PAGE:
DATE: / /

ANOVA (Analysis of Variance)

Single factor analysis

$H_0 : \mu_1 = \mu_2 = \cdots = \mu_K$

against

H_a : At least two of the μ's are different.

ANOVA Notation

K = number of populations or treatments being compared.

Population or treatment	1	2	...	K
Population mean	μ_1	μ_2	...	μ_K
Population variance	σ_1^2	σ_2^2	...	σ_K^2
Sample size	n_1	n_2	...	n_K
Sample mean	$\bar{x}_1$	$\bar{x}_2$	...	$\bar{x}_n$
Sample variance	s_1^2	s_2^2	...	s_n^2

$N = n_1 + n_2 + \cdots + n_K$ (total no of obs. in data)

T = grand total $= n_1\bar{x}_1 + n_2\bar{x}_2 + \cdots + n_K\bar{x}_K$

$\bar{\bar{x}}$ = grand mean $= \frac{T}{N}$

Assumptions for ANOVA

1. Each of the k population response distribution is normal.
2. $\sigma_1 = \sigma_2 = \dots = \sigma_k$ (The k normal distributions have identical standard deviation)
3. The observations are independent of one another.

★ Definition :

A measure of disparity among the sample means is the treatment sum of squares denoted by SSTr given by

$$SSTr = n_1(\bar{x}_1 - \bar{\bar{x}})^2 + n_2(\bar{x}_2 - \bar{\bar{x}})^2 + \dots + n_k(\bar{x}_k - \bar{\bar{x}}_k)^2$$

A measure of variation within k samples called error sum of squares

$$SSE = (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + (n_k - 1)s_k^2$$

treatment df = k-1   error df = N - k

A mean square is a sum of squares divided by its df.
In particular

mean square for treatments = $MSTr = \dfrac{SSTr}{k-1}$

mean square for error = $MSE = \dfrac{SSE}{N-k}$

F-test $F = \frac{MSTr}{MSE}$

PAGE:
DATE: / /

		Sample size	Sample Mean	Sample SD
1	N	35	10.89	0..
2	S	35	11.25	
3	U	35	11.37	
4	M	35	11.75	

	df	Sum of Sq.	Mean Sq.	F
Treatments	$k-1$	SSTr	$MSTr = \frac{SSTr}{k-1}$	$\frac{MSTr}{MSE}$
Error	$N-k$	SSE	$MSE = \frac{SSE}{N-k}$	
Total	$N-1$	SST_o		

$N = 33$ $\alpha = 0.01$

$\bar{x} = 73.30$

PAGE :
DATE : / /

Treatment		Treatment 2	Control
71	75	77	81
72	73	67	79
~~8~~75	72	79	73
80	65	78	71
60	63	81	75
65	69	72	84
63	64	71	77
78	71	84	67
		91	

Σx_i	1116		607
$\bar{x}$	69.75	$\bar{y} = 77.78$	$\bar{z}$ 75.875
	517	421.56	214.88
S_1	34.47		

$$SSTr = 16(69.75 - 73.30)^2 + 9(77.78 - 73.3)^2 + 8(75.875 - 73.30)^2$$

$$= 435.52$$

$$SSE = 1153.44$$

$$\boxed{F = \frac{217.76}{38.448}} = \boxed{5.66}$$

$$MSTr = \frac{SSTr}{K-1} = \frac{435.52}{2} = 217.76$$

$$MSE = \frac{SSE}{N-K} = \frac{1153.44}{30} = 38.448$$

(5.39, ∞)

$F_{(N-K,\ K-1)}$

$m = N-K \quad n = K-1$

$3?\ ,\ 2$

H_0 is rejected if $F \geq F_\alpha$

* Sign test is a non-parametric

No ↓

$n \leq 25$ —No→ Convert the first statistic x

$$z = \frac{(x+0.5) - (n/2)}{\sqrt{n/2}}$$

↓ Find the critical z value from table.

Yes ↓

Find the critical value from Table

↓

if $x \leq C$ → Yes Reject null hypothesis

↓ No

~~reject null~~

Fail to reject null

Notation for sign test

PAGE:
DATE: / /

x = the number of times the less frequent sign occurs

n = the total number of positive and negative signs combined.

For $n \leq 25$ x

$n > 25$ $z = \dfrac{(x+0.5) - n/2}{\frac{\sqrt{n}}{2}}$ $\dfrac{2x+1-n}{\sqrt{n}}$

Critical values

$n \leq 25$. critical x table A-7

$n > 25$ critical z A2

If two sets has equal median,
number of +ve $\approx$ number of -ve

Rege

PAGE: A
DATE: / /

$\alpha = 0.00$

g n

Regular	kiln - dried		
1903	2009	106	-
1935	1915	-20	+
1910	2011	101	-
2496	2463	-33	+
2108	2180	72	-
1961	1925	-36	+
2060	2122	62	-
1444	1482	38	-
1612	1542	-70	+
1316	1443	127	-
1551	1535	24	-

+ 4

− 7

1

~~0.2744~~

Fail to Reject.

Out of 325 types of corps, 295 are one type. Use the sign test at 0.05 significance level to test the claim.

$$z = \frac{60 + 1 - 325}{\sqrt{325}} = -14.644$$

Test hypothesis

A paired t-test

☆ A paired t-test is used to compare two population means where you have two samples in which observations in one sample can be paired with other sample.

x = test score before the module
y = test " after the "

null hypothesis : $\mu_1 = \mu_2$

1. Calculate the difference ($d_i = y_i - x_i$) between the two observations on each pair making sure you distinguish between positive and negative difference

mean difference $\bar{d}$

$$SE(\bar{d}) = \frac{S_d}{\sqrt{n}}$$

$$T = \frac{\bar{d}}{SE(\bar{d})}$$ $n-1$ degree of freedom

p-value > 0.05 not signdificant, don't reject null hypo.
p-value < 0.05 Significant. reject null hypo
p-value < 0.01 high significant, reject null hypo.

$0 < p < 1$

	pre	post	d_i		
1	18	22	4	1.95	1.05
2	21	25	4	1.95	1.05
3	16	17	1	-1.05	-1.95
4	22	24	2	-0.05	-0.95
5	19	16	-3	-5.05	-5.95
6	24	29	5	2.95	2.05
7	17	20	3	0.95	0.05
8	21	23	2	-0.05	-0.95
9	23	19	-4	-6.05	-6.95
10	18	20	2	-0.05	-0.95
11	14	15	1	-1.05	-1.95
12	16	15	-1	-3.05	-3.95
13	16	18	2	-0.05	-0.95
14	19	26	7	4.95	4.05
15	18	18	0	-2.05	-2.95
16	20	24	4	1.95	1.05
17	12	18	6	3.95	3.05
18	22	25	3	0.95	0.05
19	15	19	4	1.95	1.05
20	17	16	-1	-3.05	-3.95

$\frac{59}{20}$ 2.95

$\frac{41}{20}$

$T = \frac{2.05}{1.8}$ $\quad \bar{d}$ 2.05 $\quad sd = \frac{152.95}{8.05}$

$= 1.138$

$= 1.89$

$SDE = \frac{8.05}{\sqrt{20}} = 1.8$

PAGE :
DATE : / /

$\alpha = 0.05$

0 mg	50 mg	100 mg			
9	7	4	0.86	0.29	1
8	6	3	-0.14	0.71	0
7	6	2	~~-0.86~~ 1.44	0.71	-1
8	7	3	0.14	0.29	0
8	8	4	0.14	1.29	1
9	7	3	0.86	0.29	0
8	6	2	0.14	0.71	1
Σx_i 57	47	21	Σ 2.30	3.438	4
$\tilde{x}$ 8.14	6.71	3	2.297		

~~3.58~~ + 3.43 + 4

2.86 + 3.43 + 4

$$\tilde{\tilde{x}} = \frac{57+47+21}{\cancel{18}\ 21} = \frac{125}{\cancel{18}\ 21} = 5.95$$

SSE = 2.30 + 3.43 + 4 = ~~9.73~~ ~~11.07~~ 10.29

$$SST_s = 7\left[(8.14-5.95)^2 + (6.71-5.95)^2 + (3-5.95)^2\right]$$
$$= 7\left[4.80 + 0.58 + 8.70\right]$$
$$= 98.56$$

$$F = \frac{98.56/2}{9.73/18} = 91.17 \qquad \frac{49.28}{\cancel{0.54}\ 0.57}$$

0.57

$F_\alpha = 3.55$ Here $F >_a F_\alpha$

$\Rightarrow H_0$ is seelected