Laws of Total Expectation and Total Variance

**Definition of conditional density.** Assume an arbitrary random variable $X$ with density $f_X$. Take an event $A$ with $P(A) > 0$. Then the conditional density $f_{X|A}$ is defined as follows:

$$f_{X|A}(x) = \begin{cases} \frac{f(x)}{P(A)} & x \in A \\ 0 & x \notin A \end{cases}$$

Note that the support of $f_{X|A}$ is supported only in $A$.

**Definition of conditional expectation conditioned on an event.**

$$E(h(X)|A) = \int_A h(x) f_{X|A}(x) \, dx = \frac{1}{P(A)} \int_A h(x) f_X(x) \, dx$$

**Example.** For the random variable $X$ with density function

$$f(x) = \begin{cases} \frac{1}{64} x^3 & 0 < x < 4 \\ 0 & \text{otherwise} \end{cases}$$

calculate $E(X^2|X \geq 1)$.

**Solution.**

**Step 1.**

$$P(X \geq 1) = \int_1^4 \frac{1}{64} x^3 \, dx = \frac{1}{256} \left[ \frac{1}{4} x^4 \right]_{x=1}^{x=4} = \frac{255}{256}$$

**Step 2.**

$$E(X^2|X \geq 1) = \frac{1}{P(X \geq 1)} \int_{\{x \geq 1\}} x^2 f(x) \, dx = \frac{256}{255} \int_1^4 x^2 \left( \frac{1}{64} x^3 \right) \, dx$$

$$= \frac{256}{255} \int_1^4 \left( \frac{1}{64} x^5 \right) \, dx = \left( \frac{256}{255} \right) \left( \frac{1}{64} \right) \left[ \frac{1}{6} x^6 \right]_{x=1}^{x=4} = \frac{8192}{765}$$

**Definition of conditional variance conditioned on an event.**

$$\text{Var}(X|A) = E(X^2|A) - (E(X|A))^2$$
**Example.** For the previous example, calculate the conditional variance $\text{Var}(X|X \geq 1)$

**Solution.**

We already calculated $E(X^2 | X \geq 1)$. We only need to calculate $E(X | X \geq 1)$.

$$E(X | X \geq 1) = \frac{1}{P(X \geq 1)} \int_{x \geq 1} x f(x) \, dx = \frac{256}{255} \int_1^4 x \left( \frac{1}{64} x^3 \right) \, dx$$

$$= \frac{256}{255} \int_1^4 \left( \frac{1}{64} x^4 \right) \, dx = \left( \frac{256}{255} \right) \left( \frac{1}{64} \right) \left[ \frac{1}{5} x^5 \right]_{x=1}^{x=4} = \frac{4096}{1275}$$

Finally:

$$\text{Var}(X|X \geq 1) = E(X^2|X \geq 1) - E(X|X \geq 1)^2 = \frac{8192}{765} - \left( \frac{4096}{1275} \right)^2 = \frac{630784}{1625625}$$

**Definition of conditional expectation conditioned on a random variable.** Suppose two random variables $X$ and $Y$. To define $E(X|Y)$ we need the conditional density function. For this, let $f_{X,Y}(x,y)$ be the joint density of the pair $\{X,Y\}$. Then the conditional density $f_{X|Y}$ is defined as:

$$f_{X|Y}(x|y) = \begin{cases} \frac{f_{X,Y}(x,y)}{f_Y(y)} & f_Y(y) > 0 \\ 0 & f_Y(y) = 0 \end{cases}$$

Then $E(h(X) | Y)$ is defined to be the random variable that assigns the value $\int_{-\infty}^{\infty} h(x) f_{X|Y}(x|y) \, dx$ to $y$ in the continuous case, and assigns the value $\sum_x h(x) f_{X|Y}(x|y)$ in the discrete case.

**Example.** Take the following joint density:

<table>
<thead>
<tr>
<th></th>
<th>X=1</th>
<th>X=2</th>
<th>X=3</th>
<th>$f_Y(y)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y = 1$</td>
<td>0.07</td>
<td>0.1</td>
<td>0.23</td>
<td>0.4</td>
</tr>
<tr>
<td>$Y = 2$</td>
<td>0.25</td>
<td>0.08</td>
<td>0.07</td>
<td>0.4</td>
</tr>
<tr>
<td>$Y = 3$</td>
<td>0.05</td>
<td>0.04</td>
<td>0.11</td>
<td>0.2</td>
</tr>
<tr>
<td>$f_X(x)$</td>
<td>0.37</td>
<td>0.22</td>
<td>0.41</td>
<td>1</td>
</tr>
</tbody>
</table>

Describe the random variable $E(Y^2|X)$
Solution. \[ E(Y^2|X = 1) = \sum_y y^2 f_{Y|X}(y|1) = \sum_y y^2 \frac{f(X=1,Y=y)}{f_X(1)} \]

\[= \frac{1}{f_X(1)} \sum_y y^2 f(X = 1, Y = y) = \frac{1}{0.37} \left\{(1)^2(0.07) + (2)^2(0.25) + (3)^2(0.05)\right\} = \frac{1.52}{0.37} \]

Similarly:
\[ E(Y^2|X = 2) = \frac{0.78}{0.32} \]
\[ E(Y^2|X = 3) = \frac{0.41}{0.41} \]

So then: \[ E(Y^2|X) = \begin{cases} 1.52 \\ 0.78 \\
0.41 \end{cases} \begin{cases} \text{with probability } P(X = 1) = 0.37 \\ \text{with probability } P(X = 2) = 0.22 \\
\text{with probability } P(X = 3) = 0.41 \end{cases} \]

Law of Total Expectation.
\[ E(X) = E(E[X|Y]) \]

Law of Total Variance.
\[ \text{Var}(X) = E\left(\text{Var}[X|Y]\right) + \text{Var}\left(E[X|Y]\right) \]

Proof. By definition we have
\[ \text{Var}(X|Y) = E(X^2|Y) - (E(X|Y))^2 \]
BX taking $E$ of both sides we get:

$$E[\text{Var}(X|Y)] = E[E[X^2|Y]] - E[\{E(X|Y)\}^2]$$

$$= E(X^2) - E[\{E(X|Y)\}^2] \quad \text{law of iterated expectations}$$

$$= \left\{ E(X^2) - \{E(X)\}^2 \right\} - \left\{ E[\{E(X|Y)\}^2] - \{E(X)\}^2 \right\}$$

$$= \text{Var}(X) - \left\{ E[\{E(X|Y)\}^2] - \{E[E(X|Y)]\}^2 \right\}$$

$$= \text{Var}(X) - \text{Var}[E(X|Y)]$$

By moving terms around, the claim follows.

**Note:** Using similar arguments, one can prove the following:

**Example (from the Dean’s note):** Two urns contain a large number of balls with each ball marked with one number from the set $\{0, 2, 4\}$. The proportion of each type of ball in each urn is displayed in the table below:

<table>
<thead>
<tr>
<th>Number on Ball</th>
<th>0</th>
<th>2</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.6</td>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td>B</td>
<td>0.1</td>
<td>0.3</td>
<td>0.6</td>
</tr>
</tbody>
</table>

An urn is randomly selected and then a ball is drawn at random from the urn. The number on the ball is represented by the random variable $X$.

(a) Calculate the hypothetical means (or conditional means)

$$E[X|\theta = A] \quad \text{and} \quad E[X|\theta = B]$$

(b) Calculate the variance of the hypothetical means: $\text{Var}[E(X|\theta)]$. 

4
(c) Calculate the process variances (or conditional variances)

\[ \text{Var}[X|\theta = A] \quad \text{and} \quad \text{Var}[X|\theta = B] \]

(d) Calculate the expected value of the process variance: \( E[\text{Var}[X|\theta]] \).

(e) Calculate the total variance (or unconditional variance) \( \text{Var}[X] \) and show that it equals the sum of the quantities calculated in (b) and (d).

**Solution: Part (a)**

\[
E[X|\theta = A] = (0.6)(0) + (0.3)(2) + (0.1)(4) = 1.0
\]
\[
E[X|\theta = B] = (0.1)(0) + (0.3)(2) + (0.6)(4) = 3.0
\]

**Part (b)**

\[
E[X] = E[E[X|\theta]] = \left( \frac{1}{2} \right)(1.0) + \left( \frac{1}{2} \right)(3.0) = 2.0
\]
\[
\text{Var}[E[X|\theta]] = \left( \frac{1}{2} \right)(1.0 - 2.0)^2 + \left( \frac{1}{2} \right)(3.0 - 2.0)^2 = 1.0
\]

**Part (c)**

\[
\text{Var}[X|\theta = A] = (0.6)(0 - 1.0)^2 + (0.3)(2 - 1.0)^2 + (0.1)(4 - 1.0)^2 = 1.8
\]
\[
\text{Var}[X|\theta = B] = (0.1)(0 - 3.0)^2 + (0.3)(2 - 3.0)^2 + (0.6)(4 - 3.0)^2 = 1.8
\]

**Part (d)**

\[
E[\text{Var}[X|\theta]] = \left( \frac{1}{2} \right)(1.8) + \left( \frac{1}{2} \right)(1.8) = 1.8
\]

**Part (e)**

\[
\text{Var}[X] = \frac{1}{2} \left[ (0.6)(0 - 2.0)^2 + (0.3)(2 - 2.0)^2 + (0.1)(4 - 2.0)^2 \right]
\]
\[
+ \frac{1}{2} \left[ (0.1)(0 - 2.0)^2 + (0.3)(2 - 2.0)^2 + (0.6)(4 - 2.0)^2 \right]
\]
\[
= 2.8
\]
\[
\Rightarrow \quad \text{Var}(X) = \text{Var}[E[X|\theta]] + E[\text{Var}[X|\theta]]
\]