JOINT DISCRETE DISTRIBUTIONS

The joint pmf: The *joint probability mass function* (joint pmf) for two discrete random variables Y_1 and Y_2 is the function

$$p(y_1, y_2) \equiv p_{Y_1,Y_2}(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2) \qquad (y_1, y_2) \in \mathbb{Z} \times \mathbb{Z}.$$

We can tabulate $p_{Y_1,Y_2}(y_1, y_2)$ as follows: suppose Y_1 takes values on $\{0, 1, 2, \ldots, 8\}$ and Y_2 takes values on $\{0, 1, 2, \ldots, 6\}$, then we may write

	6	$p(0,6)$	$p(1,6)$	$p(2,6)$	$p(3,6)$	$p(4,6)$	$p(5,6)$	$p(6,6)$	$p(7,6)$	$p(8,6)$
	5	$p(0,5)$	$p(1,5)$	$p(2,5)$	$p(3,5)$	$p(4,5)$	$p(5,5)$	$p(6,5)$	$p(7,5)$	$p(8,5)$
	4	$p(0,4)$	$p(1,4)$	$p(2,4)$	$p(3,4)$	$p(4,4)$	$p(5,4)$	$p(6,4)$	$p(7,4)$	$p(8,4)$
Y_2	3	$p(0,3)$	$p(1,3)$	$p(2,3)$	$p(3,3)$	$p(4,3)$	$p(5,3)$	$p(6,3)$	$p(7,3)$	$p(8,3)$
	2	$p(0,2)$	$p(1,2)$	$p(2,2)$	$p(3,2)$	$p(4,2)$	$p(5,2)$	$p(6,2)$	$p(7,2)$	$p(8,2)$
	1	$p(0,1)$	$p(1,1)$	$p(2,1)$	$p(3,1)$	$p(4,1)$	$p(5,1)$	$p(6,1)$	$p(7,1)$	$p(8,1)$
	0	$p(0,0)$	$p(1,0)$	$p(2,0)$	$p(3,0)$	$p(4,0)$	$p(5,0)$	$p(6,0)$	$p(7,0)$	$p(8,0)$
		0	1	2	3	4	5	6	7	8
						Y_1				

For example, suppose that

$$p(y_1, y_2) = c \exp\{1 + (y_1 - 4)(y_2 - 2)/10\} \qquad 0 \le y_1 \le 8, 0 \le y_2 \le 6$$

and $p(y_1, y_2) = 0$ otherwise, for some constant $c > 0$. The constant c ensures that the joint pmf sums to 1, that is

$$c \sum_{y_1=0}^{8} \sum_{y_2=0}^{6} \exp\{1 + (y_1 - 4)(y_2 - 2)/10\} = 1$$

In this case we can compute $c = 0.012$, and compute the probabilities as follows:

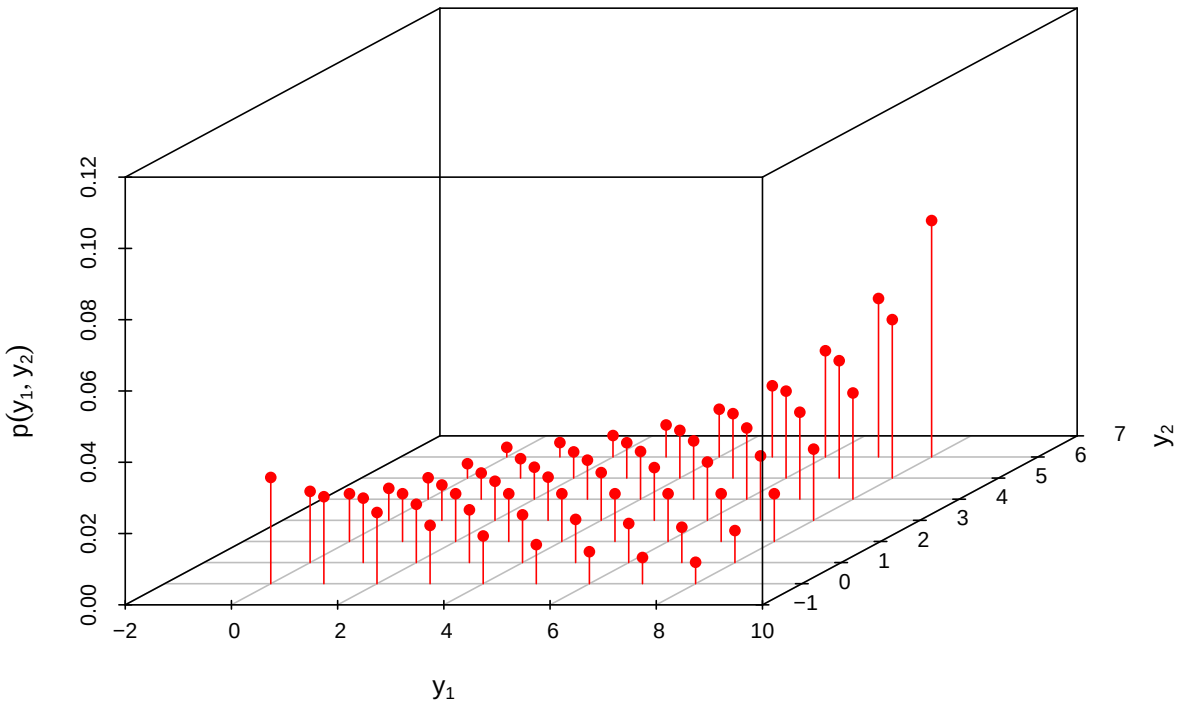
	6	0.003	0.004	0.006	0.009	0.013	0.020	0.030	0.044	0.066
	5	0.004	0.005	0.007	0.010	0.013	0.018	0.024	0.033	0.044
	4	0.006	0.007	0.009	0.011	0.013	0.016	0.020	0.024	0.030
Y_2	3	0.009	0.010	0.011	0.012	0.013	0.015	0.016	0.018	0.020
	2	0.013	0.013	0.013	0.013	0.013	0.013	0.013	0.013	0.013
	1	0.020	0.018	0.016	0.015	0.013	0.012	0.011	0.010	0.009
	0	0.030	0.024	0.020	0.016	0.013	0.011	0.009	0.007	0.006
		0	1	2	3	4	5	6	7	8
						Y_1				

Thus for example

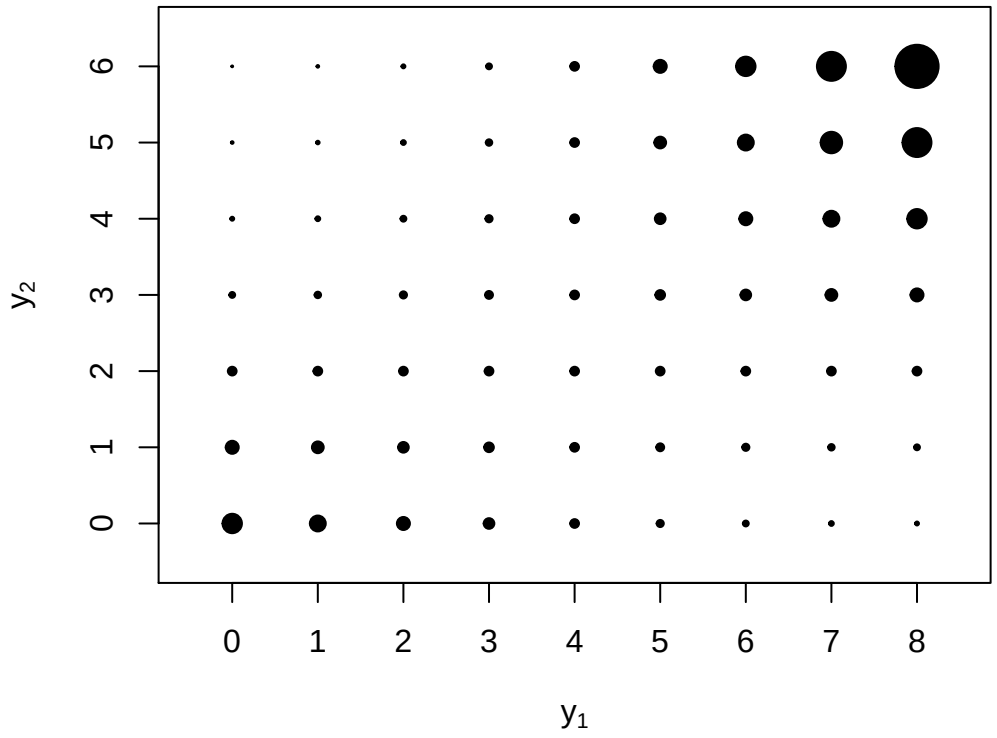
$$p(0, 3) = 0.009 \qquad p(4, 1) = 0.013 \qquad p(3, 6) = 0.009 \qquad p(8, 5) = 0.044$$

This function specifies how much probability mass is placed at each (y_1, y_2) value; the total probability mass is 1. The function can be displayed in several different ways:

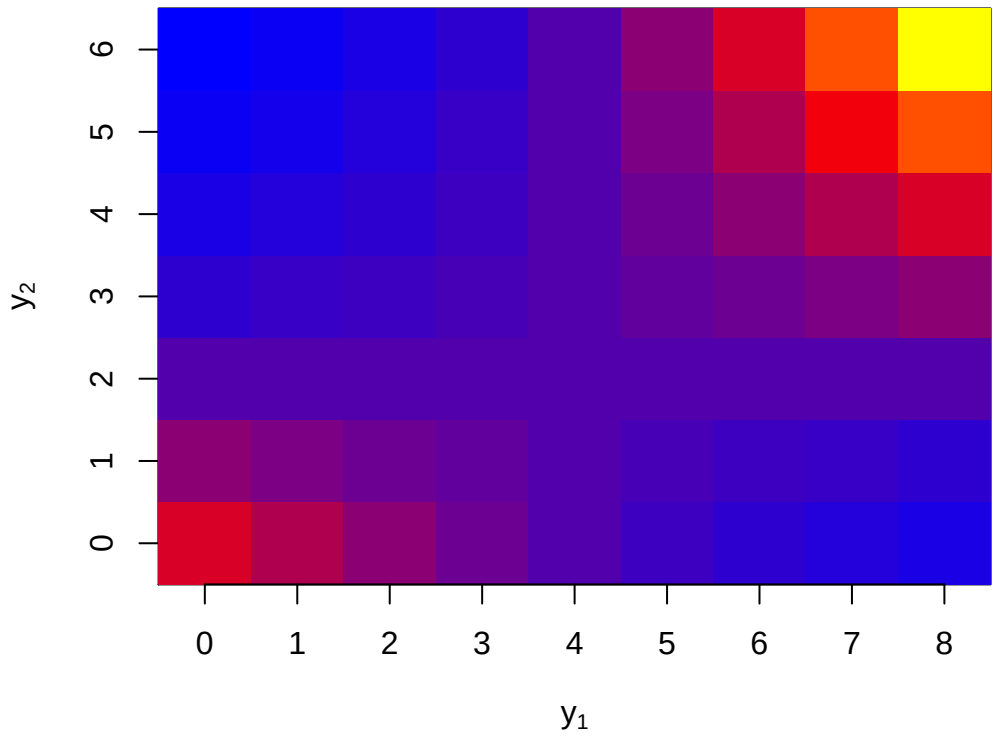
- **Surface plot:** here the probability masses are plotted vertically at the specified (y_1, y_2) positions, with a vertical line projecting down onto the (y_1, y_2)-plane.



- **Point plot:** here the radius of the point at (y_1, y_2) indicates the amount of probability placed there.



- **Image plot:** here the colours represent the probability value (yellow is high, blue is low).



From the joint pmf, we can compute the *marginal* pmf for each of the variables Y_1 and Y_2. We obtain these using the Theorem of Total Probability as follows: we have for example that

$$p_{Y_1}(y_1) = P(Y_1 = y_1) = P(Y_1 = y_1, Y_2 = 0) + P(Y_1 = y_1, Y_2 = 1) + \cdots + P(Y_1 = y_1, Y_2 = 6)$$
$$= p_{Y_1,Y_2}(y_1, 0) + p_{Y_1,Y_2}(y_1, 1) + \cdots + p_{Y_1,Y_2}(y_1, 6)$$

and so

$$Y_1: \quad p_{Y_1}(y_1) = \sum_{y_2=0}^{6} p_{Y_1,Y_2}(y_1, y_2)$$

$$Y_2: \quad p_{Y_2}(y_2) = \sum_{y_1=0}^{8} p_{Y_1,Y_2}(y_1, y_2).$$

Thus we obtain the marginal distribution for Y_1 by summing down the *columns* of the above tables, and the marginal distribution for Y_2 by summing across the *rows* of the tables. In the above numerical example, we have

y_1	0	1	2	3	4	5	6	7	8
$p_{Y_1}(y_1)$	0.085	0.083	0.083	0.087	0.094	0.106	0.124	0.151	0.189

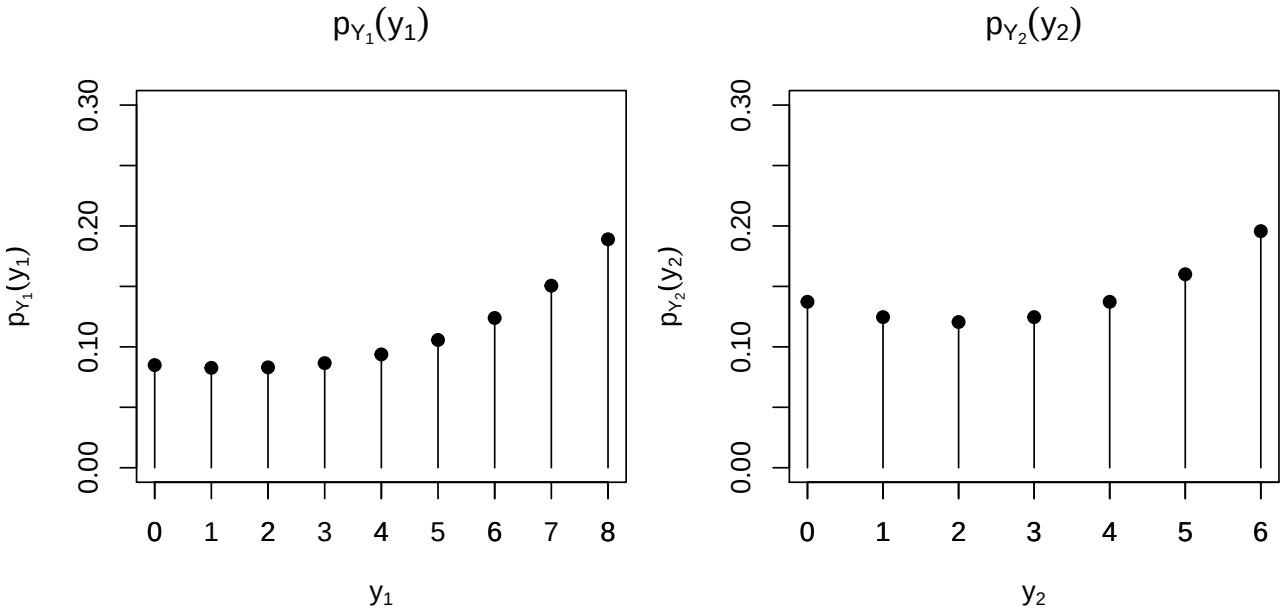
and

y_2	0	1	2	3	4	5	6
$p_{Y_2}(y_2)$	0.137	0.125	0.121	0.125	0.137	0.160	0.196

Note that these functions are both pmfs: we have

$$\sum_{y_1=0}^{8} p_{Y_1}(y_1) = \sum_{y_2=0}^{6} p_{Y_2}(y_2) = 1.$$

We may plot these pmfs in the usual way.



Finally, we can consider *conditional* pmfs: we have by the definition of conditional probability

$$P(Y_1 = y_1 | Y_2 = y_2) = \frac{P(Y_1 = y_1, Y_2 = y_2)}{P(Y_2 = y_2)} \qquad 0 \leq y_1 \leq 8$$

for each y_2, $0 \leq y_2 \leq 6$. This enables us to define the conditional pmf for Y_1, given that $Y_2 = y_2$, as

$$p_{Y_1|Y_2}(y_1|y_2) = P(Y_1 = y_1 | Y_2 = y_2) = \frac{p_{Y_1,Y_2}(y_1, y_2)}{p_{Y_2}(y_2)} \qquad 0 \leq y_1 \leq 8$$

for each y_2, $0 \leq y_2 \leq 6$. This function is a pmf in terms of y_1, that is for example

$$\sum_{y_1=0}^{8} p_{Y_1|Y_2}(y_1|y_2) = 1$$

for each y_2, $0 \leq y_2 \leq 6$. Similarly we may define the conditional pmf for Y_2, given that $Y_1 = y_1$, as

$$p_{Y_2|Y_1}(y_2|y_1) = P(Y_2 = y_2 | Y_1 = y_1) = \frac{p_{Y_1,Y_2}(y_1, y_2)}{p_{Y_1}(y_1)} \qquad 0 \leq y_2 \leq 6$$

for each y_1, $0 \leq y_1 \leq 8$. This function is a pmf in terms of y_2.

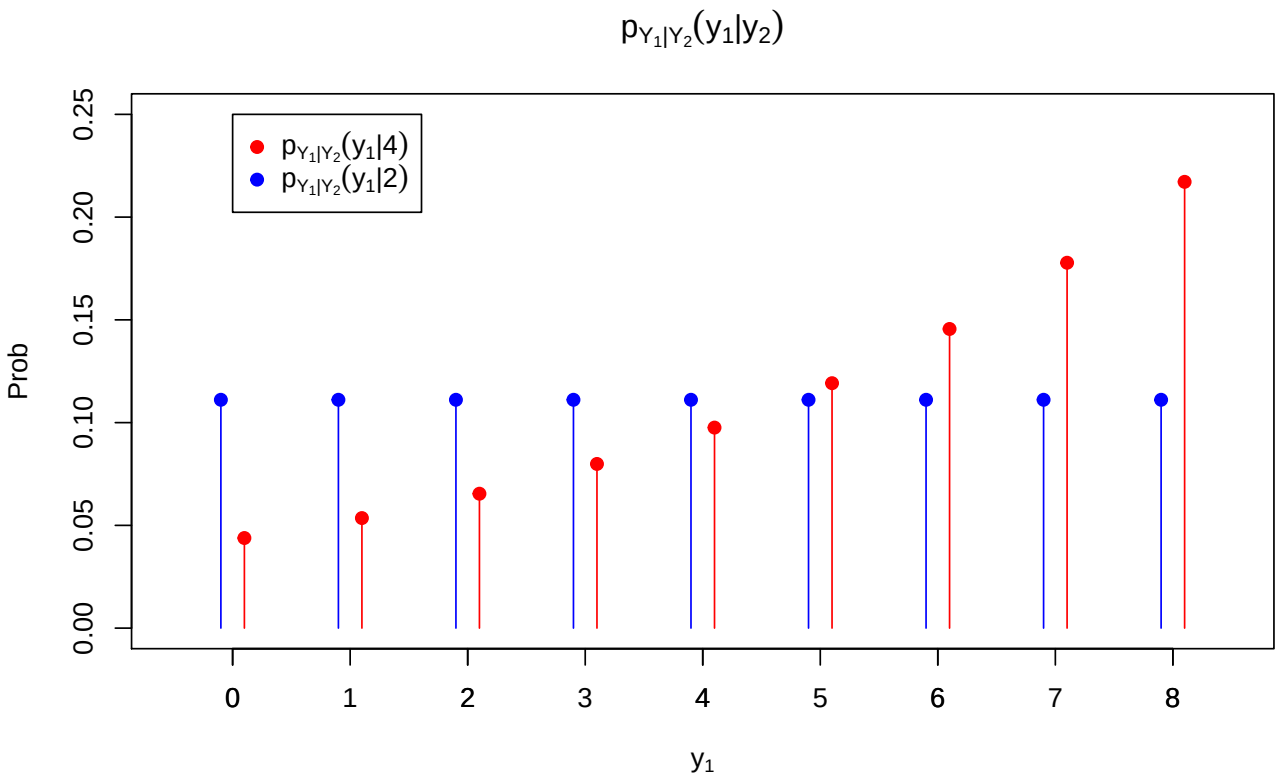
We compute the conditional pmfs by taking 'slices' through the joint pmf, that is, we examine individual rows or columns, and then standardize by dividing by the row or column sum. For example, for the conditional pmf for Y_1 given that $Y_2 = 4$, we examine the $Y_2 = 4$ row

Y_4	4	$p(0,4)$	$p(1,4)$	$p(2,4)$	$p(3,4)$	$p(4,4)$	$p(5,4)$	$p(6,4)$	$p(7,4)$	$p(8,4)$
		0	1	2	3	4	5	6	7	8
						Y_1				

and divide by $p_{Y_2}(4)$ to get the conditional pmf.

Numerically, we have for $p_{Y_1|Y_2}(y_1|4)$

Y_2 4	0.044	0.054	0.065	0.080	0.098	0.119	0.146	0.178	0.217
	0	1	2	3	4	5	6	7	8
					Y_1				



For the conditional pmf for Y_1 given that $Y_2 = 2$, we examine the $Y_2 = 2$ row and divide by $p_{Y_2}(2)$ to get the

Y_2 2	$p(0,2)$	$p(1,2)$	$p(2,2)$	$p(3,2)$	$p(4,2)$	$p(5,2)$	$p(6,2)$	$p(7,2)$	$p(8,2)$
	0	1	2	3	4	5	6	7	8
					Y_1				

conditional pmf. Numerically, we have for $p_{Y_1|Y_2}(y_1|2)$

Y_2 2	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111
	0	1	2	3	4	5	6	7	8
					Y_1				

that is, constant across all values of y_1, $0 \leq y_1 \leq 8$. In general, the function $p_{Y_1|Y_2}(y_1|y_2)$ is different for each y_2,

Similarly, to get the conditional pmf for Y_2 given that $Y_1 = 5$, we examine the $Y_1 = 5$ column and divide by $p_{Y_1}(5)$ to get the conditional pmf

Y_1	5	0.091	0.101	0.111	0.123	0.136	0.150	0.166
		0	1	2	3	4	5	6
					Y_2			

To get the conditional pmf for Y_2 given that $Y_1 = 0$, we examine the $Y_1 = 0$ column and divide by $p_{Y_1}(0)$ to get the conditional pmf

Y_1	5	0.351	0.235	0.158	0.106	0.071	0.048	0.032
		0	1	2	3	4	5	6
					Y_2			

