**Math 3029/6109, Semester 1, 2019**

**Probability Modelling with Applications**

**Mid semester practice exam**

.

This exam contains 4 problems, each worth 25 points. Answer all questions.

**Problem 1.** (a) (12 points) We draw six balls with replacement from an urn that contains three red balls, two blue balls, and one green ball. Determine the conditional probability that we draw three red balls, two blue balls and one green ball, given that we draw three red balls and at most one green ball.

(b) (13 points) Let $X : \Omega \to \mathbb{R}$ be a continuous random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $X \geq -2$, $\mathbb{P}(0 \leq X < 1) = \frac{3}{4}$ and $\mathbb{P}(X \geq 1) = \frac{1}{8}$. Suppose also that the density function $\rho_X : \mathbb{R} \to (0, \infty)$ of $X$ has the property that $\rho_X(x) \leq \rho_Y(x)$ for $x \geq 2$, where $Y$ is a random variable with the gamma distribution $\Gamma_{1,3}$. Show that

$$-\frac{1}{8} \leq \mathbb{E}(X) < 4.$$

$\gamma_{1,3} = e^{-x}\frac{x^2}{(3-1)!} = \frac{e^{-x}x^2}{2}$

*Solutions.* (a): Let $X : \Omega \to \mathbb{Z}_+^3$ be the number of balls drawn of each color. Since we draw the balls with replacement, $X$ has the multinomial distribution $\mathcal{M}_{6,\rho}$ with $\rho = (\frac{1}{2}, \frac{1}{3}, \frac{1}{6})$. The relevant probability is equal to

$$\mathbb{P}(X = (3,2,1) \mid X = (3,2,1) \text{ or } X = (3,3,0)) = \frac{\mathbb{P}(X = (3,2,1))}{\mathbb{P}(X = (3,2,1) + \mathbb{P}(X = (3,3,0))}.$$

$M_{n,p}(\{\vec{k}\}) = \frac{n!}{\prod_{a\in E} k_a!}\prod_{a\in E} p(a)^{k_a}$ E is the set of color

One has

$$\mathbb{P}(X = (3,2,1)) = \mathcal{M}_{6,\rho}(\{(3,2,1)\}) = \binom{6}{(3,2,1)}\frac{1}{2^3}\frac{1}{3^2}\frac{1}{6} = \frac{6!}{3!2!}\frac{1}{8}\frac{1}{9}\frac{1}{6} = \frac{720}{12}\frac{1}{72}\frac{1}{6} = \frac{10}{72} = \frac{5}{36}$$

$(\frac{1}{2})^3(\frac{1}{3})^2(\frac{1}{6})$

and

$$\mathbb{P}(X = (3,3,0)) = \binom{6}{(3,3,0)}\frac{1}{2^3}\frac{1}{3^3} = \frac{6!}{3!3!}\frac{1}{8}\frac{1}{27} = \frac{720}{36}\frac{1}{8}\frac{1}{27} = \frac{20}{8}\frac{1}{27} = \frac{5}{54}.$$

Hence the relevant probability is equal to

$$\frac{\mathbb{P}(X = (3,2,1))}{\mathbb{P}(X = (3,2,1) + \mathbb{P}(X = (3,3,0))} = \frac{\frac{5}{36}}{\frac{5}{36} + \frac{5}{54}} = \frac{\frac{5}{2}}{\frac{5}{2} + \frac{5}{3}} = \frac{15}{15 + 10} = \frac{3}{5}.$$

(b): One has

$$\mathbb{E}(X) = \underbrace{\int_{-2}^{0} x\rho_X(x)\mathrm{d}x}_{<0} + \int_0^1 \overset{\le 1}{x}\rho_X(x)\mathrm{d}x + \int_1^2 \overset{\le 2}{x}\rho_X(x)\mathrm{d}x + \int_2^\infty x\rho_X(x)\mathrm{d}x.$$

We first prove the upper bound. The assumptions, and the fact that $Y \geq 0$ and $\mathbb{E}(Y) = 3$, yield

$$\begin{aligned}\mathbb{E}(X) &\leq \int_0^1 \rho_X(x)\mathrm{d}x + 2\int_1^2 \rho_X(x)\mathrm{d}x + \int_2^\infty x\rho_Y(x)\mathrm{d}x \\ &\leq \mathbb{P}(0 \leq X \leq 1) + 2\mathbb{P}(1 \leq X \leq 2) + \mathbb{E}(Y) = \frac{3}{4} + \frac{1}{4} + 3 = 4.\end{aligned}$$

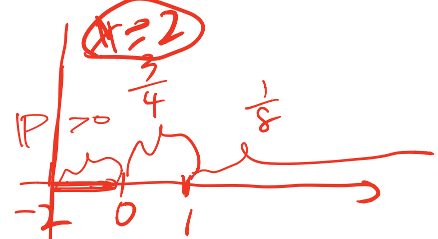
since $\rho_X \le \rho_Y$ $\forall x \ge 2$

To obtain strict inequality one can argue as follows.

One has

$$3 = \mathbb{E}(Y) = \int_0^\infty x\rho_Y(x)\mathrm{d}x > \int_2^\infty x\rho_x(x)\mathrm{d}x,$$

which leads to strict inequality.




Alternatively, because X is continuous and $\mathbb{P}(-2 \le X \le 0) > 0$, there exists an $\varepsilon > 0$ such that

$$\int_{-2}^{0} x\rho_X(x)\mathrm{d}x < -\varepsilon.$$

$<0$

This additional factor also leads to strict inequality.

For similar reasons one has

$$\int_0^1 x\rho_X(x)\mathrm{d}x \le (1-\varepsilon)\mathbb{P}(0 \le X \le 1) = (1-\varepsilon)\tfrac{3}{4} < \tfrac{3}{4}$$

and

$$\int_1^2 x\rho_X(x)\mathrm{d}x \le (1-\varepsilon)2\mathbb{P}(1 \le X \le 2) = (1-\varepsilon)\tfrac{1}{4} < \tfrac{1}{4},$$

and each of these yields strict inequality as well.

Finally, one can calculate explicitly that

$$\int_2^\infty x\rho_X(x)\mathrm{d}x \le \int_2^\infty x\rho_Y(x)\mathrm{d}x = \tfrac{1}{2}\int_2^\infty x^3e^{-x}\mathrm{d}x = \tfrac{1}{2}8e^{-2} + \tfrac{3}{2}\int_2^\infty x^2e^{-x}\mathrm{d}x$$
$$= 4e^{-2} + \tfrac{3}{2}4e^{-2} + 3\int_2^\infty xe^{-x}\mathrm{d}x = 10e^{-2} + 6e^{-2} + 3\int_2^\infty e^{-x}\mathrm{d}x = 19e^{-2}.$$

A quick calculation shows that

$$3e^2 > 3(2.7)^2 = 3(5.4 + 0.7 \cdot 2.7) > 3(5.4+1) = 19.2 > 19,$$

which also yields strict inequality. { upper bound

Either of these approaches suffices to prove the required upper bound. On the other hand, for the lower bound one has

$$\mathbb{E}(X) = \int_{-2}^{0} x\rho_X(x)\,dx + \int_0^1 x\rho_X(x)\,dx + \int_1^\infty x\rho_X(x)\,dx$$

$$\mathbb{E}(X) \ge -2\int_{-2}^{0}\rho_X(x)\mathrm{d}x + \int_1^\infty 1\cdot\rho_X(x)\mathrm{d}x = -2\mathbb{P}(-2 \le X \le 0) + \mathbb{P}(X \ge 1)$$
$$= -\tfrac{1}{4} + \tfrac{1}{8} = -\tfrac{1}{8}.$$

$-\frac{1}{4}$

$= 1 - \mathbb{P}(0<X<1) - \mathbb{P}(X \ge 1) = \frac{1}{8}$

One could obtain strict inequality here too, but I decided against including this in the question. ☐

**Problem 2.** Let $X$ be a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with cumulative distribution function $F : \mathbb{R} \to [0,1]$ given by

$$\text{cdf.}\quad F(c) = \begin{cases} 0 & \text{for } c < 1, \\ 1 - \frac{1}{2}c^{-4} & \text{for } c > 1. \end{cases}$$

$c\uparrow 1 \quad c\downarrow 1$

(a) (10 points) Determine $\mathbb{P}(X = 1)$.
(b) (15 points) Determine whether $X \in \mathcal{L}^1(\Omega)$. If so, compute $\mathbb{E}(X)$.

$\lim_{c\downarrow 1} f(c) = \lim_{\Delta \to 0} f(c+\Delta)$

*Solution.* (a): One has

$$\mathbb{P}(X=1) = \mathbb{P}(X \le 1) - \mathbb{P}(X < 1) = \lim_{c\downarrow 1}\mathbb{P}(X \le c) - \lim_{c\uparrow 1}\mathbb{P}(X < c) = \lim_{c\downarrow 1} F(c) - \lim_{c\uparrow 1} 0 = \frac{1}{2}.$$

(b): Since $F(c) = 0$ for $c < 0$, one has $X \ge 0$ almost surely, and $X \in \mathcal{L}^1(\Omega)$ if and only if $\mathbb{E}(X) < \infty$. One has

$$\mathbb{E}(X) = \int_0^\infty (1 - F(c))\mathrm{d}c - \int_{-\infty}^{0} F(c)\mathrm{d}c = \int_0^1 \mathrm{d}c + \frac{1}{2}\int_1^\infty c^{-4}\mathrm{d}c = 1 + \frac{1}{6}c^{-3}\Big|_{c=1}^{\infty} = \frac{7}{6} < \infty,$$

$\mathbb{P}(X \ge 0) = 1$

so that $X \in \mathcal{L}^1(\Omega)$.

Alternatively, one can note that $F$ is continuously differentiable on $(1,\infty)$, with derivative given by $F'(c) = 2c^{-5}$ for $c > 1$. Combined with the fact that $\mathbb{P}(X = 1) = 1/2$, this implies that the distribution $\mathbb{P}_X$ of $X$ is given by

$P(X\le 2) = P(X=1)\times P(1<X\le 2)$

在这一点，取 w

$1_A(x) = \begin{cases} 1 & \text{if } x\in A \\ 0 & \text{if } x\in A^c \end{cases}$

$\mathbb{P}_X = \frac{1}{2}\delta_1 + \frac{1}{2}\mu$, where $\delta_1$ is Dirac point mass at 1 and $\mu$ is the probability measure with Lebesgue density $2\mathbf{1}_{(1,\infty)}F'$. It is straightforward to check that $\mathbb{E}(\delta_1) = 1$, and

$E(\delta_1) = 1\cdot P(X=1) = 1\times 1 = 1$

因为前面有 1/2，所以要 ×2

$E(x) = \int x\,p(x)\,dx$

$$\mathbb{E}(\mu) = 2\int_1^\infty cF'(c)\mathrm{d}c = 4\int_1^\infty c^{-4}\mathrm{d}c = -\frac{4}{3}c^{-3}\Big|_{c=1}^{\infty} = \frac{4}{3}.$$

Hence

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{P}_X) = \frac{1}{2}\mathbb{E}(\delta_1) + \frac{1}{2}\mathbb{E}(\mu) = \frac{1}{2} + \frac{2}{3} = \frac{7}{6}.$$

$(1-\frac{1}{2}c^{-4})' = (-4)(-\frac{1}{2})c^{-5} = 2c^{-5}$

Equivalently, one can write

$$\mathbb{E}(X) = 1\cdot\mathbb{P}(X=1) + \int_1^\infty cF'(c)\mathrm{d}c = \frac{1}{2} + \frac{2}{3} = \frac{7}{6}.$$

□

**Problem 3.** Let $X$, $Y$, $S$ and $T$ be real-valued random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

(a) (13 points) Suppose that $X$ and $Y$ are independent, that $X$ has the $\beta_{1,1}$ distribution, and that $Y$ has the standard normal distribution. Determine the probability of the event $\{Y \le 1 \text{ and } X \le Y\}$.

(b) (12 points) Suppose that $S$ and $T$ are independent, that $S$ has the gamma distribution $\Gamma_{2,\frac{1}{4}}$, and that $T$ has the gamma distribution $\Gamma_{4,\frac{3}{4}}$. Determine the density of $S+2T$.

pdf

先积 variable，后积常数

*Solution.* (a): The density $\rho$ of the random variable $(X,Y)$ is the product of the densities $\rho_X$ of $X$ and $\rho_Y$ of $Y$, since $X$ and $Y$ are independent. One has

$$\rho_X(x) = 1 \qquad (x\in(0,1))$$

$B(r,s) = \frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)}$

and

$$\rho_Y(y) = \frac{1}{\sqrt{2\pi}}e^{-y^2/2} \qquad (y\in\mathbb{R}).$$

$B(1,1) = \frac{\Gamma(1)\Gamma(1)}{\Gamma(2)} = \frac{1\times 1}{1!} = 1$

So

$$\rho(x,y) = \frac{1}{\sqrt{2\pi}}e^{-y^2/2} \qquad (x\in(0,1), y\in\mathbb{R}).$$

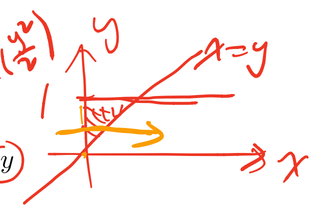
$\Gamma(n) = (n-1)!$

The probability of $\{Y\le 1 \text{ and } X\le Y\}$ is equal to the integral of $\rho$ over the region

$$D := \{(x,y)\in\mathbb{R}^2 \mid 0\le x\le y\le 1\},$$

since $\rho(x,y) = 0$ if $x<0$. So the relevant probability is equal to

$ydy = d(\frac{y^2}{2})$




$$\iint_D \rho(x,y)\mathrm{d}x\mathrm{d}y = \frac{1}{\sqrt{2\pi}}\int_0^1 e^{-y^2/2}\int_0^y \mathrm{d}x\mathrm{d}y = \frac{1}{\sqrt{2\pi}}\int_0^1 ye^{-y^2/2}\mathrm{d}y$$
$$= -\frac{1}{\sqrt{2\pi}}e^{-y^2/2}\Big|_{y=0}^1 = \frac{1}{\sqrt{2\pi}}(1-e^{-1/2}).$$

$= \frac{1}{\sqrt{2\pi}}\int_0^1 e^{-\frac{y^2}{2}}d(\frac{y^2}{2}) = \frac{1}{\sqrt{2\pi}}\int_0^{\frac{1}{2}}e^{-t}dt = \frac{1}{\sqrt{2\pi}}(-e^{-t})\Big|_{t=0}^{\frac{1}{2}} = \frac{1}{\sqrt{2\pi}}(e^0 - e^{-\frac{1}{2}})$

(b): The density $\rho_{2Y}$ of $2Y$ is given in terms of the density $\rho_Y$ of $Y$ as

$$\rho_{2Y}(y) = \frac{1}{2}\rho_Y\left(\frac{y}{2}\right) = \frac{1}{2}\frac{4^{3/4}}{\Gamma(3/4)}\left(\frac{y}{2}\right)^{-1/4}e^{-2y} = 2^{-3/4}\frac{4^{3/4}}{\Gamma(3/4)}y^{-1/4}e^{-2y} = \frac{2^{3/4}}{\Gamma(3/4)}y^{-1/4}e^{-2y}$$

for $y\in(0,\infty)$, and this is the density of the $\Gamma_{2,\frac{3}{4}}$ distribution. Moreover, $X$ and $2Y$ are independent, and by results on sums of independent random variables with gamma distributions we see that $X+2Y$ has the $\Gamma_{2,\frac{1}{4}+\frac{3}{4}} = \Gamma_{2,1} = \mathcal{E}_2$ distribution. This distribution has density $\rho:(0,\infty)\to(0,\infty)$ given by

$$\rho(x) = 2e^{-2x} \qquad (x\in(0,\infty)).$$

□

**Problem 4.** Consider a pop-up village where people arrive for lunch. For $t\ge 0$, let $N_t$ be the number of people who have arrived at the pop-up village by time $t$ (in minutes). Suppose that $(N_t)_{t\ge 0}$ is a Poisson process with rate 2.

(a) (3 points) What is the probability that at least three people arrive between the second minute and the fourth minute?

(b) (3 points) What is the expected number of people to arrive in the first three minutes?

$$S \sim \Gamma_{2,\frac{1}{4}} \quad \gamma_{2,\frac{1}{4}}(s) = \frac{e^{-2s}\, 2^{\frac{1}{4}}\, s^{-\frac{3}{4}}}{\Gamma(\frac{1}{4})}$$

$$T \sim \Gamma_{4,\frac{3}{4}}, \quad \gamma_{4,\frac{3}{4}}(t) = \frac{e^{-4t}\, 4^{3/4}\, t^{-\frac{1}{4}}}{\Gamma(\frac{3}{4})}$$

$$F_{2T}(c) = F_T\left(\frac{c}{2}\right) = \int_{-\infty}^{\frac{c}{2}} \frac{e^{-4t}\, 4^{3/4}\, t^{-\frac{1}{4}}}{\Gamma(3/4)}\, dt$$

$$= \int_{-\infty}^{c} \frac{e^{-4(\frac{t}{2})}\, 4^{3/4} \left(\frac{t}{2}\right)^{-\frac{1}{4}}}{\Gamma(\frac{3}{4})}\, d\left(\frac{t}{2}\right)$$

$$= \int_{-\infty}^{c} \underbrace{\frac{1}{2\Gamma(\frac{3}{4})}\, e^{-2t}\, 4^{3/4} \left(\frac{t}{2}\right)^{-\frac{1}{4}}}_{\rho_{2T}(t)}\, d(t) \qquad 2T \sim \Gamma_{2,\frac{3}{4}}$$

(c) (3 points) What is the expected time by which the tenth person arrives?

(d) (3 points) What is the probability that, after the tenth person arrives, for at least two minutes no one else arrives?

(e) (13 points) For $t \geq 0$, set $Z_t := (-1)^{N_t}$. Show that

$$\mathbb{P}(Z_t = Z_s) = \frac{1 + e^{-4(t-s)}}{2}$$

for $0 \leq s < t$.

*Solution.* (a): We are interested in $\mathbb{P}(N_4 - N_2 \geq 3)$. By definition of the Poisson process, the random variable $N_4 - N_2$ has the Poisson distribution $\mathcal{P}_4$. Hence

$$\begin{aligned} &\mathbb{P}(N_4 - N_2 \geq 3) = 1 - \mathbb{P}(N_4 - N_2 \leq 2) = 1 - (\mathcal{P}_4(\{0\}) + \mathcal{P}_4(\{1\}) + \mathcal{P}_4(\{2\})) \\ &= 1 - e^{-4}(1 + 4 + \tfrac{1}{2}16) = 1 - 13e^{-4}. \end{aligned}$$

(b): We know that $N_3$ has the Poisson distribution $\mathcal{P}_6$. Hence

$$\mathbb{E}(N_3) = 6.$$

(c): For $k \in \mathbb{N}$, let $T_k$ be the time by which the $k$-th person arrives. Then we are interested in $\mathbb{E}(T_{10})$. Since $T_{10}$ has the gamma distribution $\Gamma_{2,10}$, one has

$$\mathbb{E}(T_{10}) = \frac{10}{2} = 5.$$

(d): We are interested in $\mathbb{P}(T_{11} - T_{10} \geq 2)$. Now, $\underline{T_{11}} - \underline{T_{10}}$ has the exponential distribution $\mathcal{E}_2$, which has density given by $\gamma_{2,1}(x) = 2e^{-2x}$ for $x > 0$. Hence

$\Gamma_{2,11} - \Gamma_{2,10} = \Gamma_{2,1} = \mathcal{E}_2$

$$\mathbb{P}(T_{11} - T_{10}) = 2\int_2^\infty e^{-2x}\mathrm{d}x = e^{-4}.$$

(e): Note that

$$Z_t = (-1)^{N_t} = (-1)^{N_s} = Z_s$$

if and only if $N_t - N_s$ is even. Moreover, $N_t - N_s$ has the Poisson distribution $\mathcal{P}_{2(t-s)}$, by definition of the Poisson process. Hence

$$\begin{aligned} &\mathbb{P}(Z_t = Z_s) = \mathbb{P}(N_t - N_s \in 2\mathbb{Z}_+) = \sum_{k=0}^\infty \mathcal{P}_{2(t-s)}(\{2k\}) = \sum_{k=0}^\infty e^{-2(t-s)}\frac{(2(t-s))^{2k}}{(2k)!} \\ &= e^{-2(t-s)}\sum_{k=0}^\infty \frac{(2(t-s))^k + (-2(t-s))^k}{2(k!)} = \frac{e^{-2(t-s)}}{2}\Big(\sum_{k=0}^\infty \frac{(2(t-s))^k}{k!} + \sum_{k=0}^\infty \frac{(-2(t-s))^k}{k!}\Big) \\ &= \frac{e^{-2(t-s)}}{2}(e^{2(t-s)} + e^{-2(t-s)}) = \frac{1 + e^{-4(t-s)}}{2}. \end{aligned}$$

$\square$

END OF EXAMINATION