Stat 322 - Solution to Homework 4

1 Finding Antony Smith

You are trying to locate an old high school friend who lives in Chicago. Unfortunately, your friend's name is Anthony Smith and the Chicago phone book lists phone numbers for 24 different people named Anthony Smith.

a) If you call 10 of these Anthony Smith's at random, what is the probability that you will call your friend? (Assume that your friend's phone number is listed in the phone book, and that you don't call anybody twice.)

- *There are $\binom{24}{10}$ different ways to pick 10 people out of a set of 24 people. The number of possibilities, where one of them is your friend is $\binom{1}{1} \cdot \binom{23}{9}$.*
  in total this gives a probability of

$$\frac{\binom{1}{1} \cdot \binom{23}{9}}{\binom{24}{10}} = \frac{1 \cdot \frac{23!}{9!14!}}{\frac{24!}{10!14!}} = \frac{10}{24} = 0.42.$$

b) Let X be the number of calls you need to make until you find your friend. Give the probability mass function for X (note that at most you need to make 24 calls).

When looking for a pmf, we always need to think about the possible values for X first: in this example, we need one phone call at least and 24 at most. Therefore, $im(X) = \{1, 2, ..., 24\}$.

Now we are going to compute the probability mass function for the first few cases, and try to find the general principle behind it:

$$\begin{aligned}
p_X(1) &= P(X=1) = \frac{1}{24} \\
p_X(2) &= P(X=2) = \underbrace{\frac{23}{24}}_{\text{failure first}} \cdot \underbrace{\frac{1}{23}}_{\text{success in 2nd}} = \frac{1}{24} \\
p_X(3) &= P(X=3) = \underbrace{\frac{23}{24} \cdot \frac{22}{23}}_{\text{2 failures first}} \cdot \underbrace{\frac{1}{22}}_{\text{success in 3rd}} = \frac{1}{24} \\
&\dots
\end{aligned}$$

It seems, that we get a constant probability mass function:

$$p_X(k) = P(\text{ exactly } k \text{ attempts }) = \frac{1}{24}.$$

This, actually is a probability mass function (only values between 0 and 1; the sum of all is 1).

c) How many calls do you expect to make until you find your friend? (Again, assume that your friend's phone number is listed in the phone book, and that you don't call anybody twice.)

For the expected value, we need the probability mass function.

$$E[X] = \sum_{k=1}^{24} k \cdot p_X(k) = \sum_{k=1}^{24} k \cdot \frac{1}{24} = \frac{1}{24} \cdot \sum_{k=1}^{24} k = \frac{1}{24} \cdot 300 = \underline{12.5}$$

2 Snow in October

After October 1st the probability that a blizzard will occur on any particular day in the Midwest of Northern America is 0.1.
To simplify the problem, assume for the following questions that this probability is constant from Oct 1st onwards.

a) What is the probability that there won't be a blizzard in the first 28 days of October?

The probability that there will not be a blizzard is then $.9^{28} = 0.052$.

b) What is the probability that the first blizzard will occur on October 14?

$P($ *no blizzard in twelve days, blizzard on 14th* $) = 0.9^{13} \cdot 0.1 = 0.0254$.

c) What is the expected date for the first blizzard (starting October 1st)?

For the expected value, we need some random variable X *together with its probability mass function. One way to define* X *is* $X :=$ *" date of first blizzard" or, since it's not easy to compute with dates* X *= "# of days until first blizzard including day of blizzard".*

This random variable has a geometric distribution with $p = 0.1$*, we write* $X \sim Geo_{0.1}$.

For a geometric random variable we already know the expected value:

$$E[X] = \frac{1}{p} = \frac{1}{0.1} = 10.$$

The expected date for the first blizzard would therefore be Oct 10th.

3 Tossing a Coin

A fair coin is tossed 10 times. Let X be the random variable corresponding to the difference between the number of heads and the number of tails observed (i.e (*# of H*)-(*# of T*)).

a) Find the image of X and compute the probability mass function of X. (Hint: Instead of looking at X directly, define a new random variable Y, which just counts the number of heads in 10 tosses. What is the relationship between X and Y?)

If we have a look at Y*, we see that* Y *is a binomial random variable with* $n = 10$ *and* $p = 0.5$*. The image of* Y *therefore is* $\{0, 1, ..., 10\}$ *and* $p_Y(k) = \binom{10}{k} 0.5^n$.

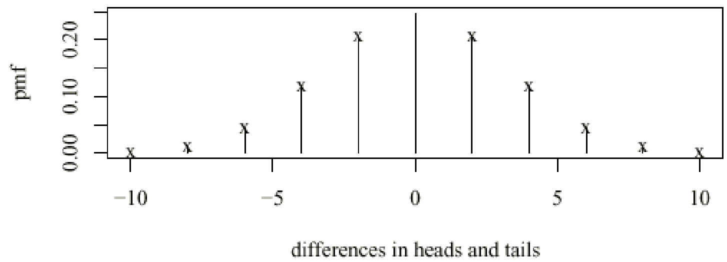
The relationship between X *and* Y *is fairly simple: If* Y *is the number of heads in 10 tosses, then we have 10-*Y *tails. The difference between heads and tails in 10 tosses is therefore* $Y - (10 - Y) = X$*. i.e* $X = 2Y - 10$.

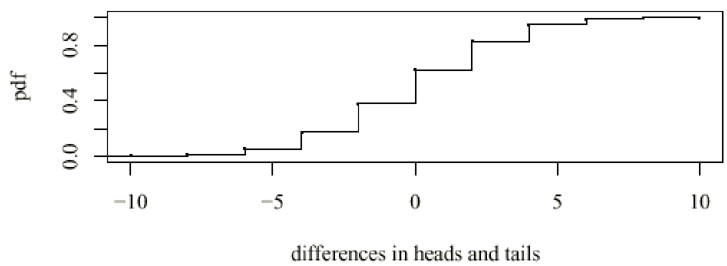
For the image of X *we therefore get all even numbers between -10 and 10,* $im(X) = \{-10, -8, -6, ..., 6, 8, 10\}$.

What about the probability mass function of X*? - we can also use, what we know about* Y*:*

$$p_X(k) \quad = \quad P(X = k) = P(2Y - 10 = k) = P(Y = \frac{k+10}{2}) = \binom{10}{\frac{k+10}{2}} 0.5^{10}.$$

b) Draw a diagram to show the probability mass function. Would you be able to tell the expected value of X from the diagram? Explain.





In this particular case, we can read the expected value from the probability mass function directly: whenever the probability mass function is symmetric, the expected value is exactly in the middle of the x-values. Here, $E[X] = 0$. *You can also get this by using properties of expectation:* $E(Y) = 5$; *so,* $E(X) = 2E(Y) - 10 = 0$.

c) Compute the variance and standard deviation of X.

Instead of working from the scratch, we, again, will use the relationship between X *and* Y. *For* Y, *we know the variance, since it is a binomial distribution:* $Var[Y] = np(1-p) = 10/4 = 2.5$.

$$Var[X] = Var[2Y - 10] = Var[2Y] + Var[10] = 4Var[Y] + 0 = 4 \cdot 2.5 = 10.$$

and therefore its standard deviation $\sigma_X = \sqrt{10}$.

4 Deaths from Horse Kicks

After analyzing data from the Prussian cavalry for a period of 20 years, statisticians came to the conclusion that the number of deaths from horse kicks follows a Poisson distribution with $\lambda = 0.6$.
Let X be the number of deaths from horse kicks in 20 years.
What is the probability that

a) no deaths occurred?

X has a $Po_{.6}$ distribution.

$$P(X = 0) = 0.6^0 \frac{e^{-0.6}}{0!} = e^{-0.6} = 0.5488$$

b) one or two deaths occurred?

$$P(1 \leq X \leq 2) = P(X = 1) + P(X = 2) = 0.6^1 \frac{e^{-0.6}}{1!} + 0.6^2 \frac{e^{-0.6}}{2!} = 0.4281$$

c) 3 or more deaths occurred?

$$P(X \geq 3) = 1 - P(X \leq 2) = 1 - 0.9769 = 0.0231$$

d) more than 3 deaths occurred?

$$P(X > 3) = 1 - P(X \leq 3) = 1 - 9966 = 0.0034$$

5 Discrete Compound PMFs

The joint pmf of two discrete r.v. X and Y is given as:

$X \backslash Y$	-1	0	1
-2	$1/16$	$1/16$	$1/16$
-1	$1/8$	$1/16$	$1/8$
1	$1/8$	$1/16$	$1/8$
2	$1/16$	$1/16$	$1/16$

(a) Find the following probabilities:

  (a) $P(X \geq 2) = 1/16 + 1/16 + 1/16 = 3/16$.

  (b) $P(X > Y) = 1/8 + 1/16 + 1/16 + 1/16 + 1/16 = 6/16$.

  (c) $P(Y > 0) = 1/16 + 1/8 + 1/8 + 1/16 = 6/16$.

(b) Are X and Y independent? *X and Y are not independent:* $p_{X,Y}(1,1) = 1/8 \neq 6/16 \cdot 5/16 = p_X(1) \cdot p_Y(1)$

(c) Are X and Y uncorrelated? *we need to check, whether the correlation between X and Y is zero (then the variables are uncorrelated).*

The definition for correlation is

$$corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

Therefore, we need to know the Variances and Covariance. For that, we need th expected values of X and Y.

For the expected values, we need the marginal probability mass functions:

X	-2	-1	1	2
p_X	$3/16$	$5/16$	$5/16$	$3/16$

Y	-1	0	1
p_Y	$6/16$	$4/16$	$6/16$

The expected values $E[X]$ and $E[Y]$ are:

$$\begin{aligned} E[X] &= -2 \cdot 3/16 + (-1) \cdot 5/16 + 1 \cdot 5/16 + 2 \cdot 3/16 = 0, \\ E[Y] &= -1 \cdot 6/16 + 0 \cdot 4/16 + 1 \cdot 6/16 = 0. \end{aligned}$$

The variances then are:

$$\begin{aligned} Var[X] &= (-2-0)^2 \cdot 3/16 + (-1)^2 \cdot 5/16 + 5/16 + 2^2 \cdot 3/16 = (12+5+5+12)/16 = 34/16, \\ Var[Y] &= 12/16 \end{aligned}$$

The covariance between X and Y is defined as $E[(X - E[X])(Y - E[Y])]$*:*

$$\begin{aligned}
Cov(X,Y) &\overset{E[X]=E[Y]=0}{=} E[X \cdot Y] = \\
&= (-2) \cdot (-1) \cdot 1/16 + (-2) \cdot 0 \cdot 1/16 + (-2) \cdot 1 \cdot 1/16 + \\
&\quad +1/8 + 0 - 1/8 + \\
&\quad -1/8 + 0 + 1/8 + \\
&\quad -2/16 + 0 + 2/16 = 0.
\end{aligned}$$

With this, we know, the correlation is 0, too. Therefore, the variables are uncorrelated.