Expectation of a Function of Two RVs and Joint Moments

Section 5.6
**Expectation of a Function of 2 RVs**

- **Recall:** If $Y$ is a function of a random variable $X$, i.e. $Y = g(X)$, then

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

- In a similar way, we can write a random variable $Z$ as a function of 2 random variables $X$ and $Y$:

$$Z = g(X,Y)$$

- In this case, we have:

$$E[Z] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{XY}(x, y) \, dx \, dy$$
Example: Linearity 1

- Let $Z = g(X, Y) = aX + bY$ for some constants $a, b$.

$$E[Z] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{XY}(x, y) \, dx \, dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (ax + by) f_{XY}(x, y) \, dx \, dy$$

$$= a \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf_{XY}(x, y) \, dx \, dy + b \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf_{XY}(x, y) \, dx \, dy$$

$$= a \int_{-\infty}^{\infty} x \left[ \int_{-\infty}^{\infty} f_{XY}(x, y) \, dy \right] \, dx + b \int_{-\infty}^{\infty} y \left[ \int_{-\infty}^{\infty} f_{XY}(x, y) \, dx \right] \, dy$$

$$= a \int_{-\infty}^{\infty} xf_X(x) \, dx + b \int_{-\infty}^{\infty} yf_Y(y) \, dy = aE[X] + bE[Y]$$

Expectation is linear!
Example: Linearity 2

- In fact, we can show that if $Z = \sum_{i=1}^{n} a_i g_i(X, Y)$ then

\[
E[Z] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left( \sum_{i=1}^{n} a_i g_i(x, y) f_{XY}(x, y) \right) dx \, dy
\]

\[
= \sum_{i=1}^{n} a_i \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_i(x, y) f_{XY}(x, y) \, dx \, dy \right]
\]

\[
= \sum_{i=1}^{n} a_i E[g_i(X, Y)]
\]

Expectation is linear over the sum of many functions.
Ex: Joint Uniform RVs

- The joint PDF is
  \[ f_{X,Y}(x, y) = \begin{cases} 
    1 & 0 \leq x, y \leq 1 \\
    0 & \text{elsewhere}
  \end{cases} \]

- Let \( Z = g(X, Y) = X^2 + Y^2 \)

\[
E[Z] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) \, dx \, dy = \int_{0}^{1} \int_{0}^{1} (x^2 + y^2) \, dx \, dy
\]

\[
= E[X^2] + E[Y^2] = 2E[X^2]
\]

\[
= 2\int_{0}^{1} x^2 \, dx = 2 \left[ \frac{x^3}{3} \right]_{x=0}^{1} = \frac{2}{3}
\]

\( X \) and \( Y \) have the same marginal distribution, so …
Joint Moments of Two RVs
Joint Moments

- Let $X$, $Y$ be 2 RVs.
- The $j$-$k^{th}$ joint moment of $X$ and $Y$ is defined as
  \[ m_{j,k} = E[X^j Y^k] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^j y^k f_{XY}(x, y) \, dx \, dy \]
- Some we already know are
  \[ m_{1,0} = E[X], \; m_{0,1} = E[Y], \; m_{2,0} = E[X^2], \; m_{0,2} = E[Y^2]. \]
- The most important new one is the correlation of $X$ and $Y$: \[ m_{1,1} = E[XY] \]
- The correlation of $X$ and $Y$ is a single real value that tries to measure how closely related $X$ and $Y$ are.
Example: Correlation of Uniform RVs

- Let \( f_{X,Y}(x, y) = \begin{cases} 1 & 0 \leq x, y \leq 1 \\ 0 & \text{elsewhere} \end{cases} \)

- The correlation of \( X \) and \( Y \) is

\[
E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf_{XY}(x, y) \, dx \, dy
\]

\[
= \int_{0}^{1} \int_{0}^{1} xy \cdot 1 \, dx \, dy = \int_{0}^{1} \left[ \frac{1}{2} x^2 \right]_0^1 \, y \, dy
\]

\[
= \int_{0}^{1} \frac{1}{2} y \, dy = \frac{1}{2} \left[ \frac{1}{2} y^2 \right]_0^1 = \frac{1}{4}
\]
Joint Central Moments

- The \( j\)-th joint central moment of \( X \) and \( Y \) is defined as

\[
\mu_{j,k} = E[(X - E[X])^j (Y - E[Y])^k]
\]

- Two of these we already know are \( \mu_{2,0} = \sigma_X^2 \), \( \mu_{0,2} = \sigma_Y^2 \).

- The most important new one is the covariance of \( X \) and \( Y \):

\[
\text{COV}(X,Y) = \mu_{1,1} = E[(X - E[X])(Y - E[Y])]
\]

- Like correlation, covariance is a single real value that measures how closely related \( X \) and \( Y \) are, but the covariance removes the effect of the means.
Covariance of $X$ and $Y$

- The covariance of $X$ and $Y$ is defined as

$$\text{COV}(X,Y) = \mu_{1,1} = E[(X - E[X])(Y - E[Y])]$$

- From linearity:


- So,

$$\text{COV}(X,Y) = E[XY] - E[X]E[Y]$$
Example: Covariance of Uniform RVs

- Let \( f_{X,Y}(x, y) = \begin{cases} 1 & 0 \leq x, y \leq 1 \\ 0 & \text{elsewhere} \end{cases} \)

- We just showed the correlation of \( X \) and \( Y \) is \( E[XY] = \frac{1}{4} \)

- We showed last time that \( X \) and \( Y \) are each uniform on \([0,1]\).

- So \( E[X] = E[Y] = \frac{1}{2} \)

- And \( \text{COV}(X,Y) = E[XY] - E[X]E[Y] = \frac{1}{4} - \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = 0 \)
Covariance of $X$ and $X$

- Note that “co-variance” is a good name for this since we see that

$$\text{COV}(X, Y) = E[XY] - E[X]E[Y]$$

implies

$$\text{COV}(X, X) = E[XX] - E[X]E[X]$$

$$\text{COV}(X, X) = E[X^2] - (E[X])^2 = \sigma_X^2$$

- The covariance of $X$ with itself is the variance of $X$.  

10/30/2014
Uncorrelated RVs

- **Definition**: RVs $X$ and $Y$ are **uncorrelated** if
  \[
  \text{COV}(X, Y) = E[XY] - E[X]E[Y] = 0
  \]

- Equivalently, $X$ and $Y$ are **uncorrelated** if
  \[
  E[XY] = E[X]E[Y]
  \]

- **Example**: We just showed that joint uniform $X, Y$ are uncorrelated, when we showed
  \[
  \text{COV}(X, Y) = E[XY] - E[X]E[Y] = \frac{1}{4} - \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = 0
  \]
Independent implies Uncorrelated

- \( X \) and \( Y \) are uncorrelated iff \( E[XY] = E[X]E[Y] \)

- **Fact:** If \( X \) and \( Y \) are independent, then they are uncorrelated:

\[
E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x, y) \, dx \, dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_X(x) f_Y(y) \, dx \, dy
\]

\[
= \left[ \int_{-\infty}^{\infty} xf_X(x) \, dx \right] \left[ \int_{-\infty}^{\infty} yf_Y(y) \, dy \right] = E[X]E[Y]
\]

- The joint uniform example illustrates that since \( X \) and \( Y \) are independent there.
Fact About Independent RVs

- **Fact:** If $X$ and $Y$ are independent RVs and we let $W = g(X)$ and $Z = h(Y)$ for some functions $g$ and $h$, then $W$ and $Z$ are also independent.

- So it is also true that if $X$ and $Y$ are independent RVs and we let $W = g(X)$ and $Z = h(Y)$ for some functions $g$ and $h$, then $W$ and $Z$ are uncorrelated.
Example: Expectation of a Function

- Find $E[X^2 e^Y]$ where $X$ and $Y$ are independent RVs, $X$ is zero-mean Gaussian with $\sigma_X^2 = 4$, and $Y$ is a uniform RV on the interval $[0,3]$.

- Let $W = g(X) = X^2$ and $Z = h(Y) = e^Y$.

- $W$ and $Z$ are uncorrelated, so

$$E[X^2 e^Y] = E[X^2] E[e^Y] = 4 E[e^Y]$$

$$= 4 E[e^Y] = 4 \int_{-\infty}^{\infty} e^y f_Y(y) \, dy = 4 \int_{0}^{3} e^y \left( \frac{1}{3} \right) \, dy$$

$$= \left( \frac{4}{3} \right) [e^3 - e^0] = \frac{4}{3} (e^3 - 1)$$
The Correlation Coefficient
Correlation Coefficient

- When we defined the jointly Gaussian PDF, we defined this parameter called the correlation coefficient:

\[ \rho_{XY} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \]

- We now see that \( \rho_{XY} = \frac{\text{COV}(X,Y)}{\sigma_X \sigma_Y} \)

- \( \rho_{XY} \) normalizes the covariance so that it is unaffected by the variances of \( X \) and \( Y \).

- In particular, we always have \( -1 \leq \rho_{XY} \leq 1 \).
Intuitive Meaning of Correlation

- The correlation coefficient was defined by
  \[ \rho_{XY} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} = \frac{\text{COV}(X,Y)}{\sigma_X \sigma_Y} \]

- And we always have \(-1 \leq \rho_{XY} \leq 1\).

- \(\text{COV}(X,Y)\) and \(\rho_{XY}\) measure the linear relationship between \(X\) and \(Y\).

- A positive value indicates that, if \(X\) is above its mean, then \(Y\) is likely to be above its mean.

- A negative value indicates that, if \(X\) is above its mean, then \(Y\) is likely to be below its mean.
Examples of Correlation

- Pairs we would expect to have positive correlation:
  - Complexity of an algorithm and the size of a VLSI implementation of that algorithm.
  - Image quality and bandwidth to transmit.
  - Height of parent and child.
  - Income and education.

- Pairs we would expect to have negative correlation:
  - Weight of a car and miles per gallon.
  - Bit rate of Internet access channel and download time for a given file.
  - Education and time in prison.
  - TV viewing and grades.
Example of Positive Correlation

- This type of graph is called a scatterplot.
- Note that the slope of the blue line is positive.
Example of Negative Correlation

<table>
<thead>
<tr>
<th>Person</th>
<th>GPA</th>
<th>TV hours per week</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>3.1</td>
<td>14</td>
</tr>
<tr>
<td>#2</td>
<td>2.4</td>
<td>10</td>
</tr>
<tr>
<td>#3</td>
<td>2.0</td>
<td>20</td>
</tr>
<tr>
<td>#4</td>
<td>3.8</td>
<td>7</td>
</tr>
<tr>
<td>#5</td>
<td>2.2</td>
<td>25</td>
</tr>
<tr>
<td>#6</td>
<td>3.4</td>
<td>9</td>
</tr>
<tr>
<td>#7</td>
<td>2.9</td>
<td>15</td>
</tr>
<tr>
<td>#8</td>
<td>3.2</td>
<td>13</td>
</tr>
<tr>
<td>#9</td>
<td>3.7</td>
<td>4</td>
</tr>
<tr>
<td>#10</td>
<td>3.5</td>
<td>21</td>
</tr>
</tbody>
</table>

- Note that the slope of the blue line is negative here.
Example: Covariance

- Consider the case where $X$ is a RV and $Y = aX + b$.

$$\text{COV}(X, Y) = E[XY] - E[X]E[Y]$$

$$= E[X(aX + b)] - E[X]E[aX + b]$$

$$= aE[X^2] + bE[X] - a(E[X])^2 - bE[X]$$

$$= aE[X^2] - a(E[X])^2$$

$$= a\sigma_X^2$$
Example: Correlation Coefficient

- Just showed that if $Y = aX + b$ then $\text{COV}(X, Y) = a\sigma_X^2$

\[
\sigma_Y^2 = E[(aX + b)^2] - (E[aX + b])^2
\]

\[
= E[a^2X^2 + 2abX + b^2] - (aE[X] + b)^2
\]

\[
= a^2E[X^2] + 2abE[X] + b^2 - a^2(E[X])^2 - 2abE[X] - b^2
\]

\[
= a^2E[X^2] - a^2(E[X])^2 = a^2\sigma_X^2
\]

- So $\sigma_Y = \sqrt{a^2\sigma_X^2} = |a|\sigma_X$

- Thus

\[
\rho_{XY} = \frac{\text{COV}(X, Y)}{\sigma_X \sigma_Y} = \frac{a\sigma_X^2}{\sigma_X |a|\sigma_X} = \frac{a}{|a|} = \begin{cases} 
1 & a > 0 \\
-1 & a < 0
\end{cases}
\]
Ex: Correlation Coefficient continued

- We just showed that if $Y = aX + b$ then
  \[
  \rho_{XY} = \begin{cases} 
  1 & a > 0 \\
  -1 & a < 0 
  \end{cases}
  \]

- Thus $X$ and $Y = aX + b$ are completely correlated.
  
  ➢ If $a$ is positive, then the correlation is completely positive: $\rho_{XY} = 1$.
  
  ➢ If $a$ is negative, then the correlation is completely negative: $\rho_{XY} = -1$.

- This makes sense since $Y$ is completely predictable if we know $a$, $b$, and the value of $X$. 
Uncorrelated vs. Independent RVs

- **Fact:** If $X$ and $Y$ are independent, then they are uncorrelated: $E[XY] = E[X]E[Y]$

- In general, the reverse is not true.

- One special case: If $X$ and $Y$ are jointly Gaussian and uncorrelated, then they are independent.

  ➢ We proved this last time: For Jointly Gaussian, $\rho_{XY} = 0$ implies $X$ and $Y$ are independent.
Example: Uncorrelated, Not Independent

- Let $X$ and $Y$ be RVs with

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{2} & |x| + |y| \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \, f_{XY}(x, y) \, dx \, dy$$

$$= \frac{1}{2} \int_{-1}^{0} \int_{-1-y}^{1+y} xy \, dx \, dy + \frac{1}{2} \int_{0}^{1} \int_{y-1}^{1-y} xy \, dx \, dy = 0.$$

These 2 are negatives of each other since everything is the same except $x$ is negated.
Ex: Uncorrelated, Not Independent continued

- $X$ and $Y$ have PDF

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{2} & |x| + |y| \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

- Now let's compute the marginal

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) \, dy = \begin{cases} \int_{x-1}^{1-x} \frac{1}{2} \, dy & 0 \leq x \leq 1 \\ \int_{-x-1}^{x+1} \frac{1}{2} \, dy & -1 \leq x \leq 0 \end{cases}$$

$$= \begin{cases} \frac{1}{2} \left[1 - x - (x - 1)\right] & 0 \leq x \leq 1 \\ \frac{1}{2} \left[x + 1 - (-x - 1)\right] & -1 \leq x \leq 0 \end{cases} = \begin{cases} 1 - x & 0 \leq x \leq 1 \\ 1 + x & -1 \leq x \leq 0 \end{cases}$$

- So $f_X(x) = 1 - |x|$ for $-1 \leq x \leq 1$. 
Ex: Uncorrelated, Not Independent  continued

- $X$ and $Y$ have PDF
  \[ f_{X,Y}(x, y) = \begin{cases} \frac{1}{2} & |x| + |y| \leq 1 \\ 0 & \text{elsewhere} \end{cases} \]

- Just showed $f_X(x) = \begin{cases} 1 - |x| & -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$.

- Since $f_X(x)$ is symmetric, we have $E[X] = 0$

- So $E[XY] = 0 = E[X]E[Y]$

- $X$ and $Y$ are uncorrelated!
Ex: Uncorrelated, Not Independent  

- $X$ and $Y$ have PDF

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{2} & |x| + |y| \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

$$f_X(x) = \begin{cases} 1 - |x| & -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- By symmetry of $X$ and $Y$, we have

$$f_Y(y) = \begin{cases} 1 - |y| & -1 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- Take $x = y = \frac{3}{4}$

- Then $f_{X,Y}(x, y) = 0$, while $f_X(x)f_Y(y) = \frac{1}{4}\left(\frac{1}{4}\right) = \frac{1}{16} \neq 0$

- $X$ and $Y$ are uncorrelated, but NOT independent!
Example from Last Time

\[ f_{X,Y}(x, y) = \begin{cases} 
  x + y & 0 < x, y \leq 1 \\
  0 & \text{elsewhere}
\end{cases} \]

Anyone want to guess \( \rho_{XY} \)?

\[
E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x, y) \, dx \, dy
\]

\[
= \int_{0}^{1} \int_{0}^{1} xy(x + y) \, dx \, dy = \int_{0}^{1} \int_{0}^{1} (x^2 y + xy^2) \, dx \, dy
\]

\[
= \int_{0}^{1} \left[ \frac{x^3 y}{3} + \frac{x^2 y^2}{2} \right]_{x=0}^{1} \, dy = \int_{0}^{1} \left[ \frac{y}{3} + \frac{y^2}{2} \right] \, dy = \left[ \frac{y^2}{6} + \frac{y^3}{6} \right]_{0}^{1}
\]

\[
= \left[ \frac{1}{6} + \frac{1}{6} \right] = \frac{1}{3}
\]
Example from Last Time continued

- We showed last time that the marginal PDFs are

\[
f_X(x) = \begin{cases} 
  x + \frac{1}{2} & 0 < x \leq 1 \\
  0 & \text{else}
\end{cases} \quad f_Y(y) = \begin{cases} 
  y + \frac{1}{2} & 0 < y \leq 1 \\
  0 & \text{else}
\end{cases}
\]

\[
E[X] = \int_{-\infty}^{\infty} x f_X(x) \, dx = \int_{0}^{1} x(x + \frac{1}{2}) \, dx = \int_{0}^{1} \left( x^2 + \frac{1}{2} x \right) \, dx
\]

\[
= \left[ \frac{x^3}{3} + \frac{x^2}{4} \right]_0^1 = \left[ \frac{1}{3} + \frac{1}{4} \right] = \frac{7}{12}
\]

- Similarly, \( E[Y] = \frac{7}{12} \)
Example from Last Time continued

\[
\text{COV}(X, Y) = E[XY] - E[X]E[Y] = \frac{1}{3} - \left(\frac{7}{12}\right)^2 \\
= \frac{1(4)12 - 49}{144} = \frac{-1}{144}
\]

\[
E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) \, dx = \int_{0}^{1} x^2(x + \frac{1}{2}) \, dx = \int_{0}^{1} \left( x^3 + \frac{1}{2} x^2 \right) \, dx \\
= \left[ \frac{x^4}{4} + \frac{x^3}{6} \right]_0^{1} = \left[ \frac{1}{4} + \frac{1}{6} \right] = \frac{5}{12}
\]

\[
\sigma_Y^2 = \sigma_X^2 = E[X^2] - E[X]^2 = \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{5(12) - 49}{144} = \frac{11}{144}
\]
Example from Last Time continued

- We have shown:

\[
\text{COV}(X, Y) = \frac{-1}{144}
\]

\[
\sigma_Y^2 = \sigma_X^2 = \frac{11}{144}
\]

- So, \[\rho_{XY} = \frac{\text{COV}(X, Y)}{\sigma_X \sigma_Y} = \frac{-1}{11}\]

So \(X\) and \(Y\) have a mild negative correlation.
2nd Moments of Sums
2nd Moments of Sums

- Let \( Z = g(X, Y) = X + Y \).

- Find \( E[Z^2] = E[(X + Y)^2] \).

\[
E[Z^2] = E[(X + Y)^2] = E[X^2 + 2XY + Y^2]
\]

\[
= E[X^2] + 2E[XY] + E[Y^2] \quad \text{by linearity.}
\]
2nd Moments of Sums continued

- Let \( Z = g(X, Y) = X + Y \).
- Find the variance of \( Z = X + Y \).

\[
\]
\[
\]
\[
\]
- So \( \text{VAR}[Z] = \text{VAR}[X + Y] = \sigma_X^2 + \sigma_Y^2 + 2\text{COV}[XY] \)
2nd Moments of Sums continued

- Let $Z = g(X, Y) = X + Y$.

- We just showed:

  \[(1) \quad \text{VAR}[Z] = \text{VAR}[X + Y] = \sigma_x^2 + \sigma_y^2 + 2\text{COV}[XY]\]

- Under what condition does the variance of the sum equal the sum of the variances?

- We see from (1) that

  $\text{VAR}[X + Y] = \sigma_x^2 + \sigma_y^2$ if and only if $\text{COV}[XY] = 0$.

- In particular, if $X, Y$ are independent, then

  $\text{VAR}[X + Y] = \sigma_x^2 + \sigma_y^2$. 