1 LLSE

Note 20

We have two bags of balls. The fractions of red balls and blue balls in bag A are $2/3$ and $1/3$ respectively. The fractions of red balls and blue balls in bag B are $1/2$ and $1/2$ respectively. Someone gives you one of the bags (unmarked) uniformly at random. You then draw 6 balls from that same bag with replacement. Let X_i be the indicator random variable that ball i is red. Now, let us define $X = \sum_{1 \le i \le 3} X_i$ and $Y = \sum_{4 \le i \le 6} X_i$.

(a) Compute $\mathbb{E}[X]$ and $\mathbb{E}[Y]$.

(b) Compute $\mathrm{Var}(X)$.

(c) Compute $\mathrm{cov}(X,Y)$. (*Hint*: Recall that covariance is bilinear.)

(d) Now, we are going to try and predict Y from a value of X. Compute $L(Y \mid X)$, the best linear estimator of Y given X. Recall that

$$L(Y \mid X) = \mathbb{E}[Y] + \frac{\mathrm{cov}(X,Y)}{\mathrm{Var}(X)}\big(X - \mathbb{E}[X]\big).$$

Solution: Although the indicator random variables are not independent, we can still apply linearity of expectation. By symmetry, we also know that each indicator follows the same distribution.

(a)

$$\mathbb{E}[X] = \mathbb{E}[Y] = 3 \cdot \mathbb{E}[X_1] = 3 \cdot \mathbb{P}[X_1 = 1] = 3 \cdot \left(\frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{2}\right) = \frac{7}{4}.$$

(b)

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{cov}\Big(\sum_{1 \le i \le 3} X_i, \sum_{1 \le j \le 3} X_j\Big) \\
&= 3 \cdot \mathrm{Var}(X_1) + 6 \cdot \mathrm{cov}(X_1, X_2) \\
&= 3\big(\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2\big) + 6 \cdot \frac{1}{144} \\
&= 3\left[\frac{7}{12} - \left(\frac{7}{12}\right)^2\right] + 6 \cdot \frac{1}{144} = \frac{111}{144}.
\end{aligned}$$

(c)

$$\begin{aligned}\operatorname{cov}(X,Y) &= \operatorname{cov}\Big(\sum_{1\le i\le 3} X_i, \sum_{4\le j\le 6} X_j\Big)\\ &= 9\cdot \operatorname{cov}(X_1,X_4)\\ &= 9\cdot \big(\mathbb{E}[X_1X_4]-\mathbb{E}[X_1]\cdot\mathbb{E}[X_4]\big)\\ &= 9\cdot\Big(\mathbb{P}[X_1=1,X_4=1]-\mathbb{P}[X_1=1]^2\Big)\\ &= 9\cdot\left(\left[\frac{1}{2}\cdot\left(\frac{2}{3}\right)^2+\frac{1}{2}\cdot\left(\frac{1}{2}\right)^2\right]-\left[\frac{1}{2}\cdot\left(\frac{2}{3}\right)+\frac{1}{2}\cdot\left(\frac{1}{2}\right)\right]^2\right)=\frac{9}{144}.\end{aligned}$$

(d)

$$L(Y\mid X)=\frac{7}{4}+\frac{9}{111}\left(X-\frac{7}{4}\right)=\frac{3}{37}X+\frac{119}{74}.$$

2 Balls in Bins Estimation

Note 20

We throw $n>0$ balls into $m\ge 2$ bins. Let X and Y represent the number of balls that land in bin 1 and 2 respectively.

(a) Calculate $\mathbb{E}[Y\mid X]$. [*Hint*: Your intuition may be more useful than formal calculations.]

(b) What is $L[Y\mid X]$ (where $L[Y\mid X]$ is the best linear estimator of Y given X)? [*Hint*: Your justification should be no more than two or three sentences, no calculations necessary! Think carefully about the meaning of the conditional expectation.]

(c) Unfortunately, your friend is not convinced by your answer to the previous part. Compute $\mathbb{E}[X]$ and $\mathbb{E}[Y]$.

(d) Compute $\operatorname{Var}(X)$.

(e) Compute $\operatorname{cov}(X,Y)$.

(f) Compute $L[Y\mid X]$ using the formula. Ensure that your answer is the same as your answer to part (b).

Solution:

(a) $\mathbb{E}[Y\mid X=x]=(n-x)/(m-1)$, because once we condition on x balls landing in bin 1, the remaining $n-x$ balls are distributed uniformly among the other $m-1$ bins. Therefore,

$$\mathbb{E}[Y\mid X]=\frac{n-X}{m-1}.$$

(b) We showed that $\mathbb{E}[Y \mid X]$ is a linear function of X. Since $\mathbb{E}[Y \mid X]$ is the best *general* estimator of Y given X, it must also be the best *linear* estimator of Y given X, i.e. $\mathbb{E}[Y \mid X]$ and $L[Y \mid X]$ coincide.

(c) Let X_i be the indicator that the ith ball falls in bin 1. Then, $X = \sum_{i=1}^n X_i$, and by linearity of expectation, $\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X_i] = n/m$, since there are n indicators and each ball has a probability $1/m$ of landing in bin 1. By symmetry, $\mathbb{E}[Y] = n/m$ as well.

(d) The number of balls that falls into the first bin is binomially distributed with parameters n and $1/m$. Hence the variance is $n(1/m)(1-1/m)$.

(e) Let X_i be as before, and let Y_i be the indicator that the ith ball falls into bin 2.

$$\operatorname{cov}(X,Y) = \sum_{i=1}^n \sum_{j=1}^n \operatorname{cov}(X_i, Y_j)$$

We can compute $\operatorname{cov}(X_i,Y_i) = \mathbb{E}[X_iY_i] - \mathbb{E}[X_i]\,\mathbb{E}[Y_i] = 0 - (1/m)(1/m) = -1/m^2$ (note that $\mathbb{E}[X_iY_i] = 0$ because it is impossible for a ball to land in both bins 1 and 2). Also, we have $\operatorname{cov}(X_i,Y_j) = 0$ because the indicator for the ith ball is independent of the indicator for the jth ball when $i \neq j$. Hence, $\operatorname{cov}(X,Y) = n(-1/m^2) = -n/m^2$.

(f)

$$\begin{aligned}
L[Y \mid X] &= \mathbb{E}[Y] + \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(X)}(X - \mathbb{E}[X]) \\
&= \frac{n}{m} + \frac{-n/m^2}{n(1/m)(1-1/m)}\left(X - \frac{n}{m}\right) \\
&= \frac{n}{m} - \frac{1}{m-1}\left(X - \frac{n}{m}\right) \\
&= \frac{mn - n - mX + n}{m(m-1)} = \frac{n-X}{m-1}
\end{aligned}$$

3 Number of Ones

Note 20

In this problem, we will revisit dice-rolling, except with conditional expectation. (*Hint*: for both of these subparts, the law of total expectation may be helpful.)

(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 $\mathbb{E}[Y \mid X = k]$. 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 in each of these $k-1$ previous rolls, giving

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

If this is confusing, we can 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 that by linearity of expectation,

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

We know that on the kth roll, we must roll a 6, so $\mathbb{E}[Y_k] = 0$. Further, by symmetry, each term in this summation has the same value; this means that we have

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

Using the law of total expectation, we now have

$$\begin{aligned}\mathbb{E}[Y] &= \sum_{k=1}^{\infty} \mathbb{E}[Y \mid X = k]\mathbb{P}[X = k] && \text{(total expectation)} \\ &= \sum_{k=1}^{\infty} \frac{1}{5}(k-1)\mathbb{P}[X = k]\end{aligned}$$

Here, we can see that this is an application of LOTUS for $f(X) = \frac{1}{5}(X-1)$, so we can simplify this to

$$\begin{aligned} &= \mathbb{E}\left[\frac{1}{5}(X-1)\right] && \text{(LOTUS)} \\ &= \frac{1}{5}(\mathbb{E}[X]-1) && \text{(linearity)}\end{aligned}$$

Since $X \sim \text{Geometric}(\frac{1}{6})$, the expected number of rolls until we roll a 6 is $\mathbb{E}[X] = 6$:

$$= \frac{1}{5}(6-1) = 1$$

Alternatively, we can use iterated expectation, along with the fact that $\mathbb{E}[Y \mid X] = \frac{1}{5}(X-1)$, to give

$$\begin{aligned}\mathbb{E}[Y] &= \mathbb{E}[\mathbb{E}[Y \mid X]] \\ &= \mathbb{E}\left[\frac{1}{5}(X-1)\right] \\ &= \frac{1}{5}(\mathbb{E}[X]-1) \\ &= \frac{1}{5}(6-1) = 1\end{aligned}$$

(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

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

Using the law of total expectation, we have

$$\begin{aligned}
\mathbb{E}[Y] &= \sum_{k=1}^{\infty} \mathbb{E}[Y \mid X = k]\mathbb{P}[X = k] && \text{(total expectation)} \\
&= \sum_{k=1}^{\infty} \frac{1}{3}(k-1)\mathbb{P}[X = k] \\
&= \mathbb{E}\left[\frac{1}{3}(X-1)\right] && \text{(LOTUS)} \\
&= \frac{1}{3}(\mathbb{E}[X]-1) && \text{(linearity)}
\end{aligned}$$

Since now $X \sim \text{Geometric}(\frac{1}{2})$, the expected number of rolls until we roll a number greater than 3 is $\mathbb{E}[X] = 2$:

$$= \frac{1}{3}(2-1) = \frac{1}{3}$$

Alternatively, we can use iterated expectation, along with the fact that $\mathbb{E}[Y \mid X] = \frac{1}{3}(X-1)$, to give

$$\begin{aligned}
\mathbb{E}[Y] &= \mathbb{E}[\mathbb{E}[Y \mid X]] \\
&= \mathbb{E}\left[\frac{1}{3}(X-1)\right] \\
&= \frac{1}{3}(\mathbb{E}[X]-1) \\
&= \frac{1}{3}(2-1) = \frac{1}{3}
\end{aligned}$$