HT to determine if mean value changes

$H_0: \mu_1 - \mu_2 = 0$

$H_a: \mu_1 - \mu_2 > 0$

TS : $t_{obs} = \dfrac{\bar{d} - D_0}{S_d/\sqrt{n}}$

Table	Di

RR : $t_{obs} > t_{n-1,\alpha}$

If true, Reject H_0

Is there difference in average of multiple

Draw the table

$H_0: \mu_1 = \mu_2 = \mu_3$

H_a: Not all are equal

TS : $F_{obs} = \dfrac{S_B^2}{S_W^2}$ $\quad S_B^2 = n.$(variance of sample means)

$S_W^2 = \dfrac{SD_1^2 + SD_2^2 + SD_3^2}{t}$ = Avg. of variances.

RR: $F_{obs} > F_{df1, df2, \alpha}$

Was there any evidence of violation in required conditions to conduct in the above.

- Check if variances are equal

$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$

H_a: Not all are equal.

TS : $F_{max} = \dfrac{S^2_{max}}{S^2_{min}}$ → highest variance → least variance (from table)

RR: $F_{max} > F_{max, df1, df2, \alpha}$ $\quad df1 = t$, $df2 = n-1$

RB/CR Design. Draw table in given

RB

Statistical Model: $y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$

μ = overall mean that's an unknown const

α_i = an effect due to treatment, " " "

β_i = an effect due to block, " " "

ε_{ij} = a Random error associated with response on treatment i, block j

	1	2	3	Sum
A	—	—	—	$y_{1\cdot}$
B	—	—	—	$y_{2\cdot}$
C	—	—	—	$y_{3\cdot}$
Sum	$y_{\cdot 1}$	$y_{\cdot 2}$	$y_{\cdot 3}$	$y_{\cdot\cdot}$

ANOVA table - $TSS = \left(\sum y_{ij}^2\right) - \dfrac{y_{\cdot\cdot}^2}{bt}$

$SST = \left(\sum_i \dfrac{y_{i\cdot}^2}{b}\right) - \dfrac{y_{\cdot\cdot}^2}{bt}$

$SSB = \left(\sum_j \dfrac{y_{\cdot j}^2}{t}\right) - \dfrac{y_{\cdot\cdot}^2}{bt}$

$SSE = TSS - SST - SSB$

$df1 = t-1, \quad df2 = b-1$

$MST = \dfrac{SST}{t-1}$, $MSB = \dfrac{SSB}{b-1}$, $MSE = \dfrac{SSE}{(b-1)(t-1)}$

$F^{①}_{obs} = \dfrac{MST}{MSE}$ $\quad F^{②}_{obs} = \dfrac{MSB}{MSE}$

Source due to	Sum of Squares (SS)	df	Mean Square (MS)	F
Treatments	SST	$t-1$	MST	MST/MSE
Blocks	SSB	$b-1$	MSB	MSB/MSE
Error	SSE	$(t-1)(b-1)$	MSE	
Total	+ ___	+ ___		×

HT to determine if there is any difference among treatment means

$H_0: \alpha_A = \alpha_B = \alpha_C = \alpha_D = 0$

H_a: Not all the above are equal to zero

TS : $F^{①}_{obs} = \dfrac{MST}{MSE}$

RR : $F^{①}_{obs} > F_{t-1,(b-1)(t-1),\alpha}$

Pvalue : $P(F_{t-1,(b-1)(t-1),\alpha} > F^{①}_{obs})$

Pvalue > α FTR H_0
Pvalue ≤ α Reject H_0

Ex - P value is greater than 0.25

HT to determine if there is any difference among block effects.

$H_0: \beta_A = \beta_B = \beta_C = \beta_D = 0$

H_a: Not all the above are equal to zero

TS : $F_{obs} = \dfrac{MSB}{MSE}$

RR: $F_{obs} > F_{b-1,(t-1)(b-1),\alpha}$

Pvalue : $P(F_{b-1,(t-1)(b-1),\alpha} > F_{obs})$

Relative efficiency $RE = \dfrac{MSE_{CR}}{MSE_{RB}} = \dfrac{(b-1)MSB + b(t-1)MSE_{RB}}{(bt-1)\,MSE_{RB}}$

Completely Randomized Design

n_T = total Samples = 12 No. of Samples in each row

$TSS = (\sum y_{ij}^2) - \frac{y_{..}^2}{n_T}$ $SSB = \left(\sum \frac{y_{i.}^2}{n_i}\right) - \frac{y_{..}^2}{n_T}$

$SSW = TSS - SSB$. $S_B^2 = \frac{SSB}{t-1}$. $S_W^2 = \frac{SSW}{n_T - t}$

ANOVA table

Source due to	Sum of Squares (SS)	df	Mean Square (MS)	F
Between Samples	SSB	$t-1$	S_B^2	Fobs
Within Samples	SSW	$n_T - t$	S_W^2	×
Total			×	×

HT to determine if there is any difference among averages

Ho: $\mu_A = \mu_B = \mu_C$
Ha: Not all are equal

TS: $F_{obs} = \frac{S_B^2}{S_W^2}$

RR: $F_{obs} > F_{df1, df2, \alpha}$ $df1 = t-1$, $df2 = n_T - t$

P value: $p(F_{2,9,0.05} > F_{obs})$
$P(F_{2,9,0.05} > 41.05)$
P value < 0.001

HT to check if there is a difference b/w brand A & B

Ho: $\mu_A - \mu_B = 0$
Ha: $\mu_A - \mu_B \neq 0$

Fischer's LSD Rule: $LSD = t_{\alpha/2, n_T - t}\left(\sqrt{S_W^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}\right)$

$= t_{0.025,9}\left(\sqrt{0.076\left(\frac{1}{4} + \frac{1}{4}\right)}\right)$

$= (\quad)(\quad)$

$= 0.44$

Mean Difference b/w A&B

$|\bar{y}_A - \bar{y}_B| = \left|\frac{10.6}{4} - \frac{15.9}{4}\right| = 1.325$

Step 3: $|\bar{y}_A - \bar{y}_B| > LSD$

Reject Ho

---

e.g. ice price(x)	Qty Sold (y)	x^2	y^2	xy

HT to determine if independent variable is used to predict dependent variable

Ho: $\beta_1 = 0$
Ha: $\beta_1 \neq 0$ $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$

TS: $t_{obs} = \frac{\hat{\beta}_1 - \beta_{10}}{S_\varepsilon / \sqrt{S_{xx}}}$ $S_\varepsilon = \sqrt{\frac{SSE}{n-2}} \rightarrow SSE = \frac{S_{xy}^2}{S_{xx}}$

$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$ $S_{yy} = \sum y_i^2 - \frac{(\sum y_i)^2}{n}$

$S_{xy} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$

RR: $|t_{obs}| > t_{n-2, \alpha/2}$

If true, Reject Ho, We've sufficient evidence to say that independent variable is useful

Check if Pooled t test / Separate t test (Equal variances)

Ho: $\sigma_1^2 = \sigma_2^2$
Ha: $\sigma_1^2 \neq \sigma_2^2$

TS: $F_{obs} = \frac{S_i^2}{S_j^2}$ where $S_i^2 > S_j^2$

$df1 = n_2 - 1$
$df_2 = n_1 - 1$

RR: $F_{obs} \geq F_{\alpha/2, df1, df2}$

high S_i value's n low S_j value's n

$< \frac{1}{F_{\alpha/2, df2, df1}}$

FTR – No sufficient evidence to conclude Ha
We conduct pooled t-test

Is there sufficient evidence that one has higher mean than the other?

Ho: $\mu_C - \mu_M = 0$
Ha: $\mu_C - \mu_M > 0$

TS: $t_{obs} = \frac{(\bar{y}_1 - \bar{y}_2) - D_0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

{Dont forget this is square & find square root of it}

$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$ where $S_1 > S_2$

RR: $t_{obs} > t_{n_1 + n_2 - 2}\ \alpha$

P value – $P(t_{n_1+n_2-2, \alpha} > t_{obs})$

Confidence Interval

$(\bar{y}_1 - \bar{y}_2) \pm t_{\alpha/2, n_1+n_2-2} \cdot S_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

X is contingent on Y

X/Y	Out	Avg	Poor	Sum Rows
				R_1
				R_2
Sum columns	C_1	C_2	C_3	whole sum

H_0: Variables X & Y are independent
H_a: Variables X & Y are dependent

$E_{11} = \frac{R_1 (C_1)}{\text{whole sum}}$ $E_{12} = \frac{R_1(C_2)}{\text{whole}}$ $E_{13} = \frac{R_1(C_3)}{\text{whole}}$

$E_{21} = \frac{R_2(C_1)}{\text{whole}}$ $E_{22} = \frac{R_2(C_2)}{\text{whole}}$ $E_{23} = \frac{R_2(C_3)}{\text{whole}}$

TS : $\chi^2_{obs} = \sum_{i=1}^{R}\sum_{j=1}^{C} \frac{(n_{ij} - E_{ij})^2}{E_{ij}}$

RR : $\chi^2_{obs} > \chi^2_{(R-1)(C-1),\alpha}$

FTR

No sufficient evidence to conclude.

X is not contingent on Y.

2F1

① Draw table & write Di Before-After & $\bar{d}-\mu_i$

X 6.75, 11.76, 7.45, 6.68, 2.21

column $\bar{d}-\mu_i \rightarrow$ 0.22, −4.79, −0.48, 0.29, 4.76

$\bar{d} = \frac{6.75+11.76+7.45+6.68+2.21}{5} = 6.97$

$S_d^2 = \frac{(0.22)^2+(-4.79)^2+(-0.48)^2+(0.29)^2+(4.76)^2}{4} = 11.49$

$S_d = \sqrt{11.49} = 3.38$

given $n=5$ μ_B = Before μ_A = After

Hypothesis ; $H_0 : \mu_B - \mu_A = 0$
$H_a : \mu_B - \mu_A > 0$

Test statistic $t_{obs} = \frac{\bar{d} - D_0}{S_d/\sqrt{n}} = \frac{6.97-0}{3.38/\sqrt{5}} = \frac{6.97}{1.51} = 4.61$

RR : $t_{obs} > t_{critical} \Rightarrow t_{obs} > t_{n-1,\alpha} \Rightarrow t_{obs} > t_{4,0.05}$

Decision: $4.61 > 2.132$ So, reject H_0

Pvalue: $P(t_{n-1} > t_{obs}) \Rightarrow P(t_4 > 4.61)$
Lies between 0.0019 0.005

Pvalue $< \alpha$ Reject H_0

There is sufficient evidence to conclude H_a.

② Draw table

a) Step 1 - Hypothesis $H_0: \mu_{Newtown} = \mu_A = \mu_B = \mu_C = \mu_D$
H_a: Not all above are equal.

$S_W^2 = \frac{99(0.3^2+0.4^2+0.4^2+0.5^2+0.5^2)}{495} = 0.182$

variance of sample means → var(8, 10, 10, 11, 11)

$\bar{y} = \frac{8+10+10+11+11}{5} = 10$

$S_d^2 = \frac{(10-8)^2+(10-10)^2+(10-10)^2+(10-11)^2+(10-11)^2}{4} = \frac{4+0+0+1+1}{4} = \frac{6}{4} = 1.5$

$S_B^2 = n$ (Variance of means) $= 100 \cdot (1.5) = 150$

Test Statistic $\Rightarrow F_{obs} = \frac{S_B^2}{S_W^2} = \frac{150}{0.182} = 824.17$

RR : $F_{critical} = F_{\alpha, df_1, df_2}$ where $df_1 = t-1$ $df_2 = nt-t$

$F_{0.05,4,495} = 2.37$

$F_{obs} > F_{\alpha, df_1, df_2} \Rightarrow F_{obs} > 2.37$

Decision : $824.17 > 2.37$ true Reject H_0

Pvalue : $P(F_{4,495} > 824.17)$
Pvalue is less than 0.001
pvalue < alpha α Reject H_0

Conclusion : There is sufficient evidence to conclude H_a

b) Violation of Required Condition → Check for equal variances

Hypothesis : $H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2$
H_a: Not all of these are equal

Test Statistic $F_{max} = \frac{S_{max}^2}{S_{min}^2} = \frac{(0.5)^2}{(0.3)^2} = 2.77$

RR ; $F_{critical} \Rightarrow F_{max,\alpha,df_1,df_2}$ $df_1 = t$ $df_2 = n-1$

$F_{max,0.05,5,99} = 2.04$ or less than that

$F_{obs} > F_{max,\alpha,df_1,df_2} \Rightarrow$

Decision $2.77 > 2.04$ Reject H_0

Conclusion: There is sufficient evidence to conclude H_a

3. Draw table & add Sum column and row in the end.

a) Here car model is extraneous variable, It is RB design

b) Statistical model $y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$: Write definitions

c) Here $t=4$ $b=4$

ANOVA table

Source due to	Sum of Squares	df	Mean Square (MS)	F
Treatments	5.7	3	1.9 = MST	0.125
Blocks	755.2	3	251.7 = MSB	16.5
Error	137.1	9	15.2 = MSE	
Total	878	15		

TSS $= \sum_{ij} y_{ij}^2 - \frac{y_{..}^2}{bt} = (15)^2 + 33^2 + 13^2 + \dots + 21.6^2 - \frac{(353.9)^2}{4(4)}$

$8725 - 7827 = 898$

SST $= \sum_{i=1} \frac{y_{i.}^2}{b} - \frac{y_{..}^2}{bt} = \frac{90.2^2 + 84.3^2 + 89.6^2 + 89.8^2}{4} - 7827.8$

$= 7833.5 - 7827.8 = 5.7$

$SSB = \sum_{j=1} \frac{y_{.j}^2}{t} - \frac{y_{..}^2}{bt} = \frac{55.8^2+125.4^2+69.3^2+103.4^2}{4} - 7827.8$

$= 8583 - 7827.8 = 755.2$

$SSE = TSS - SST - SSB = 898 - 5.7 - 755.2 = 137.1$

d) $H_0: \alpha_A = \alpha_B = \alpha_C = \alpha_D = 0$

H_a: Not all of the above are equal to zero

Test Statistic $F_{obs}^{①} = \frac{MST}{MSE} = \frac{1.9}{15.2} = 0.125$

RR: $F^{①}_{obs} > F_{t-1,(b-1)(t-1),\alpha} \Rightarrow F^{①}_{obs} > F_{3,9,0.05} \Rightarrow F^{①}_{obs} > 3.86$

Pvalue: $P(F_{t-1,(b-1)(t-1)} > F^{①}_{obs}) = P(F_{3,9} > 0.125)$

p value is greater than 0.25

Decision: p value $> \alpha \rightarrow$ FTR Ho

Conclusion: Insufficient evidence to conclude Ha

e) Car model is the extraneous variable

Hypothesis: $H_0: \beta_A = \beta_B = \beta_C = \beta_D = 0$ Ha: Not all the above are equal to zero

Test Statistic: $F^{②}_{obs} = \frac{MSB}{MSE} = \frac{251.7}{15.2} = 16.5$

RR: $F^{②}_{obs} > F_{b-1,(b-1)(t-1),\alpha} \Rightarrow F^{②}_{obs} > F_{3,9,0.05} \Rightarrow F^{②}_{obs} > 3.86$

Pvalue: $P(F_{b-1,(b-1)(t-1)} > F^{②}_{obs}) \Rightarrow P(F_{3,9} > 16.5)$ is less than 0.001

pvalue $< \alpha$ There is sufficient evidence

Decision: reject Ho Conclusion: to conclude Ha

f) $RE = \frac{MSE_{CR}}{MSE_{RB}} = \frac{(b-1)MSB + b(t-1)MSE_{RB}}{(bt-1)MSE_{RB}} = \frac{3(251.7)+4(3)(15.2)}{15(15.2)}$

$= \frac{755.1+182.4}{228} = 4.11$

4) Ozone

This is CR design

a)

					Sum
A	2.2	3	2.7	2.7	10.6
B	3.6	3.9	4.1	4.3	15.9
C	4.3	4.4	4.5	4.1	17.3
Sum	10.1	11.3	11.3	11.1	43.8

b) $t = 3$, $n_T = 12$

ANOVA

Source due to	SS	df	MS	F
Between Samples	6.245	2	$S_B^2 = 3.125$	41.11
Within Samples	0.685	9	$S_W^2 = 0.076$	
Totals	TSS 6.93	11		

$TSS = \sum_{ij} y_{ij}^2 - \frac{y_{..}^2}{n_T} \Rightarrow 2.2^2+3^2+\dots+4.1^2 - \frac{(43.8)^2}{12}$

$= 166.8 - 159.87 = 6.93$

$SSB = \frac{10.6^2}{4} + \frac{15.9^2}{4} + \frac{17.3^2}{4} - 159.87 = 6.245$

$SSW = TSS - SSB = 6.93 - 6.245 = 0.685$

c) Hypothesis: $H_0: \mu_A = \mu_B = \mu_C$

Ha: Not all are equal

TS: $\frac{S_B^2}{S_W^2} = F_{obs} = \frac{3.125}{0.076} = 41.11$

RR: $F_{obs} > F_{\alpha, t-1, n_T-t}$ $df_1 = t-1$, $df_2 = n_T - t$

$F_{obs} > F_{0.05,2,9}$ $F_{obs} > 4.26$

Pvalue: $P(F_{t-1,n_T-t} > F_{obs})$ $P(F_{2,9} > 41.11)$ less than 0.001

Decision: pvalue $< \alpha$ & $F_{obs} > 4.26$ So reject Ho

Conclusion: There is sufficient evidence to conclude Ha

d) Not Answering since it is a CR design

e) Hypothesis $H_0: \mu_A - \mu_B = 0$

$H_a: \mu_A - \mu_B \neq 0$

Fischer's LSD Rule $LSD = t_{\alpha/2, n_T-t}\left(\sqrt{S_W^2\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}\right) = t_{0.025,9}\sqrt{0.076\left(\frac{1}{4}+\frac{1}{4}\right)}$

$= 2.26\sqrt{0.076(1/2)} = 0.44 = LSD$

Mean difference b/w A & B $|\bar{y}_A - \bar{y}_B| = \left|\frac{10.6}{4} - \frac{15.9}{4}\right| = |2.65 - 3.975| = 1.325$

Step 3 $|\bar{y}_A - \bar{y}_B| = 1.325$ $|\bar{y}_A - \bar{y}_B| > LSD$ $1.325 > 0.44$

Decision: Reject Ho Conclusion: There is sufficient evidence to conclude Ha

5) a) Price - independent variable - X Qty Sold - Dependent variable = Y

Cruise Price (x)	Qty (Sold)	x^2	y^2	xy
58.7	20	3445	400	1174
59	15	3481	225	885
60.1	17	3612	289	1021.7
61.3	16	3757	256	980.8
63.2	13	3994	169	821.6
64	11	4096	121	704

b) Hypothesis $H_0: \beta_1 = 0$ $H_a: \beta_1 \neq 0$ TS: $t_{obs} = \frac{\hat\beta_1 - \beta_{10}}{S_\varepsilon / \sqrt{S_{xx}}}$

$S_{xx} = \sum_{i=1}(x_i - \bar{x})^2 \Rightarrow \sum x_i^2 - \frac{(\sum x_i)^2}{n} = 22385 - 22362 = 23$

$\sum y_i = 92$, $\sum y_i^2 = 1460$ $S_{yy} = 1460 - 1410$

$\frac{(\sum y_i)^2}{n} = 1410$ $S_{yy} = 50$

$S_{xy} = \sum_{i=1}^{n} x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n} = 5588 - \frac{366.3(92)}{6} = 5588 - 5616.6 = -28.6$

$S_\varepsilon = \sqrt{\frac{SSE}{n-2}}$ $SSE = S_{yy} - \frac{S_{xy}^2}{S_{xx}} = 50 - \frac{(-28.6)^2}{23} = 14.44$

$= \sqrt{\frac{14.44}{4}} = \sqrt{3.61} = 1.9$

TS: $t_{obs} = \frac{\hat\beta_1 - \beta_{10}}{S_\varepsilon/\sqrt{S_{xx}}} = \frac{-1.24 - 0}{1.9/\sqrt{23}} = \frac{-1.24-0}{1.9/4.79} = -3.17$

RR: $|t_{obs}| > t_{\alpha/2, n-2} \Rightarrow |t_{obs}| > t_{\alpha/2,n-2} \Rightarrow |t_{obs}| > 2.776$

Decision: $3.176 > 2.776$, Reject Ho

Conclusion: There is sufficient evidence to conclude Ha

6)

State	n	Mean	SD
Michigan	10	8458	250
California	9	9690	420

a) We need to do F test to check pooled t test / separate t test

Hypothesis $H_0: \sigma_1^2 = \sigma_2^2$ $H_a: \sigma_1^2 \neq \sigma_2^2$ TS: $F_{obs} = \frac{S_i^2}{S_j^2} = \frac{420^2}{250^2} = 2.82$

RR: $F_{obs} > F_{\alpha/2, df_1, df_2}$ $df_1 = n_2 - 1 = 8$ $df_2 = n_1 - 1 = 9$

$F_{obs} > F_{0.025,8,9} \Rightarrow F_{obs} > 4.10$

Decision: $2.82 > 4.10$ is false So we FTR Ho

Conclusion: No sufficient evidence to conclude Ha

We conduct pooled t test

b) $C > M$ $\mu_2 > \mu_1$ $\mu_2 - \mu_1 > 0$ μ_c = cali mean, μ_m - Mich mean, μ_0 - Diff in mean $\mu_c - \mu_M$

Hypothesis - $H_0: \mu_c - \mu_M = 0$ $H_a: \mu_c - \mu_M > 0$

Test Statistic - $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{9(250)^2 + 8(420)^2}{17}$

$\frac{562500 + 1411200}{17} = 116100$. $S_p = \sqrt{116100} = 340.7$

$t_{obs} = \frac{9690 - 8458 - 0}{341\sqrt{\frac{1}{10}+\frac{1}{9}}} = \frac{1232}{341(0.45)} = 8.02$

RR: $t_{obs} > t_{n_1+n_2-2,\alpha}$ $t_{obs} > t_{17,0.05}$ $t_{obs} > 1.74$

Pvalue $P(t_{17} > 8.02)$ less than 0.005 p value $< \alpha$

Decision: Reject Ho Conclusion: There is sufficient evidence to conclude Ha

i.e., california has higher mean [illegible] wrt than michigan

c) $CI = (\bar{y}_1 - \bar{y}_2) \pm t_{\alpha/2, n_1+n_2-2}(S_p)\left(\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right) = (9690 - 8455) \pm t_{0.025,17}(340.7)(0.45)$

$= 1235 \pm 2.11(340.7)(0.45) = 1235 \pm 323.7 = (911.3, 1558.7)$

2S1 ① $H_0: \mu_d = \mu_{before} - \mu_{after} = 0$ $H_a: \mu_d = \mu_{before} - \mu_{after} > 0$

Draw table with new column Difference (di)

Paired design with patients being extraneous variable

$\bar{d} = \frac{6.75+11.76+7.45+6.68+2.21}{5} = 6.97$ $s_d^2 = \frac{\Sigma(d_i-\bar{d})^2}{n-1} = 11.49$

Test Statistic $t_{obs} = \frac{\bar{d} - D_0}{s_d/\sqrt{n}} = \frac{6.97-0}{3.39/\sqrt{5}} = 4.58$ $s_d = 3.39$

RR: $t_{obs} > t_{n-1,\alpha}$. $t_{n-1,\alpha} = t_{4,0.05} = 2.132$, $t_{obs} > 2.132$

Decision: Reject Ho because 4.58 > 2.132

Conclusion: We have enough evidence to conclude that mean salt sensitivity value decreased after the patient received antihypertensive treatment i.e., μ_{before} is greater than μ_{after}

②a)

Brand	Sample Size	Sample mean	Sample SD
Low Tar	100	8	0.3
A	100	10	0.4
B	100	10	0.4
C	100	11	0.5

$H_0: \mu_{Low} = \mu_A = \mu_B = \mu_C$ Ha: Not all the above are equal

As all sample sizes are equal, we can conduct a-f method

Test Statistic: $F_{obs} = \frac{s_B^2}{s_W^2}$. $s_W^2 = \frac{s_{Low}^2 + s_A^2 + s_B^2 + s_C^2}{t} = \frac{0.3^2+0.4^2+0.4^2+0.5^2}{4}$

$s_B^2 = n \cdot$ variance of sample means $s_W^2 = 0.165$

$= 100 \cdot 1.5833 = 158.33$

$F_{obs} = \frac{s_B^2}{s_W^2} = \frac{158.33}{0.165} = 959.59$. RR: $F_{obs} > F_{df_1, df_2, \alpha}$ $df_1 = t-1 = 3$ $df_2 = nt-t = 100(4)-4 = 396$

$F_{obs} > F_{3,396,0.05} =$

$F_{obs} > 2.6$

Decision: Reject Ho

Conclusion: We have sufficient evidence to conclude that there is significant difference in average tar content from the above four brands of cigarettes.

b.) Check for equal variance condition

$H_0: \sigma_{Low}^2 = \sigma_A^2 = \sigma_B^2 = \sigma_C^2$. Ha: Not all the above are equal

Test Statistic: $F_{max} = \frac{s_{max}^2}{s_{min}^2}$ $F_{max} = \frac{(0.5)^2}{(0.3)^2} = 2.78$

Rejection Region: $F_{max} > F_{max, df_1, df_2, \alpha}$ $df_1 = t = 4$ $df_2 = n-1 = 99$

$F_{max} > F_{max\,4,99,0.05}$

$F_{max} > 1$

Decision: Reject Ho. Conclusion: We have sufficient evidence to conclude that not all variances are equal. Hence the condition of equal population variances for test in Part(a) is violated.

3.) a) Randomized Block Design

b)

Gasoline	Car Model 1	2	3	4	Sum
A	15	33	13	29.2	90.2 $y_{1.}$
B	16.3	26.4	19.1	22.5	84.3 $y_{2.}$
C	10.5	31.5	17.5	30.1	89.6 $y_{3.}$
D	14.0	34.5	19.7	21.6	89.8 $y_{4.}$
Sum	55.8 $y_{.1}$	125.4 $y_{.2}$	69.3 $y_{.3}$	103.4 $y_{.4}$	353.9 $y_{..}$

$TSS = \sum_{ij} y_{ij}^2 - \frac{y_{..}^2}{bt} = 897.18$ (given) [∵ t=4, b=4]

$SST = \left(\sum_i \frac{y_{i.}^2}{b}\right) - \frac{y_{..}^2}{bt} = 7833.68 - 7827.82 = 5.86$

$SSB = \left(\sum_j \frac{y_{.j}^2}{t}\right) - \frac{y_{..}^2}{bt} = 8583.21 - 7827.82 = 755.39$

$SSE = TSS - SST - SSB = 135.93$

$df_1 = t-1 = 3$ $df_2 = b-1 = 3$. $MST = \frac{SST}{t-1} = \frac{5.86}{3} = 1.95$

$MSB = \frac{SSB}{b-1} = \frac{755.39}{3} = 251.8$ $MSE = \frac{SSE}{(b-1)(t-1)} = \frac{135.93}{9} = 15.1$

$F^{①}_{obs} = \frac{MST}{MSE} = 0.13$ $F^{②}_{obs} = \frac{MSB}{MSE} = 16.67$

ANOVA for Randomized Block Design

Source due to	Sum of Squares (SS)	df	Mean Square (MS)	F
Treatments	5.86	3	1.95	0.13
Blocks	755.39	3	251.8	16.67
Error	135.93	9	15.1	

c) $H_0: \alpha_A = \alpha_B = \alpha_C = \alpha_D = 0$ | Ha: Not all above are equal to zero

Test Statistic: $F_{obs} = MST/MSE = 0.13$

Rejection Region: $F_{obs} > F_{t-1,(b-1)(t-1),\alpha}$

P-value $= P(F_{3,9,\cancel{0.05}} > F_{obs}) = P(F_{3,9} > 0.13)$

P value > 0.25

α > Pvalue is False

Decision: FTR Ho

Conclusion: We don't have sufficient evidence to conclude Ha

d) $H_0: \beta_A = \beta_B = \beta_C = \beta_D = 0$ | Ha: Not all above are equal to zero

Test Statistic: $F_{obs} = MSB/MSE = 16.67$

Rejection Region: $F_{obs} > F_{b-1,(t-1)(b-1),\alpha}$

Pvalue $= P(F_{3,9,0.05} > F_{obs})$ $P(F_{3,9,0.05} > 16.67)$

Pvalue < 0.001

α > p-value is true

Decision: Reject Ho. We have sufficient evidence to conclude Ha

e) Relative Efficiency $= \frac{MSE_{CR}}{MSE_{RB}} = \frac{(b-1)MSB + b(t-1)MSE_{RB}}{(bt-1)MSE_{RB}} = \frac{3(251.80)+4(3)(15.1)}{15(15.1)} = 4.135$

4)

					Sum
A	2.2	3	2.7	2.7	10.6
B	3.6	3.9	4.1	4.3	15.9
C	4.3	4.4	4.5	4.1	17.3
Sum	10.1	11.3	11.3	11.1	43.8

a) Completely Randomized Design

b) $H_0: \mu_A = \mu_B = \mu_C$ Ha: Not all of the above are equal

Test Statistic: $F_{obs} = \frac{s_B^2}{s_W^2}$. $TSS = \sum_{ij} y_{ij}^2 - \frac{y_{..}^2}{n_T} = 6.93$ (given) [∵ $n_T = 4+4+4 = 12$]

$SSB = \sum_i \frac{y_{i.}^2}{n_i} - \frac{y_{..}^2}{n_T} = 166.115 - 159.87 = 6.245$

$SSW = TSS - SSB = 0.685$ | $s_B^2 = \frac{SSB}{t-1} = \frac{6.245}{3-1} = 3.12$

$s_W^2 = \frac{SSW}{n_T - t} = \frac{0.685}{9} = 0.0076$

$F_{obs} = \frac{s_B^2}{s_W^2} = 41.05$

ANOVA for Completely Randomized design

Source due to	Sum of Squares (SS)	df	Mean Square (MS)	F
Between samples	6.245	2	3.12	41.05
Within Samples	0.685	9	0.076	
Total	6.93	11		

c) $H_0: \mu_A = \mu_B = \mu_C$ Ha: Not all the above are equal

Test Statistic: $F_{obs} = 41.05$ | Rejection Region $F_{obs} > F_{df_1, df_2, \alpha}$

$F_{df_1, df_2, \alpha} = F_{2,9,0.05} = 4.26$ $F_{obs} > 4.26$. $df_1 = t-1 = 2$ $df_2 = n_T - t = 9$

P value $= P(F_{2,9} > 41.05)$. P value < 0.0001

α > pvalue is true.

Decision: Reject Ho. We have sufficient evidence to conclude Ha

5.) a)

Price (X)	Quantity (Y)
58.7	20
59	15
60.1	17
61.3	16
63.2	13
64	11

Independent variable (X) → Price

Dependent variable (Y) → Quantity

b) $H_0: \beta_1 = 0 \quad H_a: \beta_1 \neq 0$

Test Statistic $t_{obs} = \dfrac{\hat{\beta}_1 - \beta_{10}}{S_\varepsilon / \sqrt{S_{xx}}}$ (∵ given $S_{xx} = 24$, $S_{xy} = -29.5$, $S_\varepsilon = 1.8$)

$\hat{\beta}_1 = \dfrac{S_{xy}}{S_{xx}} = \dfrac{-29.5}{24} = -1.23 \quad | \quad t_{obs} = \dfrac{-1.23 - 0}{1.8/\sqrt{24}} = -3.35$

Rejection Region : $|t_{obs}| > t_{n-2, \frac{\alpha}{2}}$

$t_{4, 0.025} = 2.776 \quad |t_{obs}| > t_{4, 0.025}$

∴ Reject Ho. We have sufficient evidence to say that independent variable is useful to predict dependent variable as slope is not equal to 0

6) a) Check the variance, in order to choose between pooled t-test and separate t-test

$H_0: \sigma_1^2 = \sigma_2^2 \quad H_a: \sigma_1^2 \neq \sigma_2^2$

Test Statistic : $F_{obs} = \dfrac{S_i^2}{S_j^2} = \dfrac{(480)^2}{(350)^2} = 1.88$

$S_1 < S_2$

$df_1 = n_2 - 1 = 8 \quad df2 = n_1 - 1 = 9$

Rejection Region : $F_{obs} > F_{df_1, df_2, \alpha/2}$ or $< \dfrac{1}{F_{df_2, df_1, \alpha/2}}$

$F_{8,9,0.025} = 4.10$

$\dfrac{1}{F_{9,8,0.025}} = \dfrac{1}{4.36} = 0.23$

$F_{obs} > F_{8,9,0.25}$ (False) or $F_{obs} < \dfrac{1}{F_{9,8,0.025}}$ (False)

Failed to Reject Ho. We conduct pooled t test

b) $H_0: \mu_C - \mu_M = 0 \quad H_a: \mu_C - \mu_M > 0$ ① - California ② - Michigan

Test Statistic ; $t_{obs} = \dfrac{(\bar{y}_1 - \bar{y}_2) - D_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

$S_p^2 = \dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2} = \dfrac{8(480)^2 + 9(350)^2}{17} = 173276.47$

$S_p = 416.26$

$t_{obs} = \dfrac{(8540 - 7358) - 0}{416.26\sqrt{\frac{1}{9} + \frac{1}{10}}} = \dfrac{1182}{192.26} = 6.18$

Rejection Region : $t_{obs} > t_{n_1+n_2-2, \alpha} \quad t_{17, 0.05} = 1.74$

∴ $t_{obs} > 1.74$

Decision : We Reject Ho. We have sufficient evidence to conclude that california has higher mean hysterectomy cost than michigan

7)

School/Rating	Outstanding	Average	Poor	
Undesirable	10	4	2	$R_1 = 16$
Adequate	12	8	2	$R_2 = 22$
	$C_1 = 22$	$C_2 = 12$	$C_3 = 4$	

$R_1 = 16 \quad R_2 = 22$

$C_1 = 22 \quad C_2 = 12 \quad C_3 = 4$

Total = 38

X = School type

Y = Rating

Ho: X and Y are not contingent

Ha: X and Y are contingent

$E_{11} = \dfrac{R_1 C_1}{n} = \dfrac{16(22)}{38} = 9.26$

$E_{12} = \dfrac{R_1 C_2}{n} = \dfrac{16(12)}{38} = 5.05$

$E_{13} = \dfrac{R_1 C_3}{n} = \dfrac{16(4)}{38} = 1.68$

$E_{21} = \dfrac{R_2 C_1}{n} = \dfrac{22(22)}{38} = 12.74$

$E_{22} = \dfrac{R_2 C_2}{n} = \dfrac{22(12)}{38} = 6.95$

$E_{23} = \dfrac{R_2 C_3}{n} = \dfrac{22(4)}{38} = 2.31$

Test Statistic : $\chi^2_{obs} = \sum_{i=1}^{r} \sum_{j=1}^{c} \dfrac{(n_{ij} - E_{ij})^2}{E_{ij}}$

$\chi^2_{obs} = \dfrac{(10-9.26)^2}{9.26} + \dfrac{(4-5.05)^2}{5.05} + \dfrac{(2-1.68)^2}{1.68} + \dfrac{(12-12.74)^2}{12.74} + \dfrac{(8-6.95)^2}{6.95} + \dfrac{(2-2.31)^2}{2.31}$

$\chi^2_{obs} = 0.06 + 0.22 + 0.06 + 0.04 + 0.16 + 0.04$

$\chi^2_{obs} = 0.58$

Rejection Region ; $\chi^2_{obs} > \chi^2_{(R-1)(C-1), \alpha}$

$\chi^2_{obs} > \chi^2_{(1)(2), 0.05}$

$\chi^2_{obs} > 5.99$

$\chi^2_{obs} > \chi^2_{2, 0.05}$ (False)

Fail to Reject Ho.

There is no sufficient evidence to conclude that Rating is contingent on School type

—X—

Put down Prediction equation

Given $S_{xx} = 20 \quad S_{xy} = 60$

$\hat{\beta} = \dfrac{S_{xy}}{S_{xx}} = \dfrac{60}{20} = 3$

$\bar{y} = \beta_0 + \hat{\beta}_1 \bar{x}$

$14 = \beta_0 + (3)(4)$

$\beta_0 = 2$

Equation is $\bar{y} = 2 + 3x$

1. Reject Ho
2. a) Reject Ho
   b) Reject Ho
3. c) FTR Ho
   d) FTR Ho
   e) RE = 1.88
4. c) Reject Ho
5. b) $\hat{y} = 2 + 3x$
   c) FTR Ho
6. a) FTR Ho
   b) Reject Ho
7. Reject Ho