① Асимптотический АБ-тест

Вино

Treatment → новый продукт P_T

Control → старое вино P_c

$$H_0: \; P_T - P_c \overset{=}{\leq} 0$$

$$H_A: \; P_T - P_c > 0$$

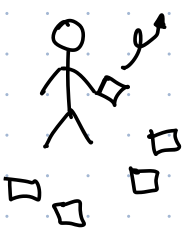
	H_0 верна	H_0 - не верна	
Не отв. H_0	ok	ошибка 2 рода	β
Отв. H_0	ошибка 1 рода	ok	

α

1) Заказал α, n
2) T_{obs}
3) T_{obs} vs T_{cr}

$1-\beta$ – мощность критерия

Упражнение



$X_1, \dots, X_n \sim iid \; Bern(p)$

p – вер. маслом вниз

$H_0: \; p = 1$

$H_A: \; p < 1$

Критерий:

$$T = \sum_{i=1}^{n} X_i$$

$T < n \Rightarrow$ отвергаю

$$\alpha = \mathbb{P}(H_A \mid H_0) = \mathbb{P}\left(\sum_{i=1}^{n} X_i < n \,\Big|\, p = 1\right) = \mathbb{P}\left(\text{хотя бы один ноль} \,\Big|\, p = 1\right) =$$

$$= 0$$

X_i	1	0
	p	$1-p$

$$\beta = \mathbb{P}(H_0 \mid H_A) = \mathbb{P}\left(\sum_{i=1}^{n} X_i = n \,\Big|\, p < 1\right) = p^n$$

$\beta \xrightarrow[n \to \infty]{} 0$ Состоятельность критерия

Предпосылки:

- Наблюдения нез.
- $Var(X_i) < \infty$ (нет выбросов)
- Big Data $n \to \infty$

$\Rightarrow$ ЦПТ

Асимпт. критерий

Союзник:

$$\bar{X}_n \overset{asy}{\sim} N\left(\mu; \frac{\hat{\sigma}^2}{n}\right)$$

X_i	0	1
	$1-p$	p

$\mu = p$

$\sigma^2 = \hat{p}(1-\hat{p})$

Данные и модель:

$X_1^c, X_2^c, \ldots, X_{n_c}^c \sim iid\ Bern(p_c)$

$X_1^T, X_2^T, \ldots, X_{n_T}^T \sim iid\ Bern(p_T)$

независимые

$$\bar{X}_{n_c} = \hat{p}_c \overset{asy}{\sim} N\left(p_c, \frac{\hat{p}_c \cdot (1-\hat{p}_c)}{n_c}\right)$$

$$\bar{X}_{n_T} = \hat{p}_T \overset{asy}{\sim} N\left(p_T, \frac{\hat{p}_T (1-\hat{p}_T)}{n_T}\right)$$

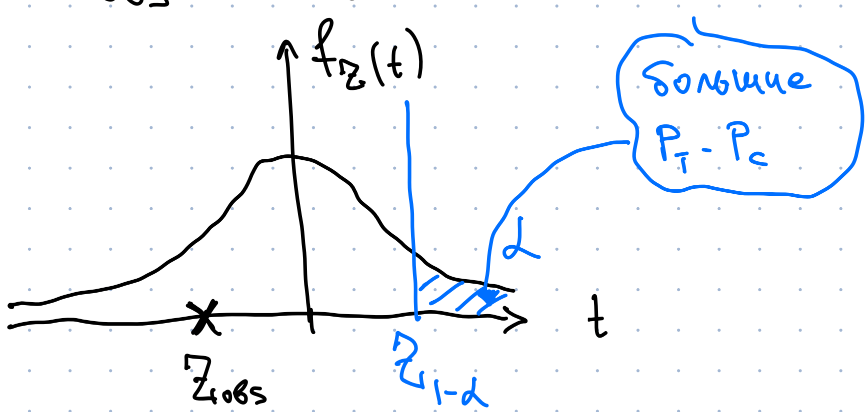
$$\hat{p}_T - \hat{p}_c \overset{asy}{\sim} N\left(p_T - p_c; \frac{\hat{p}_T(1-\hat{p}_T)}{n_T} + \frac{\hat{p}_c(1-\hat{p}_c)}{n_c}\right)$$

$\boxed{\text{Вера в } H_0:}$

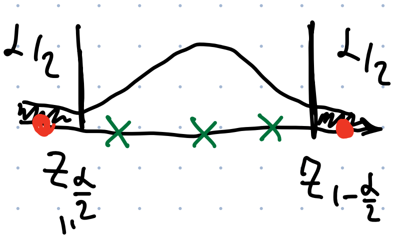
$$\hat{p}_T - \hat{p}_c \underset{H_0}{\overset{asy}{\sim}} N\left(0; \frac{\hat{p}_T(1-\hat{p}_T)}{n_T} + \frac{\hat{p}_c(1-\hat{p}_c)}{n_c}\right)$$

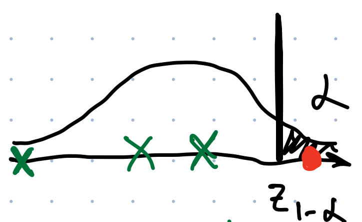
$$Z = \frac{\hat{p}_T - \hat{p}_c}{\sqrt{\frac{\hat{p}_T(1-\hat{p}_T)}{n_T} + \frac{\hat{p}_c(1-\hat{p}_c)}{n_c}}} \underset{H_0}{\overset{asy}{\sim}} N(0,1)$$

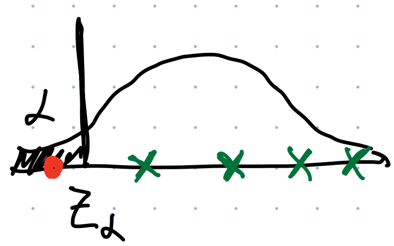
Z_{obs} — подставляем наблюдения в формулу

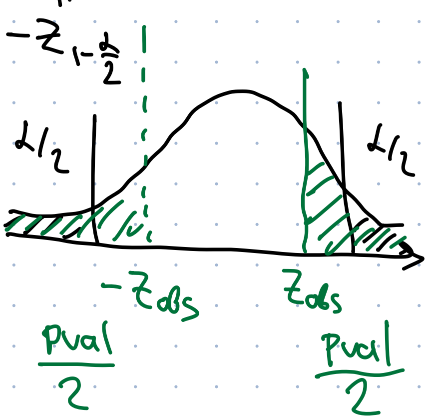


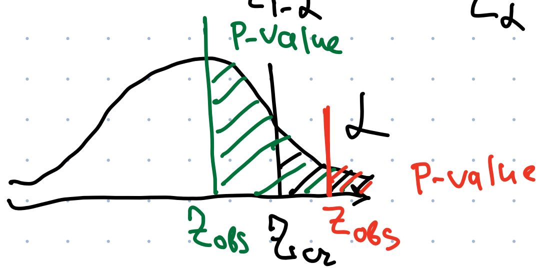
$H_0: \mu = 0$	$H_0: \mu \le 0$	$H_0: \mu \ge 0$
$H_A: \mu \ne 0$	$H_A: \mu > 0$	$H_A: \mu < 0$











Pvalue $> \alpha$	не отв.	$Z_{obs} < Z_{cr}$
Pvalue $< \alpha$	отв.	$Z_{obs} > Z_{cr}$

Опр. $P_{value} = \mathbb{P}(Z \geq Z_{obs}$ или $Z \leq -Z_{obs} \mid H_0)$

для двусторонней

выборка $\to$ Z_{obs}

$Z_{cr} = $ norm.ppf$(1-\alpha)$

Z_{obs} vs Z_{cr}.

pValue = norm.cdf$(-Z_{obs})$

1 - norm.cdf(Z_{obs})

pValue vs α

② Планирование эксперимента

T - критерий

α, β - ошибки

MDE - minimal detectable effect = минимально $P_T - P_c$

$n_T; n_c$ - число набл.

$n_T + n_c = n$ бюджет: собрать n набл.

$n_T = q \cdot n$

$n_c = (1-q) \cdot n$

Вопросы:

1) Как выбрать q ?

$\to$ от балды $q = 0.5$

$$\to \mathrm{Var}(\hat{p}_c - \hat{p}_T) = \frac{\hat{p}_T(1-\hat{p}_T)}{q \cdot n} + \frac{\hat{p}_c(1-\hat{p}_c)}{(1-q)\cdot n} \to \min_q$$

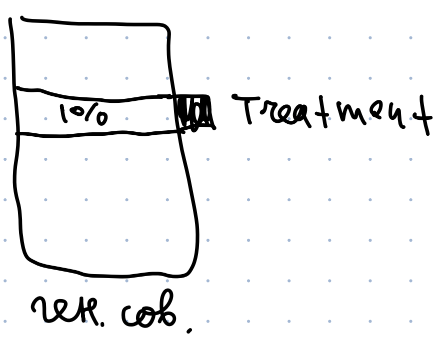
Если решить:

$$q = \frac{\sqrt{\hat{p}_T(1-\hat{p}_T)}}{\sqrt{\hat{p}_T(1-\hat{p}_T)} + \sqrt{\hat{p}_c \cdot (1-\hat{p}_c)}}$$

Скорее всего, все революционные внедрения сделаны, мы ожидаем маленькое MDE и $q \approx 0.5$ иначе

<u>2-х шаговый эксп</u>

- $q = 0.5$ $\hat{p}_T$ и $\hat{p}_c$ (неделя) $\Rightarrow q^*$
- $q = q^*$ АБ (месяц)

→ 

<u>Остальное − control</u>

Осталось понять хватает ли 1%

2) Какой брать критерий T $\quad \frac{\hat{\delta} - 0}{se(\hat{\delta})} > z_{1-\alpha} \Rightarrow$ отв.

3) Сколько надо n? $\quad n(\alpha, \beta, MDE)$

$\alpha = \mathbb{P}(H_A \mid H_0)$ фиксируем перед эксп.

MDE фиксируем из здравого смысла

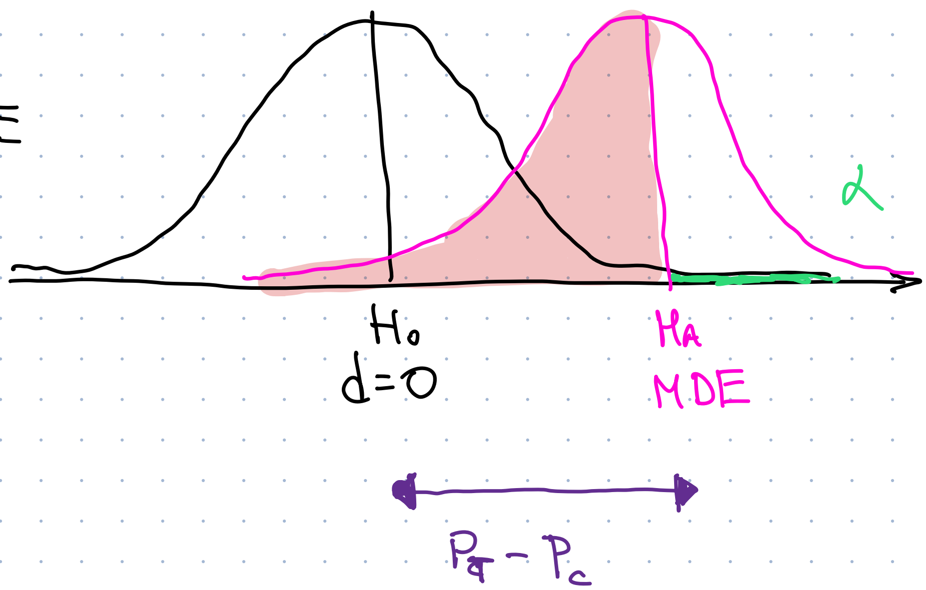
β $n - ?$

$$\beta = \mathbb{P}(H_0 \mid H_A) = \mathbb{P}\left(\hat{d} < z_{1-\alpha} \cdot se(\hat{d}) \mid d \geq MDE\right) = ⊛$$

$n \uparrow$

$H_0 : d = 0$

$H_A : d \underline{\underline{\geq}} MDE$



$$⊛ = \mathbb{P}\left(\hat{d} < z_{1-\alpha} \cdot se(\hat{d}) \mid d = MDE\right) = ⊛$$

$$\hat{d} \underset{H_0}{\overset{asy}{\sim}} \mathcal{N}\left(0, se^2(\hat{d})\right)$$

$$\hat{d} \underset{H_A}{\overset{asy}{\sim}} \mathcal{N}\left(MDE, se^2(\hat{d})\right)$$

$$⊛ = \mathbb{P}\left(\frac{\hat{d} - MDE}{se(\hat{d})} < \frac{z_{1-\alpha} \cdot se(\hat{d}) - MDE}{se(\hat{d})} \mid d = MDE\right) =$$

$$= \mathbb{P}\left(\mathcal{N}(0;1) \leq \bullet\right) = F_{\mathcal{N}(0,1)}\left(\bullet\right) = \beta$$

$$\frac{z_{1-\alpha} \cdot se(\hat{d}) - MdE}{se(\hat{d})} = z_{\beta}$$

$$\frac{MDE}{se(\hat{d})} = z_{1-\alpha} - z_{\beta}$$

а) 2-х шаговый эксп.

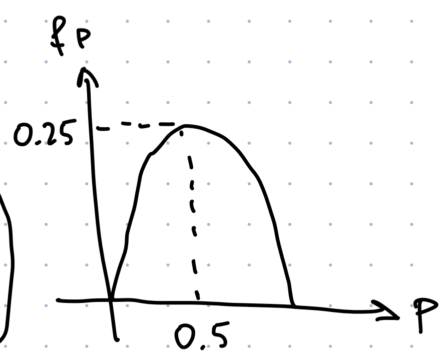
$$se(\hat{d}) = \frac{MDE}{z_{1-\alpha} + z_{1-\beta}}$$

б) ≤ 0.25 значение n будет завышено

$$Var(\hat{d}) = \left(\frac{\hat{p}_c(1-\hat{p}_c)}{q \cdot n} + \frac{\hat{p}_T(1-\hat{p}_T)}{(1-q)\cdot n}\right) = \frac{MDE^2}{(z_{1-\alpha} + z_{1-\beta})^2}$$

$q = 0.5$

$$\boxed{\begin{aligned} f(p) &= p(1-p) \\ f'(p) &= 1 - 2p = 0 \\ p &= 0.5 \end{aligned}}$$



решаем ур-е

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

Упражнение

Менеджер по книжке „Чайка"

- Хочу сделать ↑ конв.

× Какое MDE устроит?

- 0.001 сойдёт

× Какую α, β ты готов совершить

- Нулевые

× Нам нужен ∞ бюджет и ∞ времени

- А какие варики?

α \ β	0.001	0.01	0.05
0.001	n		
0.01			
0.05			

сколько/день людей на сайте

$\Rightarrow \frac{n}{avg.} = t$

$t < 7$ – не надо

③ Гипотезы vs Д.и.

$H_0: \mu_T = \mu_C$

$H_A: \mu_T \neq \mu_C$

Процедура 1:

а) $X_1, \dots, X_{n_T}$

$X_1, \dots, X_{n_C}$

б) $Z_{obs} = \dfrac{\bar X_T - \bar X_C}{\sqrt{\dfrac{\hat\sigma_T^2}{n_T} + \dfrac{\hat\sigma_C^2}{n_C}}}$

в) $Z_{obs} \overset{asy}{\underset{H_0}{\sim}} N(0;1)$

г) $Z_{1-\frac{\alpha}{2}}$

д) $|Z_{obs}| < Z_{1-\frac{\alpha}{2}}$

$-Z_{1-\frac{\alpha}{2}} < Z < Z_{1-\frac{\alpha}{2}}$

не отвергаем

Процедура 2:

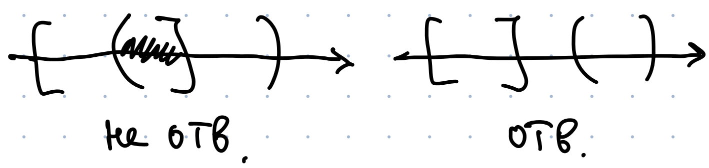
Плохая процедура

а) $X_1, \dots, X_{n_T}$

$X_1, \dots, X_{n_C}$

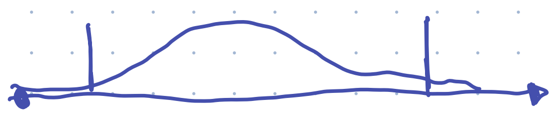
б) $\bar x_C \quad \bar x_T$

в) д.и. $\bar x_C$ д.и. $\bar X_T$



$\bar X_T - \bar X_C \overset{asy}{\sim} N(\mu_T - \mu_C, se)$

$\bar x_T - \bar x_C \pm Z_{1-\frac{\alpha}{2}} \cdot se$



Для простоты: $\hat\sigma_C^2 = \hat\sigma_T^2 = 1$

$\bar X_C > \bar X_T$

$n_C = n_T = n$

$+$ $H_0: \mu_C = \mu_T$

$H_A: \mu_C > \mu_T$

$$① \quad \beta = \mathbb{P}(H_0 \mid H_A) = \mathbb{P}\left(\frac{\bar{x}_c - \bar{x}_T}{\sqrt{2/n}} < z_{1-\alpha} \,\middle|\, \mu_c > \mu_T \right) =$$

$$= \mathbb{P}\left(\frac{\bar{x}_c - \bar{x}_T - (\mu_1 - \mu_2)}{\sqrt{2/n}} < z_{1-\alpha} - \frac{\mu_1 - \mu_2}{\sqrt{2/n}} \,\middle|\, \mu_c > \mu_T \right) =$$

$$= F_{N(0,1)}\left(z_{1-\alpha} - \frac{\mu_1 - \mu_2}{\sqrt{2/n}} \right)$$

$$② \quad \bar{x}_c \pm z_{1-\alpha} \cdot \sqrt{1/n}$$

$$\bar{x}_T \pm z_{1-\alpha} \cdot \sqrt{1/n}$$

$$\bar{x}_T + z_{1-\alpha} \cdot \sqrt{1/n} > \bar{x}_c - z_{1-\alpha} \cdot \sqrt{1/n}$$

$$\frac{\bar{x}_c - \bar{x}_T}{\sqrt{2/n}} < \cancel{2} \cdot z_{1-\alpha} \cdot \cancel{\frac{1}{\sqrt{2}}} \cdot \sqrt{2}$$

$$\mathbb{P}(H_0 \mid H_A) = \mathbb{P}\left(\frac{\bar{x}_c - \bar{x}_T - (\mu_1 - \mu_2)}{\sqrt{2/n}} < \sqrt{2} \cdot z_{1-\alpha} - \frac{\mu_1 - \mu_2}{\sqrt{2/n}} \,\middle|\, H_A \right)$$

$$= F_{N(0,1)}\left(\sqrt{2} \cdot z_{1-\alpha} - \frac{\mu_1 - \mu_2}{\sqrt{2/n}} \right)$$

$$\sqrt{2}\, z_{1-\alpha} > z_{1-\alpha}$$

$$\beta_1 > \beta_2$$

④ NPS

n – кол-во набл.

$L + D + N$

P_L, P_D, P_N

$$NPS = \frac{L - D}{n} = P_L - P_D$$

$$\hat{NPS} = \hat{P}_L - \hat{P}_D$$

$$Var(\hat{P}_L - \hat{P}_D) = \underbrace{Var(\hat{P}_L)}_{\frac{\hat{P}_L(1-\hat{P}_L)}{n}} + \underbrace{Var(\hat{P}_D)}_{\frac{\hat{P}_D(1-\hat{P}_D)}{n}} - 2\,\underbrace{Cov(\hat{P}_L, \hat{P}_D)}_{= -\frac{\hat{P}_L \hat{P}_D}{n}}$$

$$Cov(\hat{P}_L, \hat{P}_D) = Cov\left(\frac{L}{n}; \frac{D}{n}\right) = \frac{1}{n^2}\,\underbrace{Cov(L, D)}_{-n\,\hat{P}_L \hat{P}_D}$$

$$\underbrace{L + D + N}_{const} = n$$

$$Cov(L + D + N, D) = \underbrace{Cov(L, D)}_{x_{LD}} + \underbrace{Var(D)}_{\substack{V_D \\ n \cdot P_D \cdot (1 - P_D)}} + \underbrace{Cov(N, D)}_{x_{ND}} = 0$$

L

N

$$x_{ND} = -V_D - x_{LD} =$$

$$= -V_D + V_L - V_N - x_{ND}$$

$$x_{ND} = \frac{-V_D + V_L - V_N}{2}$$

$$\begin{cases} V_D + x_{LD} + x_{ND} = 0 \\ V_L + x_{LD} + x_{LN} = 0 \\ V_N + x_{LN} + x_{ND} = 0 \end{cases}$$

$$x_{LN} = -V_N - x_{ND}$$

$$x_{LD} = -V_L - (-V_N - x_{ND}) = -V_L + V_N + x_{ND}$$

$$x_{LD} = \frac{-V_D + V_N - V_L}{2}$$

$$Cov(L, D) = n \cdot \frac{-P_D(1-P_D) + P_N(1-P_N) - P_L(1-P_L)}{2} =$$

$$= n \cdot \frac{-P_D + P_D^2 + P_N - P_N^2 - P_L + P_L^2}{2} =$$

$$\boxed{\begin{aligned} &P_N = 1 - P_L - P_D \\ &P_N^2 = (1 - P_L - P_D)^2 = \\ &1 + P_L^2 + P_D^2 - 2P_L - 2P_D + 2P_L P_D \end{aligned}}$$

$$= \frac{n}{2} \cdot \Big(-2P_D - 2P_L + 1 + P_D^2 + P_L^2$$

$$= -\frac{n}{2} \cdot 2 P_L P_D = -n P_L P_D$$

Конечная ген. сов.

$$Var(\bar{X}) = \left(1 - \frac{n-1}{N-1}\right) \cdot \sigma^2$$

- <u>Бутстрэп</u>