1 Uniform Probability Space

Let $\Omega = \{1,2,3,4,5,6\}$ be a uniform probability space. Let also $X(\omega)$ and $Y(\omega)$, for $\omega \in \Omega$, be the random variables defined in the table:

Table 1: All the rows in the table correspond to random variables.

ω	1	2	3	4	5	6
$X(\omega)$	0	0	1	1	2	2
$Y(\omega)$	0	2	3	5	2	0
$X^2(\omega)$						
$Y^2(\omega)$						
$XY(\omega)$						
$L[Y \mid X](\omega)$						
$E[Y \mid X](\omega)$						

(a) Fill in the blank entries of the table.

(b) Are the variables correlated or uncorrelated? Are the variables independent or dependent?

(c) Calculate $\mathbf{E}[(Y - L[Y \mid X])^2]$ and $\mathbf{E}[(Y - \mathbf{E}[Y \mid X])^2]$. Which is smaller? Is this always true?

Solution:

(a) See the following table:

ω	1	2	3	4	5	6
$X(\omega)$	0	0	1	1	2	2
$Y(\omega)$	0	2	3	5	2	0
$X^2(\omega)$	0	0	1	1	4	4
$Y^2(\omega)$	0	4	9	25	4	0
$XY(\omega)$	0	0	3	5	4	0
$L[Y \mid X](\omega)$	2	2	2	2	2	2
$E[Y \mid X](\omega)$	1	1	4	4	1	1

The third, fourth, and fifth rows can be calculated directly from the corresponding X and Y values. Recall that

$$L[Y \mid X] = \frac{\text{cov}(X,Y)}{\text{var}(X)}(X - \mathbf{E}(X)) + \mathbf{E}(Y).$$

But $\text{cov}(X,Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y) = 2 - (1)(2) = 0$, so $L[Y \mid X] = \mathbf{E}(Y) = 2$ for all ω.

The conditional expectation can be found by averaging the values of Y for each value of X.

(b) Since $\text{cov}(X,Y) = 0$, the variables are uncorrelated. But, we see that $\Pr(Y = 0) = 1/3$ and $\Pr(Y = 0 \mid X = 3) = 0$, so the two variables are not independent. Recall that independence implies uncorrelation, but the converse is not true.

(c)

$$\mathbf{E}[(Y - L[Y \mid X])^2] = \mathbf{E}[(Y-2)^2] = \frac{4+0+1+9+0+4}{6} = 3$$

$$\mathbf{E}[(Y - \mathbf{E}[Y \mid X])^2] = \frac{1+1+1+1+1+1}{6} = 1$$

Note how the MMSE squared error is smaller than the LLSE squared error, because MMSE does not constrain its function space to linear functions. The MMSE squared error is always less than or equal to the LLSE square error, with equality holding only when the MMSE is linear.

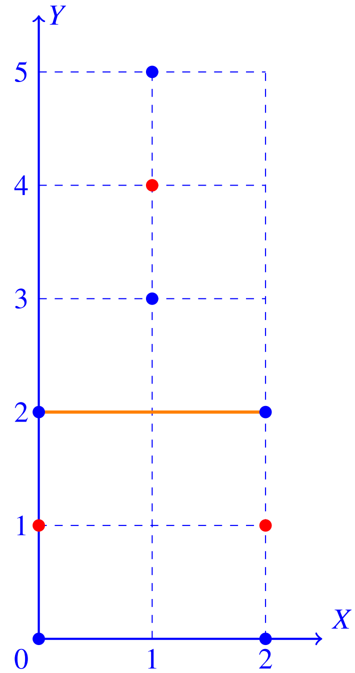

Figure 1: Visualization of the LLSE and MMSE. The circles are the (X,Y) points. The orange line is the LLSE. The red points are the MMSE.

2 Number of Ones

In this problem, we will revisit dice-rolling, except with conditional expectation.

(a) If we roll a die until we see a 6, how many ones should we expect to see?

(b) If we roll a die until we see a number greater than 3, how many ones should we expect to see?

Solution:

(a) Let Y be the number of ones we see. Let X be the number of rolls we take until we get a 6.

Let us first compute $\mathbf{E}[Y \mid X]$. We know that in each of our $k-1$ rolls before the kth, we necessarily roll a number in $\{1, 2, 3, 4, 5\}$. Thus, we have a $1/5$ chance of getting a one, meaning

$$\mathbf{E}[Y \mid X = k] = \frac{1}{5}(k-1)$$

so

$$\mathbf{E}[Y \mid X] = \frac{1}{5}(X-1).$$

If this is confusing, write Y as a sum of indicator variables.

$$Y = Y_1 + Y_2 + \cdots + Y_k$$

where Y_i is 1 if we see a one on the ith roll. This means

$$\mathbf{E}[Y \mid X = k] = \mathbf{E}[Y_1 \mid X = k] + \mathbf{E}[Y_2 \mid X = k] + \cdots + \mathbf{E}[Y_k \mid X = k].$$

We know for a fact that on the kth roll, we roll a 6, thus $\mathbf{E}[Y_k] = 0$. Thus, we actually consider

$$\begin{aligned}\mathbf{E}[Y_1 \mid X = k] + \mathbf{E}[Y_2 \mid X = k] + \cdots + \mathbf{E}[Y_{k-1} \mid X = k] &= (k-1)\mathbf{E}[Y_1 \mid X = k] \\ &= (k-1)\Pr[Y_1 = 1 \mid X = k] \\ &= (k-1)\frac{1}{5}.\end{aligned}$$

Using the Law of Total Expectation, we know that

$$\begin{aligned}\mathbf{E}[Y] = \mathbf{E}[\mathbf{E}[Y \mid X]] &= \mathbf{E}\left[\frac{1}{5}(X-1)\right] \\ &= \frac{1}{5}\mathbf{E}[X-1] \\ &= \frac{1}{5}(\mathbf{E}[X]-1).\end{aligned}$$

Since, $X \sim \text{Geom}(1/6)$, the expected number of rolls until we roll a 6 is $\mathbf{E}[X] = 6$.

$$\frac{1}{5}(\mathbf{E}[X]-1) = \frac{1}{5}(6-1) = 1.$$

(b) We use the same logic as the first part, except now each of the first $k-1$ rolls can only be 1, 2, or 3, so

$$\mathbf{E}[Y \mid X = k] = \frac{1}{3}(k-1).$$

Then

$$\begin{aligned}\mathbf{E}[Y] = \mathbf{E}[\mathbf{E}[Y \mid X]] &= \mathbf{E}\left[\frac{1}{3}(X-1)\right] \\ &= \frac{1}{3}(\mathbf{E}[X]-1).\end{aligned}$$

Since $X \sim \text{Geom}(1/2)$, we know that the expected number of rolls until we roll a number greater than 3 is $\mathbf{E}[X] = 2$. This makes $\mathbf{E}[Y] = 1/3$.

3 Marbles in a Bag

We have r red marbles, b blue marbles, and g green marbles in the same bag. If we sample marbles with replacement until we get 3 red marbles (not necessarily consecutively), how many blue marbles should we expect to see?

Solution:

Let Y be the number of blue marbles we see. Let X be the samples we take until we get 3 red marbles.

Let us first compute $\mathbf{E}[Y \mid X]$. Let Y_i be 1 if we see a blue marble on the ith sample and $Y = \sum_{i=1}^{k} Y_i$. This means

$$\begin{aligned}\mathbf{E}[Y \mid X = k] &= \mathbf{E}\left[\sum_{i=1}^{k} Y_i \,\middle|\, X = k\right] \\ &= \sum_{i=1}^{k} \mathbf{E}[Y_i \mid X = k].\end{aligned}$$

However, three Y_i (call them Y_a, Y_b, Y_c) have $\mathbf{E}[Y_i] = 0$, since there are necessarily 3 red marbles. This means the other $k-3$ marbles are necessarily blue or green.

$$\begin{aligned}\sum_{i \neq a,b,c} \mathbf{E}[Y_i \mid X = k] &= \sum_{i \neq a,b,c} \Pr[Y_i = 1 \mid X = k] \\ &= \sum_{i \neq a,b,c} \frac{b}{b+g} \\ &= (k-3)\frac{b}{b+g}.\end{aligned}$$

This means

$$\mathbf{E}[Y \mid X] = (X-3)\frac{b}{b+g}.$$

Using the Law of Total Expectation, we know that

$$\begin{aligned}\mathbf{E}[Y] = \mathbf{E}[\mathbf{E}[Y \mid X]] &= \mathbf{E}\left[\frac{b}{b+g}(X-3)\right] \\ &= \frac{b}{b+g}\mathbf{E}[X-3] \\ &= \frac{b}{b+g}(\mathbf{E}[X]-3).\end{aligned}$$

We know that $\mathbf{E}[X] = 3(r+g+b)/r$.

$$\mathbf{E}[Y] = \frac{b}{b+g}\left(3\frac{r+g+b}{r} - 3\right) = \frac{3b}{r}.$$

Alternate Solution: We notice that

$$X = X_1 + X_2 + X_3,$$

where each X_i represents the number of blue marbles seen between drawing the $(i-1)$th and ith red marble.

We know that the absolute number of marbles seen between 2 consecutive red marbles is geometric, since we want to find the number of draws until the first red marble. And given the number of marbles, n, between 2 consecutive red marbles, the number of blue marbles among these is distributed binomially.

Therefore, each X_i is drawn from a binomial distribution, where the number of trials is distributed geometrically.

We exclude the very last marble in the binomial distribution, because we know it must be red (and therefore cannot be blue). And the probability for the binomial is $b/(b+g)$ because we know that in between 2 consecutive red balls, we can only have blue or green balls. So,

$$X_i \sim \mathrm{Bin}\left(\frac{b}{b+g}, N-1\right), \qquad \text{where} \qquad N \sim \mathrm{Geom}\left(\frac{r}{r+b+g}\right).$$

And, applying the law of conditional expectation, we have

$$\begin{aligned}
\mathbf{E}(X_i) &= \mathbf{E}(\mathbf{E}(X_i \mid N)) \\
&= \mathbf{E}\left((N-1)\frac{b}{b+g}\right) \\
&= \frac{b}{b+g}\mathbf{E}(N-1) \\
&= \frac{b}{b+g}\left(\frac{r+b+g}{r} - 1\right) \\
&= \frac{b}{b+g}\left(\frac{b+g}{r}\right) \\
&= \frac{b}{r}.
\end{aligned}$$

We know that each of the X_i's is identically distributed, so

$$\mathbf{E}(X) = \mathbf{E}(X_1) + \mathbf{E}(X_2) + \mathbf{E}(X_3) = 3 \cdot \mathbf{E}(X_1) = \frac{3b}{r}.$$