

Tutorial Sheet 3

dtu1141b

February 1, 2025

Introduction

This document contains the solutions to the problems of Tutorial Sheet 3 of Probability and Statistics.

Problem 1

Question

Twenty percent of all telephones of a certain type are submitted for service while under warranty. Of these, 60% can be repaired, whereas the other 40% must be replaced with new units. If a company purchases ten of these telephones, what is the probability that exactly two will end up being replaced under warranty?

Solution

Let X be the number of telephones that need to be replaced under warranty out of the ten purchased. Since 20% of the telephones are submitted for service under warranty and 40% of these need to be replaced, the probability that a single telephone needs to be replaced is:

$$p = 0.2 \times 0.4 = 0.08$$

The probability that a single telephone does not need to be replaced is:

$$1 - p = 0.92$$

We are interested in the probability that exactly two out of ten telephones need to be replaced. This is a binomial probability problem where $n = 10$ and $k = 2$. The binomial probability formula is given by:

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

Substituting the values, we get:

$$P(X = 2) = \binom{10}{2} (0.08)^2 (0.92)^8$$

$$P(X = 2) = 45 \times (0.08)^2 \times (0.92)^8$$

$$P(X = 2) = 45 \times 0.0064 \times 0.4465 = 0.1150$$

Thus, the probability that exactly two out of the ten telephones will need to be replaced under warranty is approximately 0.1150 or 11.50%.

Problem 2

Question

A student who is trying to write a paper for a course has a choice of two topics, A and B. If topic A is chosen, the student will order two books through interlibrary loan, whereas if topic B is chosen, the student will order four books. The student believes that a good paper necessitates receiving and using at least half the books ordered for either topic chosen. If the probability that a book ordered through interlibrary loan actually arrives in time is 0.9 and books arrive independently of one another, which topic should the student choose to maximize the probability of writing a good paper? What if the arrival probability is only 0.5 instead of 0.9?

Solution

Let's denote the probability of a book arriving as p . For topic A, the student orders 2 books, and for topic B, the student orders 4 books. The student needs at least half of the books to arrive to be able to write a good paper.

Case 1: $p = 0.9$

For topic A (2 books), the student needs at least 1 book to arrive. The probability of this happening is:

$$P(\text{at least 1 book for A}) = 1 - P(\text{no books for A}) = 1 - (1 - p)^2 = 1 - (1 - 0.9)^2 = 1 - 0.01 = 0.99$$

For topic B (4 books), the student needs at least 2 books to arrive. The probability of this happening is:

$$P(\text{at least 2 books for B}) = 1 - P(0 \text{ books for B}) - P(1 \text{ book for B})$$

$$P(0 \text{ books for B}) = (1 - p)^4 = (1 - 0.9)^4 = 0.0001$$

$$P(1 \text{ book for B}) = \binom{4}{1} p(1 - p)^3 = 4 \cdot 0.9 \cdot (1 - 0.9)^3 = 4 \cdot 0.9 \cdot 0.001 = 0.0036$$

$$P(\text{at least 2 books for B}) = 1 - 0.0001 - 0.0036 = 0.9963$$

Since $0.9963 > 0.99$, the student should choose topic B if $p = 0.9$.

Case 2: $p = 0.5$

For topic A (2 books), the probability of at least 1 book arriving is:

$$P(\text{at least 1 book for A}) = 1 - P(\text{no books for A}) = 1 - (1 - p)^2 = 1 - (1 - 0.5)^2 = 1 - 0.25 = 0.75$$

For topic B (4 books), the probability of at least 2 books arriving is:

$$P(\text{at least 2 books for B}) = 1 - P(0 \text{ books for B}) - P(1 \text{ book for B})$$

$$P(0 \text{ books for B}) = (1 - p)^4 = (1 - 0.5)^4 = 0.0625$$

$$P(1 \text{ book for B}) = \binom{4}{1} p(1 - p)^3 = 4 \cdot 0.5 \cdot (1 - 0.5)^3 = 4 \cdot 0.5 \cdot 0.125 = 0.25$$

$$P(\text{at least 2 books for B}) = 1 - 0.0625 - 0.25 = 0.6875$$

Since $0.75 > 0.6875$, the student should choose topic A if $p = 0.5$

Problem 3

Question

A result called Chebyshev's inequality states that for any probability distribution of an r.v. X and any number k that is at least 1, $P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$. In words, the probability that the value of X lies at least k standard deviations from its mean is at most $\frac{1}{k^2}$.

1. Let X have possible values $-1, 0$ and 1 with probabilities $\frac{1}{18}, \frac{8}{9}$ and $\frac{1}{18}$, respectively. What is $P(|X - \mu| \geq 3\sigma)$, and how does it compare to the corresponding bound?
2. Give a distribution for which $P(|X - \mu| \geq 5\sigma) = 0.04$.
3. Using Chebyshev's inequality, calculate $P(|X - \mu| \geq k\sigma)$ for $k = 2$ and $k = 3$ when X follows Binomial distribution with parameters $n = 20, p = 0.5$ ($\text{Bin}(20, 0.5)$) and compare to the corresponding upper bound. Repeat the result for $n = 20, p = 0.75$.

Solution

(a) Calculation for X with values $-1, 0, 1$

Given the probabilities for X as $P(X = -1) = \frac{1}{18}$, $P(X = 0) = \frac{8}{9}$, and $P(X = 1) = \frac{1}{18}$, we first calculate the mean μ and variance σ^2 of X .

$$\mu = E(X) = (-1) \cdot \frac{1}{18} + 0 \cdot \frac{8}{9} + 1 \cdot \frac{1}{18} = 0$$

$$\sigma^2 = \text{Var}(X) = E(X^2) - [E(X)]^2 = (-1)^2 \cdot \frac{1}{18} + 0^2 \cdot \frac{8}{9} + 1^2 \cdot \frac{1}{18} = \frac{1}{9}$$

$$\sigma = \sqrt{\frac{1}{9}} = \frac{1}{3}$$

Now, we calculate $P(|X - \mu| \geq 3\sigma)$:

$$P(|X - 0| \geq 3 \cdot \frac{1}{3}) = P(|X| \geq 1) = 1 - P(X = 0) = 1 - \frac{8}{9} = \frac{1}{9}$$

According to Chebyshev's inequality:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} = \frac{1}{3^2} = \frac{1}{9}$$

The probability $P(|X - \mu| \geq 3\sigma)$ is exactly equal to the bound given by Chebyshev's inequality.

(b) Finding a Distribution for $P(|X - \mu| \geq 5\sigma) = 0.04$

Consider a discrete random variable X taking values $-1, 0, 1$ with corresponding probabilities p, q, p , where:

$$2p + q = 1.$$

$$\mu = E(X) = (-1)p + 0 \cdot q + 1 \cdot p.$$

Thus, the mean is $\mu = 0$. The variance is calculated as:

$$\sigma^2 = E(X^2) - [E(X)]^2.$$

Computing $E(X^2)$:

$$E(X^2) = (-1)^2 p + 0^2 q + 1^2 p = p + p = 2p.$$

If we let $\sigma = \frac{1}{5}$, we have:

$$\sigma^2 = \frac{1}{25}.$$

In such a scenario

$$2p = \frac{1}{25} \Rightarrow p = \frac{1}{50}, q = \frac{48}{50}$$

We are given:

$$P(|X - 0| \geq 5\sigma) = P(|X| \geq \frac{5}{5}) = P(|X| \geq 1).$$

Since $P(|X| \geq 1) = 2p$, we have:

$$2p = 0.04.$$

Substituting $p = \frac{1}{50}$,

$$2 \times \frac{1}{50} = \frac{2}{50} = \frac{1}{25} = 0.04,$$

which satisfies the requirement given to us by the problem.

Thus, the probability distribution is:

$$P(X = -1) = \frac{1}{50}, \quad P(X = 0) = \frac{24}{25}, \quad P(X = 1) = \frac{1}{50}.$$

In which $P(|X - \mu| \geq 5\sigma) = 0.04$.

(c) Chebyshev's inequality for Binomial distribution

For $X \sim \text{Binomial}(n = 20, p = 0.5)$:

$$\mu = np = 20 \cdot 0.5 = 10, \quad \sigma^2 = np(1 - p) = 20 \cdot 0.5 \cdot 0.5 = 5 \implies \sigma = \sqrt{5}$$

For $k = 2$:

$$P(|X - 10| \geq 2\sqrt{5}) = 1 - (F(14) - F(5)) \approx 1 - (0.97931 - 0.02069) \approx 0.04138 \leq \frac{1}{2^2} = 0.25$$

For $k = 3$:

$$P(|X - 10| \geq 3\sqrt{5}) = 1 - (F(16) - F(3)) \approx 1 - (0.99871 - 0.00129) \approx 0.00258 \leq \frac{1}{3^2} = \frac{1}{9} \approx 0.111$$

For $X \sim \text{Binomial}(n = 20, p = 0.75)$:

$$\mu = np = 20 \cdot 0.75 = 15, \quad \sigma^2 = np(1 - p) = 20 \cdot 0.75 \cdot 0.25 = 3.75 \implies \sigma = \sqrt{3.75} \approx 1.94$$

For $k = 2$:

$$P(|X - 15| \geq 2 \cdot 1.94) = 1 - (F(18) - F(11)) \approx 1 - (0.97569 - 0.04093) \approx 0.06524 \leq \frac{1}{2^2} = 0.25$$

For $k = 3$:

$$P(|X - 15| \geq 3 \cdot 1.94) = 1 - (F(20) - F(9)) \approx 1 - (1 - 0.00394) \approx 0.00394 \leq \frac{1}{3^2} = \frac{1}{9} \approx 0.111$$

Problem 4

Question

A second-stage smog alert has been called in a certain area of New Delhi in which there are 50 industrial firms. An inspector will visit 10 randomly selected firms to check for violations of regulations.

1. If 15 of the firms are actually violating at least one regulation, what is the pmf of the number of firms visited by the inspector that are in violation of at least one regulation?
2. If there are 500 firms in the area, of which 150 are in violation, approximate the pmf of part (a) by a simpler pmf.
3. For X = the number among the 10 visited that are in violation, compute $E(X)$ and $V(X)$ both for the exact pmf and the approximating pmf in part (b).

Solution

(a) PMF of the number of firms in violation

Let X be the number of firms out of the 10 inspected that are in violation of at least one regulation. Given that 15 out of 50 firms are violating regulations, the probability of a firm being in violation is $p = \frac{15}{50} = 0.3$.

Since the firms are selected randomly and without replacement, X follows a hypergeometric distribution. The probability mass function (pmf) of X is given by:

$$P(X = k) = \frac{\binom{15}{k} \binom{35}{10-k}}{\binom{50}{10}}$$

where k is the number of firms in violation, ranging from 0 to 10.

(b) Approximating the PMF

If there are 500 firms in the area, of which 150 are in violation, the probability of a firm being in violation is $p = \frac{150}{500} = 0.3$. Since the population is large and the sample size is small relative to the population, we can approximate the hypergeometric distribution with a binomial distribution. The pmf of X under this approximation is given by:

$$P(X = k) = \binom{10}{k} p^k (1-p)^{10-k}$$

where k ranges from 0 to 10.

(c) Expected Value and Variance

For the exact pmf (hypergeometric distribution):

$$E(X) = n \cdot p = 10 \cdot 0.3 = 3$$

$$V(X) = n \cdot p \cdot (1-p) \cdot \frac{N-n}{N-1} = 10 \cdot 0.3 \cdot 0.7 \cdot \frac{500-10}{500-1} = 1.89$$

For the approximating pmf (binomial distribution):

$$E(X) = n \cdot p = 10 \cdot 0.3 = 3$$

$$V(X) = n \cdot p \cdot (1-p) = 10 \cdot 0.3 \cdot 0.7 = 2.1$$

Problem 5

Question

Let X have a Poisson distribution with parameter μ . Show that $E(X) = \mu$ directly from the definition of expected value.

Solution

Let X be a Poisson random variable with parameter μ . The probability mass function (pmf) of X is given by:

$$P(X = k) = \frac{e^{-\mu} \mu^k}{k!}, \quad k = 0, 1, 2, \dots$$

The expected value $E(X)$ is calculated as:

$$E(X) = \sum_{k=0}^{\infty} k P(X = k) = \sum_{k=0}^{\infty} k \frac{e^{-\mu} \mu^k}{k!}$$

Since the $k = 0$ term is zero, we can start the sum from $k = 1$:

$$E(X) = \sum_{k=1}^{\infty} k \frac{e^{-\mu} \mu^k}{k!} = e^{-\mu} \sum_{k=1}^{\infty} \frac{\mu^k}{(k-1)!}$$

Let $j = k - 1$, then $k = j + 1$ and the sum becomes:

$$E(X) = e^{-\mu} \sum_{j=0}^{\infty} \frac{\mu^{j+1}}{j!} = e^{-\mu} \mu \sum_{j=0}^{\infty} \frac{\mu^j}{j!}$$

Recognizing the sum as the Taylor series expansion of e^{μ} , we get:

$$E(X) = e^{-\mu} \mu e^{\mu} = \mu$$

Thus, we have shown that $E(X) = \mu$ for a Poisson-distributed random variable.

Problem 6

Question

Automobiles arrive at a vehicle equipment inspection station according to a Poisson process with rate $\alpha = 10$ per hour. Suppose that with probability 0.5 an arriving vehicle will have no equipment violations.

1. What is the probability that exactly ten arrive during the hour and all ten have no violations?
2. For any fixed $y \geq 10$, what is the probability that y arrive during the hour, of which ten have no violations?
3. What is the probability that ten “no-violation vehicles” arrive during the next hour?

Solution

Given that automobiles arrive at a vehicle equipment inspection station according to a Poisson process with rate $\alpha = 10$ per hour and each vehicle has a probability of 0.5 of having no equipment violations.

(a) Probability that exactly ten arrive during the hour and all ten have no violations

The probability that exactly ten vehicles arrive during the hour is given by the Poisson distribution:

$$P(X = 10) = \frac{e^{-10} \cdot 10^{10}}{10!}$$

where X is the number of vehicles arriving.

The probability that all ten have no violations is 0.5^{10} . Therefore, the combined probability is:

$$P(\text{exactly 10 and all no violations}) = P(X = 10) \cdot 0.5^{10} = \frac{e^{-10} \cdot 10^{10}}{10!} \cdot 0.5^{10}$$

(b) Probability that y arrive during the hour, of which ten have no violations

The probability that exactly y vehicles arrive during the hour is:

$$P(X = y) = \frac{e^{-10} \cdot 10^y}{y!}$$

The probability that exactly ten out of these y vehicles have no violations follows a binomial distribution:

$$P(10 \text{ no violations out of } y) = \binom{y}{10} \cdot 0.5^{10} \cdot 0.5^{y-10}$$

Thus, the combined probability is:

$$P(\text{exactly } y \text{ arrivals, 10 no violations}) = P(X = y) \cdot \binom{y}{10} \cdot 0.5^{10} \cdot 0.5^{y-10}$$

(c) Probability that ten “no-violation vehicles” arrive during the next hour

Let Y be the number of vehicles with no violations. Y follows a Poisson distribution with parameter $\lambda = 10 \cdot 0.5 = 5$. The probability that exactly ten no-violation vehicles arrive is:

$$P(Y = 10) = \frac{e^{-5} \cdot 5^{10}}{10!}$$

Problem 7

Question

The Centers for Disease Control and Prevention reported in 2012 that 1 in 88 American children had been diagnosed with an autism spectrum disorder (ASD).

1. If a random sample of 200 American children is selected, what are the expected value and standard deviation of the number who have been diagnosed with ASD?
2. Referring back to (a), calculate the approximate probability that at least 2 children in the sample have been diagnosed with ASD?

Solution

(a) Expected Value and Standard Deviation

Given the probability of a child being diagnosed with ASD is $p = \frac{1}{88}$, we can model the number of children with ASD in a sample of 200 as a binomial random variable $X \sim \text{Binomial}(n = 200, p = \frac{1}{88})$.

The expected value $E(X)$ and standard deviation σ_X of X are given by:

$$E(X) = np = 200 \cdot \frac{1}{88} \approx 2.27$$

$$\sigma_X = \sqrt{np(1-p)} = \sqrt{200 \cdot \frac{1}{88} \cdot \left(1 - \frac{1}{88}\right)} \approx 1.484$$

(b) Probability of At Least 2 Children Diagnosed with ASD

Using the Poisson approximation, we want to find $P(X \geq 2)$, which is equivalent to $1 - P(X < 2)$.

$$P(X < 2) = P(X = 0) + P(X = 1) = e^{-\mu} + \mu e^{-\mu}$$

$$P(X < 2) = e^{-2.27} + 2.27e^{-2.27} \approx 0.103 + 0.231 = 0.334$$

Thus, the probability that at least 2 children have been diagnosed with ASD is approximately:

$$P(X \geq 2) = 1 - P(X < 2) = 1 - 0.334 = 0.666$$

Problem 8

Question

The literacy rate for a nation measures the proportion of people age 15 and over who can read and write. The literacy rate for women in that nation is 12%. Let X be the number of women you ask until one says that she is literate.

1. What is the probability distribution of X ?
2. What is the probability that you ask five women before one says she is literate?
3. Find the (i) mean and (ii) standard deviation of X .
4. Find the moment generating function of X and use it to verify the results in part (c).

Solution

Given that the literacy rate for women in a nation is $p = 0.12$, we can model the number of women X asked until one says she is literate as a geometric distribution with parameter p .

(a) Probability Distribution of X

The probability distribution of X is geometric:

$$P(X = k) = (1 - p)^{k-1}p, \quad k = 1, 2, 3, \dots$$

where $p = 0.12$ is the probability of success (a woman being literate).

(b) Probability of Asking Five Women Before One is Literate

The probability that you ask five women before one says she is literate is:

$$P(X > 5) = (1 - p)^5 = (1 - 0.12)^5$$

$$P(X > 5) = 0.88^5 \approx 0.470$$

(c) Mean and Standard Deviation of X

The mean $E(X)$ and standard deviation σ_X of a geometric distribution are given by:

$$E(X) = \frac{1}{p} = \frac{1}{0.12} \approx 8.333$$

$$\sigma_X = \sqrt{\frac{1-p}{p^2}} = \sqrt{\frac{0.88}{0.12^2}} \approx 7.23$$

(d) Moment Generating Function and Verification

The moment generating function $M_X(t)$ of X is:

$$M_X(t) = E(e^{tX}) = \frac{pe^t}{1 - (1-p)e^t} = \frac{0.12e^t}{1 - 0.88e^t}$$

To verify the mean, we take the first derivative and evaluate at $t = 0$:

$$M'_X(t) = \frac{d}{dt}M_X(t) = \frac{pe^t(1 - (1-p)e^t) - pe^t(-(1-p)e^t)}{(1 - (1-p)e^t)^2} = \frac{pe^t - pe^{2t} + p^2e^{2t} + pe^{2t} - p^2e^{2t}}{(1 - (1-p)e^t)^2}$$

$$M'_X(0) = \frac{p}{(1 - (1-p))^2} = \frac{1}{p} = \frac{1}{0.12} \approx 8.33$$

This matches our earlier calculation of the mean.

To verify the variance, we take the second derivative and evaluate at $t = 0$:

$$M''_X(t) = \frac{d}{dt}M'_X(t) = \frac{d}{dt} \left(\frac{pe^t}{(1 - (1-p)e^t)^2} \right) = \frac{pe^t(1 + (1-p)e^t)}{(1 - (1-p)e^t)^3}$$

$$M''_X(0) = \frac{p(2-p)}{(1 - (1-p))^3} = \frac{2p - p^2}{p^3}$$

$$\sigma_X = \sqrt{M''_X(0) - [M'_X(0)]^2} = \sqrt{\frac{1-p}{p^2}} = \sqrt{\frac{0.88}{0.12^2}} \approx 7.23$$

This matches our earlier calculation of the standard deviation.