ECE 534 Recitation

Chao Pan

2/10/2022

1.23 Correlation of histogram values Suppose that n fair dice are independently rolled. Let

$$X_i = \begin{cases} 1 & \text{if a 1 shows on the } i^{th} \text{ roll} \\ 0 & \text{else} \end{cases} \qquad Y_i = \begin{cases} 1 & \text{if a 2 shows on the } i^{th} \text{ roll} \\ 0 & \text{else} \end{cases}.$$

Let X denote the sum of the X_i's, which is simply the number of 1's rolled. Let Y denote the sum of the Y_i's, which is simply the number of 2's rolled. Note that if a histogram is made recording the number of occurrences of each of the six numbers, then X and Y are the heights of the first two entries in the histogram.
(a) Find $E[X_1]$ and $\text{Var}(X_1)$.
(b) Find $E[X]$ and $\text{Var}(X)$.
(c) Find $\text{Cov}(X_i, Y_j)$ if $1 \leq i, j \leq n$ (Hint: Does it make a difference if $i = j$?)
(d) Find $\text{Cov}(X, Y)$ and the correlation coefficient $\rho(X, Y)$.
(e) Find $E[Y|X = x]$ for any integer x with $0 \leq x \leq n$. Note that your answer should depend on x and n, but otherwise your answer is deterministic.

1.23 Correlation of histogram values (a) X_1 is Bernoulli($\frac{1}{6}$), so $E[X_1] = \frac{1}{6}$ and $\text{Var}(X_1) = \frac{1}{6}(1-\frac{1}{6}) = \frac{5}{36}$.
(b) $E[X] = nE[X_1] = \frac{n}{6}$ and $\text{Var}(X) = n\text{Var}(X_1) = \frac{5n}{36}$.
(c) We begin by computing $\text{Cov}(X_1, Y_1)$. Since $X_1Y_1 = 0$ with probability one, $E[X_1Y_1] = 0$. Therefore $\text{Cov}(X_1, Y_1) = E[X_1Y_1] - E[X_1]E[Y_1] = 0 - \frac{1}{6}\frac{1}{6} = \frac{-1}{36}$. So $\text{Cov}(X_i, Y_i) = \frac{-1}{36}$ for any i. On the other hand, if $i \neq j$ then X_i is independent of X_j. So

$$\text{Cov}(X_i, Y_j) = \begin{cases} \frac{-1}{36} & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

(d)

$$\text{Cov}(X, Y) = \sum_i \sum_j \text{Cov}(X_i, Y_j) = \sum_i \text{Cov}(X_i, Y_i) = n\text{Cov}(X_1, Y_1) = \frac{-n}{36}.$$

and

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{-1}{5}.$$

(e) Given that x of the dice show a 1, each of the remaining dice is equally likely to show 2,3,4,5, or 6. Thus, each of the remaining $n - x$ dice shows a 2 with conditional probability $\frac{1}{5}$. Therefore $E[Y|X = x] = \frac{n-x}{5}$.

Formal solution to (e) by definition: By definition,

$$\mathbb{E}[Y|X=x] = \mathbb{E}\left[\sum_{i=1}^{n} Y_i \middle| X=x\right] = \sum_{i=1}^{n} \mathbb{E}[Y_i|X=x].$$

If $x \le n-1$, consider term $\mathbb{E}[Y_1|X=x] = 0 \cdot \mathbb{P}(Y_1=0|X=x) + 1 \cdot \mathbb{P}(Y_1 = 1|X=x)$. For the second part, we have

$$\mathbb{P}(Y_1=1|X=x) = \frac{\mathbb{P}(Y_1=1, X=x)}{\mathbb{P}(X=x)} = \frac{\frac{1}{6}\binom{n-1}{x}\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{n-1-x}}{\binom{n}{x}\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{n-x}}$$
$$= \frac{n-x}{5n}.$$

Due to symmetry, $\mathbb{E}[Y|X=x] = n\mathbb{E}[Y_1|X=x] = \frac{n-x}{5}$.

If $x=n$, then obviously $\mathbb{E}[Y|X=x]=0$.

1.25 A function of jointly distributed random variables Suppose (U, V) is uniformly distributed over the square with corners (0,0), (1,0), (1,1), and (0,1), and let $X = UV$. Find the CDF and pdf of X.

- General solution: Find CDF first, then take derivative to get pdf. For CDF of X, we have

$$F_X(c) = \mathbb{P}(X \leq c) = \mathbb{P}(UV \leq c) = \int_{u=0}^{1} \mathbb{P}\left(V \leq \frac{c}{u}\right) f_U(u) du. \quad (7)$$

When $0 \leq c \leq 1$, if $0 \leq u \leq c$, $\frac{c}{u} \geq 1$ and thus $\mathbb{P}\left(V \leq \frac{c}{u}\right) = 1$, since V is a uniform distribution over $[0, 1]$; if $u \geq c$, $\mathbb{P}\left(V \leq \frac{c}{u}\right) = \frac{c}{u}$. Therefore, $F_X(c)$ becomes

$$\int_{u=0}^{c} 1 du + \int_{u=c}^{1} \frac{c}{u} du = c - c \ln c, \ 0 \leq c \leq 1.$$

Obviously $F_X(c) = 0$ if $c < 0$ and $F_X(c) = 1$ if $c > 1$.

Geometric Interpretation:

1.25 A function of jointly distributed random variables The square has unit area so that the joint density is unit valued within the square. The range of X is the interval $[0,1]$, so fix c in $[0,1]$ and consider the event $\{UV \leq c\}$. The probability of this event is the area of the square minus the upper right region above the curve $v = c/u$. This area is one minus the area of the region inside the square above the curve $v = c/u$. Therefore,

$$F_X(c) = \begin{cases} 0 & c \leq 0 \\ 1 - \int_c^1 (1 - \frac{c}{u})du = c - c\ln c & 0 \leq c \leq 1 \\ 1 & c \geq 1 \end{cases}$$

Differentiating yields

$$f_X(c) = \begin{cases} -\ln c & 0 < c \leq 1 \\ 0 & \text{else} \end{cases}$$

1.29 Uniform density over a union of two square regions Let the random variables X and Y be jointly uniformly distributed on the region $\{0 \leq u \leq 1, 0 \leq v \leq 1\} \cup \{-1 \leq u < 0, -1 \leq v < 0\}$. (a) Determine the value of f_{XY} on the region shown.

(b) Find f_X, the marginal pdf of X.

(c) Find the conditional pdf of Y given that $X = a$, for $0 < a \leq 1$.

(d) Find the conditional pdf of Y given that $X = a$, for $-1 \leq a < 0$.

(e) Find $E[Y|X = a]$ for $|a| \leq 1$.

(f) What is the correlation coefficient of X and Y?

(g) Are X and Y independent?

(h) What is the pdf of $Z = X + Y$?

1.29 Uniform density over a union of two square regions (a) Region has area 2 so the density function is 1/2 in the region and zero outside.

(b) $f_X(x) = \begin{cases} 0.5 & \text{if } |x| \leq 1 \\ 0 & \text{else} \end{cases}$

(c) If $0 < a \leq 1$, $f_{Y|X}(y|a) = \begin{cases} 1 & \text{if } 0 \leq y \leq 1 \\ 0 & \text{else} \end{cases}$

(d) If $-1 \leq a < 0$, $f_{Y|X}(y|a) = \begin{cases} 1 & \text{if } -1 \leq y \leq 0 \\ 0 & \text{else} \end{cases}$

(e) $E[Y|X = a] = \begin{cases} -0.5 & \text{if } -1 \leq a < 0 \\ 0.5 & \text{if } 0 < a < 1 \end{cases}$

(f) $E[X] = E[Y] = 0$, $\mathrm{Var}(X) = E[X^2] = 1/3$, $\mathrm{Var}(Y) = 1/3$, $E[XY] = \frac{1}{2}\int_0^1\int_0^1 xydxdy + \frac{1}{2}\int_{-1}^0\int_{-1}^0 xydxdy = \int_0^1\int_0^1 xydxdy = 1/4$. So $\rho_{XY} = \frac{1/4}{\sqrt{1/3 \times 1/3}} = \frac{3}{4}$.

(g) No, because $f_{XY}(x,y)$ doesn't factor into the product of a function of x and a function of y.

(h) The range of Z is $[-2, 2]$. $f_Z(z) = \begin{cases} |z|/2 & \text{if } -0 \leq |z| \leq 1 \\ 1 - |z|/2 & \text{if } 1 \leq |z| \leq 2 \\ 0 & \text{else} \end{cases}$ (Shape is two triangles.)

Suppose there are two i.i.d. random variables $X, Y \sim \text{Unif}[0, 1]$. Let $Z = X + Y$, what is the CDF and pdf of Z.

Solution: There are three types of methods to solve this problem.

- General solution: Find CDF first, then take derivative to get pdf. For CDF of Z, we have

$$\mathbb{P}(Z \le z) = \mathbb{P}(X + Y \le z) = \int_{y=0}^{1} \mathbb{P}(X \le z - y) f_Y(y) dy. \tag{1}$$

When $0 \le z \le 1$, for region $0 \le y \le z$, $\mathbb{P}(X \le z - y) = z - y$; for region $z < y \le 1$, $\mathbb{P}(X \le z - y) = 0$. So (1) becomes

$$\int_{y=0}^{z} (z - y) dy = \frac{z^2}{2}, \ 0 \le z \le 1. \tag{2}$$

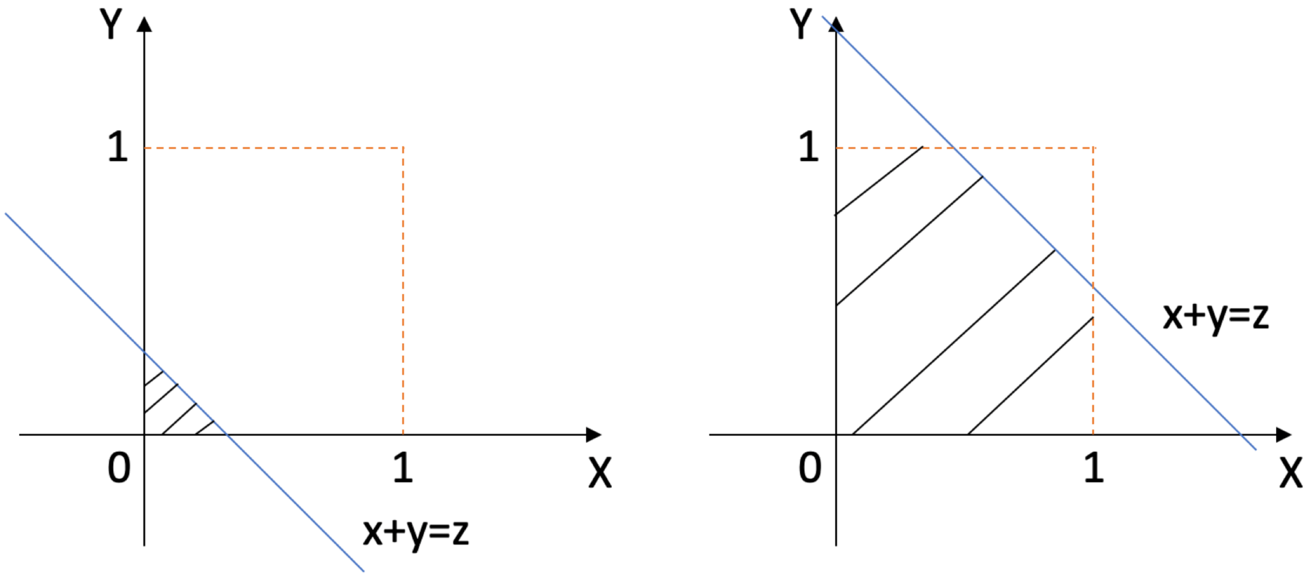
When $1 < z \le 2$, for region $0 \le y \le z - 1$, $\mathbb{P}(X \le z - y) = 1$; for region $z - 1 < y \le 1$, $\mathbb{P}(X \le z - y) = z - y$. So (1) becomes

$$\int_{y=0}^{z-1} 1 dy + \int_{y=z-1}^{1} (z - y) dy = \frac{-z^2 + 4z - 2}{2}, \ 1 < z \le 2. \tag{3}$$

For $z < 0$, $F_Z(z) = 0$ and $z > 2$, $F_Z(z) = 1$, which is obvious. To get pdf of Z we just need to take the derivative:

$$f_Z(z) = \begin{cases} 0 & z < 0 \\ z & 0 \le z \le 1 \\ 2 - z & 1 < z \le 2 \\ 0 & z > 2 \end{cases}$$

- Geometric interpretation: Consider X, Y as two axis on the 2D plane. $x + y = z$ is a line on the plane and $z + y \leq z$ corresponds to the region below this line. So the CDF becomes the area in shadow as shown in the following figure.



We can thus get the same CDF as

$$F_Z(z) = \begin{cases} 0 & z < 0 \\ \frac{z^2}{2} & 0 \leq z \leq 1 \\ 1 - \frac{(2-z)^2}{2} & 1 < z \leq 2 \\ 1 & z > 2 \end{cases}$$

Taking derivative of $F_Z(z)$, we can obtain $f_Z(z)$.

- Convolution of probability distributions (`https://en.wikipedia.org/wiki/Convolution_of_probability_distributions`): If X, Y are independent distributions, then $Z = X + Y$ has pdf

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx. \tag{4}$$

To make sure $f_X(x)$ and $f_Y(z - x)$ are both 1, it requires $0 \leq x \leq 1$ and $0 \leq z - x \leq 1$. Thus when $0 \leq z \leq 1$, (4) becomes

$$f_Z(z) = \int_0^z 1 dx = z. \tag{5}$$

When $1 < z \leq 2$, (4) becomes

$$f_Z(z) = \int_{z-1}^1 1 dx = 2 - z. \tag{6}$$

And $f_Z(z) = 0$ elsewhere. The result is the same as in the first solution.