1. Wrap-up of Lecture 16

Proof of Lemma 16.8. It suffices to prove that

\[ E(X_1 + \cdots + X_n) = n\mu \]
\[ \text{Var}(X_1 + \cdots + X_n) = n\sigma^2. \]

We prove this by induction. Indeed, this is obviously true when \( n = 1 \). Suppose it is OK for all integers \( \leq n - 1 \). We prove it for \( n \).

\[
E(X_1 + \cdots + X_n) = E(X_1 + \cdots + X_{n-1}) + EX_n \\
= (n-1)\mu + EX_n,
\]

by the induction hypothesis. Because \( EX_n = \mu \), the preceding is equal to \( n\mu \), as planned. Now we verify the more interesting variance computation.

Once again, we assume the assertion holds for all integers \( \leq n - 1 \), and strive to check it for \( n \).

Define

\[ Y = X_1 + \cdots + X_{n-1}. \]

Because \( Y \) is independent of \( X_n \), \( \text{Cov}(Y, X_n) = 0 \). Therefore, by Lecture 15,

\[
\text{Var}(X_1 + \cdots + X_n) = \text{Var}(Y + X_n) \\
= \text{Var}(Y) + \text{Var}(X_n) + \text{Cov}(Y, X_n) \\
= \text{Var}(Y) + \text{Var}(X_n).
\]

We know that \( \text{Var}(X_n) = \sigma^2 \), and by the induction hypothesis, \( \text{Var}(Y) = (n-1)\sigma^2 \). The result follows. □
2. Conditioning

2.1. Conditional mass functions. For all \( y \), define the conditional mass function of \( X \) given that \( Y = y \) as

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

provided that \( f_Y(y) > 0 \).

As a function in \( x \), \( f_{X|Y}(x \mid y) \) is a probability mass function. That is:

1. \( 0 \leq f_{X|Y}(x \mid y) \leq 1 \);
2. \( \sum_x f_{X|Y}(x \mid y) = 1 \).

**Example 17.1** (Example 14.2, Lecture 14, continued). In this example, the joint mass function of \((X,Y)\), and the resulting marginal mass functions, were given by the following:

<table>
<thead>
<tr>
<th>( x ) ( y )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>( f_X )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>16/36</td>
<td>8/36</td>
<td>1/36</td>
<td>25/36</td>
</tr>
<tr>
<td>1</td>
<td>8/36</td>
<td>2/36</td>
<td>0</td>
<td>10/36</td>
</tr>
<tr>
<td>2</td>
<td>1/36</td>
<td>0</td>
<td>0</td>
<td>1/36</td>
</tr>
<tr>
<td>( f_Y )</td>
<td>25/36</td>
<td>10/36</td>
<td>1/36</td>
<td>1</td>
</tr>
</tbody>
</table>

Let us calculate the conditional mass function of \( X \), given that \( Y = 1 \):

\[
f_{X|Y}(0 \mid 1) = \frac{f(0, 1)}{f_Y(1)} = \frac{8}{10} \\
f_{X|Y}(1 \mid 1) = \frac{f(1, 1)}{f_Y(1)} = \frac{2}{10} \\
f_{X|Y}(x \mid 1) = 0 \text{ for other values of } x.
\]

Similarly,

\[
f_{X|Y}(0 \mid 0) = \frac{16}{25} \\
f_{X|Y}(1 \mid 0) = \frac{8}{25} \\
f_{X|Y}(2 \mid 0) = \frac{1}{25} \\
f_{X|Y}(x \mid 0) = 0 \text{ for other values of } x,
\]

and

\[
f_{X|Y}(0 \mid 2) = 1 \\
f_{X|Y}(x \mid 2) = 0 \text{ for other values of } x.
\]
2.2. Conditional expectations. Define conditional expectations, as we did ordinary expectations. But use conditional probabilities in place of ordinary probabilities, viz.,

$$E(X \mid Y = y) = \sum_x xf_{X\mid Y}(x \mid y).$$ (17)

Example 17.2 (Example 17.1, continued). Here,

$$E(X \mid Y = 1) = \left(0 \times \frac{8}{10}\right) + \left(1 \times \frac{2}{10}\right) = \frac{2}{10} = \frac{1}{5}.$$ 

Similarly,

$$E(X \mid Y = 0) = \left(0 \times \frac{16}{25}\right) + \left(1 \times \frac{8}{25}\right) + \left(2 \times \frac{1}{25}\right) = \frac{10}{25} = \frac{2}{5},$$

and

$$E(X \mid Y = 2) = 0.$$ 

Note that $E(X) = 12/36 = 1/3$, which is none of the preceding. If you know that $Y = 0$, then your best bet for $X$ is $2/5$. But if you have no extra knowledge, then your best bet for $X$ is $1/3$.

However, let us note the Bayes’s formula in action:

$$E(X)$$
$$= E(X \mid Y = 0)P\{Y = 0\} + E(X \mid Y = 1)P\{Y = 1\} + E(X \mid Y = 2)P\{Y = 2\}$$
$$= \left(\frac{2}{5} \times \frac{25}{36}\right) + \left(\frac{1}{5} \times \frac{10}{36}\right) + \left(0 \times \frac{1}{36}\right)$$
$$= \frac{12}{36},$$

as it should be.

3. Sums of independent random variables

Theorem 17.3. If $X$ and $Y$ are independent, then

$$f_{X+Y}(z) = \sum_x f_X(x)f_Y(z - x).$$

Proof. We note that $X + Y = z$ if $X = x$ and $Y = z - x$ for that $x$. For example, suppose $X$ is integer-valued and $\geq 1$. Then \{\(X + Y = z\)\} = \bigcup_{x=1}^{\infty} P\{X = x, Y = z - x\}. In general,

$$f_{X+Y}(z) = \sum_x P\{X = x, Y = z - x\} = \sum_x P\{X = x\}P\{Y = z - x\}.$$ 

This is the desired result. $\square$
Example 17.4. Suppose $X = \text{Poisson}(\lambda)$ and $Y = \text{Poisson}(\gamma)$ are independent. Then, I claim that $X + Y = \text{Poisson}(\lambda + \gamma)$. We verify this by directly computing as follows: The possible values of $X + Y$ are $0, 1, \ldots$. Let $z = 0, 1, \ldots$ be a possible value, and then check that

$$f_{X+Y}(z) = \sum_{x=0}^{\infty} f_X(x) f_Y(z-x)$$

$$= \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} f_Y(z-x)$$

$$= \sum_{x=0}^{z} \frac{e^{-\lambda} \lambda^x}{x!} e^{-\gamma} \gamma^{z-x}$$

$$= \sum_{x=0}^{z} \frac{z}{x!} \binom{z}{x} \lambda^x \gamma^{z-x}$$

$$= e^{-(\lambda + \gamma)} \sum_{x=0}^{z} \binom{z}{x} \lambda^x \gamma^{z-x}$$

$$= e^{-(\lambda + \gamma)} (\lambda + \gamma)^z,$$

thanks to the binomial theorem. For other values of $z$, it is easy to see that $f_{X+Y}(z) = 0$. 