Def^n: Mean Square between $MSB = \frac{SSB}{K-1}$

Def^n: Mean Square within (or Mean Square Error) $MSW = \frac{SSW}{n-K}$
sometimes it is denoted by MSE.

In one way ANOVA, the appropriate test statistic is
$F = \frac{MSB}{MSW}$ with d.f. (K-1, n-K)

When Ho is true, SSB is about the same as SSW because all variations simply from random fluctuations. But when the groups are systematically different from one another, SSB tends to be large, relative to variation within the groups. The larger the between group variations, the greater is the likelihood that the samples do NOT come from populations with equal means.

e.g.

GP=1	GP=2	GP=3	
0	1	5	
6	4	6	
2	3	10	
4	2	8	$\bar{x}=4.0$
3	0	6	
$\bar{x}_1=3.0$	$\bar{x}_2=2.0$	$\bar{x}_3=7.0$	

$SSW = (0-3)^2+(6-3)^2+(2-3)^2+(4-3)^2+(3-3)^2+(1-2)^2+(4-2)^2+(3-2)^2$
$+(2-2)^2+(0-2)^2+(7-5)^2+(7-6)^2+(10-7)^2+(8-7)^2+(6-7)^2$
$=9+9+1+1+0+1+4+1+0+4+4+1+9+1+1=46$

$SSB = 5\times(3-4)^2+5\times(2-4)^2+5\times(7-4)^2 = 5+20+45=70$

$F = \frac{MSB}{MSW} = \frac{(\frac{70}{2})}{(46/...)} = \frac{35}{3.833} = 9.1289$ d.f. (2, 12)

6.36

$H_0: \mu_1 = \mu_2 = \mu_3$
H_A: at least two of means differ

From F-table, we have

$F_{0.05}(2, 12) = 3.89$

$F_{calc}(2, 12) = 9.1289 > 3.89$ reject H_0 at 0.05 level.

6.38

Multiple Comparisons

Dependent Variable: stress

	(I) gp Therapy Group	(J) gp Therapy Group	Mean Difference (I-J)	Std. Error	Sig.
Tukey HSD	1.00 Music Therapy	2.00 Relaxation Therapy	1.00000	1.23828	.706
		3.00 Controls	-4.00000*	1.23828	.018
	2.00 Relaxation Therapy	1.00 Music Therapy	-1.00000	1.23828	.706
		3.00 Controls	-5.00000*	1.23828	.004
	3.00 Controls	1.00 Music Therapy	4.00000*	1.23828	.018
		2.00 Relaxation Therapy	5.00000*	1.23828	.004
Scheffe	1.00 Music Therapy	2.00 Relaxation Therapy	1.00000	1.23828	.728
		3.00 Controls	-4.00000*	1.23828	.023
	2.00 Relaxation Therapy	1.00 Music Therapy	-1.00000	1.23828	.728
		3.00 Controls	-5.00000*	1.23828	.006
	3.00 Controls	1.00 Music Therapy	4.00000*	1.23828	.023
		2.00 Relaxation Therapy	5.00000*	1.23828	.006
LSD	1.00 Music Therapy	2.00 Relaxation Therapy	1.00000	1.23828	.435
		3.00 Controls	-4.00000*	1.23828	.007
	2.00 Relaxation Therapy	1.00 Music Therapy	-1.00000	1.23828	.435
		3.00 Controls	-5.00000*	1.23828	.002
	3.00 Controls	1.00 Music Therapy	4.00000*	1.23828	.007
		2.00 Relaxation Therapy	5.00000*	1.23828	.002

Multiple Comparisons

Dependent Variable: stress

	(I) gp Therapy Group	(J) gp Therapy Group	95% Confidence Interval Lower Bound	Upper Bound
Tukey HSD	1.00 Music Therapy	2.00 Relaxation Therapy	-2.3036	4.3036
		3.00 Controls	-7.3036	-.6964
	2.00 Relaxation Therapy	1.00 Music Therapy	-4.3036	2.3036
		3.00 Controls	-8.3036	-1.6964
	3.00 Controls	1.00 Music Therapy	.6964	7.3036
		2.00 Relaxation Therapy	1.6964	8.3036
Scheffe	1.00 Music Therapy	2.00 Relaxation Therapy	-2.4518	4.4518
		3.00 Controls	-7.4518	-.5482
	2.00 Relaxation Therapy	1.00 Music Therapy	-4.4518	2.4518
		3.00 Controls	-8.4518	-1.5482
	3.00 Controls	1.00 Music Therapy	.5482	7.4518
		2.00 Relaxation Therapy	1.5482	8.4518
LSD	1.00 Music Therapy	2.00 Relaxation Therapy	-1.6980	3.6980
		3.00 Controls	-6.6980	-1.3020
	2.00 Relaxation Therapy	1.00 Music Therapy	-3.6980	1.6980
		3.00 Controls	-7.6980	-2.3020
	3.00 Controls	1.00 Music Therapy	1.3020	6.6980
		2.00 Relaxation Therapy	2.3020	7.6980

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

Homogeneous Subsets

stress

	gp Therapy Group	N	Subset for alpha = .05 1	2
Student-Newman-Keuls[a]	2.00 Relaxation Therapy	5	2.0000	
	1.00 Music Therapy	5	3.0000	
	3.00 Controls	5		7.0000
	Sig.		.435	1.000
Tukey HSD[a]	2.00 Relaxation Therapy	5	2.0000	
	1.00 Music Therapy	5	3.0000	
	3.00 Controls	5		7.0000
	Sig.		.706	1.000
Scheffe[a]	2.00 Relaxation Therapy	5	2.0000	
	1.00 Music Therapy	5	3.0000	
	3.00 Controls	5		7.0000
	Sig.		.728	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 5.000.

Assumptions of oneway ANOVA:
(i) The populations have normal distribution.

(ii) Homogeneity of variance

(iii) Independence random samples

ANOVA is robust, operating well even with considerable heterogeneity of variances as long as all n_i are equal or nearly equal (Glass, Peckham & Sanders, 1972). If the n_i are quite different, then the probability of a Type I error will depart markedly from α, to a degree dependent on the magnitude of the heterogeneity (Box, 1954); if larger variances are associated with the larger samples, the prob. of a type I error will be $< \alpha$, and if they are associated with the smaller samples this prob. will be $> \alpha$ (Kohr & Games 1974; Maxwell & Delaney, 1990: 723-724; see also Horsnell, 1953). The ANOVA is also robust with respect to the assumption of the underlying populations' normality. The validity of the analysis is affected only slightly by even considerable deviations from normality (in skewness and/or kurtosis), especially as n increases (Box and Anderson, 1955; Srivastava, 1959; Tiku 1971).

references:

① Glass, G.V., P.D. Peckham & J.R. Sanders 1972 Consequences of failure to meet assumptions underlying the fixed effects analysis of variance & covariance. Rev. Edu. Res. 42: 239-288

② Box, G.E.P. Effect of inequality of variance in the one-way classification. Ann. Math. Statist 25: 290-302 1954

③ Box & Anderson, 1955. Permutation theory in the derivation of robust criteria and the study of departures from assumption. J. Royal Stat. Society B17: 1-34

④ Kohr, R.L. & P.A. Games 1974. Robustness of the analysis of variance, the Welch procedure, and a Box procedure to heterogeneous variances. J. Exp. Edu 43: 61-69.

⑤ Maxwell S.E. & H.D. Delaney 1990: Designing Expt. & Analyzing Data: A Model Comparisons Perspective. Wadsworth, Belmont, California 902.

6.41

```
get file='d:\stat601.14\polit\p171.sav'.
GLM
  cond1 cond2 cond3
  /WSFACTOR = cond 3 Polynomial
  /METHOD = SSTYPE(3)
  /EMMEANS = TABLES(cond)
  /PRINT = DESCRIPTIVE
  /CRITERIA = ALPHA(.05)
  /WSDESIGN = cond .
```

One-way repeated measure ANOVA (extension of pairwise t-test)

General Linear Model:p171.spo

Within-Subjects Factors

Measure: MEASURE_1

cond	Dependent Variable
1	cond1
2	cond2
3	cond3

Descriptive Statistics

	Mean	Std. Deviation	N
cond1	153.0000	13.08017	12
cond2	157.0000	14.50392	12
cond3	170.0000	18.57662	12

Multivariate Tests[b]

Effect		Value	F	Hypothesis df	Error df	Sig.
cond	Pillai's Trace	.870	33.428[a]	2.000	10.000	.000
	Wilks' Lambda	.130	33.428[a]	2.000	10.000	.000
	Hotelling's Trace	6.686	33.428[a]	2.000	10.000	.000
	Roy's Largest Root	6.686	33.428[a]	2.000	10.000	.000

a. Exact statistic

b.
Design: Intercept
Within Subjects Design: cond

Mauchly's Test of Sphericity[b]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
cond	.541	6.142	2	.046

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

6.42a

Mauchly's Test of Sphericity[b]

Measure: MEASURE_1

Within Subjects Effect	Epsilon[a] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
cond	.685	.750	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b.
Design: Intercept
Within Subjects Design: cond

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
cond	Sphericity Assumed	1896.000	2	948.000	56.674	.000
	Greenhouse-Geisser	1896.000	1.371	1383.060	56.674	.000
	Huynh-Feldt	1896.000	1.501	1263.406	56.674	.000
	Lower-bound	1896.000	1.000	1896.000	56.674	.000
Error(cond)	Sphericity Assumed	368.000	22	16.727		
	Greenhouse-Geisser	368.000	15.080	24.404		
	Huynh-Feldt	368.000	16.508	22.293		
	Lower-bound	368.000	11.000	33.455		

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	cond	Type III Sum of Squares	df	Mean Square	F	Sig.
cond	Linear	1734.000	1	1734.000	72.525	.000
	Quadratic	162.000	1	162.000	16.971	.002
Error(cond)	Linear	263.000	11	23.909		
	Quadratic	105.000	11	9.545		

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	921600.000	1	921600.000	1329.696	.000
Error	7624.000	11	693.091		

Estimated Marginal Means

cond

Measure: MEASURE_1

cond	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	153.000	3.776	144.689	161.311
2	157.000	4.187	147.785	166.215
3	170.000	5.363	158.197	181.803

6.42b

```
get file='d:\stat601.14\polit\p166.sav'.
UNIANOVA
  stress  BY treatmnt time
  /METHOD = SSTYPE(3)
  /INTERCEPT = INCLUDE
  /PLOT = PROFILE( time*treatmnt )
  /EMMEANS = TABLES(treatmnt*time)
  /PRINT = DESCRIPTIVE HOMOGENEITY
  /CRITERIA = ALPHA(.05)
  /DESIGN = treatmnt time treatmnt*time .
```

6.43

two-way anova (2x2 design, 2 between subjects factors, namely Tx type of 2 levels, time of 2 levels)

Univariate Analysis of Variance:p166.spo

Between-Subjects Factors

		Value Label	N
treatmnt	1.00	music	10
	2.00	relaxation	10
time time of treatment	1.00	morning	10
	2.00	evening	10

Descriptive Statistics

Dependent Variable: stress

treatmnt	time time of treatment	Mean	Std. Deviation	N
1.00 music	1.00 morning	3.0000	2.23607	5
	2.00 evening	1.0000	.70711	5
	Total	2.0000	1.88562	10
2.00 relaxation	1.00 morning	2.0000	1.58114	5
	2.00 evening	4.0000	1.58114	5
	Total	3.0000	1.82574	10
Total	1.00 morning	2.5000	1.90029	10
	2.00 evening	2.5000	1.95789	10
	Total	2.5000	1.87785	20

Levene's Test of Equality of Error Variances[a]

Dependent Variable: stress

F	df1	df2	Sig.
1.448	3	16	.266

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+treatmnt+time+treatmnt * time

6.44

Tests of Between-Subjects Effects

Dependent Variable: stress

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	25.000[a]	3	8.333	3.175	.053
Intercept	125.000	1	125.000	47.619	.000
treatmnt	5.000	1	5.000	1.905	.187
time	.000	1	.000	.000	1.000
treatmnt * time	20.000	1	20.000	7.619	.014
Error	42.000	16	2.625		
Total	192.000	20			
Corrected Total	67.000	19			

a. R Squared = .373 (Adjusted R Squared = .256)

Estimated Marginal Means

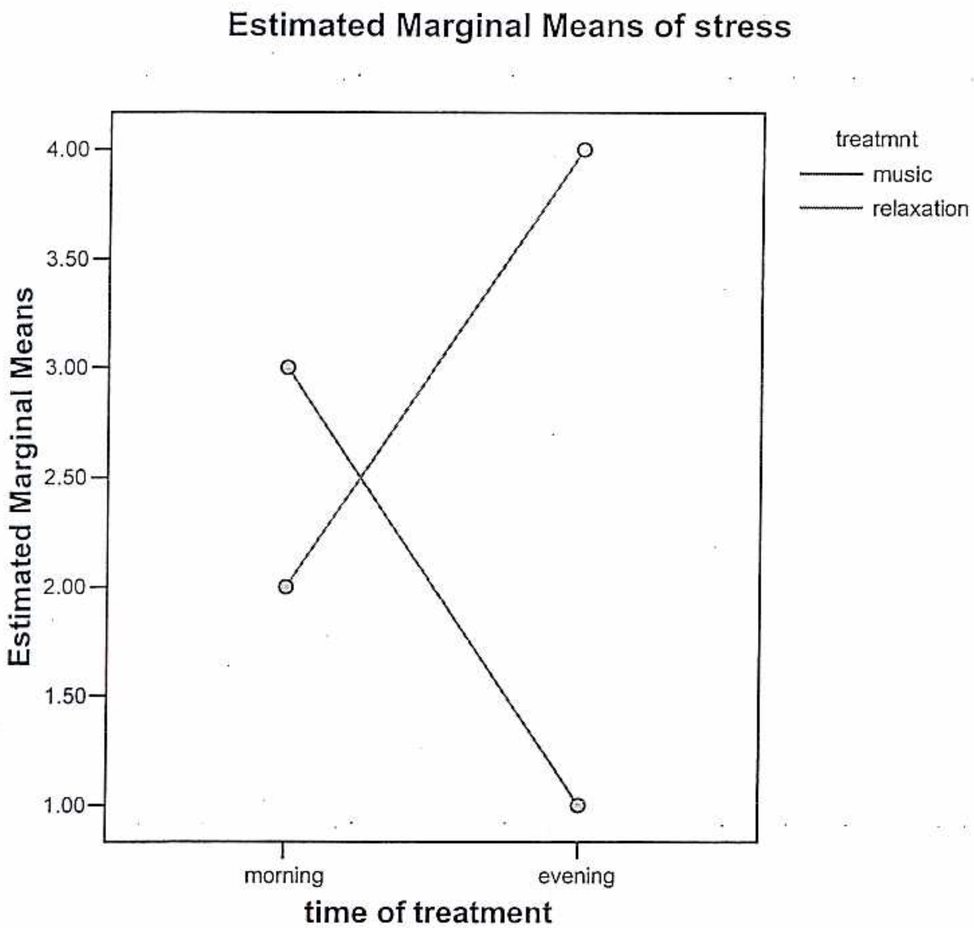
treatmnt * time of treatment

Dependent Variable: stress

treatmnt	time of treatment	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1.00 music	1.00 morning	3.000	.725	1.464	4.536
	2.00 evening	1.000	.725	-.536	2.536
2.00 relaxation	1.00 morning	2.000	.725	.464	3.536
	2.00 evening	4.000	.725	2.464	5.536

Profile Plots

6.45




```
*since there is a significant treatment & time effect.
*simple effects testings are recommended.
*For fixed level of time,we want to test treatmnt effect.
manova stress by treatmnt(1,2) time(1,2)/
    design=time, treatmnt w time(1), treatmnt w time(2).
```

Manova

```
The default error term in MANOVA has been changed from WITHIN CELLS to
WITHIN+RESIDUAL.  Note that these are the same for all full factorial
designs.
```

* * * * * * * * * * * * * * * * * * Analysis of Variance -- Design 1 *

Tests of Significance for stress using UNIQUE sums of squares

Source of Variation	SS	DF	MS	F	Sig of F
WITHIN+RESIDUAL	42.00	16	2.63		
TIME	.00	1	.00	.00	1.000
TREATMNT W TIME(1)	2.50	1	2.50	.95	.344 ← Tx effect in morning
TREATMNT W TIME(2)	22.50	1	22.50	8.57	.010 ← Tx effect in afternoon
(Model)	25.00	3	8.33	3.17	.053
(Total)	67.00	19	3.53		

R-Squared = .373
Adjusted R-Squared = .256

- -

```
*For fixed level of treatmnt effect,we want to test time effect.
manova stress by treatmnt(1,2) time(1,2)/
   design=treatmnt, time w treatmnt(1),time w treatmnt(2).
```

Manova

The default error term in MANOVA has been changed from WITHIN CELLS to WITHIN+RESIDUAL. Note that these are the same for all full factorial designs.

* * * * * * * * * * * * * * * * * * Analysis of Variance -- Design 1 *

Tests of Significance for stress using UNIQUE sums of squares

Source of Variation	SS	DF	MS	F	Sig of F
WITHIN+RESIDUAL	42.00	16	2.63		
TREATMNT	5.00	1	5.00	1.90	.187
TIME W TREATMNT(1)	10.00	1	10.00	3.81	.069 ← time effect for Tx=1
TIME W TREATMNT(2)	10.00	1	10.00	3.81	.069 ← time effect for Tx=2
(Model)	25.00	3	8.33	3.17	.053
(Total)	67.00	19	3.53		

R-Squared = .373
Adjusted R-Squared = .256

- -

```
GET
  FILE='D:\STAT601.14\POLIT\winer525.sav'.
GLM
  b1 b2 b3 b4 BY group
  /WSFACTOR = shape 4 Polynomial
  /METHOD = SSTYPE(3)
  /PLOT = PROFILE( shape*group)
  /EMMEANS = TABLES(group*shape)
  /PRINT = DESCRIPTIVE HOMOGENEITY
  /CRITERIA = ALPHA(.05)
  /WSDESIGN = shape
  /DESIGN = group .
```

Two-way anova (2 x 4 design, one between subject factor of 2 levels and one within subject factor of 4 levels)
From B. J. Winer – 2nd Edition Statistical Principles in experimental Design 1971 p. 525

General Linear Model:winer525.spo

Within-Subjects Factors

Measure: MEASURE_1

shape	Dependent Variable
1	b1
2	b2
3	b3
4	b4

Between-Subjects Factors

		N
group	1.00	3
	2.00	3

Descriptive Statistics

	group	Mean	Std. Deviation	N
b1	1.00	2.3333	2.08167	3
	2.00	5.3333	1.52753	3
	Total	3.8333	2.31661	6
b2	1.00	1.3333	1.52753	3
	2.00	4.3333	.57735	3
	Total	2.8333	1.94079	6
b3	1.00	5.3333	.57735	3
	2.00	7.0000	1.00000	3
	Total	6.1667	1.16905	6
b4	1.00	3.0000	1.00000	3
	2.00	7.6667	1.52753	3
	Total	5.3333	2.80476	6

Multivariate Tests[b]

Effect		Value	F	Hypothesis df	Error df	Sig.
shape	Pillai's Trace	.993	97.091[a]	3.000	2.000	.010
	Wilks' Lambda	.007	97.091[a]	3.000	2.000	.010
	Hotelling's Trace	145.636	97.091[a]	3.000	2.000	.010
	Roy's Largest Root	145.636	97.091[a]	3.000	2.000	.010
shape * group	Pillai's Trace	.956	14.326[a]	3.000	2.000	.066
	Wilks' Lambda	.044	14.326[a]	3.000	2.000	.066
	Hotelling's Trace	21.489	14.326[a]	3.000	2.000	.066
	Roy's Largest Root	21.489	14.326[a]	3.000	2.000	.066

a. Exact statistic

b.
Design: Intercept+group
Within Subjects Design: shape

Mauchly's Test of Sphericity[b]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
shape	.048	8.252	5	.166

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

Mauchly's Test of Sphericity[b]

Measure: MEASURE_1

	Epsilon[a]		
Within Subjects Effect	Greenhouse -Geisser	Huynh-Feldt	Lower-bound
shape	.509	.964	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b.
Design: Intercept+group
Within Subjects Design: shape

Tests of Within-Subjects Effects

6.46c

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
shape	Sphericity Assumed	40.125	3	13.375	12.506	.001
	Greenhouse-Geisser	40.125	1.526	26.292	12.506	.009
	Huynh-Feldt	40.125	2.893	13.870	12.506	.001
	Lower-bound	40.125	1.000	40.125	12.506	.024
shape * group	Sphericity Assumed	6.792	3	2.264	2.117	.151
	Greenhouse-Geisser	6.792	1.526	4.450	2.117	.200
	Huynh-Feldt	6.792	2.893	2.348	2.117	.154
	Lower-bound	6.792	1.000	6.792	2.117	.219
Error(shape)	Sphericity Assumed	12.833	12	1.069		
	Greenhouse-Geisser	12.833	6.105	2.102		
	Huynh-Feldt	12.833	11.572	1.109		
	Lower-bound	12.833	4.000	3.208		

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	shape	Type III Sum of Squares	df	Mean Square	F	Sig.
shape	Linear	18.408	1	18.408	7.488	.052
	Quadratic	.042	1	.042	.059	.820
	Cubic	21.675	1	21.675	520.200	.000
shape * group	Linear	1.008	1	1.008	.410	.557
	Quadratic	3.375	1	3.375	4.765	.094
	Cubic	2.408	1	2.408	57.800	.002
Error(shape)	Linear	9.833	4	2.458		
	Quadratic	2.833	4	.708		
	Cubic	.167	4	.042		

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
b1	.500	1	4	.519
b2	2.571	1	4	.184
b3	.400	1	4	.561
b4	.727	1	4	.442

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a.
Design: Intercept+group
Within Subjects Design: shape

6-46d

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	495.042	1	495.042	130.560	.000
group	57.042	1	57.042	15.044	.018
Error	15.167	4	3.792		

Estimated Marginal Means

group * shape

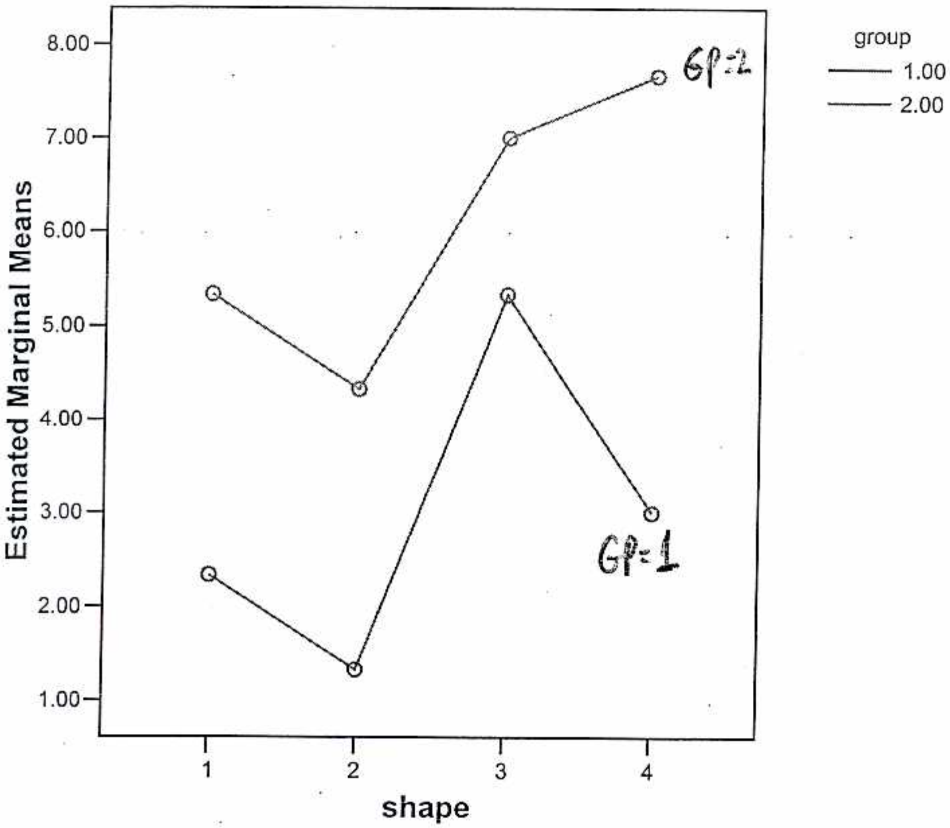
Measure: MEASURE_1

				95% Confidence Interval	
group	shape	Mean	Std. Error	Lower Bound	Upper Bound
1.00	1	2.333	1.054	-.593	5.260
	2	1.333	.667	-.518	3.184
	3	5.333	.471	4.025	6.642
	4	3.000	.745	.931	5.069
2.00	1	5.333	1.054	2.407	8.260
	2	4.333	.667	2.482	6.184
	3	7.000	.471	5.691	8.309
	4	7.667	.745	5.597	9.736

Profile Plots

6.46e

Estimated Marginal Means of MEASURE_1



Contingency table

In many situations, enumerations data are collected simultaneously for two variables and it is desired to test the hypothesis that the frequencies of occurence in the various categories of one variable are independent of the frequencies in the second variable. These two variables can be measured on the nominal or ordinal scales. The most general case is where we have two continuous variables X and Y and each variable is transformed into an ordinal scale.

e.g. 1. Let the individuals in the population W of patients be classified as either cyanotic or not and by their reaction to a particular treatment. The 3 x 2 contingency table is:

		Cyanosis YES, 1	Cyanosis NO, 2
Reaction	1 Improve		
	2 No change		
	3 Worsen		

H_0: Reaction is indep. of Cyanosis.

e.g. Let the population W of critically ill patients be stratified into two subpopulations according to whether or not they are in shock. A sample of 112 critically ill patients was taken and each patient was classified according to survival and existence or absence of shock.

		Shock Yes	Shock No	Row total
Survival	Yes	40 (49.5)	32 (22.5)	72
	No	37 (27.5)	3 (12.5)	40
		77	35	112

Ho: Survival is indep. of being in shock.
entries in parenthesis are expected frequencies.

$$\chi^2_{cal} = \frac{(40-49.5)^2}{49.5} + \frac{(32-22.5)^2}{22.5} + \frac{(37-27.5)^2}{27.5} + \frac{(3-12.5)^2}{12.5} = 16.34 \quad d.f. = 1\times1 = 1$$

or Ho: prop. of shock who survive is equal to the prop. of nonshock who survive.

$$\chi^2_{cal}(1) = 16.34 > \chi^2_{.05}(1) = 3.84$$

Reject Ho.

The population W consists of patients with a certain disease who were given a new treatment. For each patient the random variable X measured his/her age in years & the random variable Y measured the duration of his/her fever in days. The ranges of these variables were divided into 3 and 4 classes respectively.

Duration of fever	Age: ≤30	30–45	over 45
1–4			
5–6			
7–8			
9–12			

Ho: the duration of fever is independent of the age of the patient.

<u>Chi-square Test for Contingency Tables</u> (Crosstabs under SPSS)

Ho: A and B are independent.

Compute Expected frequencies by

$$E_{ij} = \frac{O_{i.} O_{.j}}{n} \qquad i = 1, \ldots r, \quad j = 1, \ldots c.$$

$$\chi^2_{calc} = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Under Ho, $\chi^2_{calc} \sim \chi^2_{(r-1)(c-1)}$

reject Ho if $\chi^2_{calc} > \chi^2_{.05, (r-1)(c-1)}$

6.482

Chisquare tests of condensed tables by SPSS:

```
-> data list free/survival shock count.

-> begin data.

-> 1 1 40
-> 1 2 32
-> 2 1 37
-> 2 2 3
-> end data.

-> weight by count.

-> value labels survival shock 1 'Yes' 2 'No'.

-> crosstabs tables=survival by shock/
->    cells=count column expected/statistics=chisquare.
```

```
SURVIVAL  by  SHOCK

                    SHOCK         Page 1 of 1
           Count   |
           Exp Val |Yes      No
           Col Pct |                  Row
                   |    1.00|    2.00| Total
SURVIVAL   --------+--------+--------+
            1.00   |    40  |    32  |    72
  Yes              |  49.5  |  22.5  | 64.3%
                   | 51.9%  | 91.4%  |
                   +--------+--------+
            2.00   |    37  |     3  |    40
  No               |  27.5  |  12.5  | 35.7%
                   | 48.1%  |  8.6%  |
                   +--------+--------+
           Column       77       35     112
            Total    68.8%    31.3%  100.0%
```

Chi-Square	Value	DF	Significance
Pearson	16.33616	1	.00005
Continuity Correction	14.66182	1	.00013
Likelihood Ratio	18.89012	1	.00001
Linear-by-Linear Association	16.19030	1	.00006
Fisher's Exact Test:			
One-Tail			.00002
Two-Tail			.00004

Minimum Expected Frequency - 12.500

Number of Missing Observations: 0

When both variables in the contingency table have only two levels – that is when we have a 2×2 table – a correction factor known as Yates' correction for continuity is sometimes use in computing χ^2.

$$\chi^2 \text{ (corrected)} = \sum_{j=1}^{2}\sum_{i=1}^{2} \frac{\left(|O_{ij}-E_{ij}| - .5\right)^2}{E_{ij}}$$

If the expected cell frequencies are large, the corrected and uncorrected are almost the same. When the expected frequencies are between 5 and 10 Yates' correction should be applied. For expected frequencies less than 5, the Fisher Exact test should be used. SPSS calculates Fisher's exact test if any expected cell value in a 2×2 table is less than 5.

e.g. H_0: Respiratory problem is indep. of Smoking

respiratory problem	Smoking	Not smoking	
Yes	46 (41.3)	13 (17.7)	59
NO	10 (14.7)	11 (6.3)	21
	56	24	80

$$\chi^2_{cal} = \frac{(|46-41.3| - .5)^2}{41.3} + \frac{(|13-17.7| - .5)^2}{17.7} + \frac{(|10-14.7| - .5)^2}{14.7} + \frac{(|11-6.3| - .5)^2}{6.3} = 5.42 >$$

$\chi^2_{.05}(1) = 3.84$ reject H_0 at 0.05 level

6 #50.

```
-> *.

-> data list free/res_prob smoking count.

-> begin data.

-> 1 1 46
-> 1 2 13
-> 2 1 10
-> 2 2 11
-> end data.

-> value labels res_prob,smoking 1 'Yes' 2 'No'.

-> weight by count.

-> crosstabs tables=res_prob by smoking/
->     cells=count column expected/statistics=chisquare.
```

RES_PROB by SMOKING

SMOKING Page 1 of 1

RES_PROB	Count / Exp Val / Col Pct	Yes 1.00	No 2.00	Row Total
Yes	1.00	46 41.3 82.1%	13 17.7 54.2%	59 73.8%
No	2.00	10 14.7 17.9%	11 6.3 45.8%	21 26.3%
	Column Total	56 70.0%	24 30.0%	80 100.0%

Chi-Square	Value	DF	Significance
Pearson	6.79196	1	.00916
Continuity Correction	5.42373	1	.01986
Likelihood Ratio	6.44802	1	.01111
Linear-by-Linear Association	6.70706	1	.00960
Fisher's Exact Test:			
One-Tail			.01123
Two-Tail			.01323

Minimum Expected Frequency - 6.300

Number of Missing Observations: 0

6.51

TABLE A.8 *F* **distribution**

Entries in the following table give F_α values, where α is the area or probability in the upper tail of the F distribution.

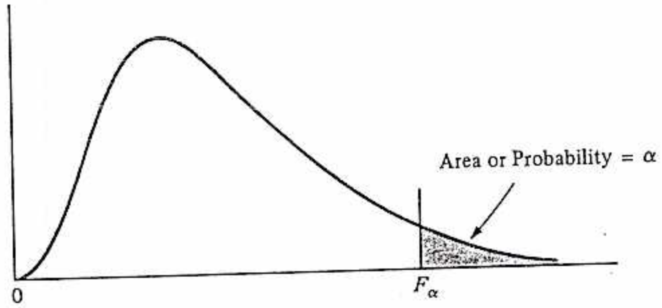


Table of $F_{.05}$ Values

Denominator Degrees of Freedom	Numerator Degrees of Freedom 1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	2
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	