Объявление

В следующую пятницу Антона не будет, я проведу обе пары (с двух часов дня, как обычно). На первой паре будет контрольная работа, на второй — разбор и еще какие-то примеры. Если все пойдет как надо, то я расскажу как строить доверительные интервалы в “сложных” случаях.

Контрольная работа будет состоять из 5 задач. Первые две задачи — построение оценок и исследование их свойств. Типичный пример: “Построить ОМП-оценку для распределения $U[0, \theta]$ и исследовать ее на несмещеность”. Еще две задачи — доверительные интервалы (для нормальной модели и для биномиальной). Нужно помнить пять формул =). Последняя задача — по теорверу, бонусная. Заменяет любую нерешенную задачу из основного набора. Скорее всего она будет идейная, “на понимание”, а не техническая.

На работе можно пользоваться любыми *пассивными* источниками. Т.е. можно иметь любые конспекты, книги, гуглить, но нельзя спрашивать у соседа или в чатике. Рассчитываю на вашу честность.

Домашнее задание

Для начала, некоторые пожелания по присылаемым задачам:

1. Присылать не только код, но и какие-то мысли по результатам его работы. Удалось ли что-то подтвердить, опровергнуть, что-то не получилось, результаты сравнения такие-то…

2. Прикладывать к коду результат его работы — таблицы и графики. Программы часто работают долго, у меня нет возможности запустить каждую.

На лекции было введено новое понятие — “доверительный интервал” (“confidence interval”, “c.i.”). Напомню, что доверительный интервал — это пара оценок $\{\widehat{\theta}^-, \widehat{\theta}^+)\}$ таких, что интервал $(\widehat{\theta}^-, \widehat{\theta}^+)$ с заданной вероятностью γ накрывает оцениваемый параметр θ. γ называется “уровнем доверия” и обычно выбирается равной 0.95. Если вероятность накрыть истинное значение не равна строго γ, а стремится к ней с ростом объема выборки N, то такой интервал называют асимтотическим. Если используется некоторая неасимтотическая аппроксимация, то такой интервал называют приближенным. Интервал, накрывающий оценку с вероятностью не менее чем заданное γ, называют консервативным.

Замечу, что оценкой (т.е. случайной величиной) здесь является сам интервал, а не оцениваемый параметр θ, поэтому говорить “оцениваемый параметр попадает в интервал с такой-то вероятностью” не совсем корректно (и вводит в заблуждение, параметр это константа).

7.1 Integral C.I.

Продолжим изучать интегрирование методом Монте-Карло. Напомню, что интеграл вычисляется как среднее одиночных оценок, где каждая оценка имеет вид

$$\eta_i = \frac{f(\xi_i)}{\rho_\xi(\xi_i)},$$

где ξ_i — случайная величина с плотностью ρ_ξ (называемой "интегрирующая плотность").

Как мы выяснили ранее, от интегрирующей плотности зависит как порядок, так и скорость сходимости. В частности, оптимальная скорость достигается тогда, когда плотность повторяет поведение модуля интегрируемой функции. Особенно это касается поведения на бесконечности и в окрестности полюсов (особенностей).

Рассмотрим интеграл $\int_0^1 (\sin x)^{-3/4} dx$. Точное его значение можно получить, например, с помощью функции `integrate()` в `R` или `scipy.integrate` в Python. Этот интеграл можно оценить с помощью метода Монте-Карло, используя следующие плотности:

1. Равномерное распределение (можно смотреть на равномерное распределение как на Бета-распределение $\mathcal{B}(1,1)$)

2. Бета-распределение $\mathcal{B}(1/4,1)$

3. Бета-распределение $\mathcal{B}(1/2,1/2)$

Уже было продемонстрировано, что первая плотность дает плохой порядок сходимости (потому что элементарная оценка не имеет дисперсии), вторая и третья дают одинаковый порядок (какой?), но сильно отличающуюся скорость.

Собственно, задание: для второй и третьей плотностей отдельно, нужно промоделировать $N = 10^7$ элементарных оценок и построить график накопленных оценок (накопленных средних). На этом графике построить два типа доверительных интервалов в нормальной модели: используя известную дисперсию элементарной оценки и используя накопленную оценку дисперсии. Уровень доверия γ следует взять 0.95. Также на графиках изобразить в виде горизонтальной линии точное значение интеграла — траектория должна к нему стремиться. *Рекомендую при построении графика при подборе масштаба по игрек-оси игнорировать первые 10% точек, потому что в начале траектория накопленных оценок очень неустойчива и имеет очень большую амплитуду, можно не разглядеть поведение в конце*

Напоминаю, что дисперсия одиночных оценок находится по формуле

$$\mathsf{D}\,\eta = \int \frac{f(x)^2}{\rho_\xi(x)} dx - \left(\int f(x) dx\right)^2.$$

Накопленные средние можно посчитать через накопленные суммы — `cumsum()` в `R` и `numpy.cumsum()` в Python.

7.2 Binomial

Рассмотрим бросание монетки, т.е. испытания Бернулли.

Фиксируем некоторое p — степень гнутости монетки, вероятность выпадения 1, фиксируем некоторое M — число серий бросания и N — число бросков в серии. M разумно взять 1000, а N мы будем перебирать в геометрической прогрессии: $10, 100, \dots, 10^5$. Для каждой серии бросков считаем количество орлов и по нему строим доверительный интервал для p четырьмя разными способами (уровень доверия фиксируем $\gamma = 0.95$).

Хинт. Не обязательно бросать монетку непосредственно и считать орлов. Можно сразу промоделировать биномиальное распределение. Это быстрее

Итак, получили M доверительных интервалов каждого типа для каждой длины выборки N. Посчитаем среднюю ширину доверительного интервала, выборочное стандартное отклонение ширины и реальный эмпирический уровень доверия, т.е. долю тех случаев, когда интервал в самом деле накрыл истинное значение p.

Проделываем все это для $p = 0.5, 0.75, 0.99$. Выводим красиво в табличку и делаем выводы.

Типы доверительных интервалов следующие:

1. Честный-честный доверительный интервал через обращение ФР для биномиального распределения (aka Clopper-Pearson, лучше использовать вид через Бета-квантили);

2. Нормальный асимтотический доверительный интервал с решением квадратного неравентсва (aka Wilson score);

3. Нормальный асимтотический доверительный интервал с простой оценкой дисперсии (aka normal approximation);

4. Консервативный доверительный интервал в нормальной аппроксимации (считаем дисперсию равной 0.25).

Про все эти интервалы можно почитать на википедии. Формулки все там есть.

Конспект

Задача (Задача из вступительной контрольной)**.** Студент получает на экзамене 5 с вероятностью 0.2; 4 с вероятностью 0.4; 3 с вероятностью 0.3 и 2 с вероятностью 0.1. За время обучения он сдает 40 экзаменов. Оценить пределы, в которых с вероятностью 0.95 лежит средний балл студента.

Решение. На первый взгляд, задача выглядит как "задача на доверительный интервал". Это впечатление обманчиво, доверительное оценивание предполагает наличие неизвестного параметра, для которого доверительный интервал строится. Здесь же все известно, нужно просто оценить квантили распределения среднего балла.

Задачу можно решить точно, найдя распределение среднего балла как распределение суммы независимых случайных величин. Удодно воспользоваться для этого аппаратом характеристических функций. Но квантили не выражаются через характеристическую функцию, придется возвращаться к вероятностям отдельных исходов...

Можно решить задачу с помощью моделирования. Просимулировать сюжет задачи много раз, получить выборку из средних баллов, построить доверительный интервал для выборки (например, взяв выборочные квантили). Трудоемкий, но универсальный и популярный на практике подход.

Но проще всего воспользоваться ЦПТ. Речь идет о распределении среднего. 40 — вполне достаточный объем выборки, чтобы пользоваться ЦПТ.

Найдем матожидание и дисперсию распределения ξ_i — оценки за i-й экзамен:

$$\mathsf{E}\xi = \sum_i p_i a_i = 0.2 \cdot 5 + 0.4 \cdot 4 + 0.3 \cdot 3 + 0.1 \cdot 2 = 3.7;$$

$$\mathsf{D}\,\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2 = 0.81 = 0.9^2.$$

Согласно ЦПТ:

$$\frac{\sum_{i=1}^N (\xi_i - 3.7)}{\sqrt{N} \cdot 0.9} \to \mathcal{N}(0, 1),$$

следовательно, $\overline{\xi}$ приближенно распределено нормально $\mathcal{N}(3.7, \sigma = 0.9/\sqrt{40})$. Найдем квантили для этого распределения:

```
> qnorm(c(0.025, 0.975), mean = 3.7, sd = 0.9 / sqrt(40))
[1] 3.421092 3.978908
```