Mathematical Statistics I

Chapter 3: Joint Distributions

Jesse Wheeler

Contents

1 Introduction 1

2 Discrete Random Variables 2

1 Introduction

Introduction

- This material is based on the textbook by Rice (2007, Chapter 3).
- Our goal is to better understand the joint probability structure of more than one random variable, defined on the same sample space.
- One reason that studying joint probabilities is an important topic is that it enables us to use what we know about one variable to study another.

Joint cdf

- Just like the univariate case, the joint behavior of two random variables, X and Y, is determined by the cumulative distribution function

$$F(x, y) = P(X \leq x, Y \leq y).$$

- This is true for both discrete and continuous random variables.
- The any set $A \subset \mathbb{R}^2$, the joint cdf can give $P\big((X, Y) \in A\big)$.
- For example, let A be the rectangle defined by $x_1 < X < x_2$, and $y_1 < Y < y_2$. (It helps to draw a picture...)
- $F(x_2, y_2)$ gives $P(X < x_2, Y < y_2)$, an area that is too big, so we subtract off pieces
  - $F(x_2, y_1) = P(X < x_2, Y < y_1)$ (we already have the area $X < x_2$, but now subtract away the area $Y < y_1$).
  - $F(x_1, y_2) = P(X < x_1, Y < y_2)$ (Now subtracting the area $X < x_1$)
  - We have "double subtracted" the area $\{X < x_1, Y < y_1\}$, so we add it back.

$$P\big((X, Y) \in A\big) = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1).$$

	y			
x	0	1	2	3
0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0
1	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$

Table 1: Frequency table for X and Y, flipping a fair coin three times.

- The definition also applies to more than two random variables.
- Let $X_1, \ldots, X_n$ be jointly distributed random variables defined on the same sample space. Then

$$F(x_1, x_2, \ldots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n).$$

- Like the univariate case, we can also define the pmf and pdf of jointly distributed random variables as well.

2 Discrete Random Variables

Discrete Random Variables

Definition: Joint pmf
Let X and Y be discrete random variables define on the same sample space, and take on values $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$, respectively. The *joint pmf* (or joint frequency function), is

$$p(x_i, y_j) = P(X = x_i, Y = y_j).$$

- For discrete RVs, it's often useful to describe the joint pmf as a frequency table.
- Suppose a fair coin is tossed 3 times. Let X denote the number of heads on the first toss, and Y the total number of heads.
- The sample space is

$$\Omega = \{hhh, hht, hth, thh, htt, tht, tth, ttt\}.$$

- The joint pmf can be expressed as the frequency table below (Table 1).

Acknowledgments

- Compiled on August 17, 2025 using R version 4.5.1.

- We acknowledge students and instructors for previous versions of this course / slides.

References

Rice JA (2007). *Mathematical statistics and data analysis*, volume 371. 3 edition. Thomson/Brooks/Cole Belmont, CA. 1