Multivariate Probability #1
ECON 670

Justin L. Tobias
Purdue University
- Preliminaries
- Marginal & Conditional Distributions
- Covariance & Correlation
- Independence
- Multivariate Distributions
Previously, we have discussed probability models and computation of probabilities for events involving only one random variable.

These are called univariate models.

Here we will talk about probability models that involve more than one random variable, called multivariate models.

**Definition**

An $n$-dimensional random vector is a function from a sample space $S$ into $\mathbb{R}^n$, the $n$-dimensional Euclidean space.
Example
A coin is tossed 3 times. Let $X_1 =$ the number of heads on the first 2 tosses, $X_2 =$ number of heads on all 3 tosses. What is the space of $(X_1, X_2)$?

In this case, the sample space is

$$S = \{TTT, TTH, THT, HTT, THH, HTH, HHT, HHH\}$$

and the space of $(X_1, X_2)$, denoted $D$ is
The joint cdf of \((X_1, X_2)\) is defined as

Since \(X_1\) and \(X_2\) are random variables, each of the events in the above intersection and the intersection of the events are events in the original sample space \(S\). Thus the expression is well defined.

We will write

\[
\Pr[\{X_1 \leq x_1\} \cap \{X_2 \leq x_2\}] \equiv \Pr[X_1 \leq x_1, X_2 \leq x_2].
\]
Result

For the two-dimensional case (i.e., bivariate cdf), we have:

- 

Definition

A random vector \((X_1, X_2)\) is **discrete** if its space \(D\) is finite or countable. The joint (bivariate) pmf of \((X_1, X_2)\) is defined by

\[
p_{X_1,X_2}(x_1, x_2) = \Pr[X_1 = x_1, X_2 = x_2] \quad \forall (x_1, x_2) \in D
\]
Properties of bivariate pmf:

1. 

2. 

Note that these are natural generalizations of what we learned in the univariate case: probabilities of events are non-negative and sum to unity.
For any event $B \in D$, we have $\Pr[(X_1, X_2) \in B] = \sum \sum_{B} p_{X_1, X_2}(x_1, x_2)$.

**Definition**

The **support** of $(X_1, X_2)$ includes all points $(x_1, x_2)$ such that $p_{X_1, X_2}(x_1, x_2) > 0$.

**Definition**

A random vector $(X_1, X_2)$ with space $D$ is of the **continuous** type if its joint cdf $F_{X_1, X_2}(x_1, x_2)$ is continuous.
Definition

The function $f_{X_1,X_2}(x_1, x_2)$ is a joint pdf of $(X_1, X_2)$ if it satisfies

1.

2.

Similarly, for any event $A \in D$,
Given a bivariate density $f$, the bivariate cdf can be obtained as:

It then follows (via Leibniz’ rule for differentiating under the integral sign:)}
Example (Bivariate #1)

(a) Verify that this is a valid (proper) density function.
(b) Derive the bivariate cdf
(c) Find Pr(0 < X_1 < 3/4, 1/3 < X_2 < 2).
As for this being a valid density, first note, for $0 < x_1 < 1$, $0 < x_2 < 1$,

$$f(x_1, x_2) > 0.$$ 

In addition,
As for the cdf, note that for $0 < a < 1$, $0 < b < 1$, 


Finally, for the given probability, note:
The bivariate, or more generally joint distribution, tells us a lot of information.

Since it describes how $X_1$ and $X_2$ move together, it can certainly be used to describe the “behavior” of $X_1$ or $X_2$ overall, as summarized by their marginal (univariate) distributions.

Formally, we can write

\[
\{X_1 \leq x_1\} = \{X_1 \leq x_1\} \cap \{-\infty < X_2 < \infty\}
\]

\[
= \{X_1 \leq x_1, -\infty < X_2 < \infty\}
\]

\[
\Rightarrow F_{X_1}(x_1) = \Pr[X_1 \leq x_1, -\infty < X_2 < \infty] \quad \forall x_1 \in \mathbb{R}
\]
Case 1: $X_1$ and $X_2$ are discrete. Let $D_{X_1}$ be the support of $X_1$. For $x_1 \in D_{X_1}$,

$$F_{X_1}(x_1) = \sum_{w_1 \leq x_1} \sum_{-\infty < x_2 < \infty} p_{X_1,X_2}(w_1, x_2)$$

$$= \sum_{w_1 \leq x_1} \left\{ \sum_{x_2 < \infty} p_{X_1,X_2}(w_1, x_2) \right\}$$

Thus, in the discrete case, to find the probability that $X_1$ is less than $x_1$, simply add up the probabilities associated with all the joint events for which $X_1 \leq x_1$. The final term in square brackets must also equal the marginal density for $X_1$ at $w_1$ (otherwise, we would have a contradiction). Thus,

$$p_{X_1}(x_1) = \sum_{x_2 < \infty} p_{X_1,X_2}(x_1, x_2) \quad \forall \ x_1 \in D_{X_1}$$
Case 2: $X_1$ and $X_2$ are continuous. Let $D_{x_1}$ be the support of $X_1$. For $x_1 \in D_{x_1}$,

$$F_{X_1}(x_1) = \int_{-\infty}^{x_1} \int_{-\infty}^{\infty} f_{X_1, X_2}(w_1, x_2) \, dx_2 \, dw_1$$

$$= \int_{-\infty}^{x_1} \left\{ \int_{-\infty}^{\infty} f_{X_1, X_2}(w_1, x_2) \, dx_2 \right\} \, dw_1$$

$$\Rightarrow f_{X_1}(w_1) = \int_{-\infty}^{\infty} f_{X_1, X_2}(w_1, x_2) \, dx_2$$

Similarly,

$$f_{X_2}(w_2) = \int_{-\infty}^{\infty} f_{X_1, X_2}(x_1, w_2) \, dx_1$$
Consider the following joint distribution of employment / unemployment and Education levels, obtained from the BLS for women in 2010:

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Unemployed</th>
<th>Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; High School</td>
<td>.01</td>
<td>.057</td>
</tr>
<tr>
<td>High School Only</td>
<td>.023</td>
<td>.245</td>
</tr>
<tr>
<td>Some College</td>
<td>.023</td>
<td>.277</td>
</tr>
<tr>
<td>BA or more</td>
<td>.018</td>
<td>.347</td>
</tr>
</tbody>
</table>

(a) What fraction of women have less than an BA degree?
(b) What is the distribution of (female) educational attainment in the population?
(c) What is the probability that a woman is unemployed in 2010?
As for (a), we need to go back to our formula for calculating a univariate cdf from the joint. Doing so reveals:

As for (b), note, for example, that:

Doing this for all categories, we obtain the joint distribution:

<table>
<thead>
<tr>
<th>&lt; High School</th>
<th>High School Only</th>
<th>Some College</th>
<th>BA or more</th>
</tr>
</thead>
</table>

And, lastly, the probability that a woman is unemployed is:
Example (Bivariate#2)

(a) Calculate the marginal pdf of $X_1$
(b) Calculate the marginal pdf of $X_2$
(c) Find $\Pr(X_1 \leq 1/2)$
(d) $\Pr(X_1 + X_2 \leq 1)$.
To obtain the marginal for, say, $X_1$, we simply integrate out $x_2$:

Similarly, the marginal for $X_2$ will be (we don’t repeat this, since it is obvious given the symmetry:)

$$f_{X_2}(x_2) = x_2 + (1/2), \quad 0 < x_2 < 1.$$ 

To obtain the probability $\Pr(X_1 < [1/2])$, we can work with either the marginal or the joint. For example,
As for the last question in (c), it is convenient to write:

\[ \Pr(X_1 + X_2 \leq 1) = \Pr(X_2 \leq 1 - X_1). \]

Then,

\[
\begin{align*}
\Pr(X_1 + X_2 \leq 1) &= \int_0^1 \int_0^{1-x_1} (x_1 + x_2) \, dx_2 \, dx_1 \\
&= \int_0^1 \left[ x_1 x_2 + \frac{1}{2} x_2^2 \right]_{x_2=0}^{x_2=1-x_1} \, dx_1 \\
&= \int_0^1 x_1 (1 - x_1) + \frac{1}{2} [1 - x_1]^2 \, dx_1 \\
&= \int_0^1 x_1 - x_1^2 + \frac{1}{2} (1 - 2x_1 + x_1^2) \, dx_1 \\
&= \int_0^1 x_1^2 - \frac{1}{2} (1/2) \, dx_1 \\
&= \left. \frac{1}{2} x_1 - \frac{1}{6} x_1^3 \right|_{x_1=0}^{x_1=1} = 1/3
\end{align*}
\]
Example

Define \( f(x, y) \) by

\[
\begin{align*}
  f(0, 0) &= f(0, 1) = 1/6 \\
  f(1, 0) &= f(1, 1) = 1/3 \\
  f(x, y) &= 0 \text{ for any other } (x, y)
\end{align*}
\]

Find the marginal pmfs of \( X \) and \( Y \).

From the table, it is clear that

- 

Similarly,

- 

Example

Define $g(x, y)$ by

$$
g(0, 0) = \frac{1}{12}, \quad g(1, 0) = \frac{5}{12} \\
g(0, 1) = g(1, 1) = \frac{3}{12} \\
g(x, y) = 0 \text{ for any other } (x, y)
$$

Find the marginal pmfs of $X$ and $Y$.

From the table, it is clear that

-  

Similarly,

-  

In the above two examples, the marginal pmfs of $X$ are same as are the marginal pmfs of $Y$.

However, the joint pmfs are different!

Hence, we cannot determine the joint distribution from the marginals alone. The joint pmf gives us additional information not present in the marginals.

The conditional densities, however, are another story ...
Expectations, Bivariate Case

- Let \((X_1, X_2)\) be a random vector and \(Y = g(X_1, X_2)\) where \(g : \mathbb{R}^2 \to \mathbb{R}\).

- If \((X_1, X_2)\) is continuous and
  \[
  \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g(x_1, x_2)| f_{X_1,X_2}(x_1, x_2) \, dx_1 \, dx_2 < \infty,
  \]
  then
  \[
  E(Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1, x_2) f_{X_1,X_2}(x_1, x_2) \, dx_1 \, dx_2
  \]

- If \((X_1, X_2)\) is discrete and
  \[
  \sum_{x_1} \sum_{x_2} |g(x_1, x_2)| p_{X_1,X_2}(x_1, x_2) < \infty,
  \]
  then
  \[
  E(Y) = \sum_{x_1} \sum_{x_2} g(x_1, x_2) p_{X_1,X_2}(x_1, x_2)
  \]
Theorem

Let \( Y_1 = g_1(X_1, X_2) \) and \( Y_2 = g_2(X_1, X_2) \) be random variables whose expectations exist. Then, for any real numbers \( k_1 \) and \( k_2 \),

\[
E(k_1 Y_1 + k_2 Y_2) = k_1 E(Y_1) + k_2 E(Y_2)
\]

We defer the proof of this theorem, as it is reasonably obvious: it follows from the fact that the integral of a sum is the sum of the integrals.

Example (2.1.13)

Consider

Find

\( (a) \ E(X_1), \ E(X_2), \ E(X_1 X_2) \)
\( (b) \ Does \ E(X_1 X_2) = E(X_1)E(X_2)? \)
Is this reasonable? Also note that, given the symmetry of the problem,
Thus, in this case

However, THIS WILL NOT BE TRUE IN GENERAL. Although the expectation of a sum is the sum of the expectations, the same is not typically true for products of random variables!