Project 2 - Cold Storage Problem

Juan Esteban Venegas

08 November 2019

Contents

Problem 1

Cold Storage started its operations in Jan 2016. They are in the business of storing Pasteurized Fresh Whole or Skimmed Milk, Sweet Cream, Flavoured Milk Drinks. To ensure that there is no change of texture, body appearance, separation of fats the optimal temperature to be maintained is between 2 - 4 C.

In the first year of business, they outsourced the plant maintenance work to a professional company with stiff penalty clauses. It was agreed that if it was statistically proven that probability of temperature going outside the 2 - 4 C during the one-year contract was above 2.5% and less than 5% then the penalty would be 10% of AMC (annual maintenance case). In case it exceeded 5% then the penalty would be 25% of the AMC fee. The average temperature data at date level is given in the file "Cold_Storage_Temp_Data.csv"

Step I.

Load libraries, specify working directory and import raw data.

```
library(data.table)
library(stats)
setwd('C:/Users/Juan Esteban Venegas/Desktop/Project_2_Cold_Storage_Problem/')

Cold_Storage_Temp_Data <- fread('Cold_Storage_Temp_Data.csv')
Cold_Storage_Mar2018 <- fread('Cold_Storage_Mar2018.csv')
```

Step II. Questions

Find mean cold storage temperature for Summer, Winter and Rainy Season

```
Cold_Storage_Temp_Data[, .(mean_temp = mean(Temperature, na.rm = TRUE)), by = .(Season)]
```

```
##    Season mean_temp
## 1: Winter  2.776423
## 2: Summer  3.147500
## 3:  Rainy  3.087705
```

Find overall mean for the full year

```
CS_mean <- mean(Cold_Storage_Temp_Data$Temperature, na.rm = TRUE)
print(CS_mean)
```

```
## [1] 3.002466
```

Find Standard Deviation for the full year

```
CS_sd <- sd(Cold_Storage_Temp_Data$Temperature, na.rm = TRUE)
print(CS_sd)
```

```
## [1] 0.4658319
```

Assume Normal distribution, what is the probability of temperature having fallen below 2 C?

```
p_below <- pnorm(2, mean = CS_mean, sd = CS_sd, lower.tail = TRUE)
print(p_below)
```

```
## [1] 0.01569906
```

Assume Normal distribution, what is the probability of temperature having gone above 4 C?

```
p_above <- pnorm(4, mean = CS_mean, sd = CS_sd, lower.tail = FALSE)
print(p_above)
```

```
## [1] 0.01612075
```

What will be the penalty for the AMC Company?

```
print(p_above + p_below)
```

```
## [1] 0.03181981
```

Penalty for AMC would be 5% of AMC (annual maintenance case) since the probability of temperature going outside the 2 - 4 C during the one-year contract is 0.0318198 which is above 2.5% and less than 5%.

Perform a one-way ANOVA test to determine if there is a significant difference in Cold Storage temperature between rainy, summer and winter seasons and comment on the findings.

```
CS_anova <- aov(Temperature~Season, Cold_Storage_Temp_Data)
summary(CS_anova)
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
```

```
## Season        2   9.70   4.848   25.32 5.08e-11 ***
## Residuals   362  69.29   0.191
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

P value for one way Anova test is equal to 5.08e-11 which means that with 99% confidence we can reject the null hyphotesis that Cold storage Temperatures are the same between seasons.

Problem 2

In Mar 2018, Cold Storage started getting complaints from their clients that they have been getting complaints from end consumers of the dairy products going sour and often smelling. On getting these complaints, the supervisor pulls out data of last 35 days' temperatures. As a safety measure, the Supervisor decides to be vigilant to maintain the temperature at 3.9 C or below.

Assume 3.9 C as the upper acceptable value for mean temperature and at alpha = 0.1. Do you feel that there is a need for some corrective action in the Cold Storage Plant or is it that the problem is from the procurement side from where Cold Storage is getting the Dairy Products? The data of the last 35 days is in "Cold_Storage_Mar2018.csv"

Which Hypothesis test shall be performed to check the if corrective action is needed at the cold storage plant? Justify your answer.

Null Hyphotesis: the mean temperature is equal to 3.9 C with 90% confidence. Alternate Hyphotesis: the mean temperature is greater than 3.9 C with 90% confidence.

Null Hyphotesis always depicts the assumption or "known situation". It is the basis we want to disprove or reject.

Perform hypothesis test and determine p-value

H0: E[Temperature] = 3.9 HA: E[Temperature] > 3.9

```
CDM_mean <- mean(Cold_Storage_Mar2018$Temperature, na.rm = TRUE)
print(CDM_mean)
```

```
## [1] 3.974286
```

```
CDM_sd <- sd(Cold_Storage_Mar2018$Temperature, na.rm = TRUE)
print(CDM_sd)
```

```
## [1] 0.159674
```

```
t_stat = (CDM_mean - 3.9)/(CDM_sd/sqrt(length(Cold_Storage_Mar2018)))
print(t_stat)
```

```
## [1] 0.930467
```

```
1-pt(t_stat,df=length(Cold_Storage_Mar2018)-1)
```

```
## [1] 0.2103822
```

```
t_test <- t.test(Cold_Storage_Mar2018$Temperature,mu=3.9,alternative='greater')
```

the p value is 0.0047112 which is lower than 10%.

Give your inference

Doing the t tests allows to compare the means observed in March with the mean established for H0. Since the p value is less than 10% we can regect H0 which states that both means are equal and thus there is a need for some corrective action in the Cold Storage Plant.