Jointly Distributed Random Variables

**Def:** Given a model, \((\Omega, P)\), random variables

\[ X, Y, Z, \ldots : \Omega \to \mathbb{R} \]

are said to be jointly distributed.

**Notation:** Then, for example,

\[ P[X \in A, Y \in B] = P(\omega : X(\omega) \in A \text{ and } Y(\omega) \in B) \]

and

\[ P[(X, Y) \in C] = P(\{\omega : (X(\omega), Y(\omega)) \in C\}) \]

for \( A, B \subseteq \mathbb{R} \) and \( C \subseteq \mathbb{R}^2 \).

---

**An Example**

If a committee of size four is selected at random from 5 Dems., 5 Inds., and 5 Reps, then

\[ X = \#\text{Dems}, \]

\[ Y = \#\text{Reps}, \]

are JDRVs for which

\[ P[X = 0, Y = 0] = \binom{5}{4} \binom{15}{4}, \]

\[ P[X = 1, Y = 1] = \binom{5}{1} \binom{7}{2} \binom{15}{4}, \]

\[ P[X = 2, Y = 2] = \binom{5}{2} \binom{13}{2} \binom{15}{4}, \]

and

\[ P[X = Y] = \frac{5 + 250 + 100}{1365} = .260. \]

---

**The Joint Probability Mass Function**

**Two RVs:** If \( X \) and \( Y \) are JD discrete RVs, then their joint probability mass function is

\[ f(x, y) = P[X = x, Y = y] \]

for \( x, y \in \mathbb{R} \).

**Several:** If \( X_1, \ldots, X_m \) are JD discrete RVs, then their joint probability mass function is

\[ f(x_1, \ldots, x_m) = P[X_1 = x_1, \ldots, X_m = x_m] \]

for \( x_1, \ldots, x_m \in \mathbb{R} \).

**Vector Notation:** Let \( \mathbf{X} = (X_1, \ldots, X_m) \) and

\[ f(\mathbf{x}) = P[\mathbf{X} = \mathbf{x}] \]

for \( \mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m \).

---

**Bivariate Hypergeometric Distributions**

Parameters \( N, R, W, \) and \( n \)

If \( n \) tickets are drawn at random \( w.o.r. \) from \( N \) tickets of which \( R \) are red and \( W \) are white, then

\[ X = \#\text{red}, \]

\[ Y = \#\text{white}, \]

are JD discrete RVs with joint PMF

\[ f(x,y) = \binom{R}{x} \binom{W}{y} \binom{N - R - W}{n - x - y} / \binom{N}{n} \]

for integers \( x, y \geq 0 \) with \( x + y \leq n \) and \( f(x,y) = 0 \) for other values of \( x, y \).

**Example:** Committees. \( N = 15, R = 5, W = 5, \) and \( n = 4. \)

**Note:** Extensions to Several Colors.
Partitions

Review

If \( n \geq 1 \) and \( n_1, \ldots, n_m \geq 0 \) are integers for which
\[
n_1 + \cdots + n_m = n,
\]
then a set of \( n \) elements may be partitioned into \( m \) subsets of sizes \( n_1, \ldots, n_m \) in
\[
\binom{n}{n_1, \ldots, n_m} = \frac{n!}{n_1! \cdots n_m!}
\]
ways.

Example: MISSISSIPPI

\[
\binom{11}{4,1,2,4} = \frac{11!}{4! \cdot 2! \cdot 4!} = 34650.
\]

Example

If a balanced (6-sided) die is rolled 12 times, then the probability that each face appears twice is
\[
\frac{12!}{(1^6)(6!)} = \frac{12!}{2^6} = 0.0034.
\]

For an outcome is
\[
\omega = (i_1, \ldots, i_{12}),
\]
where \( 1 \leq i_1, \ldots, i_{12} \leq 6 \); there are
\[
\# \Omega = 6^{12}
\]
such outcomes on
\[
\binom{12}{2,2,2,2,2,2} = \frac{12!}{2^6} = \frac{12!}{2^6}
\]
of which each face appears twice.

Multinomial Distributions

A Loaded Die: Now consider an \( m \)-sided, loaded die. Let
\[
p_i = \text{Prob}[i \text{ spots}]
\]
on a single role. So,
\[
p_1, \ldots, p_m \geq 0,
\]
\[
p_1 + \cdots + p_m = 1.
\]

Repeated Trials: Suppose that the die is rolled \( n \) times, and let
\[
X_i = \# \text{roles with } i \text{ spots}
\]
for \( i = 1, \ldots, m \). Then
\[
f(x_1, \ldots, x_m) = P[X_1 = x_1, \ldots, X_m = x_m]
\]
is
\[
\binom{n}{x_1, \ldots, x_m} p_1^{x_1} \cdots p_m^{x_m}
\]
for integers \( x_1, \ldots, x_m \geq 0 \) \((*)\)
with
\[
x_1 + \cdots + x_m = n.
\]

For the probability of any sequence with \( x_i \)’s is
\[
p_1^{x_1} \cdots p_m^{x_m}.
\]
and there are
\[
\binom{n}{x_1, \ldots, x_m}
\]
such sequences.

Thus
\[
f(x_1, \ldots, x_m) = \binom{n}{x_1, \ldots, x_m} p_1^{x_1} \cdots p_m^{x_m}
\]
if \((*)\) holds and \( f(x_1, \ldots, x_m) = 0 \) otherwise.

Def: Called multinomial with parameters \( n, m, \) and \( p_1, \ldots, p_m \).
Properties of Multivariate PMFs

If \( f \) is the joint PMF of \( X_1, \ldots, X_m \), then there is a finite or countably infinite

\[ \mathcal{X} \subseteq \mathbb{R}^m \]

for which

\[ f(x) \geq 0, \text{ for all } x \tag{1} \]

and

\[ f(x) = 0 \text{ if } x \notin \mathcal{X}, \tag{2} \]

and

\[ \sum_{x \in \mathcal{X}} f(x) = 1. \tag{3} \]

Also,

\[ P[X \in B] = \sum_{x \in B \cap \mathcal{X}} f(x) \]

for \( B \subseteq \mathbb{R}^m \).

Conversely, any \( f \) that satisfies (1), (2), and (3), is the joint PMF of some random variables \( X_1, \ldots, X_m \).

Marginal Distributions

Two Variables

Let \( X \) and \( Y \) by JD discrete RVs with joint PMF

\[ f(x, y) = P[X = x, Y = y] \]

and ranges \( \mathcal{X} \) and \( \mathcal{Y} \). So, \( f(x, y) = 0 \) unless \( x \in \mathcal{X} \) and \( y \in \mathcal{Y} \). Then \( X \) and \( Y \) have individual (marginal) PMFs

\[ f_X(x) = \sum_{y \in \mathcal{Y}} f(x, y), \]

\[ f_Y(y) = \sum_{x \in \mathcal{X}} f(x, y). \]

For

\[ \{ X = x \} = \bigcup_{y \in \mathcal{Y}} \{ X = x, Y = y \} \]

and, therefore,

\[ P[X = x] = \sum_{y \in \mathcal{Y}} P[X = x, Y = y]. \]

Example

Two tickets are drawn w.o.r. from a box with

1 ticket labelled one,
2 tickets labelled two,
3 ticket labelled three,

Let

\( X = \) label on first ticket,
\( Y = \) label on second.

Then

Table of \( f(x, y) \)

<table>
<thead>
<tr>
<th>( x, y )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>( f_X(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/5</td>
<td>2/5</td>
<td>3/5</td>
<td>1/3</td>
</tr>
<tr>
<td>2</td>
<td>3/50</td>
<td>6/50</td>
<td>9/50</td>
<td>6/5</td>
</tr>
<tr>
<td>3</td>
<td>8/50</td>
<td>12/50</td>
<td>18/50</td>
<td>3/5</td>
</tr>
</tbody>
</table>

\( f_Y(y) \)

\[ \begin{array}{c}
\frac{1}{5} \\
\frac{2}{5} \\
\frac{3}{5}
\end{array} \]

Marginal Distributions

Several Variables

Let

\( X = (X_1, \ldots, X_i) \)

and

\( Y = (Y_1, \ldots, Y_k) \)

be JD discrete RVs with joint PMF

\[ f(x, y) = P[X = x, Y = y] \]

and ranges \( \mathcal{X} = \mathbf{X}(\Omega) \) and \( \mathcal{Y} = \mathbf{Y}(\Omega) \). Then \( X \) and \( Y \) have individual (marginal) joint PMFs

\[ f_X(x) = \sum_{y \in \mathcal{Y}} f(x, y), \]

and

\[ f_Y(y) = \sum_{x \in \mathcal{X}} f(x, y). \]
Example
Multinomial Distributions

If

\((X_1, \cdots, X_m) \sim \text{Multinomial}(n, \mathbf{p})\),

and \(1 \leq k < n\), then the distribution of

\((X_1, \cdots, X_k, X_{k+1} + \cdots + X_m)\)

is

\(\text{Multinomial}(n, k + 1, \mathbf{p}_1, \cdots, \mathbf{p}_k, \mathbf{p}_{k+1}, \cdots, \mathbf{p}_m)\).

In particular,

\(X_i \sim \text{Binomial}(n, p_i)\).

Slide 13

Multiple Integrals

An Interlude

If

\(-\infty < a < b < \infty,\)

\(-\infty < c < d < \infty,\)

and

\(f : [a, b] \times [c, d] \to \mathbb{R},\)

is sufficiently nice (e.g. continuous), then

\[ \int_a^b \int_c^d f(x, y) \, dy \, dx \]

is the limit of Riemann sums.

Extensions: Under mild technical conditions,

\(a, c \to -\infty,\)

\(b, d \to \infty.\)

Other Regions: Under conditions,

\[ \int \int_C f(x, y) \, dy \, dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \mathbf{1}_C(x, y) \, dy \, dx. \]

Slide 14

Example:

\[ \int \int_C 1 \, dy \, dx = \text{Area}(C). \]

Slide 15

Reduction to Iterated Integration

Theorem: If

\(C = \{(x, y) : a \leq x \leq b, \ c(x) \leq y \leq d(x)\},\)

then

\[ \int \int_C f(x, y) \, dy \, dx = \int_a^b \left[ \int_{c(x)}^{d(x)} f(x, y) \, dy \right] \, dx, \]

under conditions.

Corollary

\[ \int_a^b \int_c^d g(x) h(y) \, dy \, dx = \left[ \int_a^b g(x) \, dx \right] \int_c^d h(y) \, dy. \]

Example: If \(T = \{(x, y) : 0 \leq x \leq 1, \ 0 \leq y \leq x\},\)

then

\[ \int \int_T 1 \, dy \, dx = \int_0^1 \left[ \int_0^x dy \right] \, dx \]

\[ = \int_0^1 x \, dx \]

\[ = \frac{1}{2}. \]
Higher Dimensions

Write \[ \int_{\mathbb{R}^m} f(x)dx \]
for
\[ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \cdots, x_m)dx_m \cdots dx_1. \]

Bivariate Densities

A function \( f : \mathbb{R}^2 \to \mathbb{R} \)

is a bivariate density if
\( f(x, y) \geq 0, \)
and
\[ \int \int_{\mathbb{R}^2} f(x, y)dydx = 1. \]

If \( f \) is a density, then JDRV X and Y have joint density \( f \)
\[ P[(X, Y) \in C] = \int \int_C f(x, y)dydx \]
for nice subsets \( C \subseteq \mathbb{R}^2. \)

Example: Uniform Distributions. If \( R \subset \mathbb{R}^2 \) and
\( 0 < \alpha = \text{Area}(R) < \infty, \) then
\[ f(x, y) = \frac{1}{\alpha} 1_{R}(x, y) \]
is a density, called the uniform density over \( R. \)

Marginal Densities

If \( X \) and \( Y \) have joint density \( f, \) then \( X \) and \( Y \)
have individual (marginal) densities
\[ f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy, \]
\[ f_Y(y) = \int_{-\infty}^{\infty} f(x, y)dx. \]

Example: Let
\[ D = \{(x, y) : x^2 + y^2 \leq 1\} \]
and
\[ f(x, y) = \frac{1}{\pi} 1_D(x, y). \]
If \(-1 < x < 1, \) then
\[ f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} dy = \frac{2}{\pi} \sqrt{1-x^2}. \]

Example

If
\[ f(x, y) = \frac{2}{(1 + x + y)^3} \]
for \( 0 \leq x, y < \infty \) and \( f(x, y) = 0 \) otherwise, then
\[ \int_0^\infty f(x, y)dy = \frac{1}{(1 + x + y)^{2\gamma}} \bigg|_{y=0}^{\infty} \]
\[ = \frac{1}{(1 + x)^2} \]
for \( 0 \leq x < \infty, \) and
\[ \int \int_{\mathbb{R}^2} f(x, y)dydx = \int_0^\infty \frac{dx}{(1 + x)^2} \bigg|_{x=0}^{\infty} \]
\[ = \frac{1}{1 + x} \bigg|_{x=0}^{\infty} \]
\[ = 1. \]
So, \( f \) is a density and
\[ f_X(x) = \frac{1}{(1 + x)^2} \]
for \( 0 \leq x < \infty. \)
**Multivariate Densities**

A function

\[ f : \mathbb{R}^m \rightarrow \mathbb{R} \]

is a multivariate density if

\[ f(x) \geq 0, \]

and

\[ \int_{\mathbb{R}^m} f(x)dx = 1. \]

If \( f \) is a density, then JDRVs \( X_1, \cdots, X_m \) have joint density \( f \)

\[ P[X \in C] = \int_C f(x)dx \]

for nice subsets \( C \subseteq \mathbb{R}^m \).

**Marginal Densities**: If \( X_1, \cdots, X_j \) and \( Y_1, \cdots, Y_k \) have joint density \( f \), then \( X_1, \cdots, X_j \) have joint density

\[ f_{X}(x) = \int_{\mathbb{R}^k} f(x,y)dy. \]

---

**Joint Distribution Functions**

**Def**: If \( X \) and \( Y \) are JDRVs, then their joint distribution function is

\[ F(a,b) = P[X \leq a, Y \leq b]. \]

**Marginal Distributions**: Then

\[ F_X(a) = \lim_{b \to \infty} F(a,b), \]

\[ F_Y(b) = \lim_{a \to -\infty} F(a,b), \]

**Notes**: a) Characteristic Properties

b) Higher Dimensions
c) Harder to Use
d) Mixed Distributions

---

**Independence**

JDRVs \( X \) and \( Y \) are independent if

\[ P[X \in A, Y \in B] = P[X \in A]P[Y \in B] \]

for all nice subsets \( A, B \subseteq \mathbb{R} \) (for example, intervals).

**Conditions for Independence**

**PMF**: If \( X \) and \( Y \) are discrete, then \( X \) and \( Y \) are independent iff

\[ f(x,y) = f_X(x)f_Y(y) \quad (*) \]

for all \( x \) and \( y \). For if \( X \) and \( Y \) are independent, then \( X = x \) iff \( X \in [x, x] \), so that

\[ f(x,y) = P[X = x, Y = y] = P[X = x]P[Y = y] = f_X(x)f_Y(y). \]

Conversely, if \( (*) \) holds, then

\[ P[X \in A, Y \in B] = \sum_{x \in A \cap X} \sum_{y \in B \cap Y} f(x,y) = \sum_{x \in A \cap X} f_X(x)f_Y(y) = [\sum_{x \in A \cap X} f(x,y)][\sum_{y \in B \cap Y} f_Y(y)] = P[X \in A]P[Y \in B], \]

where \( X \) and \( Y \) are the ranges of \( X \) and \( Y \).

**Example**: If \( E \) and \( F \) are independent events, then \( 1_E \) and \( 1_F \) are independent random variables. For example,

\[ P[1_E = 1, 1_F = 1] = P(E \cap F) = P(E)P(F) = P[1_E = 1]P[1_F = 1] \]
Other Conditions

**DFs:** $X$ and $Y$ are independent iff

$$F(a, b) = F_X(a)F_Y(b)$$

for all $a, b \in \mathbb{R}$.

**Densities:** If $X$ and $Y$ have individual densities $f_X$ and $f_Y$, then $X$ and $Y$ are independent iff $X$ and $Y$ have joint density

$$f(x, y) = f_X(x)f_Y(y).$$

---

**Example**

If

$$X \sim \text{Exp}(\lambda),$$

$$Y \sim \text{Exp}(\lambda)$$

are independent, what is

$$P[Y \geq 2X \text{ or } X \geq 2Y].$$

Here

$$f_X(x) = f_Y(y) = \lambda e^{-\lambda x}$$

for $0 \leq z < \infty$. So,

$$f(x, y) = f_X(x)f_Y(y) = \lambda^2 e^{-\lambda(x+y)}$$

for $0 \leq x, y < \infty$ and $f(x, y) = 0$ for other $x$ and $y$. So,

---

**Several Variables**

$X_1, \ldots, X_m$ are independent if

$$P[X_1 \in A_1, \ldots, X_m \in A_m] = P[X_1 \in A_1] \times \cdots \times P[X_m \in A_m]$$

**Note:** Equivalent Conditions. For example,

$$f(x_1, \ldots, x_m) = f_1(x_1) \times \cdots \times f_m(x_m).$$
The Distribution of the Maximum

Example: n light globes are placed in service at time \( t = 0 \) and allowed to burn continuously. Denote their lifetimes by \( X_1, \ldots, X_n \) and suppose that
\[
X_1, \ldots, X_n \sim \text{ind } F. \tag{*}
\]
If burned out globes are not replaced, then the room goes dark at time
\[
Y = \max[X_1, \ldots, X_n],
\]
the largest of \( X_1, \ldots, X_n \).

The Distribution of \( Y \): If (*) holds, then
\[
G(y) := P[Y \leq y] = P[X_1 \leq y, \ldots, X_n \leq y] = P[X_1 \leq y] \times \cdots \times P[X_n \leq y] = F(y)^n = F(y)^n.
\]
So, if \( F \) has density \( f \), then \( Y \) has density
\[
g(y) = \frac{d}{dy} F(y)^n = nF(y)^{n-1} f(y).
\]

Example: Revisited. If \( n = 5 \) and \( F \) is exponential with \( \lambda = 1 \) per hour, then
\[
F(t) = 1 - e^{-t},
\]
\[
G(t) = (1 - e^{-t})^5,
\]
and
\[
g(t) = 5(1 - e^{-t})^4 e^{-t}
\]
for \( 0 \leq t < \infty \). The probability that the room is still lighted after two months is
\[
P[Y \geq 2] = 1 - G(2) = 1 - (1 - e^{-2})^5 = .5167.
\]

Order Statistics

If
\[
X_1, \ldots, X_n \sim \text{ind } F,
\]
let
\[
X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}
\]
by \( X_1, \ldots, X_n \) in increasing order. Thus,
\[
X_{(1)} = \min[X_1, \ldots, X_n],
\]
\[
\ldots,
\]
\[
X_{(n)} = \max[X_1, \ldots, X_n].
\]

Notes
- Times that globes burn out—in the example.
- Can find distributions.
- Section 6.6 and Problems 9 and 10.

Sums of Independent Random Variables

Convolutions: Let \( X \) and \( Y \) be independent, integer valued random variables, and let
\[
Z = X + Y.
\]
Then
\[
f_Z(k) = \sum_{j=-\infty}^{\infty} f_X(j)f_Y(k-j). \tag{*}
\]
For
\[
\{Z = k\} = \bigcup_{j=-\infty}^{\infty} \{X = j, Y = k-j\}
\]
and, therefore,
\[
P[Z = k] = \sum_{j=-\infty}^{\infty} P[X = j]P[Y = k-j].
\]

Note: (*) is called the convolution of \( f_X \) and \( f_Y \).
Example

Poisson

If $X \sim \text{Poisson}(\alpha)$ and $Y \sim \text{Poisson}(\beta)$ are independent, then

$$f_X(j) = \frac{1}{j!} \alpha^j e^{-\alpha},$$
$$f_Y(j) = \frac{1}{j!} \beta^j e^{-\beta},$$

for $j = 0, 1, 2, \cdots$, and $f_X(j) = f_Y(j) = 0$ for $j < 0$. So, the PMF of $Z = X + Y$ is

$$f_Z(k) = \sum_{j=0}^{\infty} f_X(j)f_Y(k-j)$$
$$= \sum_{j=0}^{k} \frac{1}{j!} \alpha^j e^{-\alpha} \times \frac{1}{(k-j)!} \beta^{k-j} e^{-\beta}$$
$$= \frac{1}{k!} e^{-\alpha-\beta} \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} \alpha^j \beta^{k-j}$$
$$= \frac{1}{k!}(\alpha + \beta)^k e^{-\alpha-\beta}$$

So, if

$$X \sim \text{Poisson}(\alpha),$$
$$Y \sim \text{Poisson}(\beta),$$

are independent, then

$$X + Y \sim \text{Poisson}(\alpha + \beta).$$

By induction, if

$$X_i \sim \text{Poisson}(\lambda_i), \quad i = 1, \cdots, n,$$

are independent, then

$$X_1 + \cdots + X_n \sim \text{Poisson}(\lambda_1 + \cdots + \lambda_n).$$

Convolution

The Continuous Case

Let $X$ and $Y$ are independent with densities $f_X$ and $f_Y$, and let

$$Z = X + Y.$$ 

Then

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx.$$

Example: If $X \sim \text{Unif}[0,1]$ and $Y \sim \text{Unif}[0,1]$, then

$$f_Z(z) = \min[z, 2-z]$$

for $0 \leq z \leq 2$ and $f_Z(z) = 0$ otherwise.

In this case

$$f_X(z) = f_Y(z) = 1 \text{ for } 0 \leq z \leq 1,$$
$$f_X(z) = f_Y(z) = 0 \text{ otherwise}.$$

So, if $0 \leq z \leq 1$, for example, then

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx$$
$$= \int_{0}^{z} 1 \times 1 dx$$
$$= z.$$
Similarly

Let

\[ X_1, \ldots, X_n \] be independent

and

\[ Y = X_1 + \cdots + X_n. \]

If

\[ X_i \sim \text{Gamma}(\alpha_i, \beta), \quad i = 1, \ldots, n, \]

then

\[ Y \sim \text{Gamma}(\alpha_1 + \cdots + \alpha_n, \beta). \]

If

\[ X_i \sim \text{Normal}(\mu_i, \sigma_i^2), \quad i = 1, \ldots, n, \]

then

\[ Y \sim \text{Normal}(\mu, \sigma^2), \]

where

\[ \mu = \mu_1 + \cdots + \mu_n, \]

\[ \sigma^2 = \sigma_1^2 + \cdots + \sigma_n^2. \]

Conditional Distributions

The Discrete Case

Let \( X \) and \( Y \) have joint PMF \( f \). If \( f_X(x) > 0 \), then the conditional PMF of \( Y \) given \( X \) is

\[ f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}. \]

Thus,

\[ f_{Y|X}(y|x) = \frac{P[X = x, Y = y]}{P[X = x]} = \frac{P[Y = y|X = x]}{P[X = x]}. \]

Note: \( f_{Y|X} \) is a PMF, since

\[ \sum_{y \in \mathcal{Y}} f_{Y|X}(y|x) = \frac{1}{f_X(x)} \sum_{y \in \mathcal{Y}} f(x,y) = 1. \]

Note: Can reverse the roles of \( X \) and \( Y \).

Example

Box Ticket Models

If

\[ f(x,y) = \binom{R}{x} \binom{W}{y} \binom{N - R - W}{n - x - y} \binom{N}{n}, \]

for \( x + y \leq n \), where \( n \leq N \), \( R, W \geq 1 \) and \( R + W < N \), then

\[ f_X(x) = \binom{R}{x} \binom{N - R}{n - x} \binom{N}{n} \]

and

\[ f_{Y|X}(y|x) = \binom{W}{y} \binom{N - R - W}{n - x - y} \binom{N - R}{n - x}. \]

Note: Intuitive.

Conditional Distributions

The Continuous Case

Conditional Densities: Let \( X \) and \( Y \) have joint density \( f \). If \( f_X(x) > 0 \), then the conditional density of \( Y \) given \( X \) is

\[ f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}. \]

As above, this is a density.

Conditional Probability: Write

\[ P[Y \in B|X = x] = \int_B f_{Y|X}(y|x)dy. \]

Notes:

- New definition.
- \( P[X = x] = 0 \).
- Can reverse the roles of \( X \) and \( Y \).
Example

If
\[ f(x, y) = \frac{2}{(1 + x + y)^2} \]
for \(0 \leq x, y < \infty\), then
\[ f_X(x) = \frac{1}{(1 + x)^2} \]
for \(0 \leq x < \infty\). So,
\[ f_{Y|X}(y|x) = \frac{2(1 + x)^2}{(1 + x + y)^2} \]
and
\[ P[Y > x|X = x] = \int_x^\infty \frac{2(1 + x)^2}{(1 + x + y)^2} \, dy \]
\[ = -\frac{(1 + x)^2}{(1 + x + y)^2} \bigg|_x^\infty \]
\[ = \frac{(1 + x)^2}{(1 + x + c)^2}. \]

Bayes Theorem

In both cases (discrete and continuous),
\[ f(x, y) = f_{Y|X}(y|x)f_X(x), \]
if \(f_X(x) > 0\). In the discrete case,
\[ f_Y(y) = \sum_{x \in X} f(x, y) \quad (\ast) \]
and
\[ f_{X|Y}(x|y) = f(x, y)f_Y(y), \]
when \(f_Y(y) > 0\). In the continuous case, the sum in \((\ast)\) is replaced by an integral.

Mixed Distributions: One variable can be discrete and the other continuous.

Multivariate Extensions: \(X\) and/or \(Y\) can be vectors.

The Rule of Succession

Suppose
\[ X \sim \text{Unif}[0, 1], \]
and
\[ f_Y(y_1, \ldots, y_n|X) = x^{y_1 + \cdots + y_n}(1 - x)^{n-(y_1 + \cdots + y_n)} \]
for \(y_1, \ldots, y_n = 0\) or 1. Then
\[ P[Y_1 = 1, \ldots, Y_n = 1] = f_Y(1, \ldots, 1) \]
\[ = \int_0^1 x^n \, dx 
\[ = \frac{1}{n+1}. \]

If there were an \((n + 1)\)st \(Y_i\), then
\[ P[Y_{n+1} = 1|Y_1 = 1, \ldots, Y_n = 1] = \frac{P[Y_1 = 1, \ldots, Y_{n+1} = 1]}{P[Y_1 = 1, \ldots, Y_n = 1]} 
\[ = \frac{1/(n + 2)}{1/(n + 1)} 
\[ = \frac{n + 1}{n + 2}. \]

Note: Depends (crucially) on the distribution of \(X\).
**Transformations**

Let

$$m \geq 1,$$

$$D \subseteq \mathbb{R}^m,$$

and let

$$w : D \rightarrow E \subseteq \mathbb{R}^m$$

be surjective (one-to-one). Write

$$y = w(x)$$

as

$$y_1 = w_1(x_1, \ldots, x_m),$$

$$y_2 = w_2(x_1, \ldots, x_m),$$

$$\ldots,$$

$$y_m = w_m(x_1, \ldots, x_m),$$

**Jacobianns**

Assuming that $w$ is differentiable, let

$$J_w(x) = |\det \begin{pmatrix} \frac{\partial w_1(x)}{\partial x_1} & \cdots & \frac{\partial w_1(x)}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial w_m(x)}{\partial x_1} & \cdots & \frac{\partial w_m(x)}{\partial x_m} \end{pmatrix}|,$$

where

$$x = (x_1, \ldots, x_m).$$

Or, briefly,

$$J(x) = |\det \frac{\partial w(x)}{\partial x}|.$$

**Polar Coordinates**

Let

$$m = 2,$$

$$D = (0, \infty) \times [-\pi, \pi),$$

$$y_1 = r \cos(\theta),$$

$$y_2 = r \sin(\theta).$$

Then

$$w : D \rightarrow \mathbb{R}^2 \setminus \{0\}$$

$$\frac{\partial w_1}{\partial r} = \cos(\theta),$$

$$\frac{\partial w_1}{\partial \theta} = r \sin(\theta),$$

$$\frac{\partial w_1}{\partial \theta} = -r \sin(\theta),$$

$$\frac{\partial w_2}{\partial \theta} = r \sin(\theta),$$

So,

$$J(r, \theta) = |\det \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\theta \sin(\theta) & \theta \cos(\theta) \end{pmatrix}|.$$ 

That is,

$$J(r, \theta) = r \cos^2(\theta) + r \sin^2(\theta) = r.$$
Change of Variables

Suppose that
\[ w : D \rightarrow \mathbb{R}^n \]
is differentiable and one-to-one and let \( v : E \rightarrow D \)
be the inverse function.

**Theorem.** Under technical conditions,
\[
\int_D g(x) \, dx = \int_E g[v(y)] |J_v(y)| \, dy.
\]

**Corollary.**
\[
\int_{\mathbb{R}^2} g(x_1, x_2) \, dx_1 \, dx_2 = \int_0^\infty \int_0^{\pi} g[r \cos(\theta), r \sin(\theta)] r \, d\theta \, dr.
\]

---

**Transformations of RV's**

Now let \( X_1, \ldots, X_m \) have joint density \( f \).

Let \( \mathbf{X} = (X_1, \ldots, X_m) \).

Suppose \( P[\mathbf{X} \in D] = 1 \) and let
\[ \mathbf{Y} = w(\mathbf{X}), \]
where \( w : D \rightarrow \mathbb{R}^m \).

That is,
\[
Y_1 = w_1(X_1, \ldots, X_m), \quad Y_2 = w_2(X_1, \ldots, X_m), \quad \ldots, \quad Y_m = w_m(X_1, \ldots, X_m),
\]

---

**Example**

If \( X_1, X_2 \sim \text{ind } \Phi \),
then
\[
f(x_1, x_2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x_1^2}{2}} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{x_2^2}{2}}
\]
\[= \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2 + x_2^2)}.
\]

Then,
\[X_1 = R \cos(\Theta), \quad X_2 = R \sin(\Theta), \]
then
\[g(r, \theta) = f[r \cos(\theta), r \sin(\theta)] r
\]
for \( 0 < r < \infty \) and \( -\pi < \theta \leq \pi \).
So,

\[ g_1(r) = \int_{-\pi}^{\pi} \frac{r}{2\pi} e^{-\frac{1}{2}r^2} d\theta \]
\[ = re^{-\frac{1}{2}r^2} \]

for \(0 < r < \infty\); and

\[ g_2(\theta) = \int_{0}^{\infty} \frac{r}{2\pi} e^{-\frac{1}{2}r^2} dr \]
\[ = -\frac{1}{2\pi} e^{-\frac{1}{2}r^2} \bigg|_{r=0}^{r=\infty} \]
\[ = re^{-\frac{1}{2}r^2} ; \]

and

\[ R \text{ and } \Theta \text{ are independent.} \]