОФЕИН И РЕСПИРАТОРНЫЙ ОБМЕН

- › Бутстреп-распределение статистики $\bar{X}_{1n} - \bar{X}_{2n}$
- › Доля псевдовыборок, на которых среднее меньше либо равно нулю — 0.011
- › Это приближённый достигаемый уровень значимости бутстреп-критерия



ПЕРЕСТАНОВКИ VS БУТСТРЕП

- › Перестановочный критерий измеряет расстояние от 0 до \bar{D}_n
- › Бутстреп-критерий измеряет расстояние от \bar{D}_n до 0

ПЕРЕСТАНОВКИ VS БУТСТРЕП



- › Перестановочный критерий более точный
- › Бутстреп-критерий приближённый

ПЕРЕСТАНОВКИ VS БУТСТРЕП

➤ Перестановочный критерий проверяет

$$H_0: F_{X_1}(x) = F_{X_2}(x) \text{ против}$$

$$H_1: F_{X_1}(x) = F_{X_2}(x + \Delta), \Delta > 0$$

➤ Бутстреп-критерий проверяет

$$H_0: \mathbb{E}X_1 = \mathbb{E}X_2 \text{ против}$$

$$H_1: \mathbb{E}X_1 > \mathbb{E}X_2$$

- › Бутстреп тоже позволяет проверять гипотезы
- › Перестановочные критерии точнее, а бутстреп более гибок
- › Далее: примеры

