Student Information

Full Name : Deniz Karakoyun
Id Number : 2580678

Answer 1

a)

Probability Mass Function Normalization:

Probability mass functions must satisfy the condition that the sum of probabilities over all possible outcomes is equal to 1. In this case, we have:

$$\sum_{x=1}^{5} P(x) = N\left(\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}\right) = N\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}\right)$$

$$= N \times \left(\frac{60+30+20+15+12}{60}\right) = N \times \left(\frac{137}{60}\right)$$

For this to be equal to 1, we must have:

$$N \times \left(\frac{137}{60}\right) = 1$$

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

b)

Expected Value:

The expected value (mean) of a discrete random variable X is given by:

$$E(X) = \sum_{x} x \cdot P(x)$$

$$E(X) = 1 \cdot \frac{60}{137} + 2 \cdot \frac{30}{137} + 3 \cdot \frac{20}{137} + 4 \cdot \frac{15}{137} + 5 \cdot \frac{12}{137}$$

$$E(X) = \frac{60}{137} + \frac{60}{137} + \frac{60}{137} + \frac{60}{137} + \frac{60}{137}$$

$$E(X) = \frac{300}{137} \approx 2.910$$

c)

Variance:

The variance of a discrete random variable X is given by:

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

$$E(X^2) = \sum_x x^2 \cdot P(x)$$

$$E(X^2) = 1^2 \cdot \frac{60}{137} + 2^2 \cdot \frac{30}{137} + 3^2 \cdot \frac{20}{137} + 4^2 \cdot \frac{15}{137} + 5^2 \cdot \frac{12}{137}$$

$$E(X^2) = \frac{60}{137} + \frac{120}{137} + \frac{180}{137} + \frac{240}{137} + \frac{300}{137}$$

$$E(X^2) = \frac{900}{137}$$

$$Var(X) = \frac{900}{137} - \left(\frac{300}{137}\right)^2$$

$$Var(X) = \frac{900}{137} - \frac{90000}{18769}$$

$$Var(X) = \frac{33300}{18769} \approx 1.774$$

d)

Covariance:

Given that Y is a random variable with $P(y) = \frac{y}{15}$ and $y \in \{1, 2, 3, 4, 5\}$, and the joint distribution $P(x, y) = P(x)P(y)$, we need to calculate the covariance $Cov(X, Y) = E(XY) - E(X)E(Y)$.

First, let's calculate $E(Y)$:

$$E(Y) = \sum_{y=1}^{5} y \cdot P(y) = \sum_{y=1}^{5} y \cdot \frac{y}{15} = \frac{1}{15} \sum_{y=1}^{5} y^2$$

$$E(Y) = \frac{1}{15}\left(1^2 + 2^2 + 3^2 + 4^2 + 5^2\right) = \frac{1}{15} \cdot 55 = \frac{11}{3} \approx 3.667$$

Next, let's calculate $E(XY)$:

$$E(XY) = \sum_{x=1}^{5} \sum_{y=1}^{5} x \cdot y \cdot P(x)P(y)$$

Since $P(x)$ and $P(y)$ are independent we can say that X and ,Y are independent so we could also use :

$$E(XY) = E(X) \cdot E(Y) = \frac{3300}{411} \approx 8.030$$

So we can expect that the result will be≈ 8.030

$$E(XY) = \sum_{x=1}^{5} \sum_{y=1}^{5} x \cdot y \cdot \left(\frac{60}{137 \cdot x} \right) \cdot \left(\frac{y}{15} \right)$$

$$= \frac{4}{137} \cdot \sum_{x=1}^{5} \sum_{y=1}^{5} y^2$$

Using the values from previous calculations:

$$E(XY) = \frac{4}{137} \cdot 55 \cdot 5$$

$$= \frac{1100}{137} \approx 8.030$$

Now, let's calculate $Cov(X, Y)$:

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

$$= \frac{1100}{137} - \frac{300}{137} \cdot \frac{11}{3}$$

$$= \frac{1100}{137} - \frac{300}{137} \cdot \frac{11}{3}$$

$$= \frac{3300 - 3300}{137 \cdot 3}$$

$$= 0$$

Interpretation: The covariance $Cov(X, Y)$ indicates the degree to which X and Y vary together. In this case, As we expected we found $Cov(X, Y)$= 0 from this result we can infer that they are independent and this is what we were expecting to see.

Answer 2

a)

Probability of Success in a Single Attempt:

Let the probability of success in a single attempt be (p). The probability of failure in a single attempt is $(1 - p)$..Moreover, If probability of at least one success in 1000 trial is %95 then 1000 failure have probability of %5 because "at least one success" and "no success" are exhaustive events. The probability of at least one success in 1000 trials is equal to 1 minus the probability of failure in all 1000 trials. Since there are only 2 possible outcomes, success and failure, and each event is independent this problem can be modeled using a binomial distribution. Thus, we can express the equation as $P(x \geq 1) = 0.95 = 1 - F(0)$. Furthermore, $F(0)$ is equivalent to $P(0)$, so our equation becomes $P(x = 0) = 0.05$. Now, by substituting variables into the probability mass function (pmf) of the binomial distribution formula, we can solve this question.

$$P(X = x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}$$

$$1 - (1-p)^{1000} \cdot p^0 = 0.95$$

Solving for p:

$$P(X = 0) = (1-p)^{1000} = 0.05$$

$$1 - p = 0.05^{1/1000}$$

$$p = 1 - 0.05^{1/1000}$$

$$p \approx 1 - 0.9975$$

$$p \approx 0.003$$

b)

Winning against IM and GM:

i) Winning against an IM:

For winning against an IM, the probability is given as $p = 3 \times 10^{-3}$.

To find the likelihood of having to play more than 500 games to win twice against an IM, we can use the Binomial CDF: X = more than 500 trials needed to get 2 wins — and we can — set Y = there are less than 2 wins in 500 trials

Although it will be hard to calculate this in binomial, probability $P(X > 500)$ can be related to a binomial variable.

$$P(X > 500) = P(\text{more than 500 trials needed to get 2 wins}) = P(\text{there are less than 2 wins in 500 trials})$$

$= P(Y < 2)$ where Y is the number of wins in 500 matches, which is a binomial variable with parameters $n = 500$ and $p = 0.003$ which results in $F(1) \approx 0.558$

Therefore the answer is 0.558 and it can be easily calculated with following Octave code: binocdf(1,500,0.003)

ii) Winning against a GM:

This problem resembles the previous one, again involving a negative binomial distribution. However, to solve it, we can convert it into a binomial distribution problem. We seek $P(X >$

$10000) = 1 - F(10000)$ since we lack the cumulative distribution function (CDF) of the negative binomial distribution. Let's introduce another variable, which follows a binomial distribution.

$P(X > 10000) = P(\text{more than 10000 trials needed to get 2 wins}) = P(\text{there are less than 2 wins in 10000}$

$= P(Y < 2)$

where Y is the number of wins in 10000 matches, following a binomial distribution with parameters $n = 10000$ and $p = 0.0001$.

$$P(X > 10000) = P(Y < 2) = P(Y \leq 1) = F(1) = 0.736$$

The answer is 0.736, and it can be easily calculated using the following Octave code: 'binocdf(1, 10000, 0.0001)'.

c)

Probability of Not Feeling Sick for at Least 360 Days:

The probability of not feeling sick on any given day is 98%, or 0.98. The probability of feeling sick is $1 - 0.98 = 0.02$.

To find the probability of not feeling sick for at least 360 days out of 366 days, we can use the Poisson Approximation:

Given:

$$\text{Probability of not feeling sick on any given day} = 0.98$$
$$\text{Probability of feeling sick on any given day} = 0.02$$
$$\text{Number of days in a year} = 366$$

We can calculate the average number of days feeling sick (λ) in a year as:

$$\lambda = \text{Number of days in a year} \times \text{Probability of feeling sick on any given day} = 366 \times 0.02 = 7.32$$

Now, we want to find the probability of not feeling sick for at least 360 days in a year.

To approximate this, we use the Poisson distribution with parameter λ. The probability mass function (PMF) of the Poisson distribution is:

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

We want to find $P(X \leq 6)$. We use the cumulative distribution function (CDF) of the Poisson distribution to calculate $P(X \leq 6)$:

$$P(X \leq 6) = \sum_{i=0}^{6} \frac{e^{-\lambda}\lambda^i}{i!}$$

So, the probability of being sick for more than 6 days in a year is from the table A3 for $\lambda = 7$ our answer is 0.450 and for $\lambda = 7.5$ we get 0.378 and since 7.32 fits 7.5 more so we can say that our answer is ≈ 0.378.

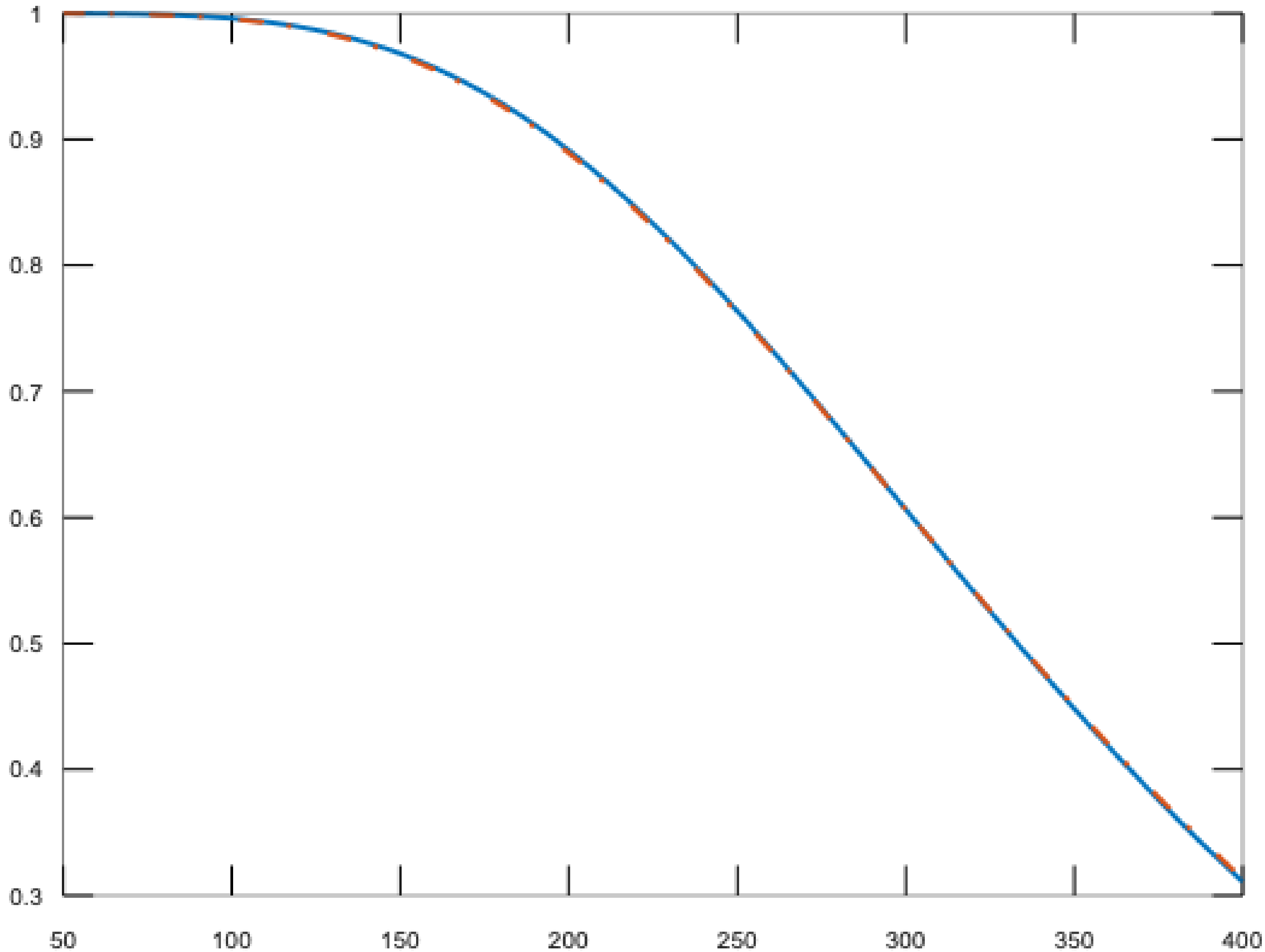
Answer 3

a)

Using Octave, we can achieve a more precise calculation. By evaluating the binomial cumulative distribution function (CDF) $F(6)$ with parameters $n = 366$ and $p = 0.02$, as mentioned earlier, we obtain the accurate result of 0.401. This outcome exceeds our initial approximation because our actual frequency is lower than the value we approximated.

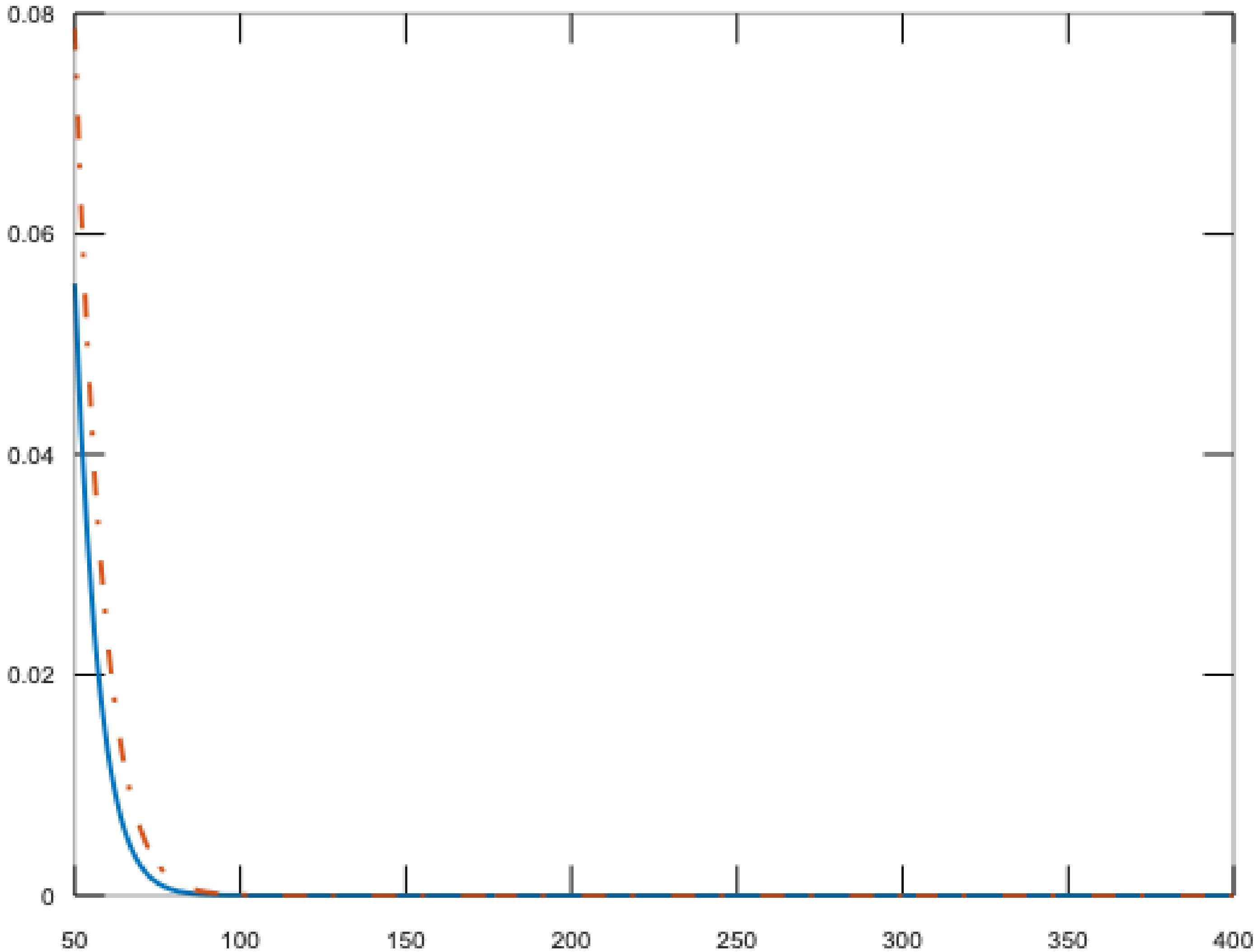
b)

Here is my Octave code for this question:
$>> p = 0.98;$
$>> ns = 50 : 400;$
$>> binomial\_probabilities = (binocdf(6, 50 : 400, 0.02));$
$>> poisson\_probabilities = (poisscdf(6, ns * 0.02));$
$>>$ close all;
$>> plot(ns,$ binomial_probabilities,$'linewidth', 2);$
$>>$ hold on;
$>> plot(ns,$ poisson_probabilities,$' - .', 'linewidth', 2);$
$>> saveas(1, "p = 0.98.png");$



c)

Here is my Octave code for this question:
>> $p = 0.78$;
>> $ns = 50 : 400$;
>> binomial_probabilities $= (binocdf(6, 50 : 400, 0.22))$;
>> $poisson\_probabilities = (poisscdf(6, ns * 0.22))$;
>> close all;
>> $plot(ns, binomial\_probabilities,' linewidth', 2)$;
>> hold on;
>> $plot(ns, poisson\_probabilities,' -.',' linewidth', 2)$;
>> $saveas(1, "p = 0.78.png")$;



Comparing two graphs, we can observe that as p increases, our approximation tends to give worse results, indicating an increasing difference between the Poisson and Binomial distributions. For instance, when $p = 0.02$, the blue and orange lines overlap, suggesting a good approximation. However, when $p = 0.22$, the blue and orange lines are separated, indicating a poorer approximation. This discrepancy arises because the Poisson approximation of the Binomial distribution performs better with smaller values of p.