Probability and Statistics I

20. Jointly Distributed Random Variables

5.1 Jointly Distributed Random Variables

Discrete Random Variables

Review: Conditional Probabilities

Recall that the conditional probability of the event A given B is:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

The two events are independent if

$$P(A|B) = P(A), \quad P(B|A) = P(A), \text{ or } P(A \cap B) = P(A)P(B).$$

Joint Probability Mass Function

Let X and Y be two discrete random variables defined on the sample space $\mathcal{S}$. The joint probability mass function $p(x, y)$ is defined for each pair of numbers (x, y) by

$$p(x, y) = P(X = x \text{ and } Y = y) = P(X = x, Y = y).$$

Examples

Let $\mathcal{S}$ be the set of days in 2024.

$$X = \begin{cases} 0 & \text{it doesn't snow} \\ 1 & \text{it snows} \end{cases}$$

and

$$Y = \text{Month of year } (1{=}\text{Jan, } 2{=}\text{Feb,} \ldots \text{,} 12{=}\text{Dec})$$

Examples

Let $\mathcal{S}$ be the set of days in cards in a standard deck.

$$X = \text{value } (1{=}\text{A}, 2{=}2, 3{=}3, \ldots, 11{=}\text{J}, 12{=}\text{Q}, 13{=}\text{K})$$

and

$$Y = \text{suit } (1{=}\text{Clubs, } 2{=}\text{Diamonds, } 3{=}\text{Hearts, } 4{=}\text{Spades})$$

Examples

Let $\mathcal{S}$ be all possible outcomes of rolling two die (e.g., (1,1) or (3,2)).

$$X = \text{sum of the two values}$$

and

$$Y = \text{difference of the two values}$$

Marginal Probability Mass Function

The marginal probability mass functions of X and Y, denoted $p_X(x)$ and $p_Y(y)$, are given by

$$p_X(x) = \sum_y p(x,y) \text{ and } p_Y(y) = \sum_x p(x,y).$$

Note that this should really say:

$$p_X(x) = \sum_{y \in \mathcal{Y}} p(x,y) \text{ and } p_Y(y) = \sum_{x \in \mathcal{X}} p(x,y)$$

where $\mathcal{X}$ and $\mathcal{Y}$ represent the range of X and Y, respectively.

Independence

Two discrete random variables X and Y are said to be independent if

$$p(x, y) = p(x)p(y)$$

for every $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.

If X and Y are not independent, then they are dependent.

Example 20.1: Berkeley Admissions Data

The following data come from a study conducted at UC Berkeley in the 1970's to assess claims of bias in graduate admissions. The data contain records from 4256 applicants to the top six different degrees (Majors). The names of the degrees were kept confidential.

	Applicants by Gender	
Major	Men	Women
A	825	108
B	560	25
C	325	593
D	417	375
E	191	393
F	373	341

Example 20.1 ctd

Define the random variables

$$X = \begin{cases} 0 & \text{an applicant is male} \\ 1 & \text{an applicant is female} \end{cases}$$

$$Y = \begin{cases} 1 & \text{an applicant applies to Major A} \\ \vdots & \\ 6 & \text{an applicant applies to Major F} \end{cases}$$

a) Estimate the joint pmf of X and Y.
b) Compute the marginal distributions of X and Y.
c) Are X and Y independent? What would this mean?

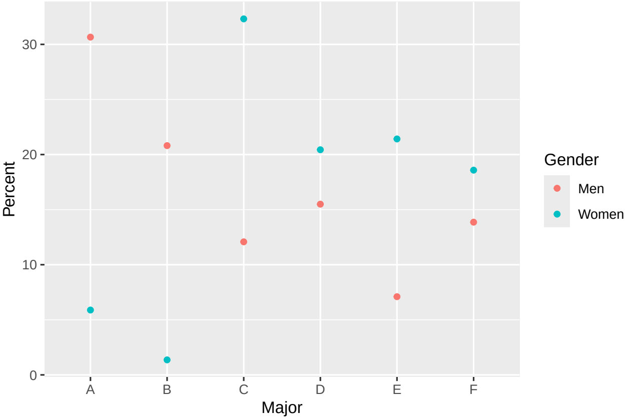
Example 20.1 ctd

	Major	Men	Women
1	A	0.182	0.024
2	B	0.124	0.006
3	C	0.072	0.131
4	D	0.092	0.083
5	E	0.042	0.087
6	F	0.082	0.075

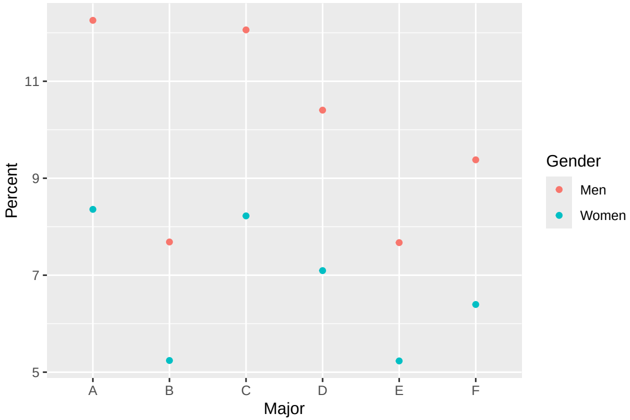
Example 20.1 ctd

Major	Men	Women	$p_Y(y)$
A	0.182	0.024	.21
B	0.124	0.006	.13
C	0.072	0.131	.20
D	0.092	0.083	.17
E	0.042	0.087	.13
F	0.082	0.075	.16
$p_X(x)$	.59	.41	

Example 20.1 ctd: $p(x, y)$



Example 20.1 ctd: $p(x)p(y)$



5.1 Jointly Distributed Random Variables

Continuous Random Variables

Joint Probability Density Function

Let X and Y be two continuous random variables. Then $f(x, y)$ is the joint probability density function for X and Y if

$$P[(X, Y) \in A] = \int \int_A f(x, y) \; dx \; dy$$

for any set $A \in \mathbb{R}^2$. In particular, if A is a rectangle with limits $a \leq x \leq b$ and $c \leq y \leq d$ then

$$P[(X, Y) \in A] = P(a \leq x \leq b, c \leq y \leq d) = \int_a^b \int_c^d f(x, y) \; dx \; dy.$$

Joint Probability Density Function

The function $f(x, y)$ is a valid joint pdf if

$$f(x, y) \geq 0 \text{ for all } (x, y) \in \mathbb{R}^2$$

and

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

Marginal Probability Density Function

The marginal probability density functions of X and Y, denoted $f_X(x)$ and $f_Y(y)$, are given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\ dy$$

and

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\ dx$$

for $x, y \in (-\infty, \infty)$.

Independence

Two continuous random variables X and Y are said to be independent if

$$f(x, y) = f(x)f(y)$$

for every $(x, y) \in \mathbb{R}^2$.

If X and Y are not independent then they are dependent.

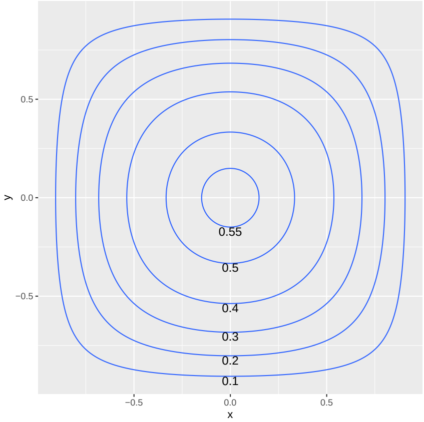
Example 20.2

Suppose that X and Y have the joint pdf

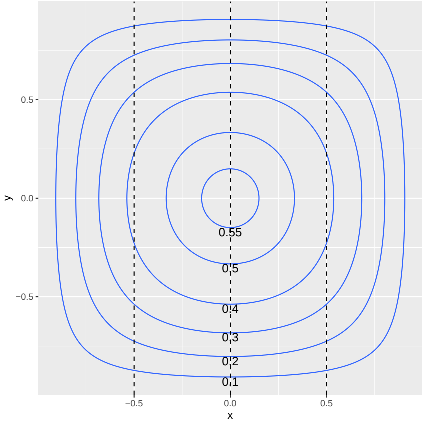
$$f(x,y) = \begin{cases} c(1-x^2)(1-y^2) & -1 < x < 1 \text{ and } -1 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

a) Find the value of c so that $f(x,y)$ is a proper pdf.
b) Find the marginal distribution of X and Y.
c) Are X and Y independent?

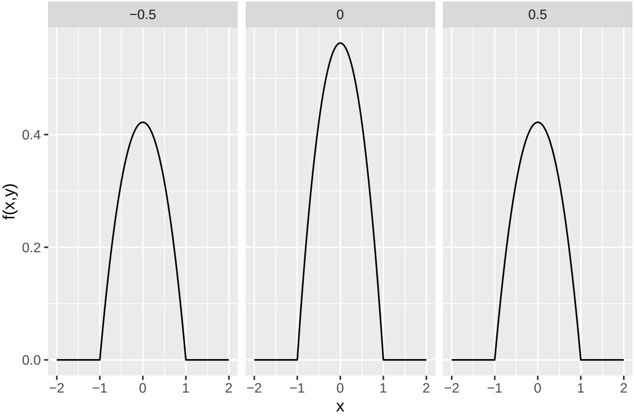
Example 20.2 ctd



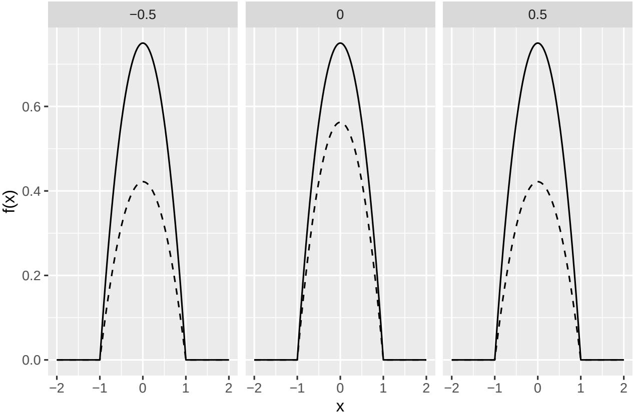
Example 20.2 ctd



Example 20.2 ctd



Example 20.2 ctd



Questions?

Exercise 20.1

Suppose that

$$f(x, y) = c \min(x, y), 0 < x, y < 1.$$

a) Find the value c so that $f(x, y)$ is a valid joint pdf.
b) Compute $P(X > .5, Y > .5)$.
c) Derive the marginal density of X and Y.
d) Are X and Y independent?