ST 501: Fundamentals of Statistical Inference

Multivariate probability distributions (Part I)

Fall 2024

Department of Statistics, North Carolina State University.

Bivariate discrete random variable

Definition
Let $X = (X_1, X_2)$ where the X_1 and X_2 are discrete random variables. Then X is said to be a discrete bivariate random variable. The joint pmf for $X = (X_1, X_2)$ is defined as

$$p(x_1, x_2) = P(X_1 = x_1, X_2 = x_2)$$

Note If $X = (X_1, X_2)$ is a discrete bivariate rv then

1. $p(x_1, x_2) \geq 0$ for all x_1, x_2.
2. $\sum p(x_1, x_2) = 1$ where the summation is over all (x_1, x_2) with $p(x_1, x_2) > 0$.
3. For any subset $A \subset \mathbb{R}^2$, $P(A) = \sum p(x_1, x_2)$ where the sum is over all pairs $(x_1, x_2) \in A$ with $p(x_1, x_2) > 0$.

Example We roll a fair die twice. Let X_1 and X_2 be the faces shown on the first and second roll, respectively. Then $X = (X_1, X_2)$ takes on $6 \times 6 = 36$ possible values with

$$p(x_1, x_2) \equiv \frac{1}{36}, \qquad (x_1, x_2) \in \{1, 2, \dots, 6\} \times \{1, 2, \dots, 6\}.$$

Example We toss a fair coin 6 times. Let X_1 and X_2 be the number of heads and tails (in these 6 tosses), respectively. Then $X = (X_1, X_2)$ takes on 7 possible values, namely

$$X \in \mathcal{S} = \{(0,6), (1,5), (2,4), (3,3), (4,2), (5,1), (6,0)\}$$

and that

$$p(x_1, x_2) = \binom{6}{x_1} 2^{-6}, \qquad (x_1, x_2) \in \mathcal{S}.$$

Joint cumulative distribution function

Definition
For any bivariate random variable $X = (X_1, X_2)$, the *joint* cdf for X is.

$$F(x) = P(X_1 \leq x_1, X_2 \leq x_2), \qquad x = (x_1, x_2) \in \mathbb{R}^2.$$

Note If $X = (X_1, X_2)$ is a bivariate discrete rv then

$$F(x) = P(X_1 \leq x_1, X_2 \leq x_2) = \sum_{y \leq x} p(y_1, y_2)$$

where the sum is over all $y = (y_1, y_2) \leq x$ with $p(y_1, y_2) > 0$.

Example When tossing a fair die twice, for $x = (2, 4)$ we have

$$\begin{aligned} P(X \leq x) &= p(1,1) + p(1,2) + p(1,3) + p(1,4) \\ &\quad + p(2,1) + p(2,2) + p(2,3) + p(2,4) = \frac{8}{36}. \end{aligned}$$

Exampe When tossing a fair coin six times with X_1 and X_2 being the number of heads and tails, if $x = (3, 5)$ then

$$P(X \leq x) = p(1,5) + p(2,4) + p(3,3) = 2^{-6}\left(\binom{6}{1} + \binom{6}{2} + \binom{6}{3}\right) = \frac{41}{64}.$$

5

A joint cdf is monotone increasing in each argument.

Proposition
Let $X = (X_1, X_2)$ *be a bivariate rv with joint cdf* F. *Then*

$$\lim_{x_1 \to -\infty} F(x_1, x_2) = \lim_{x_2 \to -\infty} F(x_1, x_2) = 0, \quad \textit{for all } x_1, x_2 \in \mathbb{R}$$

$$\lim_{x_1 \to \infty} \lim_{x_2 \to \infty} F(x_1, x_2) = 1,$$

$$F(x_1^*, x_2^*) - F(x_1, x_2^*) - F(x_1^*, x_2) + F(x_1, x_2) \geq 0, \;\; x_1^* \geq x_1, x_2^* \geq x_2.$$

6

Marginal pmf

Definition

Let $X = (X_1, X_2)$ be a discrete bivariate random variable with joint pmf $p(x_1, x_2)$. The marginal pmf of X_1 (respectively X_2) is defined as

$$p_1(x_1) = \sum_z p(x_1, z), \quad p_2(x_2) = \sum_z p(z, x_2)$$

where the sum is over all (x_1, z) (respectively (z, x_2)) pairs with $p(x_1, z) > 0$ (respectively $p(z, x_2) > 0$).

Example When rolling a fair die twice (with X_1 and X_2 being the face shown on the first and second roll), we have

$$p_1(x_1) \equiv \frac{1}{6},\ x_1 \in \{1, 2, \ldots, 6\}; \quad p_2(x_2) \equiv \frac{1}{6},\ x_2 \in \{1, 2, \ldots, 6\}.$$

Example In the setting of tossing a fair coin six times (with X_1 and X_2 being the number of heads and the number of tails), we have

$$p_1(x_1) = \binom{6}{x_1} 2^{-6}, \quad p_2(x_2) = \binom{6}{x_2} 2^{-6}$$

Example From a group of three Republicans, two Democrats, and two independent, two people are randomly selected to form a committee. Let Y_1 and Y_2 be the number of Republicans and Democrats on the committee, respectively. Find the joint pmf for $Y = (Y_1, Y_2)$ and the marginal pmf for Y_1 and Y_2.

9

Example From a group of three Republicans, two Democrats, and two independent, two people are randomly selected to form a committee. Let Y_1 and Y_2 be the number of Republicans and Democrats on the committee, respectively. Find the joint pmf for $Y = (Y_1, Y_2)$ and the marginal pmf for Y_1 and Y_2.

The joint pmf for $Y = (Y_1, Y_2)$ is

$$p(y_1, y_2) = \frac{\binom{3}{y_1}\binom{2}{y_2}\binom{2}{2-y_1-y2}}{\binom{7}{2}}, \quad y_1 \geq 0, y_2 \geq 0, y_1 + y_2 \leq 2.$$

The following table gives the joint and the marginal pmf.

	$y_1 = 0$	$y_1 = 1$	$y_1 = 2$	p_2
$y_2 = 0$	1/21	4/21	1/21	6/21
$y_2 = 1$	6/21	6/21	0	12/21
$y_2 = 2$	3/21	0	0	3/21
p_1	10/21	10/21	1/21	

9

Bivariate continuous random variable

A univariate continuous rv is defined using an integral. Analogously, a bivariate continuous rv is defined in terms of a bivariate integral.

Definition
Let $X = (X_1, X_2)$ be a bivariate random variable with joint cdf F. Then X is said to be a bivariate continuous rv if and only if there exists a function f such that

$$F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(y_1, y_2)\, \mathrm{d}y_2 \mathrm{d}y_1$$

Any function f with the above property is a joint pdf for X.

1. For a bivariate continuous rv $X = (X_1, X_2)$,

$$P(X = x) = 0 \quad \text{for all } x = (x_1, x_2).$$

2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2)\,\mathrm{d}x_1 \mathrm{d}x_2 = 1$
3. For any set A that is a countable union (or countable intersection) of rectangles

$$P(X \in A) = \int_{(x_1,x_2)\in A} f(x_1, x_2)\,\mathrm{d}x_1 \mathrm{d}x_2$$

We can visualize $f(x, y)$ as specifying a surface at height $f(x, y)$ above the point (x, y) in a three-dimensional coordinate system. Then $P[(X, Y) \in A]$ is the volume underneath this surface and above the region A, analogous to the area under a curve in the case of a single rv. This is illustrated in Figure 5.1.

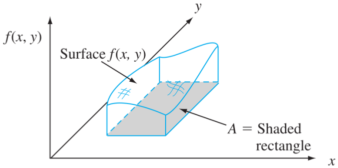

Figure 5.1 $P[(X, Y) \in A]$ = volume under density surface above A

Example Suppose $X = (X_1, X_2)$ has joint pdf

$$f(x_1, x_2) = \begin{cases} \frac{6}{5}(x_1 + x_2^2) & \text{for } 0 \le x_1 \le 1, 0 \le x_2 \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Compute $P(0 \le X_1 \le \frac{1}{4}, 0 \le X_2 \le \frac{1}{4})$.

13

Example Suppose $X = (X_1, X_2)$ has joint pdf

$$f(x_1, x_2) = \begin{cases} \frac{6}{5}(x_1 + x_2^2) & \text{for } 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Compute $P(0 \leq X_1 \leq \frac{1}{4}, 0 \leq X_2 \leq \frac{1}{4})$.

$$\begin{aligned} P(0 \leq X_1 \leq 1/4, 0 \leq X_2 \leq 1/4) &= \int_0^{1/4} \int_0^{1/4} \frac{6}{5}(x_1 + x_2^2)\, \mathrm{d}x_1\, \mathrm{d}x_2 \\ &= \frac{6}{20} \int_0^{1/4} x_1\, \mathrm{d}x_1 + \frac{6}{20} \int_0^{1/4} x_2^2\, \mathrm{d}x_2 \\ &= \frac{6}{40} x_1^2 \Big|_0^{1/4} + \frac{1}{10} x_2^3 \Big|_0^{1/4} \\ &= \frac{6}{640} + \frac{1}{640} = \frac{7}{640} \end{aligned}$$

EXAMPLE 5.4 Gasoline is to be stocked in a bulk tank once at the beginning of each week and then sold to individual customers. Let Y_1 denote the proportion of the capacity of the bulk tank that is available after the tank is stocked at the beginning of the week. Because of the limited supplies, Y_1 varies from week to week. Let Y_2 denote the proportion of the capacity of the bulk tank that is sold during the week. Because Y_1 and Y_2 are both proportions, both variables take on values between 0 and 1. Further, the amount sold, y_2, cannot exceed the amount available, y_1. Suppose that the joint density function for Y_1 and Y_2 is given by

$$f(y_1, y_2) = \begin{cases} 3y_1, & 0 \leq y_2 \leq y_1 \leq 1, \\ 0, & \text{elsewhere.} \end{cases}$$

A sketch of this function is given in Figure 5.4.

Find the probability that less than one-half of the tank will be stocked and more than one-quarter of the tank will be sold.

EXAMPLE 5.4 Gasoline is to be stocked in a bulk tank once at the beginning of each week and then sold to individual customers. Let Y_1 denote the proportion of the capacity of the bulk tank that is available after the tank is stocked at the beginning of the week. Because of the limited supplies, Y_1 varies from week to week. Let Y_2 denote the proportion of the capacity of the bulk tank that is sold during the week. Because Y_1 and Y_2 are both proportions, both variables take on values between 0 and 1. Further, the amount sold, y_2, cannot exceed the amount available, y_1. Suppose that the joint density function for Y_1 and Y_2 is given by

$$f(y_1, y_2) = \begin{cases} 3y_1, & 0 \le y_2 \le y_1 \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

A sketch of this function is given in Figure 5.4.

Find the probability that less than one-half of the tank will be stocked and more than one-quarter of the tank will be sold.

We are interested in $P(Y_1 \le 0.5, Y_2 \ge 0.25)$, i.e.,

$$\begin{aligned} P(Y_1 \le 0.5, Y_2 \ge 0.25) &= \int_0^{0.5} \int_{0.25}^{1} f(y_1, y_2)\, \mathrm{d}y_2 \mathrm{d}y_1 \\ &= \int_{0.25}^{0.5} \int_{0.25}^{y_1} 3y_1 \, \mathrm{d}y_2 \, \mathrm{d}y_1 \\ &= y_1^3 - \frac{3}{8} y_1^2 \Big|_{0.25}^{0.5} = \frac{5}{128} \end{aligned}$$

14

Marginal pmf

Definition
Let $X = (X_1, X_2)$ be a continuous bivariate rv with joint pdf $f(x_1, x_2)$. The marginal pdf of X_1 (respectively X_2) is defined as

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, z)\,\mathrm{d}z, \quad f_2(x_2) = \int_{-\infty}^{\infty} f(z, x_2)\,\mathrm{d}z.$$

Note Once again, $f_1(x_1)$ and $f_2(x_2)$ are not probabilities, i.e., $f_1(x_1) \neq P(X_1 = x_1)$ as $P(X_1 = x_1) = 0$ for all x_1. Rather

$$P(X_1 \in A) = \int_{x_1 \in A} f_1(x_1)\,\mathrm{d}x_1.$$

15

Example Let $X = (X_1, X_2)$ be a bivariate rv with joint pdf

$$f(x_1, x_2) = 2, \quad x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 1.$$

f is a valid joint pdf as

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1, x_2)\,\mathrm{d}x_1\mathrm{d}x_2 = \int_0^1\int_0^{1-x_2} 2\,\mathrm{d}x_1\mathrm{d}x_2 = \int_0^1 2(1-x_2)\,\mathrm{d}x_2 = 1$$

Q. What is the marginal pdf of X_1 and X_2 ?

16

Example Let $X = (X_1, X_2)$ be a bivariate rv with joint pdf

$$f(x_1, x_2) = 2, \quad x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 1.$$

f is a valid joint pdf as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2)\, \mathrm{d}x_1 \mathrm{d}x_2 = \int_0^1 \int_0^{1-x_2} 2\, \mathrm{d}x_1 \mathrm{d}x_2 = \int_0^1 2(1-x_2)\, \mathrm{d}x_2 = 1$$

Q. What is the marginal pdf of X_1 and X_2 ?

The marginal pdf of X_1 and X_2 are

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, z)\, \mathrm{d}z = \int_0^{1-x_1} 2\, \mathrm{d}z = 2(1 - x_1), \;\; x_1 \in [0, 1]$$

$$f_2(x_2) = \int_{-\infty}^{\infty} f(z, x_1)\, \mathrm{d}z = \int_0^{1-x_2} 2\, \mathrm{d}z = 2(1 - x_2), \;\; x_2 \in [0, 1]$$

Example Let $X = (X_1, X_2)$ be a bivarate rv with joint pdf

$$f(x_1, x_2) = c - 2(c-1)(x_1 + x_2 - 2x_1x_2) \quad x_1 \in [0,1], x_2 \in [0,1]$$

Here $c \in (0,2)$ is a constant.

Q. What is the marginal pdf of X_1 and X_2 ?

17

Example Let $X = (X_1, X_2)$ be a bivarate rv with joint pdf

$$f(x_1, x_2) = c - 2(c-1)(x_1 + x_2 - 2x_1 x_2) \quad x_1 \in [0,1], x_2 \in [0,1]$$

Here $c \in (0, 2)$ is a constant.

Q. What is the marginal pdf of X_1 and X_2 ?

The marginal pdf for X_1 and X_2 are

$$\begin{aligned} f_1(x_1) &= \int_0^1 c - 2(c-1)(x_1 + z - 2x_1 z)\,\mathrm{d}z = 1 \\ f_2(x_2) &= \int_0^1 c - 2(c-1)(z + x_2 - 2z x_2)\,\mathrm{d}z = 1 \end{aligned}$$

Note Both X_1 and X_2 are, marginally $\mathrm{Unif}(0, 1)$. However $X = (X_1, X_2)$ is not uniform on $[0, 1] \times [0, 1]$.

17

Multivariate random variable

It is straightforward to generalize the definition of a bivariate rv to a multivariate rv.

Definition
Let $X = (X_1, X_2, \ldots, X_m)$ where each of the X_i are discrete random variables. Then X is a discrete (multivariate) rv with joint pmf denoted by

$$p(x) = p(x_1, \ldots, x_m) = P(X = x) = P(X_1 = x_1, \ldots, X_m = x_m)$$

where $x = (x_1, x_2, \ldots, x_m) \in \mathbb{R}^m$.

Note For any $A \subset \mathbb{R}^m$,

$$P(A) = \sum_{x \in A} p(x)$$

where the summation is over all $x \in A$ for which $p(x) > 0$.

18

Definition

Let $X = (X_1, X_2, \ldots, X_m)$. Then X is said to be a continuous (multivariate) rv if there exists a function $f \colon \mathbb{R}^m \mapsto [0,1]$ such that for all $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$,

$$
\begin{aligned}
F(x) = P(X \leq x) &= P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_m \leq x_m) \\
&= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_m} f(z_1, z_2, \ldots, z_m) \, \mathrm{d}z_1 \mathrm{d}z_2 \cdots \mathrm{d}z_m .
\end{aligned}
$$

Note For any set $A \subset \mathbb{R}^m$ for which A is a countable union (or countable intersection) of hyper-rectangles

$$
P(A) = \int_{(x_1, \ldots, x_m) \in A} f(x_1, x_2, \ldots, x_m) \, \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_m .
$$

19

Note We rarely, if ever, need to compute the probability of an event for a (general) multivariate rv by explicitly summing or integrating over a set.

Rather, for a multivariate rv, most probability calculations are done via transforming the multivariate rv to some function (or functions) of simpler random variables (such as a collection of independent random variables).

20

Independent Random Variables

We revisit the topic of independent random variables (as previously discussed in the lecture slides for Chapter 4).

Definition
Let X_1 and X_2 be random variables with marginal cdf F_1 and F_2. Let $X = (X_1, X_2)$ has *joint* cdf F. Then X_1 and X_2 are independent if and only if for all $(x_1, x_2) \in \mathbb{R}^2$

$$F(x_1, x_2) = F_1(x_1)F_2(x_2).$$

Two rvs are said to be dependent if they are not independent.

21

We first characterize independence of two discrete rvs.

Proposition

Let X_1 and X_2 be discrete random variables with marginal pmf p_1 and p_2. Let $X = (X_1, X_2)$ has joint pmf p. Then X_1 and X_2 are independent if and only if for all $(x_1, x_2) \in \mathbb{R}^2$

$$p(x_1, x_2) = p_1(x_1)p_2(x_2).$$

22

We next characterize independence of two continuous rvs

Proposition
Let X_1 and X_2 be continuous rvs with marginal pdf f_1 and f_2. Suppose $X = (X_1, X_2)$ has joint pdf f. Then X_1 and X_2 are independent if and only if for all $(x_1, x_2) \in \mathbb{R}^2$

$$f(x_1, x_2) = f_1(x_1) f_2(x_2).$$

23

Generalizations of the previous ideas to three or more random variables are straightforward, e.g., $X_1, X_2, \ldots, X_m$ are mutually independent if and only if, for all $(x_1, x_2, \ldots, x_m)$

$$
\begin{aligned}
F(x_1, x_2, \ldots, x_m) &= P(X_1 \leq x_1, X_2 \leq x_2, \ldots X_m \leq x_m) \\
&= F_1(x_1) F_2(x_2) \cdots F_m(x_m)
\end{aligned}
$$

where the F_i are the marginal cdf.

24

Therefore, if the $X_1, X_2, \ldots, X_m$ are all discrete random variables, then they are mutually independent if and only if, for all $(x_1, x_2, \ldots, x_m)$,

$$p(x_1, x_2, \ldots, x_m) = p_1(x_1)p_2(x_2)\ldots p_m(x_m)$$

Similarly, if the $X_1, X_2, \ldots, X_m$ are all continuous random variables then they are mutually independent if and only if, for all $(x_1, x_2, \ldots, x_m)$,

$$f(x_1, x_2, \ldots, x_m) = f_1(x_1)f_2(x_2)\ldots f_m(x_m)$$

25

Example Let $Y = (Y_1, Y_2)$ has joint pdf

$$f(y_1, y_2) = \begin{cases} 6y_1y_2^2 & 0 \leq y_1 \leq 1,\ 0 \leq y_2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Show that Y_1 and Y_2 are independent.

26

Example Let $Y = (Y_1, Y_2)$ has joint pdf

$$f(y_1, y_2) = \begin{cases} 6y_1y_2^2 & 0 \leq y_1 \leq 1,\ 0 \leq y_2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Show that Y_1 and Y_2 are independent.

We first compute $f_1(y_1)$ and $f_2(y_2)$, i.e.,

$$f_1(y_1) = \int_0^1 6y_1y_2^2 \, \mathrm{d}y_2 = 2y_1y_2^3 \Big|_0^1 = 2y_1$$

$$f_2(y_2) = \int_0^1 6y_1y_2^2 \, \mathrm{d}y_2 = 3y_1^2y_2^2 \Big|_0^1 = 3y_2^2$$

and so $f(y_1, y_2) = f_1(y_1)f_2(y_2)$ for all (y_1, y_2).

26

The previous example generalizes to yield a quite useful result.

Proposition

Suppose $X = (X_1, X_2)$ has joint pdf f satisfying the conditions

1. *$f(x_1, x_2)$ is positive if and only if $(x_1, x_2) \in [a, b] \times [c, d]$.*
2. *There exists functions g and h such that*

   $$f(x_1, x_2) = g(x_1)h(x_2)$$

   for all $(x_1, x_2) \in [a, b] \times [c, d]$.

Then X_1 and X_2 are independent.

Example Let $Y = (Y_1, Y_2)$ has joint pdf

$$f(y_1, y_2) = \begin{cases} 2 & 0 \leq y_2 \leq y_1 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Show that Y_1 and Y_2 are not independent.

28

Example Let $Y = (Y_1, Y_2)$ has joint pdf

$$f(y_1, y_2) = \begin{cases} 2 & 0 \leq y_2 \leq y_1 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Show that Y_1 and Y_2 are not independent.

Let $y_2 = 0.5$ and $y_1 = 0.3$. Note that $f_1(y_1) > 0$ and $f_2(y_2) > 0$.

However, $f(y_1, y_2) = 0 \neq f_1(y_1)f_2(y_2)$.

In fact there exists (an uncountable number) of points (y_1, y_2) with $f(y_1, y_2) \neq f_1(y_1)f_2(y_2)$.

Therefore Y_1 and Y_2 are not independent.

28

10. Annie and Alvie have agreed to meet between 5:00 P.M. and 6:00 P M. for dinner at a local health-food restaurant. Let X = Annie's arrival time and Y = Alvie's arrival time. Suppose X and Y are independent with each uniformly distributed on the interval [5, 6].

a. What is the joint pdf of X and Y?

b. What is the probability that they both arrive between 5:15 and 5:45?

c. If the first one to arrive will wait only 10 min before leaving to eat elsewhere, what is the probability that they have dinner at the health-food restaurant? [*Hint*: The event of interest is $A = \{(x, y): |x - y| \leq 1/6\}$.]

29

For part (b) the answer is 1/4 as X and Y are independent and $P(5.25 \leq X \leq 5.75) = P(5.25 \leq Y \leq 5.75) = 1/2$.

For part (c) we first define $U = X - 5$ and $V = Y - 5$. We are interested in $p_* = P(|X - Y| \leq 1/6) = P(|U - V| \leq 1/6)$, i.e.,

$$
\begin{aligned}
p_* &= \int_0^{1/6} \int_0^{u+1/6} \mathrm{d}v \, \mathrm{d}u + \int_{1/6}^{5/6} \int_{v-1/6}^{v+1/6} \mathrm{d}v \, \mathrm{d}u + \int_{5/6}^{1} \int_{v-1/6}^{1} \mathrm{d}v \, \mathrm{d}u \\
&= \int_0^{1/6} (v + 1/6) \, \mathrm{d}v + \int_{1/6}^{5/6} \frac{2}{6} \, \mathrm{d}v + \int_{5/6}^{1} (7/6 - v) \, \mathrm{d}v \\
&= \left(\frac{v^2}{2} + \frac{v}{6} \Big|_0^{1/6} \right) + \left(\frac{v}{6} - \frac{v^2}{2} \Big|_{5/6}^{1} \right) + \frac{2}{9} = \frac{11}{36}.
\end{aligned}
$$

30

Example: Independent Exponential Random Variables

Q1a. Let X and Y be independent exponential rvs with rate parameters λ and ν, respectively. Find $P(X \leq Y)$.

31

Example: Independent Exponential Random Variables

Q1a. Let X and Y be independent exponential rvs with rate parameters λ and ν, respectively. Find $P(X \leq Y)$.

$$
\begin{aligned}
P(X \leq Y) &= \int_0^\infty \int_0^y \lambda \nu e^{-\lambda x} e^{-\nu y} \, \mathrm{d}x \, \mathrm{d}y \\
&= \int_0^\infty \nu e^{-\nu y} (1 - e^{-\lambda y}) \, \mathrm{d}y \\
&= \int_0^\infty \nu e^{-\nu y} - \nu e^{-(\lambda+\nu) y} \, \mathrm{d}y \\
&= -e^{-\nu y} + \frac{\nu}{\lambda + \nu} e^{-(\lambda+\nu) y} \Big|_0^\infty \\
&= 1 - \frac{\nu}{\lambda + \nu} = \frac{\lambda}{\lambda + \nu}.
\end{aligned}
$$

Q1b. Let X and Y be independent exponential rvs with rates λ and ν, respectively. Find the distribution of $X - Y$.

Q1b. Let X and Y be independent exponential rvs with rates λ and ν, respectively. Find the distribution of $X - Y$.

Let $Z = X - Y \in (\infty, \infty)$. We now compute $P(Z \leq z)$.

First suppose $z \leq 0$. Then

$$
\begin{aligned}
P(X - Y \leq z) &= P(Y \geq X - z) \\
&= P(Y \geq X - z \mid Y \geq X) \times P(Y \geq X) \\
&= P(Y \geq -z) P(Y \geq X) = \frac{\lambda e^{\nu z}}{\lambda + \nu}
\end{aligned}
$$

Next suppose $z \geq 0$. Then

$$
\begin{aligned}
P(X - Y \leq z) &= P(Y - X \geq -z) \\
&= 1 - P(Y - X \leq -z) = 1 - \frac{\nu e^{-\lambda z}}{\lambda + \nu}
\end{aligned}
$$

32

The cdf for Z is then

$$F(z) = \begin{cases} \frac{\lambda e^{\nu z}}{\lambda+\nu} & \text{if } z \leq 0 \\ 1 - \frac{\nu e^{-\lambda z}}{\lambda+\nu} & \text{if } z \geq 0 \end{cases}$$

Taking the derivative of $F(z)$ wrt z yields

$$f_Z(z) = \begin{cases} \frac{\lambda \nu e^{\nu z}}{\lambda+\nu} & \text{if } z \leq 0 \\ \frac{\lambda \nu e^{-\lambda z}}{\lambda+\nu} & \text{if } z \geq 0 \end{cases}$$

33

When $\lambda = \nu$ the previous cdf and pdf reduces to

$$F(z) = \begin{cases} \frac{1}{2}e^{\lambda z} & \text{if } z \leq 0 \\ 1 - \frac{1}{2}e^{-\lambda z} & \text{if } z \geq 0 \end{cases}, \quad f(z) = \frac{\lambda}{2}e^{-\lambda|z|}.$$

which is the Laplace distribution with mean 0 and scale λ^{-1}.

112. The article **"Error Distribution in Navigation" (*J. of the Institute of Navigation*, 1971: 429–442)** suggests that the frequency distribution of positive errors (magnitudes of errors) is well approximated by an exponential distribution. Let X = the lateral position error (nautical miles), which can be either negative or positive. Suppose the pdf of X is

$$f(x) = (.1)e^{-.2|x|} \quad -\infty < x < \infty$$

34

Example Let X and Y be independent exponential rvs with rate parameters $\lambda \neq \nu$. What is the distribution of $X + Y$?

35

Example Let X and Y be independent exponential rvs with rate parameters $\lambda \neq \nu$. What is the distribution of $X + Y$?

Let $Z = X + Y \in [0, \infty)$. Then for any $z \geq 0$,

$$
\begin{aligned}
P(Z \leq z) &= \int_0^z \int_0^{z-x} \lambda \nu e^{-\lambda x} e^{-\nu y} \, \mathrm{d}y \, \mathrm{d}x \\
&= \int_0^z \lambda e^{-\lambda x} (1 - e^{-\nu(z-x)}) \, \mathrm{d}x \\
&= 1 - e^{-\lambda z} - e^{-\nu z} \int_0^z \lambda e^{-(\lambda - \nu) x} \, \mathrm{d}x \\
&= 1 - e^{-\lambda z} + e^{-\nu z} \left(\frac{\lambda e^{-(\lambda-\nu)x}}{\lambda - \nu} \bigg|_0^z \right) \\
&= 1 - e^{-\lambda z} + \frac{\lambda}{\lambda - \nu} e^{-\lambda z} - \frac{\lambda}{\lambda - \nu} e^{-\nu z} \\
&= 1 + \frac{\nu}{\lambda - \nu} e^{-\lambda z} - \frac{\lambda}{\lambda - \nu} e^{-\nu z}
\end{aligned}
$$

35

Taking derivative of $P(Z \leq z)$ wrt z we obtain

$$f(z) = \frac{d}{\mathrm{d}z} P(Z \leq z) = \frac{\lambda\nu}{\lambda - \nu}(e^{-\nu z} - e^{-\lambda z}).$$

This is known as the hypoexponential distribution.

36

Independence of discrete and continuous rv

Recall that we defined independence of random variables in terms of the cdf and not pmf or pdf.

The use of cdf allows us to mix discrete and continuous rvs.

Example Let $X \sim \mathrm{Exp}(\lambda)$ and $Y \sim \mathrm{Bernoulli}(1/2)$ be independent rvs. What is the distribution of $Z = X(2Y - 1)$?

37

Independence of discrete and continuous rv

Recall that we defined independence of random variables in terms of the cdf and not pmf or pdf.

The use of cdf allows us to mix discrete and continuous rvs.

Example Let $X \sim \mathrm{Exp}(\lambda)$ and $Y \sim \mathrm{Bernoulli}(1/2)$ be independent rvs. What is the distribution of $Z = X(2Y - 1)$?

It is obvious that (X, Y) is a bivariate rv but there is neither a joint pmf nor a joint pdf.

37

Nevertheless, we have

$$\begin{aligned} P(Z \leq z) &= P(X(2Y-1) \leq z, Y=0) + P(X(2Y-1) \leq z, Y=1) \\ &= \frac{1}{2}P(X \geq -z) + \frac{1}{2}P(X \leq z) \quad \text{(Why ?)} \end{aligned}$$

Now suppose $z \leq 0$. Then $P(X \leq z) = 0$ and hence

$$P(Z \leq z) = \frac{1}{2}P(X \geq -z) = e^{\lambda z}$$

Similarly, if $z \geq 0$ then $P(X \geq -z) = 1$ and

$$P(Z \leq z) = \frac{1}{2} + \frac{1}{2}(1 - e^{-\lambda z}) = 1 - \frac{1}{2}e^{-\lambda z}$$

This is once again the cdf of a Laplace rv.

38

Conditional probability distributions

We first recall the definition of a conditional probability for discrete random variables.

Definition (Conditional pmf)
Let $X = (X_1, X_2)$ be a discrete bivariate random variable. The conditional pmf of X_1 given X_2 is the function

$$p(x_1 \mid x_2) = \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_2 = x_2)} = \frac{p(x_1, x_2)}{p_2(x_2)}, \quad p_2(x_2) > 0.$$

where p and p_2 are the joint pmf and the marginal pmf of X_2, respectively. Note that $p(x_1, \mid x_2) = P(X_1 = x_1 \mid X_2 = x_2)$.

39

Note The above definition implies the following obvious yet important claims.

1. For any $A \subset \mathbb{R}$

$$P(X_1 \in A \mid X_2 = x_2) = \sum_{x_1} p(x_1 \mid x_2)$$

   where the sum is over all x_1 with $p(x_1, x_2) > 0$.

2. For any $B \subset \mathbb{R}^2$

$$P((X_1, X_2) \in B) = \sum_{x_1} \sum_{x_2} p(x_1 \mid x_2) p_2(x_2)$$

   where the sum is over all $(x_1, x_2) \in B$ with $p(x_1, x_2) > 0$.

40

Example Let $X_1 \sim \mathrm{Pois}(\lambda)$ and $X_2 \sim \mathrm{Pois}(\mu)$ be independent. Then $X_1 + X_2 \sim \mathrm{Pois}(\lambda + \mu)$. The conditional pmf of X_1 given $Z = X_1 + X_2$ is

$$
\begin{aligned}
p(x_1 \mid z) &= P(X_1 = x_1 \mid X_1 + X_2 = z) \\
&= \frac{P(X_1 = x_1, X_2 = z - x_2)}{P(X_1 + X_2 = z)} \\
&= \frac{(e^{-\lambda}\lambda^{x_1}/x_1!)(e^{-\mu}\mu^{z-x_1}/(z-x_1)!)}{e^{-(\lambda+\mu)}(\lambda+\mu)^z/z!} \\
&= \frac{z!}{x_1!(z-x_1)!}\left(\frac{\lambda}{\lambda+\mu}\right)^{x_1}\left(\frac{\mu}{\lambda+\mu}\right)^{z-x_1}
\end{aligned}
$$

which is the pmf of a $\mathrm{Bin}(z, \lambda/(\lambda+\mu))$ random variable.

41

When $X = (X_1, X_2)$ is a continuous bivariate random variable, the quantity $P(X_1 = x_1 \mid X_2 = x_2)$ is possibly ill-defined as $P(X_2 = x_2) = 0$, i.e., we cannot define

$$P(X_1 = x_1 \mid X_2 = x_2) = \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_2 = x_2)}$$

as both the numerator and the denominator are equal to 0.

Q. What interpretation for $P(X_1 = x_1 \mid X_2 = x_2)$?

A. Let $P(X_1 = x_1 \mid X_2 = x_2)$ be a function $f(x_1 \mid x_2)$ such that for all $B \subset \mathbb{R}^2$

$$P((X_1, X_2) \in B) = \iint\limits_B f(x_1 \mid x_2) f_2(x_2)\, \mathrm{d}x_1 \mathrm{d}x_2$$

Definition

Let $X = (X_1, X_2)$ be a continuous bivariate random variable.

The conditional pdf of X_1 given $X_2 = x_2$ is a function $f(x_1 \mid x_2)$ such that, for all $B \subset \mathbb{R}^2$ with $P((X_1, X_2) \in B) > 0$

$$P((X_1, X_2) \in B) = \iint_B f(x_1 \mid x_2) f_2(x_2) \, \mathrm{d}x_1 \mathrm{d}x_2 \quad (*)$$

where f_2 is the marginal pdf of X_2.

43

Important Using this definition with $B = [-\infty, z] \times \mathbb{R}$ we have

$$\int_{-\infty}^{z} f_1(x_1)\,\mathrm{d}x_1 = P(X_1 \leq z) = \int_{-\infty}^{z}\int_{-\infty}^{\infty} f(x_1 \mid x_2) f_2(x_2)\,\mathrm{d}x_2\mathrm{d}x_1$$

We therefore should have

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\,\mathrm{d}x_2 = \int_{-\infty}^{\infty} f(x_1 \mid x_2) f(x_2)\,\mathrm{d}x_2$$

and thus

$$f(x_1 \mid x_2) = \frac{f(x_1, x_2)}{f(x_2)}$$

as one possible definition for $f(x_1 \mid x_2)$.

This definition turns out to be valid, i.e., $f(x_1 \mid x_2)$ as defined does satisfy condition $(*)$.

44

Summary For a continuous bivariate rv $X = (X_1, X_2)$, $f(x_1 \mid x_2)$ is different from $P(X_1 = x_1 \mid X_2 = x_2)$ as $P(X_1 = x_1 \mid X_2 = x_2)$ is ill-defined

Rather, $f(x_1 \mid x_2)$ is any function satisfying

$$\iint_B f(x_1, x_2)\,\mathrm{d}x_1\mathrm{d}x_2 = \iint_B f(x_1 \mid x_2) f_2(x_2)\,\mathrm{d}x_1\mathrm{d}x_2$$

for all $B \subset \mathbb{R}^2$ with $P((X_1, X_2) \in B) > 0$.

Canonical choice is $f(x_1 \mid x_2) = f(x_1, x_2)/f_2(x_2)$ for $f_2(x_2) > 0$.

45

EXAMPLE 5.8 A soft-drink machine has a random amount Y_2 in supply at the beginning of a given day and dispenses a random amount Y_1 during the day (with measurements in gallons). It is not resupplied during the day, and hence $Y_1 \leq Y_2$. It has been observed that Y_1 and Y_2 have a joint density given by

$$f(y_1, y_2) = \begin{cases} 1/2, & 0 \leq y_1 \leq y_2 \leq 2, \\ 0 & \text{elsewhere.} \end{cases}$$

That is, the points (y_1, y_2) are uniformly distributed over the triangle with the given boundaries. Find the conditional density of Y_1 given $Y_2 = y_2$. Evaluate the probability that less than 1/2 gallon will be sold, given that the machine contains 1.5 gallons at the start of the day.

We are interested in the conditional pdf of Y_1 given Y_2. We first need to find $f_2(y_2)$, i.e.,

$$f_2(y_2) = \int_{-\infty}^{\infty} f(z, y_2)\,\mathrm{d}z = \int_0^{y_2} \frac{1}{2}\mathrm{d}z = \frac{y_2}{2}$$

The conditional pdf of Y_1 given Y_2 is then

$$f(y_1 \mid y_2) = \frac{f(y_1, y_2)}{f_2(y_2)} = \frac{1/2}{y_2/2} = \frac{1}{y_2}.$$

46

Expected value for functions of multivariate rvs

We now formally define expected value for functions of multiple rvs. Note that we have used these ideas previously when considering the expectation of a sum of independent rvs.

Definition
Let $X = (X_1, X_2, \ldots, X_m)$ be a discrete multivariate r.v. with joint pmf $p(x_1, x_2, \ldots, x_m)$. Then for any function $g \colon \mathbb{R}^m \mapsto \mathbb{R}$,

$$\mathbb{E}[g(X_1, \ldots, X_m)] = \sum_{(x_1, \ldots, x_m)} g(x_1, \ldots, x_m) p(x_1, \ldots, x_m)$$

where the sum is over all $(x_1, \ldots, x_m)$ with $p(x_1, \ldots, x_m) > 0$.

47

Definition
Let $X = (X_1, X_2, \ldots, X_m)$ be a continuous multivariate rv with joint pdf $f(x_1, x_2, \ldots, x_m)$. Then for any function $g \colon \mathbb{R}^m \mapsto \mathbb{R}$,

$$\mathbb{E}[g(X_1, \ldots, X_m)] = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} g(x_1, \ldots, x_m) f(x_1, \ldots, x_m) \, \mathrm{d}x_1 \ldots \mathrm{d}x_m$$

Q. How do you define $\mathbb{E}[g(X_1, \ldots, X_m)]$ when the $(X_1, \ldots, X_m)$ is a mix of discrete and continuous rvs ?

48

In any case, we have the following important, albeit simple, results (mantra: expectation is a linear operation).

Let $X = (X_1, X_2, \ldots, X_m)$, then

1. $\mathbb{E}[c] = c$ for all constant c.
2. $\mathbb{E}[cg(X_1, \ldots, X_m)] = c\mathbb{E}[g(X_1, \ldots, X_m)]$ for all constant c and all functions g.
3. For any functions $g_1, g_2, \ldots, g_K$,

   $$\mathbb{E}[g_1(X_1, \ldots, X_m) + \cdots + g_K(X_1, \ldots, X_m)] = \sum_k \mathbb{E}[g_k(X_1, \ldots, X_m)]$$

Finally when X and Y are independent (possibly multivariate) random variables then, for any functions g and h

$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\mathbb{E}[h(Y)]$$

49

Example Let X and Y be independent $\mathrm{Unif}(0,1)$ random variables. What is $\mathbb{E}[|X-Y|]$?

$$
\begin{aligned}
\mathbb{E}[|X-Y|] &= \int_0^1 \int_0^1 |x-y| \,\mathrm{d}x\mathrm{d}y \\
&= \int_0^1 \Big(\int_0^y (y-x)\,\mathrm{d}x + \int_y^1 (x-y)\,\mathrm{d}x \Big) \mathrm{d}y \\
&= \int_0^1 \Big(y^2/2 + (1-y^2)/2 - y(1-y) \Big)\,\mathrm{d}y = 1/3
\end{aligned}
$$

50