**DS201**

**Statistical Programming**

**Assignment 3**

**Ashutosh Kumar Jha**

**12340390**

**DSAI**

**2nd Year**

**Semester 4**

**---------------------------------------------------------------------------**

**1. Question 1: Analysis of the new marketing strategy**

Introduction:

-----------------------

A company is testing a new marketing strategy to attract potential customers. By running several campaigns, they analyze response rates to evaluate the effectiveness of their approach. The study focuses on understanding the probability of conversions and the impact of different success probabilities on campaign performance. Furthermore, a success rate metric is introduced to measure long-term performance trends.

Data:

------------------------

The data consists of the number of potential customers contacted (n) and the probability of conversion (p) per customer. The response data follows a binomial distribution, where each contacted customer either responds positively (success) or negatively (failure). The study explores:

* A fixed scenario with n = 15 and p = 0.2.
* Varying probabilities from 0.1 to 0.9 with n = 15.
* Success rate distributions for campaigns targeting 10, 20, 50, and 200 customers with a fixed probability of 0.05.

Methodology:

—-----------------------

The methodology involves using probability mass functions (PMFs) to analyze response distributions under different conditions:

1. **Fixed Probability Analysis:** The PMF is plotted for a binomial distribution where n = 15 and p = 0.2.
2. **Varying Probability Analysis:** The PMF is computed for success probabilities ranging from 0.1 to 0.9, showing how the response distribution changes.
3. **Success Rate Analysis:** The success rate metric (Y = X/n) is studied for different campaign sizes (n = 10, 20, 50, 200) with a fixed conversion probability (p = 0.05). The probability distribution of success rates is plotted to examine trends in large-scale campaigns.

Results:

-----------------------

1. **Fixed Probability PMF:** The plot demonstrates that achieving a moderate number of successful responses is more likely than extreme values (zero or all successful responses).
2. **Varying Probability Impact:** As p increases, the probability mass shifts towards higher successful responses, indicating higher likelihoods of achieving success with more favorable conditions.
3. **Success Rate Distribution:** As the number of targeted customers increases, the success rate distribution becomes more concentrated around the mean success rate (p = 0.05). This suggests that larger campaigns yield more stable and predictable performance.

Discussions:

----------------------

* The PMF plots help visualize expected success levels under different conditions, aiding decision-making in campaign planning.
* Increasing the conversion probability directly improves campaign effectiveness, but factors influencing conversion (e.g., customer engagement, offer attractiveness) should be considered.
* Large-scale campaigns lead to a more stable success rate distribution, implying that large outreach efforts minimize variability and provide more reliable estimates of expected success.

Conclusion:

-----------------------

This study provides insights into the probability of customer conversions under different scenarios. The findings emphasize the importance of optimizing campaign strategies to improve conversion rates. Moreover, large-scale campaigns exhibit more predictable success rates, making them a valuable approach for consistent marketing performance. The analysis serves as a quantitative foundation for making data-driven marketing decisions.

**2. Question 2: Analysis for the lifetime of a battery under the gamma distribution conditions**

1. Introduction:

----------------------

The lifetime of batteries used under extremely cold conditions is crucial for performance and reliability. This study examines battery lifetime following a Gamma distribution with different parameter settings to analyze expected performance and variability. By understanding the distribution of battery lifetimes, businesses can make informed decisions on battery selection and maintenance.

2. Data

--------------------

The lifetime (in hours) of the battery follows a Gamma distribution with:

* Shape parameter (α = 5)
* Scale parameter (θ = 4)

Additionally, alternative parameter sets are analyzed:

* (α = 10, θ = 0.9)
* (α = 7, θ = 2)
* (α = 1.5, θ = 0.2)

3. Methodology

---------------------

The study involves:

1. **Expected Lifetime and Variability Calculation:** The mean, variance, and standard deviation of battery lifetime are computed based on the Gamma distribution.
2. **Median Lifetime Calculation:** The median battery lifetime is determined using the inverse cumulative distribution function (CDF), representing the point where 50% of batteries last longer and 50% last shorter.
3. **Probability Density Function (PDF) Visualization:** A PDF plot is generated for battery lifetimes ranging from 0.1 to 50 hours, providing insight into the expected lifetime distribution.
4. **Parameter Variation Analysis:** The PDF is analyzed for different parameter sets to observe how changes in α and θ affect the expected lifetime distribution.

4. Results

--------------------

1. **Expected Lifetime and Variability:**
   * The expected battery lifetime is calculated as α × θ, providing insight into the average lifespan.
   * Variance and standard deviation describe the spread of lifetimes around the mean.
2. **Median Lifetime:**
   * The median lifetime is lower than the mean in right-skewed distributions, highlighting asymmetry in battery lifetimes.
3. **PDF Analysis:**
   * The shape of the PDF indicates the likelihood of various battery lifetimes, with higher α leading to more concentrated distributions.
4. **Parameter Variation:**
   * Increasing α while reducing θ results in a more peaked distribution with lower variability.
   * Smaller values of α create more skewed distributions, indicating higher uncertainty in battery performance.

5. Discussion

-------------------

* Understanding the lifetime distribution helps optimize battery replacement strategies.
* Different α and θ values impact reliability and predictability, essential for extreme conditions.
* Manufacturers can use these insights to design batteries with optimal trade-offs between longevity and stability.

6. Conclusion

------------------  
The Gamma distribution effectively models battery lifetimes under extreme conditions. The analysis highlights how shape and scale parameters influence expected performance. Businesses can leverage these findings to improve battery selection and operational planning for cold environments.

**3. Question 3: Analysis of the lifetime of a battery which follows the chi-squared distribution**

1. Introduction

--------------------

The company has developed a new type of battery whose lifetime (in hours) follows a chi-squared distribution with 10 degrees of freedom. Understanding this distribution provides valuable insights into battery reliability and performance.

2. Data

----------

The lifetime data follows a chi-squared distribution with:

* Degrees of freedom (df) = 10

The statistical properties of the chi-squared distribution will be explored, including measures of central tendency and variability.

3. Methodology

---------------------

1. **Computation of Statistical Measures:**
   * The mean, variance, and standard deviation of the battery lifetime are computed based on chi-squared distribution properties.
   * Median, first quartile (Q1), and third quartile (Q3) are determined using the inverse cumulative distribution function (CDF).
2. **Probability Density Function (PDF) Visualization:**
   * The PDF of the battery lifetime is plotted from 0 to 24 hours.
   * Mean, median, and quartiles are highlighted on the plot to visualize their locations within the distribution.
3. **Moment Generating Function (MGF) Analysis:**
   * The first, second, and third moments of the chi-squared distribution are derived from the MGF properties.

4. Results

---------------

1. **Statistical Measures:**
   * The expected mean of the battery lifetime is equal to the degrees of freedom (df = 10).
   * Variance is computed as 2 × df, with the standard deviation derived accordingly.
   * The median and quartiles provide additional insight into the distribution’s spread and asymmetry.
2. **PDF Analysis:**
   * The PDF plot visually represents the battery lifetime distribution.
   * The placement of the mean, median, and quartiles illustrates the right-skewed nature of the chi-squared distribution.
3. **Moment Analysis:**
   * The first moment (mean) confirms the expected value.
   * The second moment accounts for variance and squared mean.
   * The third moment provides insights into skewness and higher-order distribution properties.

5. Discussion:

------------------

* The right-skewed nature of the chi-squared distribution suggests that most batteries have lifetimes clustered near the lower range, but some last significantly longer.
* Understanding quartiles helps in determining the percentage of batteries that last beyond a specific threshold, aiding in warranty and reliability estimates.
* Moment analysis offers deeper insights into how lifetime distributions behave in practical applications.

6. Conclusion

------------------

The chi-squared distribution effectively models the battery lifetime, with key statistical properties providing insights into expected longevity and variability. The results help manufacturers understand performance distribution and guide decisions on quality control and customer guarantees.

Code: [12340390 Ashutosh Asg3.ipynb](https://colab.research.google.com/drive/1mxT0zDt3UP04_TPMG9CmOzMND1sxNIMQ?usp=sharing)