Introduction to Probability

Lecture 6: Marginals and Joint Distributions

Mateja Jamnik, Thomas Sauerwald

University of Cambridge, Department of Computer Science and Technology
email: {mateja.jamnik,thomas.sauerwald}@cl.cam.ac.uk

Motivation

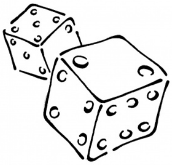
Experiments often involve several random variables, and some of them may influence each other.

To this end, we will introduce:

- Joint/Marginal distribution of two (or more) variables
- Independence of two (or more) variables
- Covariance of two variables

For simplicity, we will mainly focus on discrete random variables.

Warm-Up Exercise



Example

Let $X_1, X_2 \in \{1, 2, \ldots, 6\}$ be two independent rolls of an unbiased die. Let $S := X_1 + X_2$ and $M := \max\{X_1, X_2\}$. List the elements of the event $\{S = 7, M \leq 5\}$ and deduce the probability.

Answer

Joint Probability

Joint Probability Mass Function

The joint probability mass function of two discrete random variables X and Y is the function $p : \mathbb{R}^2 \to [0, 1]$, defined by:

$$p_{X,Y}(a, b) = \mathbf{P}[\, X = a, Y = b\,] \qquad \text{for } -\infty < a, b < \infty.$$

Joint Distribution Function

The joint distribution function of two (discrete or continuous) random variables X and Y is the function $F : \mathbb{R}^2 \to [0, 1]$, defined by:

$$F_{X,Y}(a, b) = \mathbf{P}[\, X \leq a, Y \leq b\,] \qquad \text{for } -\infty < a, b < \infty.$$

Marginal Distribution

Given a joint distribution $F_{X,Y}$ of two random variables X, Y, one obtains the marginal distribution of X for any a as follows:

$$F_X(a) = \mathbf{P}[\, X \leq a\,] = \lim_{b \to \infty} F_{X,Y}(a, b).$$

Joint Distribution contains (much) more information than the two marginals!

Discrete Example 1

Example

Let $X_1, X_2 \in \{1, 2, \ldots, 6\}$ be independent rolls of an unbiased die. Let $S := X_1 + X_2$ and $M := \max\{X_1, X_2\}$. Compute the joint probability mass function p of S and M and the marginal distributions of S and M.

Answer

	b						
a	1	2	3	4	5	6	$p_S(a)$
2	1/36	0	0	0	0	0	1/36
3	0	2/36	0	0	0	0	2/36
4	0	1/36	2/36	0	0	0	3/36
5	0	0	2/36	2/36	0	0	4/36
6	0	0	1/36	2/36	2/36	0	5/36
7	0	0	0	2/36	2/36	2/36	6/36
8	0	0	0	1/36	2/36	2/36	5/36
9	0	0	0	0	2/36	2/36	4/36
10	0	0	0	0	1/36	2/36	3/36
11	0	0	0	0	0	2/36	2/36
12	0	0	0	0	0	1/36	1/36
$p_M(b)$	1/36	3/36	5/36	7/36	9/36	11/36	1

Discrete Example 2

Example

Suppose an urn contains balls numbered $1, 2, \ldots, N$. We draw $1 \leq n \leq N$ balls uniformly and without replacement from the urn. Let $X_i \in \{1, 2, \ldots, N\}$ be the number of the ball drawn in the i-th step. What is the marginal distribution of X_i?

Answer

We first compute the joint distribution. For distinct $a_1, a_2, \ldots, a_n$,

Fix i and consider the marginal distribution of X_i:

Joint Distributions of Continuous Variables

Definition

Random variables X and Y have a joint continuous distribution if for some function $f : \mathbb{R}^2 \to \mathbb{R}$ and for all numbers $a_1 \leq b_1$ and $a_2 \leq b_2$,

$$\mathbf{P}[\, a_1 \leq X \leq b_1, a_2 \leq Y \leq b_2 \,] = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y)\, dx\, dy.$$

The function f has to satisfy $f(x, y) \geq 0$ for all x and y, and $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dxdy = 1$. We call f the joint probability density.

As in one-dimensional case we switch from F to f by differentiating (or integrating):

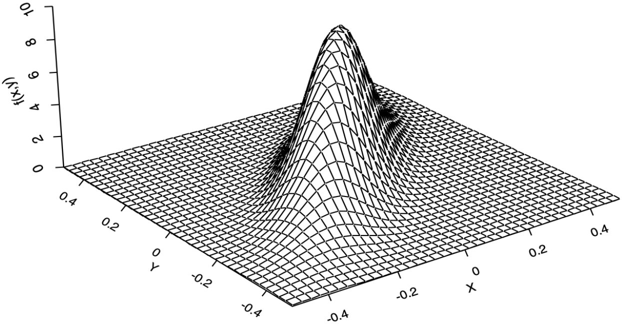
$$F(a, b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f(x, y) dx\, dy \qquad \text{and} \qquad f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y)$$

Example of a Joint Distribution of Continuous Random Variables

- Consider the density:

$$f(x, y) = \frac{30}{\pi} \cdot e^{-50x^2 - 50y^2 + 80xy},$$

  where $-\infty < x, y < \infty$.
- This is an example of a so-called bivariate normal probability density function.



Source: Modern Introduction to Statistics

Dealing with Continuous Variables

Example (1/2)

Suppose that the joint probability density of X and Y is given by

$$f(x, y) = \begin{cases} 2e^{-x}e^{-2y} & \text{for } 0 < x < \infty, 0 < y < \infty, \\ 0 & \text{otherwise.} \end{cases}$$

Compute (i) $\mathbf{P}[X > 1, Y < 1]$ and (ii) $\mathbf{P}[X < Y]$.

Answer

(i) We first compute:

$$\mathbf{P}[X > 1, Y < 1] = \int_0^1 \int_1^\infty 2e^{-x}e^{-2y}\,dxdy$$

Dealing with Continuous Variables (cont.)

Example (2/2)

Suppose that the joint probability density of X and Y is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & \text{for } 0 < x < \infty, 0 < y < \infty, \\ 0 & \text{otherwise.} \end{cases}$$

Compute (i) $\mathbf{P}[X > 1, Y < 1]$ and (ii) $\mathbf{P}[X < Y]$.

Answer

(ii) We have:

$$\mathbf{P}[X < Y] = \int_0^\infty \int_0^y 2e^{-x}e^{-2y} dxdy$$