Chapter 4
Jointly Distributed Random Variables

Discrete Multivariate Distributions

Joint Probability Function

Quite often there will be 2 or more random variables (X, Y, Z, *etc*) defined for the same random experiment.

Example 1: Portfolio Manager
X = Fund 1 has a return greater than target over a quarter.
Y = Fund 2 has a return greater than target over a quarter.
The manager of the funds is interested in the behavior of X and Y:

Example 2: A bridge hand is selected from a deck of 52 cards.
X = the number of spades in the hand (13 cards).
Y = the number of hearts in the hand (13 cards).

In these examples, we will define: $p(x,y) = P[X = x, Y = y]$. The function $p(x,y)$ is called the *joint probability function* of X and Y.

Joint Probability Function: Bridge Hand

Note:

The possible values of X are 0, 1, 2, …, 13

The possible values of Y are also 0, 1, 2, …, 13 and $X + Y \leq 13$.

$$p(x, y) = P[X = x, Y = y]$$

The number of ways of choosing the y **hearts** for the hand

The number of ways of choosing the x **spades** for the hand

$$= \frac{\binom{13}{x}\binom{13}{y}\binom{26}{13-x-y}}{\binom{52}{13}}$$

The number of ways of completing the hand with **diamonds** and **clubs.**

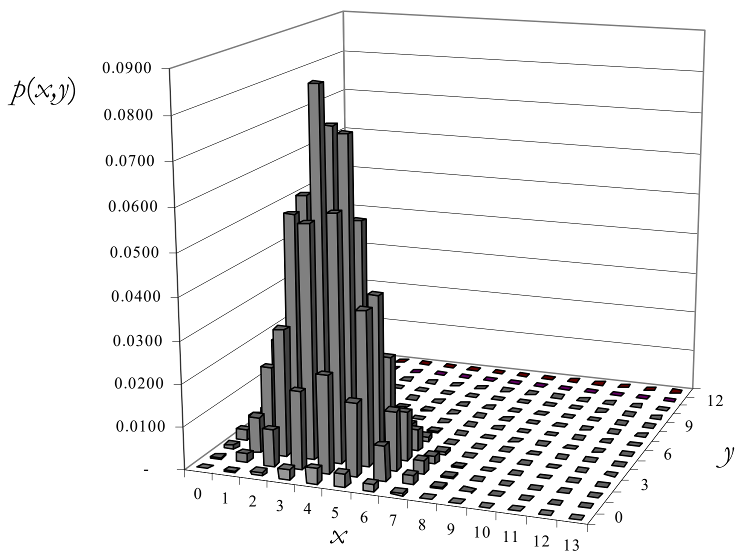
The total number of ways of choosing the 13 cards for the hand

Joint Probability Function: Bridge Hand

Table: $p(x,y)$ $= \dfrac{\binom{13}{x}\binom{13}{y}\binom{26}{13-x-y}}{\binom{52}{13}}$

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.0000	0.0002	0.0009	0.0024	0.0035	0.0032	0.0018	0.0006	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
1	0.0002	0.0021	0.0085	0.0183	0.0229	0.0173	0.0081	0.0023	0.0004	0.0000	0.0000	0.0000	0.0000	-
2	0.0009	0.0085	0.0299	0.0549	0.0578	0.0364	0.0139	0.0032	0.0004	0.0000	0.0000	0.0000	-	-
3	0.0024	0.0183	0.0549	0.0847	0.0741	0.0381	0.0116	0.0020	0.0002	0.0000	0.0000	-	-	-
4	0.0035	0.0229	0.0578	0.0741	0.0530	0.0217	0.0050	0.0006	0.0000	0.0000	-	-	-	-
5	0.0032	0.0173	0.0364	0.0381	0.0217	0.0068	0.0011	0.0001	0.0000	-	-	-	-	-
6	0.0018	0.0081	0.0139	0.0116	0.0050	0.0011	0.0001	0.0000	-	-	-	-	-	-
7	0.0006	0.0023	0.0032	0.0020	0.0006	0.0001	0.0000	-	-	-	-	-	-	-
8	0.0001	0.0004	0.0004	0.0002	0.0000	0.0000	-	-	-	-	-	-	-	-
9	0.0000	0.0000	0.0000	0.0000	0.0000	-	-	-	-	-	-	-	-	-
10	0.0000	0.0000	0.0000	0.0000	-	-	-	-	-	-	-	-	-	-
11	0.0000	0.0000	0.0000	-	-	-	-	-	-	-	-	-	-	-
12	0.0000	0.0000	-	-	-	-	-	-	-	-	-	-	-	-
13	0.0000	-	-	-	-	-	-	-	-	-	-	-	-	-

Joint Probability Function: Bar Graph *p(x,y)*



Joint Probability Function: Properties

General properties of the joint probability function

$p(x,y) = P[X = x,\ Y = y]$

1. $0 \le p(x, y) \le 1$

2. $\sum_x \sum_y p(x, y) = 1$

3. $P\left[(X, Y) \in A\right] = \sum\sum_{(x,y)\in A} p(x, y)$

Example: A die is rolled $n = 5$ times
X = the number of times a "**six**" appears.
Y = the number of times a "**five**" appears.

What is the probability that we roll more **sixes** than **fives**, that is, what is $P[X > Y]$?

Joint Probability Function: Properties

Table: $p(x,y) \;=\; \dfrac{5!}{x!\,y!\,(5-x-y)!}\left(\dfrac{1}{6}\right)^{x}\left(\dfrac{1}{6}\right)^{y}\left(\dfrac{4}{6}\right)^{5-x-y}$

	0	1	2	3	4	5
0	0.1317	0.1646	0.0823	0.0206	0.0026	0.0001
1	0.1646	0.1646	0.0617	0.0103	0.0006	0
2	0.0823	0.0617	0.0154	0.0013	0	0
3	0.0206	0.0103	0.0013	0	0	0
4	0.0026	0.0006	0	0	0	0
5	0.0001	0	0	0	0	0

$$P\left[X > Y\right] = \sum\sum_{x>y} p\left(x, y\right) = 0.3441$$

Marginal Probability

<u>Definition</u>**:** Marginal Probability

Let X and Y denote two discrete RV with joint probability function

$p(x,y) = P[X = x,\ Y = y]$

Then

$p_X(x) = P[X = x]$ is called the *marginal probability* function of X.

$p_Y(y) = P[Y = y]$ is called the *marginal probability* function of Y.

<u>Note</u>**:** Let $y_1, y_2, y_3, \ldots$ denote the possible values of Y

$$p_X\left(x\right) = P\left[X = x\right] = P\left[\{X = x, Y = y_1\} \cup \{X = x, Y = y_2\} \cup \ldots\right]$$

$$= P\left[X = x, Y = y_1\right] + P\left[X = x, Y = y_2\right] + \ldots$$

$$= p\left(x, y_1\right) + p\left(x, y_2\right) + \ldots = \sum_j p\left(x, y_j\right) = \sum_y p\left(x, y\right)$$

Marginal Probability

Thus, the marginal probability function of X, $p_X(x)$ is obtained from the joint probability function of X and Y by summing p(x,y) over the possible values of Y.

Similarly,

$$p_Y(y) = P[Y = y] = P\left[\{X = x_1, Y = y\} \cup \{X = x_2, Y = y\} \cup \dots\right]$$

$$= P[X = x_1, Y = y] + P[X = x_2, Y = y] + \dots$$

$$= p(x_1, y) + p(x_2, y) + \dots = \sum_i p(x_i, y) = \sum_x p(x, y)$$

Marginal Probability

Example: A die is rolled n = 5 times

What is the probability that we roll more **sixes** than **fives**, that is, what is $P[X > Y]$?

X = the number of times a "**six**" appears.

Y = the number of times a "**five**" appears.

	0	1	2	3	4	5	$p_X(x)$
0	0.1317	0.1646	0.0823	0.0206	0.0026	0.0001	0.4019
1	0.1646	0.1646	0.0617	0.0103	0.0006	0	0.4019
2	0.0823	0.0617	0.0154	0.0013	0	0	0.1608
3	0.0206	0.0103	0.0013	0	0	0	0.0322
4	0.0026	0.0006	0	0	0	0	0.0032
5	0.0001	0	0	0	0	0	0.0001
$p_Y(y)$	0.4019	0.4019	0.1608	0.0322	0.0032	0.0001	

$$P(X > Y) = P[X > 0, Y = 0] + P[X > 1, Y \leq 1] + P[X > 2, Y \leq 2]$$
$$= 0.2702 + 0.0727 + 0.0013 = 0.3442$$

Conditional Probability

Definition: Conditional Probability

Let X and Y denote two discrete RV with joint probability function

$$p(x, y) = P[X = x, Y = y]$$

Then,

$p_{X|Y}(x|y) = P[X = x| Y = y]$ is called the *conditional probability function of* X *given* $Y = y$

$$p_{X|Y}(x|y) = P[X = x|Y = y]$$
$$= \frac{P[X = x, Y = y]}{P[Y = y]} = \frac{p(x, y)}{p_Y(y)}$$

Similarly, $p_{Y|X}(y|x)$ is the *conditional probability function of* Y *given* $X = x$

$$p_{Y|X}(y|x) = P[Y = y|X = x]$$
$$= \frac{P[X = x, Y = y]}{P[X = x]} = \frac{p(x, y)}{p_X(x)}$$

Conditional Probability

Notes:

- Marginal distributions describe how one variable behaves ignoring the other variable.

- Conditional distributions describe how one variable behaves when the other variable is held fixed.

Conditional Probability

Example: Probability of rolling more **sixes** than **fives**, when a die is rolled $n = 5$ times

X = the number of times a "**six**" appears.

Y = the number of times a "**five**" appears.

$x \backslash y$	0	1	2	3	4	5	$p_X(x)$
0	0.1317	0.1646	0.0823	0.0206	0.0026	0.0001	0.4019
1	0.1646	0.1646	0.0617	0.0103	0.0006	0	0.4019
2	0.0823	0.0617	0.0154	0.0013	0	0	0.1608
3	0.0206	0.0103	0.0013	0	0	0	0.0322
4	0.0026	0.0006	0	0	0	0	0.0032
5	0.0001	0	0	0	0	0	0.0001
$p_Y(y)$	0.4019	0.4019	0.1608	0.0322	0.0032	0.0001	

$$p_{Y|X}(y=2|x=0) = P[Y = 2|X = 0] = p(y = 2, x = 0)/p_Y(x = 0) =$$
$$= .0823/.4019 = .20478$$

Conditional Probability

The conditional distribution of Y given $X = x$:

$$p_{Y|X}(y|x) = P[Y = y|\ X = x]$$

$x \backslash y$	0	1	2	3	4	5
0	0.3277	0.4096	0.2048	0.0512	0.0064	0.0003
1	0.4096	0.4096	0.1536	0.0256	0.0016	0.0000
2	0.5120	0.3840	0.0960	0.0080	0.0000	0.0000
3	0.6400	0.3200	0.0400	0.0000	0.0000	0.0000
4	0.8000	0.2000	0.0000	0.0000	0.0000	0.0000
5	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Marginal & Conditional Probability - Example

Example: A Bernoulli trial (**S**: p, **F**: $q = 1 - p$) is repeated until two successes have occurred.

X = trial on which the first success occurs

and

Y = trial on which the 2^{nd} success occurs.

Find the joint probability function of X, Y.

Find the marginal probability function of X and Y.

Find the conditional probability functions of Y given $X = x$ and X given $Y = y$,

Marginal & Conditional Probability - Example

Solution

A typical outcome would be:

$$\underbrace{\text{FFF}\ldots\text{F}}_{x-1}\overset{x}{\text{S}}\underbrace{\text{FFF}\ldots\text{F}}_{y-x-1}\overset{y}{\text{S}}$$

$$p(x,y) = P[X = x, Y = y]$$

$$= q^{x-1} p q^{y-x-1} p = q^{y-2} p^2 \quad \text{if } y > x$$

$$p(x,y) = \begin{cases} q^{y-2} p^2 & \text{if } y > x \\ 0 & \text{otherwise} \end{cases}$$

Marginal & Conditional Probability - Example

$p(x,y)$ - Table

		y							
		1	2	3	4	5	6	7	8
x	1	0	p^2	p^2q	p^2q^2	p^2q^3	p^2q^4	p^2q^5	p^2q^6
	2	0	0	p^2q	p^2q^2	p^2q^3	p^2q^4	p^2q^5	p^2q^6
	3	0	0	0	p^2q^2	p^2q^3	p^2q^4	p^2q^5	p^2q^6
	4	0	0	0	0	p^2q^3	p^2q^4	p^2q^5	p^2q^6
	5	0	0	0	0	0	p^2q^4	p^2q^5	p^2q^6
	6	0	0	0	0	0	0	p^2q^5	p^2q^6
	7	0	0	0	0	0	0	0	p^2q^6
	8	0	0	0	0	0	0	0	0

Marginal & Conditional Probability - Example

The marginal distribution of X

$$
\begin{aligned}
p_X(x) = P[X = x] &= \sum_y p(x,y) \\
&= \sum_{y=x+1}^{\infty} p^2 q^{y-2} \\
&= p^2q^{x-1} + p^2q^{x} + p^2q^{x+1} + p^2q^{x+2} + \ldots \\
&= p^2q^{x-1}\left(1 + q + q^2 + q^3 + \ldots\right) \\
&= p^2q^{x-1}\left(\frac{1}{1-q}\right) = pq^{x-1}
\end{aligned}
$$

This is the *geometric distribution.*

Marginal & Conditional Probability - Example

The marginal distribution of Y

$$p_Y(y) = P[Y = y] = \sum_x p(x, y)$$

$$= \begin{cases} (y-1)\, p^2 q^{y-2} & y = 2, 3, 4, \ldots \\ 0 & \text{otherwise} \end{cases}$$

This is the *negative binomial* distribution with $k = 2$. In count data models, the negative binomial ("negbin") is used when the data shows *overdispersion* –i.e., the volatility is greater than the mean, the usual assumption when using the Poisson distribution.

Marginal & Conditional Probability - Example

The conditional distribution of X given $Y = y$

$$p_{X|Y}(x|y) = P[X = x | Y = y]$$

$$= \frac{P[X = x, Y = y]}{P[Y = y]} = \frac{p(x, y)}{p_Y(y)}$$

$$= \frac{p^2 q^{y-2}}{pq^{x-1}}$$

$$= pq^{y-x-1} \quad \text{for } y = x+1, x+2, x+3 \ldots$$

This is the *geometric distribution* with time starting at x. We think of the geometric distribution as the discrete analogue of exponential distribution. It also has no memory.

Marginal & Conditional Probability - Example

The conditional distribution of Y given $X = x$:

$$p_{Y|X}\left(y|x\right) = P\left[Y = y|X = x\right]$$

$$= \frac{P\left[X = x, Y = y\right]}{P\left[X = x\right]} = \frac{p\left(x, y\right)}{p_X\left(x\right)}$$

$$= \frac{p^2 q^{y-2}}{\left(y-1\right) p^2 q^{y-2}} = \frac{1}{\left(y-1\right)} \qquad \text{for } x = 1, 2, 3, \ldots, \left(y-1\right)$$

This is the *uniform distribution* on the values 1, 2, …, $(y – 1)$