Probability 1

Daniel Yu

September 24,2024

Contents

1 Conditional Expectation

Definition 1. Let X, Y be random variables on the same space.

$$E[X|Y=y] = \sum_{k \in range(x)} k \cdot P[X=k|Y=y].$$

We can think of conditional expectations as random variables.

Note. Example
Toss two four sided fair dice, and let $X=$ 1st toss and $M=$maximum.

1. What $E[X|M=1]$? $E[X|M=1] = 1 \cdot P[X=1|M=1] = 1 \cdot \frac{P[X=1 \cap M=1]}{P[M=1]} = 1 \cdot \frac{\{(1,1)\}}{\{(1,1)\}} = 1$
2. What about $E[X|M=2]$?

$$\begin{aligned}
&= 1 \cdot P[X=1|M=2] + 2 \cdot P[X=2|M=2] \\
&= 1 \cdot \frac{P[X=1 \cap M=2]}{P[M=1]} + 2 \cdot \frac{P[X=2 \cap M=2]}{P[M=2]} \\
&= 1 \cdot \frac{P[\{(1,2)\}]}{P[\{(1,2),(2,2),(2,1)\}} + 2 \cdot \frac{P[\{(2,1),(2,2)\}}{P[\text{same}]} \\
&= \frac{\frac{1}{16}}{\frac{3}{16}} + 2 \cdot \frac{\frac{2}{16}}{\frac{3}{16}} \\
&= \frac{5}{3}.
\end{aligned}$$

3. What about $E[M|X=2]$?

$$\begin{aligned}
E[M|X=2] &= 2 \cdot \frac{P[M=2 \cap X=2]}{P[X=2]} + 3 \cdot \frac{P[M=3 \cap X=2]}{P[X=2]} + 4 \cdot \frac{P[M=4 \cap X=2]}{P[X=2]} \\
&= \frac{11}{4}.
\end{aligned}$$

Definition 2. The **law of total expectation**, when X, Y are random variables on the same space:

$$E[X] = \sum_{r \in range(x)} E[X|Y=r] \cdot P[Y=r].$$

Proof. By writing it out:

$$\begin{aligned}
E[X] &= \sum_{r \in range(x)} E[X|Y=r] \cdot P[Y=r] \\
&= \sum_{r \in range(x)} \sum_{k \in range(x)} k \cdot P[X=k|Y=r] \cdot P[Y=r] \\
&= \sum_{k \in range(x)} k \sum_{r \in range(x)} P[X=k|Y=r] \cdot P[Y=r] \\
&= \sum_{k \in range(x)} k \cdot P[X=k] \\
&= E[X].
\end{aligned}$$

$\square$

Lemma 1. Another way of writing the previous example is as follows and is known as **law of iterated expectations**:

$$E[X] = E_Y[E_X[X|Y]].$$

Remark. The use of these two expectation laws is to be able to now "condition" on a random variable.

Note. Example
Consider a rat in a maze with three intial paths to choose. Let T be the amount of time required for the mouse to leave. What is $E[T]$?

Proof. Let C = 1st choice of the mouse.

$$E[T] = E[T|C=1]\cdot P[C=1] + E[T|C=2]\cdot P[C=2] + E[T|C=3]\cdot P[C=3].$$

Since we don't know anything about the maze other than there are 3 options (this is blc the values are not relevant and just for plugging in). Let's just assume that the following values hold (but really they could be any values). For choices 2,3 the mouse goes around in a circle and arrives back assume the mouse is dumb and again independently makes a decision, i.e. the experiment resets and T' has the same distribution as T.

$$\begin{cases} c=1, T=3, E[T|C=1]=3 \\ c=2, T=4+T' \\ E[T|c=2]=4+E[T] \\ c=3, T=5+t' \\ E[T|c=3]=5+E[T] \end{cases}.$$

Substituting these values in,

$$\begin{aligned} E[T] &= 3\cdot\frac{1}{3} + (4+E[T])\cdot\frac{1}{3} + (5+E[T])\cdot\frac{1}{3} \\ &= 4 + \frac{2}{3}E[T] \\ E[T] &= 12. \end{aligned}$$

□

Note. Example
Consider a similar experiment design. We flip a coin, with $P[H]=p$, each toss independent. Let G = first time we get a H. Let X_1 be the result of the first coin toss:

Proof. Recall that G is a geometric random variable. So as expected:

$$\begin{aligned} E[G] &= E[G|X_1=H]\cdot P[X_1=H] + E[G|X_1=T]\cdot P[X_1=T] \\ &= 1\cdot p + (1+E[G])(1-p) \\ E[G] &= p + (1-p) - pE[G] \\ &= \frac{1}{p}. \end{aligned}$$

□

Remark. What if the above was modified to find $E[D]$ where D = first time we see two H's in a row.

Proof. Recall: $E[D] = E_G[E_D[D|G]]$. The **key to reasoning** about this is to consider what happens when you know G. Consider the X_{G+1} the result of the coin after G. We condition $E[D|G]$ now on X_{G+1}

$$E[D|G, X_{G+1}=H] = G+1 \text{ conditional probability reached heads then next roll heads.}$$

$$E[D|G, X_{G+1}=T] = G+1+E[D] \text{ full reset of the experiment.}$$

Then, we substituting into the law of iterated expectations gives us:

$$\begin{aligned} E[D|G] &= E[D|G, X_{G+1}=H]\cdot P[X_{G+1}=H] + E[D|G, X_{G+1}=T]\cdot P[X_{G+1}=T] \\ &= (G+1)\cdot p + (G+1+E[D])(1-p) \\ E[D|G] &= G+1+(1-p)E[D]. \end{aligned}$$

Now we substitute back for $E[D]$:

$$\begin{aligned}
E_G[E[D|G]] &= E_G[G + 1 + (1 + p)E[D]] \\
E[D] &= \left(\frac{1}{p} + 1\right) + (1 - p)E[D] \\
E[D] &= \frac{1}{p^2} + \frac{1}{p}.
\end{aligned}$$

□

Remark. What if now we consider S = first time we see HT, what is $E[S]$?

Proof. $E[S] = E_G[E_S[S|G]]$. We can keep the same definitions of G and X_{G+1} as before. But think very carefully about the pattern:

$$\begin{cases} E[S|G, X_{G+1} = T] = G + 1 \\ E[S|G, X_{G+1} = H] = G + 1 + E[Geo(1 - p)] \end{cases}.$$

This changes becasue now the pattern is $T \dots THH$ and we **don't** restart because H is the beginning of a HT pattern. So what we are really now searching for is the first time we get a tails $HH \dots HT$! This is $Geo(1 - p)$.

$$\begin{aligned}
E[S|G] &= (G + 1) \cdot (1 - p) + (G + 1 + \frac{1}{1 - p}) \cdot p \\
&= G + 1 + \frac{p}{1 - p}.
\end{aligned}$$

Then,

$$E[S] = \frac{1}{p} + 1 + \frac{p}{1 - p} = \frac{1}{p(1 - p)}.$$

□

Note. Example
Suppose N is the number of customers that walk into a bank on thursday. Each customer deposits 2 or withdraws 1, independently at random. Let Y = total cashflow on Thursday. What is $E[Y]$

Proof. Let $X_i \in \{2, -1\}$ represent the random variable of the customers behavior. Note that we can break up the expected value in this way because we are implicitly assuming (based one the problem) that the number of people N doesn't affect each customer's decision X_i

$$\begin{aligned}
E[Y|N = k] = E[\sum_{i=1}^{k} X_i] = k \cdot E[X_1] = k \cdot \left(\frac{1}{2} \cdot 2 - \frac{1}{2} \cdot 1\right) = \frac{k}{2} \\
E[Y] = \sum_{k=1}^{\infty} E[Y|N = k] \cdot P[n = k] \\
= \sum_{k=1}^{\infty} \frac{k}{2} \cdot P[N = k] \\
= \frac{1}{2} \sum_{k=1}^{\infty} k \cdot P[N = k] \\
= \frac{1}{2} E[N].
\end{aligned}$$

completely expected! □

What is the variance of Y?

Proof. First compute, $E[Y^2]$

$$\begin{aligned}
E[Y^2] &= \sum_{k=1}^{\infty} E[Y^2|N=k]\cdot P[N=k] \\
&= \sum_{k=1}^{\infty} E[\left(\sum_{i=1}^{k} X_i\right)^2 |N=k]\cdot P[N\cdot k] \\
&= \sum_{k=1}^{\infty} E[\sum_{i=1}^{k} X_i^2 + \sum_{i\neq j} X_i X_j|N=k]P[N\cdot k] \\
&= \sum_{k=1}^{\infty} [\sum_{i=1}^{k} E[X_i^2|N=k] + \sum_{i\neq j} E[X_i X_j|N=k]]\cdot P[N=k] \\
&= \sum_{k=1}^{\infty} [\sum_{i=1}^{k} E[X_i^2] + \sum_{i\neq j} E[X_i X_j]]\cdot P[N=k].
\end{aligned}$$

We know that X_i, X_j are independent, so $E[X_iX_j] = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$. Now let's analyze X_i^2 (since all X_i have the same distribution):

$$X_i = \begin{cases} 4, P[X_2 = 4] = \frac{1}{2} \\ 1, P[X_2 = 1] = \frac{1}{2} \end{cases} .$$

Then, $E[X_i]^2 = \frac{5}{2}$:

$$\begin{aligned}
E[Y^2] &= \sum_{k=1}^{\infty} [\sum_{i=1}^{k} \frac{5}{2} + \sum_{i\neq j} \frac{1}{4}]\cdot P[N=k] \\
&= \sum_{k=1}^{\infty} [k\cdot \frac{5}{2} + k\cdot(k-1)\frac{1}{4}]\cdot P[N=k] \\
&= \sum_{k=1}^{\infty} (k\cdot\frac{1}{4})\cdot P[N=k] + \sum_{k=1}^{\infty} k\cdot(\frac{5}{2}-\frac{1}{4})\cdot P[N=k] \\
&= E[X_i]^2\cdot E[N^2 + Var(X_1) + E[N]
\end{aligned}$$

.

and the $Var(Y) =$ ∎