8.1 Linearity of expectation:

Now that we have covered joint pdf and joint pmfs, we will cover 3 important concepts:

1) Expectation of sums (Linearity of Expectation)

2) Variance of sums

3) Covariance
Ex: let $X$ and $Y$ represent the values of two dice rolls. I want to find the expected value

$$E[aX + bY^2]$$

Linearity tells you

$$= aE[X] + bE[Y^2]$$

we will prove linearity in this example,

$$g(X, Y) = aX + bY^2$$

$$E[g(X, Y)] = \sum_{i_1, j} (ai + bj^2) k_{X,Y}(i_1, j)$$

$$= \sum_{i=1}^6 \sum_{j=1}^6 (ai + bj^2) \cdot \frac{1}{6} \cdot \frac{1}{6}$$
\[\sum_{i=1}^{6} \sum_{j=1}^{6} a_i \cdot \frac{1}{6} \cdot \frac{1}{6} + \sum_{i=1}^{6} \sum_{j=1}^{6} b_j \cdot \frac{1}{6} \cdot \frac{1}{6}\]

\[P = \sum_{i=1}^{6} \frac{1}{6} \left( \sum_{j=1}^{6} a_i \cdot \frac{1}{6} \right) = \sum_{i=1}^{6} a_i \cdot \frac{1}{6}\]

\[= a \sum_{i=1}^{6} \frac{1}{6} = aE[X].\]

Similarly, \[Q = bE[Y^2].\]

\[E[aX + bY^2] = aE[X] + bE[Y^2].\]

Nowhere in this calculation have we used that \(X\) and \(Y\) came from independent die rolls.
Let us state this as a theorem:

**Linearity:** Let $g_1, \ldots, g_n$ be single variable functions $x_1, \ldots, x_n$ be r.v.s. Then

$$E\left[ g(x_1) + \cdots + g(x_n) \right] = E[g(x_1)] + \cdots + E[g(x_n)]$$

In the above example we had $(x, y)$

$$g_1(x) = ax \quad g_2(y) = by^2$$
Binomial is a sum of indicators

\[ X_i = \begin{cases} 1 & \text{ if } \text{ heads on } i^{th} \text{ coin} \\ 0 & \text{ otherwise} \end{cases} \]

\[ X_1 + \cdots + X_n = S_n \text{ (binomial)} \]

\[ E[X_1 + \cdots + X_n] = E[X_1] + \cdots + E[X_n] \]

\[ = n \cdot \frac{1}{2} = E[S_n] \]

A binomial is a sum of (independent) Bernoulli's

We have mentioned this earlier, but it is worth repeating.
Ex 8.6

There is a party every month in an office with 15 coworkers. There is ONE party at the end of the month if ANY coworker has a birthday in that month. Find the average # of birthdays partyed in a year.

Stingy management; they don’t want to pay for 15 parties. So they want to combine several birthdays into one by having a single party at the end of the month.
We would like to set this average up as a sum of random variables so we can use linearity of expectation.

How do you set this up as a sum?

\[
Y = \text{total # of parties in a year.}
\]

\[
Y = J_1 + J_2 + \cdots + J_{12}
\]

\[
J_i = \text{# of parties in month } i.
\]

Each \( J_i \) takes the values 0 or 1.

Is \( J_1 \) independent of \( J_2 \)?

\[
\text{POLL}
\]

Yes | No.

Let’s compute:

\[
P(J_1 = 0) = \left( \frac{11}{12} \right)^{15}
\]

All 15 individuals do not have a birthday in month 1.
\[ P(J_2 = 0 \mid J_1 = 0) = \left( \frac{10}{11} \right)^{15} \]

(One has to think about this.)

So
\[ P(J_2 = 0 \mid J_1 = 0) = \left( \frac{10}{11} \right)^{15} \left( \frac{11}{12} \right)^{15} \]

\[ 
eq \ P(J_2 = 0) \ P(J_1 = 0) \]

\[ = \]

**BUT**

\[ E[Y] = \sum_{i=1}^{12} E[J_i] = 12 \cdot E[J_1] \]

\[ E[J_1] = 1 \cdot P(J_1 = 1) + 0 \cdot P(J_1 = 0) \]

\[ = 1 - \left( \frac{11}{12} \right)^{15} \]

Ex: Manu must pass both a written test and a road test to get a driver's license. Each time he takes the written test, he passes with a probability $\frac{4}{10}$ independently of other tests. Each time he takes the road test, he passes with probability $\frac{7}{10}$, again independently of other tests. What is the total expected # of tests Manu must take before earning his license?

Hint: Let $X_1$ be the number of written tests he takes before passing, and let $X_2$ be the # of road tests.
Ex: (Another version of the indicator method)

The sequence of coin flips $HHTHTHTTHTHT$ contains two runs of heads of length 3
one run of heads of length 2
no runs of heads of length 1

If a fair coin is flipped 100 times, what is the expected number of runs of heads of length 3?

$\Omega = \{ H_{i1}^{100} : \omega_1, \ldots, \omega_{100} : \omega_1 \in \{ H, T \} \}$

Let $I_k = \{ A \text{ run of heads of length 3 begins at toss } k \}$

$= \{ \omega : \omega_k, \omega_{k+1}, \omega_{k+2} = H, \omega_{k-1} = T, \omega_{k+4} = T \}$

Can work if $k = 1$ Cannot work if $k = 97$

$I_1 + I_{97} + \sum_{k=1}^{96} I_k = \# \text{ of runs of length 3}$

$E[I_k] = 0$

$E[I_{97}]= $

$E[I_1]= $