TEXAS McCombs
The University of Texas at Austin
McCombs School of Business

GREAT LAKES
EXECUTIVE LEARNING

GROUP ASSIGNMENT

Statistical Methods for Decision Making

Post-graduation in Business Analytics and Business Intelligence
Group -IV

By

Adith Bharath Ram
Jithesh Janardhanan
Tarique Zaman
Sarayu Pramod Kodihalli
Nishitha Ramesh

Table of Contents

1. Objective

Cold Storage has a business of storing Pasteurized Fresh Whole or Skimmed Milk, Sweet Cream, Flavored Milk Drinks. We need to analyze the given temperature data and make a hypothesis, and if any corrective action needs to be taken.

2. Problem 1 – Statement

Cold Storage started its operations in Jan 2016. They are in the business of storing Pasteurized Fresh Whole or Skimmed Milk, Sweet Cream, Flavored 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"

1. Find mean cold storage temperature for Summer, Winter and Rainy Season (3 marks)
2. Find overall mean for the full year (3marks)
3. Find Standard Deviation for the full year (3 marks)
4. Assume Normal distribution, what is the probability of temperature having fallen below 2 C? (3 marks)
5. Assume Normal distribution, what is the probability of temperature having gone above 4 C? (3 marks)
6. What will be the penalty for the AMC Company? (3 marks)

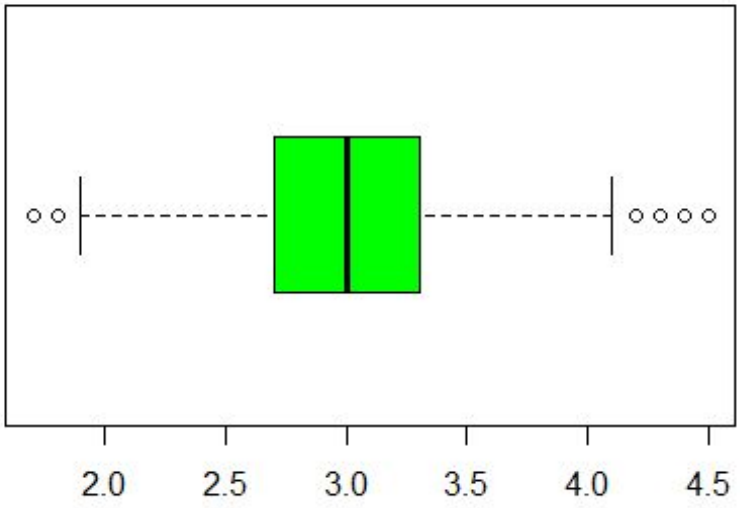
Basic descriptive analysis

```
setwd("C:\\")
TempData =
read.csv("C:\\Users\\aram4\\Downloads\\Cold_Storage_Temp_Data.csv")
View(TempData)
summary(TempData)
```

```
##     Season        Month          Date        Temperature
##  Rainy :122   Aug    : 31   Min.   : 1.00   Min.   :1.700
##  Summer:120   Dec    : 31   1st Qu.: 8.00   1st Qu.:2.700
##  Winter:123   Jan    : 31   Median :16.00   Median :3.000
##               Jul    : 31   Mean   :15.72   Mean   :3.002
##               Mar    : 31   3rd Qu.:23.00   3rd Qu.:3.300
##               May    : 31   Max.   :31.00   Max.   :4.500
##               (Other):179
```

To Analyze data and find the Out-layers.

```
boxplot(TempData$Temperature,horizontal = TRUE,col = "Green")
```



2.1. Q1 – Solution

Mean cold storage temperature for Summer: **3.1475**

Mean cold storage temperature for Winter: **2.7764**

Mean cold storage temperature for Rainy: **3.0877**

Figure 1 Mean of Summer, Winter & Rainy

```
by(TempData,INDICES = TempData$Season,FUN = summary)
```

```
## TempData$Season: Rainy
##     Season        Month         Date        Temperature
##  Rainy :122   Aug    :31   Min.   : 1.00   Min.   :1.700
##  Summer:  0   Jul    :31   1st Qu.: 8.00   1st Qu.:2.700
##  Winter:  0   Jun    :30   Median :16.00   Median :3.050
##               Sep    :30   Mean   :15.75   Mean   :3.088
##               Apr    : 0   3rd Qu.:23.00   3rd Qu.:3.400
##               Dec    : 0   Max.   :31.00   Max.   :4.500
##               (Other): 0
## ------------------------------------------------------------
## TempData$Season: Summer
##     Season        Month         Date        Temperature
##  Rainy :  0   Mar    :31   Min.   : 1.00   Min.   :2.500
##  Summer:120   May    :31   1st Qu.: 8.00   1st Qu.:2.900
##  Winter:  0   Apr    :30   Median :15.50   Median :3.200
##               Feb    :28   Mean   :15.53   Mean   :3.147
##               Aug    : 0   3rd Qu.:23.00   3rd Qu.:3.400
##               Dec    : 0   Max.   :31.00   Max.   :4.000
##               (Other): 0
## ------------------------------------------------------------
## TempData$Season: Winter
##     Season        Month         Date        Temperature
##  Rainy :  0   Dec    :31   Min.   : 1.00   Min.   :1.800
##  Summer:  0   Jan    :31   1st Qu.: 8.00   1st Qu.:2.450
##  Winter:123   Oct    :31   Median :16.00   Median :2.800
##               Nov    :30   Mean   :15.88   Mean   :2.776
##               Apr    : 0   3rd Qu.:23.50   3rd Qu.:3.000
##               Aug    : 0   Max.   :31.00   Max.   :3.900
##               (Other): 0
```

2.2. Q2 – Solution

Mean for the full year = **3.002466**

```
mean(TempData$Temperature)
```

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

2.3. Q3 – Solution

Standard Deviation for the full year = **0.4658319**

```
sd(TempData$Temperature)
```

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

2.4. Q4 – Solution

The probability of temperature having fallen below 2 C: **0.01569904**

(0.01569904 *100) = 1.57%

```
pnorm(2, 3.002466, 0.4658319)
```

```
## [1] 0.01569904
```

2.5. Q5 – Solution

The probability of temperature having gone above 4 C: **0.01612076**

(0.01612076*100) = 1.61%

```
pnorm(4, 3.002466, 0.4658319, lower.tail = FALSE)
```

```
## [1] 0.01612076
```

2.6. Q6 – Solution

The probability of temperature going outside the 2 - 4 C during the one-year contract:

0.031819

(0.031819*100) = 3.18%

Since the probability of temperature = **3.18%** which is above 2% and below 5%, the penalty will be 10% of the AMC.

```
pnorm(2, 3.002466, 0.4658319) + pnorm(4, 3.002466, 0.4658319, lower.tail =
FALSE)
```

```
## [1] 0.0318198
```

3. Problem 2 – Statement

In Mar 2018, Cold Storage started getting complaints from their Clients that they have been getting complaints from end consumers of 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 3.9 C or below. Assume 3.9 C as 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 procurement side from where Cold Storage is getting the Dairy Products. The data of the last 35 days
is in "Cold_Storage_Mar2018.csv"
1. Which Hypothesis test shall be performed to check the if corrective action is needed at the cold storage plant? Justify your answer. (7 marks)
2. State the Hypothesis, perform hypothesis test and determine p-value (7 marks)
3. Give your inference (8 marks)

3.1. Q1 – Solution

A one sample t-test is used for hypothesis testing as the population standard deviation is not given and the population is assumed to be normally distributed.

3.2. Q2 – Solution

Step1: n=35 (data given in csv Cold_Storage_Mar2018.csv)
Xbar = 3.974
Mu = 3.9
S = 0.159
α = 0.1

```
ColdStorage =
read.csv("C:\\Users\\aram4\\Downloads\\K2_Cold_Storage_Mar2018.csv")
summary(ColdStorage$Temperature)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   3.800   3.900   3.900   3.974   4.100   4.600
```

```
Xbar = mean(ColdStorage$Temperature)
Mu = 3.9
S = sd(ColdStorage$Temperature)
 = 0.1
n= length(ColdStorage$Temperature)

Xbar
```

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

```
Mu
```

```
## [1] 3.9
```

```
S
```

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

```
n
```

```
## [1] 35
```

Step2: Formulate Hypothesis:
H0: Mu<=3.9

H1: Mu>3.9

Step 3: Define test statistic

tstat = (Xbar -Mu)/(S/sqrt(n))

```
tstat  = (Xbar-Mu)/(S/sqrt(n))
tstat
```

```
## [1] 2.752359
```

Step 4: Draw diagram

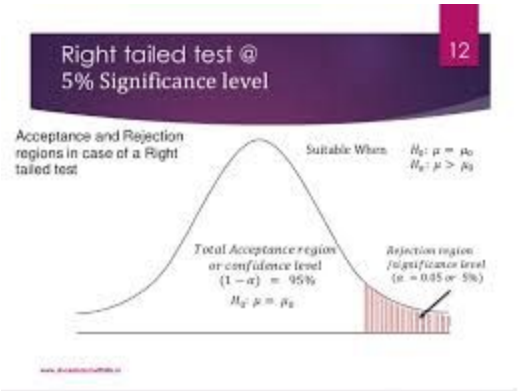


Figure Source

Step 5: (critical value approach)
Determine critical values
α = 10% = 0.1. This is one tailed test (right-tailed test).
df = degree of freedom = n-1 = 35-1=34

```
qt(0.1, df = 34, lower.tail=FALSE)
```

```
## [1] 1.306952
```

Step 6: (critical value approach)
Comparing test statistic value with critical value to make a decision.
In this case since tstat (2.75) greater than the critical value (1.30), the test statistic is falling in the rejection zone and hence we reject the null hypothesis and accept the alternate hypothesis.

P-Value approach

Step 5b (p-value approach):
• Find p-value

```
pvalue= pt(tstat,df=n-1,lower.tail = FALSE)
pvalue
```

```
## [1] 0.004711198
```

Step 6b (p-value approach):
Since p-value < α, therefore we reject the null hypothesis.

The test can also be performed using the following command.

```
t.test(ColdStorage$Temperature,Mu=3.9)
```

```
##
##  One Sample t-test
##
## data:  ColdStorage$Temperature
## t = 147.25, df = 34, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  3.919436 4.029136
## sample estimates:
## mean of x
##  3.974286
```

3.3. Q3 – Solution

Inference: It's statistically proven from the given data that cold storage is not able to keep the mean temperature below 3.9 C standard at 99% confidence interval. Corrective actions need to be taken to:

1. Reduce the customer complaints
2. To avoid the AMC penalties