Question 1

Compute E[X] and Var[X].

Since X and Y denotes the number of ones and twos in n rolls of a standard die respectively, we know that each dice roll is independent and the probability of obtaining a particular value on a roll is $\frac{1}{6}$. Using this fact, we can model X and Y as binomial distributions. $X \sim Bin(n, \frac{1}{6}), Y \sim Bin(n, \frac{1}{6})$.

Using that the expectation of a binomially distributed r.v R is $E[R] = np$, $E[X]$ can be calculated as $E[X] = np = n * \frac{1}{6} = \frac{n}{6}$.

Using that the variance of a binomially distributed r.v R is $Var[R] = np(1-p)$, $Var[X]$ can be calculated as $Var[X] = np(1-p) = n * \frac{1}{6} * \frac{5}{6} = \frac{5n}{36}$.

Are the random variables X and Y independent? Prove or give a counter-example.

If X and Y are independent, then $Pr(X = x, Y = y) = Pr(X = x)Pr(Y = y)$ must be true for each value of x and y.

I will disprove this statement for $x = n, y = n$.

$Pr(X = n, Y = n) = 0$ since if I obtain n heads, I cannot obtain n tails as well. However, $Pr(X = n)Pr(Y = n) = \binom{n}{n}\frac{1}{6}^n * \binom{n}{n}\frac{1}{6}^n = \frac{1}{6}^{2n}$.

Comparing the values of $Pr(X = n, Y = n)$ and $Pr(X = n)Pr(Y = n)$, we see they are not equal. Therefore, X and Y are not independent.

If Z = X + Y , compute E[Z] and Var[Z]

Since each roll is independent and the probability of obtaining a one or a two remains constant, I can model Z by a binomial distribution. Using $Pr(\text{rolling a one or a two}) = \frac{1}{3}$, I can say that $Z \sim Bin(n, \frac{1}{3})$.

Using the fact Z has a binomial distribution, $E[Z] = np = \frac{n}{3}$

Using the formula for variance on a binomial r.v. X, $Var[X] = np(1-p)$, $Var[Z] = n * \frac{1}{3} * \frac{2}{3} = \frac{2n}{9}$.

Use the above results to compute E[(X + Y)^2] and E[X Y]

Using the formula $Var[Z] = E[Z^2] - E[Z]^2$ and knowing that $E[(X+Y)^2] = E[Z^2]$, we can rearrange the formula for $Var[Z]$ to obtain $E[Z^2] = Var[Z] + E[Z]^2$.

Substituting the values of $Var[Z] = \frac{2n}{9}$ and $(E[Z])^2 = \frac{n^2}{9}$ into $E[Z^2] = Var[Z] + E[Z]^2$, we obtain that $E[Z^2] = \frac{2n}{9} + \frac{n^2}{9} = \frac{2n+n^2}{9}$

Expanding out $E[Z^2]$, we obtain $E[Z^2] = E[(X+Y)^2] = E[X^2 + Y^2 + 2XY]$. Rearranging the expression, we obtain $E[XY] = \frac{E[Z^2] - E[X^2] - E[Y^2]}{2}$.

Knowing that X and Y have the same distribution, finding the value of $E[X^2]$ will allow $E[Y^2]$ to be known.

Using the formula $Var[X] = E[X^2] - (E[X])^2$, the value of $E[X^2]$ is $E[X^2] = Var[X] + (E[X])^2$. Using $Var[X] = \frac{5n}{36}, E[X] = \frac{n}{6}$, we can calculate the above expression. $E[X^2] = \frac{5n}{36} + \frac{n^2}{36} = \frac{5n+n^2}{36}$.

Knowing the value of $E[X^2]$ and further aware that $E[X^2] = E[Y^2]$, I can now calculate $E[XY]$. Using the values of $E[X^2] = \frac{n^2+5n}{36}, E[Z^2] = \frac{2n+n^2}{9}$ in the expression $E[XY] = \frac{E[Z^2] - E[X^2] - E[Y^2]}{2}$, I obtain $E[XY] = \frac{\frac{2n+n^2}{9} - 2 * \frac{n^2+5n}{36}}{2} = \frac{\frac{2n+n^2}{9} - \frac{n^2+5n}{18}}{2} = \frac{\frac{4n+2n^2-n^2-5n}{18}}{2} = \frac{n^2-n}{36}$.

Use the above results to compute Cov[X, Y]

Using the formula for covariance, $Cov[X, Y] = E[XY] - E[X]E[Y]$, I have that $Cov[X, Y] = \frac{n^2-n}{36} - \frac{n^2}{36} = \frac{-n}{36}$.

Use the above results to compute Corr[X, Y]. Are X and Y positively or negatively correlated?

Using the formula for $Corr[X,Y] = \frac{Cov[X,Y]}{\sqrt{Var[X]Var[Y]}}$, I obtain that the correlation coefficient is $\frac{\frac{-n}{36}}{\sqrt{(\frac{5n}{36})^2}} = \frac{-n}{36} * \frac{36}{5n} = \frac{-1}{5}$. Since this value is negative, I conclude that X and Y are negatively correlated.

Question 2

Calculate E[X] and Var[X]

We can decompose X as $X = X_1 + X_2 + X_3 + \dots X_n$, where each X_i represents an indicator r.v signifying that the ith turn was a winning turn. Applying linearity of expectation to $E[X]$, we obtain $E[X] = E[X_1] + E[X_2] + \dots E[X_n]$.

$\forall X_i, Pr(X_i = 1) = p^i$. Applying this to $E[X]$ above, we obtain that $E[X] = p + p^2 + \dots p^n$. Recognizing this as a geometric series, we know that $\sum_{i=1}^{n} p^i = p(\frac{1-p^n}{1-p})$. Therefore, $E[X] = p(\frac{1-p^n}{1-p})$.

For calculating $Var[X]$, I can exploit the fact that for an indicator r.v Y, $Var[Y] = p(1-p) = p - p^2$. Furthermore, since all the flips are mutually independent of each other, I can assume that the results of one turn have no influence on the outcome of the other turns. Utilizing this information in calculating $Var[X]$, I have that $Var[X] = \sum_{i=1}^{n} p^i - p^{2i}$. Distributing the sum, I obtain

$Var[X] = \sum_{i=1}^{n} p^i - \sum_{i=1}^{n} p^{2i}$. Using the formula of a finite geometric sum with $a = p$ and $r = p$ for the first term and $a = p^2$ and $r = p^2$ for the second term, the variance simplifies to $Var[X] = p(\frac{1-p^n}{1-p}) - p^2\frac{1-p^{2n}}{1-p^2}$.

Calculate E[Y] and Var[Y]

I can compose Y as a collection of Y_i where each Y_i is an indicator variable r.v signifying if the ith turn is a winning turn. Applying linearity of expectation, $E[Y] = E[Y_1 + Y_2 + Y_3 + \dots Y_n] = E[Y_1] + E[Y_2] + \dots E[Y_n]$.

If I need at least one heads in order to have a winning turn, then on turn j I would have probability $1-(1-p)^j$ of obtaining at least one heads. Since this happens for each turn, the expression $E[Y]$ can be modified as $E[Y] = \sum_{j=1}^{n} 1-(1-p)^j$. Breaking apart the summation, I obtain $E[Y] = \sum_{j=1}^{n} 1 - \sum_{j=1}^{n}(1-p)^j$, which simplifies to $E[Y] = n - \sum_{j=1}^{n}(1-p)^j$. Recognizing the second term is a geometric sum, I can use the formula for a finite geometric sum ($S_n = a\frac{1-r^n}{1-r}$) to reduce the expression above into
$E[Y] = n - (1-p)\frac{1-(1-p)^n}{1-(1-p)} = n - (1-p)\frac{1-(1-p)^n}{p}$.

The expression above simplifies to $E[Y] = n - (1-p)\frac{1-(1-p)^n}{p}$.

Var[Y]

Using the fact that each of turn result is independent of the other turns, I can calculate $Var[Y]$ as $Var[Y] = \sum_{i=1}^{n} Var[Y_i]$.

Knowing that $Var[T] = p(1-p)$ for an indicator r.v T, I can apply it with Y to obtain $Var[Y] = \sum_{i=1}^{n} p_i(1-p_i)$. Substituting the value of p_i with $1-(1-p)^i$, the expression becomes
$Var[Y] = \sum_{i=1}^{n}(1-(1-p)^i)(1-(1-(1-p)^i)) = \sum_{i=1}^{n}(1-(1-p)^i)(1-p)^i = \sum_{i=1}^{n}(1-p)^i - (1-p)^{2i}$. Distributing the summation, I obtain $\sum_{i=1}^{n}(1-p)^i - \sum_{i=1}^{n}(1-p)^{2i}$. Applying the formula for a finite geometric sum on both summations, the variance reduces to
$Var[Y] = (1-p)\frac{1-(1-p)^n}{1-(1-p)} - (1-p)^2\frac{1-(1-p)^{2n}}{1-(1-p)^2} = (1-p)\frac{1-(1-p)^n}{p} - (1-p)^2\frac{1-(1-p)^{2n}}{2p-p^2}$.

The variance of Y is $(1-p)\frac{1-(1-p)^n}{p} - (1-p)^2\frac{1-(1-p)^{2n}}{2p-p^2}$.

Question three

Part one

Using $Y = (R - E[R] + a)^2$, I can apply the Markov bound since Y is a non-negative r.v. Doing this, I obtain $Pr(Y \geq (x+a)^2) \leq \frac{E[Y]}{(x+a)^2}$.

I need to calculate $E[Y]$, which I can do from expanding $Y = (R - E[R] + a)^2$. Expanding out Y, I obtain $R^2 - RE[R] + Ra - E[R]R + (E[R])^2 - E[R]a + aR - aE[R] + a^2 = R^2 - 2RE[R] + 2Ra + (E[R])^2 - 2E[R]a + a^2$. Applying linearity of expectation on Y, this results in $E[Y] = E[R^2] - E[2RE[R]] + E[2Ra] + E[(E[R])^2] - E[2E[R]a] + E[a^2]$. Knowing that a and $E[R]$ are constants, the above expression simplifies to $E[Y] = E[R^2] - 2E[R]E[R] + 2aE[R] + (E[R])^2 - 2E[R]a + a^2 = E[R^2] - 2(E[R])^2 + 2aE[R] + (E[R])^2 - 2E[R]a + a^2 = E[R^2] - (E[R])^2 + a^2$.

Using the fact that $Var[R] = E[R^2] - (E[R])^2$, the above expression can be expressed as $Var[R] + a^2$. Since the expression was simplified from calculating $E[Y]$, we can say that $E[Y] = Var[R] + a^2$.

Now that we have the value of $E[Y] = Var[R] + a^2$, we can plug it in $Pr(Y \geq (x+a)^2) \leq \frac{E[Y]}{(x+a)^2}$, resulting in $Pr(Y \geq (x+a)^2) \leq \frac{Var[R]+a^2}{(x+a)^2}$.

We are able to manipulate $R - E[R] \geq x$ to obtain $Y \geq (x+a)^2$ by adding x to both sides of the inequality and squaring each side to obtain $(R - E[R] + a)^2 \geq (x+a)^2$, which is equivalent to $Y \geq (x+a)^2$. Since we were able to obtain one expression from manipulating the other, we can say that $Pr(R - E[R] \geq x) = Pr(Y \geq (x+a)^2)$.

Given that $Pr(Y \geq (x+a)^2) \leq \frac{Var[R]+a^2}{(x+a)^2}$, we know that $Pr(R - E[R] \geq x) \leq \frac{Var[R]+a^2}{(x+a)^2}$ since $Pr(R - E[R] \geq x) = Pr(Y \geq (x+a)^2)$.

Part two

Minimizing $\frac{a^2+Var[R]}{(x+a)^2}$, we obtain $\frac{d}{da}\frac{a^2+Var[R]}{(x+a)^2} = 0$. Applying the quotient rule, we obtain $\frac{d}{da}\frac{a^2+Var[R]}{(x+a)^2} = \frac{(a+x)^2 * 2a - [a^2+Var[R]] * 2[a+x]}{(a+x)^4} = \frac{2(a+x)[(a+x)*a - (a^2+Var[R])]}{(a+x)(a+x)^3}$.

Simplifying the equation, we obtain $\frac{2((a+x)*a-(a^2+Var[R]))}{(a+x)^3)} = \frac{2(a^2+ax-a^2-Var[R])}{(a+x)^3} = \frac{2(ax-Var[R])}{(a+x)^3}$. Knowing that we must minimize a, we must solve $\frac{2(ax-Var[R])}{(a+x)^3} = 0$. Solving for a, we obtain $a = \frac{Var[R]}{x}$.

Letting $v = Var[R]$ and using $a = \frac{Var[R]}{x} = \frac{v}{x}$ into the expression $\frac{a^2+Var[R]}{(x+a)^2}$, it becomes $\frac{(\frac{v}{x})^2+v}{(x+\frac{v}{x})^2} = \frac{\frac{v^2+x^2v}{x^2}}{(\frac{x^2+v}{x})^2} = \frac{v^2+x^2v}{x^2} * \frac{x^2}{(x^2+v)^2} = \frac{v(v+x^2)}{(x^2+v)^2} = \frac{v}{x^2+v}$.

Remembering that $v = Var[R]$, we know that $\frac{v}{x^2+v} = \frac{Var[R]}{x^2+Var[R]}$, which is the one-sided Chebyshev bound.

Question four

Part 1

Let N be an r.v measuring the number of marks that a student gets on their final examination. We are given that $E[N] = 75$.

Using this information, we can apply the Markov abound to calculate $Pr(N \geq 85)$, obtaining $Pr(N \geq 85) \leq \frac{E[N]}{85}$. This simplifies to $Pr(N \geq 85) \leq \frac{75}{85} = Pr(N \geq 85) \leq \frac{15}{17}$.

Part 2

Knowing that no student gets a score below forty marks, we can modify the expression $Pr(N \geq 85)$ to $Pr(N - 40 \geq 45)$ to account for this information.

If we define $Y = N - 40$, we can equivalently state $Pr(N - 40 \geq 45)$ as $Pr(Y \geq 45)$. Since Y is non-negative, we can apply the Markov bound to calculate $Pr(Y \geq 45) \leq \frac{E[Y]}{45} = Pr(Y \geq 45) \leq \frac{E[N]-40}{45} = Pr(Y \geq 45) \leq \frac{35}{45} = Pr(Y \geq 45) \leq \frac{7}{9}$.

Part 3

We now know that $Var[N] = 25$.

We can use Chebyshev's inequality in the expression $Pr(|N - 75| \geq 11)$ to upper bound it by $\frac{Var[N]}{11^2}$. Using the value $Var[N] = 25$, we have $Pr(|N - 75| \geq 11) \leq \frac{25}{121}$.

Part 4

Using the one-sided Chebyshev's inequality, we find that $Pr(N - 75 \geq 11) \leq \frac{Var[N]}{11^2 + Var[N]} = Pr(N - 75 \geq 11) \leq \frac{25}{146}$.

Part 5

Let $Z_i = \frac{S_i - 40}{60}$.

Since we know that $S_i \in [40, 100]$, this means that $S_i - 40 \in [0, 60]$. Using this information, we know the smallest and largest value Z_i can take on is $\frac{0}{60}$ and $\frac{60}{60} = 1$. Therefore, we have that $Z_i \in [0, 1]$.

Calculating $E[Z_i]$, we obtain $E[\frac{S_i}{60} - \frac{40}{60}] = E[\frac{S_i}{60}] - \frac{2}{3} = \frac{75}{60} - \frac{2}{3} = \frac{7}{12}$.

Calculating $Var[Z_i]$, we obtain $Var[\frac{S_i}{60} - \frac{40}{60}] = Var[\frac{S_i}{60}] = \frac{Var[S_i]}{3600} = \frac{25}{3600} = \frac{1}{144}$.

Part 6

If we convert 85 into a score seen in Z, we have $\frac{85-40}{60} = \frac{45}{60} = \frac{3}{4}$. If we want the average of Z to be greater or equal than 85 for the bad case, we know that this is equivalent to $\frac{Z}{n} \geq \frac{3}{4}$. Multiplying by n, we obtain $Z \geq \frac{3n}{4}$.

If we start off with the expression $Pr(Z \geq \frac{3n}{4})$, we can subtract $E[Z]$ on both sides to obtain $Pr(Z - E[Z] \geq \frac{3n}{4} - E[Z])$, which has the same probability as the first expression. Knowing that the first expression represents the probability of the bad case occurring, that the average is greater than or equal to 85, we can lower bound it by the probability of the bad event occurring. Applying the one-sided Chebyshev's inequality, we obtain $Pr(Z - E[Z] \geq \frac{3n}{4} - E[Z]) \leq \frac{Var[Z]}{Var[Z] + (\frac{3n}{4} - E[Z])^2}$.

Substituting the value of $Var[Z] = \frac{n}{144}, E[Z] = \frac{7n}{12}$ in the expression above, we obtain
$Pr(Z - E[Z] \geq \frac{3n}{4} - E[Z]) \leq \frac{\frac{n}{144}}{\frac{n}{144} + (\frac{3n}{4} - \frac{7n}{12})^2} = \frac{\frac{n}{144}}{\frac{n}{144} + (\frac{n}{6})^2}$.

Additionally knowing that the bad case has a probability smaller than 0.001 of occurring, we can upper bound the expression $\frac{\frac{n}{144}}{\frac{n}{144}+(\frac{n}{6})^2}$ by 0.001. We obtain the inequality $\frac{\frac{n}{144}}{\frac{n}{144}+\frac{n^2}{36}} < 0.001 = \frac{\frac{1}{144}}{\frac{1}{144}+\frac{n}{36}} < 0.001$. Multiplying by the denominator, we obtain $\frac{1}{144} < 0.001(\frac{1}{144} + \frac{n}{36})$. We can further simplify the expression to obtain $\frac{1000}{144} < \frac{1}{144} + \frac{n}{36}$, which further reduces to $\frac{999}{144} * 36 < n = 249.75 < n$.

Since we can only have whole students, we can modify $n > 249.75$ to equivalently be $n \geq 250$. This means that we need at least 250 students to take the examination.

Part 7

To use the Chernoff bound, we need an expression of the form $Pr(T \geq cE[T])$, for an r.v T and $c \in \mathbb{R}$. We want to calculate $Pr(Z \geq \frac{3n}{4})$, but we need to obtain $\frac{3n}{4}$ from $E[Z] * c = \frac{7n}{12}c$. Noticing that $\frac{7n}{12} * \frac{9}{7} = \frac{3n}{4}$, we can let $c = \frac{9}{7}$. Doing this, we can calculate $Pr(Z \geq \frac{3n}{4} * \frac{9}{7}) = Pr(Z \geq \frac{3n}{4})$.

Applying Chernoff's bound, we obtain $Pr(Z \geq \frac{3n}{4}) \leq e^{-\beta(\frac{9}{7}) * \frac{7n}{12}}$. Knowing that the upper limit obtained from Chernoff's bound is bounded by 0.001, we have that $e^{-\beta(\frac{9}{7}) * \frac{7n}{12}} < 0.001$. Taking the natural logarithm of both sides, we obtain $-\beta(\frac{9}{7}) * \frac{7n}{12} < ln(0.001) = \beta(\frac{9}{7}) * n > -\frac{ln(0.001)*12}{7}$.

Using $\beta(c) = cln(c) - c + 1$ with $c = \frac{9}{7}$, I obtain $\beta(\frac{9}{7}) = \frac{9}{7} * ln(\frac{9}{7}) - \frac{9}{7} + 1 \approx 0.037$. Using the value of $\beta(\frac{9}{7})$ in $\beta(\frac{9}{7}) * n > -\frac{ln(0.001)*12}{7}$, I obtain $n \geq 317$.

Using the information above, we know that we need at least 317 people to take the examination.

Question five

Part 1

For D_1, I must consider the possibility of each file being written to the first disk. I can therefore rewrite D_1 as $D_1 = \sum_{i=1}^{1000} WF_i$, where WF_i is an indicator r.v which is 0 if F_i was not written to disk one or $1 * F_i$ if F_i was written to disk one.

Applying linearity of expectation to $E[D_1]$, I obtain $E[X] = E[WF_1] + E[WF_2] + ... E[WF_1000] = \sum_{i=1}^{1000} WF_i * Pr(WF_i = 1) = \sum_{i=1}^{1000} F_i * Pr(WF_i = 1)$. Considering that the probability of any file being written to the first disk is always $\frac{1}{4}$, I can rewrite the value for $E[D_1]$ to equivalently be $E[D_1] = \frac{1}{4}\sum_{i=1}^{1000} F_i$.

Using the fact that $\sum_{i=1}^{1000} F_i = 400$, I can simplify the summation above to $E[D_1] = \frac{1}{4} * 400 = 100$.

The expected amount of data written to disk one is 100 MB.

Part 2

Using Markov's inequality, $Pr(D_1 \geq 200) \leq \frac{E[D_1]}{200} = Pr(D_1 \geq 200) \leq \frac{100}{200} = Pr(D_1 \geq 200) \leq \frac{1}{2}$

Part 3

Since I must use the Chernoff bound, I must have an expression of the form $Pr(D_1 \geq c * E[D_1]), c \in \mathbb{R}$. Since I wish to calculate the probability that $D_1 \geq 200$ and it is known that $E[D_1] = 100$, I can set $c = 2$ to obtain $Pr(D_1 \geq 2 * E[D_1]) = Pr(D_1 \geq 200)$.

Applying the Chernoff bound $Pr(D_1 \geq c * E[D_1]) \leq e^{-\beta(c)E[T]}$ with $c = 2$, $E[D_1] = 100$, I obtain that $Pr(D_1 \geq 200) \leq e^{-\beta(2)100}$. Using the formula $\beta(c) = cln(c) - c + 1$ with $c = 2$, I obtain $\beta(2) = 2 * ln(2) - 2 + 1 = 0.39$. Using the value of $\beta(2) = 0.39$ in $Pr(D_1 \geq 200) \leq e^{-\beta(2)100}$, I obtain $Pr(D_1 \geq 200) \leq e^{-38.6} = Pr(D_1 \geq 200) \leq 1.67^{-17}$.

Part 4

Let $D = \{D_1, D_2, D_3, D_4\}$.

Using the union bound, I can upper bound $Pr(\bigcup_{D_i \in D} D_i \geq 200)$ by $\sum\limits_{D_i \in D} Pr(D_i \geq 200)$.

Knowing that $E[D_1] = E[D_2] = E[D_3] = E[D_4]$, we can say that $Pr(D_i \geq 200) \leq 1.67^{-17}$. Using this information, we can calculate $\sum\limits_{D_i \in D} Pr(D_i \geq 200)$ as $4 * 1.67^{-17}$.

Using $\sum\limits_{D_i \in D} Pr(D_i \geq 200)$ as $4 * 1.67^{-17}$ in $Pr(\bigcup_{D_i \in D} D_i \geq 200) \leq \sum\limits_{D_i \in D} Pr(D_i \geq 200)$, I obtain $Pr(\bigcup_{D_i \in D} D_i \geq 200) \leq 6.69^{-17}$.

The probability that some disk has 200 MB or more than 200 MB written to it is upper bounded by 6.69^{-17}.

Question six

Part 1

To verify that the PDF is valid, we must confirm that $\sum\limits_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} = 1$.

Starting with $\sum\limits_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!}$, we can factor out $e^{-\lambda}$ to obtain $e^{-\lambda} \sum\limits_{k=0}^{\infty} \frac{\lambda^k}{k!}$. Using the Taylor Series expansion of $e^x = \sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$ in the previous expression, it simplifies $e^{-\lambda} \sum\limits_{k=0}^{\infty} \frac{\lambda^k}{k!}$ to $e^{-\lambda}e^{\lambda} = 1$.

Since we have shown that $\sum\limits_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} = 1$, we can conclude that the PDF for the Poisson distribution is valid.

Part 2

Prove that E[X] = λ:

Calculating the expectation of the Poisson distribution, we find it is $\sum\limits_{k=0}^{\infty} k * \frac{e^{-\lambda}\lambda^k}{k!}$. Factoring out $e^{-\lambda}$ and removing the first term of the summation due to multiplying by 0, the summation changes to $e^{-\lambda} \sum\limits_{k=1}^{\infty} k * \frac{\lambda^k}{k!}$.

If we factor out a λ term from the summation and simplify $\frac{n}{n!}$ to $\frac{1}{(n-1)!}$, the summation further reduces to $e^{-\lambda}\lambda \sum\limits_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}$. In order to manipulate the summation further, let us define $y = k - 1$. Substituting y into the summation, the summation changes to $e^{-\lambda}\lambda \sum\limits_{y=0}^{\infty} \frac{\lambda^y}{(y)!}$.

Applying the Taylor Series expansion of e^x, the summation now becomes $e^{-\lambda}\lambda e^{\lambda} = \lambda$.

Since we obtained λ from calculating $E[X]$, we can state that $E[X] = \lambda$.

Prove that $Var[X] = \lambda$:

Using the formula $Var[X] = E[X^2] - (E[X])^2$, I can calculate the variance if I know the value of $E[X^2]$ and $E[X]$. I know $E[X]$, but I must obtain $E[X^2]$.

Using the formula $E[X^2] = \sum_{k=0}^{\infty} k^2 * \frac{e^{-\lambda}\lambda^k}{k!}$, we can factor out $e^{-\lambda}$ and a λ term, obtaining $e^{-\lambda}\lambda \sum_{k=0}^{\infty} k * \frac{\lambda^{k-1}}{(k-1)!}$. Let us now define $t = k-1$. Using t in the summation, we now have $e^{-\lambda}\lambda \sum_{k=0}^{\infty}(t+1) * \frac{\lambda^t}{t!} = e^{-\lambda}\lambda(\sum_{k=0}^{\infty} t * \frac{\lambda^t}{t!} + \sum_{k=0}^{\infty} \frac{\lambda^t}{t!})$.

Using our previous work, we know that $e^{-\lambda}\lambda(\sum_{k=0}^{\infty} t * \frac{\lambda^t}{t!} + \sum_{k=0}^{\infty} \frac{\lambda^t}{t!})$ simplifies to $e^{-\lambda}\lambda(\lambda e^{\lambda} + e^{\lambda}) = \lambda^2 + \lambda$.

Now that we know that $E[X^2] = \lambda^2 + \lambda$, I can calculate the variance.

Using $Var[X] = E[X^2] - (E[X])^2$ with $E[X^2] = \lambda^2 + \lambda$, $E[X] = \lambda$, I obtain $Var[X] = \lambda^2 + \lambda - \lambda^2 = \lambda$.

Part 3

Using the formula for a MGF, we have that $\phi(t) = \sum_{x \in Range(X)} e^{tx} pr(X = x)$. Knowing that X has a Poisson distribution, we can change $\phi(t)$ to $\phi(t) = \sum_{x \in Range(X)} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!}$. We can factor out $e^{-\lambda}$ to obtain $\phi(t) = e^{-\lambda} \sum_{x \in Range(X)} \frac{(e^t\lambda)^x}{x!}$. Using the Taylor Series expansion of $\sum_{x=0}^{\infty} \frac{(e^t\lambda)^x}{x!} = e^{e^t\lambda}$ in the expression above, I can simplify $\phi(t)$ to $\phi(t) = e^{-\lambda}e^{e^t\lambda} = e^{e^t\lambda - \lambda} = e^{\lambda(e^t-1)}$.

Therefore, if $X \sim Poisson(\lambda)$, this means that $\phi(t) = e^{\lambda(e^t-1)}$.

Part 4

Prove E[X] = λ.

Knowing that $E[X] = \phi'(0)$, I must first find ϕ'.

$\phi' = e^{\lambda(e^t-1)} * \lambda e^t$. Evaluating $\phi'(0)$, I obtain $\phi'(0) = e^{\lambda(e^0-1)} * \lambda e^0 = \lambda$.

Using the derivative of the MGF at the first moment, we have found that $E[X] = \lambda$.

Prove Var[X] = λ.

Using the formula $Var[X] = E[X^2] - (E[X])^2$, I can calculate this value by finding $E[X^2] = \phi''(0)$.

$\phi'' = \frac{d}{dt} e^{\lambda(e^t-1)} * \lambda e^t = \lambda e^{\lambda(e^t-1)+t} * (\lambda e^t + 1)$. Evaluating $\phi''(0)$, I obtain $\lambda e^{\lambda(e^0-1)+0} * (\lambda e^0 + 1) = \lambda(\lambda + 1) = \lambda^2 + \lambda$.

Using the values of $E[X] = \lambda$, $E[X^2] = \lambda^2 + \lambda$, I can evaluate $Var[X] = E[X^2] - (E[X])^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

We have shown that $Var[X] = \lambda$ by using the first and second moment of the MGF.