

AБ-тест

→ Яон-бо набл. / пайплайн

- Тесты:
z-test
t-test
MW-test

бүтсран

→ Ускорение АБ-теста: CUPED / статистика

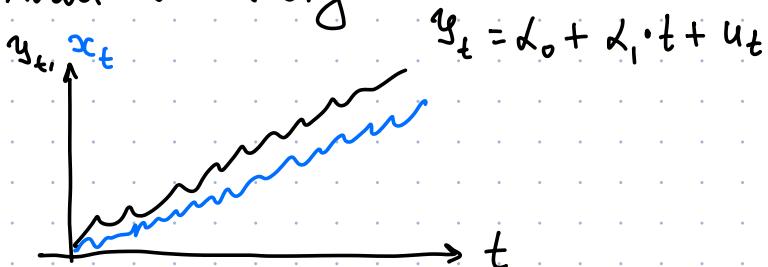
→ Онлайн АБ ML-алг ускорение АБ

Задачка:

y_t - потребление электроэнергии в Тае 1980-2020

x_t - рекорды по прыжкам в высоту

$$y_t = w_0 + w_1 \cdot x_t$$



② $w_1 > 0$

③ w_1 значима дз

$$H_0: w_1 = 0$$

$$H_A: w_1 \neq 0$$

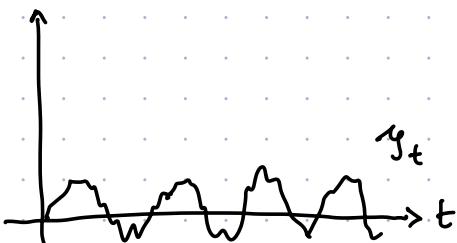
Логистическая регрессия

Логистические проблемы:

a) $y'_t = y_t - (\hat{\alpha}_0 + \hat{\alpha}_1 \cdot t)$

$$x'_t = x_t - (\hat{\beta}_0 + \hat{\beta}_1 \cdot t)$$

$$y'_t = w_0 + w_1 \cdot x'_t \quad \text{не значима}$$



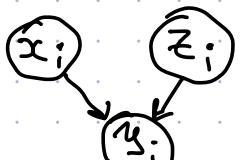
$$H_0: w_1 = 0 \quad H_A: w_1 \neq 0$$

$$y_t = w_0 + w_1 x_t + w_2 t + u_t$$

Перемінна
воздействів
Контрольні
перемінні

III. Прима-Во-Нобела

$$y_i \quad x_i \quad z_i \quad \text{Cov}(x_i, z_i) = 0$$

$$\hat{y}_i = \hat{w}_0 + \hat{w}_1 x_i + \hat{w}_2 z_i$$


$$\textcircled{1} \quad y_i = \hat{\lambda}_0 + \hat{\lambda}_1 x_i$$

$$y'_i = y_i - (\hat{\lambda}_0 + \hat{\lambda}_1 x_i)$$

$$\textcircled{2} \quad y'_i = \hat{\gamma}_0 + \hat{\gamma}_1 z_i \Rightarrow \begin{aligned} \hat{\gamma}_1 &= \hat{w}_2 \\ \hat{\lambda}_1 &= \hat{w}_1 \end{aligned}$$

A/B-тест через мінімальну розницю

$$\begin{array}{ll} y_1^c, \dots, y_{n_c}^c & \mu_c \quad \sigma_c^2 \\ y_1^t, \dots, y_{n_t}^t & \mu_t \quad \sigma_t^2 \end{array}$$

y_i - метрика

$$H_0: \mu_c = \mu_t$$

$$H_A: \mu_c \neq \mu_t$$

однаки
тому

$$H_0: w_1 = 0$$

$$H_A: w_1 \neq 0$$

$$d_i = \begin{cases} 1, \text{если } i \in \text{treatment} \\ 0, \text{иначе} \end{cases}$$

$$\mu_c + (\mu_t - \mu_c)$$

$$y_i = w_0 + w_1 d_i + u_i$$

$$d_i = 0 \quad y_i = \hat{w}_0 = \bar{y}_{\text{control}}$$

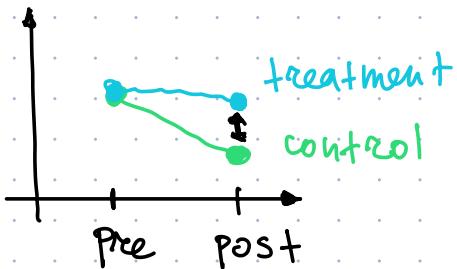
$$d_i = 1 \quad y_i = \hat{w}_0 + \hat{w}_1 = \bar{y}_{\text{treatment}}$$

$$\hat{w}_1 = \bar{y}_{\text{treatment}} - \bar{y}_{\text{control}}$$

$$\mu_t - \mu_c$$

DnD (difference in difference)

| | Pre | Post |
|-------|-----|------|
| cont | | |
| treat | | |



$$\hat{dd} = (\bar{y}_{\text{post}}^t - \bar{y}_{\text{pre}}^t) - (\bar{y}_{\text{post}}^c - \bar{y}_{\text{pre}}^c)$$

$$H_0: dd = 0$$

$$H_A: dd \neq 0$$

>
<

>To be on the causal
path!

y_i - metrika

$$d_i = \begin{cases} 1, & \text{tr.} \\ 0, & \text{cont.} \end{cases}$$

$$z_i = \begin{cases} 1, & \text{post} \\ 0, & \text{pre} \end{cases}$$

$$y_i = w_0 + w_1 d_i + w_2 z_i + w_3 d_i z_i + u_i$$

$$\hat{w}_0 = \bar{y}_{c, \text{pre}}$$

$$\hat{w}_1 = \bar{y}_{t, \text{pre}} - \bar{y}_{c, \text{pre}}$$

$$\hat{w}_2 = \bar{y}_{c, \text{post}} - \bar{y}_{c, \text{pre}}$$

$$\hat{w}_3 = \hat{dd} = (\bar{y}_{\text{post}}^t - \bar{y}_{\text{pre}}^t) - (\bar{y}_{\text{post}}^c - \bar{y}_{\text{pre}}^c)$$

$$H_0: w_3 = 0$$

$$H_A: w_3 \neq 0$$

>
<

Онлайн-AB Тестирование

онлайн маркетинг

Перекрёсток Б.у.

Перекрёсток Channel

$$y_i = w_0 + w_1 d_i + w_2 z_i + w_3 d_i \cdot z_i + w_4 x_{i1} + w_5 x_{i2} + \dots + w_k x_{ik-3} + u_i$$

контрольные

переменные

CUPED (2013) Alex Deng

Controlled-experiment using pre-experiment data

$$z_i = \frac{\bar{y}^t - \bar{y}^c}{se(\bar{y}^t - \bar{y}^c)}$$

CUPED хочет ↓
се не увер.
размер BGD.

$$\Delta \in [\bar{y}^t - \bar{y}^c - 1.96 \cdot se(\dots); \bar{y}^t - \bar{y}^c + 1.96 \cdot se(\dots)]$$

$$\Delta = \mu^t - \mu^c \quad \hat{\Delta} = \bar{y}^t - \bar{y}^c \quad \hat{\Delta}^{\text{CUPED}} = \bar{y}^t - \bar{y}^c$$

\bar{y} наименьшее x , которое не принадлежит
какому либо control / treatment

$$\hat{y} = y - \theta(x - E(x))$$

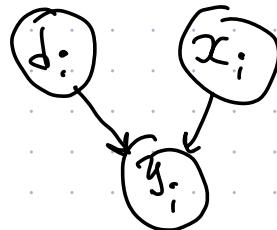
$$\text{CI: } \bar{y}^t - \bar{y}^c \pm 1.96 \cdot se(\dots)$$

$$E(\hat{y}) = E(y) - \theta(E(x) - E(x)) = E(y)$$

$$\text{Var}(\hat{y}) = \text{Var}(y) + \text{Var}(x) \theta^2 - 2\theta \text{Cov}(y, x) \rightarrow \min_{\theta}$$

$$\theta^* = \frac{s \text{Cov}(x, y)}{s \text{Var}(x)}$$

$$y_i \quad d_i \in \{0, 1\} \quad x_i$$



Re CUPED:

$$y_i = w_0 + w_1 d_i + u_i$$

МНК \hat{w}_1 СІ $H_0: w_1 = 0$
 МА: $w_1 \neq 0$

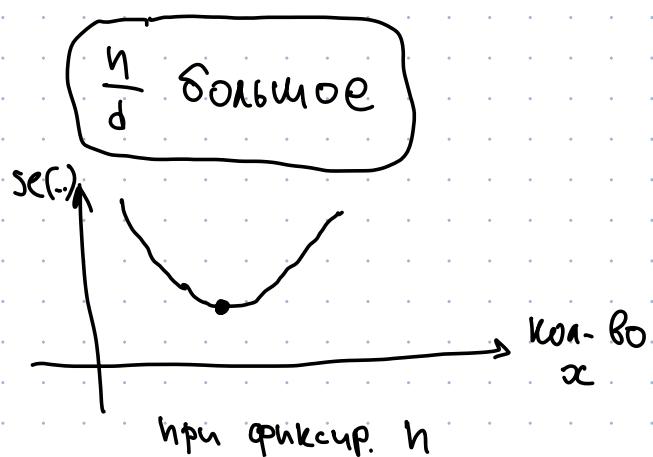
$$\begin{cases} u_i \sim \text{норм.} \\ u_i \sim \text{стн.} \end{cases}$$

Добавлены характеристики элегантности:

$$y_i = w_0 + w_1 d_i + w_2 x_i + u_i$$

МНК \hat{w}_1 СІ $H_0: w_1 = 0$
 МА: $w_1 \neq 0$

$se(\hat{w}_1)$ означает
максимум, для
 \hat{w}_1 при x_i



CUPED — оценивание МН. регр.
 + добавление в ков. матр. нулей 0 0

d_i :rez. от x_i :

$$\text{Var}(\hat{w}_i | X) = (X^T X)^{-1} \cdot \sigma^2 = \begin{bmatrix} * & * \\ * & * \\ * & 0 \end{bmatrix}$$

$$\hat{w} = \begin{pmatrix} \hat{w}_0 \\ \hat{w}_1 \\ \hat{w}_2 \end{pmatrix}$$

$$\text{Cov}(d_i, x_i) = 0$$

Пример: как сгенерировать CUPED через матрицы

$$① \hat{y}_i = \hat{w}_0 + \hat{w}_1 d_i + \hat{w}_2 x_i$$

$$z_i = y_i - \hat{w}_2 x_i$$

$$② \hat{z}_i = \hat{y}_0 + \hat{y}_1 \cdot d_i$$

$$\text{se}(\hat{y}_1) < \text{se}(\hat{w}_1)$$

$$\text{se}(\hat{w}_1)$$

каждое
члену

на то...

Это способ показать Python, что $\text{Cov}(d_i, x_i) = 0$

+ Проверка для расширенного метода

→ можно ли вместо одного x_i использовать коеф?

Да можно! CUMPED

$$① \hat{y}_i = \hat{w}_0 + \hat{w}_1 d_i + (\hat{w}_2 x_i + \hat{w}_3 h_i)$$

$$z_i = y_i - (\hat{w}_2 x_i + \hat{w}_3 h_i)$$

$$② \hat{z}_i = \hat{y}_0 + \hat{y}_1 d_i \quad \text{se}(\hat{y}_1)$$

→ A оғынға 82876 неремелілең x, h и зарапті.

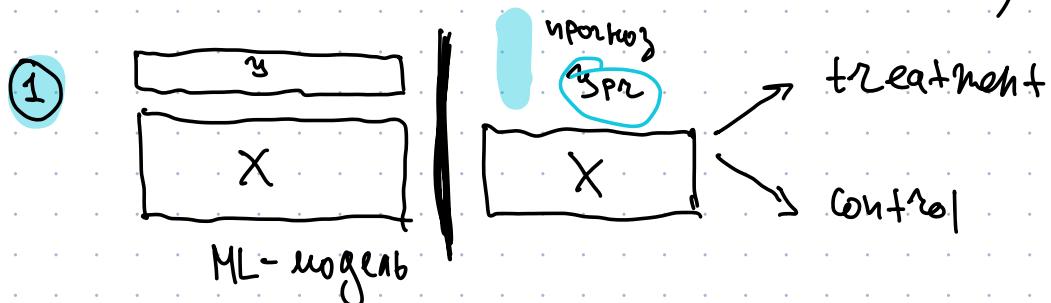
4x көз. ом ді:

y_i - нақылдаған оғын

$x_i = y_i^{\text{pre}}$ - нақылдаған "бұрса"

→ А монда оғындың сүйкітіб ML?

CUPAC (control using predictions as covariate)



$$\textcircled{1} \quad \hat{y}_i = \hat{w}_0 + \hat{w}_1 d_i + \hat{w}_2 y_{pr,i} \quad \frac{1}{n} \sum_i (y_i - \hat{y}_i) \rightarrow \min_w$$

$$2.2 \quad \hat{z}_i = y_i - \hat{w}_2 y_{pr,i}$$

$$\textcircled{3} \quad \hat{z}_i = \hat{y}_0 + \hat{y}_1 d_i \quad \text{se}(\hat{y}_1)$$

На салынғанда DNB - ӘТО Томе CUPED!

$$\textcircled{1} \quad \hat{y}_i^{\text{post}} = w_0 + w_1 d_i + \underbrace{w_2}_{\hat{z}} y_i^{\text{pre}}$$

$$z_i = \hat{y}_i^{\text{post}} - \hat{y}_i^{\text{pre}}$$

$$\textcircled{2} \quad \hat{z}_i = \hat{y}_0 + \hat{y}_1 d_i \quad \hat{y}_1 = \hat{d}_d$$

LATE - ?

Мэтрик - ?

Философский Ти Рага

→ Аук. регр. — это АБ-тест

→ ML / Эконометрика

Два великих вопроса

Что будет завтра?

Как предсказать?

→ ML

BVD

Что за фичи прогноза?

Как изменить эти же самые?

→ Статистика

$\hat{\theta}$ — техн. соч. ЗП.

$$\frac{n}{d} \ll 1$$

\Rightarrow переобучение

\Rightarrow из рукава —
регуляризация

Метод
нравдоп.

$$\frac{n}{d} \rightarrow \infty$$

\Rightarrow кайфрайтинг

