Math 180A: Final Practice Problems (Chapters 6, 8, and 9)

1. *Let X and Y be continuous random variables with joint density function*

$$f_{X,Y}(x,y) = \begin{cases} \frac{x}{2} + \frac{y}{4}, & 0 \le x \le 1,\ 0 \le y \le 2 \\ 0, & \textit{otherwise} \end{cases}$$

(a) *(5 points) Compute $\mathbb{P}(X > 2Y)$.*

(b) *(5 points) What is the marginal density f_X of X?*

(c) *(5 points) Are X and Y independent? Explain.*

Solutions. (a) Be definition of joint density

$$\begin{aligned}\mathbb{P}(X > 2Y) = \iint_{\{(x,y)\,:\,x>2y\}} f_{X,Y}(x,y)\,dxdy &= \int_0^1 \left(\int_0^{x/2} \left(\frac{x}{2} + \frac{y}{4}\right) dy \right) dx \\ &= \int_0^1 \left(\frac{x}{2}y + \frac{y^2}{8} \Big|_{y=0}^{y=x/2} \right) dx \\ &= \int_0^1 \left(\frac{x^2}{4} + \frac{x^2}{32} \right) dx = \frac{9}{32} \cdot \frac{1}{3} = \boxed{\frac{3}{32}}.\end{aligned}$$

(b) When $x \notin [0,1]$, $f_{X,Y}(x,y) = 0$ so $f_X(x) = 0$. When $x \in [0,1]$,

$$f_X(x) = \int_0^2 f_{X,Y}(x,y)\,dy = \int_0^2 \left(\frac{x}{2} + \frac{y}{4}\right) dy = \left(\frac{x}{2}y + \frac{y^2}{8} \Big|_{y=0}^{y=2} \right) = \boxed{x + \frac{1}{2}}.$$

(c) No, they are not independent. The domain is a rectangle, but the function $f_{X,Y}(x,y)$ is not a product of the form $u(x)v(y)$. Indeed, suppose it were. Then we would have to have $\frac{x}{2} = f_{X,Y}(x,0) = u(x)v(0)$, and also $\frac{x}{2} + 1 = f_{X,Y}(x,4) = u(x)v(4)$. Dividing these would show that

$$\frac{x/2}{x/2+1} = \frac{u(x)v(0)}{u(x)v(4)} = \frac{v(0)}{v(4)} = \text{constant}.$$

But that's not true: the function $\frac{x/2}{x/2+1}$ is not constant (it takes value 0 at $x = 0$ and value $\frac{1}{2}$ at $x = 2$). So $f_{X,Y}$ is not a product, and therefore X and Y are not independent.

2. *An urn contains 5 balls: two balls labeled 1, two balls labeled 2, and one ball labeled 3. You sample 2 balls from the bin, without replacement; the number on the first one is X, and the number on the second one is Y.*

(a) *(5 points) Compute the joint probability mass function of X and Y. (You may find it convenient to express it in the form of a chart.)*

(b) *(5 points) Compute the probability mass function of X and the probability mass function of Y.*

(c) *(5 points) Are X and Y independent? Justify your answer.*

Solutions. (a) The possible values of X and Y are both $\{1, 2, 3\}$, so we need to compute the 9 numbers $\mathbb{P}(X = j, Y = k)$ for $1 \le j, k \le 3$. Since Y is chosen after X, and the sampling is done without replacement, it is best to treat this as a two-stage experiment and use conditional probability:

$$\mathbb{P}(X = j, Y = k) = \mathbb{P}(Y = k|X = j)\mathbb{P}(X = j). \qquad (*)$$

For the choice of the first ball, we have $\mathbb{P}(X = 1) = \mathbb{P}(X = 2) = \frac{2}{5}$ while $\mathbb{P}(X = 3) = \frac{1}{5}$. For the second ball, we break into the 3 cases:

- $\{X = 1\}$: there are 4 balls left, one labeled 1, two labeled 2, and 1 labeled 3. So the conditional probabilities are $\mathbb{P}(Y = 1|X = 1) = \mathbb{P}(Y = 3|X = 1) = \frac{1}{4}, \mathbb{P}(Y = 2|X = 1) = \frac{2}{4} = \frac{1}{2}$.
- $\{X = 2\}$: there are 4 balls left, two labeled 1, one labeled 2, and 1 labeled 3. So the conditional probabilities are $\mathbb{P}(Y = 1|X = 2) = \frac{2}{4} = \frac{1}{2}, \mathbb{P}(Y = 2|X = 2) = \mathbb{P}(Y = 3|X = 2) = \frac{1}{4}$.
- $\{X = 3\}$: there are 4 balls left, two labeled 1 and two labeled 2. So the conditional probabilities are $\mathbb{P}(Y = 1|X = 3) = \mathbb{P}(Y = 2|X = 3) = \frac{2}{4} = \frac{1}{2}, \mathbb{P}(Y = 3|X = 3) = 0$.

Now using Equation (**??**), we can compute the joint probability mass function, which we write as the following chart:

	$Y = 1$	$Y = 2$	$Y = 3$
$X = 1$	$1/10$	$1/5$	$1/10$
$X = 2$	$1/5$	$1/10$	$1/10$
$X = 3$	$1/10$	$1/10$	0

(b) For each j, $\mathbb{P}(X = j) = \sum_{k=1}^{3} \mathbb{P}(X = j, Y = k)$, meaning we get the probability mass function of X by summing the rows of the above chart. Similarly, we get the probability mass function of Y by summing the columns. The result is

j	$\mathbb{P}(X = j)$
1	$2/5$
2	$2/5$
3	$1/5$

k	$\mathbb{P}(X = k)$
1	$2/5$
2	$2/5$
3	$1/5$

(We see that X and Y are identically distributed.)

(c) No, they are not independent. For example, note that $\mathbb{P}(X = 3, Y = 3) = 0$, while $\mathbb{P}(X = 3)\mathbb{P}(Y = 3) = \frac{1}{5} \cdot \frac{1}{5} = \frac{1}{25}$.

3. *Let $U_1, U_2, \ldots, U_n, \ldots$ be independent, identically distributed random variables, each with the* $\mathrm{Uniform}[-2,2]$ *distribution. Let $S_n = U_1 + U_2 + \cdots + U_n$.*

(a) *(5 points) Compute $\mathbb{E}(S_n)$ and $\mathrm{Var}(S_n)$.*

(b) *(5 points) For any $\epsilon > 0$, what can you say about*

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{|S_n|}{n^{2/3}} \geq \epsilon\right) ?$$

Solutions. (a) Each U_j has density $\frac{1}{4}\mathbb{1}_{[-2,2]}$, so

$$\mathbb{E}(U_j) = \int_{-2}^{2} x \cdot \frac{1}{4}\,dx = \frac{1}{8}x^2\Big|_{-2}^{2} = 0, \qquad \mathbb{E}(U_j^2) = \int_{-2}^{2} x^2 \cdot \frac{1}{4}\,dx = \frac{1}{12}x^3\Big|_{-2}^{2} = \frac{4}{3}.$$

Therefore $\mathrm{Var}(U_j) = \mathbb{E}(U_j^2) - [\mathbb{E}(U_j)]^2 = \frac{4}{3}$. The sum S_n therefore has

$$\mathbb{E}(S_n) = \sum_{j=1}^{n} \mathbb{E}(U_j) = \boxed{0}$$

and since the U_j are all independent, we have

$$\mathrm{Var}(S_n) = \sum_{j=1}^{n} \mathrm{Var}(U_j) = \boxed{\frac{4}{3}n}.$$

(b) We use Chebyshev's inequality. From part (a), $\mathbb{E}(S_n/n^{2/3}) = \frac{1}{n^{2/3}}\mathbb{E}(S_n) = 0$, while

$$\mathrm{Var}(S_n/n^{2/3}) = \frac{1}{n^{4/3}}\mathrm{Var}(S_n) = \frac{1}{n^{4/3}} \cdot \frac{4}{3}n = \frac{4}{3n^{1/3}}.$$

Thus

$$\mathbb{P}\left(\frac{|S_n|}{n^{2/3}} \geq \epsilon\right) = \mathbb{P}\left(\left|\frac{S_n}{n^{2/3}} - \mathbb{E}\left(\frac{S_n}{n^{2/3}}\right)\right| \geq \epsilon\right) \leq \frac{\mathrm{Var}(S_n/n^{2/3})}{\epsilon^2} = \frac{4}{3\epsilon^2 n^{1/3}} \to 0 \text{ as } n \to \infty.$$

Thus

$$\boxed{\lim_{n\to\infty} \mathbb{P}\left(\frac{|S_n|}{n^{2/3}} \geq \epsilon\right) = 0.}$$

4. *Let T be the triangle in $\mathbb{R}^2$ with vertices $(0,0)$, $(0,1)$, and $(1,1)$ (including the interior). Suppose that $P = (X,Y)$ is a point chosen uniformly at random inside of T.*

(a) *(5 points) What is the joint density function of (X,Y)? Use this to compute $\mathrm{Cov}(X,Y)$.*

(b) *(5 points) Determine if X and Y are independent.*

Solutions. (a) The joint density of (X,Y) is

$$f_{(X,Y)} = \begin{cases} \frac{1}{\mathrm{Area}(T)} = 2 & \text{if } (x,y) \in T; \\ 0 & \text{if } (x,y) \notin T. \end{cases}$$

We compute the covariance using the formula $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$:

$$\mathbb{E}[XY] = \int_0^1 \int_x^1 2xy\, dydx = \int_0^1 \left(xy^2 \right)\Big|_x^1 dy = \int_0^1 x - x^3\, dy = \left(\frac{x^2}{2} - \frac{x^4}{4} \right)\Big|_0^1 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}.$$

Similarly,

$$\mathbb{E}[X] = \int_0^1 \int_x^1 2x\, dydx = \int_0^1 2x - 2x^2\, dx = \left(x^2 - \frac{2x^3}{3} \right)\Big|_0^1 = 1 - \frac{2}{3} = \frac{1}{3}$$

and

$$\mathbb{E}[Y] = \int_0^1 \int_x^1 2y\, dydx = \int_0^1 \left(y^2\Big|_x^1 \right) dx = \int_0^1 1 - x^2\, dx = \left(x - \frac{x^3}{3} \right)\Big|_0^1 = 1 - \frac{1}{3} = \frac{2}{3}.$$

So, $\mathrm{Cov}(X,Y) = \frac{1}{4} - \frac{1}{3}\frac{2}{3} = \frac{1}{4} - \frac{2}{9} = \frac{1}{36} > 0$.

(b) X and Y are not independent because $\mathrm{Cov}(X,Y) \neq 0$.

5. *Suppose $X_1, X_2, \ldots, X_n, \ldots$ are i.i.d. random variables with mean $\mathbb{E}(X_j) = 0$ and variance $\mathrm{Var}(X_j) = 1$. Determine the following limits with precise justifications.*

(a) *(5 points)* $\lim_{n\to\infty} \mathbb{P}\left(-\frac{n}{4} \le X_1 + \cdots + X_n < \frac{n}{2}\right)$

(b) *(5 points)* $\lim_{n\to\infty} \mathbb{P}(X_1 + \cdots + X_n = 0)$

Solutions. (a) Note that

$$\begin{aligned} 1 &\ge \mathbb{P}\left(-\frac{n}{4} \le X_1 + \cdots + X_n < \frac{n}{2}\right) \\ &= \mathbb{P}\left(-\frac{1}{4} \le \frac{X_1 + \cdots + X_n}{n} < \frac{1}{2}\right) \\ &\ge \mathbb{P}\left(-\frac{1}{4} \le \frac{X_1 + \cdots + X_n}{n} \le \frac{1}{4}\right). \end{aligned}$$

By the law of large numbers,

$$\lim_{n\to\infty} \mathbb{P}\left(-\frac{1}{4} \le \frac{X_1 + \cdots + X_n}{n} \le \frac{1}{4}\right) = 1.$$

This implies that

$$\lim_{n\to\infty} \mathbb{P}\left(-\frac{n}{4} \le X_1 + \cdots + X_n < \frac{n}{2}\right) = 1.$$

(b)

$$\begin{aligned} \mathbb{P}(X_1 + \cdots + X_n = 0) &= \mathbb{P}\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}} = 0\right) \\ &= \mathbb{P}\left(0 \le \frac{X_1 + \cdots + X_n}{\sqrt{n}} \le 0\right). \end{aligned}$$

By the central limit theorem,

$$\lim_{n\to\infty} \mathbb{P}\left(0 \le \frac{X_1 + \cdots + X_n}{\sqrt{n}} \le 0\right) = \Phi(0) - \Phi(0) = 0.$$