Statistical Inference Chapter 4

Gallant Tsao

August 11, 2024

1. Since (X, Y) is distributed uniformly, the probabilities in this question are simply the ratio of the area satisfying the requirements to the area of the square, which is 2.

   (a) The circle $x^2 + y^2 < 1$ has area π hence the answer is $\frac{\pi}{4}$.

   (b) The line $2x - y = 0$ passes throught the origin, hence the answer is $\frac{1}{2}$.

   (c) For any (x, y) in the interior of the square, $|x + y| < 2$ hence the probability is 1.

2. This is similar to the proof of Theorem 2.2.5.

   (a)
   $$\mathbb{E}[ag_1(X, Y) + bg_2(X, Y) + c] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (ag_1(x, y) + bg_2(x, y) + c) f_{X,Y}(x, y) \ dy dx$$
   $$= a \int_{-\infty}^{\infty} g_1(x, y) f_{X,Y}(x, y) \ dy dx + b \int_{-\infty}^{\infty} g_2(x, y) f_{X,Y}(x, y) \ dy dx$$
   $$+ c \int_{-\infty}^{\infty} f_{X,Y}(x, y) \ dy dx$$
   $$= a\mathbb{E}[g_1(x, y)] + b\mathbb{E}[g_2(x, y)] + c.$$

   (b)

3. For a fair die, each sample point in the sample space of size 36 has an equally likely chance of happening. Therefore we get that

   $$P(X = 0, Y = 0) = P(\{(1, 1), (1, 3), (1, 5), (2, 1), (2, 3), (2, 5)\})$$
   $$= \frac{6}{36} = \frac{1}{6}.$$

   $$P(X = 0, Y = 1) = P(\{(1, 2), (1, 4), (1, 6), (2, 2), (2, 4), (2, 6)\})$$
   $$= \frac{6}{36} = \frac{1}{6}.$$

   $$P(X = 1, Y = 0) = P(\{(3, 1), (3, 3), (3, 5), (4, 1), (4, 3), (4, 5), (5, 1), (5, 3), (5, 5), (6, 1), (6, 3), (6, 5)\})$$
   $$= \frac{12}{36} = \frac{1}{3}.$$

   $$P(X = 1, Y = 1) = P(\{(3, 2), (3, 4), (3, 6), (4, 2), (4, 4), (4, 6), (5, 2), (5, 4), (5, 6), (6, 2), (6, 4), (6, 6)\})$$
   $$= \frac{12}{36} = \frac{1}{3}.$$

   This matches the pmf of Example 4.1.5.

4. (a)

$$\int_0^1 \int_0^2 x + 2y \ dxdy = \int_0^1 \left[\frac{1}{2}x^2 + 2xy \right]_0^2 dy$$

$$= \int_0^1 2 + 4y \ dy$$

$$= [2y + 2y^2]_0^1$$

$$= 4,$$

which directly implies that $C = \frac{1}{4}$.

(b)

$$f_X(x) = \int_0^1 \frac{1}{4}(x + 2y) \ dy$$

$$= \frac{1}{4} \left[xy + y^2 \right]_0^1$$

$$= \frac{1}{4}(x + 1), \ 0 < x < 2.$$

(c) This depends on the values of x and y. If either x or y is ≤ 0, the cdf is just 0. If $x \geq 2$ and $y \geq 1$, the cdf is 1. The remaining cases are:

- $x \in (0, 2)$ and $y \geq 1$.

$$F_{X,Y}(x, y) = \int_0^x \int_0^1 \frac{1}{4}(u + 2v) \ dvdu$$

$$= \frac{1}{4} \int_0^x [uv + v^2]_0^1 \ du$$

$$= \frac{1}{4} \int_0^x u + 1 \ du$$

$$= \frac{1}{4} \left[\frac{1}{2}u^2 + u \right]_0^x$$

$$= \frac{1}{8}x^2 + \frac{1}{4}x.$$

- $x \geq 1$ and $y \in (0, 1)$.

$$F_{X,Y}(x, y) = \int_0^y \int_0^2 \frac{1}{4}(u + 2v) \ dudv$$

$$= \frac{1}{4} \int_0^y \left[\frac{1}{2}u^2 + 2uv \right]_0^2 dv$$

$$= \frac{1}{4} \int_0^y 2 + 4v \ dv$$

$$= \frac{1}{4} [2v + 2v^2]_0^y$$

$$= \frac{1}{2}y^2 + \frac{1}{2}y.$$

2

- $x \in (0, 2)$ and $y \in (0, 1)$.

$$F_{X,Y}(x, y) = \int_0^x \int_0^y \frac{1}{4}(u + 2v)\ dv du$$

$$= \frac{1}{4}\int_0^x [uv + v^2]_0^y\ du$$

$$= \frac{1}{4}\int_0^x uy + y^2\ du$$

$$= \frac{1}{4}\left[\frac{1}{2}u^2 y + uy^2\right]_0^x$$

$$= \frac{1}{8}x^2 y + \frac{1}{4}xy^2.$$

All in all, the cdf for y is

$$F_{X,Y}(x, y) = \begin{cases} 0 & \text{if } x \leq 0 \text{ or } y \leq 0, \\ \frac{1}{8}x^2 y + \frac{1}{4}xy^2 & \text{if } 0 < x < 2 \text{ and } 0 < y < 1, \\ \frac{1}{2}y^2 + \frac{1}{2}y & \text{if } x \geq 1 \text{ and } 0 < y < 1, \\ \frac{1}{8}x^2 + \frac{1}{4}x & \text{if } 0 < x < 2 \text{ and } y \geq 1. \\ 1 & \text{if } x, y \geq 1. \end{cases}$$

(d) Since $X \in (0, 2)$, $Z = \frac{9}{(X+1)^2} \in (1, 9)$.

5. (a)

$$P(X > \sqrt{Y}) = \int_0^1 \int_0^{x^2} x + y\ dy dx$$

$$= \int_0^1 \left[xy + \frac{1}{2}y^2\right]_0^{x^2}\ dx$$

$$= \int_0^1 x^3 + \frac{1}{2}x^4\ dx$$

$$= \left[\frac{1}{4}x^4 + \frac{1}{10}x^5\right]_0^1$$

$$= \frac{7}{20}.$$

(b)

$$P(X^2 < Y < X) = \int_0^1 \int_{x^2}^x 2x\ dy dx$$

$$= \int_0^1 [2xy]_x^{x^2}\ dx$$

$$= \int_0^1 2x^2 - 2x^3\ dx$$

$$= \left[\frac{2}{3}x^3 - \frac{1}{2}x^4\right]_0^1$$

$$= \frac{1}{6}.$$

6. Let A, B be the time that A and B arrives respectively. Then $A, B \sim \text{Uniform}(1, 2)$. Moreover, A and B are independent hence their joint distribution is the product of their marginals. That is,

$$f_{A,B}(a, b) = \frac{1}{4},\ a, b \in (1, 2).$$

Let X be the amount of time that A waits for B. Then for $x \leq 0$,

7. In this question we will measure in the number of minutes past 8AM. Then $X \sim \text{Uniform}(0, 30)$, $Y \sim \text{Uniform}(40, 50)$. Since X and Y are independent, their joint pdf is

$$f_{X,Y}(x, y) = \frac{1}{300}, \ (x, y) \in (0, 30) \times (40, 50).$$

Then we get that

$$P(\text{Arrives at work before 9AM}) = P(X + Y < 60)$$
$$= \int_{40}^{50} \int_0^{60-y} \frac{1}{300} \, dxdy$$
$$= \frac{1}{2}.$$

8. (a)

9.

$$P(a \leq X \leq b, c \leq Y \leq d)$$
$$= P(X \leq b, c \leq Y \leq d) - P(X \leq a, c \leq Y \leq d)$$
$$= P(X \leq b, Y \leq d) - P(X \leq a, Y \leq d) - [P(X \leq a, Y \leq d) - P(X \leq a, Y \leq c)]$$
$$= F_{X,Y}(b, d) - F_{X,Y}(a, d) - F_{X,Y}(b, c) + F_{X,Y}(a, c)$$
$$= F_X(b)F_Y(d) - F_X(b)F_Y(d) - F_X(b)F_Y(d) + F_X(b)F_Y(d)$$
$$= (F_X(b) - F_X(a))(F_Y(d) - F_Y(c))$$
$$= P(a \leq X \leq b)P(c \leq Y \leq d).$$

10. (a) From the table, it is clear that $P(X = 2, Y = 3) = 0$. However, $P(X = 2)P(X = 3) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6} \neq 0$. Therefore X and Y are dependent.

   (b)

		U		
		1	2	3
	2	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
V	3	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
	4	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$

   This is just the product of the marginals hence X and Y have to be independent.

11. The support of U is $\{1, 2, 3, \dots\}$ while the support for V is $\{u + 1, u + 2, \dots\}$. They are clearly not random (in fact, V depends heavily on U).

12. Suppose the length of the stick is 1, and let X, Y be the two points where the stick is broken. Then $X, Y \sim \text{Uniform}(0, 1)$, and since they are independent, (X, Y) is uniformly distributed on the unit square in $\mathbb{R}^2$. For a valid triangle, the sum of the length of any two sides must be greater than the third. This is true if and only if the length of each piece is less than $\frac{1}{2}$.

   We analyze when $x < y$ as the case for $x > y$ will be identical. For $x < y$, we need that $x < \frac{1}{2}, y - x < \frac{1}{2}$, and $1 - y < \frac{1}{2}$. Graphing this on the unit square and calculating the area gives a triangle with vertices $(0, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, 1)$, and its area is $\frac{1}{8}$. Now adding the case where $x > y$, the probability that the pieces can be put together into a triangle is just $2 \cdot \frac{1}{8} = \frac{1}{4}$.

13. (a)

14.

15. For the distribution of $X|X + Y$,

16. 4.16

4

17. 4.17

18. We can see that $f(x, y) \geq 0$ for all $x, y > 0$: The numerator is always nonnegative as $g(x)$ is nonnegative, and the denominator is always greater than 0. Now we show that $f(x, y)$ integrates to 1:

$$
\begin{aligned}
\int_0^\infty \int_0^\infty f(x, y) \, dxdy &= \int_0^\infty \int_0^\infty \frac{2g(\sqrt{x^2 + y^2})}{\pi \sqrt{x^2 + y^2}} \, dxdy \\
&= \frac{2}{\pi} \int_0^\infty \int_0^{\pi/2} \frac{g(r)}{r} \cdot r \, d\theta dr \\
&= \frac{2}{\pi} \int_0^{\pi/2} \int_0^\infty g(r) \, dr d\theta \\
&= \frac{2}{\pi} \int_0^{\pi/2} 1 \, d\theta \\
&= 1.
\end{aligned}
$$

19. 4.20

20. 4.21ß