Law of Iterated Expectations

Guillem Riambau. YSS211. Econometrics, Yale-NUS. Spring 2016.

The Law of Iterated Expectations states that:

$$E(X) = E(E(X|Y)) \tag{1}$$

This document tries to give some intuition to the L.I.E. Sometimes you may see it written as $E(X) = E_y(E_x(X|Y))$. At the end of the document it is explained why (note, both mean exactly the same!). We will first start with a simple and numerical example, then proceed to the proof.

Suppose that we take a random sample of the Singapore population (N=1,000), and we ask people to tell us how many durians they eat per month. For simplicity, suppose that there is no variation within each category (i.e. all young angmoh give the same answer, all young locals give the same answer etc.). These are the quantities regarding durians that we find:

Young	Old	
4	6	Local
0	4	Angmoh

Table 1: Number of durians eaten per month

That is, young locals eat 4 durians per month, young angmoh eat none per month, etc. Next, the frequencies (number of respondents in each category)

Young	Old	
300	500	Local
50	150	Angmoh

Table 2: Frequencies in the sample

Notation: D= number of durians, D_{young}= number of durians eaten by young, D_{old}= number of durians eaten by old, $D_{local,old}$= number of durians eaten by and old uncle/auntie, etc.

We can first ask, what is the expected number of durians (D) eaten, conditional on being a local? That is, find $E(D|local)$.

For that, we need to know the conditional probabilities: that is, P(young | local) and P(old | local). In words, the probability of being young conditional on being local, and the probability of being old conditional on being local.

P(young | local) $= \frac{300}{500+300} = \frac{3}{8}$. Given that there are only two categories, we can infer that P(old | local)=1-P(young | local) $= \frac{5}{8}$.

Now we can compute the conditional expectation:

(2)

$$E(D|local) = P(young|local) \times D_{young,local} + P(old|local) \times D_{old,local} = \frac{3}{8}4 + \frac{5}{8}6 = \frac{12+30}{8} = \frac{21}{4}$$

Similarly,

(3)

$$E(D|angmoh) = P(young|angmoh) \times D_{young,angmoh} + P(old|angmoh) \times D_{old,angmoh} = \frac{1}{4}0 + \frac{3}{4}4 = 3$$

Suppose now that we want to compute the expected number of durians eaten for the whole population. We can do it in two ways. [Notation: P(*young, local*)=Probability of being both young and local.]

- First way. Easy but longer:

$$\begin{aligned}
&E(D) = \\
&P(young,\ local) \times D_{young,local} + \\
&P(young,\ angmoh) \times D_{young,angmoh} + \\
&P(old,\ local) \times D_{old,local} \\
&P(old,\ angmoh) \times D_{old,angmoh} = \\
&0.3 \times 4 + 0.05 \times 0 + 0.5 \times 6 + 0.15 \times 4 = 4.8
\end{aligned} \tag{4}$$

- Second way. Complex but much faster. Given that we have the two conditional expectations, we only need to take the expectation of them. That is,

(5)

$$E(D) = E(E(D|origin)) = P(local) \times E(D|local) + P(angmoh) \times E(D|angmoh) =$$
$$\frac{8}{10} \times \frac{21}{4} + \frac{2}{10} \times 3 = \frac{168}{40} + \frac{6}{10} = \frac{42}{10} + \frac{6}{10} = \frac{48}{10} = 4.8 \qquad \text{as above.}$$

Now lets see the formal proof. Recall some notation:

- $P(Y = y)$: Probability that the random variable Y takes value y.
- $\sum_y$: Sum over all possible values y that Y can take.
- $E(X|Y = y)$: Expected value of X given that Y takes value y.
- $P(X = x|Y = y)$: Probability that X takes value x given that Y takes value y.

Also recall that: $\sum_{i=1}^{N}\sum_{j=1}^{M} x_i y_j = \sum_{i=1}^{N} x_i \sum_{j=1}^{M} y_j = \sum_{j=1}^{M}\sum_{i=1}^{N} y_j x_i$

(Note: in the problem set, you are asked to identify in which step Bayes' Law is used. You are also asked to show that the equality just above holds).

Proof of the Law of Iterated Expectations:

$$
\begin{aligned}
& E(X) = E(E(X|Y)) \\
& = \sum_y E(X|Y=y)P(Y=y) \\
& = \sum_y \sum_x xP(X=x|Y=y)P(Y=y) \\
& = \sum_y \sum_x xP(Y=y|X=x)P(X=x) \\
& = \sum_y \sum_x xP(X=x)P(Y=y|X=x) \\
& = \sum_x xP(X=x) \sum_y P(Y=y|X=x) \\
& (\text{Note: } \sum_y P(Y=y|X=x) = 1) \\
& = \sum_x xP(X=x) \\
& = E(X)
\end{aligned} \tag{6}
$$

Note the first step could be written as $E(X) = E_y(E_x(X|Y))$, because for the 'inner' expectation we are using the probabilities over the x-outcomes, and for the 'outer' expectation we use the probabilities over the y-outcomes.

In the example above, it would be $E_{or}(E_d(D|origin))$, since the 'inner' expectation is taken over the conditional probabilities of eating durian (hence, subscript 'd'), and the 'outer' expectation is taken over the probabilities of being 'local' or 'angmoh' (hence, subscript 'or' for origin).