Quiz 15 : Linearity of Expectation Solutions

written by Alvin Wan . alvinwan.com/cs70 . Monday, October 31, 2016

This quiz does not count towards your grade. It exists to simply gauge your understanding. Treat this as though it were a portion of your midterm or final exam.

1 Independence

For the following, construct each of these examples if it is possible. If it not possible, write "impossible".

1. Construct a sample space and events X, Y so that $E[X]E[Y] = E[XY]$ but X, Y are not independent.

   Solution: Consider unit square centered about the origin. First, we know that X is not independent of Y, because for a given X, we know that Y either ranges on $[-1, 1]$ or it is in $\{-1, 1\}$. We know that $E[X] = 0, E[Y] = 0, E[XY] = 0$.

2. Construct a sample space and events X, Y so that $\text{var}(X+Y) = \text{var}(X)+\text{var}(Y)$, but X, Y are not independent.

   Solution: We can use the same example as above, with the unit square. This is because $\text{cov}(X, Y) = E[XY]-E[X]E[Y] = 0$, so $\text{var}(X+Y) = \text{var}(X)+\text{var}(Y)-2\text{cov}(X, Y) = \text{var}(X) + \text{var}(Y)$, as desired.

2 Linearity of Expectation

Bob is throwing a biased coin p. Let X be the number of heads that he throws. Compute $\text{var}(X)$. Do not apply linearity of variance (although it is valid here). Note that we are deriving the variance for a Binomial distribution.

Solution: First and foremost, we could have recognized that this is a Binomial distribution, and simply written the answer from memory (or copied from Wikipedia). This is an important takeaway: Recognizing distributions is just as important as understanding the calculations. In probability, there are often two ways of doing the problem: The straightforward way, or the clever way. This solution will take you through the straightforward way, to illustrate concepts you will need for future problems that may not fall into a pre-determined distribution.

We will introduce concepts as if for the first time. First, recall what an indicator random variable X_i is: a random variable that assumes the value 1 with probability p and 0 with probability $1-p$. Note that $E[X_i] = 1p + 0(1-p) = p$. In other words, the expectation of an indicator variable is *always* the probability of success.

In this way, X_i is 1 if the ith trial is a success. As a result, we can count successes by summing X_i.

$$X = \sum_i X_i$$

We first compute $E[X]$. Knowing that linearity of expectation always holds, regardless of the independence of X_is, we can do the following.

$$E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i]$$

Now, let's compute X_i. We note that this is a general result, for all indicator variables. Recall that an indicator variable has value 1 with probability p and 0 otherwise.

$$E[X_i] = 1p + 0(1-p) = p$$

In other words, the *expectation of a random variable is always its probability of success.* Thus, we have our result.

$$E[X] = \sum_{i=1}^{n} p = np$$

Let us now compute the nastier term, $E[X^2]$. Note that this is $E[(\sum_{i=1}^{n} X_i)^2]$, which seems quite ugly to compute. Note however, that when multiplying $(X_1 + X_2 \ldots X_n)(X_1 +$

$X_2 \dots X_n$), we have a bunch of terms of the form $X_1^2 + X_1X_2 + \dots X_1X_n + X_2^2 \dots$. Considering all X_iX_j, we either have that $i = j$ or $i \neq j$. In other words:

$$
\begin{aligned}
E[(\sum_{i=1}^{n} X_i)^2] &= E[\sum_{i=1}^{n} X_i^2 + \sum_{i \neq j} X_iX_j] \\
&= \sum_{i=1}^{n} E[X_i^2] + \sum_{i \neq j} E[X_iX_j]
\end{aligned}
$$

We can compute $E[X_iX_j]$ first. We can note that the term is only non-zero when both the ith and jth trials are heads. Alternatively, we can expand this out.

$$
\begin{aligned}
E[X_iX_j] &= \sum_{i,j} x_ix_j \Pr(X_i, X_j) \\
&= 1(1)p^2 + 1(0)p(1-p) + 0(1)p(1-p) + 0(0)(1-p)^2 \\
&= p^2
\end{aligned}
$$

Finally, we reason about how many terms $i \neq j$, terms there are. We can pick any X_i (n possibilities), then simply pick an X_j where $i \neq j$ (n-1 possibilities). As a result, we have $n(n-1)$ total values.

$$
\sum_{i \neq j} E[X_iX_j] = n(n-1)p^2
$$

Let us now compute $E[\sum_i X_i^2]$.

$$
\sum_i E[X_i^2] = \sum_i (1^2 p + 0^2(1-p)) = np
$$

Finally, we have an expression for $E[X^2]$.

$$
E[X^2] = np + n(n-1)p^2
$$

Given both $E[X^2]$ and $E[X]^2$, we can now compute $\mathrm{var}(X)$.

$$\begin{aligned}
\mathrm{var}(X) &= E[X^2] - E[X]^2 \\
&= np + n(n-1)p^2 - n^2p^2 \\
&= np + n^2p^2 - np^2 - n^2p^2 \\
&= np(1-p)
\end{aligned}$$