Probability and Statistics I

STAT 3600 – Fall 2021

Le Chen
lzc0090@auburn.edu

Last updated on

July 4, 2021

Auburn University
Auburn AL

Chapter 4. Bivariate Distributions

Definition 4.1-1 Let X and Y be two random variables defined on a discrete probability space. Let S denote the corresponding two-dimensional space of X and Y, the two random variables of the discrete type. The probability that $X = x$ and $Y = y$ is denoted by

$$f(x, y) = \mathbb{P}(X = x, Y = y).$$

The function $f(x, y)$ is called the *joint probability mass function (joint pmf)* of X and Y and has the following properties:
(a) $0 \leq f(x, y) \leq 1$.
(b) $\sum \sum_{(x,y) \in S} f(x, y) = 1$.
(c) $\mathbb{P}[(X, Y) \in A] = \sum \sum_{(x,y) \in A} f(x, y)$, where A is a subset of the space S.

Definition 4.1-2 The *joint cdf* of a discrete bivariate r.v. (X, Y) is given by

$$F_{XY}(x, y) = \mathbb{P}(X \leq x, Y \leq y) = \sum_{x_i \leq x} \sum_{y_i \leq y} f(x_i, y_i).$$

Example 4.1-1 Consider an experiment of tossing a fair coin twice. Let (X, Y) be a bivariate r.v., where X is the number of heads that occurs in the two tosses and Y is the number of tails that occurs in the two tosses.
(a) What is the range R_X of X?
(b) What is the range R_Y of Y?
(c) Find the range R_{XY} of (X, Y).
(d) Find $\mathbb{P}(X = 2, Y = 0)$, $P(X = 0, Y = 2)$ and $\mathbb{P}(X = 1, Y = 1)$.

Example 4.1-2 Roll a pair of unbiased dice. For each of the 36 sample points with probability $1/36$, let X denote the smaller and Y the larger outcome on the dice. For example, if the outcome is $(3, 2)$, then the observed values are $X = 2$, $Y = 3$. The event $\{X = 2, Y = 3\}$ could occur in one of two ways - $(3, 2)$ or $(2, 3)$- so its probability is

$$\frac{1}{36} + \frac{1}{36} = \frac{2}{36}.$$

If the outcome is $(2, 2)$, then the observed values are $X = 2$, $Y = 2$. Since the event $\{X = 2, Y = 2\}$ can occur in only one way, $\mathbb{P}(X = 2, Y = 2) = 1/36$. The joint pmf of X and Y is given by the probabilities

$$f(x, y) = \begin{cases} \frac{1}{36}, & \text{if } 1 \leq x = y \leq 6 \\ \frac{2}{36}, & \text{if } 1 \leq x < y \leq 6. \end{cases}$$

when x and y are integers.

Definition 4.1-3 Let X and Y have the joint probability mass function $f(x, y)$ with space S. The probability mass function of X alone, which is called the *marginal probability mass function of X*, is defined by

$$f_1(x) = f_X(x) = \sum_y f(x, y) = \mathbb{P}(X = x), \qquad x \in S_1$$

where the summation is taken over all possible y values for each given x in the x space S_1. That is, the summation is over all (x, y) in S with a given x value.

Similarly, the *marginal probability mass function of Y* is defined by

$$f_2(y) = f_Y(y) = \sum_x f(x, y) = \mathbb{P}(Y = y), \qquad y \in S_2,$$

where the summation is taken over all possible x values for each given y in the y space S_2.

Definition 4.1-4 The random variables X and Y are *independent* if and only if

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y), \qquad \forall x \in S_1, \quad y \in S_2,$$

or equivalently,

$$f(x, y) = f_1(x)f_2(y), \qquad \forall x \in S_1, \quad y \in S_2.$$

Otherwise X and Y are said to be *dependent*.

Example 4.1-3 Let the joint pmf of X and Y be defined by

$$f(x, y) = \frac{x + y}{32}, \qquad x = 1, 2, \quad y = 1, 2, 3, 4.$$

(a) Find $f_1(x)$, the marginal pmf of X.
(b) Find $f_2(y)$, the marginal pmf of Y.
(c) Find $\mathbb{P}(X > Y)$.
(d) Find $\mathbb{P}(Y = 2X)$.
(e) Find $\mathbb{P}(X + Y = 3)$.
(f) Find $\mathbb{P}(X \leq 3 - Y)$.
(g) Are X and Y independent or dependent? Why or why not?
Ans.
(a)
(b)
(c)
(d)
(e)
(f)
(g)

Example 4.1-4 The joint pmf of a bivariate r.v. (X, Y) is given by

$$f(x, y) = \begin{cases} k(2x + y), & \text{if } x = 1, 2; y = 1, 2 \\ 0, & \text{otherwise.} \end{cases}$$

where k is a constant.
(a) Find the value of k.
(b) Find the marginal pmf's of X and Y.
(c) Are X and Y independent?
Ans.
(a)
(b)
(c)

Definition 4.1-5 If X_1 and X_2 are random variables of discrete type with the joint pmf $f(x_1, x_2)$ on the space S. If $u(X_1, X_2)$ is a function of these two random variables, then

$$\mathbb{E}[u(X_1, X_2)] = \sum\sum_{(x_1, x_2) \in S} u(x_1, x_2) f(x_1, x_2),$$

if it exists, is called the *mathematical expectation (or expected value)* of $u(X_1, X_2)$.

Example 4.1-5 There are eight chips in a bowl: three marked $(0,0)$, two marked $(1,0)$, two marked $(0,1)$, and one marked $(1,1)$. A player selects a chip at random and is given the sum of the coordinates in dollars. If X_1 and X_2 represent those two coordinates, respectively, their joint pmf is

$$f(x_1, x_2) = \frac{3 - x_1 - x_2}{8}, \qquad x_1 = 0, 1; x_2 = 0, 1.$$

Find the expected payoff.

Solution. Let $u(X_1, X_2) = \text{payoff} = \$(X_1 + X_2)$. Thus, the expected payoff is given by

$$\begin{aligned} E(u(X_1, X_2)) = \mathbb{E}(X_1 + X_2) &= \sum_{x_2=0}^{1} \sum_{x_1=0}^{1} (x_1 + x_2)\left(\frac{3 - x_1 - x_2}{8}\right) \\ &= (0)\left(\frac{3}{8}\right) + (1)\left(\frac{2}{8}\right) + (1)\left(\frac{2}{8}\right) + (2)\left(\frac{1}{8}\right) \\ &= \frac{3}{4}. \end{aligned}$$

That is, the expected payoff is 75 cents. □

The following mathematical expectations, if they exist, have special names:

(a) If $u_1(X_1, X_2) = X_i$, then

$$\mathbb{E}[u_1(X_1, X_2)] = \mathbb{E}(X_i) = \mu_i$$

is called the mean of $X_i, i = 1, 2$.

(b) If $u_2(X_1, X_2) = (X_i - \mu_i)^2$, then

$$\mathbb{E}[u_2(X_1, X_2)] = \mathbb{E}[(X_i - \mu_i)^2] = \sigma_i^2 = Var(X_i)$$

is called the variance of $X_i, i = 1, 2$.

Remark 4.1-1 The mean μ_i and the variance σ_i^2 can be computed either from the joint pmf $f(x_1, x_2)$ or from the marginal pmf $f_i(x_i), i = 1, 2$.

We now extend the binomial distribution to a trinomial distribution. Here we have three mutually exclusive and exhaustive ways for an experiment to terminate: perfect, "seconds," and defective. We repeat the experiment n independent times, and the probabilities $p_1, p_2, p_3 = 1 - p_1 - p_2$ of perfect, seconds, and defective,respectively, remain the same from trial to trial. In the n trials, let X_1 =number of perfect items, $X_2 =$ number of seconds, and $X_3 = n - X_1 - X_2 =$ number of defectives. If x_1 and x_2 are nonnegative integers such that $x_1 + x_2 \leq n$, then the probability of having x_1 perfect, x_2 seconds, and $n - x_1 - x_2$ defectives, in that order, is

$$p_1^{x_1} p_2^{x_2} (1 - p_1 - p_2)^{n - x_1 - x_2}.$$

However, if we want $\mathbb{P}(X_1 = x_1, X_2 = x_2)$, then we must recognize that $X_1 = x_1, X_2 = x_2$ can be achieved in

$$\binom{n}{x_1, x_2, n - x_1 - x_2} = \frac{n!}{x_1! x_2! (n - x_1 - x_2)!}$$

different ways. Hence, the **trinomial** pmf is given by

$$\begin{aligned} f(x_1, x_2) &= P(X_1 = x_1, X_2 = x_2) \\ &= \frac{n!}{x_1! x_2! (n - x_1 - x_2)!} p_1^{x_1} p_2^{x_2} (1 - p_1 - p_2)^{n - x_1 - x_2}, \end{aligned}$$

where x_1 and x_2 are nonnegative integers such that $x_1 + x_2 \leq n$. Without summing, we know that X_1 is $b(n, p_1)$ and X_2 is $b(n, p_2)$; thus, X_1 and X_2 are dependent, as the product of these marginal probability mass function is not equal to $f(x_1, x_2)$.

Example 4.1-6 A manufactured item is classified as good, a "second," or defective with probability $6/10, 3/10$, and $1/10$, respectively. Fifteen such items are selected at random from the production line. Let X denote the number of good items, Y the number of seconds, and $15 - X - Y$ the number of defective items.
(a) Give the joint pmf of X and Y, $f(x, y)$.
(b) Sketch the set of points for which $f(x, y) > 0$. From the shape of this region, can X and Y be independent? Why or why not?
(c) Find $\mathbb{P}(X = 10, Y = 4)$.
(d) Give the marginal pmf of X.
(e) Find $\mathbb{P}(X \leq 11)$.

Ans.
(a) $f(x, y) = \frac{15!}{x!y!(15-x-y)!}(0.6)^x(0.3)^y(0.1)^{n-x-y}$
(b) no, because the space is not rectangular.
(c) 0.0735.
(d) X is $b(15, 0.6)$.
(e) 0.9095.

Exercises from textbook: 4.1-2, 4.1-3, 4.1-4,4.1-5, 4.1-6, 4.1-8,

Definition 4.2-1 The *covariance* of X and Y, denoted by $\mathrm{Cov}(X, Y)$ or σ_{XY}, is defined by

$$\mathrm{Cov}(X, Y) = \sigma_{XY} = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)] = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

If $\mathrm{Cov}(X, Y) = 0$, then we say that X and Y are *uncorrelated*. X and Y are uncorrelated if and only if

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

Remark 4.2-1 Note that if X and Y are independent, then it can be show that they are uncorrelated, namely,

$$\text{independent} \quad \Rightarrow \quad \text{uncorrelated.}$$

However, the converse is not true in general; that is, the fact that X and Y are uncorrelated does not, in general, imply that they are independent:

$$\text{independent} \quad \nLeftarrow \quad \text{uncorrelated.}$$

Definition 4.2-2 For two random variables X and Y, the *correlation coefficient*, denoted by ρ_{XY}, is defined by

$$\rho_{XY} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Remark 4.2-2 It can be shown that

$$|\rho| \leq 1 \quad \text{or} \quad -1 \leq \rho \leq 1.$$

Remark 4.2-3 Note that the correlation coefficient of X and Y is a measure of linear dependence between X and Y. The *least square regression line* (the line that describes linear relationship between X and Y) is given by

$$y = \mu_Y + \rho_{XY} \frac{\sigma_Y}{\sigma_X} (x - \mu_X).$$

Example 4.2-1 Let X and Y have the joint pmf

$$f(x, y) = \frac{x + y}{32}, \quad x = 1, 2, y = 1, 2, 3, 4.$$

Find the mean μ_X and μ_Y, the variances σ_X^2 and σ_Y^2, the correlation coefficient ρ, and the equation of the least square regression line. Are X and Y independent?

Ans:
$\mu_X = 25/16$
$\mu_Y = 45/16$
$\sigma_X^2 = 63/256$
$\sigma_Y^2 = 295/256$
$\mathrm{Cov}(X, Y) = -5/256$
$\rho = -0.0367$ dependent.

Example 4.2-2 Let X and Y be random variables of the continuous type having the joint pdf

$$f(x, y) = 2, \qquad 0 \leq y \leq x \leq 1.$$

(a) Find the marginal pdf of X and Y.
(b) Compute $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \text{Cov}(X, Y)$, and ρ.
(c) Determine the equation of the least square regression line.
Ans:
(a) $f_1(x) = 2x$ for $0 \leq x \leq 1$; $f_2(y) = 2(1 - y)$;
(b) $\mu_X = \mathbb{E}(X) = 2/3$ and $\mu_Y = \mathbb{E}(Y) = 1/3$.

Exercises from textbook: 4.2-1, 4.2-2, 4.2-3, 4.2-7, 4.2-9.

Definition 4.3-1 Let X and Y have a joint discrete distribution with pmf $f(x, y)$ on space S. Say the marginal probability mass functions are $f_1(x)$, and $f_2(y)$ with space S_1 and S_2, respectively. The *conditional probability mass function of X*, given that $Y = y$, is defined by

$$g(x|y) = \frac{f(x, y)}{f_2(y)} \quad \text{provided that } f_2(y) > 0.$$

Similarly, the *conditional probability mass function of Y*, given that $X = x$, is defined by

$$h(y|x) = \frac{f(x, y)}{f_1(x)} \quad \text{provided that } f_1(x) > 0.$$

Definition 4.3-2 If (X, Y) is a discrete bivariate r.v. with joint pmf $f(x, y)$, then the *conditional mean (or conditional expectation)* of Y, given that $X = x$, is defined by

$$\mu_{Y|x} = \mathbb{E}(Y|x) = \sum_{y} yh(y|x),$$

and the *conditional variance* of Y, given that $X = x$, is defined by

$$\sigma^2_{Y|x} = \mathbb{E}\{[Y - \mathbb{E}(Y|x)]^2|x\} = \sum_{y}[y - \mathbb{E}(Y|x)]^2 h(y|x)$$

which can be reduced to

$$\text{Var}(Y|x) = \sigma^2_{Y|x} = \mathbb{E}(Y^2|x) - [E(Y|x)]^2 = \mathbb{E}(Y^2|x) - (\mu_{Y|x})^2.$$

Example 4.3-1 Let X and Y have a uniform distribution on the set of points with integer coordinates in $S = \{(x, y) : 0 \leq x \leq 7, x \leq y \leq x + 2\}$. That is, $f(x, y) = 1/24, (x, y) \in S$, and both x and y are integers.
(a) Find $f_1(x)$.
(b) Find $h(y|x)$.
(c) Find $\mu_{Y|x} = \mathbb{E}(Y|x)$.
(d) Find $\sigma^2_{Y|x}$.
Ans.
(a) $f_1(x) = 1/8, x = 0, 1, \cdots 7;$
(b) $h(y|x) = 1/3, y = x, x + 1, x + 2$, for $x = 0, 1, \cdots, 7;$
(c) $\mu_{Y|x} = x + 1, x = 0, 1, \cdots, 7.$
(d) $\sigma^2_{Y|x} = 2/3.$

Exercises from textbook: 4.3-1, 4.3-2, 4.3-5, 4.3-6, 4.3-10

Definition 4.4-1 *The joint probability density function (joint pdf)* of two continuous-type random variables is an integrable function $f(x, y)$ with the following properties:

(a) $f(x, y) \geq 0$, where $f(x, y) = 0$ only when (x, y) is not in the support (space) S of X and Y.

(b) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dxdy = 1.$

(c) $\mathbb{P}[(X, Y) \in A] = \iint_A f(x, y)\rangle dxdy$ where $\{(X, Y) \in A\}$ is the event defined in the plane.

Example 4.4-1

The joint pdf of a bivariate r.v. (X, Y) is given by

$$f(x, y) = \begin{cases} kxy, & \text{if } 0 < x < 1, 0 < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

where k is a constant.

(a) Find the value of k.

(b) Are X and Y independent?

(c) Find $\mathbb{P}(X + Y < 1)$

(d) Compute $\mu_X; \mu_Y; \sigma_X^2; \sigma_Y^2$.

Ans.

(a) $k = 4$

(b) they are independent.

(c) $1/6$.

Definition 4.4-2 If (X, Y) is a continuous bivariate r.v. with joint pdf $f(x, y)$, then the *conditional pdf of X*, given that $Y = y$, is defined by

$$g(x|y) = \frac{f(x, y)}{f_2(y)}, \quad f_2(y) > 0.$$

Similarly, the *conditional pdf of Y*, given that $X = x$, is defined by

$$h(y|x) = \frac{f(x, y)}{f_1(x)}, \quad f_1(x) > 0.$$

Definition 4.4-3 If (X, Y) is a continuous bivariate r.v. with joint pdf $f(x, y)$, the *conditional mean* of Y, given that $X = x$ is defined by

$$\mu_{Y|x} = \mathbb{E}(Y|x) = \int_{-\infty}^{\infty} yh(y|x)dy.$$

The *conditional variance* of Y, given that $X = x$, is defined by

$$\sigma^2_{Y|x} = Var(Y|x) = \mathbb{E}[(Y - \mu_{Y|x})^2|x] = \int_{-\infty}^{\infty} (y - \mu_{Y|x})^2 h(y|x)dy$$

which can be reduced to

$$\text{Var}(Y|x) = \mathbb{E}(Y^2|x) - (\mu_{Y|x})^2.$$

Example 4.4-2 Let $f(x, y) = 1/40, 0 \le x \le 10, 10 - x \le y \le 14 - x$, be the joint pdf of X and Y.
(a) Find $f_1(x)$, the marginal pdf of X.
(b) Determine $h(y|x)$, the conditional pdf of Y, given that $X = x$.
(c) Calculate $\mathbb{E}(Y|x)$, the conditional mean of Y, given that $X = x$.
(d) Find $\mathbb{P}(9 \le Y \le 11|X = 2)$.
Ans.
(a) $f_1(x) = 1/10, 0 \le x \le 10$;
(b) $h(y|x) = 1/4, 10 - x \le y \le 14 - x$ for $0 \le x \le 10$;
(c) $\mathbb{E}(Y|x) = 12 - x$.

Definition 4.5-1 A bivariate r.v. (X, Y) is called the *bivariate normal(or gaussian) distribution* if the joint pdf is given by

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left[-\frac{q(x,y)}{2}\right],$$

where

$$q(x,y) = \frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right].$$

A joint pdf of this form is called a *bivariate normal pdf*.

The marginal pdf of X is

$$f_1(x) = f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}}\exp\left[-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right]$$

and so the conditional distribution of Y, given that $X = x$, is a normal distribution with conditional mean

$$\mathbb{E}(Y|x) = \mu_{Y|x} = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)$$

and conditional variance

$$\mathrm{Var}(Y|x) = \sigma_{Y|x}^2 = \sigma_Y^2(1-\rho^2).$$

Example 4.5-1 Let X and Y have a bivariate normal distribution with parameters $\mu_X = -3, \mu_Y = 10, \sigma_X^2 = 25, \sigma_Y^2 = 9$, and $\rho = 3/5$.
(a) Compute $\mathbb{P}(-5 < X < 5)$.
(b) Compute $\mathbb{P}(-5 < X < 5 | Y = 13)$.
(c) Compute $\mathbb{P}(7 < Y < 16)$.
(d) Compute $\mathbb{P}(7 < Y < 16 | X = 2)$.
Ans.
(a) 0.6006;
(b) 0.7888;
(c) 0.8185;
(d) 0.9371.

Theorem 4.5-1 If X and Y have a bivariate distribution with correlation coefficient ρ, then X and Y are independent if and only if $\rho = 0$.

Exercises from textbook: 4.5-1, 4.5-3, 4.5-6, 4.5-7, 4.5-8.