

Computational Statistics Project I

MSc in Statistics, 2023-2024

A company wants to market sticker collections for children. The collection they want to issue consists of n different stickers and each pack contains exactly k stickers, never containing the same stickers. The problem of the company is to find the appropriate combination of values for n and k so as to optimize their profits.

The limitations that exist are

- The possible values of n are 200, 300 and 400 while for k are 5,6,7 and 8.
- About the production cost for each sticker (the album for the stickers is given free of charge by the company), a collection of 200 stickers costs 4 Euros, while respectively for 300 and 400 stickers the cost is 7 and 11 euros respectively.
- Each sticker costs the company 0.02 euros, therefore a pack of k stickers costs to the company $0.02k$ euro.
- If the pack has 5,6,7 or 8 stickers the selling price are 0.12, 0.15, 0.18 and 0.20 Euro respectively.

The questions to be answered are:

- If we assume that a customer buys until to complete the collection, which combination of n and k will give the greater profits in the company?
- For each combination of n and k what is the probability one buyer with a total amount of 40 euros to complete the collection?
- What does it change if there can be twice (or more) the same sticker in each pack? How likely is this to happen?

Notes

The cost for a collection of n stickers with packs containing k stickers is $K(n) + 0.02kY$ where

$$K(n) = \begin{cases} 4 & n = 200 \\ 7 & n = 300 \\ 11 & n = 400 \end{cases}$$

Y is the number of packs someone will buy until complete the collection. The amount that will be collected from each customer is $M(k)Y$ where

$$M(k) = \begin{cases} 0.12 & k = 5 \\ 0.15 & k = 6 \\ 0.18 & k = 7 \\ 0.20 & k = 8 \end{cases}$$

You should submit a fairly detailed report so help the company make the best decision, taking into account all factors of uncertainty.