Frequently asked questions.

1. Define Gauss-Markoff Linear Model

Consider a set of n independent random variables y_i whose expectations are given as linear functions of p unknown parameters say $\beta_1, \beta_2, \ldots \beta_p$, ($p \leq n$) with known coefficients a_{ij}'s and whose variances are a constant σ^2.

Then $E(y_i) = a_{i1}\beta_1 + a_{i2}\beta_2 + \ldots a_{ip}\beta_p$, i=1…n (1)

$V(y_i) = \sigma^2$, $Cov(y_i, y_j) = 0$, $i \neq j$.

or

Writing $y = (y_1, y_2, \ldots y_n)^1$; $\beta = (\beta_1, \beta_2, \ldots \beta_p)^1$; $\alpha_j = (a_{1j}, a_{2j}, \ldots a_{nj})^1$, $j = 1 \ldots p (\leq n)$

$A = (\alpha_1, \alpha_2, \ldots \alpha_p)$

Then model (1) can be written as

$E(y) = A\beta$ (2)

$D(y) = \sigma^2 I$ where D(y) denote the dispersion matrix of y and I is identity matrix of order n.

Model (1) or (2) or (3) is known as Gauss Markov Model.

2. Define Estimable function:

A linear function $a^1\beta$ of parameters with known "a" is said to be an estimable parametric function (or estimable) if there exists a linear function $L^1y$of Y such that $E(L^1y) = a^1\beta$ for all β. [where $L = (l_1, l_2, \ldots l_p)^1$; $y = (y_1, y_2, \ldots y_n)^1$, $a = (a_1, a_2, \ldots a_p)^1$, $\beta = (\beta_1, \beta_2, \ldots \beta_p)^1$]

3. State Gauss Markoff Theorem.

If the linear parametric function $L^1\beta$ is estimable then the linear estimator $L^1\hat{\beta}$ where $\hat{\beta}$ is a solution of $X^1X\hat{\beta} = X^1Y$ is the best linear unbiased estimator of $L^1\beta$ in the sense of having minimum variance in the class of all linear and unbiased estimators of $L^1\beta$

4. What is Analysis of Variance (ANOVA)

Analysis of variance involves breaking down of total variation into orthogonal components. Each orthogonal components represents the variation due to a particular factor contributing in the total variation.

5. **What are the assumptions involved in Analysis of Variance** (ANOVA)

1. The observations are independent
2. Parent population from which observation are taken is normal
3. Various treatments(factors) and environmental effects are additive in nature
4. The distributions of the residuals are normal. i.e. random errors are distributed as iid normal with mean 0 and variance σ^2.

6. What are fixed effect models:
If the levels of the factor are fixed, the model considered is called fixed effect models.

7. **Define One way ANOVA Model.**: We define the model

$$Y_{ij} = \mu + \alpha_i + €_{ij},\ i=1,2\ldots p,\ j=1,2,\ldots n_i \qquad (4)$$

Where μ general effect

α_i – i^{th} treatment effect;

ϵ_{ij}- error term which are independent and identically distributed normal random variables with mean 0 and variance σ2

8. **State the hypothesis tested under One way ANOVA Model**

We test the hypotheses

H0: $\alpha_1 = \alpha_2 = \ldots = \alpha_p = 0$(i.e. all the treatment has the same effect on the yield)

Against the alternative

H1: atleast one $\alpha_i \neq \alpha_j$, for all I,j (i.e. the treatment effects are different)

9. **Split the total variation into different component parte in one way anova**

Consider

$$SST = \sum_{i=1}^{p} \sum_{j=1}^{ni} (y_{ij} - \overline{y_{..}})^2 = \sum_{i=1}^{p} \sum_{j=1}^{ni} (y_{ij} - \overline{y_{i..}} + \overline{y_{i..}} - \overline{y_{..}})^2$$

$$= \sum_{i=1}^{p} n_i(\overline{yi.} - \overline{y_{..}})^2 + \sum_{i=1}^{p}\sum_{j=1}^{ni}(yij - \overline{yi_{..}})^2 + 2$$

$$\sum_{i=1}^{p}\sum_{j=1}^{ni}(\overline{y_i.} - \overline{y_{..}})(y_{ij} - \overline{y_{i..}})$$

The third term becomes zero hence we have

$$\sum_{i=1}^{p}\sum_{j=1}^{ni}(y_{ij} - \overline{y_{..}})^2 = \sum_{i=1}^{p} n_i(\overline{yi.} - \overline{y_{..}})^2 + \sum_{i=1}^{p}\sum_{j=1}^{ni}(yij - \overline{yi_{..}})^2$$

SST = SSTR + SSE where SSTR is called the sum of squares due to treatments .and SSE is called the sum of squares due to error.

10. **Write the test statistics used to test the hypothesis in one way anova**.

The test statistics is $F_\alpha = \dfrac{SSTR/(p-1)}{SSE/(N-p)} = \dfrac{MSSTR}{MSSE}$

(8)

is distributed as F with p-1 and N-p degrees of freedom.

11. **Write the Analysis of variance Table(ANOVA Table) in one way ANOVA**

Sources of variation	Degree of freedom	Sum of squares	Mean sum of squares	F-Value
Treatments	p-1	SSTR	MSSTR = SSTR/(p-1)	MSSTR/ MSSE
Error	N-p	SSE	MSSE = SSE/(p-1) (k-1)	
total	N-1	TSS		

Compare the table value of F and compare it with calculated F value.

12. **Write the procedure of critical difference to Compare pairs of treatment means when the hypothesis is rejected**

Case of rejection of H_0.:

if $F_0 > F_{\alpha,,p-1, N-p,}$ then

H_0: $\alpha_1 = \alpha_2 = \ldots = \alpha_p = 0$

is rejected. This means that at least one αi

is different from other effects which is responsible for the rejection of the H0.

This can be done by pairwise testing of αi's.

i.e. we test H_0: αi = αi , (i≠j) against H1: αi≠ αi

This can be tested using the following t-statistics:

T = $\dfrac{\overline{y_{i.}} - \overline{y_{j.}}}{\sqrt{MESS\left(\dfrac{1}{n_i}+\dfrac{1}{n_j}\right)}}$ which follows t distribution with N-p degrees of freedom

under H0.

Thus the decision rule to reject H0 at α if the observed difference

$\overline{y_{i.}} - \overline{y_{j.}} > t_{\alpha,n-p}\sqrt{MESS\left(\dfrac{1}{n_i}+\dfrac{1}{n_j}\right)}$ The quantity $t_{\alpha,n-p}\sqrt{MESS\left(\dfrac{1}{n_i}+\dfrac{1}{n_j}\right)}$ is

called critical difference. Where $t_{\alpha, N-p}$ is the table value referring to t table at

α level significance corresponding to N-p degrees freedom.

13. **Fill the missing values.**

Sources of variation	Degree of freedom	Sum of squares	Mean sum of squares	F-Value
Treatments (foodstuffs)	3	??	8527.717	12.571
Error	??	??	??	
total	21	37793.318		

Solution:

Sources of variation	Degree of freedom	Sum of squares	Mean sum of squares	F-Value
Treatments (foodstuffs)	3	25583.152	8527.717	12.571
Error	18	12210.167	678.343	
total	21	37793.318		