PROJECT-1

Course: Mathematical and Statistical Methods in Chemical Engineering

PROBLEM-1:

Source: Having prior knowledge about Alzhemer's disease motivated us to do the statistical analysis on NAA and Mi concentration levels in Cingulate gyrus, which are crucial in diagnosing the Alzhemer's diseace. We went to the Nanavati hospital and got data from them.

Problem statement: Statistical analysis of NAA and Mi concentration levels in Cingulate gyrus

Introduction: The subject data is obtained from Nanavati hospital. The concentrations of Sodium Acetyl Aspartate (NAA) and Myoinositol (Mi) of the subjects have been obtained from the cingulate gyrus by performing Magnetic Resonance Spectroscopy. 1cm*1cm*1cm voxel has been selected across the grey matter of occipital region and PRESS (Point Resolved Spectroscopy), Single Voxel Spectroscopy technique has been utilized on 1.5T MRI machine at Nanavati. The subjects have been diagnosed for Alzheimer's disease on the basis of the concentration levels of NAA and MI from the spectrum obtained via PRESS.

Note: 1. when the concentration of NAA goes below the normal level and when the concentration of Mi goes beyond the normal level, the subject is diagnosed to be of Alzheimer's type.

2. We have considered only non-Alzheimer's type subjects. Above 65 years of age, the NAA and Mi levels of the normal subjects decrease.

Data: We considered Population X : < 65 years of age and

Population Y: >above 65 years of age.

Assumption: NAA and Mi levels of population X and population Y follow Normal distribution.

Sample from Population X				**Sample from Population Y**			
Sr. No.	**Age**	**NAA concentration**	**Mi concentration**	**Sr. No.**	**Age**	**NAA concentration**	**Mi concentration**
1.	58	1.53	0.72	1.	66	1.13	0.62
2.	53	1.43	0.68	2.	80	1.09	0.56
3.	49	1.45	0.62	3.	66	1.30	0.63
4.	47	1.47	0.72	4.	69	1.33	0.62
5.	59	1.27	0.67	5.	74	1.33	0.67
6.	59	1.43	0.70	6.	81	1.21	0.62
7.	52	1.24	0.76	7.	81	1.34	0.62
8.	56	1.48	0.57	8.	81	1.34	0.79
9.	62	1.41	0.66	9.	67	1.52	0.61
10.	45	1.5	0.67	10.	66	1.43	0.7
11.	48	1.32	0.68				

Question 1) Study of variance of NAA levels of Population X and Population Y

- **Null Hypothesis H_0**:

(Variance of NAA levels of Population X) ^{2}= (Variance of NAA levels of Population Y) ^{2}

- **Alternate Hypothesis H_1**:

(Variance of NAA levels of Population X) ^{2}/ (Variance of NAA levels of Population Y) $^2 \neq 1$

Significance value (α): 0.05 (5%)

Analysis: Using the statistical software R, we find the following:
Test Statistic: F=0.4853
Numerator dof=10
Denominator dof=9
p-value = 0.2753.
95% Confidence Interval (C.I.)= [0.1224257 1.8338515]
Ratio of Variances=0.4852791

Conclusion: Since, p-value >α, we fail to reject H0. The variance of NAA levels of population X could be equal to the variance of NAA levels of population Y.

Question 2) Study of variance of Mi levels of Population X and Population Y

- **Null Hypothesis H_0**:

(Variance of Mi levels of Population X)^2= (Variance of Mi levels of Population Y)^2

- **Alternate Hypothesis H_1**:

(Variance of Mi levels of Population X)^2/(Variance of Mi levels of Population Y)$^2 \neq 1$

Significance value (α): 0.05 (5%)

Analysis: Using the statistical software R, we find the following:
Test Statistic: F=0.6584
Numerator dof=10
Denominator dof=9
p-value = 0.5234.
95% Confidence Interval (C.I.) = [0.1660956 2.487996]
Ratio of Variances=0.6583807
Conclusion: Since, p-value >α, we fail to reject H0. The variance of Mi levels of population X could be equal to the variance of Mi levels of population Y.

Question 3) Study of Mean of NAA levels of Population X and Population Y

- **Null Hypothesis H_0**:
  Mean of NAA levels of Population X= Mean of NAA levels of Population Y
- **Alternate Hypothesis H_1**:
  Mean of NAA levels of Population X is greater than Mean of NAA levels of Population Y

Significance value (α): 0.05 (5%)
We considered variance of NAA levels of population of X and population Y are equal.
Analysis: Using the statistical software R, we find the following:
Test Statistic: t=2.1693
Dof: 19
p-value = 0.02148
95% Confidence Interval (C.I.) = [0.02134239 ∞]
Mean of NAA levels of Sample X =1.408182
Mean of NAA levels of Sample Y =1.3030

Conclusion: Since, p-value<α, we reject H0. The Mean of NAA levels of population X are greater than the mean of NAA levels of population Y. Therefore, we conclude that the concentration of NAA levels would decrease after the age of 65 years.

Question 4) Study of Mean of Mi levels of Population X and Population Y

- **Null Hypothesis H0**:

  Mean of Mi levels of Population X= Mean of Mi levels of Population Y

- **Alternate Hypothesis H1**:

  Mean of Mi levels of Population X is greater than Mean of Mi levels of Population Y
  Significance value (α): 0.05 (5%)
  We considered variance of Mi levels of population of X and population Y are equal.
  Analysis: Using the statistical software R, we find the following:
  Test Statistic: t=1.3325
  Dof: 19
  p-value = 0.09924
  95% Confidence Interval (C.I.) = [-0.0099 ∞]
  Mean of Mi levels of Sample X =0.6772727
  Mean of Mi levels of Sample Y =0.644

  Conclusion: Since, p-value>α, we fail to reject H0 but as per the found p-value at 0.1 significance level, we can reject H0. Still, it's a strong statement to state that the Mean of Mi levels of population X are greater than the mean of Mi levels of population Y. Therefore, we have strong evidence to conclude that the concentration of Mi levels would decrease after the age of 65 years.

  Final Conclusion: Variance of NAA and Mi levels of Population X and Population Y could be equal (It's a weak statement). We have strong evidence that NAA and Mi levels would decrease after the age of 65 years.

PROBLEM-2:

Source: We have taken the following unsolved two factor anova problem (14.3) from Montgomery 5th edition.

Problem statement: The results of an experiment involving a storage battery used in the launching mechanism of a shoulder-fired ground-to-air missile were presented. Three material types can be used to make the battery plates. The objective is to design a battery that is relatively unaffected by the ambient temperature. The output response from the battery is effective life in hours. Three temperature levels are selected and a factorial experiment with four replicates is run.

We are interested in

a) drawing conclusions using anova with the significance value 0.05

b) analysing the interaction graphically

The data are as follows:

	Temperature (°F)					
Material	Low		Medium		High	
1	130	155	34	40	20	70
	74	180	80	75	82	58
2	150	188	136	122	25	70
	159	126	106	115	58	45
3	138	110	174	120	96	104
	168	160	150	139	82	60

a)

$Y_{ijk}= \mu + \Gamma_i + \beta_j + (\Gamma\ \beta)_{ij} + \varepsilon_{ijk}$ (Model)

μ– the overall mean

Γ_i – the effect of material i, i=1, 2, 3.

β_j– the effect of temperature j, j= 1(low),2(medium),3(high)

$(\Gamma\ \beta)_{ij}$– the effect of interaction between material i and temperature j

ε_{ijk}–random error having normal distribution

Analysis:

1. H_0: $\Gamma_1= \Gamma_2= \Gamma_3=0$ (no main effect of material)

   H_1: at least one $\Gamma_i \neq 0$

2. H_0: $\beta_1= \beta_2= \beta_3=0$ (no main effect of temperature)

H_1: at least one $\beta_j \neq 0$

3. H_0: $(\Gamma\ \beta)_{11}=(\Gamma\ \beta)_{11=\ldots=}(\Gamma\ \beta)_{33}=0$ (no interaction)

   H_1: at least one $(\Gamma\ \beta)_{ij} \neq 0$

   After running the code in R, we got the following output (used print screen).

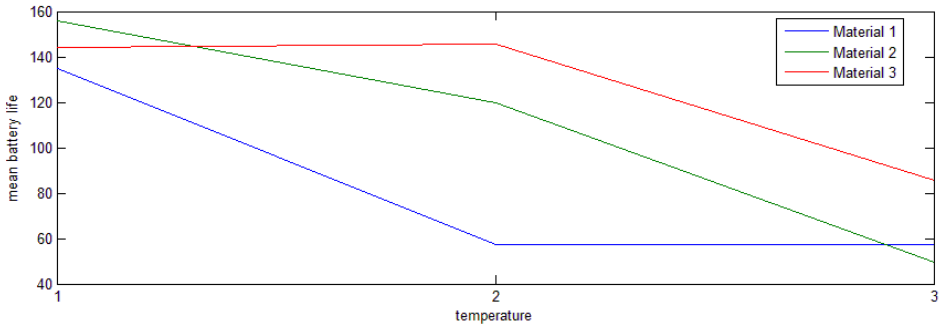
```
Analysis of Variance Table

Response: Y
          Df Sum Sq Mean Sq F value    Pr(>F)
Mat        2  11800  5900.2  6.4286  0.005207 **
Temp       2  40085 20042.7 21.8376 2.282e-06 ***
Mat:Temp   4  12597  3149.3  3.4313  0.021604 *
Residuals 27  24781   917.8
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Mat=Material, Temp=Temperature, Mat:Temp=Interaction between temperature and material

Conclusion:

Since in three cases p-values are less than our significance value 0.05, we reject all Null Hypotheses. From the above result, it is evident that the materials, temperatures and the interaction between materials and temperatures significantly affect the battery life.

b)



Conclusion: the mean battery life of material2 is highest at low temperature and lowest at high temperature. All three materials at low temperature are giving mean life with less variability, but at medium temperature they are giving mean life with more variability. This shows the interaction between temperature and material is strong.