3.5.1 Marginal Distributions

$$
\begin{array}{cc|ccc|c}
 & & & x & & \\
 & & 0 & 1 & 2 & \\
\hline
 & 0 & \frac{6}{36} & \frac{12}{36} & \frac{3}{36} & \frac{21}{36} \\
y & 1 & \frac{8}{36} & \frac{6}{36} & & \frac{14}{36} \\
 & 2 & \frac{1}{36} & & & \frac{1}{36} \\
\hline
 & & \frac{15}{36} & \frac{18}{36} & \frac{3}{36} &
\end{array}
$$

Returning to the caplet example, sum the rows and columns of the table of probabilities.

The column sums are the probabilities that $X = 0, 1, 2$ respectively, and the row sums are probabilities that $Y = 0, 1, 2$ respectively.

Therefore, the column totals are the probability distribution for X: for $x = 0, 1, 2$

$$g(x) = P(X = x) = \sum_{y=0}^{2} f(x, y),$$

and the row totals are the probability distribution for Y: for $y = 0, 1, 2$

$$h(y) = P(Y = y) = \sum_{x=0}^{2} f(x, y).$$

If X and Y are discrete random variables, and $f(x, y)$ is their joint probability distribution, then the function

$$g(x) = \sum_y f(x, y)$$

is called the **marginal distribution of** X and the function

$$h(y) = \sum_x f(x, y)$$

is called the **marginal distribution of** Y. The sums are over all values of either y or x respectively.

If X and Y are jointly continuous random variables, and $f(x, y)$ is their joint probability density function, then

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

is called the **marginal density of** X and the function

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

is called the **marginal density of** Y. These functions are defined for all $x \in \mathbb{R}$,$y \in \mathbb{R}$ respectively.

Example 3.5.8 *Find the marginal densities of X and Y given their joint probability density*

$$f(x, y) = \begin{cases} \frac{2}{3}(x + 2y) & \textit{for } 0 < x < 1, 0 < y < 1 \\ 0 & \textit{elsewhere} \end{cases}$$

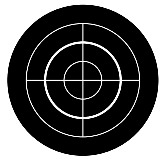
Marginal density of X:

$$g(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \int_0^1 \frac{2}{3}(x + 2y)\, dy = \ldots$$

Marginal density of Y:

$$h(y) = \int_{-\infty}^{\infty} f(x, y)\, dx = \int_0^1 \frac{2}{3}(x + 2y)\, dx = \ldots$$

Note that joint marginal distributions can be defined in the case of 3 or more random variables, and that is beyond the scope of this course too.



Example 3.5.9 *A circular biathlon target for prone position has a diameter of 45mm. Suppose that each point on the target has equally likely probability of being hit by a shot.*

Let $(0,0)$ *be the centre of the target, and define random variables* X *and* Y*, so that* (X,Y) *denotes the coordinates (in millimetres) of the shot fired.*

The joint density function for X *and* Y *is then, for some constant* k*,*

$$f(x,y) = \begin{cases} k & \textit{for } x^2+y^2 \le (22.5)^2 \\ 0 & \textit{elsewhere} \end{cases}$$

It follows that $k = \frac{1}{(22.5)^2\pi}$, so the integral of the joint density function equals 1 over the area of the circle.

To find the marginal density for X, integrate over all y values:

$$x^2 + y^2 \le (22.5)^2 \Rightarrow \quad y^2 \le (22.5)^2 - x^2$$

$$\Rightarrow \quad -\sqrt{(22.5)^2 - x^2} \le y \le \sqrt{(22.5)^2 - x^2}$$

Thus

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_{-\sqrt{(22.5)^2-x^2}}^{\sqrt{(22.5)^2-x^2}} \frac{1}{(22.5)^2\pi}\,dy$$

$$=$$

The marginal density of X can be used to find the probability the shot will land any horizontal distance x from the centre, regardless of its vertical position.

Notice that $g(x)$ is largest when $x = 0$ and gets smaller as x gets near the boundary of the target.

3.5.2 Conditional Distributions

Recall: Conditional probability of event A given event B:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad (P(B) \neq 0)$$

In terms of random variables: If A is the event $X = x$ and B is the event $Y = y$ then

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}.$$

For discrete random variables with joint probability distribution $f(x, y)$ we have

$$P(X = x, Y = y) = \frac{f(x, y)}{h(y)},$$

where $h(y) \neq 0$ is the marginal distribution of Y.

If X and Y are discrete random variables with joint probability distribution $f(x, y)$, and respective marginal distributions $g(x)$ and $h(y)$, the function

$$f(x|y) = \frac{f(x, y)}{h(y)}$$

is called the **conditional distribution of X given $Y = y$**, provided $h(y) \neq 0$. The function

$$f(y|x) = \frac{f(x, y)}{g(x)}$$

is called the **conditional distribution of Y given $X = x$**, provided $g(x) \neq 0$.

Example 3.5.10

$$
\begin{array}{cc|ccc|c}
 & & & x & & \\
 & & \mathit{0} & \mathit{1} & \mathit{2} & \\
\hline
 & \mathit{0} & \frac{6}{36} & \frac{12}{36} & \frac{3}{36} & \frac{21}{36} \\
y & \mathit{1} & \frac{8}{36} & \frac{6}{36} & & \frac{14}{36} \\
 & \mathit{2} & \frac{1}{36} & & & \frac{1}{36} \\
\hline
 & & \frac{15}{36} & \frac{18}{36} & \frac{3}{36} &
\end{array}
$$

Caplet example: The conditional distribution of X given $Y = 1$ is, $f(x|1) = \frac{f(x,1)}{h(1)}$.

Its values are:

$$f(0|1) = \frac{f(0,1)}{h(1)} = \dots$$

$$f(1|1) = \frac{f(1,1)}{h(1)} = \dots$$

$$f(2|1) = \dots$$

Conditional distribution for discrete random variables is extended to the idea of conditional density for jointly continuous random variables:

For jointly continuous random variables X and Y with joint density $f(x, y)$, and marginal densities $g(x)$ and $h(y)$:

The function

$$f(x|y) = \frac{f(x, y)}{h(y)}$$

is called the **conditional density of X given $Y = y$**, provided $h(y) \neq 0$.

The function

$$f(y|x) = \frac{f(x, y)}{g(x)}$$

is called the **conditional density of Y given $X = x$**, provided $g(x) \neq 0$.