Домашняя работа 2 от d2bCvlM (дедлайн – 15:00 3.10.19)

October 10, 2019

Задача 1 (3 балла)

Пусть $X_1, \dots, X_n$ – независимые одинаково распределенные случайные величины (н.о.р.с.в.) с функцией распределения $F(x)$. Найти функции распределения $\max\limits_{1 \le i \le n} X_i$ и $\min\limits_{1 \le i \le n} X_i$.

Решение

Для максимума: $Y = maxX_i$,

$$F_y(x) = P(Y < x) = P(maxX_i < x) = P(X_i < x) = F_x(x)^n$$

Для минимума:

$$F_y(x) = P(minX_i < x) = 1 - P(minX_i > x) = 1 - P(X_i > x) = 1 - (1 - F_x(x))^n$$

Задача 2 (3 балла) Случайная величина ξ имеет функцию распределения $F(x)$. Найти функцию распределения случайной величины $\frac{1}{2}(\xi + |\xi|)$

Решение

$$F_\xi(x) = \mathbb{P}(\xi < x)$$

$$\eta = \frac{1}{2}(\xi + |\xi|)$$

$$F_\eta(x) = \mathbb{P}(\eta < x) = \mathbb{P}(\frac{1}{2}(\xi + |\xi|) < x) = \mathbb{P}(\xi < 0) \cdot \mathbb{P}(0 < x) + \mathbb{P}(\xi \ge 0) \cdot \mathbb{P}(\xi < x) =$$

$$= F_\xi(0)\mathbb{P}(0 < x) + (1 - F_\xi(0))F_\xi(x)$$

Не совсем ясно что делать с $\mathbb{P}(0 < x)$, но в целом ответ получен разумно.

Задача 3 (3 балла)

В круглой комнате произвольным образом провели диаметр и в одном из концов этого диаметра поставили прожектор так, чтобы он мог светить внутрь комнаты (направление, в котором светит прожектор задаётся углом α, который отсчитывается от направления проведённого диаметра; $\alpha \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ – отрицательный угол отсчитывается по часовой стрелке от направления диаметра, положительный – против часовой). Найти функцию распределения длины луча от прожектора до стены, если угол α – это равномерно распределённая случайная величина на отрезке $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

Решение

$l = d \cdot \cos\alpha$ Задача симметричная, из за этого в конце возникает 2, но потом исчезет из за того, как именно мы берем углы

$$F_l(x) = \mathbb{P}(l < x) = \mathbb{P}(d\cos\alpha < x) = \mathbb{P}(\cos\alpha < \frac{x}{d}) = 2\mathbb{P}(\alpha < \arccos\frac{x}{d}) = \frac{\frac{\pi}{2} - \arccos\frac{x}{d}}{\frac{\pi}{2}}$$

Задача 4 (4 балла)

Допустим, что вероятность столкновения молекулы с другими молекулами в промежутке времени $[t, t+\Delta t)$ равна $p = \lambda\Delta t + o(\Delta t)$ и не зависит от времени, прошедшего после предыдущего столкновения ($\lambda = const$). Найти распределение времени свободного пробега молекулы (показательное распределение) и вероятность того, что это время превысит заданную величину $t^\star$

Решение Вдохновило: $http://online.mephi.ru/courses/physics/molecular_physics/data/course/4/4.1.1.html$
По ссылке не совсем то что нам надо, но есть интересные мысли. $p(\Delta t) = \Delta\lambda + o(\Delta)$
Найдем распределение вероятности пролететь дельта t без столкновений, тогда 1 - p будет ответом к нашей задаче. $P(x+dx)$ - вероятность пролтетье ч + dx без столкновений, в силу малости величин: $P(x+dx) = P(x) + \frac{dP}{dx}dx$ - вероятность пролететь x+dx без столновение , это же равно (введем $\alpha = \frac{1}{\lambda}$, тогда

$$p(\Delta t) = \Delta/\alpha + o(\Delta)$$

$$P(x) + \frac{dP}{dx}dx = P(x) \cdot (1 - dx/\alpha - o(dx)) = P(x) \cdot (1 - dx/\alpha)$$

Откуда получаем уравнение на альфа и на лямбда соотвественно:

$$\frac{\frac{dP}{dx}}{P(x)} = -1/\alpha$$

Откуда

$$P(x) = e^{-\frac{x}{\alpha}} = e^{-x\lambda}$$

наша искомое распределение вероятностей F(x) = 1 - P(x)

$$F(x) = 1 - e^{-x\lambda}$$

Соовественно вероятность превышения равна в таком случае

$$F'(x) = \lambda e^{-x\lambda}$$

Неизвестно что попытка доказать таким образом успешна, но вроде как норм.

Задача 5 (2 балла)

Диаметр круга измерен приближенно. Считая, что его величина равномерно распределена в отрезке $[a, b]$, найти распределение площади круга, её среднее значение и дисперсию.

Решение

$$\mathbb{E}S = \mathbb{E}\frac{\pi d^2}{4} = \frac{\pi}{4}\mathbb{E}d^2 = \frac{\pi}{4}\int_{-\infty}^{+\infty} x^2 f(x)dx = \frac{\pi}{4}\int_a^b x^2 \frac{1}{b-a}dx = \frac{\pi}{4(b-a)}\frac{x^3}{3}\Big|_a^b = \frac{\pi}{12}(a^2 + ab + b^2)$$

$$\mathbb{D}S = \mathbb{E}S^2 - (\mathbb{E}S)^2$$

$$\mathbb{E}S^2 = \mathbb{E}\frac{\pi^2 d^4}{16} = \frac{\pi^2}{16}\int_{-\infty}^{+\infty} x^4 f(x)dx = \frac{\pi^2}{16(b-a)}\int_a^b x^4 dx = \frac{\pi^2}{16(b-a)}\frac{x^5}{5}\Big|_a^b = \frac{\pi^2(b^5 - a^5)}{80(b-a)}$$

$$\mathbb{D}S = \frac{\pi^2(b^5 - a^5)}{80(b-a)} - \frac{\pi^2(b^3 - a^3)^2}{144(b-a)^2}$$

Задача 6 (2 балла)

Пусть случайные величины ξ и η независимы и $\mathbb{E}\xi = 1$, $\mathbb{E}\eta = 2$, $\mathbb{D}\xi = 1$, $\mathbb{D}\eta = 4$. Найти математические ожидания случайных величин: **Решение** а) $\xi^2 + 2\eta^2 - \xi\eta - 4\xi + \eta + 4$;

$$\mathbb{E}(\xi^2 + 2\eta^2 - \xi\eta - 4\xi + \eta + 4) = \mathbb{E}\xi^2 + 2\mathbb{E}\eta^2 - \mathbb{E}\xi\eta - 4\mathbb{E}\xi + \mathbb{E}\eta + \mathbb{E}4$$

- $\mathbb{E}\xi^2 = \mathbb{D}\xi + (\mathbb{E}\xi)^2 = 1 + 1 = 2$
- $\mathbb{E}\eta^2 = \mathbb{D}\eta + (\mathbb{E}\eta)^2 = 4 + 2^2 = 8$
- $\mathbb{E}\xi\eta = \mathbb{E}\xi \cdot \mathbb{E}\eta = 2$ в силу независимости

Тогда ответ:

$$\mathbb{E}(\xi^2 + 2\eta^2 - \xi\eta - 4\xi + \eta + 4) = \mathbb{E}\xi^2 + 2\mathbb{E}\eta^2 - \mathbb{E}\xi\eta - 4\mathbb{E}\xi + \mathbb{E}\eta + \mathbb{E}4 = 2 + 2 \cdot 8 - 2 - 4 \cdot 1 + 2 + 4 = 18$$

б) $(\xi + \eta + 1)^2$

$$\mathbb{E}(\xi + \eta + 1)^2 = \mathbb{D}(\xi + \eta + 1) + (\mathbb{E}(\xi + \eta + 1))^2$$

- $\mathbb{D}(\xi+\eta+1) = \mathbb{D}\xi + \mathbb{D}\eta = 1 + 4 = 5$
- $\mathbb{E}(\xi+\eta+1) = \mathbb{E}\eta + \mathbb{E}\xi + 1 = 1 + 2 + 1 = 4$

  $\mathbb{E}(\xi+\eta+1)^2 = \mathbb{D}(\xi+\eta+1) + (\mathbb{E}(\xi+\eta+1))^2 = 5 + 4^2 = 21$

Задача 7 (3 балла)

а) Пусть ξ – положительная невырожденная случайная величина с конечным математическим ожиданием. Доказать, что

$$\frac{1}{\mathbb{E}\xi} \le \mathbb{E}\frac{1}{\xi}$$

Решение Воспользуемся неравенством Йенсена:

$$f(\mathbb{E}\xi) \le \mathbb{E}f(\xi), f \in CONVEX$$

В нашем случае $f = \frac{1}{x}$, то есть $f(\mathbb{E}\xi) = \frac{1}{\mathbb{E}\xi}$, $\mathbb{E}f(\xi) = \mathbb{E}\frac{1}{\xi}$ Критерий выпуклости функции - f выпукла на C, если для любых $x, y : x, y \in C$:

$$f(\frac{x+y}{2}) \le \frac{f(x)+f(y)}{2}$$

$$\frac{1}{\frac{x+y}{2}} \le \frac{\frac{1}{x}+\frac{1}{y}}{2} \iff \frac{2}{x+y} \le \frac{x+y}{2xy} \iff 4xy \le (x+y)^2 \iff 2xy \le x^2+y^2 \iff 0 \le (x-y)^2$$

Доказали выпуклость, теперь по неравенству Йенсена получаем доказательство утверждения.

б) Пусть ξ и η – независимые положительные случайная величины, с конечным математическим ожиданием. Доказать, что

$$\mathbb{E}\Big(\frac{\xi}{\eta}\Big)^r \ge \frac{\mathbb{E}\xi^r}{\mathbb{E}\eta^r}$$

Для независимых величин справедливо

$$\mathbb{E}ab = \mathbb{E}a \cdot \mathbb{E}b$$

$$a = \xi^r, \quad b = 1/\eta^r$$

$$\mathbb{E}(\frac{\xi}{\eta})^r = \mathbb{E}\xi^r \cdot \mathbb{E}\frac{1}{\eta^r}$$

В предыдущем пункте доказали, что

$$\mathbb{E}\frac{1}{\eta^r} \ge \frac{1}{\mathbb{E}\eta^r}$$

ПОдстановка:

$$\mathbb{E}(\frac{\xi}{\eta})^r = \mathbb{E}\xi^r \cdot \mathbb{E}\frac{1}{\eta^r} \ge \mathbb{E}\xi^r \cdot \frac{1}{\mathbb{E}\eta^r}$$

ЧТД

Задача 8 (5 баллов)

На небольшом кластере GPU прямо сейчас очередь на обучение из 30 нейросетей. Единовременно на кластере может обучаться только одна сеть. За каждую нейросеть отвечают разные ML-инженеры. Нейросети имеют разные свойства: 10 из них больших (время их обучения 15 часов) и 20 маленьких (время их обучения 1 час). Пока не наступил момент начала обучения, разработчик переживает и внимательно следит за очередью, бесполезно растрачивая время. Посчитайте математическое ожидание, сколько человеко-часов будет потрачено на переживания разработчиков ровно с текущего момента (кластер освободился и начинает обрабатывать очередь, описанную выше), если задачи в очереди расположены в случайном порядке

Решение

Я могу ошибиться в арифметике, но основной посыл - сначала раскидать маленькие сетки , а потом докидывать большие, которые будут увеличивать время сразу у группы саентологов.

Разобьем задачу на две подзадачи. Сначала представим что у нас всего 20 машинистов и 20 сетей по одному часу соотвественно, тогда по определению матождания их суммарного ожидания(по факту мы его знаем, ведь это просто постоянная сумма):

$$\mathbb{E}T_{SmallNet_{all}} = 20 \cdot \sum_{i=1}^{20} \frac{1}{20} \cdot (i-1) = 20 \cdot \sum_{i=0}^{19} \frac{1}{20} \cdot i$$

i-1 потому что если он учит третьим, то ждать будет 2 часа.
Теперь будем как то докидывать большие сетки в очередь и смотреть как увеличивается время ожидания у инженеров, которые оказались в очереди после вставленной. Можем так делать в силу линейности матожидания. По факту у нас есть какие то эффективные раскидки, а есть не очень, поэтому запишем почестному. Пусть у нас сейчас уже вставлено j больших нейронок, смотрим сколько докинет новая.

$$\mathbb{E}T_{net_j} = \sum_{i=0}^{i=20+j} \frac{1}{21+j} 15i$$

Обьяснимся - если у нас сейчас уже 20 + j нейронок и инженеров в очереди, то у нас 20+j+1 место для новой нейронки соотвественное. Теперь обратимся к 15i. Пусть наша нейронка встала на последнее место, тогда время ожидания не увеличится, это случай $i = 0$. Или же наша сетка наоборот встала на нулевое место, тогда время ожидания увеличилось на 15 часов для каждого из 20 + j инженеров на 15 часов, получается $15(20 + j)$. То есть наше i пробегает от 0 до 20+j.
Так как у нас 10 больших нейронок, нужно сделать еще одну сумму

$$\mathbb{E}T_{BigNet_{all}} = \sum_{j=1}^{j=10} \sum_{i=0}^{i=20+j} \frac{1}{21+j} 15i$$

Тогда ответ:

$$\mathbb{E}T = \mathbb{E}T_{SmallNet_{all}} + \mathbb{E}T_{BigNet_{all}} = \sum_{i=0}^{19} i + \sum_{j=1}^{j=10} \sum_{i=0}^{i=20+j} \frac{1}{21+j} 15i = \frac{19}{2} \cdot 20 + \sum_{j=1}^{j=10} \frac{15}{21+j} \sum_{i=0}^{i=20+j} i$$

$$\mathbb{E}T = \frac{380}{2} + \sum_{j=1}^{j=10} \frac{15}{21+j} \cdot \frac{20+j}{2} \cdot (21+j) = \frac{380}{2} + 15 \sum_{j=1}^{j=10} \frac{20+j}{2} = \frac{380}{2} + \frac{15}{2} \cdot \frac{21+30}{2} \cdot 10$$

$$\mathbb{E}T = \frac{380}{2} + \frac{75 \cdot 51}{2} = 2102.5$$