BONTA AALAYA
12340530
DS250 - Statistical Programming

1.

Introduction

In this experiment, we aim to study and analyse pmf of successful customer responses, and observe how varying sample sizes and conversion probabilities affects the overall output of the marketing campaign

Data

The data was generated using inbuilt functions in matlab (Statistics and Machine Learning Toolbox used)

Parameters n, p, and number of simulations are defined within each context

Methodology

(a) By setting parameters, simulating distributions, and plotting the pmf of successful responses against the number of successful responses, we can anaylse the trends in expected marketing outcomes
(b) We go through all p values {0.1, 0.2, … , 0.9} and plot the pmf, overlaying pmfs for different values of p
(c) By taking different sample sizes and plotting them individually, and then combining the graphs into one graph, we can analyse how sample size affects overall success rate

Results

Below are the graphs obtained from matlab

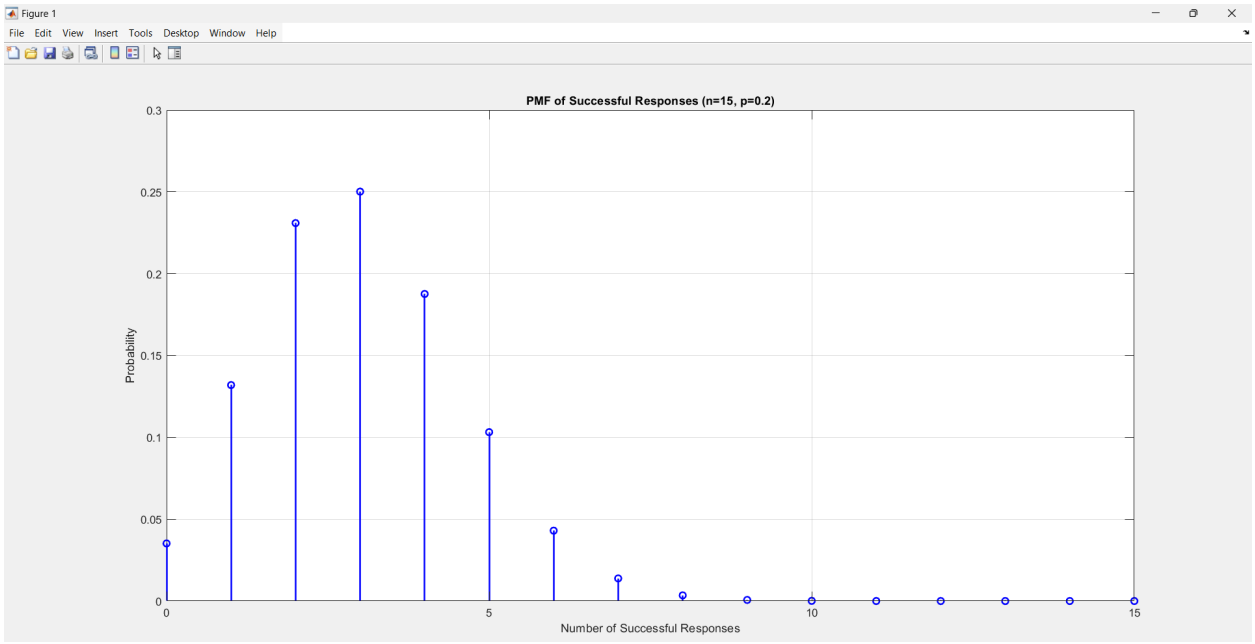


Figure 1
File Edit View Insert Tools Desktop Window Help
PMF of Successful Responses (n=15, p=0.2)
Probability
Number of Successful Responses
0
0.05
0.1
0.15
0.2
0.25
0.3
0
5
10
15

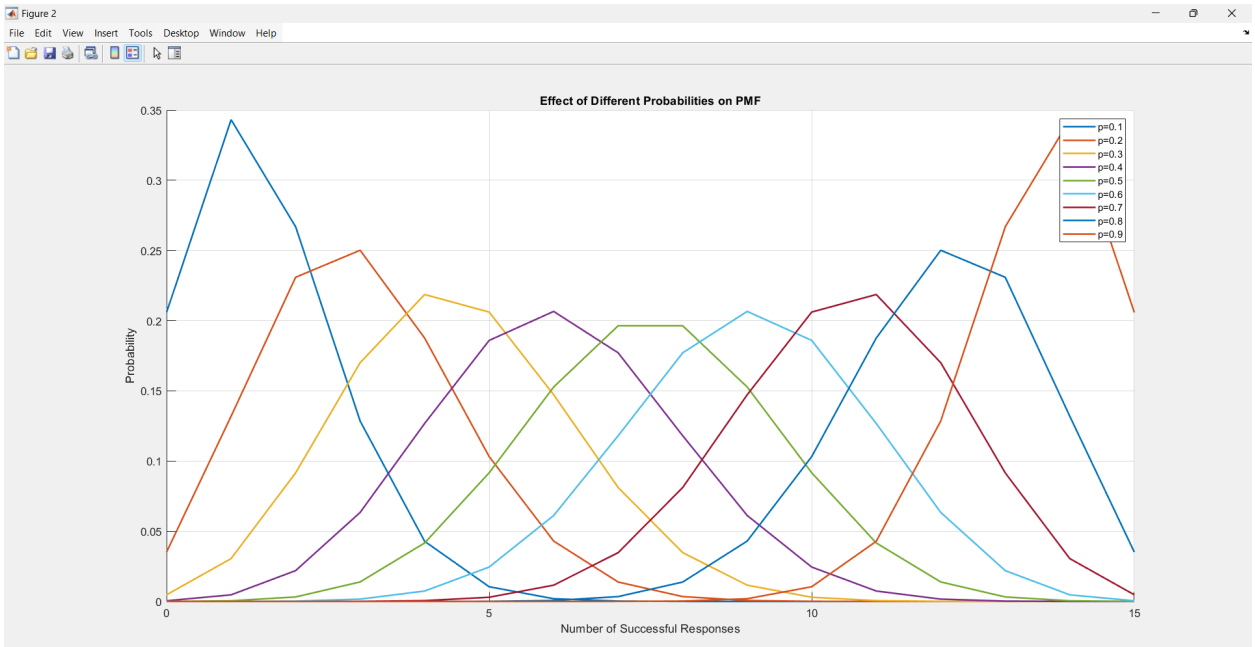
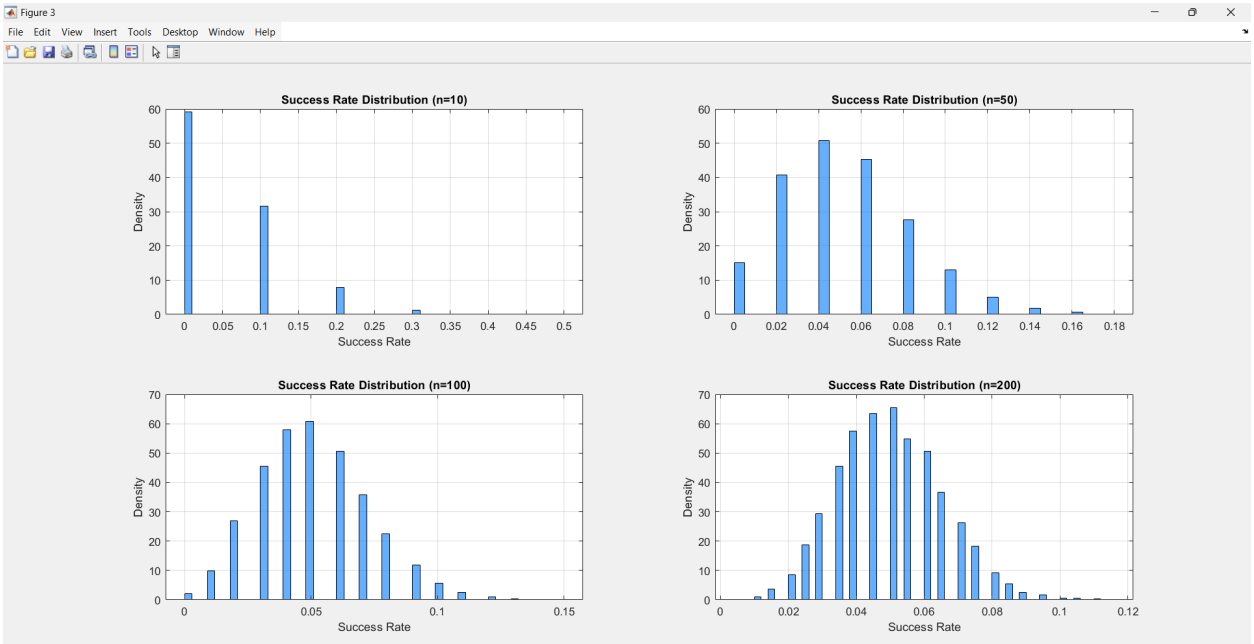
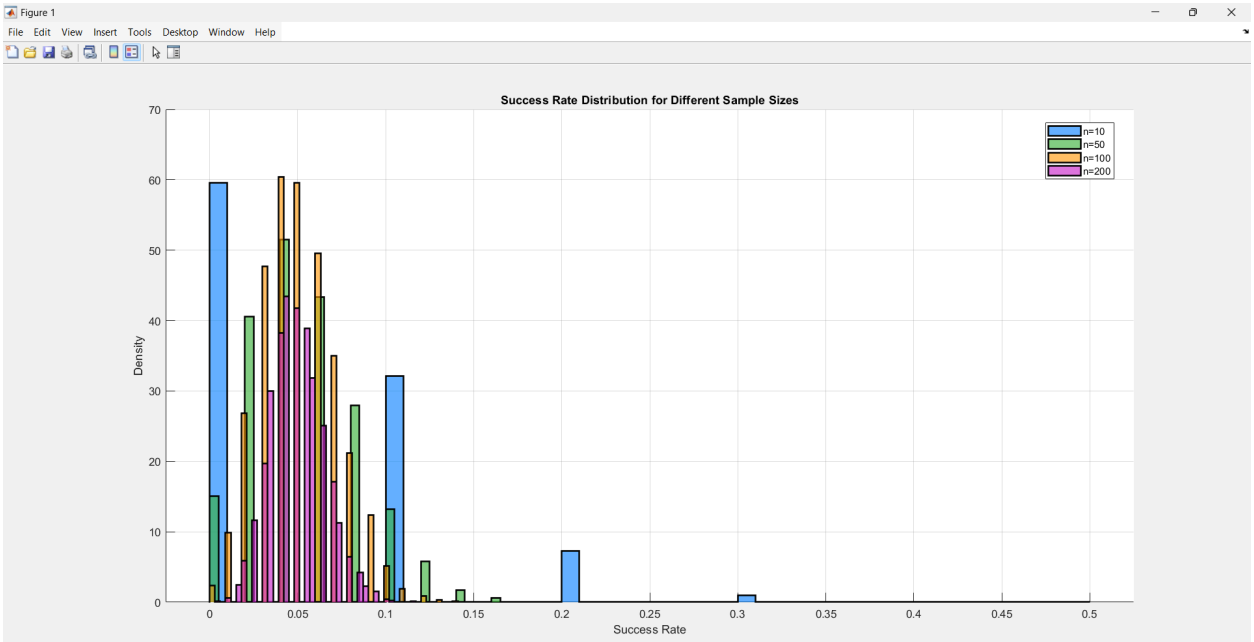


Figure 2
File Edit View Insert Tools Desktop Window Help
Effect of Different Probabilities on PMF
Probability
Number of Successful Responses
p=0.1
p=0.2
p=0.3
p=0.4
p=0.5
p=0.6
p=0.7
p=0.8
p=0.9
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0
5
10
15



Overlaid output:



Discussion

For a fixed number of customers, the probability of achieving higher successful conversions increases as the conversion probability rises. We can also observe that smaller campaigns show

more variability in success rates, while larger campaigns provide more stable performance (relatively less variability). While increasing the sample size stabilizes results, the success rate remains dependent on the underlying probability of conversion. Campaigns with low response probabilities continue to show limited success rates, showing the importance of targeting the right audience and good marketing. For lower probabilities, pmf is more skewed towards zero, showing a high likelihood of failure in smaller campaigns. As probability increases, the distribution flattens and shifts to the right, showing a greater likelihood of moderate to high success.

Conclusion

We can predict customer response patterns and optimize campaigns by analysing the binomial pdfs for different sample sizes and success probabilities. The findings suggest that increasing campaign size stabilizes success rates, making large-scale marketing efforts more reliable. However, simply expanding the number of customers reached is not enough; ensuring a high conversion probability (p) is also a key factor in determining this success rate. Results show that audience segmentation and message enhancement should be focused on.

2,3

Introduction

In this experiment, we use the knowledge that the lifetime of a battery used in extremely cold conditions follows a given set of parameters. With this in mind, we compute the required statistics according to the conditions mentioned in the question

Data

The required distributions are generated using builtin functions in matlab and the parameters used differ depending on the sub-question (Statistics and Machine Learning Toolbox used)

Methodology

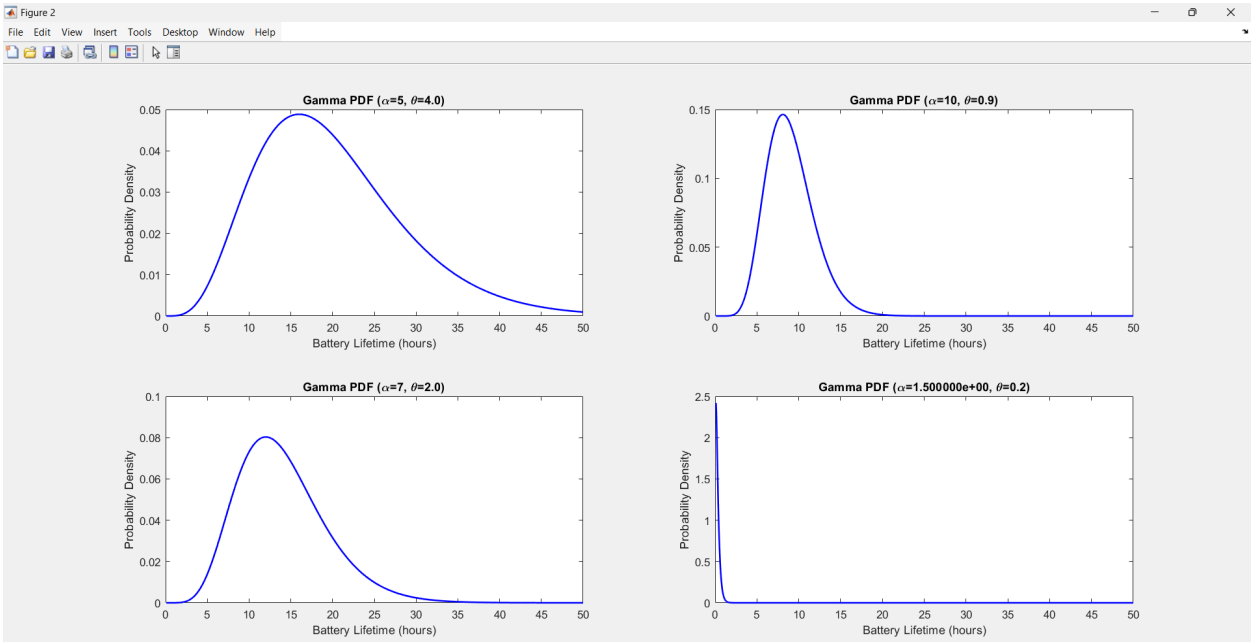
To calculate Gamma and Chi squared distributions, we use builtin function makedist and chi2pdf to generate distributions and compute moments using built-in moment function for Gamma

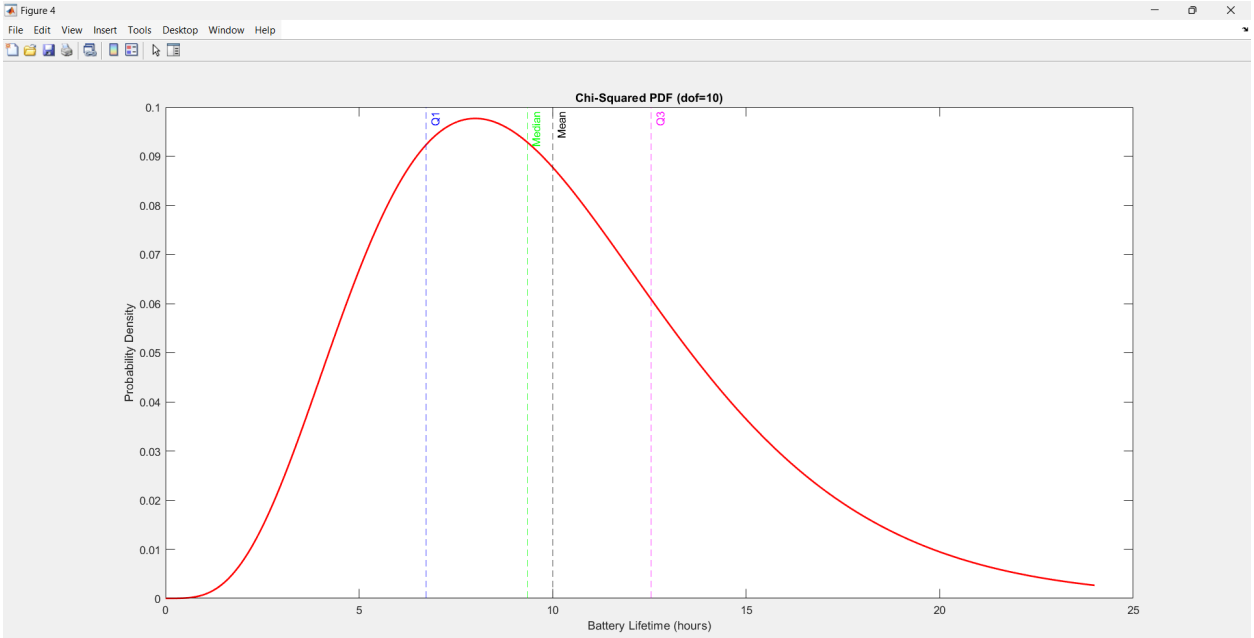
distribution, and use formulas from [1] for Chi square distribution because matlab doesn't support moment function for chi square distribution

For Gamma distribution, mean lifetime is $\alpha \times \theta$ and variance is $\alpha \times \theta^2$, and median is found by using invcdf
For Chi square distribution, mean = dof, and variance is = 2 x dof . Median and quartiles are found using builtin functions

Results

The graphs obtained from implementing code in matlab are as given below





First moment: 10
Second moment: 120
Third moment: 1680

Discussion

The plots show how battery lifetime varies according to different parameters. For Gamma distribution, we can see that scale and shape parameters affect longevity. Higher α values show less variance/randomness and different θ values show different rates of failure. Further, we can say from the Chi square plot that battery lifetimes are positively skewed, and variance is high, which means that some batteries will die out faster than mean, but some have extended lifetime (much more than mean)

Conclusion

High α value indicates that the battery life is more predictable, and θ value is responsible for rate of failure. The chi square plot not being symmetric reinforces this idea that a certain skewness is present, and that most batteries will die out before mean (i.e, 10 hours), but a few batteries will last much longer (as seen from the long tail). Analysis of the moments helps us understand how lifetimes are spread out

References

[1] https://statproofbook.github.io/P/chi2-mom.html