Точната формула за оценка се формира в зависимост от резултатите. За приблизителна, може да използвате 2 + брой точки. Време за работа: 3 часа. Успех.

С m, n, k ще бележим неотрицателни цели числа.

Задача 1. Нека X е случайна величина с функция на разпределение

$$F_X(x) = \begin{cases} 0, & \text{за } x \leq 0; \\ ax^2 + bx, & \text{за } 0 < x \leq 1; \\ 1, & \text{за } x > 1. \end{cases}$$

1. (0.25 т.) Отговорете дали X е непрекъсната случайна величина, като изведете нейната плътност в случай, че смятате, че е такава.
2. (0.25 т.) Намерете константите a и b, ако знаете в допълнение, че $\mathbb{P}(X > 0.75) = 0.5$.
3. (0.25 т.) Ако очакването на случайна величина Y съществува, то за него е вярно, че

$$\mathbb{E}[Y] = \int_0^\infty \mathbb{P}(Y > t)\mathbb{D}t - \int_{-\infty}^0 \mathbb{P}(Y \leq t)\mathbb{D}t.$$

Докажете това равенство, ако знаете, че Y е непрекъсната случайна величина с плътност f_Y.

4. (0.25 т.) Използвайте горната формула, за да намерите очакването на X.

Решение 1.

1. Да, непрекъсната е, понеже приема неизброимо безкрайно много стойности. Освен това е абсолютно непрекъсната, понеже дадената функция на разпределение е диференцируема:

$$f_X(x) = \frac{d}{dx}F_X(x) = \begin{cases} 0, & \text{за } x \leq 0; \\ 2ax + b, & \text{за } 0 < x < 1; \\ 0, & \text{за } x > 1. \end{cases} = (2ax + b)\mathbb{1}_{\{0<x<1\}}$$

2. Допълнително ни е дадено, че $\mathbb{P}(X > 0.75) = 1 - F_X(0.75) = 0.5$. Отделно от дефиницията горе знаем, че например $1 = \mathbb{P}(X < 1) = F_X(1) = a + b$. Решаваме системата:

$$\left| \begin{array}{l} 1 - F_X\left(\frac{3}{4}\right) = \frac{1}{2} \\ F_X(1) = 1 \end{array} \right. \Leftrightarrow \left| \begin{array}{l} F_X\left(\frac{3}{4}\right) = \frac{1}{2} \\ F_X(1) = 1 \end{array} \right. \Leftrightarrow \left| \begin{array}{l} \frac{9}{16}a + \frac{3}{4}b = \frac{1}{2} \\ a + b = 1 \end{array} \right.$$

И тъй $(a, b) = \left(\frac{1}{6}, \frac{5}{6}\right)$

3. Стандартно, търсената алтернативна формулировка на очакването можем да получим като просто разменим реда на интегриране. Ще покажем, че дясната страна е равна на лявата. Да разгледаме първото лявото събираемо:

$$\int_0^\infty \mathbb{P}(Y > t)dt = \int_0^\infty \left(\int_t^\infty f_Y(y)dy\right)dt = \int\limits_{\{0<y<t\}} f_Y(y)d(t, y) = \int_0^\infty \left(\int_0^y f_Y(y)dt\right)dy =$$

$$\int_0^\infty \left(\int_0^y dt\right) f_Y(y)dy = \int_0^\infty y f_Y(y)dy$$

Аналогично получаваме и за второто събираемо, че:

$$-\int_{-\infty}^0 \mathbb{P}(Y \leq t)dt = \int_{-\infty}^0 y f_Y(y)dy$$

С което доказателството е завършено.

Формално: използвахме теоремата на Фубини, която ни позволява да разменяме реда на интегриране както ни е удобно, в случай, че интегралът върху областта е абсолютно сходящ. Помислете защо можем да го правим тук.

4.

$$\begin{aligned}
\mathbb{E}[X] &= \int_0^\infty \mathbb{P}(X > t)dt - \int_{-\infty}^0 \mathbb{P}(X < t)dt \\
&= \int_0^\infty (1 - F_X(t))dt - \underbrace{\int_{-\infty}^0 F_X(t)dt}_{=0} \\
&= \int_0^1 (1 - ax^2 - bx)dx + \underbrace{\int_1^\infty (1 - F_X(x))dx}_{=0} \\
&= 1 - \frac{a}{3} - \frac{b}{2} \\
&= \frac{19}{36}
\end{aligned}$$

Ако не си вярваме на сметките, можем и да си направим проверка дали стойността съвпада с конвенцоналния подход: $\int_0^1 x f_X(x)dx$.

Задача 2. Нека $U_1, U_2, U_3 \sim U(0,1)$ са независими. Нека $U_{(1)}, U_{(2)}, U_{(3)}$ са наредените в нарастващ ред U_1, U_2 и U_3.

Например, ако дадена реализация е $U_1 = 0.23$, $U_2 = 0.88$, $U_3 = 0.1$, *то* $U_{(1)} = 0.1$, $U_{(2)} = 0.23$, $U_{(3)} = 0.88$.

1. (0.25 т.) Намерете плътностите и очакванията на $U_{(1)} := \min\{U_1, U_2, U_3\}$ и $U_{(3)} := \max\{U_1, U_2, U_3\}$.
2. (0.25 т.) Вярно ли е, че $U_{(2)} \sim U(0,1)$. Ако да, го докажете, а ако не - намерете разпределението му.
3. (0.25 т.) Намерете дисперсията на $X := -\ln(U_1)$.

Нека $X_1, X_2, X_3, X_4 \sim Exp(1)$ са независими.

4. (0.25 т.) Намерете плътностите, очакванията и дисперсиите на $S_2 := X_1 + X_2$, $S_3 := X_1 + X_2 + X_3$ и $S_4 := X_1 + \cdots + X_4$.
5. (0.5 т.) Нека $Y_k := S_k / S_4$. Докажете, че $Y_{(k)}$ и $U_{(k)}$ имат еднакви разпределения за $k = 1, 2, 3$.

Решение 2. 1. Подхождаме стандартно (разглежданията ни са при $0 < t < 1$, другото е ясно по стандартна конструкция):

$$\begin{aligned}
F_{U_{(1)}}(t) &= \mathbb{P}(U_{(1)} < t) = 1 - \mathbb{P}(U_{(1)} > t) = 1 - \mathbb{P}(U_1 > t)^3 = 1 - (1-t)^3 \\
\Rightarrow f_{U_{(1)}}(t) &= \frac{d}{dt} F_{U_{(1)}}(t) = -3(1-t)^2 \cdot (-1) = 3(1-t)^2 \\
\Rightarrow E[U_{(1)}] &= \int_0^1 t \cdot 3(1-t)^2 dt = 3\int_0^1 (t^3 - 2t^2 + t)dt = \cdots = \frac{1}{4}
\end{aligned}$$

$$\begin{aligned}
F_{U_{(3)}}(t) &= \mathbb{P}(U_{(3)} < t) = \mathbb{P}(U_1 < t)^3 = t^3 \\
\Rightarrow f_{U_{(3)}}(t) &= 3t^2 \\
\Rightarrow E[U_{(3)}] &= \int_0^1 t \cdot 3t^2 dt = 3\int_0^1 t^3 dt = \cdots = \frac{3}{4}
\end{aligned}$$

2. Бива да помислим защо не е хубаво да се подвеждаме, но straight to the point:

$$\begin{aligned}
F_{U_{(2)}}(t) &= \mathbb{P}(U_{(2)} < t) \\
&= \mathbb{P}\left(\text{поне две от трите независими реализации са } < t\right) \\
&= \mathbb{P}(Bin(3,t) \geq 2) \\
&= \binom{3}{2} t^2(1-t) + t^3 \\
&= 3t^2 - 2t^3 \\
\Rightarrow f_{U_{(2)}}(t) &= 6t - 6t^2 = 6t(1-t)
\end{aligned}$$

3. За улеснение ще отбележим $U := U_1$. Имаме $X := -\ln(U) =: g(U) \in [0, +\infty)$. Функцията g е строго монотонно намаляваща, значи обратима, и обратната $h(x) := g^{-1}(x) = e^{-x}$ е диференцируема. Сменяме променливите:

$$f_X(x) = \underbrace{f_U(h(x))}_{=1} \cdot |\underbrace{h'(x)}_{-e^{-x}}| = e^{-x}, \text{ разбира се, при } 0 < x$$

Следователно $X \sim Exp(1)$ и можем да използваме, че $\mathbb{E}[X] = DX = 1$ или да пресметнем директно:

$$\begin{aligned}\mathbb{E}[X] &= \int_0^\infty xe^{-x}dx = \cdots = 1\\ \mathbb{E}[X^2] &= \int_0^\infty x^2e^{-x}dx = \cdots = 2\\ \Rightarrow \mathbb{D}[X] &= 2 - 1 = 1\end{aligned}$$

4. Ако сме се сетили от лекции, знаем, че сума на независими еднакво разпределени експоненциални е именно гама разпределена случайна величина, т.е. ако $X_i \sim Exp(\lambda)$ са независими, то:

$$S_n := \sum_{i=1}^n X_i \sim \Gamma(n, \lambda)$$

Оттам вече: в общия случай и (по-вдясно) конкретно при нас, за $t > 0$:

$$\begin{aligned}f_{S_n}(t) &= \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!} && = \frac{t^{n-1}}{(n-1)!}e^{-t}\\ \mathbb{E}[S_n] &= \frac{n}{\lambda} && = n\\ \mathbb{D}[S_n] &= \frac{n}{\lambda^2} && = n\end{aligned}$$

Разбира се, докато очакването и дисперсията можем да изведем директно от независимостта на експоненциалните в сумата, то и без да намесваме гама разпределението можем също да изведем плътността чрез формулата за конволюция (сума) на две независими случайни величини:

$$\begin{aligned}f_{S_2}(t) &\overset{t\geq 0}{=} \int_{-\infty}^{+\infty} \underbrace{f_{X_1}(x)}_{=e^{-x}\cdot \mathbb{1}_{\{0<x\}}} \underbrace{f_{X_2}(t-x)}_{e^{-(t-x)}\cdot \mathbb{1}_{\{0<t-x\}}} dx = \int_0^t e^{-x}e^{x-t}dx = e^{-t}\int_0^t dx && = te^{-t}\\ f_{S_3}(t) &\overset{t\geq 0}{=} \int_{-\infty}^{+\infty} \underbrace{f_{S_2}(x)}_{=xe^{-x}\cdot \mathbb{1}_{\{0<x\}}} \underbrace{f_{X_3}(t-x)}_{e^{-(t-x)}\cdot \mathbb{1}_{\{0<t-x\}}} dx = \int_0^t xe^{-x}e^{x-t}dx = e^{-t}\int_0^t xdx && = \frac{t^2}{2}e^{-t}\\ f_{S_4}(t) &\overset{t\geq 0}{=} \ldots \text{ по индукция}\end{aligned}$$

5. За последната точка, след като представим

$$Y_k = \frac{X_1 + \cdots + X_k}{(X_1 + \cdots + X_k) + (X_{k+1} + \cdots + X_4)} = \frac{S_k}{S_k + \widetilde{S}_{4-k}},$$

за $\widetilde{S}_{4-k}$ независимо от S_k копие на S_{4-k}, можем да продължим по стандарния метод за намиране на плътност на трансформация (двумерна смяна, якобиан и т.н.).

Задача 3. 1. (0.25 т.) В производството на специализирани автомобилни компоненти, дължината на определени детайли се измерва, и отклоненията от препоръчителната дължина се моделират с очакване 0 и дисперсия 9. Компонентът се счита за приемлив, ако отклонението от препоръчителната дължина е по-малко от зададен праг s. Определете какъв процент от грешките по дължина попадат в приемливия обхват, когато $s = 1$. Намерете стойността на s, при която 95% от компонентите отговарят на критериите за приемливост.

2. (0.5 т.) Нека $X \sim N(1, 1)$. С точност 0.01, намерете $\mathbb{E}|X|$.

Решение 3. 1.

2.

$$\begin{aligned}
\mathbb{E}[|X|] &= \int_{-\infty}^{+\infty} |x| f_X(x)dx \overset{t:=x-1}{=} \int_{-\infty}^{+\infty} |t+1| \underbrace{f_X(t+1)}_{f_Z(t)} dt \\
&= \int_{-\infty}^{-1} -(t+1) f_Z(t)dt + \int_{-1}^{+\infty} (t+1) f_Z(t)dt \\
&= \underbrace{-\int_{-\infty}^{-1} f_Z(t)dt}_{=-\Phi(-1)} + \underbrace{\int_{-1}^{+\infty} f_Z(t)dt}_{=1-\Phi(-1)} - \underbrace{\int_{-\infty}^{-1} t f_Z(t)dt}_{t\mapsto -t} + \underbrace{\int_{-1}^{+\infty} t f_Z(t)dt}_{=\int_{-1}^{-1}+\int_1^{\infty}} \\
&= 1 - 2\underbrace{\Phi(-1)}_{1-\Phi(1)} + 2\int_1^{\infty} t f_Z(t)dt + \underbrace{\int_{-1}^{1} t f_Z(t)dt}_{=0\text{ (нечетна)}} \\
&= 2\Phi(1) - 1 + 2J
\end{aligned}$$

Остана да пресметнем:

$$J := \int_1^{\infty} t f_Z(t)dt = \int_1^{\infty} t \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = -\frac{1}{\sqrt{2\pi}} \int_1^{\infty} e^{-\frac{t^2}{2}} d\left(-\frac{t^2}{2}\right) = -\frac{1}{\sqrt{2\pi}} \left[e^{-\frac{t^2}{2}}\right]_{t=1}^{t\to+\infty} = \frac{1}{\sqrt{2\pi e}}$$

И в крайна сметка:

$$\mathbb{E}[|X|] = -1 + 2\underbrace{\Phi(1)}_{\approx 0.841} + \underbrace{\frac{2}{\sqrt{2\pi e}}}_{\approx 0.484} \approx 1.166$$

Задача 4. Равнината е разграфена на успоредни прави, като разстоянието между всеки две съседни е $2t$. Хвърляме върху нея игла с дължина $2l \leq 2t$, която се приземява на случайна позиция.

1. (0.5 т.) Каква е вероятността иглата да докосва някоя от правите?

   Упътване: Параметризирайте по позицията на центъра на иглата и ъгъла, който сключва с правите.

2. (0.25 т.) Да допуснем, че $l = t$. Колко най-малко игли би трябвало да хвърлим, независимо една от друга, така че с вероятност поне 95% поне половината от тях да пресичат права? Можете ли да обясните как хвърлянето на игли по описания горе начин би могло да ни даде оценка за π и защо?

3. *(0.25 т.) Бонус: Решете 1. в случая $l > t$.*

Решение 4. Задачата е позната като Иглата на Буфон. Детайлно решение има в посочения линк.

1. Нека

$$\begin{aligned}
X &:= \text{ разстоянието от центъра на иглата до най-близката права} && \sim U(0,t) \\
\theta &:= \text{ острият ъгъл, сключен между иглата и най-близката права} && \sim U\left(0, \frac{\pi}{2}\right)
\end{aligned}$$

   Тогава иглата пресича правата т.т.к. $\{X \leq l\sin\theta\}$, значи търсим:

$$p := \mathbb{P}(X \leq l\sin\theta) = \int_0^{\frac{\pi}{2}} \int_0^{l\sin\theta} \frac{2}{\pi t} dt d\theta = \frac{2l}{\pi t}$$

   Решението в Уикипедия.

2. Понеже $l = t$, то $p = \frac{2}{\pi}$. Хвърлянето на игли представяме като бернулиев експеримент, т.е. ще разгледаме $S_n := \sum_{i=1}^n X_i$, където $X_i \sim Ber(p)$ са независими хвърляния на една игла, отчитащи „успех“, ако тя пресича права. Сега търсим такава оценка за n, че

$$\mathbb{P}\left(S_n \geq \frac{1}{2}n\right) \geq 95\%$$

   Стандартно, ще използваме ЦГТ. Нека отбележим предварително, че $\mu := \mathbb{E}[X_1] = p = \frac{2}{\pi}$ и $\sigma := \sqrt{\mathbb{D}[X]} = \sqrt{p(1-p)}$. И така, сега искаме:

$$\mathbb{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \geq \frac{\frac{n}{2} - n\mu}{\sigma\sqrt{n}}\right) = \mathbb{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \geq \frac{\sqrt{n}\left(\frac{1}{2} - \mu\right)}{\sigma}\right) \overset{\text{ЦГТ}}{\approx} \mathbb{P}\left(Z \geq \underbrace{\frac{\sqrt{n}\left(\frac{1}{2} - \mu\right)}{\sigma}}_{=:C_n}\right) \geq 95\%$$

$$\begin{aligned}
1 - \Phi\left(C_n\right) &\geq 95\% \\
\Phi(C_n) &\leq 5\% && \Phi \text{ е монотонно растяща, значи също и } \Phi^{-1} \\
C_n &\leq \Phi^{-1}(5\%) \\
\frac{\sqrt{n}\left(\frac{1}{2} - \mu\right)}{\sigma} &\leq \Phi^{-1}(5\%) && \text{но } C_n < 0 \text{ и е дори намаляваща по } n \text{, понеже } \frac{1}{2} < \mu \\
\sqrt{n} &\geq \frac{\sigma\Phi^{-1}(5\%)}{\frac{1}{2} - \mu} && \text{да отбележим, че и } \Phi^{-1}(5\%) < 0 \\
n &\geq \left(\frac{\sigma\Phi^{-1}(5\%)}{\frac{1}{2} - \mu}\right)^2
\end{aligned}$$

Приближаваме:

$$\begin{aligned}
\Phi^{-1}(5\%) &\approx -1.645 \\
\sigma &\approx 0.481 \\
\mu &\approx 0.637 \\
\Rightarrow \left(\frac{\sigma\Phi^{-1}(5\%)}{\frac{1}{2} - \mu}\right)^2 &\approx 33.5387
\end{aligned}$$

И в крайна сметка $n \geq 34$.

А по въпроса как това ни дава оценка за π отговор ни дава (У)ЗГЧ. Понеже $\frac{1}{n}S_n \xrightarrow{\text{п.с.}} \mu = \frac{2}{\pi}$, то нашата оценка за π би била $\frac{2}{\frac{1}{n}S_n}$. Можем да си мислим колко голямо искаме да е n, за да достигнем желана точност.

3. Отново, решението в Уикипедия.