5 Multivariate Probability Distributions

5.1 Joint Probability Distributions

5.2 Expected Values of Functions of $rv$s

5.3 Covariance

5.4 Linear Functions of $rv$s

5.5 Conditional Expectation
5 Multivariate Probability Distributions

5.1 Joint Probability Distributions

Joint Distribution of Two Discrete rvs

Let \((X, Y)\) be 2 discrete rvs, then their joint probability function is given as:

\[
p(x, y) = P(X = x \text{ and } Y = y)
\]

and is defined over the range of \((X, Y)\).

If \(A\) is any subset of the range of \((X, Y)\) then

\[
P\{(X, Y) \in A\} = \sum_{(x, y) \in A} p(x, y)
\]
The joint probability function must satisfy:

- \( p(x, y) \geq 0 \) for all \((x, y)\) in the range of \((X, Y)\).
- \[
\sum_{\text{range of } (X,Y)} p(x, y) = 1
\]

**Example 5.1.1** The following is the joint pmf for \((X, Y)\):

<table>
<thead>
<tr>
<th>(p(x, y))</th>
<th>(y)</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x)</td>
<td>1</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td></td>
</tr>
</tbody>
</table>

Calculate

1. \( P(X = 1, Y = 1) \)
2. $P(X = 1, Y \leq 1)$

3. $P(X = Y)$
Marginal Distributions

The marginal distribution of $X$ based on the joint distribution of $(X,Y)$ is denoted by:

$$P(X = x) = p_X(x) \text{ for range of } X$$

and can be obtained from the joint probability function:

$$p_X(x) = \sum_{\text{range of } y} p(x, y)$$

Similarly, the marginal distribution of $Y$ can be obtained by:

$$p_Y(y) = \sum_{\text{range of } x} p(x, y)$$
Example 5.1.2 For the \((X, Y)\) in Example 5.1.1, calculate \(p_X(1)\) and \(p_Y(1)\).

<table>
<thead>
<tr>
<th>(p(x, y))</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x = 1)</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>(x = 2)</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
</tr>
</tbody>
</table>

Note: The sum of all probabilities in the table should equal 1.0, which is indicated by the last row.
Conditional Probability Function

Let $p(x, y)$ be the joint pmf for $(X, Y)$ and $p_Y(y)$ be the marginal pmf for $Y$. Then the conditional pmf for $X$ given $Y = y$ is

$$p_X(x|y) = \frac{p(x, y)}{p_Y(y)}$$

for values of $y$ for which $p_Y(y) > 0$. That is,

$$P(X = x|Y = y) = p_X(x|y) = \frac{p(x, y)}{p_Y(y)}.$$

Similarly,

$$P(Y = y|X = x) = p_Y(y|x) = \frac{p(x, y)}{p_X(x)}.$$

**Note:** $p_X(x|y)$ is a pmf of $X$ for the subpopulation with $Y = y$; $p_Y(y|x)$ is a pmf of $Y$ for the subpopulation with $X = x$. 
Example 5.1.3 For the \((X, Y)\) in Example 5.1.1, calculate \(p_Y(y|1)\) for \(y = 0, 1, 2\).

\[
\begin{array}{c|ccc}
\text{p}(x, y) & y \\
& 0 & 1 & 2 \\
\hline
x & & & \\
1 & 0.3 & 0.2 & 0.1 \\
2 & 0.1 & 0.2 & 0.1 \\
\hline
\end{array}
\]

\[
\begin{array}{c|c}
\text{p}(y|1) & 1.0 \\
\hline
\end{array}
\]
Independence of Discrete Random Variables

If the joint pmf for $(X, Y)$ is $p(x, y)$, then rv $X$ and $Y$ are said to be independent if, and only if,

$$p(x, y) = p_X(x)p_Y(y)$$

for ALL pairs $(x, y)$ in the range of $(X, Y)$.

Otherwise we say $X$ and $Y$ are dependent.

**Example 5.1.4** Are $(X, Y)$ in Example 5.1.1 independent?

<table>
<thead>
<tr>
<th>$p(x, y)$</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>1.0</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary of bivariate discrete distribution.

<table>
<thead>
<tr>
<th>Joint pmf</th>
<th>$p(x, y) = P(X = x, Y = y)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P{(X, Y) \in A} = \sum_{all (x,y) \in A} p(x, y)$</td>
<td>[= \sum_{all (x,y) \in A} p(x, y)]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Marginal pmf</th>
<th>$P_X(x) = P(X = x) = \sum_{all y} p(x, y)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_Y(y) = P(Y = y) = \sum_{all x} p(x, y)$</td>
<td>[= \sum_{all x} p(x, y)]</td>
</tr>
</tbody>
</table>

| Conditional pmf            | $P_X(x|y) = P(X = x|Y = y) = \frac{p(x,y)}{P_Y(y)}$ |
|----------------------------|----------------------------------------------------|
| $P_Y(y|x) = P(Y = y|X = x) = \frac{p(x,y)}{P_X(x)}$ | \[= \frac{p(x,y)}{P_X(x)}\] |

| Independence              | If $p(x, y) = p_X(x) \cdot p_Y(y)$ for all $(x,y)$ in the range. |
Example 5.1.5 Suppose \((X, Y)\) has the following joint pmf. Are \((X, Y)\) independent?

<table>
<thead>
<tr>
<th>(p(x, y))</th>
<th>(\quad y)</th>
<th>(\quad 0)</th>
<th>(\quad 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x)</td>
<td>1</td>
<td>0.12</td>
<td>0.18</td>
</tr>
<tr>
<td>2</td>
<td>0.28</td>
<td>0.42</td>
<td></td>
</tr>
</tbody>
</table>

\[1.0\]
Example 5.1.6 Let $Y_1 = 0$ if child survived; $Y_1 = 1$ if not. Let $Y_2 = 0$ if no belt used; 1 if adult belt used; 2 if car-seat belt used.

| $p(y_1, y_2)$ | $y_1$ |  
|---|---|---|
| $y_2$ | 0 | 0.38 0.17 |
| | 1 | 0.14 0.02 |
| | 2 | 0.24 0.05 |
| | | 1.0 |

1. Give the marginal distributions for $Y_1$ and $Y_2$.

2. Give the conditional distribution of $Y_2$ given $Y_1 = 0$.

3. What is the probability that a child survived given that he/she was in a car-seat belt.

4. Are $Y_1$ and $Y_2$ independent of each other?
Joint Distribution of Two Continuous rvs

Let $X$ and $Y$ be continuous random variables. Then their joint probability density function, $f(x, y)$, has the property that for events of the form:

$$A = \{a \leq X \leq b, c \leq Y \leq d\},$$

$$P\{(X, Y) \in A\} = P(a \leq X \leq b, c \leq Y \leq d) = \int_{a}^{b} \int_{c}^{d} f(x, y) dy dx.$$ 

Properties of joint pdf $f(x, y)$:

- $f(x, y) \geq 0$
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dy dx = 1.$
Joint Cumulative Distribution Function

For rvs \((X, Y)\) (can be either discrete or continuous), the joint cumulative distribution function (cdf), \(F(x, y)\), is defined as

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

for \(-\infty < x < +\infty\) and \(-\infty < y < +\infty\).

**Example 5.1.7** Suppose the joint pdf for rv \(X\) and \(Y\) is:

\[
f(x, y) = 0.25, \ 0 < x < 2, \ 0 < y < 2.
\]

**Calculate** \(P(X \leq 1, Y \leq 0.5)\)
Example 5.1.8 Suppose the joint pdf for rv $X$ and $Y$ is:

$$f(x, y) = x + y, 0 < x < 1, 0 < y < 1.$$ 

Calculate $P(1/4 \leq X \leq 3/4, 1/4 \leq Y \leq 3/4)$
Example 5.1.9 (5.12) Suppose the joint pdf for rv $X$ and $Y$ is:

$$f(x, y) = 2, \ 0 \leq x \leq 1, \ 0 \leq y \leq 1, \ 0 \leq x + y \leq 1.$$  

Calculate $P(X \leq 3/4, Y \leq 3/4)$ and $P(X \leq 1/2, Y \leq 1/2)$. 
Marginal Distributions of Continuous \( rvs \) Let \( X \) and \( Y \) be continuous random variables with joint density function \( f(x, y) \), then the marginal distribution of \( X \) is denoted by \( f_X(x) \) and can be be derived from the joint pdf as

\[
f_X(x) = \int_{-\infty}^{+\infty} f(x, y) \, dy.
\]

Similarly, the marginal distribution of \( Y \) is denoted by \( f_Y(y) \) and can be be derived from the joint pdf as

\[
f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) \, dx.
\]
Example 5.1.10 Suppose the joint pdf for rvs \((X, Y)\) is

\[ f(x, y) = x + y, \quad 0 < x < 1, \quad 0 < y < 1. \]

Derive the marginal distribution of \(Y\)
Conditional Distribution for Continuous rv$s

Let $f(x, y)$ be the joint pdf for continuous rv$s$ $(X, Y)$. Then the conditional distribution of $X$ given $Y = y$ is defined as:

$$f(x|Y = y) = \frac{f(x, y)}{f_Y(y)}$$

for $y$ values such that $f_Y(y) > 0$. Here $f_Y(y)$ is the marginal pdf of $Y$.

**Note:** the $f(x|Y = y)$ is a pdf of $X$ for the subpopulation with $Y = y$. 
Example 5.1.11 Suppose the joint pdf for rvs \((X, Y)\) is

\[
f(x, y) = x + y, \quad 0 < x < 1, \quad 0 < y < 1.
\]

Drive \(f_X(x|Y = 1/4)\) and calculate \(P(X \leq 1/2|Y = 1/4)\).
Example 5.1.12 Suppose \( rvs \ (X, Y) \) has joint pdf:
\[ f(x, y) = \frac{1}{2}, \quad 0 < x < y < 2. \]
Derive the conditional distribution of \( Y \) given \( X = x \).
Indepedence of Continuous rvs

Two continuous rvs, \((X, Y)\), are \textbf{independent} if, and only if,

\[ f(x, y) = f_X(x)f_Y(y), \quad \text{for all } (x, y), \]

otherwise \(X\) and \(Y\) are said to be \textbf{dependent}.

\textbf{Example 5.1.13} Suppose the joint pdf for rvs \((X, Y)\) is:

\[ f(x, y) = x + y, 0 < x < 1, 0 < y < 1. \]

Show that \(X\) and \(Y\) are not independent.
Example 5.1.14 Suppose the joint pdf for rvs \((X, Y)\) is:

\[
f(x, y) = \lambda^2 e^{-\lambda(x+y)}, \quad 0 < x < +\infty, \quad 0 < y < +\infty
\]

and \(0 < \lambda < +\infty\) is known. Are \(X\) and \(Y\) independent?
### Summary of bivariate distribution.

<table>
<thead>
<tr>
<th></th>
<th>discrete</th>
<th>continuous</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Joint pmf/pdf</strong></td>
<td>$p(x, y) = P(X = x, Y = y)$</td>
<td>$f(x, y):$ density</td>
</tr>
<tr>
<td>$P{ (X, Y) \in A } = \sum_{(x,y) \in A} p(x,y)$</td>
<td>$P{ (X, Y) \in A } = \int \int_A f(x,y) , dx , dy$</td>
<td></td>
</tr>
<tr>
<td><strong>Marginal pmf/pdf</strong></td>
<td>$P_X(x) = P(X = x) = \sum_{all , y} p(x,y)$</td>
<td>$f_X(x) = \int f(x,y) , dy$</td>
</tr>
<tr>
<td>$P_Y(y) = P(Y = y) = \sum_{all , x} p(x,y)$</td>
<td>$f_Y(y) = \int f(x,y) , dx$</td>
<td></td>
</tr>
<tr>
<td><strong>Conditional pmf/pdf</strong></td>
<td>$P_X(x</td>
<td>y) = \frac{p(x,y)}{P_Y(y)}$</td>
</tr>
<tr>
<td>$P_Y(y</td>
<td>x) = \frac{p(x,y)}{P_X(x)}$</td>
<td>$f_Y(y</td>
</tr>
<tr>
<td><strong>Independence</strong></td>
<td>If $p(x, y) = p_X(x) \cdot p_Y(y)$</td>
<td>If $f(x, y) = f_X(x)f_Y(y)$</td>
</tr>
<tr>
<td>for all $(x, y)$ in the range</td>
<td>for all $(x, y)$</td>
<td></td>
</tr>
</tbody>
</table>
Example 5.1.15 Suppose \((X, Y)\) has the joint pdf 
\[ f(x, y) = 30xy^2, \ x - 1 \leq y \leq 1 - x, \ 0 \leq x \leq 1. \]

1. Calculate \(F(1/2, 1/2)\). Answer: \(9/16\)
2. Calculate \(F(1/2, 2)\). Answer: \(13/16\)
3. Calculate \(P(X > Y)\). Answer: \(21/32\)
Example 5.1.16 Suppose \((X, Y)\) has the joint pdf
\[ f(x, y) = 6(1 - y), \quad 0 \leq x \leq y \leq 1. \] Are \(X\) and \(Y\) independent?
5.2 Expected Values of Functions of rvs

Expectations of Functions of Discrete rvs

Let $Y_1, Y_2, \cdots, Y_k$ be $k$ discrete rvs with joint pmf $p(y_1, y_2, \cdots, y_k)$, and let $U = g(Y_1, Y_2, \cdots, Y_k)$ be a function of $Y_1, Y_2, \cdots, Y_k$. Then the expectation of $U$ is defined as:

$$E(U) = \sum_{y_1} \sum_{y_2} \cdots \sum_{y_k} g(y_1, y_2, \cdots, g_k)p(y_1, y_2, \cdots, y_k).$$

Expectations of Functions of Continuous rvs

Let $Y_1, Y_2, \cdots, Y_k$ be $k$ continuous rvs with joint pdf $f(y_1, y_2, \cdots, y_k)$, and let $U = g(Y_1, Y_2, \cdots, Y_k)$ be a function of $Y_1, Y_2, \cdots, Y_k$. Then the expectation of $U$ is defined as:

$$E(U) = \int_{y_1} \int_{y_2} \cdots \int_{y_k} g(y_1, y_2, \cdots, g_k)f(y_1, y_2, \cdots, y_k)dy_k \cdots dy_2 dy_1.$$
**Example 5.2.1** The following is the joint pmf of $(X, Y)$:

<table>
<thead>
<tr>
<th>$p(x, y)$</th>
<th>$y$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>$x$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Calculate $E(XY)$ and $E\{\max(X, Y)\}$. 
Example 5.2.2 Suppose rvs $(X, Y)$ has joint pdf:

$$f(x, y) = \frac{1}{2}, \quad 0 \leq x < y \leq 2.$$  

Calculate (1) $E(XY)$; (2) $E(Y)$.  

Expected Values of Special Functions

Let $c, c_1, \cdots, c_k$ be constants, $Y_1$ and $Y_2$ be rvs, and $g_1(Y_1, Y_2), g_2(Y_1, Y_2), \cdots, g_k(Y_1, Y_2)$ be $k$ functions of $Y_1$ and $Y_2$. Then by definition we can show:

- $E(c) = c$
- $E \{cg_1(Y_1, Y_2)\} = cE \{g_1(Y_1, Y_2)\}$.
- $E \{c_1 g_1(Y_1, Y_2) + \cdots + c_k g_k(Y_1, Y_2)\} = c_1 E \{g_1(Y_1, Y_2)\} + \cdots + c_k E \{g_k(Y_1, Y_2)\}$.

See Chapters 3 and 4 for the analogous results for univariate cases.
Expectation of Products of Functions of Independent rv's

Theorem 1 Let $X$ and $Y$ be independent rv's. Let $g(X)$ and $h(Y)$ be functions of $X$ and $Y$, respectively. Then

$$E \{g(X)h(Y)\} = E\{g(X)\}E\{h(Y)\}.$$ 

Proof:
5.3 Covariance

If $X$ and $Y$ are rvs, the covariance of $X$ and $Y$ is defined as:

$$\text{Cov}(X, Y) = E \left[ (X - E(X))(Y - E(Y)) \right]$$

**Theorem 2** The computing formula for covariance is

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = E(XY) - \mu_X \mu_Y$$

Proof:
Example 5.3.1 Suppose the joint pdf for rvs \((X, Y)\) is:

\[ f(x, y) = x + y, 0 < x < 1, 0 < y < 1. \]

Calculate \(Cov(X, Y)\)
Properties of Covariance:

- $Cov(Y, Y) = Var(Y, Y)$
- $Cov(X, Y) = Cov(Y, X)$
- Covariance measures the linear dependence between two random variables.
- $Cov(X, Y) = 0 \Rightarrow X$ and $Y$ are **uncorrelated**
- Independence $\Rightarrow$ Uncorrelated (prove this)
- Uncorrelated $\not\Rightarrow$ Independence
Example 5.3.2 The following is the joint pmf of \((X, Y)\):

<table>
<thead>
<tr>
<th>(p(x, y))</th>
<th>(-1)</th>
<th>0</th>
<th>+1</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>0</td>
<td>1/4</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1/4</td>
<td>0</td>
<td>1/4</td>
</tr>
<tr>
<td>+1</td>
<td>0</td>
<td>1/4</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1/4</td>
<td>1/2</td>
<td>1/4</td>
</tr>
</tbody>
</table>

Verify that \(X\) and \(Y\) are uncorrelated but not independent.
Correlation

The correlation between rvs, X and Y, is defined as

\[ \rho_{X,Y} = \text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}, \]

where \( \sigma_X \) and \( \sigma_Y \) are the standard deviations of \( X \) and \( Y \), respectively.

**Properties of Correlation:**

- \( -1 \leq \text{Corr}(X, Y) \leq 1 \)
- Values of \( \text{Corr}(X, Y) \) near 0 indicate weak linear association between \( X \) and \( Y \)
- \( X \) and \( Y \) are independent \( \Rightarrow \text{Corr}(X, Y) = 0 \)
- Values of \( \text{Corr}(X, Y) \) near +1 (-1) indicate strong positive (negative) linear association between \( X \) and \( Y \)
- If high values of \( X \) associate with high values of \( Y \), while low
values of $X$ associate with low values of $Y$, then $\text{Corr}(X, Y) > 0$

- If high values of $X$ associate with low values of $Y$, while low values of $X$ associate with high values of $Y$, then $\text{Corr}(X, Y) < 0$

- Scale-free measure of association

**Example 5.3.3** Suppose the joint pdf for rvs $(X, Y)$ is

$$f(x, y) = x + y, 0 < x < 1, 0 < y < 1.$$  

Calculate $\text{Corr}(X, Y)$.  

5.4 Linear Functions of \textit{rvs}

Let $Y_1, Y_2, \cdots, Y_n$ be $n$ random variables, and $a_1, a_2, \cdots, a_n$ be $n$ constants. Then the function

$$U = a_1 Y_1 + a_2 Y_2 + \cdots + a_n Y_n = \sum_{i=1}^{n} a_i Y_i$$

is called a linear function (or a linear combination) of $Y_1, Y_2, \cdots, Y_n$. Since $Y_1, Y_2, \cdots, Y_n$ are random variables, $U$ is also a random variable.
Theorem 3 (Linear Functions of \(r\)vs) Let \(Y_1, Y_2, \ldots, Y_n\) be random variables, and \(a_1, a_2, \ldots, a_n\) be constants. Let 
\[ U = \sum_{i=1}^{n} a_i Y_i. \]
Then
\[
E(U) = \sum_{i=1}^{n} a_i E(Y_i)
\]
\[
V(U) = \sum_{i=1}^{n} a_i^2 V(Y_i) + 2 \sum_{i<j} a_i a_j Cov(Y_i, Y_j)
\]

Theorem 4 (Linear Functions of Uncorrelated \(r\)vs) Let 
\(Y_1, Y_2, \ldots, Y_n\) be uncorrelated random variables, and \(a_1, a_2, \ldots, a_n\) be constants. Let 
\[ U = \sum_{i=1}^{n} a_i Y_i. \]
Then
\[
E(U) = \sum_{i=1}^{n} a_i E(Y_i)
\]
\[
V(U) = \sum_{i=1}^{n} a_i^2 V(Y_i)
\]
**Example 5.4.1** Suppose the joint pdf for rvs \((X, Y)\) is

\[ f(x, y) = x + y, \quad 0 < x < 1, \quad 0 < y < 1. \]

**Given:**

\[ E(X) = E(Y) = \frac{7}{12}, \quad Cov(X, Y) = -\frac{1}{144} \]

\[ E(X^2) = E(Y^2) = \frac{5}{12}. \]

Let \(U = (X + Y)/2\). *Calculate* \(E(U)\) *and* \(V(U)\).
5.5 Conditional Expectation

Definition. Let $X$ and $Y$ be continuous random variables with conditional pdf of $X$ given $Y = y$ as:

$$f(x|Y = y)$$

Then the conditional expectation of $X$ given $Y = y$ is:

$$E(X|Y = y) = \int_{-\infty}^{+\infty} x f(x|Y = y) dx$$

If $X$ and $Y$ are discrete rvs with conditional pmf: $p_x(x|Y = y)$. Then the conditional expectation of $X$ given $Y = y$ is:

$$E(X|Y = y) = \sum_{\text{range of } x} x p_x(x|y) dx$$
Example 5.5.1 Suppose the joint pdf for rvs \((X, Y)\) is

\[ f(x, y) = x + y, \quad 0 < x < 1, \quad 0 < y < 1. \]

Calculate \(E(X|Y = 1/4)\).
Theorem 5  Law of Iterated Expectations

\[ E(X) = E\{E(X|Y)\} \]

Similarly for the variance:

\[ V(X) = E\{V(X|Y)\} + V\{E(X|Y)\} \]
Example 5.5.2 A quality control plan for an assembly line involves sampling \( n = 10 \) finished items per day and counting \( Y \), the number of defectives. Let \( p \) denote the probability of observing defective, then \( Y \sim \text{Binomial}(n, p) \). But \( p \) varies from day to day and is assumed to follow \( U(0, 1/4) \) distribution. Find \( E(Y) \) and \( V(Y) \). What if \( p \sim \text{Beta}(\alpha = 1, \beta = 7) \)?