Statistics Department
For Master student /Lesson (4)

Analysis of Experimental Design ((ONE WAY ANOVA (CRD))

Assnt.prof.
Dr. Suhad Ali Shaheed

2020/2019

Chapter two CH-2

Experiments with a single factor

One-Way classification Analysis of variance

The Completely Randomized Design.

Definitions.

Single factor (CRD) is an experiment design involving one factor with (P) levels (treatment) the treatment are assigned to the experimental units with no restriction on the assignment expect that each treatment will be used a specified number of times.

Thus for the CRD with p treatments and n_i experimental units for the ith treatment the model is as follows :-

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \qquad \begin{array}{l} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, n_i \end{array}$$

where

y_{ij} : is the response to the jth experimental unit of the ith treatment.

μ : is the overall average response to all treatments.

τ_i : is the effect of the ith treatment

- parameter $\sum_{i=1}^{p} \hat{\tau}_i = 0$ (fixed Model)
- random variables $\tau_i \sim NID(0, \sigma_\tau^2)$

(Random Model)

ϵ_{ij} : is the random error associated with the jth experimental unit of the ith treatment, we assume that $E(\epsilon_{ij}) = 0$ and $var(\epsilon_{ij}) = \sigma^2$ $\forall i$ and j. further more we assume the ϵ_{ij} are uncorrelated.

- Fixed Models-

with this model for the observation (Y_{ij}) we have

$$E(Y_{ij}) = \mu + \tau_i \quad \text{and} \quad var(Y_{ij}) = \sigma^2$$

One of our primary objectives in the Analysis of a (CRD) will be to compare the means of (P) treatments for example the hypothesis $H_0: \tau_1 = \tau_2 = \tau_3 = \text{---} = \tau_p$ which states that the (p) treatments means are equal will generally be of major interest to us.

In general previous model can be written in matrix notation by letting.

$$y = \begin{bmatrix} y_{11} \\ \vdots \\ y_{1n_1} \\ y_{21} \\ \vdots \\ y_{2n_2} \\ \vdots \\ y_{p1} \\ \vdots \\ y_{pn_p} \end{bmatrix}, \quad X = \begin{array}{l} \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & \cdots & 1 \end{bmatrix} \end{array} \begin{array}{l} \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} n_1 \text{ rows} \\ \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} n_2 \text{ rows} \\ \vdots \\ \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} n_p \text{ rows} \end{array}$$

$$\beta = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ \vdots \\ \tau_p \end{bmatrix} \qquad \epsilon = \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \vdots \\ \epsilon_{1n_1} \\ \epsilon_{21} \\ \epsilon_{22} \\ \vdots \\ \epsilon_{2n_2} \\ \vdots \\ \epsilon_{p1} \\ \vdots \\ \epsilon_{pn_p} \end{bmatrix}$$

In these representation Y and ϵ are $n\times 1$ random vectors, β is a $(P+1)\times 1$ vector of unknown constants (parameter) and X is a $(n\times(P+1))$ matrix of specified constants (0 or 1), our CRD model is now seen to be a Linear Model.

, $Y = X\beta + \epsilon$ with assumptions

$E(\epsilon)=0$ and $cov(\epsilon) = \sigma^2 I$

In the matrix (X) called the design matrix the sum of the last (P) columns equal the first columns, Therefore X is not full Rank (less than full rank)

ماهي تداعيات مصفوفة X عندما تكون Less than Full Rank

- Least squares Results :-

To find least squares estimators, we first find a solution to the normal equations for a CRD $X'Xb = X'Y$

$$\begin{bmatrix} n & n_1 & n_2 & \cdots & n_p \\ n_1 & n_1 & 0 & & 0 \\ n_2 & 0 & n_2 & & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ n_p & 0 & 0 & & n_p \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\tau}_2 \\ \vdots \\ \hat{\tau}_p \end{bmatrix} = \begin{bmatrix} Y_{..} \\ Y_{1.} \\ Y_{2.} \\ \vdots \\ Y_{p.} \end{bmatrix}$$

where $y_{..} = \sum_{i=1}^{p} \sum_{j=1}^{n_i} y_{ij}$ and $y_{i.} = \sum_{j=1}^{n_i} y_{ij}$

These normal equations yield the $p+1$ equation

$$
\begin{aligned}
n\hat{\mu} + n_1\hat{\tau}_1 + n_2\hat{\tau}_2 + \cdots + n_p\hat{\tau}_p &= y_{..} \\
n_1\hat{\mu} + n_1\hat{\tau}_1 \qquad\qquad &= y_{1.} \\
n_2\hat{\mu} \qquad + n_2\hat{\tau}_2 \qquad &= y_{2.} \\
\vdots \qquad\qquad\qquad & \\
n_p\hat{\mu} \qquad\qquad + n_p\hat{\tau}_p &= y_{p.}
\end{aligned}
$$

if we impose the side condition $\sum_{i=1}^{p} n_i\hat{\tau}_i = 0$

on the system we obtian the solution set

$\hat{\mu} = \frac{y_{..}}{n} = \bar{y}_{..}$

$\hat{\tau}_i = \frac{y_{i.}}{n_i} - \hat{\mu} = \bar{y}_{i.} - \bar{y}_{..} \qquad i = 1, 2, \cdots, p$

Note that our conditions $\sum_{i=1}^{p} n_i\hat{\tau}_i = 0$

is of the form $Hb = 0$ where H is

the $1 \times (p+1)$ vector $(0 \; n_1 \; n_2 \; \cdots \; n_p)$

Next we find the estimable functions for the CRD. using [in the linear model $y = X\beta + \epsilon$] with $E(\epsilon) = 0$. $X\beta$ is aset of estimable functions using X and β for a CRD

we find that the distinct entires in $X\beta$ are $\mu+\tau_i$.
Hence $\mu+\tau_i$ is estimable for $i=1,2,\dots,p$
obviously $\mu, \tau_1, \tau_2, \dots, \tau_p$ are not individually estimable **, i.e none of the parameters in the CRD model is estimable, Important function like the treatment means $(\mu+\tau_i)$ and the differenc between treatment means are estimable.

$$\mu+\tau_1 - (\mu+\tau_2) = \tau_1-\tau_2 \quad \text{are estimable.}$$

Def.
The function $\sum_{i=1}^{p} c_i\tau_i$ is called treatment contrast if

$$\sum_{i=1}^{p} c_i = 0$$

*** all treatments contrasts are estiamble
BLUE for The CRD Model.

function	BLUE
$\mu+\tau_i$	$\bar{y}_{i.}$
$\sum_{i=1}^{p} c_i\tau_i$ wher $\sum_{i=1}^{p} c_i = 0$	$\sum_{i=1}^{p} c_i\bar{y}_{i.}$

we find that

$$\hat{y}_{ij} = \hat{\mu}+\hat{\tau}_i = \bar{y}_{i.} \qquad i=1,2,\dots,p \quad j=1,2,\dots,n_i$$

where

$$\hat{y}_{ij} = \hat{\mu} + \hat{\tau}_i = \bar{y}_{..} + \bar{y}_{i.} - \bar{y}_{..} = \bar{y}_{i.}$$

$$sse = \sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \hat{y}_{ij})^2$$

$$S^2 = \frac{sse}{n-p}$$

for CRD

$$S^2 = \sum_{i=1}^{p} \sum_{j=1}^{n_i} \frac{(y_{ij} - \bar{y}_{i.})^2}{n-p}$$

H.W homework.

Consider the daphnia life time data find BLUE for the $\mu + \tau_3$, $\tau_1 - \tau_2$ and S^2.

- Reparametrization :-

In regression to treatment of one way classification the following defination will be used

τ_i The parameter of the ith level of the factor

$\mu + \tau_i$ The mean of the ith level of the

-Analysis of Variance and F-test (ANOVA):-

To derive the test statistic for

$H_o: \tau_1 = \tau_2 = \cdots = \tau_p$

The null hypothesis H_o can be represented in the form of the general Linear hypothesis

$H_o: L\beta = 0$

$$\begin{bmatrix} 0 & -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 0 & 1 & 0 & & 0 \\ 0 & -1 & 0 & 0 & 1 & & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & -1 & 0 & 0 & 0 & & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ \vdots \\ \tau_p \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

where $L_{(p-1)(p+1)}$ of rank $(p-1)$

$L\beta$ a set of $(p-1)$ Linearly estimable fun.

$$L\beta = \begin{matrix} \tau_2 - \tau_1 \\ \tau_3 - \tau_1 \\ \vdots \end{matrix}$$

H_o $\mu + \tau_i$ are all equal so we can set $\mu^* = \mu + \tau_i$ for $i = 1, 2, \cdots, p$

The reduced Model becomes انموذج مختزل

$$Y_{ij} = \mu^* + \epsilon_{ij}^* \quad , i = 1, 2, \cdots, p \quad j = 1, 2, \cdots, n_i$$

The least squares estimate for μ^* is $\bar{y}_{..}$ so that

$$\hat{y}_{ij} = \bar{y}_{..}$$

As a result, the error sum of squares of the full model is.

$$SSe = \sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2$$

and for the reduced model is

$$SSe^* = \sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2$$

Now

$$SSe^* - SSe = \sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2 - \sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2$$

$$= \sum_{i=1}^{p} n_i (\bar{y}_{i.} - \bar{y}_{..})^2 = SStr$$

we obtian the test statistics

$$F = \frac{(SSe^* - SSe)/q}{S^2} = \frac{\sum_{i=1}^{p} n_i (\bar{y}_{i.} - \bar{y}_{..})^2 / p-1}{SSe / n-p}$$

$$F = \frac{MStr}{MSe} \Rightarrow \text{we reject } H_0 \text{ if } F > f_{\alpha (p-1)(n-p)}$$

ANOVA Table for CRD

S.O.V	d.f	S.S	M.S	F
treatments	$p-1$	$\sum_{i=1}^{p} n_i(\bar{y}_{i.}-\bar{y}_{..})^2 = SStr$	$MStr = SStr/p-1$	$\frac{MStr}{S^2}$
Error	$n-p$	$\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i.})^2 = SSe$	$S^2 = MSe = SSe/_{n-p}$	
Total	$n-1$	$\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{..})^2 = SST$		

-Confidence Intervals and Tests:-

It is evident that to find confidence interval and perform t-test we need the least Square estimates and the standard errors for the estimable functions. Now from the elementary statistics we know that

$$var(\bar{y}_{i.}) = \sigma^2/n_i \Rightarrow Se(\bar{y}_{i.}) = \sqrt{S^2/n_i}$$

$$var(\bar{y}_{i.}-\bar{y}_{k.}) = \frac{\sigma^2}{n_i} + \frac{\sigma^2}{n_k} = \sigma^2\left(\frac{1}{n_i}+\frac{1}{n_k}\right)$$

$$\left[Se(\bar{y}_{i.}-\bar{y}_{k.}) \quad \sqrt{S^2\left(\frac{1}{n_i}+\frac{1}{n_k}\right)}\right]$$

Finally

$$var\left(\sum_{i=1}^{p} c_i\bar{y}_{i.}\right) = \sum_{i=1}^{p} c_i^2\, var(\bar{y}_{i.})$$

$$= \sigma^2 \sum_{i=1}^{p} \frac{c_i^2}{n_i}$$

$$\left[Se\left(\sum_{i=1}^{p} c_i \bar{y}_{i\cdot}\right) = \sqrt{S^2 \sum_{i=1}^{p} \frac{c_i^2}{n_i}} \right]$$

- $100(1-\alpha)\%$ confidence Interval formulas for CRD :-

function	Interval
mean : $\mu + \tau_i$	$\bar{y}_{i\cdot} \mp t_{\alpha/2, n-p} \sqrt{S^2/n_i}$
Difference : $\tau_i - \tau_k$	$(\bar{y}_{i\cdot} - \bar{y}_{k\cdot}) \mp t_{\alpha/2, n-p} \sqrt{S^2\left(\frac{1}{n_i} + \frac{1}{n_k}\right)}$
contrast : $\sum_{i=1}^{p} c_i \tau_i$, $\sum_{i=1}^{p} c_i = 0$	$\sum_{i=1}^{p} c_i \bar{y}_{i\cdot} \mp t_{\alpha/2, n-p} \sqrt{S^2 \sum_{i=1}^{p} c_i^2/n_i}$

- Test statistics Associated with a CRD :-

null hypothesis	Test statistics
$H_0 : \mu + \tau_i = \gamma$	$t = (\bar{y}_{i\cdot} - \gamma) / \sqrt{S^2/n_i}$
$H_0 : \tau_i - \tau_k = \gamma$	$t = (\bar{y}_{i\cdot} - \bar{y}_{k\cdot} - \gamma) / \sqrt{S^2\left(\frac{1}{n_i} + \frac{1}{n_k}\right)}$
$H_0 : \sum_{i=1}^{p} c_i \tau_i = \gamma$, $\sum_{i=1}^{p} c_i = 0$	$t = \left(\sum_{i=1}^{p} c_i \bar{y}_{i\cdot} - \gamma\right) / \sqrt{S^2 \sum_{i=1}^{p} \frac{c_i^2}{n_i}}$

- Expected Mean Squares :-

Theorm ①

Consider the linear model $Y = X\beta + \epsilon$ where ϵ is $MN(0, \sigma^2 I)$, let SS be the sum of squars in the ANOVA table with q degress of fredom. The corresponding expected mean square, denoted by $E(MS)$, is given by:

$$E(MS) = \sigma^2 + \frac{1}{q} W$$

where, W is the expression for SS when the Y's are replaced by their expected values.

Expected Mean Squares in CRD Analysis of variance, is given by

$$E(MSE) = \sigma^2$$

we utilize theorem ① to find the expected mean square for treatments $E(MSTR)$ from

$$SSTR = \sum_{i=1}^{p} n_i (\bar{Y}_{i.} - \bar{Y}_{..})^2$$

So

$$W = \sum_{i=1}^{p} n_i \left[E(\bar{Y}_{i.}) - E(\bar{Y}_{..}) \right]^2$$

$$E(\bar{Y}_{i.}) = \mu + \tau_i$$

$$E(\bar{Y}_{..}) = \mu + \bar{\tau} \qquad \text{where } \bar{\tau} = \sum_{i=1}^{p} n_i \tau_i / n$$

Then

$$W = \sum_{i=1}^{p} n_i \left[(\mu + \tau_i) - (\mu + \bar{\tau}) \right]^2$$

$$= \sum_{i=1}^{p} n_i (\tau_i - \bar{\tau})^2$$

This gives us

$$E(MSTR) = \sigma^2 + \frac{1}{p-1} \sum_{i=1}^{p} n_i (\tau_i - \bar{\tau})^2$$

At times it is advantageous to extend the ANOVA table to include an expected mean square column

S.O.V	d.f	S.S	M.S	E(MS)
treatment	$p-1$	SStr	MSTR	$\sigma^2 + \frac{1}{p-1} \sum_{i=1}^{p} n_i (\tau_i - \bar{\tau})^2$
Error	$n-p$	SSe	MSE	σ^2
Total	$n-1$	SSt		

when $H_0: \tau_1 = \tau_2 = \cdots = \tau_p$ is true $(\tau_i - \bar{\tau} = 0)$ and then

$$\sum_{i=1}^{p} n_i (\tau_i - \bar{\tau})^2 = 0$$

Thus when H_0 is true $E(MSTR) = \sigma^2 = E(MSE)$

when H_0 is false $\sum_{i=1}^{p} n_i (\tau_i - \bar{\tau})^2 > 0$

resulting in $E(MSTR) > E(MSE)$

That is whenever an F random variable is used to test Ho, under Ho the ratio of the expected values of the numerator and denominator of F is one, This ratio is greater than on Ho is not true.

$$E(F) = \frac{E(MSTR)}{E(MSE)} = \frac{\sigma^2 + \frac{1}{p-1}\sum_{i=1}^{p} n_i(\tau_i - \bar{\tau})^2}{\sigma^2}$$

$$= 1 + \boxed{\frac{1}{p-1} \frac{\sum_{i=1}^{p} n_i(\tau_i - \bar{\tau})^2}{\sigma^2}}$$

Non-Central Term

i.e F — central under Ho

— Non Central under H_1

نهاية الجزء الاول يتبع لطفا

الجزء الثاني

The Restricted Model:

In Some textbooks the model given for CRD include certain restrictions on the parameters of the model. The most common restriction used is $\sum_{i=1}^{p} n_i \tau_i = 0$. When this restriction is included as a part of the model we refer to the model as the restricted model.

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij} \quad ; i = 1, 2, \ldots, p \; ; j = 1, 2, \ldots, n_i$$

$$\sum_{i=1}^{p} n_i \tau_i = 0$$

use of the restricted Model not alter the Analysi we have described for the unrestricted Model (we call the traditional Model the unrestricted CRD model) if we refer back, we see that the side condition used to obtian a unique solution to the Normal equations is $\sum_{i=1}^{p} n_i \hat{\tau}_i = 0$.

As a result when that restriction becomes part of the Model ① all parameters are individually estimable we see that the BLUE for μ is $\bar{y}_{..}$ and the BLUE fo τ_i is $\bar{y}_{i.} - \bar{y}_{..}$ (from the unrestricted CRD Model $\mu + \tau_i$ is estimable but μ and τ_i separately are not estimable)

② The ANOVA null hypothesis reduced to $H_0 : \tau_i = 0$ for all i for the restricted model in words Ho is "no treatment effects" ($H_0 : \tau_1 = \tau_2 = \cdots = \tau_p$ associeted with the unrestricted model in words Ho is "no difference in treatment effects")

③ for the restricted model the formula for $E(MSTR)$ becomes $\sigma^2 + \frac{1}{p-1} \sum_{i=1}^{p} n_i \tau_i^2$

* show the propritIes of the estimator

Note: when all treatment sample size are the same $n_1 = n_2 = \cdots = n_p$, this situation is sometimes referred to a (balanced) of the an experimental design, in general the computations required in analyzing balanced experimental design data are easily Carried out.

- A Random effect Model :-

• Consider a CRD with (p) treatments and (n_i) experimental units assigned to treatment (i) we assume that the treatment factor is Random that is the (p) levels are randomly chosen from a Large population since the treatment factor is random the treatment effects (τ_i) will be r.v.

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij} \qquad \begin{array}{l} i = 1, 2, \dots, p \\ j = 1, 2, \dots, n_i \end{array}$$

where the τ_i and ϵ_{ij} are uncorrelated r.v's

$E(\tau_i) = 0$, $var(\tau_i) = \sigma_\tau^2$ $\forall i$

$E(\epsilon_{ij}) = 0$, $var(\epsilon_{ij}) = \sigma^2$ $\forall i, j$

To perform infrence we will require the normality assumptions, which can be written as (τ_i) are independent $N(0, \sigma^2)$ and the τ_i and ϵ_{ij} are independent with these normality assumption we note that $Y_{ij} \sim N(\mu, \sigma_\tau^2 + \sigma^2)$.

We also note that in random effect model some observation are dependent while others might be independent (observation within a treatment are depended while observation in different treatment are independent).

That is

$cov(y_{ij}, y_{ik}) \neq 0 \quad$ for $j \neq k$

and

$cov(y_{ij}, y_{km}) = 0 \quad$ for $i \neq k$

Random effects models in general differ basically from fixed effects model in that under the underlying assumptions.

1. All the observation have the Same expectation $E(y_{ij}) = \mu \quad \forall i, j$.

2. The observation are not statistically independent.

Infrence that are made involve parameters σ_τ^2 and σ^2, which are sometimes called the variance components we test

$H_0: \sigma_\tau^2 = 0$

if thes hypothesis is true then the variation in the observeation (y_{ij}) is due only to random error term and the treatment have no effect on the response variable if $\hat{\sigma}_\tau^2 > 0$ then treatment have an effect on the observation.

Analysis of variance:-

The total variation can be decompose into two parts (variance components) which are the treatment and error effects (the same analysis for the fixed model)

$$\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{..})^2=\sum_{i=1}^{p}n_i(\bar{y}_{i.}-\bar{y}_{..})^2+\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i.})^2$$

$$SST = SStr + SSe.$$

As before if normalty assumed.

$$\frac{SSe}{\sigma^2}=\frac{(n-p)\,MSe}{\sigma^2}\sim\chi^2_{(n-p)}$$

$$=\frac{\cancel{(n-p)}E(MSe)}{\sigma^2}=\cancel{(n-p)}$$

$$\therefore E(MSe)=\sigma^2 \Rightarrow \frac{E(MSe)}{\sigma^2}=1 \Rightarrow \text{central } \chi^2_{(n-p)}$$

Since when (τ_i) are r.v. the distribution of (sstr) is very complected generally but if $n_i=n$ $\forall i$ more elegint treatment can be given

$$sstr=n\sum_{i=1}^{p}(\bar{y}_{i.}-\bar{y}_{..})^2$$

$$=n\sum_{i=1}^{p}\Big[(\underbrace{\tau_i+\bar{\epsilon}_{i.}}_{z_i})-(\underbrace{\bar{\tau}+\bar{\epsilon}_{..}}_{\bar{z}})\Big]^2$$

$$sstr=n\sum_{i=1}^{p}(z_i-\bar{z})^2 \quad \text{wher } z\sim N(0,\sigma^2_\tau+\frac{\sigma^2}{n})$$

$$\frac{\sum_{i=1}^{p}(z_i-\bar{z})^2}{\sigma^2_\tau+\frac{\sigma^2}{n}}\sim\chi^2_{(p-1)}$$

Thus $$\frac{sstr}{n\left(\sigma_\tau^2 + \frac{\sigma^2}{n}\right)} \sim \chi^2_{(p-1)} \Rightarrow \frac{sstr}{n\sigma_\tau^2 + \sigma^2} \sim \chi^2_{(p-1)}$$

$$sstr \sim \chi^2_{(p-1)}(n\sigma_\tau^2 + \sigma^2)$$

$$Mstr \sim \chi^2_{(p-1)} \cdot \frac{n\sigma_\tau^2 + \sigma^2}{(p-1)}$$

$$E(MStr) = \frac{p-1}{p-1}(n\sigma_\tau^2 + \sigma^2) \Rightarrow \frac{EMstr}{\sigma^2} = 1 + \frac{n\sigma_\tau^2}{\sigma^2}$$

if $\sigma_\tau^2 = 0 \Rightarrow \frac{E(MStr)}{\sigma^2} = 1$

i.e MStr have central χ^2 distribution if $\sigma_\tau^2 = 0$
since the hypothesis for the $H_0 : \sigma_\tau^2 = 0$.

$$MStr \begin{cases} \text{central } \chi^2 \text{ under } H_0 \\ \text{Non-central } \chi^2 \text{ under } H_1 \end{cases}$$

since sstr and sse are independent.

$$\frac{\dfrac{sstr/p-1}{\sigma^2 + n\sigma_\tau^2}}{\dfrac{sse/n-p}{\sigma^2}} \sim F_{(p-1)(n-p)}$$

$$\frac{Mstr}{MSe} \cdot \frac{\sigma^2}{\sigma^2 + n\sigma_\tau^2} \sim F_{(p-1)(n-p)}$$

$$\frac{MStr}{MSe} \sim F_{(p-1)(n-p)} \Rightarrow \frac{\sigma^2 + n\sigma_\tau^2}{\sigma^2} \text{ Non-Central tearm.}$$

The Ratio $MStr/MSe$ is always related to a central F under Ho.

However the non-central distribution of MStr under H1 and Hence the Ratio MStr/MSe dose differ from centrality.

- components of variance estimated :-

To estimate σ_τ^2, σ^2 we will use the Analysis of variance method as follows.

$MSe = \hat{\sigma}^2 \Rightarrow$ MSe is unbiased estimator for σ^2

$MStr = n\hat{\sigma}_\tau^2 + \hat{\sigma}^2 \Rightarrow$ MStr is unbiased estimator for σ^2 under Ho only.

$$\therefore \hat{\sigma}_\tau^2 = \frac{MStr - MSe}{n} \quad \text{balanced}$$

$$\hat{\sigma}_\tau^2 = \frac{MStr - MSe}{\left(\frac{1}{p-1}\right)\left(n - \frac{\sum n_i^2}{n}\right)} \quad \text{unbalanced.}$$

The variance of $\hat{\sigma}^2$ & $\hat{\sigma}_\tau^2$ are :-

$$var(\hat{\sigma}^2) = var(MSe) \Rightarrow \frac{(n-p)MSe}{\sigma^2} \sim \chi^2_{(n-p)}$$

$$var\left(\frac{(n-p)MSE}{\hat{\sigma}^2}\right) = 2(n-p)$$

$$\frac{(n-p)^{\cancel{2}}}{\sigma^4} var(MSE) = 2\cancel{(n-p)}$$

$$\boxed{\therefore var(\hat{\sigma}^2) = \frac{2\sigma^4}{n-p}}$$

$$var(\hat{\sigma}_\tau^2) = \frac{1}{n^2} var(MStr - MSE)$$

indep.

$$= \frac{1}{n^2}\left[var(MStr) + var(MSE) - 2cov(MStr, MSE)\right]$$

since $var(MStr) = 2(p-1)\left[\frac{(n\sigma_\tau^2+\sigma^2)}{(p-1)}\right]^2 = \frac{2(n\sigma_\tau^2+\sigma^2)}{p-1}$

$$\boxed{\therefore var(\hat{\sigma}_\tau^2) = \frac{1}{n^2}\left[\frac{2(n\sigma_\tau^2+\sigma^2)^2}{p-1} + \frac{2\sigma^4}{n-p} - 0\right]}$$

we can find.

$$S^2(\hat{\sigma}^2) = \frac{2\hat{\sigma}^4}{n-p} = \frac{2S^4}{n-p}$$

$$S^2(\hat{\sigma}_\tau^2) = \frac{1}{n^2}\left[\frac{2(nS_\tau^2+S^2)^2}{p-1} + \frac{2S^4}{n-p}\right]$$

H.W find the cov $(\hat{\sigma}^2, \hat{\sigma}_\tau^2)$

- Confidence Intervales for the Ratio σ^2_τ / σ^2:-

we know that

$$\underbrace{\frac{MStr}{MSe}}_{R} \cdot \frac{\sigma^2}{n\hat{\sigma}^2_\tau + \sigma^2} \sim F_{(p-1)(n-p)}$$

Thus $pr\left\{ a \leq R \cdot \frac{\sigma^2}{n\hat{\sigma}^2_\tau + \sigma^2} \leq b \right\} 1-\alpha$

where $a = F(p-1, n-p, 1-\alpha/2)$

$b = F(p-1, n-p, 1-\alpha/2)$

$$pr\left(\frac{a}{R} \leq \left(\frac{n\sigma^2_\tau + \sigma^2}{\sigma^2}\right)^{-1} \leq \frac{b}{R}\right) = 1-\alpha$$

$$pr\left(\frac{a}{R} \leq \left(n\frac{\sigma^2_\tau}{\sigma^2} + 1\right)^{-1} \leq \frac{b}{R}\right) = 1-\alpha$$

$$pr\left(\frac{R}{b} \leq n\frac{\sigma^2_\tau}{\sigma^2} + 1 \leq \frac{R}{a}\right) = 1-\alpha$$

$$pr\left(\frac{1}{n}\left(\frac{R}{b} - 1\right) \leq \frac{\sigma^2_\tau}{\sigma^2} \leq \frac{1}{n}\left(\frac{R}{a} - 1\right)\right) = 1-\alpha$$

- Expected Mean Squares :-

$$E(MSE) = \frac{1}{n-p} E\left[\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i\cdot})^2\right]$$

$$= \frac{1}{n-p} E\left[\sum_{i=1}^{p}\sum_{j=1}^{n_i}(\mu + \tau_i + \epsilon_{ij} - \mu - \tau_i - \bar{\epsilon}_{i\cdot})^2\right]$$

$$= \frac{1}{n-p} E\left[\sum_{i=1}^{p}\sum_{j=1}^{n_i}(\epsilon_{ij} - \bar{\epsilon}_{i\cdot})^2\right]$$

$$= \frac{1}{n-p} \sum_{i=1}^{p}\sum_{j=1}^{n_i} E(\epsilon_{ij} - \bar{\epsilon}_{i\cdot})^2$$

$$= \frac{1}{n-p} \sum_{i=1}^{p}\sum_{j=1}^{n_i} \left[var(\epsilon_{ij} - \bar{\epsilon}_{i\cdot}) + \left[E(\epsilon_{ij} - \bar{\epsilon}_{i\cdot})\right]^2\right]$$

$$= \frac{1}{n-p} \sum_{i=1}^{p}\sum_{j=1}^{n_i} \left[var(\epsilon_{ij}) + var(\bar{\epsilon}_{i\cdot}) - 2cov(\bar{\epsilon}_{i\cdot}, \epsilon_{ij}) + \left[E(\epsilon_{ij}) - E(\bar{\epsilon}_{i\cdot})\right]^2\right]$$

$$= \frac{1}{n-p} \sum_{i=1}^{p}\sum_{j=1}^{n_i} \left[\sigma^2 + \frac{\sigma^2}{n_i} - 2\frac{\sigma^2}{n_i} + 0\right]$$

$$= \frac{1}{n-p}\left[n\sigma^2 - \sum_{i=1}^{p} n_i \frac{\sigma^2}{n_i}\right]$$

$$E(MSE) = \frac{1}{n-p}(n-p)\sigma^2 = \boxed{\sigma^2}$$

or

$$E(MSE) = \frac{1}{n-p} E\sum_{i=1}^{p}\sum_{j=1}^{n_i}(\epsilon_{ij} - \bar{\epsilon}_{i\cdot})^2$$

$$= \frac{1}{n-p}\left[\sum_{i=1}^{p}\sum_{j=1}^{n_i} E(\epsilon_{ij}^2) + \sum_{i=1}^{p} n_i E(\bar{\epsilon}_{i\cdot}^2) - 2\sum_{i=1}^{p}\sum_{j=1}^{n_i} E(\epsilon_{ij}\bar{\epsilon}_{i\cdot})\right]$$

$$= \frac{1}{n-p}\left[\sum_{i=1}^{p}\sum_{j=1}^{n_i}\left[var(\epsilon_{ij}) + [E(\epsilon_{ij})]^2\right]\right] - \sum_{i=1}^{p} n_i E(\bar{\epsilon}_{i.}$$

$$= \frac{1}{n-p}\left[n\sigma^2 - \sum_{i=1}^{p} n_i \frac{\sigma^2}{n_i}\right]$$

$$E(MSE) = \frac{1}{n-p}(n-p)\sigma^2 = \boxed{\sigma^2}$$

$$MStr = \frac{1}{p-1}\sum_{i=1}^{p} n_i(\bar{y}_{i.} - \bar{y}_{..})^2$$

$$\bar{y}_{..} = \frac{\sum_{i=1}^{p}\sum_{j=1}^{n_i} y_{ij}}{n}$$

$$= \sum_{i=1}^{p}\sum_{j=1}^{n_i}(\mu + \tau_i + \epsilon_{ij})/n$$

$$= \mu + \sum_{i=1}^{p} n_i \tau_i/n + \sum_{i=1}^{p}\sum_{j=1}^{n_i}\epsilon_{ij}/n$$

$$= \mu + \bar{\tau} + \bar{\epsilon}_{..}$$

$$MStr = \frac{1}{p-1}\sum_{i=1}^{p} n_i(\mu + \tau_i + \bar{\epsilon}_{i.} - \mu - \bar{\tau} - \bar{\epsilon}_{..})^2$$

$$= \frac{1}{p-1}\sum_{i=1}^{p} n_i\left[(\tau_i - \bar{\tau}) + (\bar{\epsilon}_{i.} - \bar{\epsilon}_{..})\right]^2$$

$$E(MStr) = \frac{1}{p-1}\sum_{i=1}^{p} n_i E\left[(\tau_i - \bar{\tau})^2 + (\bar{\epsilon}_{i.} - \bar{\epsilon}_{..})^2 + 2(\tau_i - \bar{\tau})(\bar{\epsilon}_{i.} - \bar{\epsilon}_{..})\right]$$

$$E(\tau_i-\bar{\tau})^2 = var(\tau_i-\bar{\tau})+\left[E(\tau_i-\bar{\tau})\right]^2$$

$$= var(\tau_i)+var(\bar{\tau})-2cov(\tau_i\,\bar{\tau})$$

$$= \sigma_\tau^2 + \sum_{i=1}^{p} n_i^2\sigma_\tau^2/n^2 - 2\sum_{i=1}^{p} n_i^2\sigma_\tau^2/n^2$$

$$= \sigma_\tau^2 - \sum_{i=1}^{p} n_i^2\sigma_\tau^2/n^2$$

$$\sum_{i=1}^{p} n_i E(\tau_i-\bar{\tau})^2 = \sum_{i=1}^{p} n_i\sigma_\tau^2 - n\sum_{i=1}^{p} n_i^2\sigma_\tau^2/n^2$$

$$= n\sigma_\tau^2 - \frac{\sigma_\tau^2}{n}\sum_{i=1}^{p} n_i^2$$

$$\boxed{\sum_{i=1}^{p} n_i E(\tau_i-\bar{\tau})^2 = \sigma_\tau^2\left[n - \frac{\sum_{i=1}^{p} n_i^2}{n}\right]}$$

$$E(\bar{\epsilon}_{i.}-\bar{\epsilon}_{..})^2 = var(\bar{\epsilon}_{i.}-\bar{\epsilon}_{..})+\left[E(\bar{\epsilon}_{i.}-\bar{\epsilon}_{..})\right]^2$$

$$= var(\bar{\epsilon}_{i.})+var(\bar{\epsilon}_{..})-2cov(\bar{\epsilon}_{i.},\bar{\epsilon}_{..})$$

$$= \frac{\sigma^2}{n_i}+\frac{\sigma^2}{n}-2cov\left(\bar{\epsilon}_{i.},\frac{1}{n}\sum_{i=1}^{p} n_i\bar{\epsilon}_{i.}\right)$$

$$= \frac{\sigma^2}{n_i}+\frac{\sigma^2}{n}-2cov\begin{cases} 0 & \forall i\neq i' \\ \sum_{i=1}^{p}\frac{n_i^2}{n^2}\,var(\bar{\epsilon}_{i.}) & \forall i=i' \end{cases}$$

$$\therefore E(\bar{\epsilon}_{i.}-\bar{\epsilon}_{..})^2 = \frac{\sigma^2}{n_i}+\frac{\sigma^2}{n}-2\sum_{i=1}^{p}\frac{\sigma^2}{\cancel{n_i}}\cdot\frac{n_i^{\cancel{2}}}{n^2}$$

$$= \frac{\sigma^2}{n_i}+\frac{\sigma^2}{n}-2\frac{\sigma^2}{n^2}\left(\sum_{i=1}^{p} n_i\right) \quad n$$

$$E(\bar{\epsilon}_{i.}-\bar{\epsilon}_{..})^2 = \frac{\sigma^2}{n_i} - \frac{\sigma^2}{n}$$

$$\sum_{i=1}^{p} n_i E(\bar{\epsilon}_{i.}-\bar{\epsilon}_{..})^2 = \sum_{i=1}^{p} n_i \frac{\sigma^2}{n_i} - \frac{\sigma^2}{n} \sum_{i=1}^{p} n_i \quad n$$

$$\boxed{\sum_{i=1}^{p} n_i E(\bar{\epsilon}_{i.}-\bar{\epsilon}_{..})^2 = p\sigma^2 - \sigma^2 = (p-1)\sigma^2}$$

$$2E(\tau_i-\bar{\tau})(\bar{\epsilon}_{i.}-\bar{\epsilon}_{..}) = 0$$

because τ_i, ϵ_{ij} uncorrelated.

$$\therefore E(MStr) = \frac{1}{p-1}\left[\sigma_\tau^2\left(n - \frac{\sum_{i=1}^{p} n_i^2}{n}\right) + (p-1)\sigma^2\right]$$

$$= \sigma^2 + \underbrace{\left[\frac{1}{p-1}\sigma_\tau^2\left(n - \frac{\sum_{i=1}^{p} n_i^2}{n}\right)\right]}_{\text{non-central term}}$$

if $\sigma_\tau^2 = 0$

$\therefore E(MStr) = \sigma^2$

$\therefore MStr \sim \chi^2_{(p-1)}$ central.

$\boxed{\text{or}}$

$$E(MStr) = \frac{1}{p-1} E\left[\sum_{i=1}^{p} n_i(\bar{y}_{i.}-\bar{y}_{..})^2\right]$$

$$= \frac{1}{p-1} E\left[\sum_{i=1}^{p} n_i(\mu+\tau_i+\bar{\epsilon}_{i.}-\mu-\bar{\tau}-\bar{\epsilon}_{..})^2\right]$$

$$= \frac{1}{p-1} E\sum_{i=1}^{p} n_i\left[(\tau_i-\bar{\tau})+(\bar{\epsilon}_{i.}-\bar{\epsilon}_{..})\right]^2$$

$$= \frac{1}{p-1}\sum_{i=1}^{p} n_i\left[var(\tau_i)+var(\bar{\tau})-2cov(\tau_i\bar{\tau}) + var(\bar{\epsilon}_{i.})+var(\bar{\epsilon}_{i.})-2cov(\bar{\epsilon}_{i.},\bar{\epsilon}_{..})\right]$$

$$E(MStr) = \frac{1}{p-1}\left[n\sigma_\tau^2 + \frac{\sigma_\tau^2}{n}\sum_{i=1}^{p} n_i^2 - 2\frac{\sigma_\tau^2}{n}\sum_{i=1}^{p} n_i + p\sigma^2 + \sigma^2 - 2\sigma^2\right]$$

$$\boxed{= \sigma^2 + \sigma_\tau^2\left(\frac{1}{p-1}\right)\left(n - \frac{\sum_{i=1}^{p} n_i}{n}\right)}$$ non-central

$$E\left(\frac{MStr}{MSe}\right) = \begin{cases} 1 & \text{if } H_0 \text{ is true} \rightarrow \text{central} \\ >1 & \text{if } H_1 \text{ is true} \rightarrow \text{Non-central} \end{cases}$$

$$\frac{E(MStr)}{\sigma^2} = 1 + \underbrace{\frac{\sigma_\tau^2}{\sigma^2}\cdot\frac{1}{p-1}\left(n - \frac{\sum_{i=1}^{p} n_i^2}{n}\right)}_{\text{non-central tearm.}} \sim \text{Non-central } \chi^2$$

$$\frac{E(MSe)}{\sigma^2} = \frac{\sigma^2}{\sigma^2} = 1 \sim \text{central } \chi^2$$

$H_0: \sigma_\tau^2 = 0$

$$E(R) = \frac{E(MStr)}{E(MSe)} \begin{cases} 1 & \text{under } H_0: \sigma_\tau^2 = 0 \text{ central} \\ >1 & \text{under } H_1: \sigma_\tau^2 \neq 0 \text{ non-central} \end{cases}$$

Thanks for lessening

Please Read Carefully and write the answer of above questions