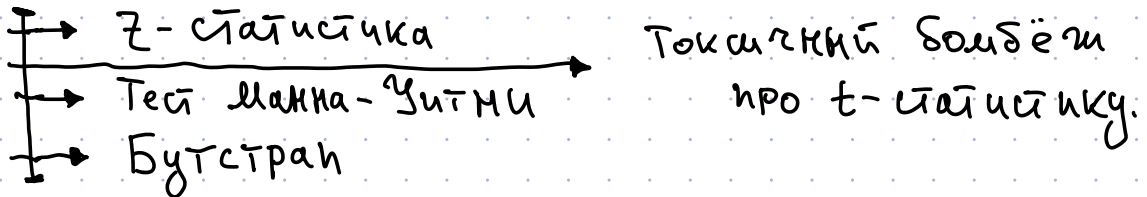


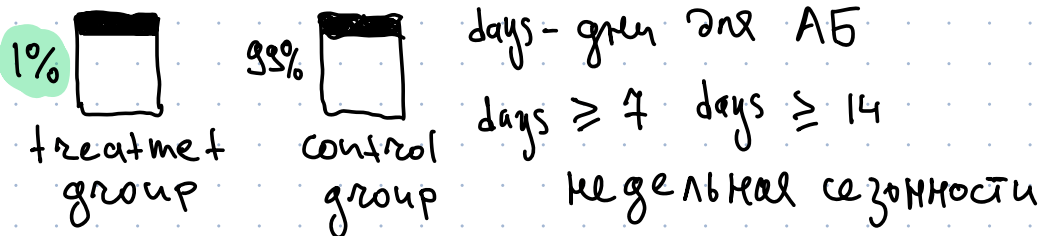
АБ-тесты

- Планирование АБ-теста
- Основные критерии из прода:



- Оффлайн истории

① Планирование АБ-теста



$X_1, \dots, X_{n_T} \sim iid$ $Y_1, \dots, Y_{n_C} \sim iid$
 $n_T - ?$ $n_T < n_C$

$$days = \frac{n_T}{DAU \cdot 0.05}$$

Параметры:

α - ош. 1 рода
 β - ош. 2 рода

	H_0	H_A
exp. H_0	True	β
exp. H_A	α	True

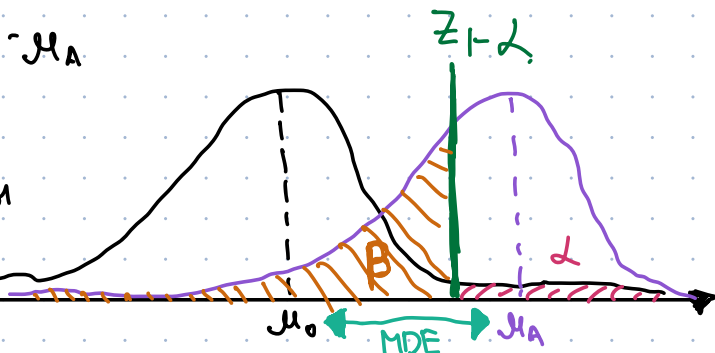
$1 - \beta$
 Recall

α
 FPR

MDE - minimal detectable effect
 $\mu_0 - \mu_A$

σ_T^2, σ_C^2 - known

$$n_T = f(\alpha, \beta, MDE, \sigma_C^2, \sigma_T^2)$$



② Z-Тест

$$H_0: \mu_0 = \mu_A$$

$$H_A: \mu_0 < \mu_A$$

$>$
 \neq

$$\begin{matrix} Z_{obs} & \text{vs} & Z_{1-\alpha} \\ p\text{-val} & \text{vs} & \alpha \end{matrix}$$

Число наблюдений:

$$n = \frac{\hat{\sigma}_c^2 + \hat{\sigma}_T^2}{0.5 \cdot MDE} \cdot (Z_{1-\alpha} + Z_{1-\beta})^2$$

$$n_T = 0.01 \cdot n$$



$$\alpha \uparrow \Rightarrow Z_{1-\alpha} \downarrow \Rightarrow n \downarrow$$

Откуда взять $\hat{\sigma}_T^2$, $\hat{\sigma}_c^2$?

1) Вариация: неделя/день/час

$$\hat{\sigma}_T^2, \hat{\sigma}_c^2$$

MDE

можно и по истории данных

2) $n_T = \dots$ days = ...

3) Тестовая выборка ка days.

4) Оборот теста

→ нез. набл. и выборки

NPS

0 1 2 3 4 5 6 7 8 9 10

||
P-d

d

p

выб. P. зав. от
выб. ↓

→ большое n (big data)

→ нет вопросов

$$\bar{X}_T - \bar{X}_c \underset{H_0}{\sim}^{asy} \mathcal{N}\left(0, \frac{\hat{\sigma}_T^2}{n_T} + \frac{\hat{\sigma}_c^2}{n_c}\right)$$

$$Z = \frac{\bar{X}_T - \bar{X}_c}{\sqrt{\frac{\hat{\sigma}_T^2}{n_T} + \frac{\hat{\sigma}_c^2}{n_c}}} \underset{H_0}{\sim}^{asy} \mathcal{N}(0, 1)$$

Таксисты "допомогательства" → Санюрт
 ШАЛЬНАЯ ИМПЕРАТРИЦА

"Обучение"

$P \downarrow$

$n = \frac{\text{маленький}}{\text{большой}}$

$\Rightarrow n_T$ большая

days — несколько лет

③ Токсичный боаёт про t -статистику

t -тест

- 1 выборка

→ Нез. набл. и выд.

→ обе выборки $N(\mu, \sigma^2)$

$$t = \frac{\bar{x} - \mu_0}{\sqrt{s^2/n}} \sim t(n-1)$$

- две связанные выборки

1 объект: $X_i \mid Y_i$ $d_i = Y_i - X_i$

- две выборки

$$\sigma_T^2 = \sigma_c^2$$

Полти
никогда
не бывает

$$\dots \sim t(h_T + h_c - 2)$$

- Если все те выборки две, то они маленькие

Тест Уэлча: $z = \frac{\bar{x}_T - \bar{x}_c}{\sqrt{\frac{s_T^2}{n_T} + \frac{s_c^2}{n_c}}} \approx t(1)$

\uparrow = стрелная формула

④ Тест Манна-Уитни

⊗ Тест знаков (1 выборка)

мэр: медиана < 1 млн. р

x_1, x_2, \dots, x_n

$\oplus, \ominus, \dots, \oplus$

$$\bar{x} \quad \mu$$

$$T = \sum_{i=1}^n [\tilde{x}_i = \oplus] \sim \text{Bin}(p, n)$$

$$H_0: \text{Med}(x) = 1$$

$$H_A: \text{Med}(x) < 1$$

● 2 нез. выборки

x_1, \dots, x_{n_T}

y_1, \dots, y_{n_c}

$$U = \sum_{i=1}^{n_T} \sum_{j=1}^{n_c} [x_i > y_j] \sim ?$$

Тест М.-У.

~ Табличное

$$\sim \begin{matrix} n_T \rightarrow \infty \\ n_c \rightarrow \infty \end{matrix}$$

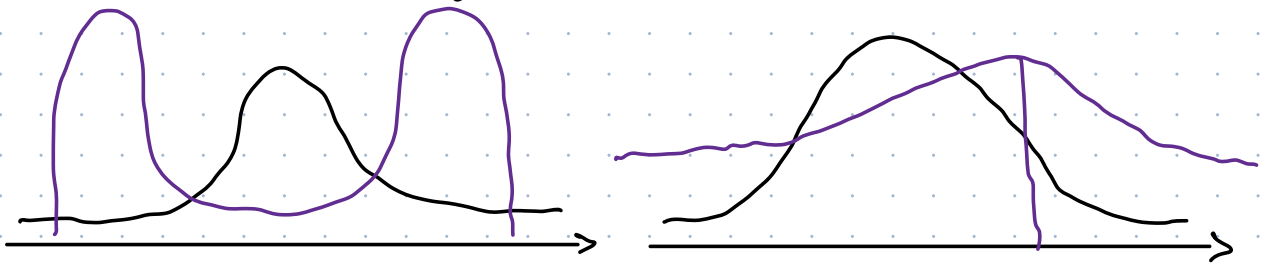
$$\mathcal{N}\left(\frac{n_x \cdot n_y}{2}; \text{большая}\right)$$

$$H_0: f_x(x) = f_y(y)$$

$$H_A: f_x(x) = f_y(y + \Delta), \Delta \neq 0$$



Когда тест развивается



ММ: качество не
изменяется

⑤ Бутстреп