

1 Marginal Probability Distribution

Let X and Y be random variables whose distribution can be obtained by summing or integrating over all relevant values of the other random variable. This provides the **marginal** p.m.f or p.d.f

$$f_X(x) = \sum_y f_{XY}(x, y) \quad f_Y(y) = \sum_x f_{XY}(x, y) \quad (1)$$

$$f_X(x) = \int_y f_{XY}(x, y) dy \quad f_Y(y) = \int_x f_{XY}(x, y) dx \quad (2)$$

Example 5.3

Let the columns X be the number of bars of service. Let the rows Y be the response time from cell towers

	1	2	3
4	0.15	0.1	0.05
3	0.02	0.1	0.05
2	0.02	0.03	0.2
1	0.01	0.02	0.25

From this we see:

x	$f_X(x)$
1	$0.15 + 0.02 + 0.02 + 0.01$
2	$0.1 + 0.1 + 0.03 + 0.02$
3	$0.05 + 0.05 + 0.2 + 0.25$

and

y	$f_Y(y)$
1	$0.01 + 0.02 + 0.25$
2	$0.02 + 0.03 + 0.2$
3	$0.02 + 0.1 + 0.05$
4	$0.15 + 0.1 + 0.05$

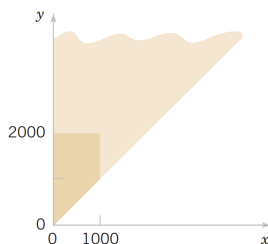
We can also find the probabilities between intervals for each random variable separately. For instance

$$P(a < X \leq b) = \int_{x=a}^b f_X(x) dx = \int_{x=a}^b \left(\int_{y=-\infty}^{\infty} f_{XY}(x, y) dy \right) dx \quad (3)$$

Example 5.4 Find $P(Y > 2000)$ if the joint p.d.f of (X, Y) is

$$f_{XY} = \begin{cases} 6(10^{-6} \exp(-0.001x - 0.002y)) & 0 < x < y < \infty \\ 0 & \text{else} \end{cases}$$

Draw a diagram to find the limits of integration



Example 5.4 Contd. From the diagram we see

$$\int_{y=2000}^{\infty} \int_{x=0}^y 6(10^{-6} \exp(-0.001x - 0.002y)) dx dy = 0.0500$$

2 Independent Random Variables

Recall that for independent events E and F , the probability of $P(E \cap F) = P(E)P(F)$. Similarly if X and Y are random variable and independent if and only if

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B) \quad (4)$$

In other words, any event involving X is independent of any event involving Y . Also recall the following is true

$$P(X \in A, Y \in B) = f_{XY}(x, y) = f_X(x)f_Y(y) \text{ for all } x, y \quad (5)$$

Independence requires a rectangular range.

Independence depends on the physical situation being modelled and needs to be justified, for instance Binomial random variables require independent outcomes. Whereas the Hypergeometric has dependent outcomes.

Example 5.11 This following example is the same as the previous, but we are going to modify the bounds to expand the rectangular region of integration. Notice the coefficient is smaller, this is due to the larger bounds (total probability $\int_0^{\infty} \int_0^{\infty} f_{XY}(x, y) dx dy = 1$). Show that X and Y are independent and find $P(X > 1000, Y < 1000)$

$$f_{XY} = \begin{cases} 2(10^{-6} \exp(-0.001x - 0.002y)) & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{else} \end{cases}$$

Proof

$$f_X = \int_0^{\infty} 2(10^{-6} \exp(-0.001x - 0.002y)) dy = 0.001e^{-0.001x}$$

$$f_Y = \int_0^{\infty} 2(10^{-6} \exp(-0.001x - 0.002y)) dx = 0.002e^{-0.002y}$$

$$f_X f_Y = 0.002 \exp(-0.002y) * 0.001e^{-0.001x} = f_{XY}$$

Thus we have shown the variables are independent of each other

$$\begin{aligned} P(X > 1000, Y < 1000) &= P(X > 1000)P(Y < 1000) \\ &= \int_{1000}^{\infty} 0.001e^{-0.001x} dx \int_0^{\infty} 0.002e^{-0.002y} dy \\ &= 0.3181 \end{aligned}$$

Example 5.12 Suppose (X, Y) are independent

$$X \sim N(10.5, 0.0025) \quad Y \sim N(3.2, 0.0036)$$

Find $P(10.4 < X < 10.6, 3.15 < Y < 3.25)$

Ans.

$$\begin{aligned} P &= P(10.4 < X < 10.6)P(3.15 < Y < 3.25) \\ &= P\left(\frac{10.4 - 10.5}{\sqrt{0.0025}} < Z < \frac{10.6 - 10.5}{\sqrt{0.0025}}\right)P\left(\frac{3.15 - 3.2}{\sqrt{0.0036}} < Z < \frac{3.25 - 3.2}{\sqrt{0.0036}}\right) \\ &= P(-2 < Z < 2)P(-0.837 < Z < 0.833) \\ &= (P(Z < 2) - P(Z < -2))(P(Z < 0.833) - P(Z < -0.837)) \\ &= 0.586 \end{aligned}$$