Assignment 3
DSL253 - Statistical Programming

Amay Dixit - 12340220

Submitted to Dr. Anil Kumar Sao

Links

- Notebook Link:
  `https://colab.research.google.com/drive/18plkYZjrGsSDY_YMIkT3WtDE3kWGzbwy?usp=sharing`
- Github Link:
  `https://github.com/amaydixit11/Academics/tree/main/DSL253/assignment_3`

1 Introduction

In this report, we analyze the performance of different marketing strategies and the behavior of battery lifetimes under various statistical distributions. The report is divided into two main parts.

- The first part examines the response rates to a marketing campaign, exploring the probability of success under various conditions.
- The second part focuses on the lifetimes of batteries, analyzed using Gamma and Chi-Squared distributions to estimate their expected behavior. This report also discusses the methodology used in calculating the key statistics, visualizing the distributions, and making inferences based on the results obtained.

2 Data

2.1 Marketing Campaign Data

The marketing team targets potential customers, with each customer having a fixed probability of responding positively to the campaign. The data for this analysis involves the number of successful responses (out of a fixed total number of customers) under different success probabilities. The team first reaches out to 15 customers with a fixed probability of success (20%), and then the analysis is extended to different probabilities (10%, 20%, ..., 90%).

2.2 Battery Lifetime Data

The battery lifetime is assumed to follow a Gamma distribution under extreme cold conditions with parameters:

$$\text{Shape parameter, } \alpha = 5, \quad \text{Scale parameter, } \theta = 4$$

We also explore how the lifetime distribution changes when the parameters are varied:

$$\alpha = 10, \theta = 0.9, \quad \alpha = 7, \theta = 2, \quad \alpha = 1.5, \theta = 0.2$$

Additionally, a new type of battery with a Chi-Squared distribution is analyzed, with 10 degrees of freedom.

3 Methodology

3.1 Marketing Campaign Analysis

In the marketing campaign analysis, we model the number of successful responses using a Binomial Distribution.

3.1.1 Binomial Distribution:

The number of successful responses X in a marketing campaign can be modeled by a Binomial Distribution with parameters n (the number of customers contacted) and p (the probability of success for each customer). The probability mass function (PMF) is given by:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \ldots, n$$

where:

- n is the total number of trials (customers),
- k is the number of successes (responses),
- p is the probability of success on each trial.

For part (a), we compute the PMF for $n = 15$ and various values of p (from 0.1 to 0.9) using the formula above. This gives us the probability of getting exactly k successes (successful responses) for each value of p.

3.1.2 Success Rate Distribution:

For part (c), the success rate Y is defined as:

$$Y = \frac{X}{n}$$

where X is the number of successful responses and n is the total number of customers contacted. The success rate is computed for different campaign sizes n (10, 20, 50, 200) and a fixed probability of success $p = 0.05$. The distribution of success rates is visualized by normalizing the number of successes by n.

3.2 Battery Lifetime Analysis

3.2.1 Gamma Distribution:

The lifetime of the battery is assumed to follow a Gamma Distribution with shape parameter α and scale parameter θ. The probability density function (PDF) of the Gamma distribution is given by:

$$f(x; \alpha, \theta) = \frac{x^{\alpha-1} e^{-x/\theta}}{\theta^{\alpha} \Gamma(\alpha)}, \quad x > 0$$

where:

- α is the shape parameter,
- θ is the scale parameter,
- $\Gamma(\alpha)$ is the Gamma function, defined as $\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt$.

The mean μ and variance σ^2 of the Gamma distribution are calculated as:

$$\mu = \alpha \times \theta, \quad \sigma^2 = \alpha \times \theta^2$$

For the Gamma distribution, the median is determined by finding the value x_{median} such that:

$$P(X \leq x_{\text{median}}) = 0.5$$

This is computed using the percent-point function (PPF) of the Gamma distribution. The PDF is plotted over a range of values (from 0.1 to 50 hours) to visualize the distribution of battery lifetime.

3.2.2 Chi-Squared Distribution:

For the second part of the analysis, the battery lifetime is assumed to follow a Chi-Squared Distribution with $\nu = 10$ degrees of freedom. The probability density function (PDF) of the Chi-Squared distribution is given by:

$$f(x;\nu) = \frac{x^{\frac{\nu}{2}-1} e^{-x/2}}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)}, \quad x > 0$$

For the Chi-Squared Distribution with 10 degrees of freedom, we compute the following statistics:

- Mean: $\mu = \nu$,
- Variance: $\sigma^2 = 2\nu$,
- Median: x_{median} such that $P(X \leq x_{\text{median}}) = 0.5$.

The PDF of the Chi-Squared distribution is plotted over the range from 0 to 24 hours.

Finally, we calculate the moments (mean, second moment, and third moment) of the Chi-Squared distribution using the Moment Generating Function (MGF). The MGF for the Chi-Squared distribution with ν degrees of freedom is given by:

$$M_X(t) = (1-2t)^{-\frac{\nu}{2}}, \quad \text{for } t < \frac{1}{2}$$

The moments are obtained by differentiating the MGF and evaluating at $t = 0$. The first derivative gives the mean:

$$E[X] = M_X'(0) = \nu$$

The second derivative provides the second moment:

$$E[X^2] = M_X''(0) = \nu(\nu + 2)$$

Using this, the variance is given by:

$$\mathrm{Var}(X) = E[X^2] - (E[X])^2 = 2\nu$$

Similarly, the third moment is obtained by taking the third derivative:

$$E[X^3] = M_X'''(0) = \nu(\nu + 2)(\nu + 4)$$

4 Results

4.1 Marketing Campaign Results

4.1.1 (a) PMF for 15 Customers (p = 0.2)

The PMF of successful responses for 15 customers with a conversion probability of 20% is as follows: The probability of exactly 3 successful responses is the highest, with a gradual decline as the number of successes increases or decreases from this point.

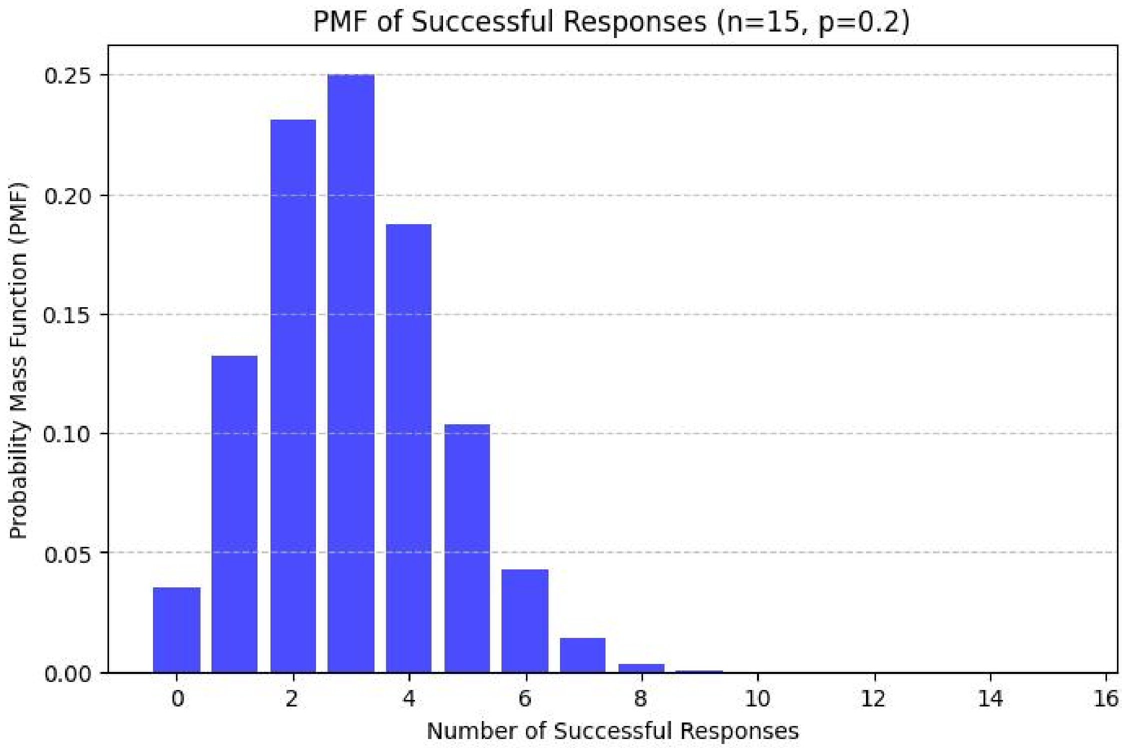

Figure 1: PMF for 15 Customers (p = 0.2)

4.1.2 (b) PMF for Different Conversion Probabilities

The PMF for various conversion probabilities (10%, 20%, ..., 90%) shows the following trends:

- As the probability of conversion increases, the distribution shifts towards higher numbers of successful responses.
- For higher probabilities (such as 70% and above), the majority of the mass is concentrated on the higher number of successes.

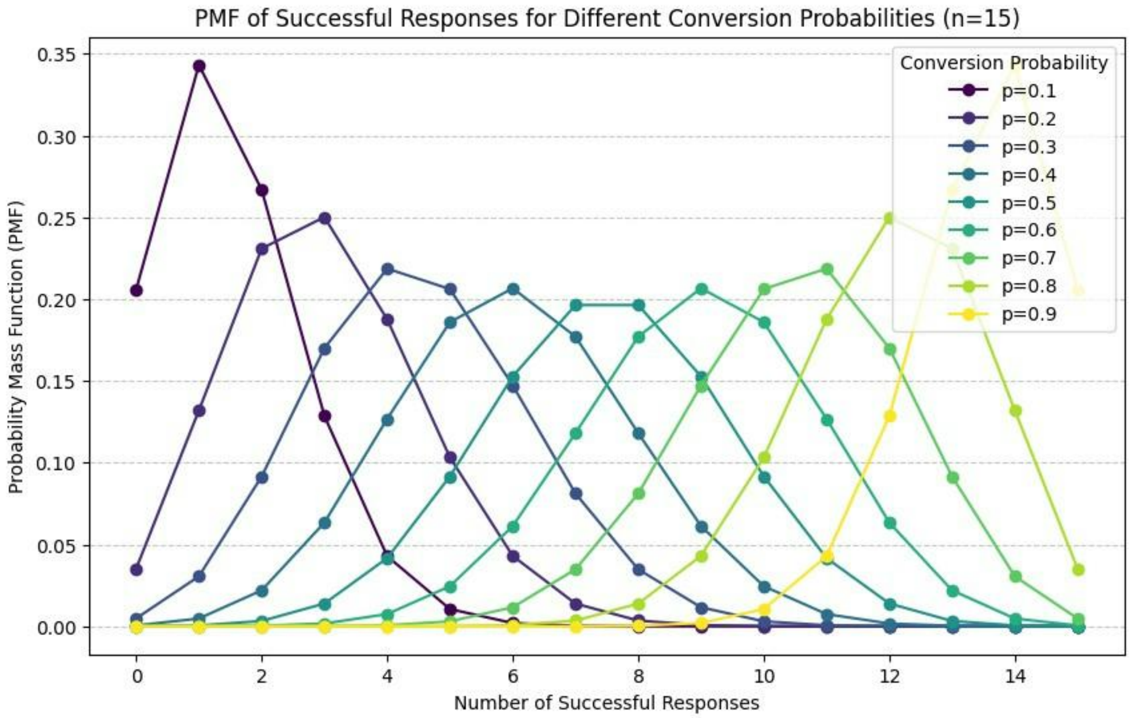

Figure 2: PMF for Different Conversion Probabilities

4.1.3 (c) Success Rate Distribution for Different Campaign Sizes

The success rate distribution for campaigns targeting 10, 20, 50, and 200 customers with a 5% conversion probability shows that as the campaign size increases, the success rate becomes more stable around the expected mean of 5%.

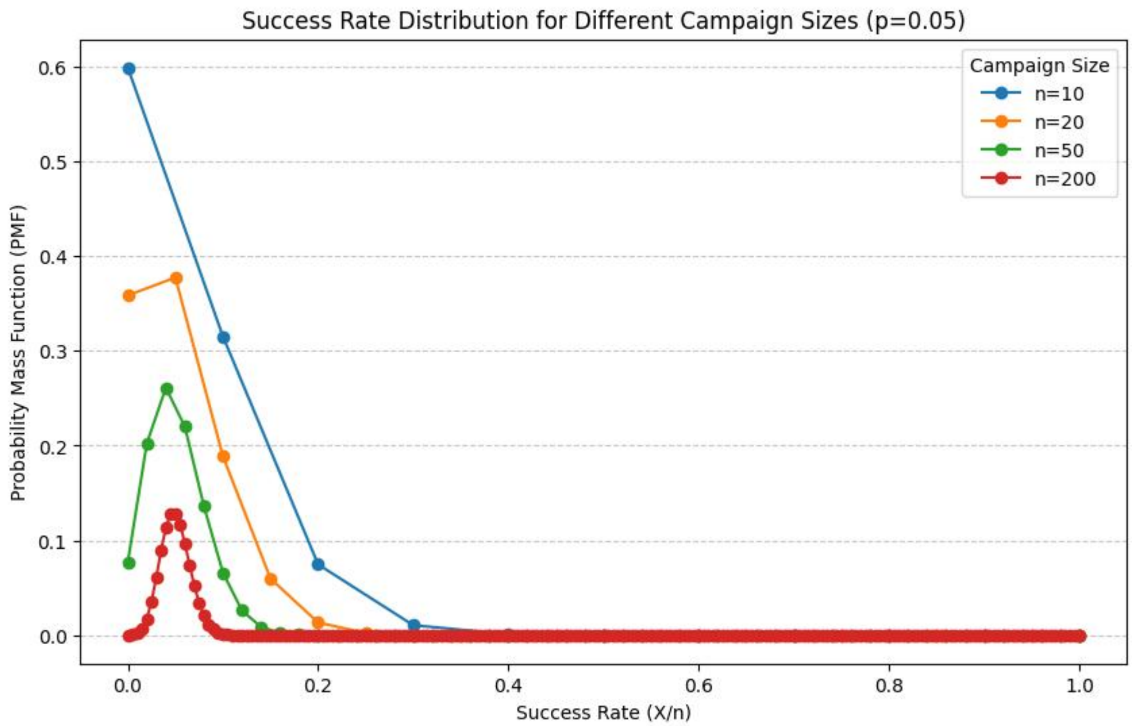

Figure 3: Success Rate Distribution for Different Campaign Sizes

4.2 Battery Lifetime Results

4.2.1 (a) Mean and Variance of Battery Lifetime

For the Gamma distribution with $\alpha = 5$ and $\theta = 4$:

- Mean Lifetime: 20 hours
- Variance: 80 hours^2

4.2.2 (b) Median Lifetime

The median lifetime for the given parameters ($\alpha = 5$, $\theta = 4$) is calculated as approximately 18.68 hours.

4.2.3 (c) PDF of Battery Lifetime (Gamma Distribution)

The PDF of the battery lifetime with the parameters $\alpha = 5$ and $\theta = 4$ shows a skewed distribution with a peak around 17 hours.

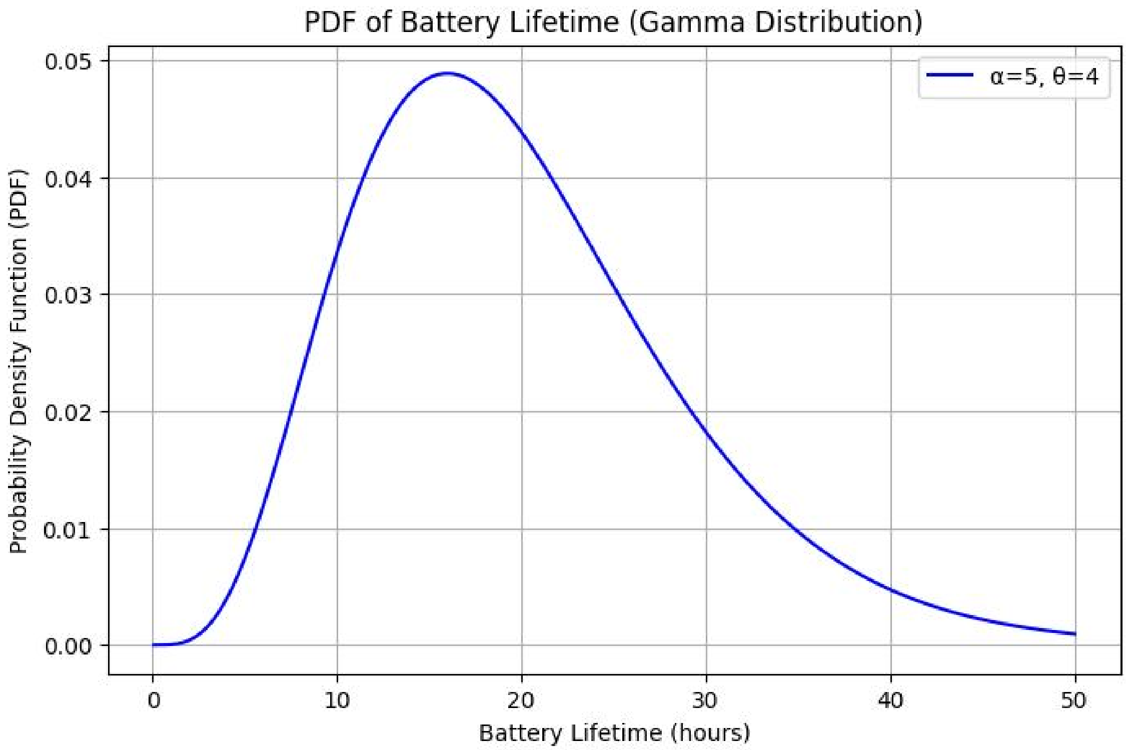

Figure 4: PDF of Battery Lifetime (Gamma Distribution)

4.2.4 (d) Comparison of Different Parameter Sets

The PDFs for different sets of α and θ parameters show the following patterns:

- As α increases, the distribution becomes less skewed and more symmetric. This is because the shape of the Gamma distribution approaches a Normal distribution for large α due to the **Central Limit Theorem**.
- For smaller values of α, the distribution is **positively skewed**, meaning it has a long right tail with a peak near zero. This happens because lower values of α correspond to a higher probability density near small values of x.

- A larger θ value results in a more spread-out distribution, as the scale parameter stretches the distribution along the x-axis.
- A smaller θ value results in a more concentrated distribution around the mean.

Mathematically, the skewness and kurtosis of the Gamma distribution can be expressed as:

$$\text{Skewness} = \frac{2}{\sqrt{\alpha}}$$

$$\text{Kurtosis (Excess)} = \frac{6}{\alpha}$$

From these formulas: - The **skewness** decreases as α increases, making the distribution more symmetric. - The **kurtosis** (which measures the "peakedness" of the distribution) also decreases as α increases, meaning the distribution becomes more bell-shaped.

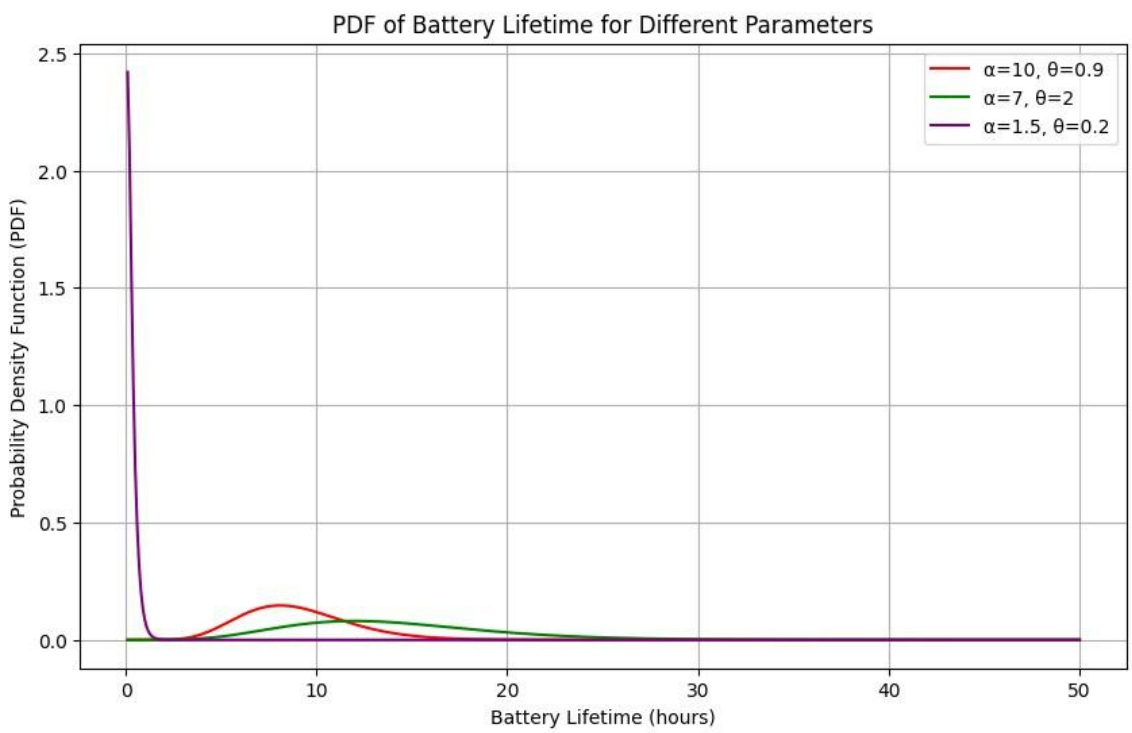

Figure 5: Comparison of Different Parameter Sets

4.3 Chi-Squared Distribution Results

4.3.1 (a) Mean, Standard Deviation, Median, Q1, and Q3

For the Chi-Squared distribution with 10 degrees of freedom:

- Mean: 10
- Standard Deviation: 4.47
- Median: 9.34
- Q1: 6.74
- Q3: 12.55

4.3.2 (b) PDF with Key Statistics

The Chi-Squared distribution with 10 degrees of freedom is plotted, with markers indicating the mean, median, and quartiles.

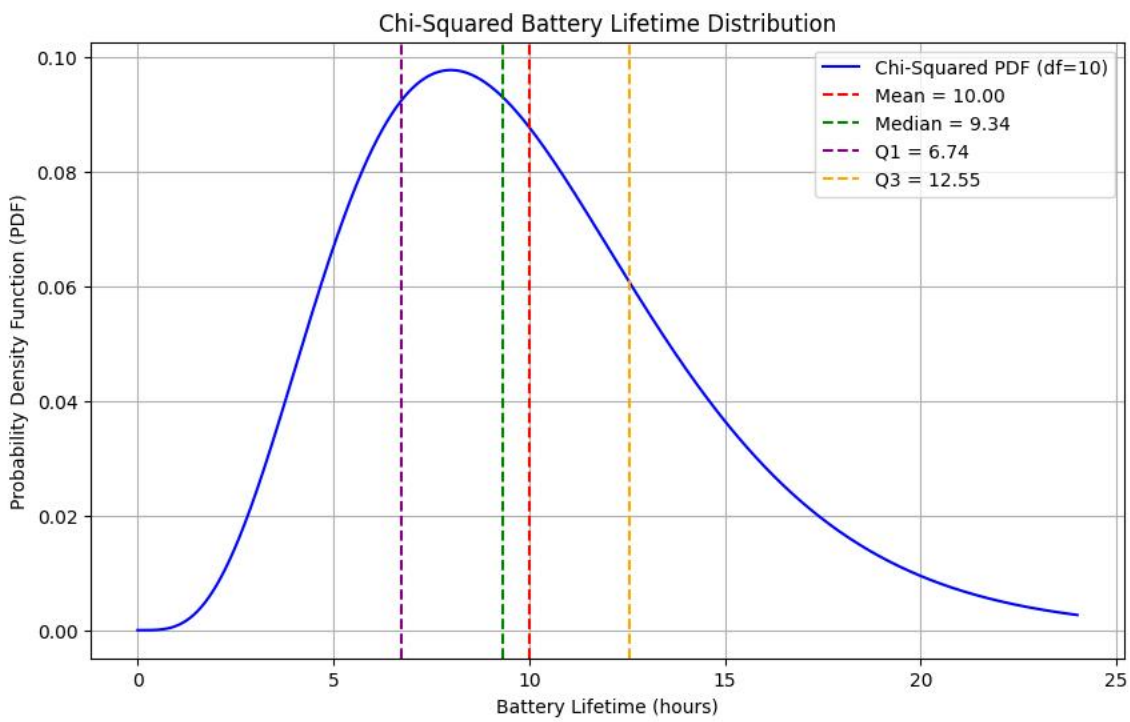

Figure 6: PDF with Key Statistics (Chi-Squared Distribution)

4.3.3 (c) Moments of the Chi-Squared Distribution

Using the Moment Generating Function (MGF), the moments are:

- First Moment (Mean): 10.00
- Second Moment: 120.00
- Third Moment: 1540.00

5 Discussion

5.1 Marketing Campaign Insights

The probability mass function (PMF) results for the marketing campaign provide key insights into the effectiveness of different strategies:

- The PMF for 15 customers with a conversion probability of 20% shows that the most likely outcome is around 3 successful conversions, with the probability decreasing symmetrically for both higher and lower success counts. This highlights the expected success range for small-scale campaigns.
- When varying the conversion probability, the PMF distributions shift accordingly. Lower probabilities result in a distribution concentrated near zero, indicating a high chance of very few conversions. Conversely, higher probabilities (e.g., 70% or more) push the distribution towards higher success counts, meaning that campaigns with strong conversion rates are more predictable in outcome.
- Analyzing success rate distributions across different campaign sizes (10, 20, 50, and 200 customers) reveals a key trend: larger campaigns exhibit lower variance in their success rates, meaning they stabilize around the expected mean of 5% conversion. This suggests that larger campaigns provide more consistent results, reducing uncertainty in marketing performance.

5.2 Battery Lifetime Insights

- The Gamma-distributed battery lifetime results indicate an expected mean of 20 hours, but with high variance (80 hours^2), implying substantial variation in battery performance across different units.

- The median lifetime of 18.68 hours is slightly lower than the mean, reinforcing the right-skewed nature of the distribution, where some batteries last significantly longer but the majority fall below the mean.
- The probability density function (PDF) illustrates that most batteries operate around 17 hours before failure, with a long right tail indicating occasional longer lifetimes.
- Parameter sensitivity analysis shows that increasing the shape parameter (α) makes the distribution more symmetric, resembling a normal distribution. This suggests that for larger α, the battery lifetimes become more predictable. Conversely, smaller α leads to a highly skewed distribution, emphasizing early failures.
- The scale parameter (θ) controls dispersion: higher θ spreads the lifetime distribution over a wider range, while lower θ concentrates lifetimes closer to the mean. This highlights the trade-off between battery consistency and longevity based on design parameters.

5.3 Chi-Squared Distribution Insights

- The Chi-Squared distribution with 10 degrees of freedom has a mean of 10, with a moderate right skew. This suggests that while most values cluster around the mean, higher values occur with decreasing probability.
- The quartiles (Q1 = 6.74, Median = 9.34, Q3 = 12.55) indicate that half of the values fall between 6.74 and 12.55, providing a clear range for expected outcomes.
- The PDF visualizes key statistics, reinforcing the concentration of values around the mean, with a noticeable right-skewed tail.
- Using the Moment Generating Function (MGF), we computed the first three moments:
  - The first moment (mean) confirms the expected value of 10.
  - The second moment (120) helps in understanding variance and spread.
  - The third moment (1540) provides insights into the shape of the distribution, reinforcing its positive skew.

These moments confirm that while the Chi-Squared distribution provides a reasonable prediction model for variance-heavy phenomena, it exhibits a notable tail effect, which must be accounted for in practical applications.

5.4 Overall Implications

The results across all analyses highlight key insights relevant to real-world applications:

- Marketing campaigns benefit from larger sample sizes, as they stabilize the success rate and reduce unpredictability. Higher conversion probabilities dramatically shift success likelihood, emphasizing the importance of optimizing campaign strategies.
- Battery lifetime predictions must consider parameter sensitivities, as shape and scale adjustments influence skewness and reliability. Manufacturers can design for either long lifetimes (higher θ) or more consistent performance (higher α).
- Chi-squared distributed variables exhibit moderate skewness, meaning that while most outcomes cluster around the mean, extreme values remain possible. This insight is crucial when using Chi-Squared models for uncertainty quantification.

Overall, the statistical analysis provides a deeper understanding of campaign performance, battery reliability, and distribution behaviors, which can guide decision-making in marketing, engineering, and data analysis contexts.

6 Conclusion

In conclusion, the marketing analysis shows that as the conversion probability increases, the outcome becomes more predictable, which can guide future campaign strategies. Additionally, larger campaigns will likely yield success rates closer to the expected value. For battery lifetimes, both the Gamma and Chi-Squared distributions provide valuable insights into the expected behavior of the batteries under extreme conditions. The results help to better understand the reliability of the batteries, which is crucial for their application in cold environments.