Student Information

Full Name : Yarkın Özcan
Id Number : 2580835

Answer 1

a)

Since X is a discrete random variable, by summing probabilities for all possible X values, we should get 1.

$$\sum_{k=1}^{5} P(X=k) = P(X=1) + P(X=2) + P(X=3) + P(X=4) + P(X=5) = 1$$

$$\frac{N}{1} + \frac{N}{2} + \frac{N}{3} + \frac{N}{4} + \frac{N}{5} = 1$$

$$N * \frac{(60+30+20+15+12)}{60} = 1$$

$$N = \frac{60}{137} = 0.438$$

b)

We can calculate the expected value of X by using the formula;

$$E(X) = \sum_{x=1}^{5} x * P(X=x) = 1 * \frac{N}{1} + 2 * \frac{N}{2} + 3 * \frac{N}{3} + 4 * \frac{N}{4} + 5 * \frac{N}{5}$$

$$5 * N = 5 * \frac{60}{137} = \frac{300}{137} = 2.190$$

c)

We can calculate the variance of X by using the formula;

$$Var(X) = E[(X-\mu)^2] = E(X^2) - \mu^2$$

From the previous question we found that μ is equal to $\frac{300}{137}$. We only need to find $E(X^2)$.

$$E(X^2) = \sum_{x=1}^{5} x^2 * P(X=x) = 1 * \frac{N}{1} + 2^2 * \frac{N}{2} + 3^2 * \frac{N}{3} + 4^2 * \frac{N}{4} + 5^2 * \frac{N}{5}$$

$$E(X^2) = N * \frac{5*6}{2} = \frac{900}{137}$$

$$Var(X) = \frac{900}{137} - (\frac{300}{137})^2 = \frac{33300}{18769} \approx 1.774$$

d)

Since both X and Y are discrete random variables, we can calculate $E(XY)$, $E(X)$ and $E(Y)$ using these formulas;

$$E(XY) = \sum_{y=1}^{5} \sum_{x=1}^{5} x * y * P(X = x, Y = y)$$

$$P(X, Y) = P(X) * P(Y) = \frac{N}{x} * \frac{y}{15}$$

$$E(XY) = \sum_{y=1}^{5} \sum_{x=1}^{5} N * \frac{y^2}{15}$$

$$E(XY) = 5 * \frac{N}{15} * \sum_{y=1}^{5} y^2 = \frac{N}{3} * (1 + 4 + 9 + 16 + 25) = \frac{20}{137} * 55 = \frac{1100}{137}$$

$$E(Y) = \sum_{y=1}^{5} y * P(Y = y) = \sum_{y=1}^{5} y * \frac{y}{15} = \frac{55}{15} = \frac{11}{3}$$

$$E(X) * E(Y) = \frac{300}{137} * \frac{11}{3} = \frac{1100}{137}$$

$$Cov(X, Y) = E(XY) - E(X) * E(Y) = 0$$

Covariance of 2 random variables is the measure of strength/relationship between those variables. So Cov(X, Y) being 0 means that X and Y are two independent variables. That is there is no correlation between them.

Answer 2

a)

Let X be a random variable representing the number of successes in 1000 trials and let p be the probability of success in a single attempt. We need to find P$(X \geq 1)$ which is equal to 1 - P(X=0).

1 - P(X=0) = 0.95

P(X=0) = 0.05

(1-p)^{1000} = 0.05

p= 0.003

b)

i-)

If we need more than 500 games in order to win twice, then it means that we have at most 1 win in the first 500 games. So, we need to find P(X≤1) = $F_x(1)$ where X is the number of wins, n is the number of matches, p is the probability of winning against a IM and F_x is the Binomial CDF. In this case, n = 500, p = 0.003 and we can use Octave by using the keyword binocdf(1, 500, 0.003) in order to find our answer. Eventually we get the answer $F_x(1) = 0.558$.

ii-)

Again if we have to play more than 10000 matches in order to get 2 wins, it means that in the first 10000 matches we have at most 1 win, so we have to find P(X≤1) = $F_x(1)$ where X is the number of wins, n is the number of matches, p is the probability of winning against a GM and F_x is the Binomial CDF. In this case, n = 10000, p = 0.0001 and we can use Octave by using the keyword binocdf(1, 10000, 0.0001) in order to find our answer. Eventually we get the answer $F_x(1) = 0.736$.

c)

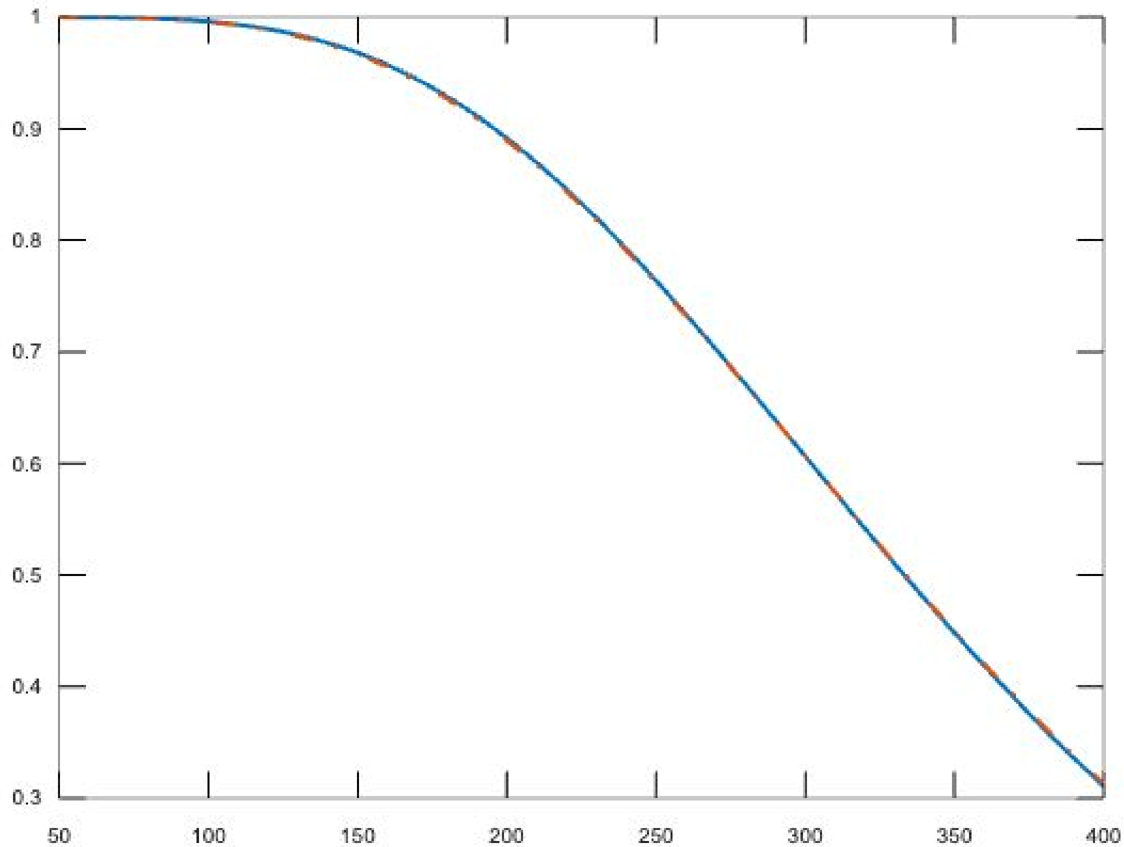
If we assume feeling sick in a day as a success, the probability of success p = 1 - 0.98 = 0.02. There are 366 days in the year, so number of trials n = 366. If we want to feel healthy at least 360 days in a year, it means that we want to feel sick at most 6 days in a year. So we need to find P($X \leq 6$) where X is the number of days that we feel sick. In our case, since n = 366 $\geq$ 30 and $p = 0.02 \leq 0.05$, we can use Poisson Approximation of Binomial Distribution. In order to use Poisson Distribution, firstly we need to find λ (the average number of events). $\lambda = E(X)$ = n*p = 366 * 0.02 = 7.32. So $\lambda = 7.32$ and $X \leq 6$. Using the table in the book, the closest λ value to ours is 7.50. So if we check the value of Poisson Distribution CDF from the table when λ is 7.50 and $X \leq 6$, it is equal to 0.378.

Answer 3

a)

If we assume feeling sick in a day as a success, the probability of success p = 1 - 0.98 = 0.02. There are 366 days in the year, so number of trials n = 366. If we want to feel healthy at least 360 days in a year, it means that we want to feel sick at most 6 days in a year. So we need to find P($X \leq 6$) where X is the number of days that we feel sick. Here we can use Binomial CDF in order to find P($X \leq 6$) = $F_x(6)$. In Octave Online, we can write binocdf(6, 366, 0.02). This gives the answer 0.4013. We can see that the result we found is actually higher than the result we found using Poisson Approximation of Binomial Distribution with $\lambda = 7.50$. This is because the actual λ = n*p was 7.32. Since the λ we used is higher than the original frequency, the probability of X being lower than 6 became harder.

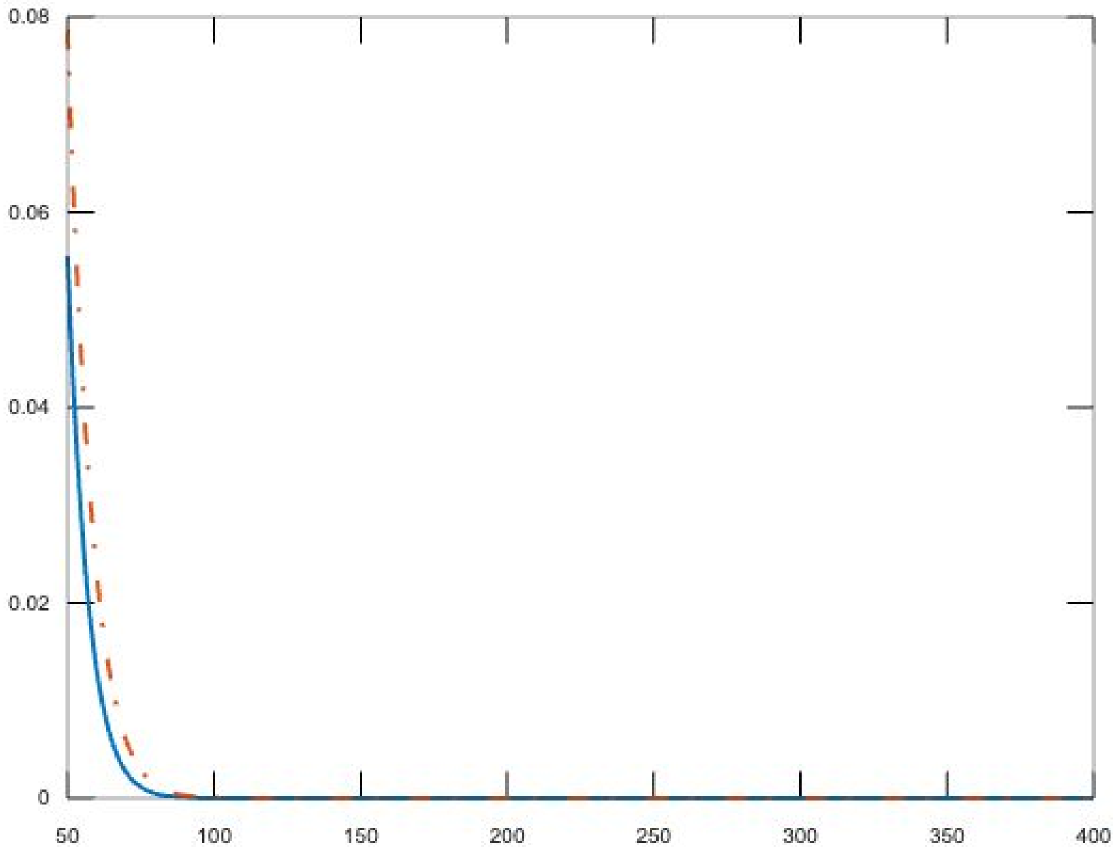
b)



Blue line represents Binomial CDF and Red line represents Poisson CDF. Here is the Octave Online code to create this graph;

```
octave:1> p = 0.02;
ns = 50:400;
binomial_probabilities = (binocdf(6,ns,p));
poisson_probabilities = (poisscdf(6,ns*p));
close all;
plot(ns, binomial_probabilities, 'linewidth', 2);
hold on;
plot(ns, poisson_probabilities, '-.', 'linewidth', 2);
saveas(1,"p=0.98.png");
```

c)



Blue line represents Binomial CDF and Red line represents Poisson CDF. Here is the Octave Online code to create this graph;

```
octave:1> p = 0.22;
ns = 50:400;
binomial_probabilities = (binocdf(6,ns,p));
poisson_probabilities = (poisscdf(6,ns*p));
close all;
plot(ns, binomial_probabilities, 'linewidth', 2);
hold on;
plot(ns, poisson_probabilities, '-.', 'linewidth', 2);
saveas(1,"p=0.78.png");
```

We can see that as p got smaller, the difference between Binomial CDF and Poisson CDF became larger. This is because Poisson distribution's approximation to the Binomial distribution is more accurate when the number of trials n is larger and the success rate p (in our case it was initially 0.02 then it became 0.22) is smaller.