

Про приближенное моделирование произвольных распределений с помощью кубика

Роман Гученко

19 сентября 2020 г.

Встречали ли вы в интернете вопросы на смекалочку типа «как смоделировать бросание шестигранного кубика, если есть семигранный»? В этой курсовой работе предлагается познакомиться с основами теории вероятностей¹, а также с двумя универсальными способами моделирования распределений², методом обратных функций и методом отбора, и ответить на несколько более сложный вопрос: как с помощью бросания n -гранного кубика приближенно смоделировать распределение Барта Симпсона^{3,4}, то есть распределение с плотностью

$$f(x) = 0.5 \phi(x, 0, 1) + 0.1 \sum_{i=0}^4 \phi(x, 0.5 i - 1, 0.1),$$

где $\phi(x, \mu, \sigma)$ — плотность нормального распределения со средним μ и стандартным отклонением σ ? [Ответ на этот вопрос оценивается в 40 баллов для 1го курса и в 20 баллов для 2го.] Сколько потребуется бросков, чтобы функция распределения для смоделированного распределения отличалась от заданной (скажем, в смысле максимума модуля разности между функциями распределения) не более чем на фиксированное ε ? [Тут надо разобрать два случая: с методом обратных функций и с методом отбора; каждый случай — это 30/15 баллов в зависимости от курса]

¹Например, по первым двум главам книги [1]. Если со второй главой, где используется теория меры, возникнут проблемы, то подберем книжку попроще.

²Элементарное изложение которых можно найти в книге [2].

³Смотри книгу [3], страницы 125-126. Распределение так называется потому, что его плотность напоминает прическу соответствующего персонажа из популярного мультсериала.

⁴Ну, или любое другое распределение...

В процессе также (возможно) придется ответить на нижеследующие (наводящие?) вопросы. Как с помощью бросания n -мерного кубика приближенно смоделировать равномерное распределение на отрезке $[0, 1]$? Как такое приближенное моделирование связано с системами счисления? Как равномерное распределение на сетке из равноотстоящих узлов (на равномерной сетке) отличается от равномерного распределения на отрезке (в смысле близости функций распределения)? Как с помощью равномерного распределения на неравномерной сетке приближать произвольное распределение и при чем здесь метод обратных функций? Что делать, если функция распределения не считается аналитически, но очень нужна нам в методе обратных функций? Как двумерные равномерные сетки «портят» идеальный метод отбора?... В общем и целом эта тема для тех, кто хочет достаточно глубоко окунуться в моделирование распределений немного раньше срока, используя только базовый аппарат математического анализа, доступный на младших курсах.

Список литературы

- [1] Ширяев А.Н. Вероятность - 1. Москва : МЦНМО, 2007, 552 стр.
- [2] Rubinstein R.Y., Kroese D.P. Simulation and the Monte Carlo method. Hoboken: Wiley, 2008.
- [3] Wasserman L. All of Nonparametric Statistics. Springer, 2006.