

Причинно-следственные связи в данных

Георгий Калашнов, Ольга Сучкова

16 марта 2020 г.

План на сегодня

Тестирование гипотез

Снижение дисперсии с помощью контрольных переменных

Пример

Ковариаты

Table of Contents

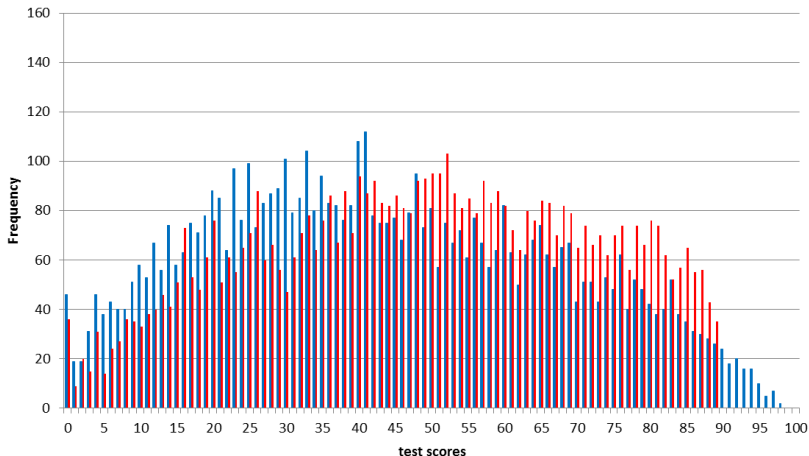
Тестирование гипотез

Снижение дисперсии с помощью контрольных переменных

Пример

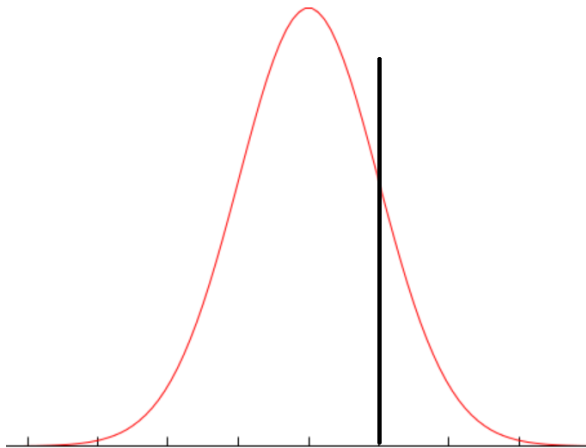
Ковариаты

Оценки школьников в Мумбаи: $ATE^{\wedge} = 6$. Можно ли доверять результату?



Здесь и далее графики – см из п.2 списка литературы на последнем слайде

Точечная и интервальная оценка. Доверительный интервал



«Презумпция нуля»

- «Бремя доказательства» лежит на исследователе: важно доказать, что эффект значим.
- Аналогично именно на исследователе лежит бремя доказательства, что оценка показывает именно заявленный эффект, что отсутствует «влияние третьих факторов» и т.д.

Ошибки первого и второго рода

	Тритмент-эффект есть	Тритмент-эффекта нет
Тест в пользу H1 Оценка тритмент- эффекта значимая	Ок	Ошибка 1 рода (вероятность = альфа)
Тест в пользу H0 Оценка тритмент- эффекта незначимая	Ошибка 2 рода (вероятность = β)	Ок

Уровень значимости

- Уровень значимости 5% означает, вероятность случайно получить отличный от нуля результат составляет 5%.

Мощность

- Это вероятность того, что если истинный эффект - размера b , то наш эксперимент будет в состоянии разграничить оценку этого эффекта и ноль.
- Это вероятность избежать ошибки 2 рода ($= 1 - k$).

Связь: мощность и...

- предполагаемый размер эффекта
- вариативность показателя
- размер выборки
- пропорция между тритмент- и контрольной группой

MDE (minimum detectable effect) size

The diagram shows the formula for Minimum Detectable Effect (MDE) size. Red arrows point from labels to specific parts of the formula: 'Effect Size' points to the entire formula; 'Power' points to the $t_{(1-\kappa)}$ term; 'Significance Level' points to the t_α term; 'Proportion in Treatment' points to the P in the denominator of the first square root; 'Variance' points to the σ^2 in the second square root; and 'Sample Size' points to the N in the denominator of the second square root.

$$EffectSize = \left(t_{(1-\kappa)} + t_\alpha \right) * \sqrt{\frac{1}{P(1-P)}} * \sqrt{\frac{\sigma^2}{N}}$$

Источник: Glennerster, Takavarasha "Running randomized evaluations. A practical Guide", ch 6.

Пример: проект NERICA

- L. Casaburi, R. Glennerster, S. Kamara, T. Suri (2014) "Providing collateral and improving product market access for smallholder farmers in Sierra Leone"
- Части фермеров в Сьерра Леоне раздали новый сорт риса на посев.
- Выросла ли урожайность и выручка фермеров? – Опрос фермеров, но затраты на дорогу!
- Размер эффекта – по результатам «лабораторных испытаний» - как должна отличаться урожайность (97 кг с гектара)
- Уровень значимости и мощность – заданы
- Расчёт необходимого количества наблюдений

MDE: аналогия

$$EffectSize = (t_{(1-\kappa)} + t_{\alpha}) * \sqrt{\frac{1}{P(1-P)}} * \sqrt{\frac{\sigma^2}{N}}$$

Стандартная ошибка
оценки коэффициента

- Аналогия: расчётная t-статистика в тесте на значимость коэффициента показывает, на сколько стандартных ошибок оценка отличается от нуля.

Minimum detectable effect: аналогия

=>

Чтобы быть значимым на 5% уровне, коэффициент должен быть как минимум в “t_табличное” раз больше, чем его стандартная ошибка

=>

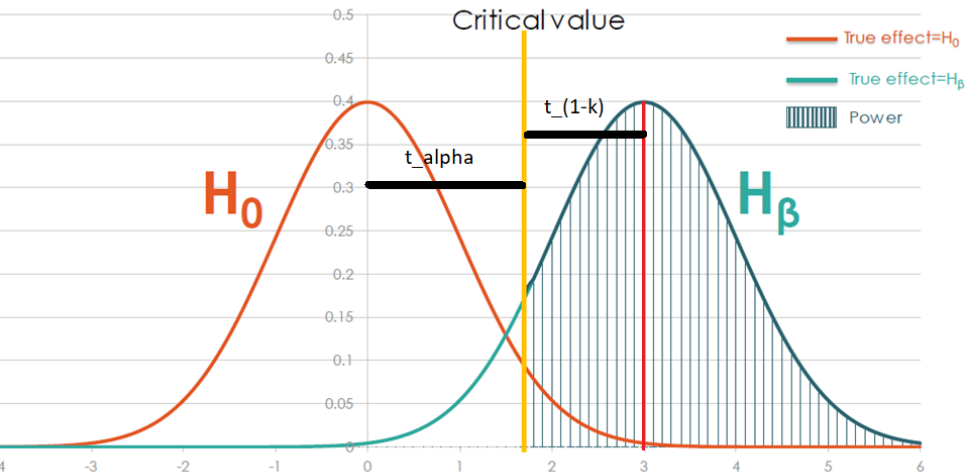
Минимально отличимый от нуля на alpha %-ном уровне коэффициент должен быть $= t_{\alpha/2} * s.e(b^{\wedge})$

Minimum detectable effect: наш случай

- Чтобы быть отличным от нуля при заданной мощности $(1-k)$ и уровне значимости (α) , коэффициент должен быть как минимум в $t_{(1-k)} + t_{\alpha}$ раз больше, чем s.e.

$$EffectSize = (t_{(1-\kappa)} + t_{\alpha}) * \sqrt{\frac{1}{P(1-P)}} * \sqrt{\frac{\sigma^2}{N}}$$

MDE графически



Что почитать

- Gerber, Green “Field experiments”, ch.4, 10
- Glennerster, Takavarasha “Running randomized evaluations. A practical Guide”, ch. 6
- Owen Ozier, лекция + выкладки + примеры
http://economics.ozier.com/owen/slides/ozier_powercalc_talk_20100914a.pdf
- Кратко + код: <https://egap.org/methods-guides/10-things-you-need-know-about-statistical-power>

Table of Contents

Тестирование гипотез

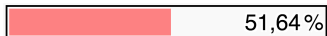
Снижение дисперсии с помощью контрольных переменных

Пример

Ковариаты

Президентские выборы во Франции 2012 (Pons 2018)

Франсуа Олланд



Николя Саркози



Президентские выборы во Франции 2012 (Pons 2018)

Франсуа Олланд



Николя Саркози



Эффект от агитации в 0.5 процентных пункта.

Как суметь поймать такой эффект?

2 метода снижения дисперсии:

1. Контрольные переменные (**covariates**)
2. Грамотное планирование эксперимента и престратификация (в другой раз)

Ковариаты¹

- ▶ Y_1, Y_0 – потенциальные исходы (**potential outcomes**)
- ▶ $T = 1$, если наблюдение в эксперименте и 0 иначе (**treatment variable**)
- ▶ X – Независимые переменные (**covariates**)

Ковариаты X меняются вместе с Y_1 и Y_0 (коварируют или (мульти) коррелируют)

$$\text{Cov}(X, Y_1) > 0, \text{Cov}(X, Y_0) > 0$$

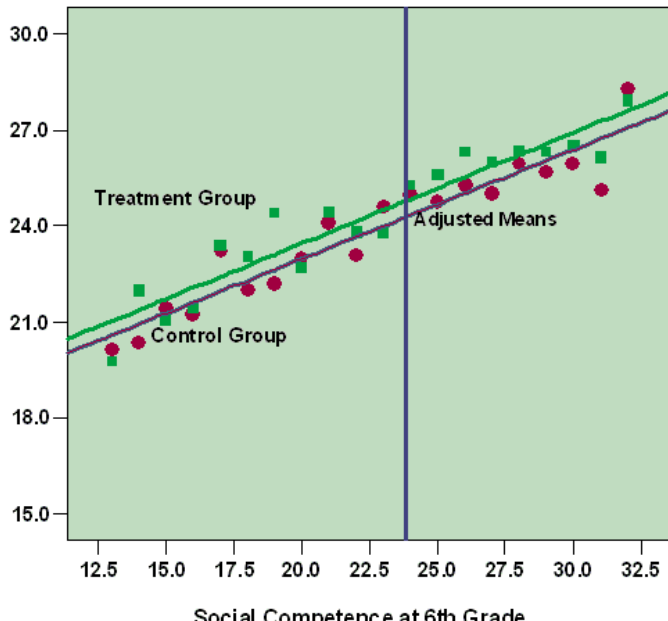
¹Angrist и Pischke 2008, Раздел 3.1.1, Imbens и Rubin 2015, Глава 7.5-7.8

Маленький пример

	Y_1	Y_0	X
Пациент 1	-	36.6	Из Европы
Пациент 2	36.6	-	Из Европы
Пациент 3	35	-	Из Европы
Пациент 4	-	36	Из Европы
Пациент 5	37.3	-	Из Азии
Пациент 6	-	39.3	Из Азии
Пациент 7	37.2	-	Из Азии
Пациент 8	-	39.2	Из Азии

- ▶ *Повторение.* На глаз. Выполнена ли предпосылка экзогенности? $(Y_1, Y_0, X)_i \perp T_i$
- ▶ Высока ли дисперсия с одним и тем же X : $\mathbb{V}(Y|X)$?
- ▶ У кого в среднем выше температура? $\text{Cov}(X, Y_0) - ?$
- ▶ Высока ли общая дисперсия: $\mathbb{V}(Y)$?

То же, но на картинке



Контроль, снижающий дисперсию

$$\begin{aligned}\mathbb{V}(Y) &= \mathbb{E}((Y - \mathbb{E}Y)^2) = \\ &\mathbb{E}\left((Y - \mathbb{E}(Y|X) + \mathbb{E}(Y|X) - \mathbb{E}Y)^2\right) = \\ &\mathbb{E}((Y - \mathbb{E}(Y|X))^2) + \mathbb{E}((\mathbb{E}(Y|X) - \mathbb{E}Y)^2) + \\ &\quad 2\mathbb{E}((\mathbb{E}(Y|X) - \mathbb{E}Y)(Y - \mathbb{E}(Y|X))) = \\ &\mathbb{E}(\mathbb{V}(Y|X)) + \mathbb{V}(\mathbb{E}(Y|X)) + 0\end{aligned}$$

Результаты выборов во Франции

TABLE 5—IMPACT ON HOLLANDE’S VOTE SHARE

	First round			Second round			Average of first and second rounds		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>Panel A. ITT estimation</i>									
Treatment	0.0063 (0.0023)	0.0050 (0.0019)	0.0044 (0.0018)	0.0048 (0.0028)	0.0053 (0.0019)	0.0046 (0.0018)	0.0056 (0.0024)	0.0049 (0.0017)	0.0043 (0.0016)
Strata fixed effects	X	X	X	X	X	X	X	X	X
Control for past outcome and PO		X	X		X	X		X	X
Additional controls			X			X			X
Observations	3,390	2,660	2,660	3,390	2,660	2,660	3,390	2,660	2,660
R^2	0.003	0.516	0.528	0.001	0.632	0.645	0.002	0.645	0.655
Mean in control group	0.3157	0.2994	0.2994	0.5757	0.5597	0.5597	0.4457	0.4295	0.4295
<i>Panel B. Instrumental variable estimation: “allocated to canvassers” instrumented with “treatment”</i>									
Allocated to canvassers	0.0112 (0.0041)	0.0094 (0.0036)	0.0084 (0.0035)	0.0084 (0.0050)	0.0099 (0.0036)	0.0087 (0.0035)	0.0098 (0.0042)	0.0092 (0.0031)	0.0081 (0.0030)
Strata fixed effects	X	X	X	X	X	X	X	X	X
Control for past outcome and PO		X	X		X	X		X	X
Additional controls			X			X			X
Observations	3,390	2,660	2,660	3,390	2,660	2,660	3,390	2,660	2,660

Повторение. А это зачем?

TABLE 2—SUMMARY STATISTICS

	Control group		Treatment group		<i>p</i> -value treatment = control	Number of observations
	Mean	SD	Mean	SD		
<i>Panel A. Electoral outcomes</i>						
Randomization at precinct level	0.504	0.500	0.504	0.500	0.992	3,397
Number of registered citizens	1,014.3	1,097.6	1,133.8	1,605.3	0.022	3,397
Potential to win votes, PO	0.089	0.035	0.089	0.033	0.970	3,397
Voter turnout, 2007 pres. election, first round	0.843	0.050	0.840	0.048	0.231	2,665
Voter turnout, 2007 pres. election, second round	0.837	0.045	0.836	0.045	0.675	2,665
PS vote share, 2007 pres. election, first round	0.274	0.081	0.279	0.081	0.172	2,665
PS vote share, 2007 pres. election, second round	0.515	0.103	0.516	0.101	0.743	2,665

Повторение. А это зачем?

	First round			Second round			Average of first and second rounds		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>Panel A. ITT estimation</i>									
Treatment	0.0050 (0.0027)	0.0015 (0.0024)	0.0007 (0.0023)	0.0009 (0.0029)	-0.0013 (0.0024)	-0.0014 (0.0023)	0.0029 (0.0026)	-0.0002 (0.0020)	-0.0006 (0.0020)
Strata fixed effects	X	X	X	X	X	X	X	X	X
Control for past outcome		X	X		X	X		X	X
Additional controls			X			X			X
Observations	2,660	2,133	2,133	2,660	2,133	2,133	2,660	2,133	2,133
R^2	0.002	0.371	0.395	0.000	0.495	0.509	0.001	0.525	0.541
Mean in control group	0.2740	0.2620	0.2620	0.5146	0.5056	0.5056	0.3943	0.3838	0.3838
<i>Panel B. Instrumental variable estimation: "allocated to canvassers" instrumented with "treatment"</i>									
Allocated to canvassers	0.0093 (0.0049)	0.0028 (0.0045)	0.0015 (0.0045)	0.0016 (0.0054)	-0.0025 (0.0046)	-0.0028 (0.0046)	0.0055 (0.0048)	-0.0004 (0.0039)	-0.0011 (0.0039)
Strata fixed effects	X	X	X	X	X	X	X	X	X
Control for past outcome		X	X		X	X		X	X
Additional controls			X			X			X
Observations	2,660	2,133	2,133	2,660	2,133	2,133	2,660	2,133	2,133

Литература: книжки



Angrist, Joshua D и Jörn-Steffen Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.



Imbens, Guido W и Donald B Rubin (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Литература: статьи



Pons, Vincent (2018). «Will a five-minute discussion change your mind? A countrywide experiment on voter choice in France». в: *American Economic Review* 108.6, с. 1322—63.