Lecture 14: Joint Distributions

Sta230 / Mth230

Colin Rundel

March 19, 2014

Joint Distribution - Example

Draw two socks at random, without replacement, from a drawer full of twelve colored socks:

6 black, 4 white, 2 purple

Let B be the number of Black socks, W the number of White socks drawn, then the distributions of B and W are given by:

	0	1	2
P(B=k)	$\frac{6}{12}\frac{5}{11} = \frac{15}{66}$	$2\frac{6}{12}\frac{6}{11} = \frac{36}{66}$	$\frac{6}{12}\frac{5}{11} = \frac{15}{66}$
P(W=k)	$\frac{8}{12}\frac{7}{11} = \frac{28}{66}$	$2\frac{4}{12}\frac{8}{11} = \frac{32}{66}$	$\frac{4}{12}\frac{3}{11} = \frac{6}{66}$

Note - $B \sim \text{HyperGeo}(12, 6, 2) = \frac{\binom{6}{k}\binom{6}{2-k}}{\binom{12}{2}}$ and $W \sim \text{HyperGeo}(12, 4, 2) = \frac{\binom{4}{k}\binom{8}{2-k}}{\binom{12}{2}}$

Joint Distribution - Example, cont.

Let B be the number of Black socks and W the number of White socks drawn, then the joint distribution of B and W is given by:

		W = 0	W = 1	W = 2	
B	0	$\frac{1}{66}$	$\frac{8}{66}$	$\frac{6}{66}$	$\frac{15}{66}$
	1	$\frac{12}{66}$	$\frac{24}{66}$	0	$\frac{36}{66}$
	2	$\frac{15}{66}$	0	0	$\frac{15}{66}$
		$\frac{28}{66}$	$\frac{32}{66}$	$\frac{6}{66}$	$\frac{66}{66}$

$$P(B = b, W = w) = \begin{cases} 1/66 & \text{If b=0,w=0} \\ 8/66 & \text{If b=0,w=1} \\ 6/66 & \text{If b=0,w=2} \\ 12/66 & \text{If b=1,w=0} \\ 24/66 & \text{If b=1,w=1} \\ 0/66 & \text{If b=1,w=2} \\ 15/66 & \text{If b=2,w=0} \\ 0/66 & \text{If b=2,w=1} \\ 0/66 & \text{If b=2,w=2} \\ 0 & \text{otherwise} \end{cases}$$

$$P(B = b, W = w) = \frac{\binom{6}{b}\binom{4}{w}\binom{2}{2-b-w}}{\binom{12}{2}}, \text{ for } 0 \le b, w \le 2 \text{ and } b + w \le 2$$

Marginal Distributions

Note that the column and row sums are the distributions of B and W respectively.

$$P(B = b) = P(B = b, W = 0) + P(B = b, W = 1) + P(B = b, W = 2)$$
$$P(W = w) = P(B = 0, W = w) + P(B = 1, W = w) + P(B = 2, W = w)$$

These are the *marginal* distributions of B and W. In general,

$$P(X = x) = \sum_y P(X = x, Y = y) = \sum_y P(X = x|Y = y)P(Y = y)$$

Conditional Distribution

Conditional distributions are defined as we have seen previously with

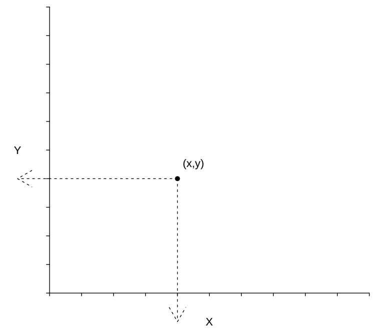
$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{\text{joint pmf}}{\text{marginal pmf}}$$

Therefore the pmf for white socks given no black socks were drawn is

$$P(W = w|B = 0) = \frac{P(W = w, B = 0)}{P(B = 0)} = \begin{cases} \frac{1}{66} \Big/ \frac{15}{66} = \frac{1}{15} & \text{if } W = 0 \\ \frac{8}{66} \Big/ \frac{15}{66} = \frac{8}{15} & \text{if } W = 1 \\ \frac{6}{66} \Big/ \frac{15}{66} = \frac{6}{15} & \text{if } W = 2 \end{cases}$$

Joint CDF

$$\begin{aligned} F(x, y) &= P[X \le x, Y \le y] \\ &= P[(X, Y) \text{ lies south-west of the point } (x, y)] \end{aligned}$$



Joint CDF, cont.

The joint Cumulative distribution function follows the same rules as the univariate CDF,

Univariate definition:

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(z)dz$$

$$\lim_{x \to -\infty} F(x) = 0 \qquad \lim_{x \to \infty} F(x) = 1 \qquad x \le y \Rightarrow F(x) \le F(y)$$

Bivariate definition:

$$F(x, y) = P(X \le x, Y \le y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(x, y)\; dx\; dy$$

$$\lim_{x,y \to -\infty} F(x, y) = 0 \qquad \lim_{x,y \to \infty} F(x, y) = 1 \qquad x \le x', y \le y' \Rightarrow F(x, y) \le F(x', y')$$

Marginal Distributions

We can define marginal CDFs using the joint CDF by setting one of the values to infinity:

$$\begin{aligned} F(x, \infty) = P(X \le x, Y \le \infty) &= \int_{-\infty}^{x} \int_{-\infty}^{\infty} f(x, y)\; dy\; dy \\ &= P(X \le x) = F_X(x) \end{aligned}$$

$$\begin{aligned} F(\infty, y) = P(X \le \infty, Y \le y) &= \int_{-\infty}^{y} \int_{-\infty}^{\infty} f(x, y)\; dx\; dy \\ &= P(Y \le y) = F_Y(y) \end{aligned}$$

Joint pdf

Similar to the CDF the probability density function follows the same general rules in two dimensions,

Univariate definition:

$f(x) \geq 0$ for all x $\qquad f(x) = \frac{d}{dx}F(x) \qquad \int_{-\infty}^{\infty} f(x)dx = 1$

Bivariate definition:

$$f(x,y) \geq 0 \text{ for all } (x,y)$$

$$f(x,y) = \frac{\partial}{\partial x}\frac{\partial}{\partial y}F(x,y)$$

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

Marginal pdfs

Marginal pdfs are derived by integrating out one of the random variables.

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

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

Previously we defined independence in terms of, X and Y are independent if and only if $E(XY) = E(X)E(Y)$.

An equivalent definitions is, X and Y are independent if and only if $f(x,y) = f_X(x)f_Y(y)$.

Probability and Expectation

Univariate definition:

$$P(X \in A) = \int_A f(x)\;dx$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\cdot f(x)\;dx$$

Bivariate definition:

$$P(X \in A, Y \in B) = \int_A\int_B f(x,y)\;dx\;dy$$

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\cdot f(x,y)\;dx\;dy$$

Example 1 - Joint Uniforms

Let $X, Y \sim \text{Unif}(0,1)$, find $F(x,y)$.

Example 1, cont.

Using the $F(x, y)$ we just found, find $f(x, y)$

Example 1, cont.

Check that $f(x, y)$ produces the correct marginal densities for X and Y ($f_X(x)$ and $f_Y(y)$)

Example 1, cont.

Find the expected value of X and Y

Example 1, cont.

Find the expected value of XY

Example 1, another way

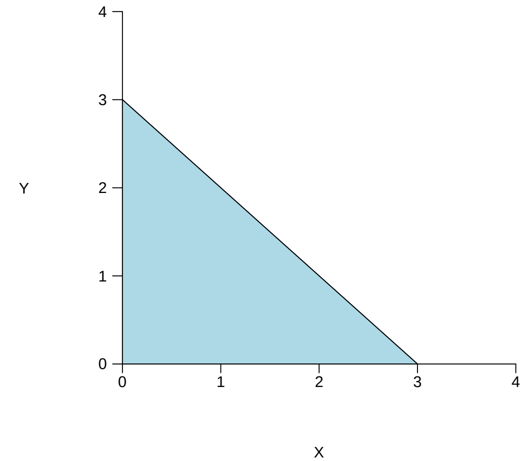
If we did not feel comfortable coming up with the graphical arguments for $F(x, y)$ we can also use the fact that the pdf is constant on $(0,1) \times (0,1)$ to derive the same distribution / density.

Let $f(x, y) = c$ for $x \in (0,1)$, $y \in (0,1)$ then

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\; dx\; dy \\ &= \int_0^1\int_0^1 c\; dx\; dy \\ &= \int_0^1 \left(cx|_0^1\right) dy = \int_0^1 c\; dy \\ &= cy|_0^1 = c \end{aligned}$$

Example 2

Let X and Y be drawn uniformly from the triangle below



Find the joint pdf, cdf, and marginals.

Example 2, cont.

Since the joint density is constant then

$$f(x, y) = c = \frac{2}{9}, \quad \text{for } 0 \le x + y \le 3$$

based on the area of the triangle, but we need to be careful to define on what range. We can define the range in two ways since X and Y depend on each other, so we can define the range of X in terms of Y or Y in terms of X.

$$\begin{aligned} f(x,y) &= \begin{cases} \frac{2}{9} & \text{if } y \in (0,3), x \in (0, 3-y) \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} \frac{2}{9} & \text{if } x \in (0,3), y \in (0, 3-x) \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Example 2, cont.

Depending on which range definition you choose it makes life easier when evaluating the marginal densities.

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)\; dy \\ &= \int_0^{3-x} \frac{2}{9}\; dy \\ &= \frac{2}{9}(3-x) \quad \text{for } x \in (0,3) \end{aligned}$$

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)\; dy \\ &= \int_0^{3-y} \frac{2}{9}\; dx \\ &= \frac{2}{9}(3-y) \quad \text{for } y \in (0,3) \end{aligned}$$

Example 2, cont.

Find the CDF

Example 3

Let $f(x,y) = cx^2y$ for $x^2 \le y \le 1$.

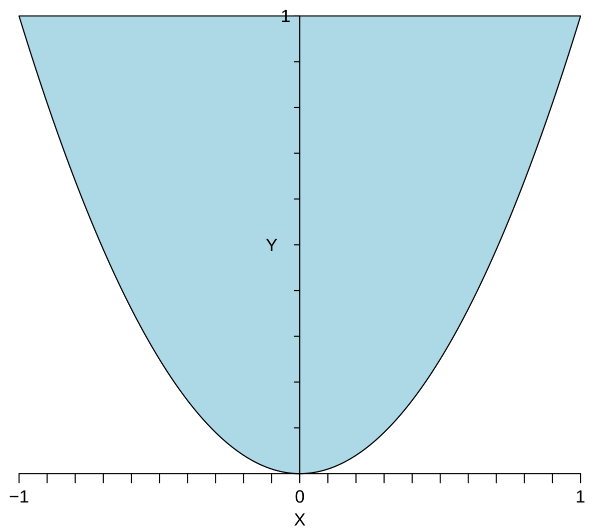
Find:

a) c

b) $P[X \ge Y]$

c) $f_X(x)$ and $f_Y(y)$

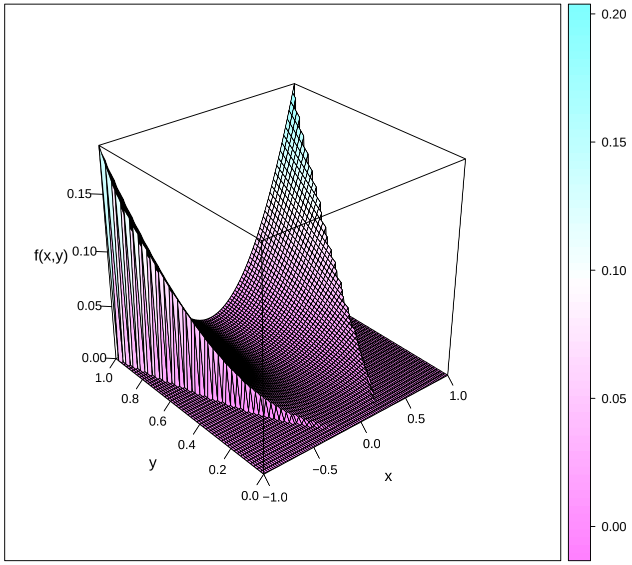
Example 3 - Range



Example 3.a

Find c

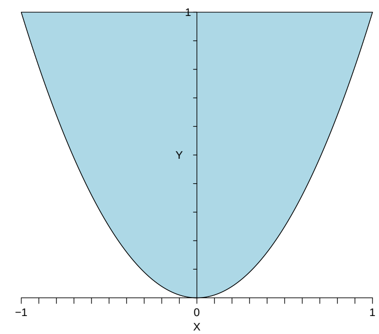
Example 3 - pdf

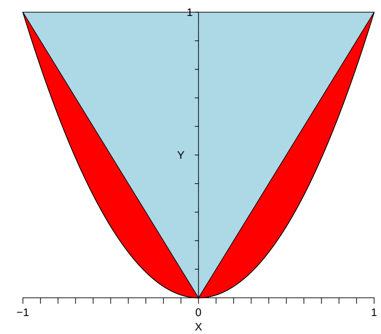


Example 3.b

Find $P(X \geq Y)$.

To do this we need to integrate over the region where $x^2 \leq y \leq 1$ and $x \geq y$ which is indicated in red below





Example 3.b, cont.

Example 3.c

Find the marginal densities

Example 3.c, cont.

It is always a good idea to check that the marginals are proper densities.

Example 4

Let Y be the rate of calls at a help desk, and X the number of calls between 2 pm and 4 pm one day; Let's say that:

$$f(x,y) = \frac{(2y)^x}{x!} e^{-3y}$$

for $y > 0$, $x = 0, 1, 2, \ldots$.

Find:

a) $P(X = 0)$

b) $P(Y > 2)$

Example 4.a

Find $P(X = 0)$

Example 4.b

Find $P(Y > 2)$