Math 224 **Spring 2018**
Homework 4 **Drew Armstrong**

---

1 (St. Petersburg Paradox). I am running a lottery. I will let you flip a fair coin until you get heads. If the first head shows up on the k-th flip I will pay you r^k dollars.

(a) Compute your expected winnings when $r = 1$.
(b) Compute your expected winnings when $r = 1.5$.
(c) Compute your expected winnings when $r = 2$. Does this make any sense? How much would you be willing to pay me to play this game?

[Moral of the Story: The expected value is not always meaningful.]

Let X be the number of flips until you see heads. This is a geometric random variable with

$$P(X = k) = P(T)^{k-1} \cdot P(H) = \left(\frac{1}{2}\right)^{k-1}\left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^{k}.$$

Thus your total expected winnings is

$$E[\text{winnings}] = \sum_{k=1}^{\infty} r^k \cdot P(X = k) = \sum_{k=1}^{\infty} r^k \left(\frac{1}{2^k}\right) = \sum_{k=1}^{\infty} \left(\frac{r}{2}\right)^k.$$

As long as $|r/2| < 1$ then we can simplify this with a geometric series:

$$\begin{aligned} E[\text{winnings}] &= (r/2) + (r/2)^2 + (r/2)^3 + \cdots \\ &= (r/2) \cdot \left[1 + (r/2) + (r/2)^2 + \cdots\right] \\ &= (r/2) \cdot \frac{1}{1 - r/2} \\ &= (r/2) \cdot \frac{1}{(2-r)/2} = \frac{r}{2-r}. \end{aligned}$$

(a) Plugging in $r = 1$ gives $E[\text{winnings}] = \$1$.

(b) Plugging in $r = 1.5$ gives $E[\text{winnings}] = \$3$.

(c) If $r = 2$ then the formula doesn't work. Instead we get a divergent series:

$$E[\text{winnings}] = \sum_{k=1}^{\infty} \left(\frac{2}{2}\right)^k = 1 + 1 + 1 + \cdots.$$

This seems to indicate that your expected winnings are infinite. But that doesn't make any sense because there is no such thing as an infinite amount of money. In reality, your winnings will be 2^k only for small values of k. Then when 2^k exceeds all the money I have, I'll just give you all of my money. For example, suppose that I have $2^5 = \$32$ in the bank. Then your expected winnings are

$$\begin{aligned}
E[\text{winnings}] &= \frac{2}{2} + \frac{2^2}{2^2} + \frac{2^3}{2^3} + \frac{2^4}{2^4} + \frac{2^5}{2^5} + \frac{2^5}{2^6} + \frac{2^5}{2^7} + \frac{2^5}{2^8} + \cdots \\
&= 1 + 1 + 1 + 1 + 1 + 2^5\left(\frac{1}{2^6} + \frac{1}{2^7} + \frac{1}{2^8} + \cdots\right) \\
&= 1 + 1 + 1 + 1 + 1 + \frac{1}{2}\left(1 + \frac{1}{2} + \frac{1}{4} + \cdots\right) \\
&= 1 + 1 + 1 + 1 + 1 + \frac{1}{2}(2) \\
&= \$6.
\end{aligned}$$

More generally, if I have 2^n dollars in the bank, then you should expect to win $n+1$ dollars from this game.

2. I am running a lottery. I will sell 50 million tickets, 5 million of which will be winners.

(a) If you purchase 10 tickets, what is the probability of getting at least one winner?
(b) If you purchase 15 tickets, what is the probability of getting at least one winner?
(c) If you purchase n tickets, what is the probability of getting at least one winner?
(d) What is the smallest value of n such that your probability of getting a winner is greater than 50%? What is the smallest value of n that gives you a 95% chance of winning?

[Hint: If n is small, then each ticket is approximately a coin flip with $P(H) = 1/10$. In other words, for small values of n we have the approximation

$$\binom{45,000,000}{n} \Big/ \binom{50,000,000}{n} \approx (9/10)^n.]$$

Suppose that you purchase n tickets and let X be the number of winning tickets you get. Then X is a hypergeometric random variable with pmf

$$P(X = k) = \frac{\binom{5,000,000}{k}\binom{45,000,000}{n-k}}{\binom{50,000,000}{n}}.$$

However, if n is small[1] then it is reasonable to treat each ticket as a coin flip with $P(\text{win}) = 1/10$ and $P(\text{lose}) = 9/10$. Then our hypergeometric random variable is approximately binomial:

$$P(X = k) = \frac{\binom{5,000,000}{k}\binom{45,000,000}{n-k}}{\binom{50,000,000}{n}} \approx \binom{n}{k}\left(\frac{1}{10}\right)^k\left(\frac{9}{10}\right)^{n-k}.$$

(a) If $n = 10$ then $P(X \geq 1) = 1 - P(X = 0) \approx 1 - (9/10)^{10} = 65.13\%$.

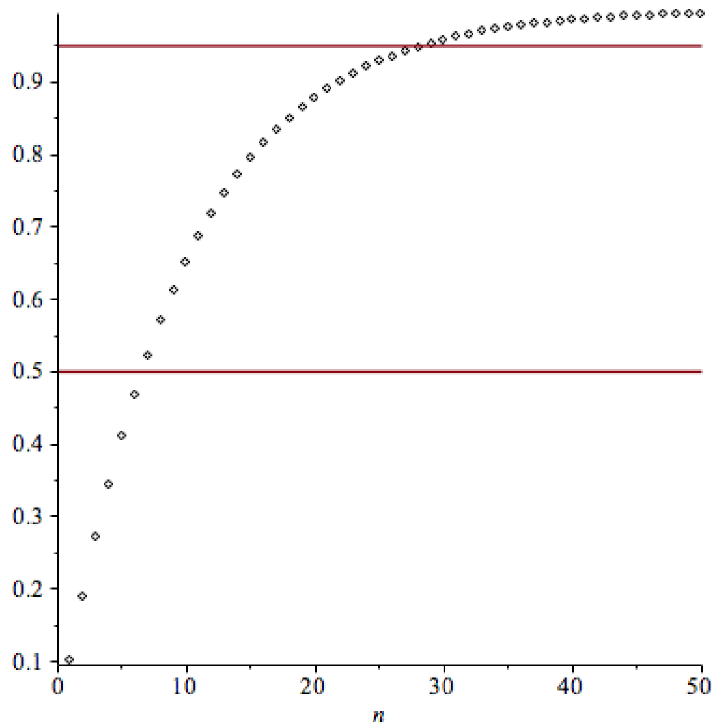
(b) If $n = 15$ then $P(X \geq 1) = 1 - P(X = 0) \approx 1 - (9/10)^{15} = 79.41\%$.

(c) For general small values of n, the probability of getting at least 1 winning ticket is

$$P(X \geq 1) = 1 - P(X = 0) \approx 1 - (9/10)^n.$$

Here is a plot of the probability $P(X \geq 1)$ for values of n from 1 to 50:

[1] in comparison to $5,000,000$



From the diagram it seems that the probability crosses 0.5 between $n = 6$ and $n = 7$, and the probability crosses 0.95 when n is around 30. To be precise, we have

$$\begin{aligned} n = 6 &\quad \to \quad P(X \geq 1) \approx 46.86\% \\ n = 7 &\quad \to \quad P(X \geq 1) \approx 52.17\% \\ n = 28 &\quad \to \quad P(X \geq 1) \approx 94.77\% \\ n = 29 &\quad \to \quad P(X \geq 1) \approx 95.29\%. \end{aligned}$$

3. Flip a fair coin 3 times and let

$$X = \text{“number of heads squared, minus the number of tails.”}$$

(a) Write down a table showing the pmf of X.
(b) Compute the expected value $\mu = E[X]$.
(c) Compute the variance $\sigma^2 = \mathrm{Var}(X)$.
(d) Draw the line graph of the pmf. Indicate the values of $\mu - \sigma, \mu, \mu + \sigma$ in your picture.

(a) Let Y be the number of heads, so that $3 - Y$ is the number of tails, and $X = Y^2 - (3 - Y)$. The probability of getting $Y = \ell$ is the same as the probability of $X = k$ where $k = \ell^2 - (3 - \ell)$. Therefore we have the following pmf table:

ℓ	0	1	2	3
$P(Y = \ell)$	1/8	3/8	3/8	1/8
k	-3	-1	3	9
$P(X = k)$	1/8	3/8	3/8	1/8

(b) By definition we have

$$\mu = E[X] = \sum_k k \cdot P(X = k) = (-3)\frac{1}{8} + (-1)\frac{3}{8} + (3)\frac{3}{8} + (9)\frac{1}{8} = \frac{12}{8} = 1.5.$$

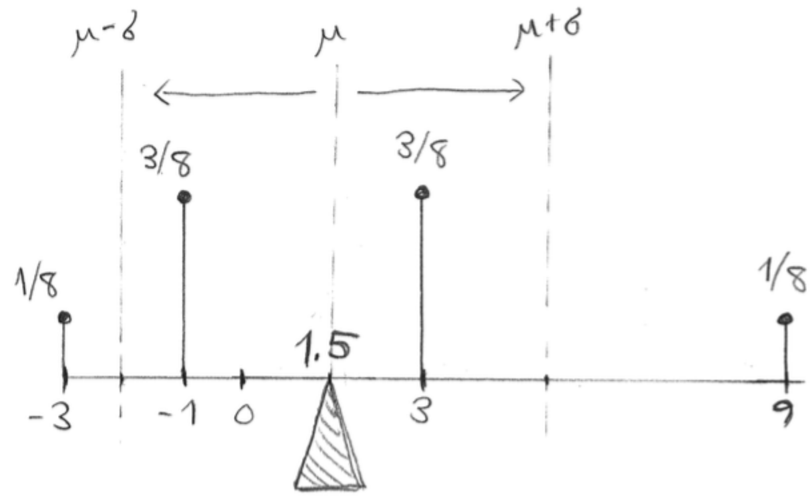
(c) First we compute the second moment:

$$E[X^2] = \sum_k k^2 \cdot P(X = k) = (-3)^2\frac{1}{8} + (-1)^2\frac{3}{8} + (3)^2\frac{3}{8} + (9)^2\frac{1}{8} = \frac{120}{8} = 15$$

Then we compute the variance:

$$\sigma^2 = \mathrm{Var}(X) = E[X^2] - E[X]^2 = 15 - (1.5)^2 = 12.75.$$

(d) The standard deviation is $\sigma = \sqrt{12.75} = 3.57$. Here is the line graph for the pmf of X, with the interval $\mu \pm \sigma$ indicated:



4. Let X and Y be random variables with supports $S_X = \{1, 2\}$ and $S_Y = \{1, 2, 3, 4\}$, and with joint pmf given by the formula

$$f_{XY}(k, \ell) = P(X = k, Y = \ell) = \frac{k + \ell}{32}.$$

(a) Draw the joint pmf table, showing the marginal probabilities in the margins.
(b) Compute the following probabilities directly from the table:

$$P(X > Y), \quad P(X \leq Y), \quad P(Y = 2X), \quad P(X + Y \leq 3), \quad P(X + Y > 3).$$

(c) Use the marginal distributions to compute $E[X], \mathrm{Var}(X)$ and $E[Y], \mathrm{Var}(Y)$.
(d) Use the table to compute the pmf of XY. Use this to compute $E[XY]$ and $\mathrm{Cov}(X, Y)$.
(e) Compute the correlation coefficient:

$$\rho_{XY} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}.$$

Are the random variables X, Y independent? Why or why not?

(a) Here is the joint pmf table:

X \ Y	1	2	3	4	
1	$\frac{2}{32}$	$\frac{3}{32}$	$\frac{4}{32}$	$\frac{5}{32}$	$\frac{14}{32}$
2	$\frac{3}{32}$	$\frac{4}{32}$	$\frac{5}{32}$	$\frac{6}{32}$	$\frac{18}{32}$
	$\frac{5}{32}$	$\frac{7}{32}$	$\frac{9}{32}$	$\frac{11}{32}$	

(b) To compute these probabilities we just add the probabilities from the relevant cells:

$$\begin{aligned} P(X > Y) &= 3/32 \\ P(X \leq Y) &= 1 - P(X > Y) = 29/32 \\ P(Y = 2X) &= 3/32 + 9/32 = 9/32 \\ P(X + Y \leq 3) &= 2/32 + 3/32 + 3/32 = 8/32 \\ P(X + Y > 3) &= 1 - P(X + Y \leq 3) = 24/32. \end{aligned}$$

(c) We use the marginal pmf of X to compute:

$$\begin{aligned} E[X] &= (1)\frac{14}{32} + (2)\frac{18}{32} = \frac{25}{16} = 1.56 \\ E[X^2] &= (1)^2\frac{14}{32} + (2)^2\frac{18}{32} = \frac{43}{16} = 2.69 \\ \mathrm{Var}(X) &= E[X^2] - E[X]^2 = \frac{43}{16} - \left(\frac{25}{16}\right)^2 = \frac{63}{256} = 0.25 \end{aligned}$$

And we use the marginal pmf of Y to compute:

$$\begin{aligned} E[Y] &= (1)\frac{5}{32} + (2)\frac{7}{32} + (3)\frac{9}{32} + (4)\frac{11}{32} = \frac{45}{16} = 2.81 \\ E[Y^2] &= (1)^2\frac{5}{32} + (2)^2\frac{7}{32} + (3)^2\frac{9}{32} + (4)^2\frac{11}{32} = \frac{145}{16} = 9.06 \\ \mathrm{Var}(Y) &= E[Y^2] - E[Y]^2 = \frac{43}{16} - \left(\frac{25}{16}\right)^2 = \frac{295}{256} = 1.15 \end{aligned}$$

(d) First we circle the blobs corresponding to different values of XY:

$X \setminus Y$	1	2	3	4	
1	$\frac{2}{32}$	$\frac{3}{32}$	$\frac{4}{32}$	$\frac{5}{32}$	$\frac{14}{32}$
2	$\frac{3}{32}$	$\frac{4}{32}$	$\frac{5}{32}$	$\frac{6}{32}$	$\frac{18}{32}$
	$\frac{5}{32}$	$\frac{7}{32}$	$\frac{9}{32}$	$\frac{11}{32}$	

Then we add them up:

k	1	2	3	4	6	8
$P(XY = k)$	$\frac{2}{32}$	$\frac{6}{32}$	$\frac{4}{32}$	$\frac{9}{32}$	$\frac{5}{32}$	$\frac{6}{32}$

We use this compute:

$$E[XY] = (1)\frac{2}{32} + (2)\frac{6}{32} + (3)\frac{4}{32} + (4)\frac{9}{32} + (6)\frac{5}{32} + (8)\frac{6}{32} = \frac{140}{32} = 4.38$$
$$\mathrm{Cov}(X,Y) = E[XY] - E[X] \cdot E[Y] = \frac{140}{32} - \frac{25}{16} \cdot \frac{45}{16} = -\frac{5}{256} = -0.02.$$

(e) The coefficient of correlation is

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \cdot \sigma_Y} = \frac{-5/256}{\sqrt{63/256} \cdot \sqrt{295/256}} = \frac{-5}{\sqrt{63} \cdot \sqrt{295}} = -0.037,$$

hence X and Y are (ever so slightly) negatively correlated.

5. Let X and Y be random variables with the following joint distribution:

$X \setminus Y$	-1	0	1
-1	0	0	1/4
0	1/2	0	0
1	0	0	1/4

(a) Compute the numbers $E[X], \mathrm{Var}(X)$ and $E[Y], \mathrm{Var}(Y)$.
(b) Compute the expected value $E[XY]$ and the covariance $\mathrm{Cov}(X,Y)$.
(c) Are the random variables X, Y independent? Why or why not?

[Moral of the Story: Uncorrelated does not always imply independent.]

First we fill in the marginal distributions:

$X \setminus Y$	-1	0	1	
-1	0	0	1/4	1/4
0	1/2	0	0	1/2
1	0	0	1/4	1/4
	1/2	0	1/2	

(a) Now we use the marginal distributions to compute:

$$\begin{aligned}
E[X] &= (-1)\frac{1}{4} + (0)\frac{1}{2} + (1)\frac{1}{4} = 0 \\
E[X^2] &= (-1)^2\frac{1}{4} + (0)^2\frac{1}{2} + (1)^2\frac{1}{4} = \frac{1}{2} \\
\mathrm{Var}(X) &= E[X^2] - E[X]^2 = \frac{1}{2} \\
E[Y] &= (-1)\frac{1}{2} + (0)0 + (1)\frac{1}{2} = 0 \\
E[Y^2] &= (-1)^2\frac{1}{2} + (0)0 + (1)^2\frac{1}{2} = 1 \\
\mathrm{Var}(Y) &= E[Y^2] - E[Y]^2 = 1.
\end{aligned}$$

(b) Instead of first computing the pmf of XY I will compute $E[XY]$ directly by summing over the whole table:

$$\begin{aligned}
E[XY] &= \sum_{k,\ell} k \cdot \ell \cdot P(X = k, Y = \ell) \\
&= (-1)(-1)0 + (-1)(0)0 + (-1)(1)1/4 \\
&\quad (0)(-1)1/2 + (0)(0)0 + (0)(1)1/2 \\
&\quad (1)(-1)0 + (1)(0)0 + (1)(1)1/4 \\
&= -1/4 + 1/4 = 0.
\end{aligned}$$

Hence we have

$$\mathrm{Cov}(X, Y) = E[XY] - E[X] \cdot E[Y] = 0 - 0 \cdot 0 = 0.$$

(c) However, the random variables X, Y are **not** independent because, for example,

$$P(X = 0, Y = 1) = 0 \neq \frac{1}{2} \cdot \frac{1}{2} = P(X = 0) \cdot P(Y = 1).$$

6. Roll a fair 6-sided die twice. Let X be the number that shows up on the first roll and let Y be the number that shows up on the second roll. You may assume that X and Y are independent.

(a) Compute the covariance $\mathrm{Cov}(X, Y)$.
(b) Compute the covariance $\mathrm{Cov}(X, X + Y)$.
(c) Compute the covariance $\mathrm{Cov}(X, 2X + 3Y)$.

Here is the pmf of X:

k	1	2	3	4	5	6
$P(X = k)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Hence we have

$$E[X] = (1)\frac{1}{6} + (2)\frac{1}{6} + (3)\frac{1}{6} + (4)\frac{1}{6} + (5)\frac{1}{6} + (6)\frac{1}{6} = \frac{7}{2} = 3.5$$

$$E[X^2] = (1)^2\frac{1}{6} + (2)^2\frac{1}{6} + (3)^2\frac{1}{6} + (4)^2\frac{1}{6} + (5)^2\frac{1}{6} + (6)^2\frac{1}{6} = \frac{91}{6} = 15.17$$

$$\mathrm{Var}(X) = E[X^2] - E[X]^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12} = 2.92.$$

The distribution of Y is the same.

(a) Since X and Y are independent we have $\mathrm{Cov}(X, Y) = 0$.

(b) Using bilinearity of covariance gives

$$\mathrm{Cov}(X, X + Y) = \mathrm{Cov}(X, X) + \mathrm{Cov}(X, Y) = \mathrm{Var}(X) + 0 = \frac{35}{12}.$$

(c) Using bilinearity of covariance gives

$$\mathrm{Cov}(X, 2X + 3Y) = 2 \cdot \mathrm{Cov}(X, X) + 3 \cdot Cov(X, Y) = 2 \cdot \mathrm{Var}(X) + 3 \cdot 0 = \frac{35}{6}.$$

7. Let X_1 and X_2 be independent samples from a distribution with the following pmf:

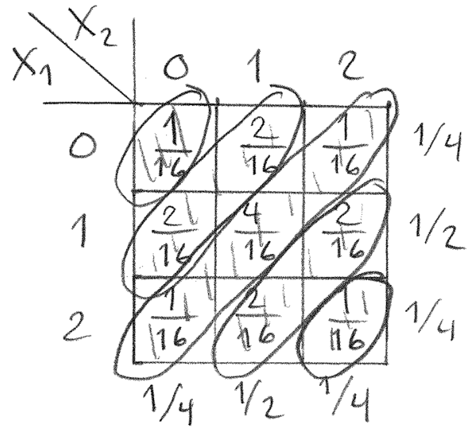
k	0	1	2
$f(k)$	1/4	1/2	1/4

(a) Draw the joint pmf table of X_1 and X_2.
(b) Use your table to compute the pmf of $X_1 + X_2$.
(c) Compute the variance $\mathrm{Var}(X_1 + X_2)$ in two different ways.

(a) Since X_1 and X_2 are independent we can multiply their marginal pmf's to obtain the joint pmf table:

$X_1 \backslash X_2$	0	1	2	
0	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	1/4
1	$\frac{2}{16}$	$\frac{4}{16}$	$\frac{2}{16}$	1/2
2	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	1/4
	1/4	1/2	1/4	

(b) First we draw blobs for the different values of $X_1 + X_2$:



Then we add the entries in the blobs to obtain the pmf table of $X_1 + X_2$:

k	0	1	2	3	4
$P(X_1 + X_2 = k)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

(c) We can compute the variance $\text{Var}(X_1) + \text{Var}(X_2)$ in two ways. One way is directly from the pmf table:

$$\begin{aligned}
E[X_1 + X_2] &= (0)\frac{1}{16} + (1)\frac{4}{16} + (2)\frac{6}{16} + (3)\frac{4}{16} + (4)\frac{1}{16} = 2 \\
E[(X_1 + X_2)^2] &= (0)^2\frac{1}{16} + (1)^2\frac{4}{16} + (2)^2\frac{6}{16} + (3)^2\frac{4}{16} + (4)^2\frac{1}{16} = 5 \\
\text{Var}(X_1 + X_2) &= E[(X_1 + X_2)^2] - E[X_1 + X_2]^2 = 5 - 2^2 = 1.
\end{aligned}$$

Alternatively, we can compute the variance of X_i first:

$$\begin{aligned}
E[X_i] &= (0)\frac{1}{4} + (1)\frac{1}{2} + (2)\frac{1}{4} = 1 \\
E[X_i^2] &= (0)^2\frac{1}{4} + (1)^2\frac{1}{2} + (2)^2\frac{1}{4} = 3/2 \\
\text{Var}(X_i) &= E[X_i^2] - E[X_i]^2 = 3/2 - 1^2 = 1/2.
\end{aligned}$$

Then we use the fact that X_1 and X_2 are independent to compute

$$\begin{aligned}
\text{Var}(X_1 + X_2) &= \text{Var}(X_1) + \text{Var}(X_2) - 2 \cdot \text{Cov}(X_1, X_2) \\
&= \text{Var}(X_1) + \text{Var}(X_2) + 0 \\
&= \frac{1}{2} + \frac{1}{2} = 1.
\end{aligned}$$

Actually, there is a third way. We could observe that $X_1 + X_2$ has a binomial distribution with parameters $n = 4$ and $p = 1/2$. Thus the variance is

$$\text{Var}(X_1 + X_2) = npq = 4 \cdot \frac{1}{2} \cdot \frac{1}{2} = 1.$$

8. Each box of a certain brand of cereal comes with a toy inside. If there are n possible toys and if the toys are distributed randomly, how many boxes do you expect to buy before you get them all?

(a) Assuming that you already have ℓ of the toys, let X_ℓ be the number of boxes you need to purchase until you get a new toy that you don't already have. Compute the expected value $E[X_\ell]$. [Hint: We can think of each new box purchased as a "coin flip" where $H =$"we get a new toy" and $T =$"we don't get a new toy." Thus X_ℓ is a geometric random variable. What is $P(H)$?]

(b) Let X be the number of boxes you purchase until you get all n toys. Thus we have

$$X = X_0 + X_1 + X_2 + \cdots + X_{n-1}.$$

Use part (a) and linearity to compute the expected value $E[X]$.

(c) Application: Suppose you continue to roll a fair 6-sided die until you see all six sides. How many rolls do you expect to make?

(a) Fix $\ell \in \{0, 1, \ldots, n-1\}$ and assume that you already have ℓ of the toys. Until you get a new toy, each box of cereal can temporarily be thought of as a coin flip with $H =$ "a new toy" and $T =$ "an old toy." Since the n toys are distributed randomly this means that

$$P(H) = p = \frac{n-\ell}{n} \qquad \text{and} \qquad P(T) = q = \frac{\ell}{n}.$$

If X_ℓ is the "number of boxes until a new toy" (i.e., the "number of flips until heads") then we observe that X_ℓ is a geometric random variable. Hence we expect to get a new toy after purchasing $1/p$ boxes:

$$E[X_\ell] = \frac{1}{p} = \frac{n}{n-\ell}.$$

(b) If X is the total number of boxes we need to purchase until getting all n toys then we can express this as a sum of the random variables from part (a):

$$X = X_0 + X_1 + X_2 + \cdots + X_{n-1}.$$

Now we can apply linearity to compute the expected value:

$$\begin{aligned} E[X] &= E[X_0] + E[X_1] + E[X_2] + \cdots + E[X_{n-1}] \\ &= \frac{n}{n-0} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1}. \end{aligned}$$

Sadly, this formula does not simplify.

(c) Application: Start rolling a fair 6-sided die and let X be the number of rolls until you see all six faces. Then according to part (b) we have

$$\begin{aligned} E[X] &= \frac{6}{6} + \frac{6}{5} + \frac{6}{4} + \frac{6}{3} + \frac{6}{2} + \frac{6}{1} \\ &= 1 + 1.2 + 1.5 + 2 + 3 + 6 = 14.7. \end{aligned}$$

In other words, we see our 1st new face on the first roll. Then it takes 1.2 rolls to see the 2nd face and 1.5 rolls to see the 3rd face, etc. After we have seen 5 faces it will take on average 6 rolls until we see the final face. In total, we expect to roll the die 14.7 times.

[Moral of the Story: Linearity of expectation is useful.]