# Теория вероятностей и математическая статистика

## Тестирование статистических гипотез II.

Глеб Карпов

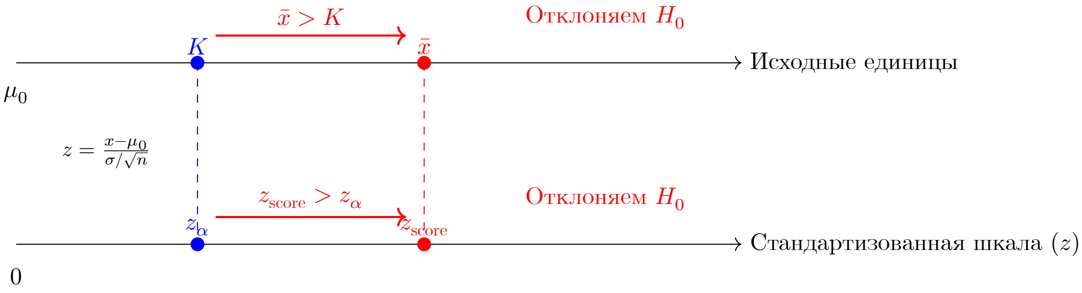
ВШБ Бизнес-информатика

# Test-score: два эквивалентных подхода

- В предыдущей лекции мы сравнивали наблюдаемое значение (например, $\bar{x}$ или $\tilde{p}$) с границей принятия решения $K$ в исходных единицах измерения.
- Альтернативный подход: преобразовать и наблюдаемое значение, и границу решения в стандартизованную шкалу.
- **Ключевая идея**: стандартизация — это монотонное преобразование, которое сохраняет отношения "больше/меньше" между числами.
- Тип score зависит от распределения:
  - $z_{\text{score}}$ — когда используем стандартное нормальное распределение (тестирование математических ожиданий при известной дисперсии или тест долей признаков)
  - $t_{\text{score}}$ — когда используем $t$-распределение (неизвестная дисперсия)
- Если наблюдаемое значение $> K$ в исходных единицах, то score $>$ критическое значение в стандартизованной шкале.
- Если наблюдаемое значение $< K$ в исходных единицах, то score $<$ критическое значение в стандартизованной шкале.

# Test-score: визуализация преобразования

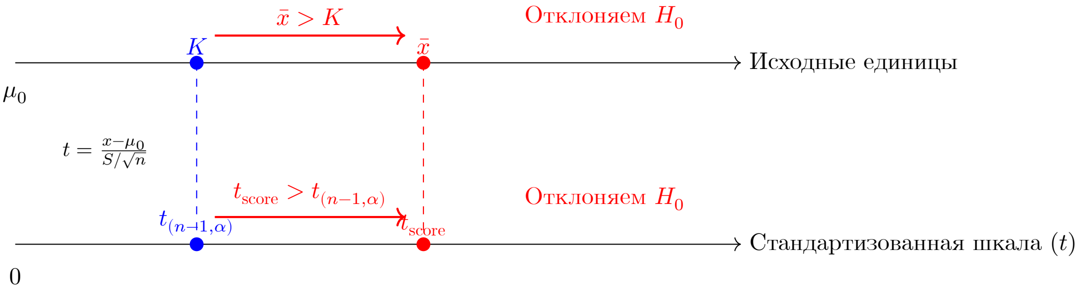
Пример: $z$-score для тестирования математического ожидания




Преобразование сохраняет отношение порядка. Решение одинаково в обеих шкалах.

# Test-score: визуализация преобразования

Пример: $t$-score для тестирования математического ожидания




Преобразование сохраняет отношение порядка. Решение одинаково в обеих шкалах.

# Test-score: преимущества подхода

- **Универсальность**: одна и та же критическая точка ($z_\alpha$ или $t_{df,\alpha}$) используется для всех тестов с соответствующим распределением.
- **Удобство**: не нужно вычислять границу решения $K$ для каждого конкретного случая.
- **Эквивалентность**: оба подхода (сравнение в исходных единицах и сравнение score) дают одинаковый результат, так как стандартизация — монотонное преобразование.
- **Гибкость**: подход работает для разных типов тестов (односторонние, двусторонние) и разных распределений.

# p-value: концепция

- **p-value** (p-значение) — это вероятность получить наблюдаемое значение статистики или более экстремальное значение *при условии, что нулевая гипотеза верна.*
- p-value показывает, насколько "необычны" полученные нами данные при предполагамой верной нулевой гипотезе.
- **Правило решения через p-value**: Отклонить $H_0$, если $p$-value $< \alpha$.
- Чем меньше p-value, тем сильнее свидетельство против нулевой гипотезы.
- p-value — это альтернативный способ принятия решения, эквивалентный сравнению score с критическим значением.

# p-value: пример

## Пример: Анализ длины шеи жирафа

Исследовательская команда изучает популяции жирафов. Исследователи верят, что средняя длина шеи жирафов огромная — около 20 метров. Однако в выборке было получено выборочное среднее 2.8 метров. Команда хочет проверить, действительно ли средняя длина шеи меньше того значения, в которое они верят.

- Случайная величина для анализа: $X$ — длина шеи. Дисперсию $X$ предположим известной и правдивой.
- Предполагается, что длины шеи следуют нормальному распределению.

Данные:

- Выборка: $n = 20$ жирафов
- Выборочное среднее: $\bar{x} = 2.8$ метров
- Известная дисперсия $X$: $\mathrm{Var}\left[X\right] = 0.16$.

Гипотезы:

- $H_0 : \mu = 20$
- $H_1 : \mu < 20$

## p-value: решение через score

- **Гипотезы**: $H_0 : \mu = 20$ против $H_1 : \mu < 20$ (левосторонний тест)
- **Данные**: $\mu_0 = 20$, $\sigma = 0.4$, $n = 20$, $\bar{x} = 2.8$, $\alpha = 0.05$
- ***z*-score**:

$$z_{\text{score}} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{2.8 - 20}{0.4/\sqrt{20}} = \frac{-17.2}{0.0894} \approx -192.4$$

- **Критическое значение**: $z_{0.05} = 1.645$
- **Решение через score**: $z_{\text{score}} = -192.4 < -z_{0.05} = -1.645$, поэтому отклоняем $H_0$.

# p-value: вычисление и интерпретация

- **p-value** для левостороннего теста:

$$p\text{-value} = P_{H_0}(Z < z_{\text{score}}) = P(Z < -192.4) \approx 0$$

- p-value практически равна нулю, что означает крайне малую вероятность получить такое экстремальное значение при верной нулевой гипотезе.
- **Решение через p-value**: $p$-value $\approx 0 < \alpha = 0.05$, поэтому отклоняем $H_0$.
- **Вывод**: Средняя длина шеи жирафов статистически значимо меньше 20 метров. Наблюдаемое значение 2.8 метра крайне противоречит гипотезе о том, что средняя длина равна 20 метрам.
- **Эквивалентность**: Оба подхода (сравнение $z_{\text{score}}$ с $z_\alpha$ и сравнение $p$-value с $\alpha$) дают одинаковый результат.