Background

A manufacturing company produces light bulbs and wants to ensure consistent quality in its production process. The company management has hired you as a statistical analyst to evaluate production efficiency, defect rates, and overall product quality. You will apply statistical methods learned in Weeks 5-8 to analyze data and draw meaningful conclusions.

Core Issue

The company produces 10,000 light bulbs daily, and quality control checks are performed regularly. Defective bulbs must be identified and categorized. Your task is to analyze the data provided using appropriate probability distributions and statistical methods.

Assume the following statistics have been gathered from the plant's quality control department:

- The probability that a randomly chosen bulb is defective is 0.03.
- A sample of 100 bulbs is randomly selected for inspection.
- Historical data shows the defects follow specific probability distributions based on defect type.

Assignment Tasks

Part 1: Discrete Random Variables and Distributions

Binomial Distribution:

Question 1 - Determine the probability that exactly 5 out of 100 inspected bulbs are defective.
Question 2 - Calculate the probability that fewer than 3 bulbs are defective.

Geometric Distribution:

Question 3 - Find the probability that the first defective bulb appears on the 4th inspection.

Hypergeometric Distribution:

Question 4 - If the sample of 100 bulbs includes 10 defective bulbs, what is the probability of selecting 3 defective bulbs in a sub-sample of 10?

Poisson Distribution:

Question 5 - Given that defects occur at an average rate of 3 per hour in a continuous production line, what is the probability of observing exactly 5 defects in a given hour?

Part 2: Continuous Random Variables and Distributions

Normal Distribution:

Question 1 - The lifetime of light bulbs follows a normal distribution with a mean of 1,200 hours and a standard deviation of 200 hours. Find the probability that a randomly selected bulb lasts longer than 1,500 hours.
Question 2 - Determine the number of hours at which only 5% of the bulbs will have failed.

Standard Normal Distribution & Estimations:

Question 3 - Convert a raw score of 1,500 hours into a z-score and interpret its meaning.
Question 4 - Using the normal approximation to the binomial distribution, estimate the probability of at least 10 defective bulbs in the sample of 100.

Exponential Distribution:

Question 5 - If the time between defects follows an exponential distribution with an average of 30 minutes, what is the probability that the next defect occurs within 20 minutes?

Part 3: The Central Limit Theorem and Hypothesis Testing

Central Limit Theorem:

Question 1 - If repeated samples of 36 bulbs are taken from the production line, what is the standard deviation of the sample mean?
Question 2 - What is the probability that the average lifespan of a sample of 36 bulbs is less than 1,150 hours?

Hypothesis Testing:

Question 3 - The company claims that less than 4% of the bulbs produced are defective. Perform a hypothesis test at a 5% significance level to verify this claim using the sample data.

Finite Population Correction Factor:

Question 4 - If the total production in a day is considered a finite population of 10,000 bulbs, calculate the standard error of the sample mean when sampling 100 bulbs.

Answer

Part 1: Discrete Random Variables and Distributions

Question 1: Binomial Distribution
Determine the probability that exactly 5 out of 100 inspected bulbs are defective.

We use the Binomial distribution formula:
$$
P(X = k) = \binom{n}{k} p^k (1 - p)^{n - k}
$$
Where:
- $n = 100$ (the number of trials)
- $k = 5$ (the number of successes)
- $p = 0.03$ (the probability of success)

Calculating it:
$$
P(X = 5) = \binom{100}{5} (0.03)^5 (0.97)^{95}
$$

Calculating $\binom{100}{5}$:
$$
\binom{100}{5} = \frac{100!}{5!(100-5)!} = 252,887,520

\]
Now calculating \(P(X = 5) \):
\[
P(X = 5) = 252,887,520 \times (0.03)^5 \times (0.97)^{95}
\]
\[
P(X = 5) = 252,887,520 \times 0.00000243 \times 0.1159 \approx 0.0527
\]

Answer: \(P(X = 5) \approx 0.0527 \) (or 5.27%).

Question 2: Binomial Distribution
Calculate the probability that fewer than 3 bulbs are defective.

We have to calculate \(P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2) \).

Calculating each:
1. **For \(k = 0 \)**:
\[
P(X = 0) = \binom{100}{0} (0.03)^0 (0.97)^{100} \approx (0.97)^{100} \approx 0.0507
\]

2. **For \(k = 1 \)**:
\[
P(X = 1) = \binom{100}{1} (0.03)^1 (0.97)^{99} = 100 \times 0.03 \times (0.97)^{99} \approx 100 \times 0.03 \times 0.0520 \approx 0.1562
\]

3. **For \(k = 2 \)**:
\[
P(X = 2) = \binom{100}{2} (0.03)^2 (0.97)^{98} \approx 4950 \times 0.0009 \times (0.97)^{98} \approx 4950 \times 0.0009 \times 0.0535 \approx 0.2385
\]

Now summing these probabilities:
\[
P(X < 3) \approx 0.0507 + 0.1562 + 0.2385 \approx 0.4454
\]

Answer: \(P(X < 3) \approx 0.4454 \) (or 44.54%).

Question 3: Geometric Distribution
Find the probability that the first defective bulb appears on the 4th inspection.

We use the Geometric distribution formula:

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

Where $p = 0.03$ and $k = 4$.

Calculating it:

$$
P(X = 4) = (0.97)^{3} \times (0.03) \approx 0.912673 \times 0.03 \approx 0.0274
$$

Answer: $P(X = 4) \approx 0.0274$ (or 2.74%).

Question 4: Hypergeometric Distribution

If the sample of 100 bulbs includes 10 defective bulbs, what is the probability of selecting 3 defective bulbs in a sub-sample of 10?

Using the Hypergeometric distribution formula:

$$
P(X = k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}
$$

Where:
- $N = 100$ (total bulbs)
- $K = 10$ (defective bulbs in the population)
- $n = 10$ (sample size)
- $k = 3$ (number of defective bulbs in the sample)

Calculating it:

$$
P(X = 3) = \frac{\binom{10}{3} \binom{90}{7}}{\binom{100}{10}}
$$

Calculating each part:
- $\binom{10}{3} = 120$
- $\binom{90}{7} \approx 4,097,238,60$
- $\binom{100}{10} \approx 17,310,909,170$

Now substituting these values:

$$
P(X = 3) = \frac{120 \times 4,097,238,60}{17,310,909,170} \approx 0.2869
$$

Answer: $P(X = 3) \approx 0.2869$ (or 28.69%).

Question 5: Poisson Distribution

Given that defects occur at an average rate of 3 per hour in a continuous production line, what is the probability of observing exactly 5 defects in a given hour?

Using the Poisson distribution formula:
\[
P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}
\]
Where \(\lambda = 3 \) and \(k = 5 \).

Calculating it:
\[
P(X = 5) = \frac{e^{-3} \cdot 3^5}{5!} = \frac{e^{-3} \cdot 243}{120}
\]
Using \(e^{-3} \approx 0.0498 \):
\[
P(X = 5) = \frac{0.0498 \cdot 243}{120} \approx 0.1009
\]

Answer: \(P(X = 5) \approx 0.1009 \) (or 10.09%).

Part 2: Continuous Random Variables and Distributions

Question 1: Normal Distribution
Find the probability that a randomly selected bulb lasts longer than 1,500 hours.

To find this probability, we first standardize the value using the z-score formula:
\[
z = \frac{X - \mu}{\sigma}
\]
Where:
- \(X = 1500 \) hours
- \(\mu = 1200 \) hours (mean)
- \(\sigma = 200 \) hours (standard deviation)

Calculating the z-score:
\[
z = \frac{1500 - 1200}{200} = \frac{300}{200} = 1.5
\]

Now, we look up the z-score in the standard normal distribution table or use a calculator to find \(P(Z > 1.5) \).

Using the standard normal distribution, we find:
\[
P(Z < 1.5) \approx 0.9332
\]

Thus,
$$
P(Z > 1.5) = 1 - P(Z < 1.5) \approx 1 - 0.9332 = 0.0668
$$

Answer: $P(X > 1500) \approx 0.0668$ (or 6.68%).

Question 2: Normal Distribution
Determine the number of hours at which only 5% of the bulbs will have failed.

We need to find the value of X such that:
$$
P(X < x) = 0.05
$$
From the standard normal distribution table, the z-score corresponding to the lower 5% is approximately $z \approx -1.645$.

Now, we convert the z-score back to the original X value using the formula:
$$
X = \mu + z \cdot \sigma
$$
Substituting in the known values:
$$
X = 1200 + (-1.645) \cdot 200 = 1200 - 329 = 871
$$

Answer: Only 5% of the bulbs will have failed at approximately **871 hours**.

Question 3: Standard Normal Distribution & Estimations
Convert a raw score of 1,500 hours into a z-score and interpret its meaning.

Using the z-score formula:
$$
z = \frac{X - \mu}{\sigma}
$$
Substituting the values:
$$
z = \frac{1500 - 1200}{200} = \frac{300}{200} = 1.5
$$

Interpretation of the z-score: A z-score of 1.5 indicates that a bulb lasting 1,500 hours is 1.5 standard deviations above the mean lifetime of the bulbs. This implies that it is relatively rare, as it corresponds to a higher percentile in the distribution.

Question 4: Normal Approximation to the Binomial Distribution
Estimate the probability of at least 10 defective bulbs in the sample of 100.

Assuming the number of defective bulbs follows a binomial distribution with $n = 100$ and $p = 0.03$, we can use the normal approximation. The mean (μ) and standard deviation (σ) of the binomial distribution are calculated as:

$$
\mu = np = 100 \times 0.03 = 3
$$
$$
\sigma = \sqrt{np(1-p)} = \sqrt{100 \times 0.03 \times 0.97} \approx \sqrt{2.91} \approx 1.70
$$

We want to find $P(X \geq 10)$. Using the continuity correction, we calculate $P(X \geq 9.5)$:

Finding the z-score:
$$
z = \frac{9.5 - 3}{1.70} \approx \frac{6.5}{1.70} \approx 3.82
$$

Looking up this z-score in the standard normal distribution table:
$$
P(Z \geq 3.82) \approx 0.00007
$$

Answer: The probability of at least 10 defective bulbs is approximately $P(X \geq 10) \approx 0.00007$ (or 0.007%).

Question 5: Exponential Distribution
If the time between defects follows an exponential distribution with an average of 30 minutes, what is the probability that the next defect occurs within 20 minutes?

For an exponential distribution, we use the formula:
$$
P(X < x) = 1 - e^{-\lambda x}
$$
Where $\lambda = \frac{1}{\text{mean}} = \frac{1}{30}$ minutes.

Plugging in $x = 20$:
$$
P(X < 20) = 1 - e^{-\frac{1}{30} \cdot 20} = 1 - e^{-\frac{20}{30}} = 1 - e^{-\frac{2}{3}}
$$

Calculating $e^{-\frac{2}{3}}$:
Using a calculator, $e^{-\frac{2}{3}} \approx 0.5134$:

$$
P(X < 20) = 1 - 0.5134 \approx 0.4866
$$

Answer: The probability that the next defect occurs within 20 minutes is approximately $P(X < 20) \approx 0.4866$ (or 48.66%).

Let’s go through each question regarding the Central Limit Theorem and hypothesis testing step by step.

Part 3: The Central Limit Theorem and Hypothesis Testing

Question 1: Central Limit Theorem
If repeated samples of 36 bulbs are taken from the production line, what is the standard deviation of the sample mean?

The standard deviation of the sample mean (often called the standard error) is given by the formula:

$$
\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}
$$

Where:
- $\sigma = 200$ hours (the population standard deviation)
- $n = 36$ (the sample size)

Calculating the standard error:

$$
\sigma_{\bar{x}} = \frac{200}{\sqrt{36}} = \frac{200}{6} \approx 33.33
$$

Answer: The standard deviation of the sample mean is approximately **33.33 hours**.

Question 2: Central Limit Theorem
What is the probability that the average lifespan of a sample of 36 bulbs is less than 1,150 hours?

First, we calculate the z-score:

$$
z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}
$$

Where:
- $\bar{x} = 1150$ hours (the sample mean we are testing)

- $\mu = 1200$ hours (the population mean)
- $\sigma_{\bar{x}} \approx 33.33$ hours (standard deviation of the sample mean calculated previously)

Now calculating:

$$
z = \frac{1150 - 1200}{33.33} = \frac{-50}{33.33} \approx -1.5
$$

Next, we look up the z-score in the standard normal distribution table:

$$
P(Z < -1.5) \approx 0.0668
$$

Answer: The probability that the average lifespan of a sample of 36 bulbs is less than 1,150 hours is approximately **0.0668** (or 6.68%).

Question 3: Hypothesis Testing
The company claims that less than 4% of the bulbs produced are defective. Perform a hypothesis test at a 5% significance level to verify this claim using the sample data.

We will use a one-sample proportion test. Let’s assume the sample shows that out of 100 sampled bulbs, 2 are defective.

1. **Set up the hypotheses:**
   - Null hypothesis $H_0: p = 0.04$ (the proportion of defective bulbs is 4%)
   - Alternative hypothesis $H_a: p < 0.04$

2. **Calculate the sample proportion ($\hat{p}$):**

$$
\hat{p} = \frac{x}{n} = \frac{2}{100} = 0.02
$$

3. **Calculate the standard error (SE) for the sample proportion:**

$$
SE = \sqrt{\frac{p_0(1 - p_0)}{n}} = \sqrt{\frac{0.04(1 - 0.04)}{100}} = \sqrt{\frac{0.04 \times 0.96}{100}} \approx \sqrt{0.000384} \approx 0.0196
$$

4. **Calculate the z-score:**

$$
z = \frac{\hat{p} - p_0}{SE} = \frac{0.02 - 0.04}{0.0196} \approx \frac{-0.02}{0.0196} \approx -1.02
$$

5. **Determine the critical z-value at the 5% significance level:**

The critical z-value for a one-tailed test at $\alpha = 0.05$ is approximately -1.645.

6. **Decision rule:**
   If $z < -1.645$, we reject the null hypothesis.

Since -1.02 is greater than -1.645, we fail to reject the null hypothesis.

Answer: There is not enough evidence at the 5% significance level to support the claim that less than 4% of the bulbs produced are defective.

Question 4: Finite Population Correction Factor
Calculate the standard error of the sample mean when sampling 100 bulbs from a total production of 10,000 bulbs.

Given the finite population correction (FPC) factor is used when sampling without replacement, the formula for the standard error (SE) becomes:

$$
SE = \frac{\sigma}{\sqrt{n}} \times \sqrt{\frac{N - n}{N - 1}}
$$

Where:
- $\sigma = 200$ hours (population standard deviation)
- $n = 100$ (sample size)
- $N = 10,000$ (population size)

Substituting the values into the formula:

$$
SE = \frac{200}{\sqrt{100}} \times \sqrt{\frac{10000 - 100}{10000 - 1}} = \frac{200}{10} \times \sqrt{\frac{9900}{9999}} \approx 20 \times \sqrt{0.989} \approx 20 \times 0.9945 \approx 19.89
$$

Answer: The standard error of the sample mean when sampling 100 bulbs from a total production of 10,000 bulbs is approximately **19.89 hours**.