ANOVA of a 2^n Experiment

Consider performing an ANOVA on a 2^3 experiment with factors *A*, *B*, & *C*.

ANOVA Table

Source of Variation	Degrees of Freedom
A	1
B	1
*A*B*	1
C	1
*A*C*	1
*B*C*	1
*A*B*C*	1
Error	0

One can use SAS-ANOVA and/or SAS-GLM. The basic commands for either procedure are:

```
PROC ANOVA;
 CLASSES A B C;
 MODEL dep var = A B A*B C A*C B*C A*B*C;
RUN;
```

Equivalently, one can use the following commands:

```
PROC ANOVA;
 CLASSES A B C;
 MODEL dep var = A|B|C;
RUN;
```

QUESTION:

If one performs an ANOVA on a 2^3 experiment, how do the Effect Mean Squares compare to our Effects and/or our Standardized Effects?

ANSWER:

The Effect Mean Squares are the squares of our Standardized Effects. That is,

(Standardized Effect) ^{2} ***= Effect Mean Square***

or

$$|\textit{Standardized Effect}| = \sqrt{\boldsymbol{Effect\ Mean\ Square}}$$

```
TITLE 'A 2**3 EXPERIMENT';
DATA ONE;
 INPUT TEMP CONC CATLST $ Y;
 LINES;
160 20 C1 60
180 20 C1 72
160 40 C1 54
180 40 C1 68
160 20 C2 52
180 20 C2 83
160 40 C2 45
180 40 C2 80
```

This program is on the Course website as ST722_1.sas

```
ODS RTF FILE='C:TEMP.RTF';
PROC PRINT;
 TITLE2 'THE OBSERVED DATA';
 RUN;
PROC ANOVA;
 CLASSES TEMP CONC CATLST;
 MODEL Y=TEMP|CONC|CATLST;
 RUN;
 ODS RTF CLOSE;
```

Print of the Data for a 2^3 Experiment

Obs	TEMP	CONC	CATLST	Y
1	160	20	C1	60
2	180	20	C1	72
3	160	40	C1	54
4	180	40	C1	68
5	160	20	C2	52
6	180	20	C2	83
7	160	40	C2	45
8	180	40	C2	80

Class Level Information

Class	**Levels**	**Values**
TEMP	2	160 180
CONC	2	20 40
CATLST	2	C1 C2

ANOVA for the Full Model from SAS

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	1317.50000	188.214286	.	.
Error	0	0.000000	.		
Corrected Total	7	1317.50000			

ANOVA Table from SAS

Source	DF	Anova SS	Mean Square	F Value	Pr > F
TEMP	1	1058.000	1058.000	.	.
CONC	1	50.0000	50.00000	.	.
TEMP*CONC	1	4.5000	4.50000	.	.
CATLST	1	4.5000	4.500000	.	.
TEMP*CATLST	1	200.0000	200.000000	.	.
CONC*CATLST	1	0.0000	0.000000	.	.
TEMP*CONC*CATLST	1	0.5000	0.500000	.	.

Here are the Standardized Effects from our Yates' Analysis to compare to the SAS results.

Standardized Effect	EFFECT
32.527	*A*
-7.071	*B*
2.121	*A*B*
2.121	*C*
14.142	*A*C*
0.000	*B*C*
0.707	*A*B*C*

General Definitions:

Consider a 2^n experiment with factors

$A_1, A_2, \ldots, A_n$.

Effect Type	Effects	Number of Effects
Main Effects	$A_1, A_2, \ldots, A_n$	$n = \binom{n}{1}$
Two-Way Interactions	$A_1A_2, A_1A_3, \ldots, A_{n-1}A_n$	$\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n(n-1)}{2}$
Three-Way Interactions	$A_1A_2A_3, A_1A_2A_4 \ldots, A_{n-2}A_{n-1}A_n$	$\binom{n}{3} = \frac{n!}{3!(n-3)!}$
. . .		
n – Way Interaction	$A_1A_2 \ldots A_n$	$\binom{n}{n} = 1$

Example: Consider a 2^6 experiment.

Effect Type	Number
Main	6
2-Way Interactions	15
3-Way Interactions	20
4-Way Interaction	15
5-Way Interactions	6
6-Way Interaction	1
Total	63

Testing for Statistical Significance

CASE 1: Have some independent replications of one or more of the treatment combinations.

If there are independent replications, then σ^2 can be estimated in a traditional manner. Suppose we denote such an estimate by $\hat{\sigma}^2$, and suppose that the estimate is based on ν degrees of freedom.

Then a Least Significant Difference (LSD) for a standardized effect is

$$LSD_{\alpha} = t_{\alpha/2,\nu} \cdot \hat{\sigma}$$

If |Standardized Effect| > $LSD_{\alpha} = t_{\alpha/2,\nu} \cdot \hat{\sigma}$,

then the Effect is said to be statistically significant at the $\alpha \cdot 100\%$ significance level.

Equivalently, an Effect is statistically significant at this level if

$$F = \frac{\text{Effect MS}}{\hat{\sigma}^2} > F_{\alpha,\,1,\,\nu}.$$

Definition: An Effect is said to be standardized if the VAR(EFFECT) = σ^2 .

Remark: When there are r independent replications of each of the treatment combinations in a 2^n experiment, we can apply Yates' method to the treatment combination means. The variance of each of these means is $\frac{\sigma^2}{r}$.

Therefore, to compute the standardized effects, we would divide the values in the Step n column by $(2^n/r)^{1/2}$.

Remark: When there are r independent replications of each of the treatment combinations in a 2^n experiment, we could also apply Yates' method to the treatment combination totals. The variance of each of these totals is $r\sigma^2$

Therefore, to compute the standardized effects from the treatment combination totals, we would divide the values in the Step n column of Yates' Method by $[r(2^n)]^{1/2}$.

Example: An experiment was conducted to determine the effects of manure (*m*), nitrogen (*n*), phosphorus (*p*), and potassium (*k*) on the yield of prairie hay.

Example from Cochran & Cox.

	Replicate			
Trt Comb	1	2	3	4
(1)	32	43	27	19
m	47	41	48	45
n	26	36	24	18
mn	61	76	56	64
p	29	39	27	28
mp	51	34	40	48
np	36	31	32	30
mnp	76	65	70	63
k	35	42	56	35
mk	63	41	60	53
nk	80	68	75	67
mnk	100	68	87	66
pk	40	44	53	36
mpk	64	39	75	72
npk	105	99	74	73
mnpk	90	82	89	101

	Replicate				
Trt Comb	1	2	3	4	Total
(1)	32	43	27	19	121
m	47	41	48	45	181
n	26	36	24	18	104
mn	61	76	56	64	257
p	29	39	27	28	123
mp	51	34	40	48	173
np	36	31	32	30	129
mnp	76	65	70	63	274
k	35	42	56	35	168
mk	63	41	60	53	217
nk	80	68	75	67	290
mnk	100	68	87	66	321
pk	40	44	53	36	173
mpk	64	39	75	72	250
npk	105	99	74	73	351
mnpk	90	82	89	101	362

Trt Comb	Total				
(1)	121				
m	181				
n	104				
mn	257				
p	123				
mp	173				
np	129				
mnp	274				
k	168				
mk	217				
nk	290				
mnk	321				
pk	173				
mpk	250				
npk	351				
mnpk	362				

Trt Comb	Total	Step 1	Step 2	Step 3	Step 4
(1)	121	302	663	1362	3494
m	181	361	699	2132	576
n	104	296	996	408	682
mn	257	403	1136	168	104
p	123	385	213	166	176
mp	173	611	195	516	-10
np	129	423	80	188	112
mnp	274	713	88	-84	-46
k	168	60	59	36	770
mk	217	153	107	140	-240
nk	290	50	226	-18	350
mnk	321	145	290	8	-272
pk	173	49	93	48	104
mpk	250	31	95	64	26
npk	351	77	-18	2	16
mnpk	362	11	-66	-48	-50

To standardize the results in Step 4, we divide by

Sqrt[$(4)(2^4)$] = 8.

Trt Comb	Total	Step 1	Step 2	Step 3	Step 4	Std Eff
(1)	121	302	663	1362	3494	
m	181	361	699	2132	576	72.00
n	104	296	996	408	682	85.25
mn	257	403	1136	168	104	13.00
p	123	385	213	166	176	22.00
mp	173	611	195	516	-10	-1.25
np	129	423	80	188	112	14.00
mnp	274	713	88	-84	-46	-5.75
k	168	60	59	36	770	96.25
mk	217	153	107	140	-240	-30.00
nk	290	50	226	-18	350	43.75
mnk	321	145	290	8	-272	-34.00
pk	173	49	93	48	104	13.00
mpk	250	31	95	64	26	3.25
npk	351	77	-18	2	16	2.00
mnpk	362	11	-66	-48	-50	-6.25

Trt Comb	Total	Std Eff	**Effect**
(1)	121		
m	181	72.00	M
n	104	85.25	N
mn	257	13.00	M*N
p	123	22.00	P
mp	173	-1.25	M*P
np	129	14.00	N*P
mnp	274	-5.75	M*N*P
k	168	96.25	K
mk	217	-30.00	M*K
nk	290	43.75	N*K
mnk	321	-34.00	M*N*K
pk	173	13.00	P*K
mpk	250	3.25	M*P*K
npk	351	2.00	N*P*K
mnpk	362	-6.25	M*N*P*K

Next, we need to determine the estimate of the experimental error variance. How this is done will depend on the experimental design that was used.

If the design is a completely randomized design, we can get an estimate of σ^2 from each set of replications. Each estimate would be based on $r - 1$ degrees of freedom. Then a pooled estimate can be obtained from the average of the individual estimates. The pooled estimate will have $(r - 1)(2^n)$ degrees of freedom associated with it.

Trt Comb	1	2	3	4	Variance
(1)	32	43	27	19	100.917
m	47	41	48	45	
n	26	36	24	18	
mn	61	76	56	64	
p	29	39	27	28	
mp	51	34	40	48	
np	36	31	32	30	
mnp	76	65	70	63	
k	35	42	56	35	
mk	63	41	60	53	
nk	80	68	75	67	
mnk	100	68	87	66	
pk	40	44	53	36	
mpk	64	39	75	72	
npk	105	99	74	73	
mnpk	90	82	89	101	

Trt Comb	1	2	3	4	Variance
(1)	32	43	27	19	100.917
m	47	41	48	45	9.583
n	26	36	24	18	56.000
mn	61	76	56	64	72.250
p	29	39	27	28	30.917
mp	51	34	40	48	59.583
np	36	31	32	30	6.917
mnp	76	65	70	63	33.667
k	35	42	56	35	98.000
mk	63	41	60	53	95.583
nk	80	68	75	67	37.667
mnk	100	68	87	66	262.917
pk	40	44	53	36	52.917
mpk	64	39	75	72	267.000
npk	105	99	74	73	276.917
mnpk	90	82	89	101	61.667

Trt Comb	1	2	3	4	Variance
(1)	32	43	27	19	100.917
m	47	41	48	45	9.583
n	26	36	24	18	56.000
mn	61	76	56	64	72.250
p	29	39	27	28	30.917
mp	51	34	40	48	59.583
np	36	31	32	30	6.917
mnp	76	65	70	63	33.667
k	35	42	56	35	98.000
mk	63	41	60	53	95.583
nk	80	68	75	67	37.667
mnk	100	68	87	66	262.917
pk	40	44	53	36	52.917
mpk	64	39	75	72	267.000
npk	105	99	74	73	276.917
mnpk	90	82	89	101	61.667
					95.156

Trt Comb	1	2	3	4	Variance
(1)	32	43	27	19	100.917
m	47	41	48	45	9.583
n	26	36	24	18	56.000
mn	61	76	56	64	72.250
p	29	39	27	28	30.917
mp	51	34	40	48	59.583
np	36	31	32	30	6.917
mnp	76	65	70	63	33.667
k	35	42	56	35	98.000
mk	63	41	60	53	95.583
nk	80	68	75	67	37.667
mnk	100	68	87	66	262.917
pk	40	44	53	36	52.917
mpk	64	39	75	72	267.000
npk	105	99	74	73	276.917
mnpk	90	82	89	101	61.667
					95.156

$$\hat{\sigma}^2 = 95.156 \quad \text{with} \quad (16)(3) = 48 \ \text{degrees of freedom}$$

$$\hat{\sigma}^2 = 95.156 \quad \text{with} \quad (16)(3) = 48 \ \text{degrees of freedom}$$

$$\hat{\sigma} = 9.755$$

$$\hat{\sigma}^2 = 95.156 \quad \text{with} \quad (16)(3) = 48 \ \text{degrees of freedom}$$

$$\hat{\sigma} = 9.755$$

$$LSD_{0.05} = t_{0.025,\,48} \cdot \hat{\sigma} = (2.011) \cdot (9.756) = 19.619$$

Std Eff		EFFECT
72	*	*M*
85.25	*	*N*
13		*M*N*
22		*P*
-1.25		*M*P*
14		*N*P*
-5.75		*M*N*P*
96.25	*	*K*
-30	*	*M*K*
43.75	*	*N*K*
-34	*	*M*N*K*
13		*P*K*
3.25		*M*P*K*
2		*N*P*K*
-6.25		*M*N*P*K*

* Denotes statistical significance at the 5% level.

Actually, the previous experiment was conducted in a Randomized Complete Block Design. Thus the Block Sum of Squares should be removed from the within treatment combinations sum of squares.

	Replicate				
Trt Comb	1	2	3	4	Total
(1)	32	43	27	19	121
m	47	41	48	45	181
n	26	36	24	18	104
mn	61	76	56	64	257
p	29	39	27	28	123
mp	51	34	40	48	173
np	36	31	32	30	129
mnp	76	65	70	63	274
k	35	42	56	35	168
mk	63	41	60	53	217
nk	80	68	75	67	290
mnk	100	68	87	66	321
pk	40	44	53	36	173
mpk	64	39	75	72	250
npk	105	99	74	73	351
mnpk	90	82	89	101	362
Totals	935	848	893	818	3494

$$Blk\ SS = \frac{935^2 + 848^2 + 893^2 + 818^2}{16} - \frac{(3494)^2}{64} = 493.3125$$

$$Within\ Trt\ SS = (48)(95.156) = 4567.488$$

$$Error\ SS = 4567.488 - 493.3125 = 4074.1755$$

The above SS has 48 - 3 = 45 degrees of freedom associated with it.

$$\hat{\sigma}^2 = \frac{4074.1755}{45} = 90.5372 \quad \text{and} \quad \hat{\sigma} = 9.515$$

The LSD for the experiment when considering the blocks is:

$$LSD_{0.05} = t_{0.025,45} \cdot \hat{\sigma} = (2.014) \cdot (9.515) = 19.163$$

The previous data was analyzed with SAS using the following commands. The SAS program is on the course website under the name ST722_2.SAS.

```
PROC ANOVA;
 CLASSES M N P K BLK;
 MODEL YIELD = BLK M|N|P|K;
RUN;
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	18	27285.25000	1515.84722	16.74	<.0001
Error	45	4074.18750	90.53750		
Corrected Total	63	31359.43750			

Source	DF	Anova SS	Mean Square	F Value	Pr > F
BLK	3	493.312500	164.437500	1.82	0.1578
M	1	5184.000000	5184.000000	57.26	<.0001
N	1	7267.562500	7267.562500	80.27	<.0001
M*N	1	169.000000	169.000000	1.87	0.1787
P	1	484.000000	484.000000	5.35	0.0254
M*P	1	1.562500	1.562500	0.02	0.8961
N*P	1	196.000000	196.000000	2.16	0.1482
M*N*P	1	33.062500	33.062500	0.37	0.5487
K	1	9264.062500	9264.062500	102.32	<.0001
M*K	1	900.000000	900.000000	9.94	0.0029
N*K	1	1914.062500	1914.062500	21.14	<.0001
M*N*K	1	1156.000000	1156.000000	12.77	0.0009
P*K	1	169.000000	169.000000	1.87	0.1787
M*P*K	1	10.562500	10.562500	0.12	0.7343
N*P*K	1	4.000000	4.000000	0.04	0.8345
M*N*P*K	1	39.062500	39.062500	0.43	0.5146