

Spis treści

Przedmowa	11
Część I. Podstawy	19
1. Wprowadzenie do wnioskowania przyczynowego	21
Czym jest wnioskowanie przyczynowe?	21
Po co stosować wnioskowanie przyczynowe?	22
Uczenie maszynowe i wnioskowanie przyczynowe	23
Asocjacja a związek przyczynowy	24
Oddziaływanie i wynik	25
Podstawowy problem z wnioskowaniem przyczynowym	25
Modele przyczynowe	26
Interwencje	28
Indywidualny efekt oddziaływania	30
Wyniki potencjalne	30
Spójność i stabilna wartość oddziaływania jednostkowego	31
Interesujące wartości przyczynowe	32
Wartości przyczynowe — przykład	33
Błąd systematyczny	35
Wzór na błąd systematyczny	36
Wizualny przewodnik po błędzie systematycznym	37
Identyfikowanie efektu oddziaływania	40
Założenie o niezależności	41
Identyfikacja przy randomizacji	41
Najważniejsze zagadnienia	45
2. Eksperymenty z randomizacją i przegląd elementów statystyki	46
„Siłowe” zapewnianie niezależności za pomocą randomizacji	46
Przykładowy test A/B	48
Idealny eksperyment	51

Najbardziej niebezpieczne równanie	52
Błąd standardowy szacunków	55
Przedziały ufności	56
Testowanie hipotez	63
Hipoteza zerowa	64
Statystyki testowe	66
Wartości p	67
Moc testu	68
Obliczanie wielkości próby	70
Najważniejsze zagadnienia	71
3. Graficzne modele przyczynowe.....	73
Myślenie o przyczynowości	73
Wizualizacja związków przyczynowych	75
Czy konsultanci są warci swojej ceny?	77
Błyskawiczny kurs z zakresu modeli graficznych	78
Łańcuchy	78
Rozgałęzienia	80
Kolider	81
Ściąga dotycząca przepływu asocjacji	82
Tworzenie zapytań dotyczących grafu w Pythonie	83
Jeszcze o identyfikacji	86
Założenie o warunkowej niezależności i formuła korygująca	87
Założenie o dodatniości prawdopodobieństwa	88
Przykład identyfikacji na podstawie danych	89
Błąd spowodowany zmiennymi zakłócającymi	91
Zastępcze zmienne zakłócające	92
Jeszcze o randomizacji	92
Błąd doboru	93
Warunki dotyczące kolidera	94
Korygowanie błędu doboru	97
Warunek dotyczący mediatora	100
Najważniejsze zagadnienia	101
Część II. Uwzględnianie błędu systematycznego	103
4. Zaskakująca skuteczność regresji liniowej.....	105
Potrzebujesz tylko regresji liniowej	105
Dlaczego potrzebujemy modeli?	106
Regresja w testach A/B	107
Korygowanie za pomocą regresji	109

Teoria regresji	113
Regresja liniowa pojedynczej zmiennej	114
Wielozmiennowa regresja liniowa	114
Twierdzenie Frischa-Waughana-Lovella i ortogonalizacja	115
Etap eliminowania błędu systematycznego	116
Etap eliminowania szumu	118
Błąd standardowy estymatora regresji	119
Ostateczny model wyników	120
Podsumowanie na temat twierdzenia FWL	120
Regresja jako model wyników	122
Dodatniość prawdopodobieństwa i ekstrapolacja	124
Nieliniowość w regresji liniowej	125
Linearyzacja oddziaływania	127
Nieliniowe twierdzenie FWL i eliminowanie błędu systematycznego	129
Regresja z użyciem zmiennych zastępczych	130
Eksperymenty warunkowo losowe	130
Zmienne zastępcze	132
Nasycony model regresji	135
Regresja jako średnia ważona wariancją	137
Odejmowanie średniej i efekty stałe	139
Błąd systematyczny spowodowany pominiętą zmienną — zmienne zakłócające w kontekście regresji	141
Neutralne zmienne kontrolne	143
Zmienne kontrolne powodujące szum	144
Dobór cech: kompromis między błędem systematycznym a wariancją	145
Najważniejsze zagadnienia	147
5. Wskaźnik skłonności	148
Wpływ szkoleń menedżerskich	148
Uwzględnianie zmiennych za pomocą regresji	150
Wskaźnik skłonności	151
Szacowanie wskaźnika skłonności	152
Wynik skłonności i ortogonalizacja	152
Technika PSM	153
Wagi będące odwrotnością wskaźnika skłonności	155
Wariancja w metodzie IPW	157
Stabilizowane wagi oparte na wskaźniku skłonności	160
Pseudopopulacje	162
Błąd doboru	163
Kompromis między błędem systematycznym a wariancją	163
Dodatniość prawdopodobieństwa	164
Identyfikacja oparta na projekcie a identyfikacja oparta na modelu	167

Podwójnie odporna estymacja	167
Oddziaływanie łatwe do modelowania	169
Wynik łatwy do modelowania	172
Uogólniony wskaźnik skłonności dla oddziaływania ciągłego	173
Najważniejsze zagadnienia	179
Część III. Niejednorodność efektu i personalizacja	181
6. Niejednorodność efektu.....	183
OdATE do CATE	183
Dlaczego predykcje nie są rozwiązaniem	184
Obliczanie wartości CATE za pomocą regresji	187
Ocena predykcji wartości CATE	190
Ocena efektu na podstawie kwantyla z modelu	191
Skumulowany efekt	195
Skumulowany wzrost	196
Transformacja celu	199
Kiedy modele predykcyjne są dobre w porządkowaniu efektów?	201
Krańcowo malejące zwroty	201
Wyniki binarne	202
Wykorzystanie wartości CATE do podejmowania decyzji	203
Najważniejsze zagadnienia	206
7. Systemy metauczące.....	208
Systemy metauczące dla oddziaływania dyskretnego	208
T-learner	210
X-learner	213
Systemy metauczące dla oddziaływania ciągłego	217
S-learner	218
Podejście DDML	222
Najważniejsze zagadnienia	230
Część IV. Dane panelowe	231
8. Metoda różnicy w różnicach.....	233
Dane panelowe	234
Kanoniczna postać metody różnicy w różnicach	236
Różnica w różnicach ze wzrostem wyniku	238
Obliczanie różnicy w różnicach na podstawie błędu średniokwadratowego	240
Różnica w różnicach z efektami stałymi	241
Wiele przedziałów czasowych	241
Wnioskowanie	243

Założenia związane z identyfikacją	245
Trendy równoległe	246
Założenie o braku antycypacji	
i założenie o stabilnej wartości jednostki oddziaływania	248
Ścisła egzogeniczność	248
Brak czynników zakłócających zmiennych w czasie	249
Brak sprzężenia zwrotnego	250
Brak efektu przeniesienia i brak przesuniętej w czasie zmiennej zależnej	251
Dynamika efektu w czasie	251
Metoda różnicy w różnicach ze zmiennymi towarzyszącymi	253
Podwójnie odporna wersja metody różnicy w różnicach	256
Model oparty na wskaźniku skłonności	257
Model zmiany wyników	257
Łączenie wszystkich elementów	257
Stopniowe wprowadzanie oddziaływania	259
Niejednorodność efektu w czasie	264
Zmienne towarzyszące	267
Najważniejsze zagadnienia	268
9. Metoda syntetycznej kontroli.....	270
Zestaw danych dotyczących marketingu internetowego	270
Reprezentacja macierzowa	273
Metoda kontroli syntetycznej jako regresja pozioma	275
Kanoniczna wersja metody kontroli syntetycznej	278
Metoda kontroli syntetycznej ze zmiennymi towarzyszącymi	281
Eliminowanie błędu systematycznego w metodzie kontroli syntetycznej	285
Wnioskowanie	289
Metoda syntetycznej różnicy w różnicach	291
Jeszcze o metodzie różnicy w różnicach	292
Jeszcze o metodzie syntetycznej kontroli	292
Szacowanie wag związanych z czasem	294
Metoda syntetycznej kontroli i metoda różnicy w różnicach	296
Najważniejsze zagadnienia	298
 Część V. Inne projekty eksperymentów	 301
10. Eksperymenty geograficzne i eksperymenty z przełączaniem oddziaływania.	303
Eksperymenty geograficzne	304
Projektowanie z syntetyczną grupą kontrolną	305
Próba z losową grupą jednostek poddanych oddziaływaniu	307
Wyszukiwanie losowe	309

Eksperyment z przełączaniem oddziaływania	312
Potencjalne wyniki dla sekwencji	314
Szacowanie stopnia efektu przeniesienia	314
Szacowanie oparte na projekcie	316
Optymalny projekt eksperymentów z przełączaniem oddziaływania	320
Odporna wariancja	322
Najważniejsze zagadnienia	325
 11. Niezgodność ze schematem przydziału oddziaływania i zmienne instrumentalne	327
Niezgodność	327
Rozszerzanie notacji potencjalnych wyników	329
Założenia związane z identyfikacją zmiennych instrumentalnych	332
Pierwszy etap	334
Postać zredukowana	335
Dwuetapowa metoda najmniejszych kwadratów	336
Błąd standardowy	336
Dodatkowe zmienne kontrolne i instrumentalne	339
Ręczne stosowanie dwuetapowej metody najmniejszych kwadratów	340
Implementacja macierzowa	341
Projekt z nieciągłością	342
Założenia w projekcie z nieciągłością	343
Efekt zamiaru oddziaływania	344
Oszacowanie zmiennej instrumentalnej	346
Skupiska wartości	347
Najważniejsze zagadnienia	347
 12. Dalsze kroki	350
Odkrywanie relacji przyczynowych	350
Sekwencyjne podejmowanie decyzji	351
Przyczynowe uczenie ze wzmacnianiem	352
Prognozowanie przyczynowe	352
Adaptacja domeny	353
Uwagi końcowe	353
 Skorowidz	355