One assignment submitted per group. Show all of your work. Please staple in the upper left hand corner.

1. Identify the most plausible distribution (Bernoulli, Binomial, Geometric, Poisson) to model the following random variables and identify the expected value, the variance, and the probability that the observed outcome is 1:

   (a) **(Pts: 4)** The number of multiple choice questions a bot randomly guessing will get correct (20 questions, 4 choices for each question).
   Answer:
   $X \sim Binomial(n = 20, p = 1/4)$.
   $E[X] = n * p = 20 * (1/4) = 5$
   $Var[X] = n * p * (1 - p) = 20 * (1/4) * (3/4) = 3.75$
   $P(X = 1) = \binom{20}{1}(1/4)^1(3/4)^{19} = 0.02114141$

   (b) **(Pts: 4)** How many printing jobs a printer receives in an hour (on average it receives 50 printing jobs in an 8 hour day).
   Answer:
   $X \sim Poisson(\lambda = 6.25)$.
   $E[X] = \lambda = 6.25$
   $Var[X] = \lambda = 6.25$
   $P(X = 1) = \frac{e^{-6.25}6.25^1}{1!} = 0.01206534$

   (c) **(Pts: 4)** The number of people you need to meet before meeting somebody with the same birthday as you. Assume birthdays are perfectly uniformly distributed among the 365 days of the year (don't include February 29).
   Answer:
   $X \sim Geometric(p = 1/365)$.
   $E[X] = 1/p = 365$
   $Var[X] = (1 - p)/p^2 = 132860$
   $P(X = 1) = (1 - p)^0 p = 1/365 = 0.002739726$

   (d) **(Pts: 4)** Whether my dog will be excited to see me when I come home today (Use $X = 0$ for not excited, $X = 1$ for excited). Assume there is a probability of .9999 that he will be excited.
   Answer:
   $X \sim Bernoulli(p = .9999)$. ($X \sim B(1, .9999)$ also acceptable)
   $E[X] = p = .9999$
   $Var[X] = p * (1 - p) = .00009999$
   $P(X = 1) = p^1(1 - p)^0 = p = .9999$

2. Suppose I have a random variable X with probability mass function $f(x) = cx^2$ for $x = 1, 2, 3, 4, 5$ and $f(x) = 0$ for all other values.

   (a) **(Pts: 2)** What value does c need to be to make this a valid probability mass function?
   Answer: $1/55$
   We need $\sum_x f(x) = 1$
   We have $\sum_x f(x) = 1^2c + 2^2c + 3^2c + 4^2c + 5^2c = 55c$ which implies we need $c = \frac{1}{55}$

   (b) **(Pts: 4)**Find the CDF corresponding to this probability mass function.
   Answer:
   $F(x) = P(X \leq x)$ so we just need to sum the probabilities of the values of X less than or equal to the input to the CDF.

$$F(x) = \begin{cases} 0 & : x < 1 \\ 1/55 & : 1 \leq x < 2 \\ 5/55 & : 2 \leq x < 3 \\ 14/55 & : 3 \leq x < 4 \\ 30/55 & : 4 \leq x < 5 \\ 1 & : 5 \leq x \end{cases}$$

(c) **(Pts: 1)** What is $P(X = 3)$?
Answer: 9/55
$P(X = 3) = f(3) = 3^2 c = 9/55$

(d) **(Pts: 2)** What is $P(|X - 3| > 1)$?
Answer: 26/55
The only values of x that meet the condition $|X - 3| > 1$ and have non-zero probability are $X = 1$ and $X = 5$ so this probability statement is the same as

$$P(X = 1 or X = 5) = P(X = 1) + P(X = 5) = f(1) + f(5) = 1/55 + 25/55 = 26/55$$

(e) **(Pts: 2)** What is $P(X = 2|X \leq 3)$?
Answer: 4/14
Recall the definition of conditional probability: $P(A|B) = \frac{P(A and B)}{P(B)}$
So we have

$$\begin{aligned} P(X = 2|X \leq 3) &= \frac{P(X = 2 and X \leq 3)}{P(X \leq 3)} \\ &= \frac{P(X = 2)}{P(X \leq 3)} \\ &= \frac{f(2)}{F(3)} \\ &= \frac{4/55}{14/55} \\ &= 4/14 = .2857 \end{aligned}$$

We go from the first line to the second by recognizing that the only value of X that meet the conditions X=2 and X≤3 is X=2.

(f) **(Pts: 2)** What is $E[X]$?
Answer: 4.0909

$$\begin{aligned} E[X] &= \sum_{x=1}^{5} x f(x) \\ &= 1 * (1/55) + 2 * (4/55) + 3 * (9/55) + 4 * (16/55) + 5 * (25/55) \\ &= 225/55 = 4.090909 \end{aligned}$$

(g) **(Pts: 2)** What is $E[e^X]$?
Answer: 87.2172

$$\begin{aligned} E[e^X] &= \sum_{x=1} e^x f(x) \\ &= e^1 * (1/55) + e^2 * (2]4/55) + e^3 * (9/55) + e^4 * (16/55) + e^5 * (25/55) \\ &= 87.21716 \end{aligned}$$

(h) **(Pts: 3)** What is $SD[X]$?

Answer: 17.8

First we want $E[X^2]$.

$$\begin{aligned} E[X^2] &= \sum_{x=1}^{5} x^2 f(x) \\ &= 1^2 * (1/55) + 2^2 * (4/55) + 3^2 * (9/55) + 4^2 * (16/55) + 5^2 * (25/55) \\ &= 979/55 = 17.8 \end{aligned}$$

Now

$$SD[X] = \sqrt{Var(X)} = \sqrt{E[X^2] - E[X]^2} = \sqrt{17.8 - 4.090909^2} = 1.031728$$

3. The current Jackpot for the Powerball lottery is 70 million dollars (top line in the table). You can win the same amount of money in more than one way (depending on if you get a match on the special red ball). The table of the prizes and corresponding probabilities are as follows (make the assumption that nobody else would win the Jackpot so you wouldn't have to split the winnings):

Prize ($)	Probability
70,000,000	1/175223510.00
1,000,000	1/5153632.65
10,000	1/648975.96
100	1/19087.53
100	1/12244.83
7	1/360.14
7	1/706.43
4	1/110.81
4	1/55.41
0	c

(a) **(Pts: 2)** What must c be to make this a valid probability distribution?

Answer: $c = .9686002$

```
> x <- c(175223510, 5153632.65, 648975.96, 19087.53, 12244.83, 360.14, 706.43, 110.81, 55.41)
> probs <- 1/x
> sum(probs)
[1] 0.0313998
> # value for c
> 1 - sum(probs)
[1] 0.9686002
> c.val <- 1 - sum(probs)
> probs <- c(probs, c.val)
```

(b) **(Pts: 2)** What is the expected Prize amount of a Powerball ticket? (don't include ticket price)

Answer: 0.7599751

```
> prizes <- c(70000000, 1000000, 10000, 100, 100, 7, 7, 4, 4, 0)
> sum(prizes * probs)
[1] 0.7599751
```

(c) (**Pts:** 2) Note that a Powerball ticket has a cost of \$2. What is the expected payout of a Powerball ticket?

Answer: -1.240025

By the linearity of expectation E[Payout] = E[Prize-2] = E[Prize] - 2 = .7599751 - 2 = -1.240025

(d) (**Pts:** 3) What is the variance and standard deviation of the payout of a Powerball ticket?

Answer: Variance = 28158477, Standard deviation = 5306.456

Note that it doesn't matter if we calculate the variance of the payout or the variance of the prize amount since they only differ by a constant

```
> # Directly
> ex <- sum(prizes * probs)
> (var <- sum((prizes - ex)^2 * probs))
[1] 28158477
> # Other way
> ex2 <- sum(prizes^2 * probs)
> (var <- ex2 - ex^2)
[1] 28158477
> (stddev <- sqrt(var))
[1] 5306.456
```

(e) (**Pts:** 3) If Tim plays the Powerball 2 weeks in a row what is the probability that Tim doesn't lose money?

Answer: 0.06181366

Note that 2 tickets costs 4 dollars so if Tim wins either of the 2 times he plays then he doesn't lose money. The only way for Tim to lose money is to win nothing both weeks in a row. The probability of winning nothing is the value of c computed in part (a). The probability of winning nothing two weeks in a row is $c^2 = 0.9686002^2$ and so the probability that Tim doesn't lose money is $1 - 0.9686002^2 = 0.06181366$.

(f) (**Pts:** 3) What would the Jackpot need to be for the expected value of the payout of a ticket to rise to \$0?

Answer: 287281513

Let x be the value of the jackpot. Then we want 0 = E[payout].

$$\begin{aligned} 0 =&(x-2)*1/175223510.00 + (1,000,000-2)*1/5153632.65 + (10,000-2)*1/648975.96+ \\ &(100-2)*1/19087.53 + (100-2)*1/12244.83 + (7-2)*1/360.14 + (7-2)*1/706.43+ \\ &(4-2)*1/110.81 + (4-2)*1/55.41 + (0-2)*0.9686002 \\ =&x*1/175223510 - 1.639515 \end{aligned}$$

So $x = 1.639515/(1/175223510) = 287281511$ (rounding will cause this number to be a little off for most). But since we solved for the payout of the jackpot we need to add \$2 to get what the jackpot needs to be so the answer is 287281513.

Alternatively one could use a similar method and solve for the value of the jackpot that would give E[Prize] = 2.

4. A statistics instructor wrote an exam and on it they put a question that they believed each student would have a probability of .8 of getting correct. There were 55 exams that were graded and 38 students correctly answered the question.

   (a) (**Pts:** 2) What is the pro

      i. Define random variables to denote the number of print jobs each printer gets and state their distributions.

      ii. What is the expected number of print jobs for Printer1 from 1:00pm-1:25pm?

      iii. What is the probability that Printer2 receives 10 or more print jobs between 1pm and 2pm?

iv. What is the distribution of the number of print jobs requested in that computer lab from 1pm-2pm?

v. What is the probability that there are 0 print jobs requested in the computer lab between 1pm and 2pm?

bability that if the true probability of a student getting the question correct was 80% that 38 students out of the 55 would answer correctly?

Answer: 0.0186

Let $X \sim B(55, .8)$ then we want $f(38) = .0186$

$$f(38) = \binom{55}{38} .8^{38} .2^{55-38}$$

```
> dbinom(38, 55, .8)
[1] 0.01857895
```

(b) **(Pts: 3)** What is the probability that 38 or less students would have answered the question correctly if the true probability of success was .8?

Answer: 0.0367

We want $F_X(38) = \sum_{x=0}^{38} f(x) = .0367$

```
> pbinom(38, 55, .8)
[1] 0.03673059
```

(c) **(Pts: 1)** Using the previous two answers do you think the professor's expectations were too high?

Answer: Yes

The expected number of students that would answer correctly is 44 if the true proportion was 0.8. So we observed less than we would expect; however, this doesn't say anything by itself since we only observe the actual expected value about 13% of the time. However, since we observed less than we expected and the probability of observing what we actually observed or something less than what we observed is only .0367 this tells us that it's very unlikely that we would see a result this low if the true probability of success was 0.8. For these reasons we think that the professor's expectations were too high.

5. In a computer lab there are two different printers. From 1pm-2pm Printer1 receives on average 12 print jobs and Printer2 receives on average 5 print jobs. Assume the number of print jobs requested on Printer1 and on Printer2 are independent of each other.

(a) **(Pts: 1)** Define random variables to denote the number of print jobs each printer gets and state their distributions.

Answer:

Let X_1 be the number of print jobs Printer 1 receives between 1 and 2pm. Then $X_1 \sim Pois(12)$. Let X_2 be the number of print jobs Printer 2 receives between 1 and 2pm. Then $X_2 \sim Pois(5)$.

(b) **(Pts: 2)** What is the expected number of print jobs for Printer1 from 1:00pm-1:25pm?

Answer: 5

Let Y bet the number of print jobs for Printer1 between 1 and 1:25. Then $Y \sim Pois(12 * (25/60))$ so $E[Y] = 12 * (25/60) = 5$.

(c) **(Pts: 3)** What is the probability that Printer2 receives 10 or more print jobs between 1pm and 2pm?

Answer: 0.03182806

We want $P(X_2 \geq 10) = 1 - P(X_2 \leq 9) = 1 - F_{X_2}(9)$. Using R:

```
> 1 - ppois(9, 5)
[1] 0.03182806
```

(d) **(Pts: 1)** What is the distribution of the number of print jobs requested in that computer lab from 1pm-2pm?

Answer: $Pois(17)$

Students need only state the correct distribution - no justification necessary.

Let Z be a random variable associated with the number of print jobs requested in the computer lab from 1pm-2pm. Then $Z = X_1 + X_2$ since those are the only two printers in the lab. Since X_1 and X_2 are independent poisson random variables we know that their sum is a poisson distribution with a expected value equal to the sum of the two corresponding expected values ($12 + 5 = 17$).

(e) **(Pts: 2)** What is the probability that there are 0 print jobs requested in the computer lab between 1pm and 2pm?

Answer: $\exp(-17) \approx .000000041399377$

We are looking for $P(Z = 0) = \frac{e^{-17}17^0}{0!} = e^{-17}$

Alternatively we know that the only way for there to be 0 print jobs is for both printers to have no print jobs. So

$$\begin{aligned}
P(nojobs) &= P(\text{no jobs on printer 1 and no jobs on printer 2}) \\
&= P(\text{no jobs on printer 1})P(\text{no jobs on printer 2}) \\
&= \left(\frac{e^{-12}12^0}{0!}\right)\left(\frac{e^{-5}5^0}{0!}\right) \\
&= e^{-17}
\end{aligned}$$