Introduction to Statistics and Data Science using *eStat*

Chapter 9 Testing Hypothesis for Several Population Means

9.2 Design of Experiments for Sampling
9.2.1 Completely Randomized Design
9.2.2 Randomized block design

Jung Jin Lee
Professor of Soongsil University, Korea
Visiting Professor of ADA University, Azerbaijan

9.2 Design of Experiments for Sampling

9.2.1 Completely Randomized Design

- Design of experiments to have little impact from other factors.
- One way to do this is to make the whole experiments random.
- Example: Compare gas milage of three cars (A, B, C) with 5 drivers

Driver	1	2	3	4	5
Car Type	B	A	B	C	A
	B	C	A	A	C
	C	B	A	B	C

9.2 Design of Experiments for Sampling

9.2.2 Randomized Block Design

Driver	1	2	3	4	5
Car Type (gas mileage)	A(22.4)	B(12.6)	C(18.7)	A(21.1)	A(24.5)
	C(20.2)	C(15.2)	A(19.7)	B(17.8)	C(23.8)
	B(16.3)	A(16.1)	B(15.9)	C(18.9)	B(21.0)

9.2 Design of Experiments for Sampling

9.2.2 Randomized Block Design

- **Statistical model of the randomized block design:**

$$Y_{ij} = \mu + \alpha_i + B_j + \epsilon_{ij}, \qquad i = 1, 2, \dots, k, j = 1, 2, \dots, b$$

B_j : effect of j^{th} level of the block variable

- **In the randomized block design, the total variation is divided into as follows:**

$$Y_{ij} - \overline{Y}_{..} = (Y_{ij} - \overline{Y}_{i.} - \overline{Y}_{.j} + \overline{Y}_{..}) + (\overline{Y}_{i.} - \overline{Y}_{..}) + (\overline{Y}_{.j} - \overline{Y}_{..})$$

9.2 Design of Experiments for Sampling

9.2.2 Randomized Block Design

Division of sum of squares and degree of freedom

Sum of squares : $SST = SSE + SSTr + SSB$

Degree of freedom : $bk-1 = (b-1)(k-1) + (k-1) + (b-1)$

Table 9.2.3 Analysis of Variance Table of the randomized block design

Variation	Sum of Squares	Degree of freedom	Mean Squares	F value
Treatment	SSTr	$k-1$	$MSTr = \frac{SSTr}{k-1}$	$F_0 = \frac{MSTr}{MSE}$
Block	SSB	$b-1$	$MSB = \frac{SSB}{b-1}$	
Error	SSE	$(b-1)(k-1)$	$MSE = \frac{SSE}{(b-1)(k-1)}$	
Total	SST	$bk-1$		

Total sum of squares, degree of freedom $bk-1$

$$SST = \sum_{i=1}^{k}\sum_{j=1}^{b}(Y_{ij} - \overline{Y}_{..})^2$$

Error sum of squares, degree of freedom $(b-1)(k-1)$

$$SSE = \sum_{i=1}^{k}\sum_{j=1}^{b}(Y_{ij} - \overline{Y}_{i.} - \overline{Y}_{.j} + \overline{Y}_{..})^2$$

Treatment sum of squares, degree of freedom $k-1$

$$SSTr = \sum_{i=1}^{k}\sum_{j=1}^{b}(\overline{Y}_{i.} - \overline{Y}_{..})^2 = b\sum_{i=1}^{k}(\overline{Y}_{i.} - \overline{Y}_{..})^2$$

Block sum of squares, degree of freedom $b-1$

$$SSB = \sum_{i=1}^{k}\sum_{j=1}^{b}(\overline{Y}_{.j} - \overline{Y}_{..})^2 = k\sum_{j=1}^{b}(\overline{Y}_{.j} - \overline{Y}_{..})^2$$

9.2 Design of Experiments for Sampling

[Example 9.2.1] Table 9.2.4 is the rearrangement of the fuel mileage data in Table 9.2.2 measured by five drivers and car types.

Driver		1	2	3	4	5	Average
Car Type	A	22.4	16.1	19.7	21.1	24.5	20.76
	B	16.3	12.6	15.9	17.8	21.0	16.72
	C	20.2	15.2	18.7	18.9	23.8	19.36
Average		19.63	14.63	18.10	19.27	23.10	18.947

1) Assuming that this data has been measured by the completely design, use『eStat』to do the analysis of variance whether the three car types have the same fuel mileage.

2) Assuming that this data has been measured by the randomized block design, use『eStat』to do the analysis of variance whether the three car types have the same fuel mileage.

9.2 Design of Experiments for Sampling

<Answer of [Example 9.2.1]>

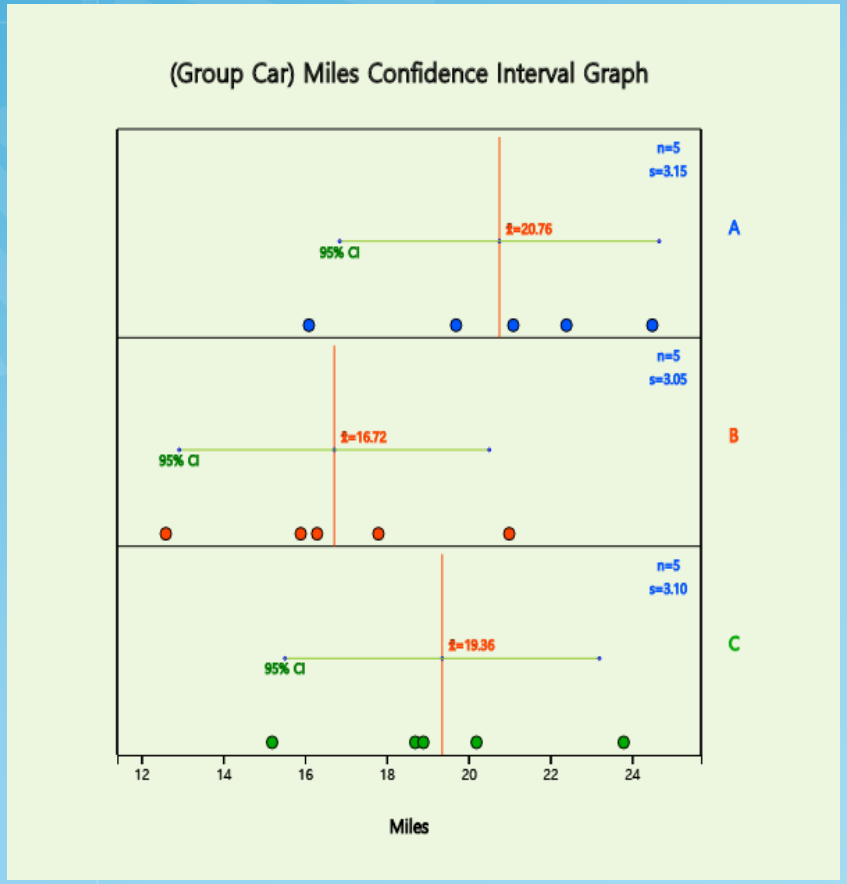
File: Ex921GasMileage.csv

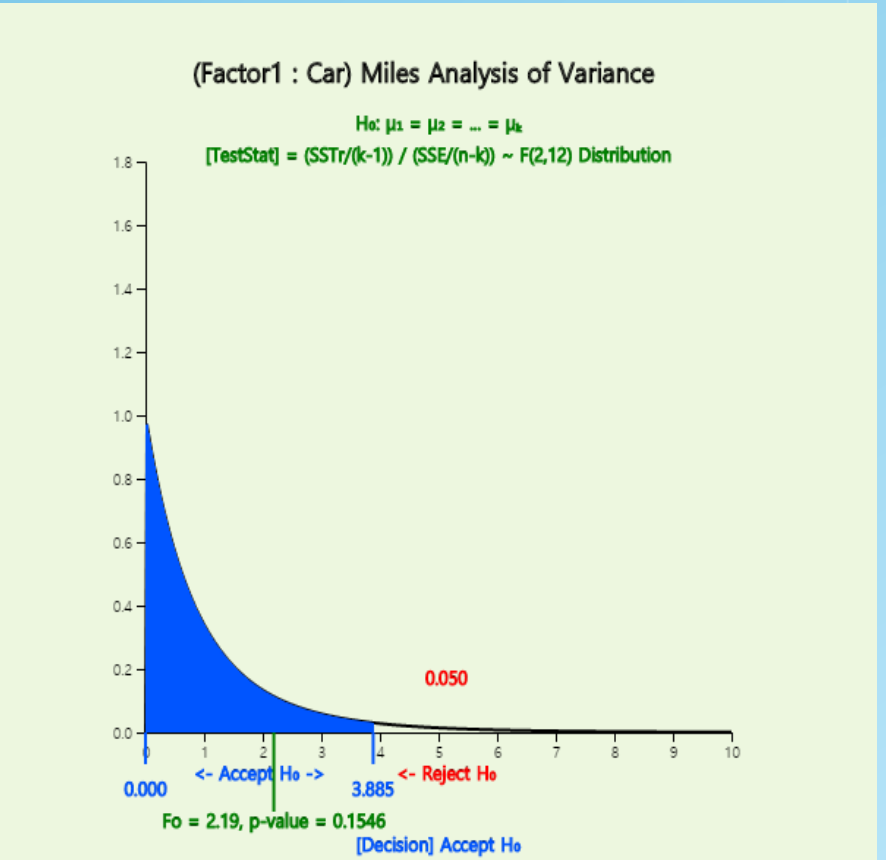
Analysis Var: 3: Miles | by Group: 1: Car

(Selected data: Raw Data) (Select up to two g

SelectedVar: V3 by V1,

	Car	Driver	Miles	V4
1	A	1	22.4	
2	A	2	16.1	
3	A	3	19.7	
4	A	4	21.1	
5	A	5	24.5	
6	B	1	16.3	
7	B	2	12.6	
8	B	3	15.9	
9	B	4	17.8	
10	B	5	21.0	
11	C	1	20.2	
12	C	2	15.2	
13	C	3	18.7	
14	C	4	18.9	
15	C	5	23.8	





9.2 Design of Experiments for Sampling

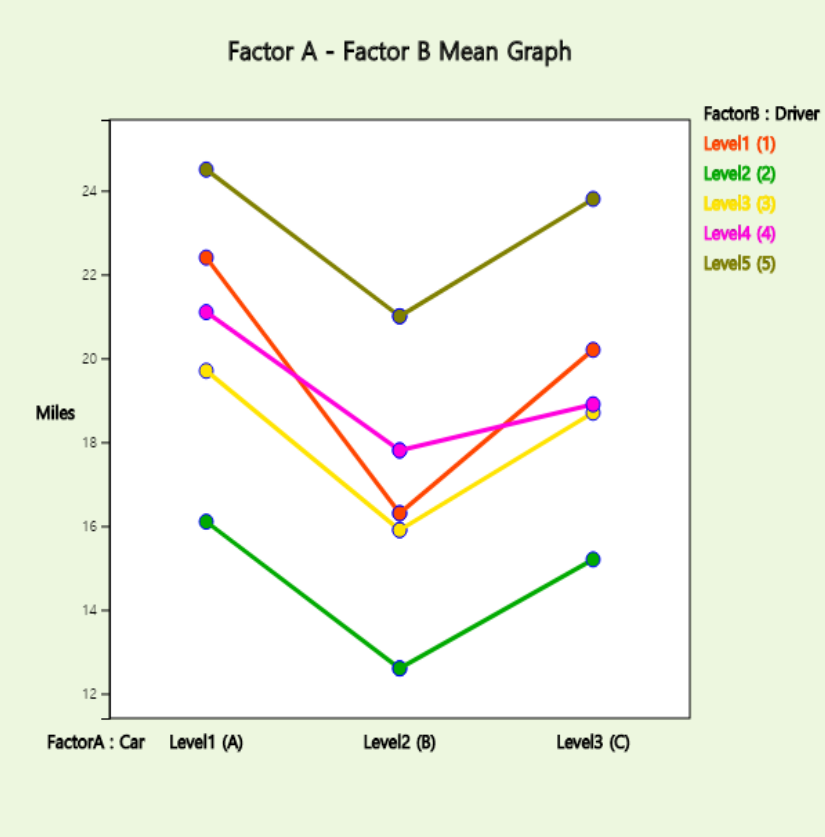
<Answer of [Example 9.2.1]>

Analysis of Variance					
Factor	Sum of Squares	deg of freedom	Mean Squares	F value	p value
Treatment	42.085	2	21.043	2.190	0.1546
Error	115.312	12	9.609		
Total	157.397	14			

Multiple Comparison	Analysis Var	(Miles)	Group Name	(Car)
Mean Difference (95%HSD)	1 (A) 20.76	2 (B) 16.72	3 (C) 19.36	
1 (A) 20.76		4.04 (5.23)	1.40 (5.23)	
2 (B) 16.72	4.04 (5.23)		2.64 (5.23)	
3 (C) 19.36	1.40 (5.23)	2.64 (5.23)		

9.2 Design of Experiments for Sampling

<Answer of [Example 9.2.1]>



Two-dimension Statistics

Observation Mean Std Dev	Factor B (Driver) Level1 (1)	Factor B (Driver) Level2 (2)	Factor B (Driver) Level3 (3)	Factor B (Driver) Level4 (4)	Factor B (Driver) Level5 (5)	Factor A Level i Total
FactorA (Car) Level1 (A)	1 22.400 NaN	1 16.100 NaN	1 19.700 NaN	1 21.100 NaN	1 24.500 NaN	5 20.760 3.148
FactorA (Car) Level2 (B)	1 16.300 NaN	1 12.600 NaN	1 15.900 NaN	1 17.800 NaN	1 21.000 NaN	5 16.720 3.054
FactorA (Car) Level3 (C)	1 20.200 NaN	1 15.200 NaN	1 18.700 NaN	1 18.900 NaN	1 23.800 NaN	5 19.360 3.097
Factor B Level j Total	3 19.633 3.089	3 14.633 1.818	3 18.100 1.970	3 19.267 1.680	3 23.100 1.852	15 18.947 3.353
Missing Observations	0					

Analysis of Variance

Factor	Sum of Squares	deg of freedom	Mean Squares	F value	p value
Factor A (Car)	42.085	2	21.043	43.447	< 0.0001
Factor B (Driver)	111.437	4	27.859	57.521	< 0.0001
Error	3.875	8	0.484		
Total	157.397	14			

Thank you