7 Спецкурс "Дополнительные главы теории вероятностей"

Вариант определяется по первой букве имени – А-К отправляется в первый вариант, Л-Я – во второй. Первый вариант получает пункты а), а второй – пункты б).
Решение всех задач со звездочкой добавляет +0.5 балла к сумме оценок за два семестра спецкурса "Дополнительные главы".

7.1 Первый семестр

1. Теорема Севастьянова.
   Ниже описан ряд величин. Построить гистограмму распределения соответствующей величины. Соответствует ли это теореме Севастьянова? Чтобы сопоставить модель теореме Севастьянова – рассмотрите индикаторы того, что в данном месте заканичивается некий удовлетворяющий условию блок.
   Симулировать многократно
   (a) серии из $c2^n$ бросков симметричной монеты и подсчитать число блоков из орлов длины не менее чем n бросков в каждой из серий, $c = 2$, $n = 14$;
   (b) серии розыгрышей из $cn!n!(2n+1)$ н.о.р. случайных величин из $R[0, 1]$ и подсчитать в каждой серии число лесенок длины $2n + 1$ ($a_1 < a_2 < \ldots < a_{n+1} > a_{n+2} > \ldots > a_{2n+1}$), $n = 5$, $c = 2$;
   (c) * проверить в каждом из случаев условия теоремы Севастьянова.
2. Пустые ячейки и высоковероятные слова.
   (a) Многократно симулировать распределение n частиц по n ячейкам, где $n = 50, 100, 200$, в каждом случае подсчитать число пустых ячеек. Произвести нормировку соответственно ЦПТ о размещении частиц по ячейкам и визуально продемонстрировать сходимость к $\mathcal{N}(0, 1)$ распределению.
   (b) Сгенерировать $N = 1000000$ последовательностей слов из случайного алфавита $\{a, b\}$ с вероятностями 0.4 и 0.6 длины $T = 18$ и подсчитать частоты встречаемости различных слов. Какие частоты получили самые вероятные слова? Какая частота у большинства слов?
   (c) * Сравнить результаты с предельными теоремами, описанными в курсе.
3. * Расстояние по вариации.
   Пусть X, Y – случайные величины с распределениями F, G. Задать генератор двух зависимых случайных величин X', Y' с ф.р. $F_{1,1}$, $F_{1,2}$, для которых $\rho(X, Y) = P(X' \neq Y')$, где ρ – расстояние по вариации.
   (a) $F = Bern(1/2)$, $G = Bern(1/3)$.
   (b) $F = R[0, 1]$, $G = x^2 I_{[0,1)} + I_{x>1}$
4. Теорема Линдеберга.
   Моделировать данные $X_i \sim P_i$, $i \leq n$. Исследовать предельное поведение $(S_n - \mathbf{E}S_n)/\sqrt{DS_n}$.
   (a) а) $P_n(n) = P_n(-n) = 1/(2n^2)$, $P_n(0) = 1 - 1/n^2$; б) $P_n(n) = P_n(-n) = 1/4$, $P_n(0) = 1/2$.
   (b) а) $P_n(\sqrt{n}) = P_n(-\sqrt{n}) = 1/2$, если n – полный квадрат, $P_n(1) = P_n(-1) = 1/2$ иначе, б) $P_n(n) = P_n(-n) = 1/(2\sqrt{n})$, $P_n(0) = 1 - 1/\sqrt{n}$.
   (c) * Проверить во всех случаях выполнение теоремы Линдеберга и сравнить полученные результаты с моделированием.
5. Безгранично делимые распределения.
   Построить предложенные ниже схемы серий и в каждой из них эмпирически определить предельное безграничное делимое распределение:
   (a) $X_{n,i} \sim Geom(1 - 1/n)$.

(b)

$$X_{n,i} = \begin{cases} \frac{(1-1/i)}{\sqrt{\ln n}}, & \frac{1}{i} \\ -\frac{1}{i\sqrt{\ln n}}, & 1-\frac{1}{i} \end{cases}$$

(c) $X_{n,i} \sim NegBinom(1/n, 1/n)/n$.