Подведение итогов 13 Колмогоровской студенческой олимпиады по теории вероятностей

Москва,
23 апреля 2014

Общая информация

Всего участников — 37
Младшая возрастная категория — 28, старшая — 9
Юношей — 32, девушек — 5

Университеты, факультеты и кафедры

Кафедра \ Курс	1	2	3	4	5	всего
ММ МГУ, без кафедры	1	19	-	-	-	20
ТВ	-	-	0	1	0	1
ВысАлг	-	-	1	0	0	1
МатАн	-	-	1	0	0	1
МатСтат	-	-	1	0	0	1
МатЛог	-	-	1	0	0	1
ММ МГУ, всего	1	19	4	1	0	25
ВМК МГУ	0	5	1	0	0	6
СПбГУ	0	0	3	1	0	4
МФТИ	0	0	0	0	1	1
РЭШ	0	0	0	0	1	1
Всего	1	24	8	2	2	56

Всего предлагалось 8 задач в младшей возрастной категории и 9 задач в старшей. Для каждой задачи ниже приведено число ее решивших.

Задача 1

Задача 1 (18,5). Пусть X и Y — н.о.р. неотрицательные случайные величины, каждая из которых принимает два значения, $\mathsf{E}X = 1$. Найти минимальное значение, которое может принимать вероятность $\mathsf{P}(X + Y < 3)$.
Решение. Пусть $\mathsf{P}(X = a) = p$, $\mathsf{P}(X = b) = 1 - p$ и $a \leq b$. Очевидно, $\mathsf{E}X = ap + b(1 - p) = 1$, $a \leq 1$ и имеет смысл рассматривать только $b \geq 3/2$ (в противном случае рассматриваемая вероятность равна 1). Возможны два случая:
1 случай: $a + b < 3$. Тогда рассматриваемая вероятность равна

$$\mathsf{P}(X = Y = a) + 2\mathsf{P}(X = a, Y = b) = 1 - (1 - p)^2.$$

Так как по условию $p = (b - 1)/(b - a)$, то указанное выражение возрастает по a и b, т.е. его нижняя грань достигается при $a = 0$ и $b = 3/2$. При этом $p = 1/3$ и $1 - (1 - p)^2 = 5/9$.
2 случай: $a + b \geq 3$. Тогда рассматриваемая вероятность равна

$$\mathsf{P}(X = Y = a) = p^2 = \frac{(b - 1)^2}{(b - a)^2}.$$

По той же причине нижняя грань достигается при $b = 3$ и $a = 0$, и тогда $p^2 = 4/9 < 5/9$, так что ответ $4/9$.

Задача 2

Задача 2 (11). Имеется калькулятор, позволяющий вычислять функцию $\Phi(t) = \mathrm{P}(Z \le t)$, где $Z \sim N(0,1)$ (а также все элементарные функции). Найти $\int_{\mathbb{R}} (2\pi)^{-1/2} e^{-t^2/2} \Phi(t+1) dt$.

Решение. Имеем

$$\int_{\mathbb{R}} (2\pi)^{-1/2} e^{-t^2/2} \Phi(t+1) dt = \frac{1}{2\pi} \int_{\mathbb{R}} \int_{-\infty}^{t+1} e^{-t^2/2 - y^2/2} dy\, dt =$$

$$= \frac{1}{2\pi} \iint_{y-t \le 1} e^{-t^2/2 - y^2/2} dy\, dt = \mathrm{P}(X - Y \le 1),$$

где $X, Y \sim N(0,1)$ и независимы. Так как $X - Y \sim N(0,2)$, то эта вероятность равна $\mathrm{P}((X-Y)/\sqrt{2} \le 1/\sqrt{2}) = \Phi(1/\sqrt{2})$.

Задача 3

Задача 3 (6,5). За бесконечным (в обе стороны) столом сидит счетное число школьников, решающих задачу. Для каждого школьника вероятность решить ее самостоятельно равна $1/2$. Также с вероятностью $1/4$ ему удастся подсмотреть в тетрадь соседа слева, и с вероятностью $1/4$ — в тетрадь соседа справа. Все эти события (для всех школьников) независимы между собой. Если кто-то из школьников получил решение задачи, то каждый подсмотревший к нему в тетрадь также его получает. Найти вероятность, что сидящий в ряду школьник Вася получит решение задачи.

Решение. Сначала пусть стол бесконечен в одну сторону, и пусть q — вероятность, что школьник, сидящий с краю, получит решение. С вероятностью $1/2$ он решит задачу сам, а с вероятностью $1/2 \cdot 1/4$ ему удастся заглянуть в тетрадь единственного соседа. Так как для этого соседа вероятность получить решение, не пользуясь помощью крайнего школьника, тоже равна q, то $q = 1/2 + q/8 \Rightarrow q = 4/7$. Далее, у Васи также есть возможность решить задачу самому (вероятность этого $1/2$). Пусть A и B — события, состоящие в том, что Вася не решил задачу сам, но смог подсмотреть у соседа соответственно справа и слева. Тогда $P(A) = P(B) = 1/2 \cdot 1/4 \cdot q = 1/14$, $\mathrm{P}(AB) = 1/2 \cdot 1/4 \cdot 1/4 \cdot q^2 = 1/98$. В итоге вероятность равна $1/2 + 2P(A) - P(AB) = 31/49$.

Задача 4

Задача 4 (2). Пусть неотрицательная случайная величина X такова, что $Law(X) = Law(1 + UX)$, где U равномерно распределена на $[0, 1]$ и не зависит от X. Доказать, что дисперсия X конечна.
Решение. Пусть $U_1, U_2, \ldots$ — последовательность н.о.р. равномерных на $[0, 1]$ случайных величин, не зависящая от X. Тогда

$$Law(X) = Law(1+U_1X) = Law(1+U_1+U_1U_2X) = Law(1+U_1+U_1U_2+U_1U_2U_3X) = \ldots.$$

Так как $U_1 \ldots U_n \to 0$ п.н. при $n \to \infty$, то получается, что $X - 1$ распределена как сумма

$$\sum_{n=1}^{\infty} \prod_{i=1}^{n} U_i.$$

Второй момент этой суммы имеет вид

$$\sum_{n,m=1}^{\infty} \mathrm{E} \prod_{i=1}^{n} U_i \prod_{j=1}^{m} U_j \le 2 \sum_{n=1}^{\infty} \sum_{m=n}^{\infty} \mathrm{E} \prod_{i=1}^{n} U_i^2 \prod_{j=n+1}^{m} U_j = 2 \sum_{n=1}^{\infty} \sum_{m=n}^{\infty} 3^{-n} 2^{-(m-n)} = 4 \sum_{n=1}^{\infty} 3^{-n} < \infty.$$

Задача 4, продолжение

Замечание. Второе решение состоит в исследовании характеристической функции X. По теореме Фубини имеем

$$\varphi_X(t) = e^{it} \int_0^1 \varphi_X(tu)du = \frac{e^{it}}{t} \int_0^t \varphi_X(s)ds,$$

откуда следует, что

$$\varphi_X'(t) = \varphi_X(t) \left(\frac{e^{it} - 1}{t} + i \right).$$

Исследуя решение этого дифференциального уравнения, несложно показать, что оно дважды дифференцируемо, так что дисперсия X конечна. Более того, решение несложно выписать через степенные ряды:

$$\varphi_X(t) = \exp\left\{ it + \sum_{k=1}^{\infty} \frac{(it)^k}{k!k} \right\}.$$

Отсюда следует, что φ_X аналитична на всей числовой оси, и, следовательно, у X существуют все экспоненциальные моменты.

Задача 5

Задача 5 (6,5). Вы имеете право сыграть в такую игру. Есть несимметричная монета, которая выпадает гербом с вероятностью $1/3$. Ее бросают до тех пор, пока не выпадет третий герб. После каждого из гербов вы можете прекратить игру и получить столько тугриков, сколько испытаний потребовалось для получения этого герба (не считая предыдущий герб, если он был, но считая текущий). Как следует играть, чтобы математическое ожидание выигрыша было максимально возможным, и чему оно равно?

Решение. Прежде всего докажем, что если X и Y — независимые $\mathbb{N}$-значные случайные величины, и мы сначала разыгрываем X, а потом либо забираем X тугриков, либо разыгрываем Y, то оптимальная тактика следующая: надо остановиться, если мы уже получили $\mathrm{E}Y$ или больше, и продолжать в противном случае.

В самом деле, пусть $C \subset \mathbb{N}$ — те значения X, при выпадении которых мы продолжим игру. Тогда математическое ожидание выигрыша равно

$$\sum_{k \in C} (\mathrm{E}Y)\mathrm{P}(X = k) + \sum_{k \notin C} k\mathrm{P}(X = k).$$

Отсюда видно, что это выражение максимально, если

$$k \in C \Leftrightarrow \mathrm{E}Y > k.$$

Задача 5, продолжение

Пусть ξ_1, ξ_2, ξ_3 — это 3 независимые геометрические случайные величины с параметром $1/3$, равные длинам первой, второй и третьей серий соответственно. Сначала предположим, что мы уже продолжаем игру после первого герба. Согласно приведенному соображению, нам следует остановиться на втором гербе, если $\xi_2 \geq \mathsf{E}\xi_3 = 3$, и продолжить игру в противном случае. При этом наш средний выигрыш равен

$$\sum_{k=1}^{2} 3\mathsf{P}(\xi_2 = k) + \sum_{k=3}^{\infty} k\mathsf{P}(\xi_2 = k) = \sum_{k=1}^{\infty} k\mathsf{P}(\xi_2 = k) + 2\mathsf{P}(\xi_2 = 1) + \mathsf{P}(\xi_2 = 2) =$$

$$= 3 + 2\frac{1}{3} + \frac{2}{3}\frac{1}{3} = \frac{35}{9}.$$

Это число меньше 4, поэтому на первом шаге нам нужно остановиться, если $\xi_1 \geq 4$, и продолжить игру в противном случае. При этом наш средний выигрыш равен

$$\sum_{k=1}^{3} \frac{35}{9}\mathsf{P}(\xi_1 = k) + \sum_{k=4}^{\infty} k\mathsf{P}(\xi_1 = k) = \sum_{k=1}^{\infty} k\mathsf{P}(\xi_1 = k) + \frac{26}{9}\mathsf{P}(\xi_1 = 1) +$$

$$+ \frac{17}{9}\mathsf{P}(\xi_1 = 2) + \frac{8}{9}\mathsf{P}(\xi_1 = 3) = 3 + \frac{26}{9}\frac{1}{3} + \frac{17}{9}\frac{2}{9} + \frac{8}{9}\frac{4}{27} == 4 + \frac{125}{243} \approx 4,51.$$

Задача 6

Задача 6 (0). Фильтрация $\{\mathcal{F}_s, s \geq 0\}$ называется непрерывной справа, если $\cap_{s>t}\mathcal{F}_s = \mathcal{F}_t$ при любом $t \geq 0$. Верно ли, что естественная фильтрация случайного процесса с непрерывными траекториями непрерывна справа?
Решение. Неверно. Рассмотреть процесс $X(t) = \xi t$, $t \geq 0$, здесь $\xi \sim N(0,1)$. Тогда $\mathcal{F}_0 = \{\varnothing, \Omega\}$, но при любом $s > 0$ имеем $\mathcal{F}_s = \sigma(\xi)$.

Задача 7

Задача 7. Для случайной величины X, принимающей целые неотрицательные значения, введем функционал $H(X)=\sum_{n=0}^{\infty} p_n \ln p_n$ (здесь $p_n = \mathrm{P}(X=n)$ и считается, что $0\ln 0 = 0$). а) (10) Доказать, что для независимых X и Y мы имеем $H(X+Y) \geq H(X)+H(Y)$; б) (6) при каком дополнительном условии достигается равенство?

Решение. а) Имеем

$$H(X+Y)=\sum_{n=0}^{\infty} \mathrm{P}(X+Y=n)\ln \mathrm{P}(X+Y=n)=$$

$$\sum_{n=0}^{\infty}\sum_{k=0}^{n} \mathrm{P}(X=k, Y=n-k)\ln \sum_{j=0}^{n} \mathrm{P}(X=j, Y=n-j)=$$

$$=\sum_{n=0}^{\infty}\sum_{k=0}^{n} \mathrm{P}(X=k)\mathrm{P}(Y=n-k)\ln \sum_{j=0}^{n} \mathrm{P}(X=j)\mathrm{P}(Y=n-j)\geq$$

$$\geq \sum_{n=0}^{\infty}\sum_{k=0}^{n} \mathrm{P}(X=k)\mathrm{P}(Y=n-k)\ln \mathrm{P}(X=k)\mathrm{P}(Y=n-k)=$$

Задача 7, продолжение

$$= \sum_{n=0}^{\infty} \sum_{k=0}^{n} P(X = k)P(Y = n-k) \ln P(X = k) +$$

$$+ \sum_{n=0}^{\infty} \sum_{k=0}^{n} P(X = k)P(Y = n-k) \ln P(Y = n-k) =$$

$$= \sum_{k=0}^{\infty} P(X = k) \ln P(X = k) \sum_{n=k}^{\infty} P(Y = n-k) +$$

$$+ \sum_{k=0}^{n} P(X = k) \sum_{n=k}^{\infty} P(Y = n-k) \ln P(Y = n-k) = H(X) + H(Y).$$

б) Равенство достигается тогда, когда в приведенном выше решении использованная оценка

$$\sum_{j=0}^{n} P(X = j, Y = n-j) \geq P(X = k, Y = n-k)$$

обращается в равенство, т.е. когда не существует таких $k, k_1, m, m_1 \in \mathbb{Z}_+$, что

$$k + m = k_1 + m_1 \;\; \text{и} \;\; P(X = k)P(Y = m)P(X = k_1)P(Y = m_1) > 0.$$

Задача 7, продолжение

Комментарий С.А.Молчанова. Выражение $-H(X)$ называется энтропией случайной величины X. Для абсолютно непрерывных случайных величин ее обычно полагают равной

$$-\int_{\mathbb{R}} p(x) \ln p(x) dx.$$

А.Н.Колмогоров подчеркивал, что энтропия непрерывной случайной величины равна ∞ и приведенное выражение — это только ее ренормализация, которая теряет многие важные свойства истинной энтропии.

Задача 8

Задача 8 (4,5). Найти $\lim_{n\to\infty} \sum_{k=n}^{3n} C_{k-1}^{n-1}(1/3)^n(2/3)^{k-n}$.
Решение. Под знаком предела стоит вероятность того, что в схеме Бернулли с вероятностью успеха $1/3$ успех номер n случился не позже испытания номер $3n$. Другими словами, если ввести независимые случайные величины $\xi_1, \xi_2, \ldots$, имеющие геометрическое распределение с параметром $1/3$, то рассматривается вероятность

$$\mathrm{P}\left(\sum_{i=1}^{n} \xi_i \leq 3n\right) = \mathrm{P}\left(\sum_{i=1}^{n}(\xi_i - \mathrm{E}\xi_i) \leq 0\right) = \mathrm{P}\left(\frac{\sum_{i=1}^{n}(\xi_i - \mathrm{E}\xi_i)}{\sqrt{n\mathrm{D}\xi_i}} \leq 0\right),$$

которая по центральной предельной теореме стремится к $1/2$.