Lecture 10: Joint Distributions & Order Statistics

Sta 111

Colin Rundel

May 28, 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:

B \ W	0	1	2	
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

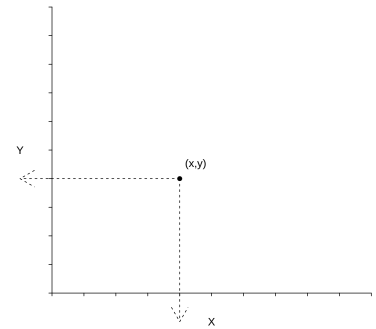
$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{\text{joint pmf}}{\text{marginal pmf}}$$

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

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



Joint CDF, cont.

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

Univariate definition:

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

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

Bivariate definition:

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

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

Marginal CDFs

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

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

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

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 $f(x,y) = c$ for $x \in (0,1)$, $y \in (0,1)$. Find c.

Example 1, cont

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

Example 1, cont.

Verify that the $F(x,y)$ we just found gives the correct $f(x,y)$,

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 2

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

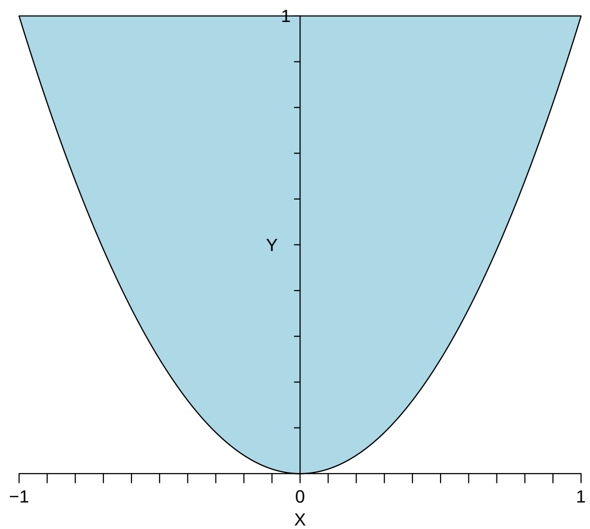
Find:

a) c

b) $P(|X| \ge Y)$

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

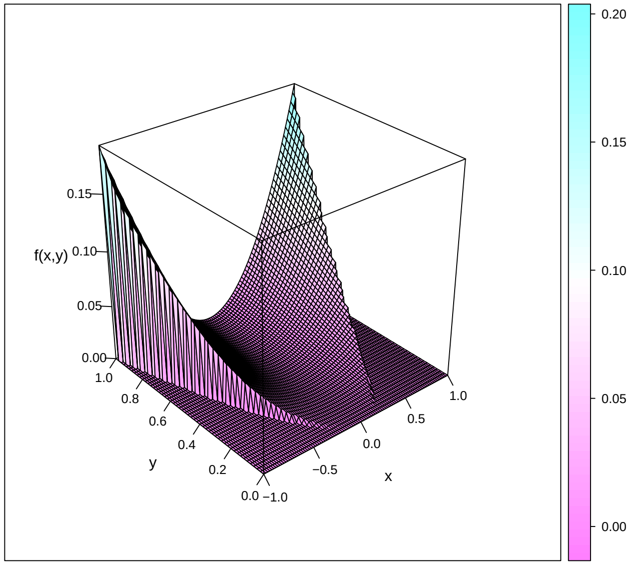
Example 2 - Range



Example 2.a

Find c

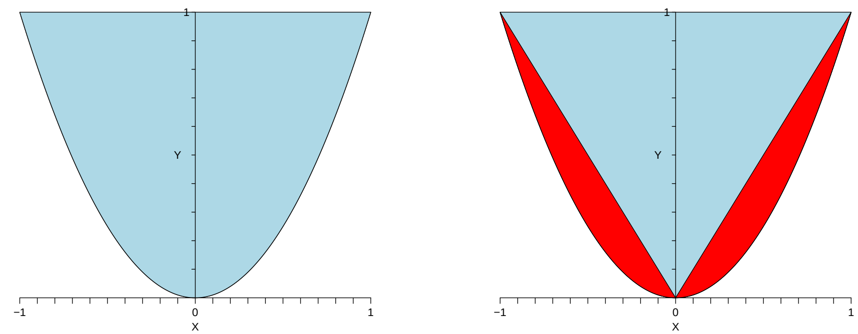
Example 2 - pdf



Example 2.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 2.b, cont.

Example 2.c

Find the marginal densities

Example 2.c, cont.

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

Example 3

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 3.a

Find $P(X = 0)$

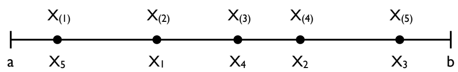
Example 3.b

Find $P(Y > 2)$

Order Statistics

Let X_1, X_2, X_3, X_4, X_5 be iid random variables with a distribution f on the range (a, b). We can relabel these X's such that their labels correspond to arranging them in increasing order

$$X_{(1)} \le X_{(2)} \le X_{(3)} \le X_{(4)} \le X_{(5)}.$$



In the case where the distribution f is continuous we can make the stronger statement

$$X_{(1)} < X_{(2)} < X_{(3)} < X_{(4)} < X_{(5)}.$$

Since $P(X_i = X_j) = 0$ for all $i \neq j$ for continuous random variables.

Order Statistics, cont.

For $X_1, X_2, \ldots, X_n$ iid random variables $X_{(k)}$ is the kth smallest X, usually called the kth order statistic.

The first order statistic, $X_{(1)}$ is therefore the smallest X_i and

$$X_{(1)} = \min(X_1, \ldots, X_n)$$

Similarly, the nth order statistic, $X_{(n)}$ is the largest X_i and

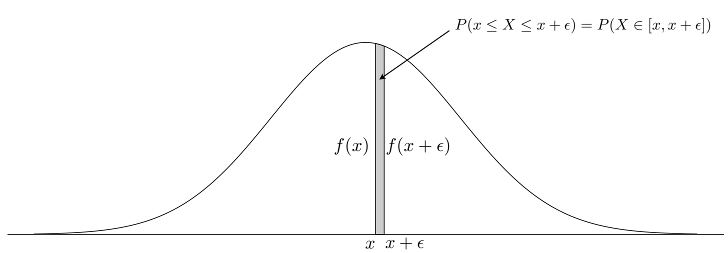
$$X_{(n)} = \max(X_1, \ldots, X_n)$$

Notational Detour

For a continuous random variable

$$f(x)\epsilon \approx P(x \le X \le x + \epsilon) = P(X \in [x, x + \epsilon])$$
$$f(x) \approx P(x \le X \le x + \epsilon) = P(X \in [x, x + \epsilon])/\epsilon$$

$$f(x) = \lim_{\epsilon \to 0} P(X \in [x, x + \epsilon])/\epsilon$$



Density of the maximum

For $X_1, X_2, \ldots, X_n$ iid continuous random variables with pdf f and cdf F the density of the maximum is

$$\begin{aligned}
P(X_{(n)} \in [x, x+\epsilon]) &= P(\text{one of the } X\text{'s} \in [x, x+\epsilon] \text{ and all others} < x) \\
&= \sum_{i=1}^{n} P(X_i \in [x, x+\epsilon] \text{ and all others} < x) \\
&= nP(X_1 \in [x, x+\epsilon] \text{ and all others} < x) \\
&= nP(X_1 \in [x, x+\epsilon])P(\text{all others} < x) \\
&= nP(X_1 \in [x, x+\epsilon])P(X_2 < x) \cdots P(X_n < x) \\
&= nf(x)\epsilon F(x)^{n-1}
\end{aligned}$$

$$f_{(n)}(x) = nf(x)F(x)^{n-1}$$

Density of the minimum

For $X_1, X_2, \ldots, X_n$ iid continuous random variables with pdf f and cdf F the density of the minimum is

$$\begin{aligned}
P(X_{(1)} \in [x, x+\epsilon]) &= P(\text{one of the } X\text{'s} \in [x, x+\epsilon] \text{ and all others} > x) \\
&= \sum_{i=1}^{n} P(X_i \in [x, x+\epsilon] \text{ and all others} > x) \\
&= nP(X_1 \in [x, x+\epsilon] \text{ and all others} > x) \\
&= nP(X_1 \in [x, x+\epsilon])P(\text{all others} > x) \\
&= nP(X_1 \in [x, x+\epsilon])P(X_2 > x) \cdots P(X_n > x) \\
&= nf(x)\epsilon(1 - F(x))^{n-1}
\end{aligned}$$

$$f_{(1)}(x) = nf(x)(1 - F(x))^{n-1}$$

Density of the kth Order Statistic

For $X_1, X_2, \ldots, X_n$ iid continuous random variables with pdf f and cdf F the density of the kth order statistic is

$$\begin{aligned}
P(X_{(k)} \in [x, x+\epsilon]) &= P(\text{one of the } X\text{'s} \in [x, x+\epsilon] \text{ and exactly } k-1 \text{ of the others} < x) \\
&= \sum_{i=1}^{n} P(X_i \in [x, x+\epsilon] \text{ and exactly } k-1 \text{ of the others} < x) \\
&= nP(X_1 \in [x, x+\epsilon] \text{ and exactly } k-1 \text{ of the others} < x) \\
&= nP(X_1 \in [x, x+\epsilon])P(\text{exactly } k-1 \text{ of the others} < x) \\
&= nP(X_1 \in [x, x+\epsilon]) \left(\binom{n-1}{k-1} P(X < x)^{k-1} P(X > x)^{n-k} \right) \quad = n
\end{aligned}$$

$$f_{(k)}(x) = nf(x)\binom{n-1}{k-1} F(x)^{k-1}(1 - F(x))^{n-k}$$

Cumulative Distribution of the min and max

For $X_1, X_2, \ldots, X_n$ iid continuous random variables with pdf f and cdf F the density of the kth order statistic is

$$\begin{aligned}
F_{(1)}(x) &= P(X_{(1)} < x) = 1 - P(X_{(1)} > x) \\
&= 1 - P(X_1 > x, \ldots, X_n > x) = 1 - P(X_1 > x) \cdots P(X_n > x) \\
&= 1 - (1 - F(x))^n
\end{aligned}$$

$$\begin{aligned}
F_{(n)}(x) &= P(X_{(n)} < x) = 1 - P(X_{(n)} > x) \\
&= P(X_1 < x, \ldots, X_n < x) = P(X_1 < x) \cdots P(X_n < x) \\
&= F(x)^n
\end{aligned}$$

$$f_{(1)}(x) = \frac{d}{dx}(1 - F(x))^n = n(1 - F(x))^{n-1}\frac{dF(x)}{dx} = nf(x)(1 - F(x))^{n-1}$$

$$f_{(n)}(x) = \frac{d}{dx}F(x)^n = nF(x)^{n-1}\frac{dF(x)}{dx} = nf(x)F(x)^{n-1}$$

Order Statistic of Standard Uniforms

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Unif}(0, 1)$ then the density of $X_{(n)}$ is given by

$$\begin{aligned}
f_{(k)}(x) &= nf(x)\binom{n-1}{k-1} F(x)^{k-1}(1 - F(x))^{n-k} \\
&= \begin{cases} n\binom{n-1}{k-1} x^{k-1}(1-x)^{n-k} & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

This is an example of the Beta distribution where $r = k$ and $s = n - k + 1$.

Beta Distribution

The Beta distribution is a continuous distribution defined on the range $(0,1)$ where the density is given by

$$f(x) = \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1}$$

where $B(r,s)$ is called the Beta function and it is a normalizing constant which ensures the density integrates to 1.

$$\begin{aligned}
1 &= \int_0^1 f(x)dx \\
1 &= \int_0^1 \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1}dx \\
1 &= \frac{1}{B(r,s)} \int_0^1 x^{r-1}(1-x)^{s-1}dx \\
B(r,s) &= \int_0^1 x^{r-1}(1-x)^{s-1}dx
\end{aligned}$$

Beta Function

The connection between the Beta distribution and the kth order statistic of n standard Uniform random variables allows us to simplify the Beta function.

$$\begin{aligned}
B(r,s) &= \int_0^1 x^{r-1}(1-x)^{s-1}dx \\
B(k, n-k+1) &= \frac{1}{n\binom{n-1}{k-1}} \\
&= \frac{(k-1)!(n-1-k+1)!}{n(n-1)!} \\
&= \frac{(r-1)!(n-k)!}{n!} \\
&= \frac{(r-1)!(s-1)!}{(r+s-1)!} = \frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)}
\end{aligned}$$

Beta Function - Expectation

Let $X \sim \text{Beta}(r,s)$ then

$$\begin{aligned}
E(X) &= \int_0^1 x \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1}dx \\
&= \frac{1}{B(r,s)} \int_0^1 x^{(r+1)-1}(1-x)^{s-1}dx \\
&= \frac{B(r+1,s)}{B(r,s)} \\
&= \frac{r!(s-1)!}{(r+s)!} \frac{(r+s-1)!}{(r-1)!(s-1)!} \\
&= \frac{r!}{(r-1)!} \frac{(r+s-1)!}{(r+s)!} \\
&= \frac{r}{r+s}
\end{aligned}$$

Beta Function - Variance

Let $X \sim \text{Beta}(r,s)$ then

$$\begin{aligned}
E(X^2) &= \int_0^1 x^2 \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1}dx \\
&= \frac{B(r+2,s)}{B(r,s)} = \frac{(r+1)!(s-1)!}{(r+s+1)!} \frac{(r+s-1)!}{(r-1)!(s-1)!} \\
&= \frac{(r+1)r}{(r+s+1)(r+s)}
\end{aligned}$$

$$\begin{aligned}
Var(X) &= E(X^2) - E(X)^2 \\
&= \frac{(r+1)r}{(r+s+1)(r+s)} - \frac{r^2}{(r+s)^2} \\
&= \frac{(r+1)r(r+s) - r^2(r+s+1)}{(r+s+1)(r+s)^2} \\
&= \frac{rs}{(r+s+1)(r+s)^2}
\end{aligned}$$

Beta Distribution - Summary

If $X \sim \text{Beta}(r, s)$ then

$$f(x) = \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1}$$
$$F(x) = \int_0^x \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1} dx = \frac{B_x(r,s)}{B(r,s)}$$

$$B(r,s) = \int_0^1 x^{r-1}(1-x)^{s-1} dx = \frac{(r-1)!(s-1)!}{(r+s-1)!} = \frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)}$$
$$B_x(r,s) = \int_0^x x^{r-1}(1-x)^{s-1} dx$$

$$E(X) = \frac{r}{r+s}$$
$$Var(X) = \frac{rs}{(r+s)^2(r+s+1)}$$