

STATS604_project3

2025-10-23

Research Question

How do fridge location (by layer) and bagging affect the freshness of cilantro?

Factors and Experiments

We consider two factor variables. The first is location with 4 levels (“Outside”, “Layer 1”, “Layer 2”, “Layer 3”). The last 3 different levels are placed in the different layers in the refrigerator with different temperatures. The second factor is bag with 2 levels (“In Bag”, “Out of Bag”). We set 5 bunches of cilantro for each condition. The total number of cilantro bunch $N=4(\text{locations}) \times 2(\text{bag types}) \times 5(\text{replicates})=40$.

Measurement

We measure the shelf-life in days for each individual cilantro for each bunch in each level. Also we use a binary event indicator to denote whether the herb is still acceptable. If the cilantro bunch is unacceptable (e.g., the yellow and brown leaves are >50% or the freshness score is less than 2), we denote status = 1. Otherwise, we denote the status as 0, which means the herb is still acceptable.

Analysis Plan

EDA

We will compute the mean survival time for each of the 8 groups and visualize the data by plotting K-M survival curves.

Regression Model

To test the main effects and their interaction, we will build a regression model. The model is as below:

$$Y = \beta_1 \times Location + \beta_2 \times Bag + \beta_3 \times Location : Bag$$

Here Y can be the survival time or the status indicating whether the herbs are still acceptable.

Log-rank Test

Next we will use log-rank test. First we want to compare different layers in the fridge. Every three days, we count the number of acceptable herbs and unacceptable ones and then we can construct a 2×3 contingency table.

We calculate the Observed (O) number of failures in each group and the Expected (E) number of failures (under the null hypothesis). The final test statistic follows chi-square distribution with 2 degrees of freedom.