Stat 543
4-30-15

Analysis of Variance (ANOVA)

Compare the means of several populations (1)

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

H_1: the population means are not all equal

Procedure: decompose the total sum of squares into 2 parts < Treatment, Error

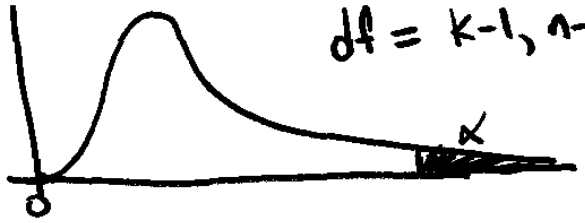
(2)

ANOVA table (SS: sum of squares; df: degrees of freedom; MS: mean square)

Source	SS	df	MS	F
TRT	SS_{TRT}	$K-1$	$\frac{SS_{TRT}}{K-1}$	$\frac{MS_{TRT}}{MSE}$
ERR	SS_{ERR}	$n-k$	$\frac{SSE}{n-k}$	—
TOT	SS_{TOT}	$n-1$	—	—

Compare the F test stat to the F table

$df = k-1, n-k$



(3)

In any ANOVA, when H_0 is rejected,

a post hoc analysis is required

The post hoc analysis consists of pairwise comparisons.

Fisher's Least Significant Difference (LSD) (overly liberal)

Bonferroni method (overly conservative)

Tukey's method - moderate

(4)

Underline method for summarizing the post-hoc comparisons.

Write the sample means in increasing order

①	②	③	
36.75	46.58	60.25	Tukey

①	②	③	
36.75	46.58	60.25	Fisher

HW #5 follows this page

15. Concrete As concrete cures, it gains strength. The following data represent the 7-day and 28-day strength (in pounds per square inch) of a certain type of concrete:

7-Day Strength, x	28-Day Strength, y	7-Day Strength, x	28-Day Strength, y
2,300	4,070	2,480	4,120
3,390	5,220	3,380	5,020
2,430	4,640	2,660	4,890
2,890	4,620	2,620	4,190
3,330	4,850	3,340	4,630

(a) Treating the 7-day strength as the explanatory variable, x, determine the estimates of β_0 and β_1.

(b) What is the estimated mean 28-day strength of this concrete if the 7-day strength is 3,000 psi?

18. Heat Index A researcher wanted to determine whether there was a linear relation among heat index, air temperature, and dew point. The following data show the heat index, air temperature (in degrees Fahrenheit), and dew point for various days.

Air Temp.	Dew Point	Heat Index
90	64	93
90	68	95
94	66	99
94	70	102
96	70	105
96	76	111
99	68	107
99	72	111
100	74	114
100	80	123
93	72	103
93	78	109
97	80	118
92	82	114
95	66	100
95	82	118

Find the least-squares regression equation $\hat{y} = b_0 + b_1x_1 + b_2x_2$, where x_1 is air temperature, x_2 is dew point, and y is the response variable "heat index."

	turnout	type
1	33	1
2	78	1
3	32	1
4	28	1
5	10	1
6	12	1
7	61	1
8	28	1
9	29	1
10	45	1
11	44	1
12	41	1
13	35	2
14	56	2
15	35	2
16	40	2
17	45	2
18	42	2
19	65	2
20	62	2
21	25	2
22	47	2
23	52	2
24	55	2
25	42	3
26	40	3
27	52	3
28	66	3
29	78	3
30	62	3
31	57	3
32	75	3
33	72	3
34	51	3
35	69	3
36	59	3

Oneway

ANOVA

TURNOUT

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3342.889	2	1671.444	7.592	.002
Within Groups	7265.417	33	220.164		
Total	10608.306	35			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: TURNOUT

	(I) TYPE	(J) TYPE	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	1	2	-9.83	6.058	.250	-24.70	5.03
		3	-23.50*	6.058	.001	-38.36	-8.64
	2	1	9.83	6.058	.250	-5.03	24.70
		3	-13.67	6.058	.077	-28.53	1.20
	3	1	23.50*	6.058	.001	8.64	38.36
		2	13.67	6.058	.077	-1.20	28.53
LSD	1	2	-9.83	6.058	.114	-22.16	2.49
		3	-23.50*	6.058	.000	-35.82	-11.18
	2	1	9.83	6.058	.114	-2.49	22.16
		3	-13.67*	6.058	.031	-25.99	-1.34
	3	1	23.50*	6.058	.000	11.18	35.82
		2	13.67*	6.058	.031	1.34	25.99

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

TURNOUT

	TYPE	N	Subset for alpha = .05	
			1	2
Tukey HSD[a]	1	12	36.75	
	2	12	46.58	46.58
	3	12		60.25
	Sig.		.250	.077

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 12.000.

Results for: Sheet1

One-way ANOVA: turnout versus type

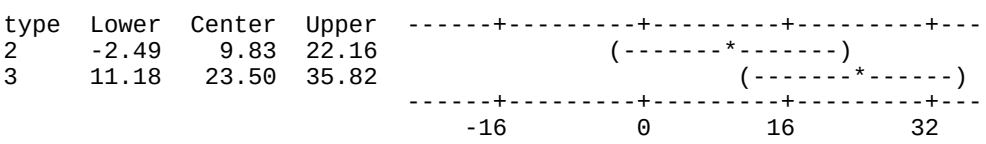
```
Source  DF     SS    MS     F      P
type     2   3343  1671  7.59  0.002
Error   33   7265   220
Total   35  10608

S = 14.84   R-Sq = 31.51%   R-Sq(adj) = 27.36%


                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -------+---------+---------+---------+--
1      12  36.75  19.08  (-------*------)
2      12  46.58  11.94           (------*------)
3      12  60.25  12.40                    (------*------)
                          -------+---------+---------+---------+--
                                36        48        60        72

Pooled StDev = 14.84


Grouping Information Using Tukey Method

type   N   Mean  Grouping
3     12  60.25  A
2     12  46.58  A B
1     12  36.75    B

Means that do not share a letter are significantly different.


Tukey 95% Simultaneous Confidence Intervals
All Pairwise Comparisons among Levels of type

Individual confidence level = 98.04%


type = 1 subtracted from:

type  Lower  Center  Upper  ----+---------+---------+---------+-----
2     -5.03    9.83  24.70             (-------*------)
3      8.64   23.50  38.36                     (-------*------)
                           ----+---------+---------+---------+-----
                             -20         0        20        40


type = 2 subtracted from:

type  Lower  Center  Upper  ----+---------+---------+---------+-----
3     -1.20   13.67  28.53               (-------*------)
                           ----+---------+---------+---------+-----
                             -20         0        20        40


Grouping Information Using Fisher Method

type   N   Mean  Grouping
3     12  60.25  A
2     12  46.58    B
1     12  36.75    B
```

Means that do not share a letter are significantly different.

Fisher 95% Individual Confidence Intervals
All Pairwise Comparisons among Levels of type

Simultaneous confidence level = 88.02%

type = 1 subtracted from:



type = 2 subtracted from: