Joint Probability

The function p(x, y) is a joint probability mass function of the discrete random variables X and Y if

1. $p(x, y) \geq 0$ for all (x, y)
2. $$\sum_{x}\sum_{y} p(x, y) = 1$$
3. $P(X = x, Y = y) = p(x, y)$

For any region A in the xy plane,

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

Example

Two refills for a ballpoint pen are selected at random from a box that contains 3 blue refills, 2 red refills and 3 green refills. If X is the number of blue refills and Y is the number of red refills selected, find

a. The joint probability mass function p(x, y), and

b. $P[(X, Y) \in A]$, where A is the region $\{(x, y) | x + y \leq 1\}$

Example - Solution

The possible pairs of values (x, y) are (0, 0), (0, 1), (1, 0), (1, 1), (0, 2), and (2, 0), where p(0, 1), for example represents the probability that a red and a green refill are selected. The total number of equally likely ways of selecting any 2 refills from the 8 is:

$$\binom{8}{2} = \frac{8!}{2!6!} = 28$$

The number of ways of selecting 1 red from 2 red refills and 1 green from 3 green refills is

$$\binom{2}{1}\binom{3}{1} = 6$$

Example - Solution

Hence, p(0, 1) = 6/28 = 3/14. Similar calculations yield the probabilities for the other cases, which are presented in the following table. Note that the probabilities sum to 1.

		x			
		0	1	2	Row Totals
	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
y	1	$\frac{3}{14}$	$\frac{3}{14}$		$\frac{3}{7}$
	2	$\frac{1}{28}$			$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Example - Solution

$$p(x, y) = \frac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{\binom{8}{2}}$$

for x = 0, 1, 2; y = 0, 1, 2; $0 \le x + y \le 2$.

b. $P[(X, Y) \in A] = P(X + Y \le 1)$

$= p(0, 0) + p(0, 1) + p(1, 0)$

$$= \frac{3}{28} + \frac{3}{14} + \frac{9}{28}$$

$$= \frac{9}{14}$$

Joint Density Functions

The function f(x, y) is a joint probability density function of the continuous random variables X and Y if

1. $f(x, y) \geq 0$ for all (x, y)

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

3. $$P[(X, Y) \in A] = \iint_A f(x, y)\,dx\,dy$$

For any region A in the xy plane.

Marginal Distributions

The marginal probability mass functions of x alone and of Y alone are

$$g(x) = \sum_{y} p(x, y) \quad \text{and} \quad h(y) = \sum_{x} p(x, y)$$

for the discrete case.

Marginal Distributions

The marginal probability density functions of x alone and of Y alone are

$$g(x) = \int_{-\infty}^{\infty} f(x, y)dxdy \quad \text{and} \quad h(y) = \int_{-\infty}^{\infty} f(x, y)dxdy$$

for the continuous case.

Conditional Probability Distributions

Let X and Y be two discrete random variables, with joint probability mass function p(x,y) and marginal probability mass functions m(x) and n(y). The conditional probability mass function of the random variable Y, given that X = x, is

$$l(y \mid x) = \frac{p(x, y)}{m(x)} \quad , \qquad m(x) > 0$$

Similarly, the conditional probability mass function of the random variable X, given that Y = y, is

$$l(x \mid y) = \frac{p(x, y)}{n(y)} \quad , \qquad n(y) > 0$$

Statistical Independence

Let X and Y be two discrete random variables, with joint probability mass function p(x, y) and marginal probability mass functions m(x) and n(y), respectively. The random variables X and Y are said to be statistically independent if and only if

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

for all (x, y) within their range.

Statistically Independent

Let $X_1, X_2, \ldots, X_n$ be n discrete random variables, with joint probability mass functions $p(x_1, x_2, \ldots, x_n)$ and marginal probability mass functions $p_1(x_1), p_2(x_2), \ldots, p_n(x_n)$, respectively. The random variables $X_1, X_2, \ldots, X_n$ are said to be mutually statistically independent if and only if

$$p(x_1, x_2, \ldots, x_n) = p_1(x_1), p_2(x_2), \ldots, p_n(x_n)$$

for all $(x_1, x_2, \ldots, x_n)$ within their range.

Example

A candy company distributed boxes of chocolates with a mixture of creams, toffees, and nuts coated in both light and dark chocolate. For a randomly selected box, let X and Y, respectively, be the proportions of the light and dark chocolates that are creams and suppose that the joint density function is

$$f(x, y) = \begin{cases} \frac{2}{5}(2x+3y), & 0 \le x \le 1, 0 \le y \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

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

b) Find $P[(X,Y) \in A]$, where A is the region $\{(x,y) \mid 0<x<\frac{1}{2}, \frac{1}{4}<y<\frac{1}{2}\}$.

Example – Solution

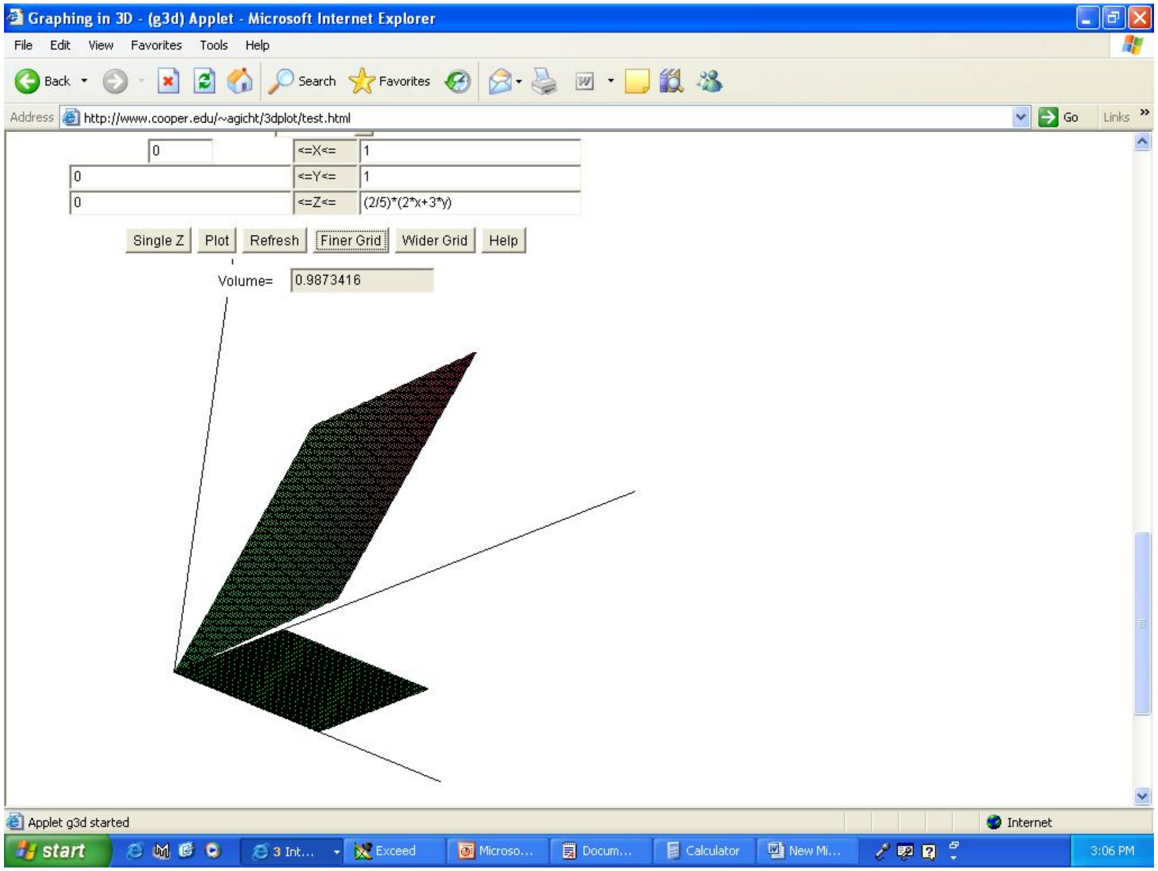
a) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \int_0^1\int_0^1 \frac{2}{5}(2x+3y)\,dx\,dy$$

$$= \int_0^1 \left.\frac{2x^2}{5} + \frac{6xy}{5}\right|_{x=0}^{x=1} dy$$

$$= \int_0^1 \left(\frac{2}{5} + \frac{6y}{5}\right) dy = \left.\frac{2y}{5} + \frac{3y^2}{5}\right|_0^1$$

$$= \frac{2}{5} + \frac{3}{5} = 1$$

3D plotting for example problem



Example – Solution

b) $P\left[(X,Y)\in A\right] = P(0 < X < \tfrac{1}{2}, \tfrac{1}{4} < Y < \tfrac{1}{2})$

$$= \int_{\frac{1}{4}}^{\frac{1}{2}} \int_{0}^{\frac{1}{2}} \frac{2}{5}(2x+3y)\,dx\,dy$$

$$= \int_{\frac{1}{4}}^{\frac{1}{2}} \left. \frac{2x^2}{5} + \frac{6xy}{5} \right|_{x=0}^{x=\frac{1}{2}} dy$$

$$= \int_{\frac{1}{4}}^{\frac{1}{2}} \left(\frac{1}{10} + \frac{3y}{5} \right) dy = \left. \frac{y}{10} + \frac{3y^2}{10} \right|_{\frac{1}{4}}^{\frac{1}{2}}$$

$$= \frac{1}{10}\left[\left(\frac{1}{2} + \frac{3}{4} \right) - \left(\frac{1}{4} + \frac{3}{16} \right) \right] = \frac{13}{160}$$

Example

Show that the column and row totals of the following table give the marginal distribution of X alone and of Y alone.

		X			
		0	1	2	Row Totals
	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
y	1	$\frac{3}{14}$	$\frac{3}{14}$		$\frac{3}{7}$
	2	$\frac{1}{28}$			$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Example – Solution

For the random variable X, we see that

$$P(X=0)=g(0)=\sum_{y=0}^{2} f(0,y)= f(0,0)+ f(0,1)+ f(0,2)$$

$$=\frac{3}{28}+\frac{3}{14}+\frac{1}{28}=\frac{5}{14}$$

$$P(X=1)=g(1)=\sum_{y=0}^{2} f(1,y)= f(1,0)+ f(1,1)+ f(1,2)$$

$$=\frac{9}{28}+\frac{3}{14}+0=\frac{15}{28}$$

$$P(X=2)=g(2)=\sum_{y=0}^{2} f(2,y)= f(2,0)+ f(2,1)+ f(2,2)$$

$$=\frac{3}{28}+0+0=\frac{3}{28}$$

Example

Find g(x) and h(y) for the joint density function of the previous example :

$$f(x,y)=\begin{cases}\frac{2}{5}(2x+3y), & 0\le x\le 1, 0\le y\le 1\\ 0, & \text{elsewhere}\end{cases}$$

Example – Solution

By definition,

$$g(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^1 \frac{2}{5}(2x+3y)dy = \frac{4xy}{5} + \frac{6y^2}{10}\Bigg|_{y=0}^{y=1} = \frac{4x+3}{5}$$

For 0≤x ≤1, and g(x)=0 elsewhere. Similarly,

$$h(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_0^1 \frac{2}{5}(2x+3y)dx = \frac{4(1+3y)}{5}$$

For 0 ≤y ≤1, and h(y)=0 elsewhere.

Conditional Probability Distribution

Let X and Y be two continuous random variables, with joint probability density function f(x,y) and marginal probability density functions g(x) and h(y). The conditional probability density function of the random variable Y, given that X = x, is

$$f(y \mid x) = \frac{f(x, y)}{g(x)}, \qquad g(x) > 0$$

Similarly, the conditional probability density function of the random variable X, given that Y = y, is

$$f(x \mid y) = \frac{f(x, y)}{h(y)}, \qquad h(y) > 0$$

Example

Given the joint density function

$$f(x,y)=\begin{cases}\dfrac{x(1+3y^2)}{4}, & 0<x<2, 0<y<1\\ 0, & \text{elsewhere}\end{cases}$$

Find g(x), h(y), f(x|y), and evaluate P(¼<X<½|Y=1/3).

Example – Solution

By definition,

$$g(x)=\int_{-\infty}^{\infty} f(x,y)dy=\int_0^1 \frac{x(1+3y^2)}{4}dy=\frac{xy}{4}+\frac{xy^3}{4}\bigg|_{y=0}^{y=1}=\frac{x}{2},\quad 0<\mathrm{x}<2$$

and

$$h(y)=\int_{-\infty}^{\infty} f(x,y)dx=\int_0^2 \frac{x(1+3y^2)}{4}dy=\frac{x^2}{8}+\frac{3x^2y^2}{8}\bigg|_{x=0}^{x=2}=\frac{1+3y^2}{2},\quad 0<\mathrm{y}<1$$

therefore,

$$f(x\,|\,y)=\frac{f(x,y)}{h(y)}=\frac{x(1+3y^2)/4}{(1+3y^2)/2}=\frac{x}{2},\quad 0<\mathrm{x}<2$$

and

$$P\left(\frac{1}{4}<X<\frac{1}{2}\,|\,Y=\frac{1}{3}\right)=\int_{\frac{1}{4}}^{\frac{1}{2}} \frac{x}{2}dx=\frac{3}{64}.$$

Statistical Independence

Let X and Y be two continuous random variables, with joint probability density function f(x, y) and marginal probability density functions g(x) and h(y), respectively. The random variables X and Y are said to be statistically independent if and only if

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

for all (x, y) within their range.

Statistically Independent

Let $X_1, X_2, \ldots, X_n$ be n continuous random variables, with joint probability density functions $f(x_1, x_2, \ldots, x_n)$ and marginal probability functions $f_1(x_1), f_2(x_2), \ldots, f_n(x_n)$, respectively. The random variables $X_1, X_2, \ldots, X_n$ are said to be mutually statistically independent if and only if

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1), f_2(x_2), \ldots, f_n(x_n)$$

for all $(x_1, x_2, \ldots, x_n)$ within their range.

Example

Two refills for a ballpoint pen are selected at random from a box that contains 3 blue refills, 2 red refills and 3 green refills. If X is the number of blue refills and Y is the number of red refills selected, Show that the random variables are not statistically independent.

Example – Solution

Let us consider the point (0,1). From the following table:

		x			
		0	1	2	Row Totals
	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
y	1	$\frac{3}{14}$	$\frac{3}{14}$		$\frac{3}{7}$
	2	$\frac{1}{28}$			$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Example – Solution

We find the three probabilities f(0,1), g(0), and h(1) to be

$$f(0,1) = \frac{3}{14},$$

$$g(0) = \sum_{y=0}^{2} f(0,y) = \frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14},$$

$$h(1) = \sum_{y=0}^{2} f(x,1) = \frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{17}$$

Clearly, $f(0,1) \neq g(0)h(1)$ and therefore X and Y are not statistically independent.