

Содержание

Вступительное слово	10
Об авторе	12
О рецензентах	13
Предисловие	14
Глава 1. Вероятностное мышление	19
Статистика, модели и подход, принятый в этой книге.....	19
Работа с данными	21
Байесовское моделирование	22
Теория вероятностей.....	23
Объяснение смысла вероятностей	23
Определение вероятности	25
Байесовский вывод с одним параметром.....	34
Задача о подбрасывании монеты.....	35
Взаимодействие с байесовским анализом.....	46
Нотация и визуализация модели	46
Обобщение апостериорного распределения.....	47
Проверки апостериорного прогнозируемого распределения.....	49
Резюме	50
Упражнения.....	52
Глава 2. Вероятностное программирование	54
Вероятностное программирование.....	55
Основы использования библиотеки PyMC3	56
Решение задачи о подбрасывании монет с использованием библиотеки PyMC3.....	57
Обобщение апостериорного распределения	59
Решения на основе апостериорного распределения	61
Гауссова модель в подробном изложении	67
Гауссовы статистические выводы	68
Надежные статистические выводы	73
Сравнение групп.....	79
d-мера Коэна.....	81
Вероятность превосходства	82
Набор данных tips.....	82
Иерархические модели	86

Редуцирование.....	91
Еще один пример.....	94
Резюме.....	96
Упражнения.....	99
Глава 3. Моделирование с использованием линейной регрессии.....	101
Простая линейная регрессия	102
Связь с машинным обучением.....	102
Сущность моделей линейной регрессии.....	103
Линейные модели и сильная автокорреляция	108
Интерпретация и визуальное представление апостериорного распределения	111
Коэффициент корреляции Пирсона.....	114
Робастная линейная регрессия	118
Иерархическая линейная регрессия.....	122
Корреляция, причинно-следственная связь и беспорядочность жизни ...	128
Полиномиальная регрессия	130
Интерпретация параметров полиномиальной регрессии.....	131
Является ли полиномиальная регрессия конечной моделью.....	132
Множественная линейная регрессия	133
Спутывающие переменные и избыточные переменные	137
Мультиколлинеарность или слишком сильная корреляция.....	140
Маскировочный эффект переменных.....	144
Добавление взаимодействий.....	146
Дисперсия переменной.....	147
Резюме.....	150
Упражнения.....	151
Глава 4. Обобщение линейных моделей	154
Обобщенные линейные модели	154
Логистическая регрессия	156
Логистическая модель.....	157
Набор данных iris.....	157
Множественная логистическая регрессия	163
Граница решения.....	163
Реализация модели.....	164
Интерпретация коэффициентов логистической регрессии.....	165
Обработка коррелирующих переменных	167
Работа с несбалансированными классами	169
Регрессия с использованием функции softmax.....	171
Дискриминативные и порождающие модели	173
Регрессия Пуассона.....	176
Распределение Пуассона	176
Модель Пуассона с дополнением нулевыми значениями	178

Регрессия Пуассона и модель Пуассона с дополнением нулевыми значениями	179
Робастная логистическая регрессия	181
Модуль GLM	183
Резюме	184
Упражнения	185
Глава 5. Сравнение моделей	188
Проверки прогнозируемого апостериорного распределения	188
Лезвие Оккама – простота и точность	194
Лишние параметры приводят к перепогонке	196
Недостаточное количество параметров приводит к недопогонке	197
Баланс между простотой и точностью	197
Измерения прогнозируемой точности	198
Информационные критерии	200
Логарифмическая функция правдоподобия и отклонение	201
Информационный критерий Акаике	202
Часто применяемый информационный критерий	202
Парето-сглаженная выборка по значимости для перекрестной проверки LOOCV	203
Другие информационные критерии	203
Сравнение моделей с помощью библиотеки РумСЗ	204
Усреднение моделей	207
Коэффициенты Байеса	210
Некоторые дополнительные замечания	212
Коэффициенты Байеса и информационные критерии	216
Регуляризация априорных распределений	220
Более подробно об информационном критерии WAIC	222
Энтропия	222
Расхождение Кульбака–Лейблера	224
Резюме	227
Упражнения	228
Глава 6. Смешанные модели	230
Смешанные модели	231
Конечные смешанные модели	232
Категориальное распределение	234
Распределение Дирихле	235
Неидентифицируемость смешанных моделей	238
Как правильно выбрать число К	241
Смешанные модели и кластеризация	245
Смешанные модели с бесконечной размерностью	246
Процесс Дирихле	246
Непрерывные смешанные модели	253

Биномиальное бета-распределение и отрицательное биномиальное распределение	254
t-распределение Стьюдента.....	255
Резюме.....	255
Упражнения.....	257
Глава 7. Гауссовы процессы.....	258
Линейные модели и нелинейные данные	258
Функции моделирования.....	259
Многомерные гауссовы распределения и функции.....	261
Ковариационные функции и ядра.....	261
Гауссовы процессы	264
Регрессия на основе гауссовых процессов	265
Регрессия с пространственной автокорреляцией	270
Классификация с использованием гауссова процесса	277
Процессы Кокса.....	283
Модель катастроф в угледобывающей промышленности	284
Набор данных redwood	286
Резюме	289
Упражнения.....	289
Глава 8. Механизмы статистического вывода	291
Механизмы статистического вывода	292
Немарковские методы.....	293
Грид-вычисления.....	293
Метод квадратической аппроксимации	296
Вариационные методы	298
Марковские методы.....	301
Метод Монте-Карло.....	303
Цепи Маркова	305
Алгоритм Метрополиса–Гастингса	305
Метод Монте-Карло с использованием механики Гамильтона	310
Последовательный метод Монте-Карло	312
Диагностирование выборок.....	314
Сходимость.....	316
Ошибка метода Монте-Карло	319
Автокорреляция.....	320
Эффективный размер выборки	321
Расхождения	322
Резюме	326
Упражнения.....	326
Глава 9. Что дальше?	328
Предметный указатель.....	332