We are often interested in the probabilistic description of how two or more random variables behave together. We will consider the case of two random variables.

**Definition** Suppose $X$ and $Y$ are two discrete r.v.'s defined on the same sample space, having possible values in the sets $\mathcal{X}$ and $\mathcal{Y}$, respectively. The *joint probability mass function* (pmf) of $X$ and $Y$ is defined by

$$p(x, y) = P(X = x, Y = y)$$

**Remark:** $P(X = x, Y = y)$ is a simplified notation for the probability

$$P(X = x \text{ AND } Y = y) = P\{X = x\} \cap \{Y = y\}$$

**Properties of the joint pmf**

1. $p(x, y) \geq 0$ for all $x$ and $y$.
2. $p(x, y) = 0$ if $x \notin \mathcal{X}$ or $y \notin \mathcal{Y}$.
3. $\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) = 1$

**Proof:** Follows the same lines as the proof of $\sum_{x \in \mathcal{X}} P(X = x) = 1$.

4. For any $A \subset \mathbb{R}^2$,

$$P((X, Y) \in A) = \sum_{x \in \mathcal{X}, y \in \mathcal{Y} \cap (x, y) \in A} p(x, y)$$

**Proof:** Follows the same lines as the proof of $P(X \in B) = \sum_{x \in \mathcal{X}, x \in B} P(X = x)$.

Given the joint pmf, we can recover the individual pmf’s of $X$ and $Y$:

$$p_X(x) = P(X = x) = P(X = x \text{ AND } Y \in \mathcal{Y}) = \sum_{y \in \mathcal{Y}} P(X = x, Y = y) = \sum_{y \in \mathcal{Y}} p(x, y)$$

Thus

$$p_X(x) = \sum_{y \in \mathcal{Y}} p(x, y)$$

Similarly,

$$p_Y(y) = \sum_{x \in \mathcal{X}} p(x, y)$$

The pmf’s $p_X(x)$ and $p_Y(y)$ are called the *marginal pmf’s* of $X$ and $Y$. 

**Discrete Random Variables**
**Example:** Five cards are drawn from a standard deck of 52 cards without replacement. Let $X$ be the number of spades drawn and $Y$ the number of clubs drawn. Find the joint pmf of $X$ and $Y$.

**Solution:** We have $X = Y = \{0, 1, 2, 3, 4, 5\}$. If $x + y > 5$, then $P(X = x, Y = y) = 0$. Otherwise

$$P(X = x, Y = y) = \frac{\binom{13}{x} \binom{13}{y} \binom{26}{5-x-y}}{\binom{52}{5}}$$

Thus

$$p(x, y) = \begin{cases} 
\frac{\binom{13}{x} \binom{13}{y} \binom{26}{5-x-y}}{\binom{52}{5}} & \text{if } 0 \leq x, y \leq 5, \: x + y \leq 5 \\
0 & \text{otherwise}
\end{cases}$$

We obtain

<table>
<thead>
<tr>
<th>$y$</th>
<th>$x$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>$p_Y(y)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>0.1</td>
<td>0.05</td>
<td>0.05</td>
<td>0.2</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td>0.6</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
<td>0.2</td>
</tr>
<tr>
<td>$p_X(x)$</td>
<td>0.5</td>
<td>0.35</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(This is why $p_X(x)$ and $p_Y(y)$ are called the marginal pmf’s.)

$$P(X + Y \leq 2) = \sum_{x+y\leq2} p(x, y)$$

$$= p(0,0) + p(0,1) + p(1,0) + p(1,1) + p(0,2) + p(2,0)$$

$$= 0.1 + 0.3 + 0.05 + 0.2 + 0.1 + 0.05$$

$$= 0.8$$

**Example:** The joint pmf $p(x, y)$ of $X$ and $Y$ are given in the table below. Find the marginal pmf’s and $P(X + Y \leq 2)$.

<table>
<thead>
<tr>
<th>$y$</th>
<th>$x$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>$p_Y(y)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>0.1</td>
<td>0.05</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

**Solution:**

$$p_X(0) = 0.1 + 0.3 + 0.1 = 0.5$$

$$p_X(1) = 0.05 + 0.2 + 0.1 = 0.35$$

$$p_X(2) = 0.05 + 0.1 + 0 = 0.15$$

Similarly,

$$p_Y(0) = 0.2, \quad p_Y(1) = 0.6, \quad p_Y(2) = 0.2$$

**Remark:** The joint pmf $p(x, y)$ uniquely determines the marginal pmf’s through the relationship

$$p_X(x) = \sum_{y \in Y} p(x, y), \quad p_Y(y) = \sum_{x \in X} p(x, y)$$

However, the reverse is not true. For given $p_X(x)$ and $p_Y(y)$ there are (infinitely) many joint pmf’s whose marginals are $p_X(x)$ and $p_Y(y)$.

For example, compare this table and the previous example:

<table>
<thead>
<tr>
<th>$y$</th>
<th>$x$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>$p_Y(y)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>0.1</td>
<td>0.07</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>0.3</td>
<td>0.21</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.1</td>
<td>0.07</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>$p_X(x)$</td>
<td>0.5</td>
<td>0.35</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Expected value

Note that from the joint pmf we can compute the marginal pmf’s and so the expected values of $X$ and $Y$:

\[ E(X) = \sum_{x \in X} x p_X(x) \quad \text{and} \quad E(Y) = \sum_{y \in Y} y p_Y(y) \]

If $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a real function of two variables, $h(X, Y)$ is a discrete random variable.

**Theorem 1**

If $X$ and $Y$ are discrete r.v.’s with joint pmf $p(x, y)$ and $h(x, y)$ is a real function of two variables, then

\[ E[h(X, Y)] = \sum_{x \in X} \sum_{y \in Y} h(x, y) p(x, y) \]

**Proof:** Follows the lines of the proof of $E[g(X)] = \sum_{x \in X} g(x) P(X = x)$.

**Corollary 2**

If $X$ and $Y$ are discrete r.v.’s, then

\[ E(X + Y) = E(X) + E(Y) \]

**Proof:** With $h(x, y) = x + y$ we have

\[
E(X + Y) = \sum_{x \in X} \sum_{y \in Y} (x + y) p(x, y) \\
= \sum_{x \in X} \sum_{y \in Y} xp(x, y) + \sum_{x \in X} \sum_{y \in Y} yp(x, y) \\
= \sum_{x \in X} x \sum_{y \in Y} p(x, y) + \sum_{y \in Y} y \sum_{x \in X} p(x, y) \\
= \sum_{x \in X} xp_X(x) + \sum_{y \in Y} yp_Y(y) \\
= E(X) + E(Y)
\]

Jointly Continuous Random Variables

**Definition** The random variables $X$ and $Y$ are said to be jointly continuous if there exists a nonnegative function $f(x, y)$ such that for any “reasonable” planar set $C \subset \mathbb{R}^2$,

\[ P((X, Y) \in C) = \iint_C f(x, y) \, dx \, dy \]

The function $f(x, y)$ is called the joint probability density function (joint pdf) of $X$ and $Y$.

**Properties of jointly continuous r.v.’s**

1. $f(x, y) \geq 0$ (by definition).
2. \[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx \, dy = 1
\]

**Proof:**

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx \, dy = \int_{\mathbb{R}^2} f(x, y) \, dx \, dy \\
= P((X, Y) \in \mathbb{R}^2) \\
= 1
\]
3. For any \( a \leq b \) and \( c \leq d \),
\[
P(a \leq X \leq b, c \leq Y \leq d) = \int_c^d \int_a^b f(x, y) \, dx \, dy
\]

**Proof:** Let \( C = \{(x, y) : a \leq x \leq b, c \leq y \leq d\} \). Then
\[
P(a \leq X \leq b, c \leq Y \leq d) = P((X, Y) \in C)
= \int_C f(x, y) \, dx \, dy
= \int_c^d \int_a^b f(x, y) \, dx \, dy \quad \square
\]

**Note:** Letting \( a = b = u \) and \( c = d = v \), this shows that
\[
P(X = u, Y = v) = 0
\]
for all \( u \) and \( v \).

**Example:** Let the joint pdf of \( X \) and \( Y \) be given by
\[
f(x, y) = \begin{cases} 
  kxy & 0 \leq x, y \leq 1 \\
  0 & \text{otherwise}
\end{cases}
\]

(a) Determine \( k \).

**Solution:**
\[
1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx \, dy
= \int_0^1 \int_0^1 kxy \, dx \, dy
= k \left( \int_0^1 x \, dx \right) \left( \int_0^1 y \, dy \right)
= k \left( \frac{1}{2} \right)^2 = k \frac{1}{4}
\]
Thus \( k = 4 \).

Similarly to the discrete case, the *marginal* pdf’s can be obtained from the joint pdf:
\[
f_X(x) = \int_{-\infty}^{\infty} f(x, y) \, dy \quad \text{and} \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y) \, dx
\]

**Proof sketch:**
\[
P(a \leq X \leq b) = P(a \leq X \leq b, -\infty < Y < \infty)
= \int_{-\infty}^{\infty} \int_a^b f(x, y) \, dx \, dy
= \int_a^b \left( \int_{-\infty}^{\infty} f(x, y) \, dy \right) \, dx
\]

Thus the function \( g(x) = \int_{-\infty}^{\infty} f(x, y) \, dy \) satisfies
\[
P(a \leq X \leq b) = \int_a^b g(x) \, dx
\]
for all \( a < b \), and so it must equal to the pdf \( f_X \) of \( X \).

(b) Find the marginal pdf’s.

**Solution:** For \( 0 \leq x \leq 1 \),
\[
f_X(x) = \int_{-\infty}^{\infty} f(x, y) \, dy
= \int_0^1 4xy \, dy
= 2x
\]
Note that \( f(x, y) = 0 \) if \( x \notin [0, 1] \), so \( f_X(x) = 0 \) if \( x \notin [0, 1] \). Thus
\[
f_X(x) = \begin{cases} 
  2x & 0 \leq x \leq 1 \\
  0 & \text{otherwise}
\end{cases}
\]

By symmetry
\[
f_Y(y) = \begin{cases} 
  2y & 0 \leq y \leq 1 \\
  0 & \text{otherwise}
\end{cases}
\]
(c) Find \( P(X + Y \leq 1) \).

**Solution:** Note that \( P(X + Y \leq 1) = P((X, Y) \in C) \), where

\[
C = \{(x, y) : x + y \leq 1\}
\]

Thus

\[
P(X + Y \leq 1) = \iint_C f(x, y) \, dx \, dy = \int_0^1 \int_0^{1-y} f(x, y) \, dx \, dy
\]

\[
= \int_0^1 \int_0^{1-y} 4xy \, dx \, dy
\]

\[
= \int_0^1 2y(1-y)^2 \, dy
\]

\[
= 2 \int_0^1 (y - 2y^2 + y^3) \, dy = 2 \left[ \frac{y^2}{2} - \frac{2y^3}{3} + \frac{y^4}{4} \right]_0^1
\]

\[
= 2 \left( \frac{1}{2} - \frac{2}{3} + \frac{1}{4} \right) = \frac{1}{6}
\]

---

(b) Find the marginal pdf’s \( f_X(x) \) and \( f_Y(y) \).

**Solution:** For \( x \in [0, 1] \),

\[
f_X(x) = \int_{-\infty}^\infty f(x, y) \, dy = \int_x^1 10xy^2 \, dy = \frac{10}{3} x(1 - x^3)
\]

Thus

\[
f_X(x) = \begin{cases} 
\frac{10}{3} x(1 - x^3) & 0 \leq x \leq 1 \\
0 & \text{otherwise}
\end{cases}
\]

Similarly, for \( y \in [0, 1] \),

\[
f_Y(y) = \int_{-\infty}^\infty f(x, y) \, dx = \int_y^\infty 10xy^2 \, dx = 5y^4
\]

hence

\[
f_Y(y) = \begin{cases} 
5y^4 & 0 \leq y \leq 1 \\
0 & \text{otherwise}
\end{cases}
\]

---

**Example:** Suppose the joint pdf of \( X \) and \( Y \) is given by

\[
f(x, y) = \begin{cases} 
10xy^2 & 0 \leq x \leq y \leq 1 \\
0 & \text{otherwise}
\end{cases}
\]

(a) Find \( P(Y \leq 2X) \).

**Solution:** Note that \( f(x, y) = 0 \) outside the region

\[
A = \{(x, y) : 0 \leq x \leq y \leq 1\}.
\]

Letting \( C = \{(x, y) : y \geq 2x\} \), we have

\[
P(Y \leq 2X) = \iint_C f(x, y) \, dx \, dy = \int_{C \cap A} 10xy^2 \, dx \, dy
\]

\[
= \int_0^1 \int_{y/2}^y 10xy^2 \, dx \, dy
\]

\[
= 5 \int_0^1 y^2 \left( y^2 - \frac{y^2}{4} \right) \, dy
\]

\[
= \frac{15}{4} \int_0^1 y^4 \, dy = \frac{3}{4}
\]

---

**Geometric Probability**

Suppose a point is drawn randomly from a bounded planar region \( B \subset \mathbb{R}^2 \). The mathematical model is a pair of random variables \((X, Y)\) with joint pdf given by

\[
f(x, y) = \begin{cases} 
c & \text{if } (x, y) \in B \\
0 & \text{otherwise}
\end{cases}
\]

where \( c > 0 \) is a constant.

The constant is determined by

\[
1 = \iint_{\mathbb{R}^2} f(x, y) \, dx \, dy = \iint_{B} c \, dx \, dy \quad \text{(since } f(x, y) = 0 \text{ outside } B)
\]

\[
= c \cdot \text{area}(B)
\]

so that \( c = \frac{1}{\text{area}(B)} \).
Let’s compute \( P((X, Y) \in A) \) for some \( A \subset \mathbb{R}^2 \):

\[
P((X, Y) \in A) = \int_{A} f(x, y) \, dx \, dy
\]

\[
= \int_{A \cap B} \frac{1}{\text{area}(B)} \, dx \, dy
\]

\[
= \frac{\text{area}(A \cap B)}{\text{area}(B)}
\]

Thus if \((X, Y)\) is a point randomly drawn from \(B\), then for all “reasonable” \(A \subset \mathbb{R}^2\),

\[
P((X, Y) \in A) = \frac{\text{area}(A \cap B)}{\text{area}(B)}
\]

**Expected value**

Just as in the discrete case, using the joint pdf we can compute the marginal pdf’s and thus the expected values of \(X\) and \(Y\):

\[
E(X) = \int_{-\infty}^{\infty} x f_X(x) \, dx \quad \text{and} \quad E(Y) = \int_{-\infty}^{\infty} y f_Y(y) \, dy
\]

The following theorem shows how to calculated the expected value of the r.v. \(h(X, Y)\).

**Theorem 3**

*If \(X\) and \(Y\) are continuous r.v.‘s with joint pdf \(f(x, y)\) and \(h(x, y)\) is a real function of two variables, then*

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

**Example**

Let \((X, Y)\) be a point randomly drawn from the unit square \(B = \{(x, y) : 0 \leq x, y \leq 1\}\). Calculate \(P(|X - Y| \geq 1/2)\).

**Solution**

Note that \(\text{area}(B) = 1\). The probability can be calculated by integrating the pdf

\[
f(x, y) = \begin{cases} 1 & \text{if } 0 \leq x, y \leq 1 \\ 0 & \text{otherwise} \end{cases}
\]

over the region \(A = \{(x, y) : |x - y| \geq 1/2\}\).

However, since \(A \cap B\) is a union of two congruent isosceles right triangles of side length 1/2, we have

\[
P(|X - Y| \geq 1/2) = \frac{\text{area}(A \cap B)}{\text{area}(B)} = \frac{2 \cdot \frac{1}{2} \cdot \left(\frac{1}{2}\right)^2}{1} = \frac{1}{4}
\]

**Corollary 4**

*If \(X\) and \(Y\) are jointly continuous r.v.’s, then*

\[
E(X + Y) = E(X) + E(Y)
\]

**Proof**

Similar to the discrete case, but the sums are replaced with integrals. Letting \(h(x, y) = x + y\) we have

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

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

\[
= \int_{-\infty}^{\infty} x \left( \int_{-\infty}^{\infty} f(x, y) \, dy \right) \, dx + \int_{-\infty}^{\infty} y \left( \int_{-\infty}^{\infty} f(x, y) \, dx \right) \, dy
\]

\[
= E(X) + E(Y)
\]
Remark: The main theorem about the expected value of \( h(X,Y) \) can be used for both discrete and jointly continuous r.v.’s to show the following “linearity of expectation” property:

\[
E[\alpha g_1(X,Y) + \beta g_2(X,Y)] = \alpha E[g_1(X,Y)] + \beta E[g_2(X,Y)]
\]

for any two functions \( g_1, g_2 : \mathbb{R}^2 \rightarrow \mathbb{R} \).

Example: Let \( X \) and \( Y \) have joint pdf

\[
f(x,y) = \begin{cases} 
10xy^2 & 0 \leq x \leq y \leq 1 \\
0 & \text{otherwise}
\end{cases}
\]

(a) Find \( E(X^2 + Y^2) \).

Solution: Using the linearity of expectation,

\[
E(X^2 + Y^2) = E(X^2) + E(Y^2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 f(x,y) \, dx \, dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y^2 f(x,y) \, dx \, dy \\
= \int_{0}^{1} \int_{0}^{y} x^2 10xy^2 \, dx \, dy + \int_{0}^{1} \int_{0}^{y} y^2 10xy^2 \, dx \, dy \\
= \frac{5}{14} \cdot \frac{5}{7} = \frac{25}{98}
\]

(b) Find \( E(X^2Y^2) \).

Solution:

\[
E(X^2Y^2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2y^2 f(x,y) \, dx \, dy \\
= \int_{0}^{1} \int_{0}^{y} x^2y^210xy^2 \, dx \, dy \\
= 10 \int_{0}^{1} \int_{0}^{y} x^3y^4 \, dx \, dy \\
= \frac{5}{18}
\]

Note that \( E(X^2)E(Y^2) = \frac{5}{14} \cdot \frac{5}{7} = \frac{25}{98} \), so that

\( E(X^2Y^2) \neq E(X^2)E(Y^2) \).

Independence of Random Variables

Definition: Two random variables \( X \) and \( Y \) defined on the same probability space are independent if for all “reasonable” \( A, B \subset \mathbb{R} \), the events \( \{ X \in A \} \) and \( \{ Y \in B \} \) are independent, i.e.,

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

Let \( F(s,t) \) denote the joint distribution function of \( X \) and \( Y \):

\[
F(s,t) = P(X \leq s, Y \leq t)
\]

It can be shown using the axioms of probability that \( X \) and \( Y \) are independent if and only if

\[
F(s,t) = F_X(s)F_Y(t)
\]

i.e., the joint distribution function is the product of the marginal distribution functions.
The following theorem gives an easy to check characterization of independence for discrete random variables.

**Theorem 5**

Let $X$ and $Y$ be discrete r.v.’s with joint pmf $p(x, y)$. Then $X$ and $Y$ are independent if and only if

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

i.e., the joint pmf is the product of the marginal pmf’s.

**Proof** If $X$ and $Y$ are independent, then $\{X = x\}$ and $\{Y = y\}$ are independent events, so $p(x, y) = p_X(x)p_Y(y)$ follows.

The proof that $p(x, y) = p_X(x)p_Y(y)$ for all $x, y$ implies that $X$ and $Y$ are independent is left as an exercise. □

**Example:** If $X$ and $Y$ have pmf given below, are they independent?

<table>
<thead>
<tr>
<th>$y$</th>
<th>$x$</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.1</td>
<td>0.07</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.3</td>
<td>0.21</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>0.07</td>
<td>0.03</td>
<td></td>
</tr>
</tbody>
</table>

**Solution:** They are independent. Calculate the marginal pmf’s and check that $p(x, y) = p_X(x)p_Y(y)$ for all $x$ and $y$:

<table>
<thead>
<tr>
<th>$y$</th>
<th>$x$</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.1</td>
<td>0.07</td>
<td>0.03</td>
<td><strong>0.2</strong></td>
</tr>
<tr>
<td>1</td>
<td>0.3</td>
<td>0.21</td>
<td>0.09</td>
<td><strong>0.6</strong></td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>0.07</td>
<td>0.03</td>
<td><strong>0.2</strong></td>
</tr>
</tbody>
</table>

$p_X(x)$ | **0.5** | 0.35 | 0.15 |

**Example:** The joint pmf of $X$ and $Y$ is given by

<table>
<thead>
<tr>
<th>$y$</th>
<th>$x$</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.1</td>
<td>0.05</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>0.1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Are $X$ and $Y$ independent?

**Solution:** The marginal pmf’s are

$p_X(0) = 0.5$, $p_X(1) = 0.35$, $p_X(2) = 0.15$

and

$p_Y(0) = 0.2$, $p_Y(1) = 0.6$, $p_Y(2) = 0.2$

Clearly, $X$ and $Y$ are not independent since e.g.,

$p_X(2)p_Y(2) = 0.15 \cdot 0.2 = 0.03 \neq 0 = p(2,2)$

**Independence of continuous r.v.’s**

The following characterizes the independence of jointly continuous random variables in term of their pdf’s:

**Theorem 6**

Let $X$ and $Y$ be jointly continuous r.v.’s with joint pdf $f(x, y)$. Then $X$ and $Y$ are independent if and only if

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

i.e., the joint pdf is the product of the marginal pdf’s.
Example: Let $\Omega$ be the planar region defined by

$$\Omega = \{(x, y) : 0 \leq x + y \leq 1, 0 \leq x \leq 1, 0 \leq y \leq 1\}$$

and the joint pdf of $X$ and $Y$ be given by

$$f(x, y) = \begin{cases} 6x & (x, y) \in \Omega \\ 0 & \text{otherwise} \end{cases}$$

Are $X$ and $Y$ independent?

Solution: Let’s calculate $f_X(x)$ and $f_Y(y)$:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) \, dy = \begin{cases} \int_0^{1-x} 6x \, dy = 6x(1-x) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

and

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) \, dx = \begin{cases} \int_0^{1-y} 6x \, dx = 3(1-y)^2 & 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Clearly, $f(x, y) \neq f_X(x)f_Y(y)$. For example,

$$f_X(x)f_Y(y) > 0 \quad \text{for all } (x, y) \in (0, 1)^2 - \Omega$$

but

$$f(x, y) = 0 \quad \text{if } (x, y) \in (0, 1)^2 - \Omega$$

Thus $X$ and $Y$ are not independent.

Example: A point $(X, Y)$ is selected at random from the rectangle

$$R = \{(x, y) : 0 \leq x \leq a, 0 \leq y \leq b\}$$

Are $X$ are $Y$ independent?

Solution: Recall that the joint pdf is given by

$$f(x, y) = \begin{cases} \frac{1}{\text{area}(R)} = \frac{1}{ab} & (x, y) \in R \\ 0 & \text{otherwise} \end{cases}$$

Thus

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) \, dy = \begin{cases} \int_0^b \frac{1}{ab} \, dy = \frac{1}{a} & 0 \leq x \leq a \\ 0 & \text{otherwise} \end{cases}$$

Similarly

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) \, dx = \begin{cases} \int_0^a \frac{1}{ab} \, dx = \frac{1}{b} & 0 \leq y \leq b \\ 0 & \text{otherwise} \end{cases}$$

Clearly, $f(x, y) = f_X(x)f_Y(y)$ for all $x$ and $y$, so $X$ and $Y$ are independent.
**Example:** Let $X$ and $Y$ be independent exponential r.v.’s with pdf’s

$$f_X(x) = \begin{cases} \alpha e^{-\alpha x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}, \quad f_Y(y) = \begin{cases} \beta e^{-\beta y} & y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Find $P(X < Y)$.

**Solution:** Since $X$ and $Y$ are independent,

$$f(x, y) = f_X(x)f_Y(y) = \begin{cases} \alpha \beta e^{-\alpha x-\beta y} & x, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

We will calculate

$$P(X < Y) = \int \int_{(x,y): x < y} f(x, y) \, dx \, dy$$

**Some consequences of independence**

**Theorem 7**

Let $X$ and $Y$ be discrete or jointly continuous independent random variables and $h(x)$ and $g(y)$ real functions. Then $h(X)$ and $g(Y)$ are independent r.v.’s.

For example, if $X$ and $Y$ are independent, then $\sin X$ and $\cos Y$ are independent, $\cos X$ and $\sin Y$ are independent, $X^2$ and $e^{-Y}$ are independent, etc.

**Theorem 8**

Let $X$ and $Y$ be discrete or jointly continuous independent random variables and $h(x)$ and $g(y)$ real functions. Then

$$E[h(X)g(Y)] = E[h(X)]E[g(Y)]$$

**Proof** Assume $X$ and $Y$ are jointly continuous; the proof for the discrete case is similar:

$$E[h(X)g(Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x)g(y)f_X(x)f_Y(y) \, dx \, dy$$

(by independence)

$$= \int_{-\infty}^{\infty} g(y)f_Y(y) \left( \int_{-\infty}^{\infty} h(x)f_X(x) \, dx \right) \, dy$$

$$= E[h(X)] \int_{-\infty}^{\infty} g(y)f_Y(y) \, dy$$

$$= E[h(X)]E[g(Y)] \square$$
Corollary 9
If $X$ and $Y$ are independent random variables, then
$$E(XY) = E(X)E(Y)$$

Remark: The converse of the corollary is not true, i.e., the fact that $E(XY) = E(X)E(Y)$ does not imply that $X$ and $Y$ are independent.

We have
$$E(X) = E(Y) = 0$$
and
$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x,y)\,dxdy$$
$$= \int \int_{R_1 \cup R_2 \cup R_3 \cup R_4} xy\,dxdy$$
$$= \sum_{i=1}^{4} \int \int_{R_i} xy\,dxdy$$
$$= \frac{1}{2} \int_{0}^{1/2} \int_{0}^{1/2} xy\,dxdy + \int_{0}^{1/2} \int_{-1}^{-1/2} xy\,dxdy$$
$$+ \int_{-1}^{-1/2} \int_{0}^{1/2} xy\,dxdy + \int_{-1/2}^{1} \int_{-1/2}^{1} xy\,dxdy$$
$$= \left(\frac{1}{8}\right)\left(\frac{3}{8}\right) + \left(\frac{1}{8}\right)\left(-\frac{3}{8}\right) + \left(\frac{1}{8}\right)\left(-\frac{3}{8}\right) + \left(\frac{3}{8}\right)\left(\frac{1}{8}\right) = 0$$
Thus $E(X)E(Y) = E(XY)$, but $X$ and $Y$ are not independent.

Example: Let $R_1$, $R_2$, $R_3$, and $R_4$ be the square regions defined by
$$R_1 = \{0 \leq x \leq \frac{1}{2}, \frac{1}{2} \leq y \leq 1\}, R_2 = \{-1 \leq x \leq -\frac{1}{2}, 0 \leq y \leq \frac{1}{2}\}$$
$$R_3 = \{-\frac{1}{2} \leq x \leq 0, -1 \leq y \leq -\frac{1}{2}\}, R_4 = \{\frac{1}{2} \leq x \leq 1, -\frac{1}{2} \leq y \leq 0\}$$
and assume $(X,Y)$ is randomly drawn from $A = R_1 \cup R_2 \cup R_3 \cup R_4$
Then the joint pdf is
$$f(x,y) = \begin{cases} 
\frac{1}{\text{area}(A)} = 1 & (x,y) \in A \\
0 & \text{otherwise}
\end{cases}$$
It is easy to see that the marginal pdf’s are
$$f_X(x) = \begin{cases} 
\frac{1}{2} & -1 \leq x \leq 1 \\
0 & \text{otherwise}
\end{cases}$$
$$f_Y(y) = \begin{cases} 
\frac{1}{2} & -1 \leq y \leq 1 \\
0 & \text{otherwise}
\end{cases}$$
Clearly, $f(x,y) \neq f_X(x)f_Y(y)$, so $X$ and $Y$ are not independent.

Example: Let $X$ and $Y$ be independent uniform r.v.’s on the interval $(0,1)$. Find the pdf of $Z = XY$ and use it to calculate $E(Z)$.

Solution: The joint pdf $f(x,y) = f_X(x)f_Y(y)$ is given by
$$f(x,y) = \begin{cases} 
1 & 0 < x, y < 1 \\
0 & \text{otherwise}
\end{cases}$$
Since $Z = XY$ and $X,Y \in (0,1)$, we have $Z = XY \in (0,1)$, and so
$$F_Z(t) = P(Z \leq t) = \begin{cases} 
0 & t \leq 0 \\
\int_{0}^{t} P(XY \leq t) & 0 < t < 1 \\
1 & t \geq 1
\end{cases}$$
For $0 < t < 1$ we have from the geometry of the problem

$$P(XY \leq t) = P\left( Y \leq \frac{t}{X} \right) = t \cdot 1 + \int_t^1 \frac{t}{x} dy dx$$

$$= t + \int_t^1 \frac{t}{x} dx = t - t \ln t$$

Thus the pdf of $Z = XY$ is given by

$$f_Z(t) = F'_Z(t) = \begin{cases} \frac{d}{dt}(t - t \ln t) & 0 < t < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} -t \ln t & 0 < t < 1 \\ 0 & \text{otherwise} \end{cases}$$

We can calculate $E(Z)$ as

$$E(Z) = \int_{-\infty}^{\infty} tf_Z(t) dt = \int_0^1 (-t \ln t) dt$$

Using integration by parts with $u = -\ln t$ and $dv = t$, we obtain

$$E(Z) = \int_0^1 (-t \ln t) dt = \left[ -\frac{t^2}{2} \ln t \right]_0^1 + \int_0^1 \frac{t^2}{2} \cdot \frac{1}{t} dt$$

$$= 0 - 0 + \int_0^1 \frac{t}{2} dt$$

$$= \frac{1}{4}$$

Compare the above with the following simple calculation which uses the fact that $E(XY) = E(X)E(Y)$ since $X$ and $Y$ are independent:

$$E(XY) = E(X)E(Y) = \left( \int_0^1 x dx \right) \left( \int_0^1 y dy \right)$$

$$= \left( \frac{1}{2} \right)^2 = \frac{1}{4}$$

### Conditional Distributions

Recall the definition of conditional probability of an event $A$ given another event $B$ (such that $P(B) > 0$):

$$P(A|B) = \frac{P(AB)}{P(B)}$$

We want to extend this notion to a pair of random variables $X$ and $Y$. Specifically, we are interested in how the knowledge of the value of one of them “affects” the probability distribution of the other.

### Discrete distributions

Let $X$ and $Y$ be discrete random variables with joint pmf $p(x, y)$. If we don’t know anything about the value of $Y$, the probabilities concerning $X$ are calculated from the marginal pmf

$$p_X(x) = \sum_{y \in Y} p(x, y)$$

Now assume that we know that $Y = y$. Then we have extra knowledge about the probabilities concerning $X$ in the form of the conditional probabilities

$$P(X = x | Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p(x, y)}{p_Y(y)}$$
**Definition** Let $X$ and $Y$ be discrete r.v.'s with joint pmf $p(x, y)$. The conditional pmf of $X$ given $Y = y$ is defined by

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

whenever $p_Y(y) > 0$.

**Note:** For fixed $y \in \mathcal{Y}$, the function $p_{X|Y}(x|y)$ is a pmf in $x$. Indeed, $p_{X|Y}(x|y) \geq 0$ and

$$\sum_{x \in \mathcal{X}} p_{X|Y}(x|y) = \sum_{x \in \mathcal{X}} \frac{p(x, y)}{p_Y(y)} = \frac{1}{p_Y(y)} \sum_{x \in \mathcal{X}} p(x, y) = \frac{1}{p_Y(y)} p_Y(y) = 1$$

**Example:** Let $X$ be the number of spades and $Y$ the number of clubs in a randomly drawn poker hand. We have seen that

$$p(x, y) = \begin{cases} \binom{13}{x} \binom{13}{y} \binom{26}{5-x-y} \binom{52}{5} & \text{if } 0 \leq x, y \leq 5, \ x + y \leq 5 \\ 0 & \text{otherwise} \end{cases}$$

Let's calculate the conditional pmf of $X$ given $Y$. We have

$$p_Y(y) = \binom{13}{y} \binom{39}{5-y} \binom{52}{5} \quad 0 \leq y \leq 5$$

Thus

$$p_{X|Y}(x|y) = \frac{p(x, y)}{p_Y(y)} = \frac{\binom{13}{x} \binom{13}{y} \binom{26}{5-x-y} \binom{52}{5}}{\binom{13}{y} \binom{39}{5-y} \binom{52}{5}}$$

$$= \frac{\binom{13}{5-x-y} \binom{39}{y}}{\binom{52}{5-y}} \quad 0 \leq x \leq 5 - y$$

**Remarks:**

1. The conditional pmf $p_{Y|X}(y|x)$ is similarly defined.
2. The conditional pmf can be used to calculate the conditional probability of events in the form $\{X \in C\}$ given $\{Y = y\}$:

$$P(X \in C|Y = y) = \sum_{x \in C} p_{X|Y}(x|y)$$

**Proof** This is essentially the law of total probability. The formal proof is left as an exercise.

3. If $X$ and $Y$ are independent, then $p(x, y) = p_X(x)p_Y(y)$, so

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

Thus in this case $p_{X|Y}(x|y) = p_X(x)$. The converse is also true, and in fact $X$ and $Y$ are independent if and only if $p_{X|Y}(x|y) = p_X(x)$ for all $x$ and $y$ such that $p_Y(y) > 0$.

Let's calculate $P(X \leq 1|Y = 3)$:

$$P(X \leq 1|Y = 3) = \sum_{x \leq 1} p_{X|Y}(x|3)$$

$$= p_{X|Y}(0|3) + p_{X|Y}(1|3)$$

$$= \binom{13}{0} \binom{26}{2} \binom{39}{5} + \binom{13}{1} \binom{26}{1} \binom{39}{4}$$

$$\approx 0.895$$
Continuous distributions

It is not immediately clear how to condition on the value of a continuous random variable \( Y \) since \( P(Y = y) = 0 \) for all \( y \).

**Definition** Let \( X \) and \( Y \) be jointly continuous r.v.’s with joint pdf \( f(x, y) \). The *conditional pdf* of \( X \) given \( Y = y \) is defined by

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

whenever \( f_Y(y) > 0 \). Similarly, the conditional pdf of \( Y \) given \( X = x \) is

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

whenever \( f_X(x) > 0 \).

**Remarks:**
1. Just as in the discrete case, it is easy to show that \( f_{X|Y}(x|y) \) is a valid pdf for fixed \( y \):

\[
 \int_{-\infty}^{\infty} f_{X|Y}(x|y) \, dx = \int_{-\infty}^{\infty} \frac{f(x, y)}{f_Y(y)} \, dx = \frac{1}{f_Y(y)} \int_{-\infty}^{\infty} f(x, y) \, dx = \frac{1}{f_Y(y)} f_Y(y) = 1
\]

2. Similarly to the discrete case, it can be shown that \( X \) and \( Y \) are independent if and only if

\[
 f_{X|Y}(x|y) = f_X(x)
\]

for all \( x \) and \( y \) such that \( f_Y(y) > 0 \).

**Definition** For a reasonable set \( B \subset \mathbb{R} \), the conditional probability of the event \( \{Y \in B\} \) given \( X = x \) is defined by

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

**Note.** The meaning of \( P(Y \in B|X = x) \) is not as obvious as in the discrete case, since we are conditioning on the event \( \{X = x\} \) of probability zero. We will see later how useful \( P(Y \in B|X = x) \) can be.

**Example:** Let the joint pdf of \( X \) and \( Y \) be

\[
 f(x, y) = \begin{cases} 
 10xy^2 & 0 \leq x \leq y \leq 1 \\
 0 & \text{otherwise}
\end{cases}
\]

(a) Find \( f_{X|Y}(x|y) \) and \( f_{Y|X}(y|x) \).

**Solution:** We have seen that

\[
 f_Y(y) = \begin{cases} 
 5y^4 & 0 \leq y \leq 1 \\
 0 & \text{otherwise}
\end{cases}
\]

We have \( f_Y(y) > 0 \) if \( y \in (0, 1) \). Thus for all \( y \in (0, 1] \),

\[
 f_{X|Y}(x|y) = f(x, y) f_Y(y) = \begin{cases} 
 10xy^2 & 0 \leq x < y \\
 \frac{2x}{y^2} & 0 \leq x < y \\
 0 & \text{otherwise}
\end{cases}
\]
Previously we have also calculated $f_X(x)$:

$$f_X(x) = \begin{cases} \frac{10}{3} x (1 - x^3) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Thus for $x \in (0, 1)$ (so that $f_X(x) > 0$),

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} = \begin{cases} \frac{10 xy^2}{\frac{10}{3} x (1 - x^3)} = \frac{3 y^2}{1 - x^3} & x \leq y < 1 \\ 0 & \text{otherwise} \end{cases}$$

The conditional probability $P(Y \in B|X = x)$ can be very useful in calculating probabilities via the following version of the law of total probability:

**Theorem 10 (Law of total probability)**

*Let $X$ and $Y$ be discrete random variables. Then*

$$P(Y \in B) = \sum_{x \in X} P(Y \in B|X = x)p_X(x)$$

*If $X$ and $Y$ are jointly continuous random variables, then*

$$P(Y \in B) = \int_{-\infty}^{\infty} P(Y \in B|X = x)f_X(x) \, dx$$

(b) Find $P(X \leq \frac{1}{2}|Y = \frac{3}{4})$.

**Solution:** We have

$$f_{X|Y}(x|\frac{3}{4}) = \begin{cases} \frac{2x}{\left(\frac{3}{4}\right)^2} = \frac{32}{9} x & 0 \leq x \leq \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

Thus

$$P(X \leq \frac{1}{2}|Y = \frac{3}{4}) = \int_{-\infty}^{1/2} f_{X|Y}(x|y) \, dx$$

$$= \frac{32}{9} \int_{0}^{1/2} x \, dx$$

$$= \frac{32}{9} \cdot \frac{1}{8} = \frac{4}{9}$$

**Proof** We only do the continuous case:

$$\int_{-\infty}^{\infty} P(Y \in B|X = x)f_X(x) \, dx$$

$$= \int_{-\infty}^{\infty} \left( \int_{B} f_{Y|X}(y|x) \, dy \right) f_X(x) \, dx$$

$$= \int_{B} \left( \int_{-\infty}^{\infty} f_{Y|X}(y|x) f_X(x) \, dx \right) dy$$

$$= \int_{B} f_Y(y) \, dy$$

$$= P(Y \in B) \quad \Box$$
**Example:** Suppose $X$ is uniformly distributed on $[0,1]$. Given $X = x$, let $Y$ be a random point in the interval $[0,x]$. Calculate the probability $P(Y \geq 1/2)$.

**Solution:** We could calculate $f_{X,Y}(x,y)$ and obtain $P(Y \geq 1/2)$ from a double integral. Instead, we will use the law of total probability.

We have

$$f_{Y|X}(y|x) = \begin{cases} \frac{1}{x} & 0 \leq y \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

and so

$$P(Y \geq 1/2|X = x) = \begin{cases} \frac{x - 1/2}{x} & 1/2 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Thus from the law of total probability we have

$$P(Y \geq 1/2) = \int_{-\infty}^{\infty} P(Y \geq 1/2|X = x) f_X(x) \, dx$$

$$= \int_{1/2}^{1} \frac{x - 1/2}{x} \, dx$$

$$= \left[ x - \frac{1}{2} \ln x \right]_{1/2}^{1}$$

$$= \frac{1}{2} (1 - \ln 2)$$

---

**Conditional Expectation**

We can generalize the expected value to conditional distributions in the following way.

**Definition** Let $X$ and $Y$ be random variables defined on the same sample space. If $X$ and $Y$ are discrete, then the **conditional expectation** of $X$ given $Y = y$ is defined by

$$E(X|Y = y) = \sum_{x \in \mathcal{X}} x p_{X|Y}(x|y)$$

whenever $p_Y(y) > 0$. If $X$ and $Y$ are jointly continuous, the corresponding definition is

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

whenever $f_Y(y) > 0$.

**Note:** For fixed $y$, the conditional expectation of $X$ given $Y = y$ is simply an expectation calculated according to the *conditional distribution* of $X$ given $Y = y$.

Thus all the basic properties we derived for (unconditional) expectations still hold. For example,

$$E(h(X)|Y = y) = \sum_{x \in \mathcal{X}} h(x) p_{X|Y}(x|y)$$

for discrete r.v.’s, and

$$E(h(X)|Y = y) = \int_{-\infty}^{\infty} h(x) f_{X|Y}(x|y) \, dx$$

for continuous r.v.’s.
Example: Let the joint pdf of $X$ and $Y$ be

$$f(x, y) = \begin{cases} 10xy^2 & 0 \leq x \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find $E(X|Y = y)$.

Solution: We previously derived

$$f_{X|Y}(x|y) = \begin{cases} \frac{2x}{y^2} & 0 \leq x < y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Thus for all $y \in [0, 1]$ we need to calculate

$$E(X|Y = y) = \int_0^y x f_{X|Y}(x|y) \, dx = \int_0^y x \frac{2x}{y^2} \, dx$$

We obtain for all $y \in [0, 1]$,

$$E(X|Y = y) = \int_0^y \frac{2x^2}{y^2} \, dx = 2 \int_0^y x^2 \, dx$$
$$= \frac{2}{y^2} \cdot \frac{y^3}{3} = \frac{2}{3} y$$

(b) Find $E(Y|X = x)$.

Solution: From previous calculations,

$$f_{Y|X}(y|x) = \begin{cases} \frac{3y^2}{1 - x^3} & 0 \leq x < y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Hence for $x \in [0, 1)$

$$E(Y|X = x) = \int_0^1 y f_{Y|X}(y|x) \, dy = \int_x^1 \frac{3y^3}{1 - x^3}$$
$$= \frac{3}{1 - x^3} \int_x^1 y^3 \, dy$$
$$= \frac{3}{1 - x^3} \cdot \frac{1 - x^4}{4}$$
$$= \frac{3}{4} \cdot \frac{1 - x^4}{1 - x^3}$$

Conditional expectation can greatly simplify the calculation of expected value.

**Theorem 11 (Law of total expectation)**

Let $X$ and $Y$ be random variables defined on the same sample space. If $X$ and $Y$ are discrete, then

$$E(Y) = \sum_{x \in X} E(Y|X = x) \mu_X(x)$$

If $X$ and $Y$ are jointly continuous, then

$$E(Y) = \int_{-\infty}^{\infty} E(Y|X = x) f_X(x) \, dx$$

**Note:** The proof is very similar to the proof of the law of total probability and is left as an exercise. The theorem also holds if the roles of $X$ and $Y$ are exchanged.
Example: Suppose $X$ is uniformly distributed on $[0, 1]$. Given $X = x$, let $Y$ be a random point in the interval $[0, x]$. Calculate $E(Y)$.

Solution: Since $Y$ is uniformly distributed on $[0, x]$ given $X = x$, we have

$$E(Y | X = x) = \frac{x}{2}$$

Thus by the law of total expectation

$$E(Y) = \int_{-\infty}^{\infty} E(Y | X = x) f_X(x) \, dx = \int_0^1 \frac{x}{2} \, dx = \frac{1}{4}$$

Exercise: Calculate $E(Y)$ by finding $f_Y(y)$ first. Which solution is simpler?

1. The mapping $(h_1, h_2) : \mathbb{R}^2 \to \mathbb{R}^2$ is invertible, i.e., for any $u$ and $v$ the pair of equations

$$h_1(x, y) = u, \quad h_2(x, y) = v$$

has at most one solution

$$x = g_1(u, v), \quad y = g_2(u, v)$$

2. The functions $h_1$ and $h_2$ have continuous partial derivatives

$$\frac{\partial h_1}{\partial x} \quad \frac{\partial h_1}{\partial y} \quad \frac{\partial h_2}{\partial x} \quad \frac{\partial h_2}{\partial y}$$

3. The Jacobian of $(h_1, h_2)$ is nonzero everywhere, i.e., the following determinant is nonzero for all $x$ and $y$:

$$J(x, y) = \det \begin{bmatrix} \frac{\partial h_1}{\partial x} & \frac{\partial h_1}{\partial y} \\ \frac{\partial h_2}{\partial x} & \frac{\partial h_2}{\partial y} \end{bmatrix} = \frac{\partial h_1}{\partial x} \frac{\partial h_2}{\partial y} - \frac{\partial h_1}{\partial y} \frac{\partial h_2}{\partial x} \neq 0$$

Let $\tilde{J}$ denote the Jacobian of $(g_1, g_2)$, i.e.,

$$\tilde{J}(u, v) = \det \begin{bmatrix} \frac{\partial g_1}{\partial u} & \frac{\partial g_1}{\partial v} \\ \frac{\partial g_2}{\partial u} & \frac{\partial g_2}{\partial v} \end{bmatrix} = \frac{1}{J(g_1(u, v), g_2(u, v))}$$

Theorem 12 (Joint pdf of transformed random variables)

Assume $X$ and $Y$ have joint pdf $f_{X,Y}(x, y)$. If $h_1$ and $h_2$ satisfy the above conditions, then the joint pdf of $U = h_1(X, Y)$ and $V = h_2(X, Y)$ is given by

$$f_{U,V}(u, v) = f_{X,Y}(g_1(u, v), g_2(u, v))|\tilde{J}(u, v)|$$

Remark: The proof of the theorem relies on the change of variable formula for double integrals.
**Example:** Let $X$ and $Y$ be independent standard normal r.v.’s. Let $R$ and $\Theta$ be the polar coordinates of the point $(X, Y)$, i.e.,

$$R = \sqrt{X^2 + Y^2}, \quad \Theta = \arctan(Y/X)$$

Find the joint pdf of $R$ and $\Theta$.

**Solution:** We have $h_1(x, y) = \sqrt{x^2 + y^2}$ and $h_2(x, y) = \arctan(y/x)$.

We have to solve

$$r = \sqrt{x^2 + y^2}, \quad \theta = \arctan(y/x)$$

for $x$ and $y$. The solution (of course) is the well-known expression

$$x = r \cos \theta = g_1(r, \theta), \quad y = r \sin \theta = g_2(r, \theta)$$

Note that $-\infty < x < \infty$ and $-\infty < y < \infty$, while $0 \leq r < \infty$ and $0 \leq \theta < 2\pi$.

Since $X$ and $Y$ are independent,

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

$$= \frac{1}{2\pi} e^{-(x^2+y^2)/2}$$

Therefore

$$f_{R,\Theta}(r, \theta) = f_{X,Y}(g_1(r, \theta), g_2(r, \theta))|J(r, \theta)|$$

$$= f_{X,Y}(r \cos \theta, r \sin \theta)|J(r, \theta)|$$

$$= \frac{1}{2\pi} e^{-r^2/2} r$$

for all $r > 0$ and $0 \leq \theta < 2\pi$.

The partial derivatives of $g_1$ and $g_2$ are

$$\frac{\partial g_1}{\partial r} = \frac{\partial}{\partial r} (r \cos \theta) = \cos \theta, \quad \frac{\partial g_1}{\partial \theta} = -r \sin \theta$$

and

$$\frac{\partial g_2}{\partial r} = \frac{\partial}{\partial r} (r \sin \theta) = \sin \theta, \quad \frac{\partial g_2}{\partial \theta} = r \cos \theta$$

Thus the Jacobian of $(g_1, g_2)$ is

$$J(r, \theta) = \begin{vmatrix} \cos \theta & -r \sin \theta \\ r \sin \theta & r \cos \theta \end{vmatrix} = r \cos^2 \theta + r \sin^2 \theta = r$$

since $\cos^2 \theta + \sin^2 \theta = 1$.

In conclusion, we obtained that

$$f_{R,\Theta}(r, \theta) = \begin{cases} \frac{1}{2\pi} r e^{-r^2/2} & \text{if } r > 0, \ 0 \leq \theta < 2\pi \\ 0 & \text{otherwise} \end{cases}$$

The marginal pdf’s are

$$f_R(r) = \begin{cases} re^{-r^2/2} & \text{if } r > 0 \\ 0 & \text{otherwise} \end{cases} \quad f_{\Theta}(\theta) = \begin{cases} \frac{1}{2\pi} & \text{if } 0 \leq \theta < 2\pi \\ 0 & \text{otherwise} \end{cases}$$

Since $f_{R,\Theta}(r, \theta) = f_R(r)f_{\Theta}(\theta)$, the random variables $R$ and $\Theta$ are independent. Also note that $\Theta$ is uniformly distributed on $[0, 2\pi)$.
Sum of two independent random variables

Assume $X$ and $Y$ are independent and let $U = X + Y$. We want to find the pdf of $U$ in terms of the pdf’s of $X$ and $Y$.

One way to do this is to calculate the joint pdf $f_{U,V}(u,v)$ of the pair $U = X + Y$, $V = X - Y$, and then find $f_U(u)$ as the marginal of $f_{U,V}(u,v)$.

Instead, we follow a more direct approach:

\[
F_U(u) = P(U \leq u) = P(X + Y \leq u) = \int \int_{x+y\leq u} f_{X,Y}(x,y) \, dx \, dy = \int_{-\infty}^{\infty} \int_{-\infty}^{u-y} f_X(x)f_Y(y) \, dx \, dy
\]

Thus

\[
f_U(u) = F'_U(u) = \frac{d}{du} \int_{-\infty}^{\infty} \int_{-\infty}^{u-y} f_X(x)f_Y(y) \, dx \, dy
\]

We obtained

\[
f_U(u) = \int_{-\infty}^{\infty} f_X(u-y)f_Y(y) \, dy
\]

By symmetry, the same proof also implies that

\[
f_U(u) = \int_{-\infty}^{\infty} f_X(x)f_Y(u-x) \, dx
\]

Remark: The integral

\[
\int_{-\infty}^{\infty} f_X(u-y)f_Y(y) \, dy
\]

is called the convolution of $f_X(x)$ and $f_Y(y)$ and is denoted by $f_X \ast f_Y$.

Example: Let $X$ and $Y$ be independent uniform r.v.’s on $[0,1]$. Find the pdf of $U = X + Y$.

Solution: We have $f_U = f_X \ast f_Y$, where

\[
f_X(t) = f_Y(t) = \begin{cases} 1 & 0 \leq t \leq 1 \\ 0 & \text{otherwise} \end{cases}
\]

Thus

\[
f_U(u) = \int_{-\infty}^{\infty} f_X(u-y)f_Y(y) \, dy = \int_{0}^{1} f_X(u-y) \, dy
\]

If $u < 0$ or $u > 2$, then $f_U(u) = 0$. For $0 \leq u < 1$ we have

\[
\int_{0}^{1} f_X(u-y) \, dy = \int_{0}^{u} dy = u
\]

For $1 \leq u \leq 2$,

\[
\int_{0}^{1} f_X(u-y) \, dy = \int_{u-1}^{1} dy = 2 - u
\]

Putting these together, we obtain

\[
f_U(u) = \begin{cases} u & 0 \leq u < 1 \\ 2 - u & 1 \leq u \leq 2 \\ 0 & \text{otherwise} \end{cases}
\]

This function is called the triangular pdf.