Chapter 13: ANOVA for 2-way classifications (2 of 2) Fixed and Random factors, Model I, Model II, and Model III (mixed model) ANOVA

Chapter 14: Unreplicated Factorial & Nested Designs

Class 23, 5/4/09 M

Slide 1 Chapter 13: ANOVA for 2-way classifications (2 of 2) Fixed and Random factors, Model I, Model II, and Model III (mixed model) ANOVA

Chapter 14:

Unreplicated Factorial & Nested Designs

NOTES:

HW 15 due Weds 5/6/09 10 am

Submit as Myname-HW15.doc (or *.rtf)

- Read Chapter 14 Multifactor studies without replication
- For Weds read Chapter 23: Elements of Research Design
- For Monday Chapters 18-19: Comparisons of Proportions or Odds
- Final Class: Weds May 13 Experimental Designs
- Class schedule May 6 (Nesting and Experimental Designs), May 11 (Overview of generalized linear models) Exptl design May 13 W Last class
- Wimba Sessions: new times to get help on HW15
  - Tues night (5/5/09) 10 pm New day
  - Thus afternoon pm New Time
- HW15: Due Weds 5/6/09 10 am
  - 14.17 Tennessee Corn Yields
  - Note that there is insufficient replication to test the full factorial model (use custom model in GLM/Univariate to test only main effects. What must you assume? You can test White vs. Yellow using linear contrasts – must use syntax in GLM/Univariate - see Fish tail example as a guide)
- HW16: Final Homework Exercise 23.20
- Final Exam 5/22 8-11 am

Slide 2 HW 15 due Weds 5/6/09 10 am

NOTES:

Case 13.2 Pygmalion Effect

Slide 3 Case 13.2 Pygmalion Effect

NOTES:

Pygmalion effect

A study to avoid interpersonal interactions

- Tracking in schools:
  - Good students get better and poor students get worse
  - Self-fulfilling prophecies
- Goal of the study by Dov Eden: Pygmalion without interpersonal contrast effects
- Ten companies selected (9 in data), 3 platoons in each company, 1 platoon leader out of 3 told he had an exceptional group

Slide 4 Pygmalion effect

NOTES:

Pygmalion Effect

Mean scores for the platoons to be contrasted

Display 13.3

Average scores of soldiers on the Practical Specialty Test, for platoons given the Pygmalion treatment and for control platoons

Company	Pygmalion	Control	
1	80.0	63.2	69.2
2	83.9	63.1	81.5
3	68.2	76.2	
4	76.5	59.5	73.5
5	87.8	73.9	78.5
6	89.8	78.9	84.7
7	76.1	60.6	69.6
8	71.5	67.8	73.2
9	69.5	72.3	73.9
10	83.7	63.7	77.7

gn

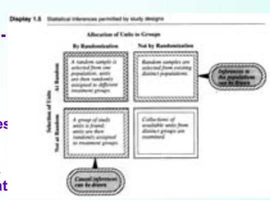
Slide 5 Pygmalion Effect

NOTES:

Pygmalion results

Note: Gallagher added results of random effects model

- Pygmalion treatment added 7.2 (±5.4) points to a platoon's score
- Very strong evidence that the Pygmalion effect is real (Fixed effect, randomized block ANOVA, $F_{1,18}=7.8$; 1-sided p = 0.006)
- Because of the randomized design, a causal inference can be made for this group of 10 companies
- [Gallagher analysis: If these companies are representative of all army companies, the Pygamalion treatment added 6.84 (±6.42) units to a platoon's score. There is moderate evidence that the effect would be found throughout Army companies (Linear contrast estimate of Pygmalion effect p=0.02)]



Slide 6 Pygmalion results

NOTES:

Strategies for factorial analysis

- *Decide at the design stage whether factors are fixed or random*
- Analyze the data graphically for outliers, need for transformation
- Fit the rich model (saturated model) examine the residual plots
- With interactions, graphically display the data or use multiway tables
- Look at particular terms in the additive model to examine particular effects
- ANOVA F-test for additivity, Interaction MS over error MS
  - Use appropriate rules for pooling:
  - Pool only if p>0.25 and only if df for MSE is < 5
- *Test main effects over appropriate error term for fixed or random effects model*

Slide 7 Strategies for factorial analysis

NOTES:

Additive and non-additive models

- Both Ch 13 Case Studies can be viewed as additive models
  - 13.1 Area + predator effects (no intxn)
  - 13.2: Block (Company) + Pygmalion effect
- Additive model: both block and factor add fixed amount

Most recent statistics texts, esp. in ecology, accept the reduced (additive) model if the interaction p values > 0.25 or 0.5

Display 13.4

Hypothetical mean scores on the Practical Specialty Test, illustrating additivity of treatment and company effects

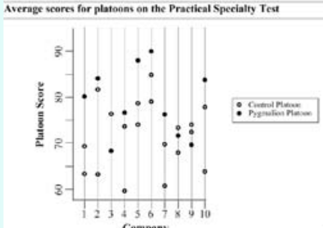
	Treatments		Treatment Effects
Company	Pygmalion	Control	(Pygmalion - Control)
1	78	70	8
2	85	77	8
3	79	71	8
4	72	64	8
5	84	76	8
6	89	81	8
7	73	65	8
8	76	68	8
9	75	67	8
10	82	74	8

Slide 8 Additive and non-additive models

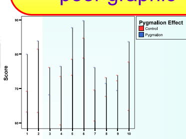
NOTES:

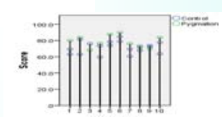
Display 13.14

Average scores for platoons on the Practical Specialty Test



Tufte data-ink rule violated here: a really poor graphic

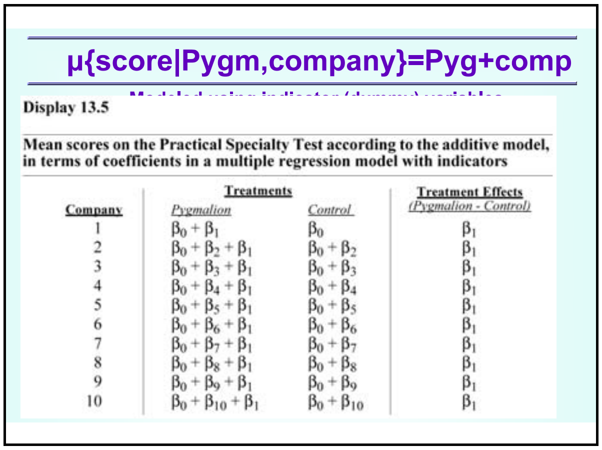




Tufte's data-ink rule: Every graphical element should convey meaning; avoid 'chart junk'

Slide 9

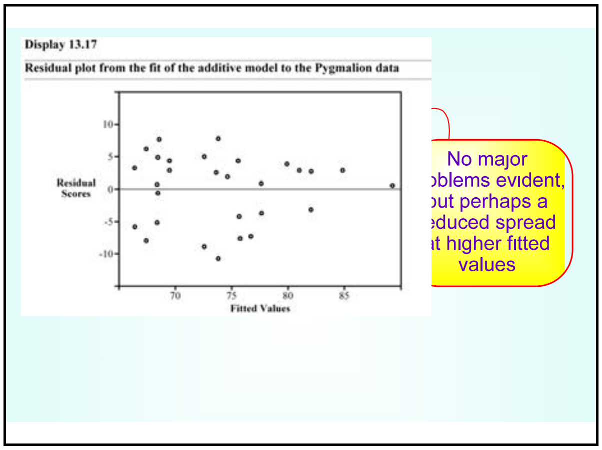
NOTES:



Company	Treatments: Pygmalion	Treatments: Control	Treatment Effects (Pygmalion - Control)
1	$\beta_0+\beta_1$	β_0	β_1
2	$\beta_0+\beta_2+\beta_1$	$\beta_0+\beta_2$	β_1
3	$\beta_0+\beta_3+\beta_1$	$\beta_0+\beta_3$	β_1
4	$\beta_0+\beta_4+\beta_1$	$\beta_0+\beta_4$	β_1
5	$\beta_0+\beta_5+\beta_1$	$\beta_0+\beta_5$	β_1
6	$\beta_0+\beta_6+\beta_1$	$\beta_0+\beta_6$	β_1
7	$\beta_0+\beta_7+\beta_1$	$\beta_0+\beta_7$	β_1
8	$\beta_0+\beta_8+\beta_1$	$\beta_0+\beta_8$	β_1
9	$\beta_0+\beta_9+\beta_1$	$\beta_0+\beta_9$	β_1
10	$\beta_0+\beta_{10}+\beta_1$	$\beta_0+\beta_{10}$	β_1

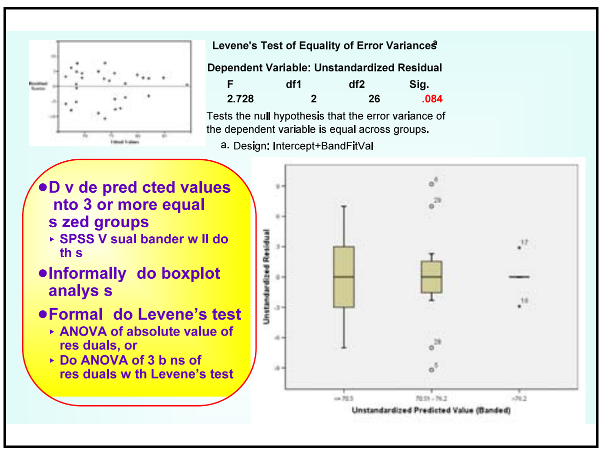
Slide 10
μ{score|Pygm,company}=Pyg+comp

NOTES:



Slide 11

NOTES:



Slide 12

NOTES:

Visual binning to examine residuals

Available in SPSS

```
* Visual Binning.
*PRE_1.
RECODE  PRE_1 (MISSING=COPY) (LO THRU 70.500000000000=1) (LO THRU 76.200000000000=2) (LO THRU
HI=3) (ELSE=SYSMIS) INTO BandFitVal.
VARIABLE LABELS  BandFitVal 'Unstandardized Predicted Value (Binned)'.
FORMAT  BandFitVal (F5.0).
VALUE LABELS  BandFitVal 1 '<= 70.50000' 2 '70.50001 - 76.20000' 3 '76.20001+'.
MISSING VALUES  BandFitVal ( ).
VARIABLE LEVEL  BandFitVal (ORDINAL).
EXECUTE .
EXAMINE
VARIABLES=RES_1 BY BandFitVal
/PLOT=BOXPLOT/STATISTICS=NONE/NOTOTAL/ID=case .
```

```
MISSING VALUES BandFitVal ( ).
VARIABLE LEVEL BandFitVal ( ORDINAL ).
EXECUTE.
* Do an ANOVA on the residuals, examining only the Levene test (identical result).
UNIANOVA
RES_1  BY BandFitVal
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/PRINT = HOMOGENEITY
/CRITERIA = ALPHA(.05)
/DESIGN = BandFitVal .
```

Slide 13 Visual binning to examine residuals

NOTES:

μ{score|Pygm,company} = Pyg+company+Pyg x company

The saturated model (includes 9 interaction terms)

Display 13.6

9 interaction terms

Mean scores on the Practical Specialty Test, in terms of the parameters in a saturated multiple linear regression model with interaction

Company	Treatments: Pygmalion	Treatments: Control	Treatment Effects (Pygmalion - Control)
1	$\beta_0 + \beta_1$	β_0	β_1
2	$\beta_0 + \beta_2 + \beta_1 + \beta_{11}$	$\beta_0 + \beta_2$	$\beta_1 + \beta_{11}$
3	$\beta_0 + \beta_3 + \beta_1 + \beta_{12}$	$\beta_0 + \beta_3$	$\beta_1 + \beta_{12}$
4	$\beta_0 + \beta_4 + \beta_1 + \beta_{13}$	$\beta_0 + \beta_4$	$\beta_1 + \beta_{13}$
5	$\beta_0 + \beta_5 + \beta_1 + \beta_{14}$	$\beta_0 + \beta_5$	$\beta_1 + \beta_{14}$
6	$\beta_0 + \beta_6 + \beta_1 + \beta_{15}$	$\beta_0 + \beta_6$	$\beta_1 + \beta_{15}$
7	$\beta_0 + \beta_7 + \beta_1 + \beta_{16}$	$\beta_0 + \beta_7$	$\beta_1 + \beta_{16}$
8	$\beta_0 + \beta_8 + \beta_1 + \beta_{17}$	$\beta_0 + \beta_8$	$\beta_1 + \beta_{17}$
9	$\beta_0 + \beta_9 + \beta_1 + \beta_{18}$	$\beta_0 + \beta_9$	$\beta_1 + \beta_{18}$
10	$\beta_0 + \beta_{10} + \beta_1 + \beta_{19}$	$\beta_0 + \beta_{10}$	$\beta_1 + \beta_{19}$

Slide 14 μ{score|Pygm,company} = Pyg+company+Pyg x company

NOTES:

Display 13.16

F-test for interactions between companies and treatment; Pygmalion data

Analysis of variance table from regression fit to the full, non-additive model,
PYG + COMP + PYG ×COMP:

Source of Variation	Sum of Squares	df	Mean Square	F-Statistic	p-value
Regression	1321.3221	19	69.5433	1.3401	0.1747
Residual	467.04	9	51.8933		
Total	1,788.3621	28			

Analysis of variance table from regression fit to the additive model,
PYG + COMP:

Source of Variation	Sum of Squares	df	Mean Square	F-Statistic	p-value
Regression	1,009.8581	10	100.9858	2.3349	0.0564
Residual	778.5039	18	43.2502		
Total	1,788.3621	28			

Residual Sum of Squares from reduced model

Residual Sum of Squares from full model

$$\text{F-Statistic} = \frac{(778.5039 - 467.0400)/(18-9)}{51.8933} = \frac{34.6071}{51.8933} = 0.667,$$

Residual Mean Square (est. of σ^2) from full model

p-value for interaction = $\Pr(F_{9,9} > 0.667) = .72$

There is no reason to keep the 9 interaction terms (Extra sum of Squares F text: p = 0.72). This meets the criteria (p>0.5) established by Underwood, Quinn & Keough, Sokal & Rohlf.

Slide 15

NOTES:

Extra sum of squares F test

Enter 3 models hiearchically using /Analyze/Regression
The 9 interaction terms do not explain a significant portion of the residual variation.

Model Summary[d]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics: R Square Change	F Change	df1	df2	Sig. F Change
1	**.428[a]**	**.183**	**.153**	**7.3561**	**.183**	**6.049**	**1**	**27**	**.021**
2	**.751[b]**	**.565**	**.323**	**6.5765**	**.382**	**1.753**	**9**	18	.148
3	.860[c]	.739	.188	7.2037	.174	.667	9	9	.722

a. Predictors: (Constant), **Pyg**
b. Predictors: (Constant), **Pyg**, CMP10, CMP9, CMP3, CMP8, CMP7, CMP6, CMP5, CMP2, CMP4
c. Predictors: (Constant), **Pyg**, CMP10, CMP9, CMP3, CMP8, CMP7, CMP6, CMP5, CMP2, CMP4, INT9, INT8, INT6, INT5, INT2, INT7, INT4, INT10, INT3
d. Dependent Variable: Score

The 9 block x interaction terms, with a p value of 0.72 can be dropped

Slide 16 Extra sum of squares F test

NOTES:

Display 13.18

Multiple linear regression output from the fit of the additive model to the Pygmalion data: $\mu\{score \mid PYG, COMP\} = PYG + COMPANY$

Variable	Coefficient	Standard Error	t-Statistic	2-Sided p-Value
CONSTANT	75.6137	4.1682	18.1405	<.0001
pyg	**7.2205**	**2.5795**	**2.7992**	**.0119**
cmp2	5.3667	5.3697	0.9994	.3308
cmp3	0.1966	6.0189	0.0327	.9743
cmp4	-0.9667	5.3697	-0.1800	.8591
cmp5	9.2667	5.3697	1.7257	.1015
cmp6	13.6667	5.3697	2.5452	.0203
cmp7	-2.0333	5.3697	-0.3787	.7094
cmp8	0.0333	5.3697	0.0062	.9951
cmp9	1.1000	5.3697	0.2049	.8400
cmp10	4.2333	5.3697	0.7884	.4407

Estimated SD = 6.576 on 18 d.f.

The Pygmalion effect adds 7.2 (± 5.4) to the score of the typical platoon

Slide 17

NOTES:

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B: Lower Bound	Upper Bound
1	(Constant)	71.63	1.69		42.45	0.00	68.17	75.09
	Pygmalion Effect	7.07	2.87	0.43	2.46	0.02	1.17	12.97
2	(Constant)	68.39	3.89		17.57	0.00	60.21	76.57
	Pygmalion Effect	7.22	2.58	0.44	2.80	0.01	1.80	12.64
	cmp2	5.37	5.37	0.21	1.00	0.33	-5.91	16.65
	cmp3	.20	6.02	0.01	0.03	0.97	-12.45	12.84
	cmp4	-.97	5.37	-0.04	-0.18	0.86	-12.25	10.31
	cmp5	9.27	5.37	0.36	1.73	0.10	-2.01	20.55
	cmp6	13.67	5.37	0.53	2.55	0.02	2.39	24.95
	cmp7	-2.03	5.37	-0.08	-0.38	0.71	-13.31	9.25
	cmp8	.03	5.37	0.00	0.01	1.00	-11.25	11.31
	cmp9	1.10	5.37	0.04	0.20	0.84	-10.18	12.38
	cmp10	4.23	5.37	0.16	0.79	0.44	-7.05	15.51

a. Dependent Variable: Score

Slide 18

NOTES:

Unbalanced designs & effect sizes

Different estimates of the treatment effect, from each company and from the combined data ignoring company differences

	i	Averages: Pygmalion	Averages: Control	Difference $\hat{\delta}_i$
Company	1	80.0	66.2	13.8
	2	83.9	72.3	11.6
	3	68.2	76.2	-8.0
	4	76.5	66.5	10.0
	5	87.8	76.2	11.6
	6	89.8	81.8	8.0
	7	76.1	65.1	11.0
	8	71.5	70.5	1.0
	9	69.5	73.1	-3.6
	10	83.7	70.7	13.0
	All	78.7000	71.6316	7.0684

Taking into ccount company fects, effect size 7.22 (not 7.07) d standard error of estimate is lower.

...the multiple linear regression estimate (of SD) will always give the most efficient weighting to estimates from different levels of a confounding variable in unbalanced situations. Sleuth p. 397

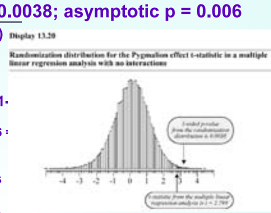
Slide 19 Unbalanced designs & effect sizes

NOTES:

Exact Test for Pygamalion Effect

p=150/2x3^9 =150/39,366≈0.0038; asymptotic p = 0.006

- Pygmalion treatment added 7.2 (±5.4) points to a platoon's score
- Very strong evidence that the Pygmalion effect is real (Fixed effect, randomized block ANOVA, $F_{1,18}$=7.8; 1-sided p = 0.006)
  - Exact p value from randomization =150/39366 = 0.0038
- Because of the randomized design, a causal inference can be made for this group of 10 companies
- If these companies are representative of all army companies, the Pygmalion treatment added 6.84 (±6.42) units to a platoon's score. There is moderate evidence that the effect would be found throughout Army companies (Linear contrast estimate of Pygmalion effect p=0.02)



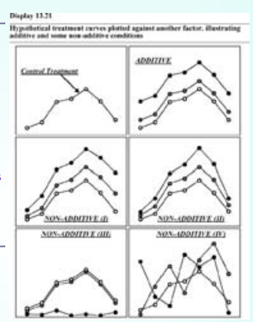
There are 10 companies, the Pygmalion platoon must be randomly assigned **within** each company

Slide 20 Exact Test for Pygamalion Effect

NOTES:

Nonadditivities & interactions

- If there are significant interaction terms, you should usually just present plots of the data
- Some effort should still be made to estimate the effect size
- Non-additive (I) handled with interaction terms
- Non- additive (II) can often be changed to an additive model by transformations
- Non-additive (III) handle separately
- Non-additive (IV) just plot the data (and wave your hands)



Slide 21 Nonadditivities & interactions

NOTES:

Fixed *vs.* Random Factors

Are intertidal areas in Case 13.1 fixed or random and does it matter?
Are the 10 companies in the Case 13.2 (Pygmalion experiment) random or representative samples of all companies, and does it matter?

Yes, it does matter The statistical tests and scope of inference are different

Slide 22 Fixed vs. Random Factors

NOTES:

Fixed vs. Random factors

Expected mean squares from Underwood (1997)

$X_{ij} = \mu + A_i + e_{ij}$

where X_{ij} is jth replicate in ith treatment (ith level of factor A; $i = 1 \ldots a$), A_i is difference between ith level of factor A and overall mean of all levels (μ), e_{ij} is the deviation of replicate j in ith sample from the mean of that population.

One way ANOVA Fixed factor (or Model I)

Fixed factor:
By definition:

$$\sum_{i=1}^{a} A_i = 0$$

(see Section 7.6).

Analysis of variance	Mean square estimates
Among treatments	$\sigma_e^2 + \frac{n\sum_{i=1}^{a}(A_i - \bar{A})^2}{(a-1)}$ or $\sigma_e^2 + nk_A^2$
Within treatments	σ_e^2

where k_A^2 indicates fixed differences, all sampled in the experiment.

Slide 23 Fixed vs. Random factors

NOTES:

Fixed *vs.* Random factors

Tables from Underwood (1997)

One way ANOVA Random factor Model II

Random factor:

$$E\left(\sum_{i=1}^{a} A_i\right) = 0$$

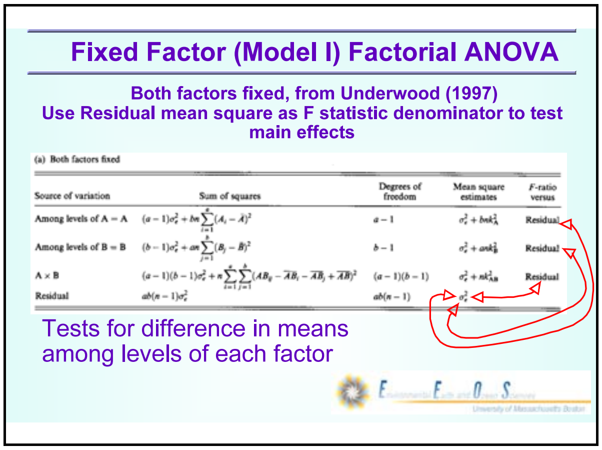
Meaning you *expect* $\sum_{i=1}^{a} A_i = 0$ on average, over many experiments, but in a single experiment, A_i values as sampled may not sum to zero.

Analysis of variance	Mean square estimates
Among treatments	$\sigma_e^2 + n\sigma_A^2$
Within treatments	σ_e^2

where σ_A^2 is the variance of the population of A_i values sampled in your experiment.

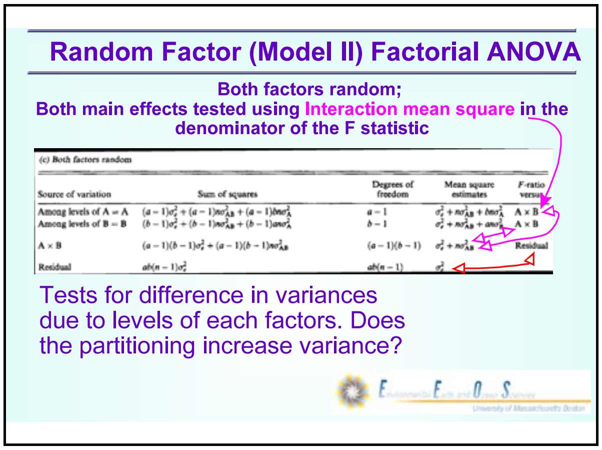
Slide 24 Fixed vs. Random factors

NOTES:



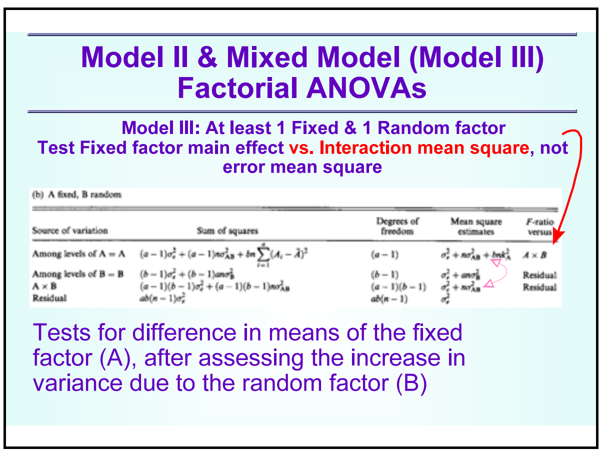
Slide 25 Fixed Factor (Model I) Factorial ANOVA

NOTES:



Slide 26 Random Factor (Model II) Factorial ANOVA

NOTES:



Slide 27 Model II & Mixed Model (Model III) Factorial ANOVAs

NOTES:

SPSS mixed effects ANOVA

If companies were randomly selected; Use if inferences are to be made to a larger population

```
UNIANOVA
score BY pyg company
/RANDOM = company
/CONTRAST (pyg)=Simple
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/PLOT = PROFILE( pyg*company )
/EMMEANS = TABLES(pyg) COMPARE ADJ(LSD)
/PLOT = SPREADLEVEL RESIDUALS
/CRITERIA = ALPHA(.05)
/DESIGN = pyg company pyg*company.
```

If the interaction terms (pyg *company) are not included in the model, then the mixed effects ANOVA is identical to the fixed effects ANOVA

Slide 28 SPSS mixed effects ANOVA

NOTES:

When should a factor be regarded as random instead of fixed?

- Winer *et al.* (1991)
  - If the number of levels of a factor, *p*, is a very small fraction of the number of possible levels of a factor ($P_{effective}$), $p/P_{effective} \approx 0$ and the factor should be regarded as random
  - If the number of levels of a factor *p* is a large fraction of the total number of possible levels, then $p/P_{effective} \approx 1$ and the factor should be regarded as fixed
  - If the levels are random samples of the possible levels, then the factor should be considered random.

Slide 29 When should a factor be regarded as random instead of fixed?

NOTES:

13.2 Companies as a random effect

Test Pygmalion main effect over interaction Mean Square
p value increased from 0.012 to 0.016

Tests of Between-Subjects Effects

Dependent Variable: Score

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	146247.185	1	146247.185	1981.329	<.0000001
	Error	670.688	9.086	73.813[a]		
pyg	Hypothesis	301.843	1	301.843	8.692	**.016**
	Error	318.928	**9.185**	**34.724**[b]		
company	Hypothesis	665.663	9	73.963	2.137	.137
	Error	311.464	9	34.607[c]		
pyg * company	Hypothesis	311.464	9	34.607	.667	.722
	Error	467.040	9	51.893[d]		

a. .993 MS(company) + .007 MS(Error)
b. .993 MS(pyg * company) + .007 MS(Error)
c. MS(pyg * company)
d. MS(Error)

Non-integer *df* due to the unbalanced design

Slide 30 13.2 Companies as a random effect

NOTES:

Effect size of Pygmalion treatment

Two of the 10 companies had a negative Pygmalion effect

Pairwise Comparisons

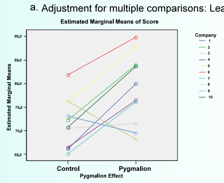
Dependent Variable: Score

(I) Pygmalion Effect	(J) Pygmalion Effect	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
Control	Pygmalion	**-6.840***	**2.836**	**.039**	**-13.256**	**-.424**
Pygmalion	Control	**6.840***	2.836	**.039**	.424	13.256

Based on estimated marginal means

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

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).



With the interaction terms, the Pygmalion effect is reduced from 7.2 to 6.8 units [Identical effect in both fixed & mixed effects models] and p value to 0.02 (1-sided)

Slide 31 Effect size of Pygmalion treatment

NOTES:

Conclusions to Case 13.2

- If the goal is to make inferences to all Army companies and platoons, then companies should have been randomly selected from the statistical population of 'all possible companies' and companies should be treated as a random factor
  - The Pygmalion effect is tested vs. 'Pyg x company' interaction instead of error MS
  - The effect still offers evidence against the no-effect null hypothesis (p=0.008, one-tailed), but the p value is slightly larger than if a fixed effect model were used (p=0.006)

Slide 32 Conclusions to Case 13.2

NOTES:

Conclusion from Chapter 13

- Factorial ANOVA models are a subset of the general linear model
  - Can be analyzed using ANOVA, Regression, or GLM/Univariate
  - The results are mathematically identical
- Fisher noted that factorial ANOVA is superior to testing 1 factor at a time
- Interactions: factors have synergistic effects
  - Interactions must be assessed
    - Note that transforms can eliminate interaction effects
  - Pooling
    - Sleuth doesn't properly cover the problem of pooling interaction terms: use caution when pooling
    - Inappropriate pooling is an example of pseudoreplication & can give rise to Type II error (concluding no interaction or block effect when such effects exist)
    - At the least, use p>0.25 rule
- Random vs. Fixed factors in ANOVA designs
  - The choice should be made *a priori*
  - Interaction MS used as denominator to test main effects in Model II and Model III (mixed model) Factorial ANOVA

Slide 33 Conclusion from Chapter 13

NOTES:

Chapter 14: Multifactor studies without replication

& Nested ANOVA

ECOS611

Slide 34 Chapter 14: Multifactor studies without replication

NOTES:

Strategies for analyzing tables with one observation per cell

- Often it is more important to evaluate different levels of factors than to provide replicates
- Without replication, only non-saturated models can be fit:
  - The interaction terms can not be estimated
  - you must make assumptions (e.g., linear relation, no interactions) and test them where possible
- Approach
  - Graphical displays of the data
  - If any of the explanatory variables can be treated as continuous, attempt to fit this simpler model
  - For categorical variables, test for additivity

Slide 35 Strategies for analyzing tables with one observation per cell

NOTES:

Case 14.01

Chimpanzee learning

- 4 chimps, including 2 males
- Time to learn 10 words
- Test chimp-to-chimp differences and word - learning differences
- No replication possible

Display 14.1

Minutes to acquisition of American Sign Language signs by four chimps

	listen	drink	shoe	key	more	food	fruit	hat	look	string
Booee	12	15	14	10	10	80	80	78	115	129
Cindy	10	25	18	25	15	55	20	99	54	476
Bruno	2	36	60	40	225	14	177	178	345	287
Thelma	15	18	20	40	24	190	195	297	420	372

Slide 36 Case 14.01

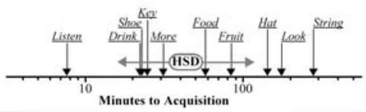
NOTES:

Statistical results, Case 14.1

Tukey-Kramer *a posteriori* test, based on studentized range

Display 14.2

Multiple comparisons of sign means on the log scale



- Inconclusive evidence that some chimps are faster learners than others (p = 0.064, $F_{3,27}$)
- Some words (shown above) easier to learn than others (p = 0.00002, $F_{9,27}$)

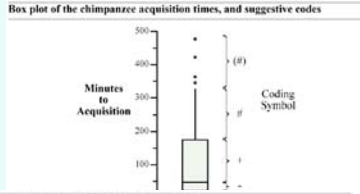
Slide 37 Statistical results, Case 14.1

NOTES:

Coded boxplot

A poor graphic: too complicated

Box plot of the chimpanzee acquisition times, and suggestive codes



Coded two-way table for the chimpanzee data

	Listen	Drink	Shoe	Key	More	Food	Fruit	Hat	Look	String
Booee	=	=	=	=	=	+	+	+	+	+
Cindy	=	-	-	-	=	+	-	+	+	(#)
Bruno	=	-	+	-	#	=	+	#	(#)	#
Thelma	=	-	-	-	-	#	#	#	(#)	(#)

Slide 38 Coded boxplot

NOTES:

Table of estimated means

Available in GLM: Sign effect and Chimp effect

Observed values, fitted values, and residuals for the additive fit to the chimpanzee sign acquisition times (all in minutes)

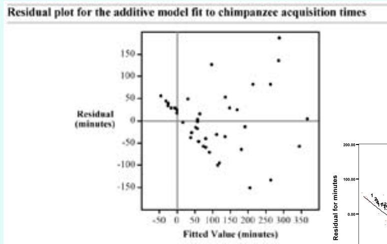
Y fitted value residual	Booee	Cindy	Bruno	Thelma	Average	Sign Effect
listen	12 -43.325 +55.325	10 -17.925 +27.925	2 38.775 -36.775	15 61.475 -46.475	9.75	-97.625
drink	13 -29.575 +42.575	25 -4.175 +29.175	36 52.525 -16.525	18 75.225 -57.225	23.50	-83.875
shoe	14 -25.075 +39.075	18 0.325 +17.675	60 57.025 +2.975	20 79.725 -59.725	28.00	-79.375
key	10 -24.325 +34.325	25 1.075 +23.925	40 57.775 -17.775	40 80.475 -40.475	28.75	-78.625
more	10 15.425 -5.425	15 40.825 -25.825	225 97.525 +127.475	24 120.225 -96.225	68.50	-38.875
food	80 31.675 +48.325	55 57.075 -2.075	14 113.775 -99.775	190 136.475 +53.525	84.75	-22.625
fruit	80 64.925 +15.075	20 90.325 -70.325	177 147.025 +29.975	195 169.725 +25.275	118.00	+10.625
hat	78 109.925 -31.925	99 135.325 -36.325	178 192.025 -14.025	297 214.725 +82.275	163.00	+55.625
look	115 180.425 -65.425	54 205.825 -151.825	345 262.525 +82.475	420 285.225 +134.775	233.50	+126.125
string	129 262.925 -133.925	476 288.325 +187.675	287 345.025 -58.025	372 367.725 +4.275	316.00	+208.625
Average	54.3	79.7	136.4	159.1	107.375	
Chimp Effect	-53.075	-27.675	+29.025	+51.725		

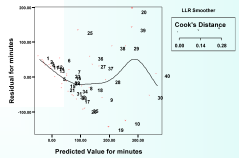
Slide 39 Table of estimated means

NOTES:

Horn-shaped residuals

ANOVA & regression share the same assumptions



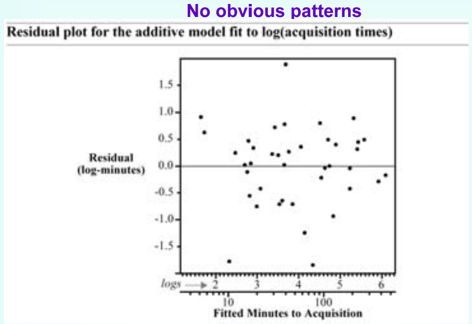


Slide 40 Horn-shaped residuals

NOTES:

Log-transformed learning times

No obvious patterns



Slide 41 Log-transformed learning times

NOTES:

No replication: no interaction test

Tests of Between-Subjects Effects

Dependent Variable: ln(Minutes)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	**68.675**[a]	**39**	**1.761**	.	.
Intercept	615.158	1	615.158	.	.
chimp	5.333	3	1.778	.	.
sign	45.690	9	5.077	.	.
chimp * sign	17.653	27	.654	.	.
Error	.000	0	.		
Total	683.834	40			
Corrected Total	68.675	39			

a. R Squared = 1.000 (Adjusted R Squared = .)

Insufficient *df* to estimate all parameters in the full factorial (saturated) model with no replication

Slide 42 No replication: no interaction test

NOTES:

Acquisition time as a blocked ANOVA, chimp as a blocking variable

Modest evidence for chimp-to-chimp differences, p=0.06

Display 14.9

Analysis of variance for the additive model fit to log(acquisition times)

Source of Variation	Sum of Squares	df	Mean Square	F-Statistic	p-value
Signs	45.6900	9	5.0767	7.7649	0.00001
Chimpanzees	5.3329	3	1.7776	2.7190	0.0642
Residual	17.6526	27	0.6538		
Total	68.6755	39			

R-squared = 74.3% **Estimated SD** = 0.8086

One assumes that there are no block x treatment interactions

Slide 43 Acquisition time as a blocked ANOVA, chimp as a blocking variable

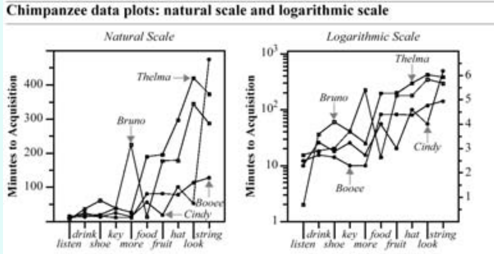
NOTES:

Log-transformation & Interaction

Despite lines crossing, qualitatively argue for no interactions. There is Tukey's additivity test (Quinn & Keough p. 278)

Display 14.10

Chimpanzee data plots: natural scale and logarithmic scale



Slide 44 Log-transformation & Interaction

NOTES:

Tukey one degree of freedom test

= Tukey test for additivity in factorial ANOVA with n=1

- Tukey developed a test for interaction in 2-factor designs, with n=1 (single replicates of each factor combination)
  - Neter *et al.* (1996, p. 882) discuss the test
  - If variables are quantitative, use regression
- Tukey's additivity test in SPSS reliability not appropriate for factorial ANOVA
- If the Tukey additivity test is positive, then the interaction effect can't be ignored or assumed away
  - Transform the variables to achieve an additive model
  - Neter et al. (1996) cite Johnson & Graybill (1972) for an approximate 2-factor test if Tukey's test for interactions is positive

Slide 45 Tukey one degree of freedom test

NOTES:

Matlab’s Tukey additivity test

Available on the Mathworks file exchange

- % Trujillo-Ortiz, A., R. Hernandez-Walls and R. Castro-Valdez. (2003).
- % adTukeyAOV2: Tukey's test of additivity for a two-way classification
- % Analysis of Variance. A MATLAB file.
- % [WWW document]. URL http://www.mathworks.com/matlabcentral/fileexchange/

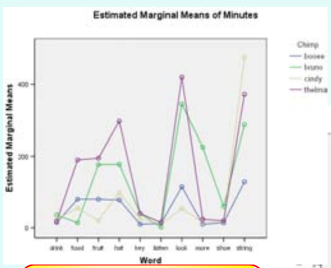
```
Chimps and learning words

The number of levels of factor 1 are:10
The number of levels of factor 2 are: 4
```

SOV	SS	df	MS	F	P
Residual	17.6463	27			
Nonaddit.	0.1671	1	0.167	0.25	0.6223
Remainder	17.4792	26	0.672		

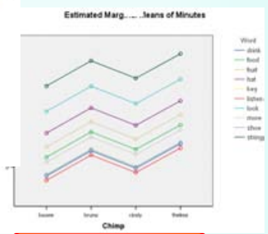
```
The hypothesis of additivity is tenable.
```

Slide 46 Matlab’s Tukey additivity test

NOTES:



With interactions



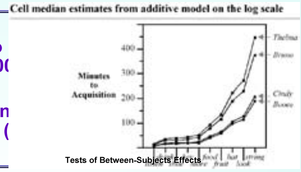
No interactions

Slide 47

NOTES:

Summary of statistical findings

- Convincing evidence that some signs take longer to learn than others (p=0.000 $F_{9,27}$)
- Only slight evidence for an difference among chimps (0.064, $F_{3,27}$)



Tests of Between-Subjects Effects

Dependent Variable: ln(Minutes)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	51.023[a]	12	4.252	6.503	2.78E-005
Intercept	615.158	1	615.158	940.896	1.59E-022
chimp	5.333	3	1.778	2.719	.064
sign	45.690	9	5.077	7.765	1.50E-005
Error	17.653	27	.654		
Total	683.834	40			
Corrected Total	68.675	39			

a. R Squared = .743 (Adjusted R Squared = .629)

Slide 48 Summary of statistical findings

NOTES:

Case 14.2

Effects of ozone in conjunction with sulfur dioxide and water stress on soybean yield – a randomized experiment

Case 14.2 Not covered in 2009: Replace a categorical variable with a continuous variable to free up df to test interactions

Slide 49 Case 14.2

NOTES:

Case 14.02

Soybean yield = f (ozone, moisture, sulfur dioxide)

- **Two different soybean strains**
- **Split-plot design**
  - Chambers as whole plots, cultivars as split plots
  - Handled here as 2 separate analyses
- **3 factors**
  - Ozone (5 levels), previously documented
  - Sulfur dioxide (3 levels)
  - Water stress (2 levels)
- **3 hypotheses**
  - Does soil moisture stress affect yield in addition to SO_2 and ozone?
  - 3-way interactions?
  - Does stress effect the 2 cultivars differently?

Display 14.3 p. 400

Seed yields for soybean cultivars *Forrest* and *Williams* from chambers kept under varying conditions of ozone, sulphur dioxide and water stress

Water Stress	SO_2 (µL/L)	O_3 (µL/L)	Yields (kg/ha) Forrest	Yields (kg/ha) Williams
Well-Watered (WW) (-0.05 MPa)	0.0045	0.017	4376	5561
		0.049	4544	5947
		0.067	2806	4273
		0.084	3339	3470
		0.099	3320	3080
	0.0170	0.017	3747	5092
		0.049	4570	4752
		0.067	4635	4232
		0.084	3613	2867
		0.099	3259	3106
	0.0590	0.017	4179	4736
		0.049	5077	3672
		0.067	3401	3386
		0.084	3371	2854
		0.099	2158	2557
Soil Moisture Stress (SMS) (-0.40 MPa)	0.0045	0.017	4977	4520
		0.049	3780	3047
		0.067	3804	3526
		0.084	3941	3357
		0.099	2863	2663
	0.0170	0.017	3573	4869
		0.049	3555	3774
		0.067	3340	2955
		0.084	3243	2513
		0.099	2802	2838
	0.0590	0.017	4589	4056
		0.049	3250	2758
		0.067	3045	3094
		0.084	2827	2398
		0.099	2979	2101

Slide 50 Case 14.02

NOTES:

Summary of statistical findings

Case study 14.2 Forrest cultivar; Originally a split-plot design

- Strong effect of ozone on yield
  - Fixed effect ANOVA F, 2-sided $p < 0.001$
  - A 0.01 µl/L increase in ozone decreased median yield by 5.3% (95% CI: 3.4 to 7%)
- No effect of SO_2 on yield
  - 2-sided $p = 0.13$
  - Effect: 1.6% reduction (-0.5% to 3.5%)
- No effect of water stress
  - Fixed effects ANOVA F, 2-sided $p = 0.55$
  - Effect: 3.3% increase (-7.3% to 15.3%)
- No interactions, but weak power
  - Ozone effect when SO_2 is 0.059 is 14.7% of effect when ozone is 0.0045 µl/L (0.16 to 1365%)
  - Ozone effect with water stress is 450% vs. Watered (9.2% to 22000%)

Slide 51 Summary of statistical findings

NOTES:

Summary of statistical findings

Case study 14.2 Williams cultivar (differences from Forrest)

- Strong effect of ozone on yield
  - Fixed effect ANOVA F, 2-sided $p < 0.001$
  - A 0.01 μl/L increase in ozone decreased median yield by 6.6% (95% CI: 5.3 to 7.9%)
- Strong effect of SO_2 on yield
  - 2-sided $p < 0.0001$
  - Effect: a 0.01 μl/L increase in SO_2 results in a 3.5% reduction in median yield (2% to 4.9%)
- Strong effect of water stress
  - Fixed effects ANOVA F, 2-sided p = 0.0001
  - Effect: -0.04MPa water stress reduces median yield 19.4% (10% to 30%)
- No interactions, but weak power
  - Ozone effect when SO_2 is 0.059 μl/L is 40.7% of effect when ozone is 0.0045 μl/L (1.47 to 384.5%)
  - Ozone effect with water stress (-0.40 Mpa) is 24% vs. Watered (1.4% to 390%)

Slide 52 Summary of statistical findings

NOTES:

Strategies for analyzing tables with one observation per cell

- Often it is more important to evaluate different levels of factors than to provide replicates
- Without replication, only non-saturated models can be fit:
  - The interaction terms can **not** be estimated
  - you must make assumptions (e.g., linear relation, no interactions) and test them where possible
- Approach
  - Graphical displays of the data
  - **If any of the explanatory variables can be treated as numerical (i.e., treat as a continuous covariate), attempt to fit this simpler model**

Slide 53 Strategies for analyzing tables with one observation per cell

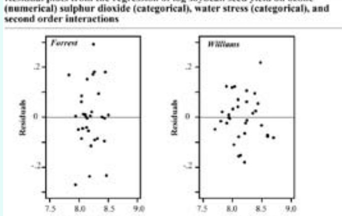
NOTES:

Residual plots to assess model misspecification

Log (yield): No obvious differences between cultivars

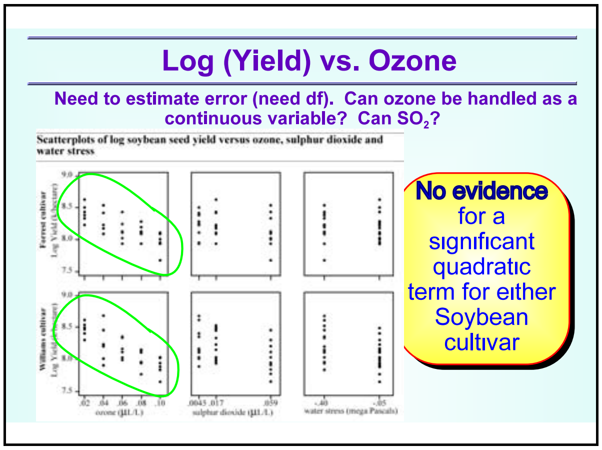
Display 14.13

Residual plots from the regression of log soybean seed yield on ozone (numerical) sulphur dioxide (categorical), water stress (categorical), and second order interactions



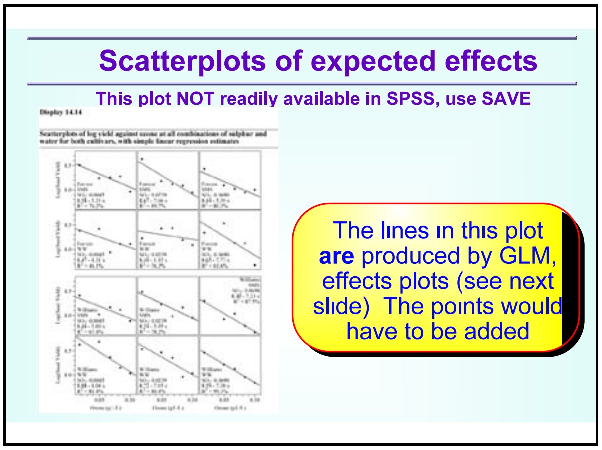
Slide 54 Residual plots to assess model misspecification

NOTES:



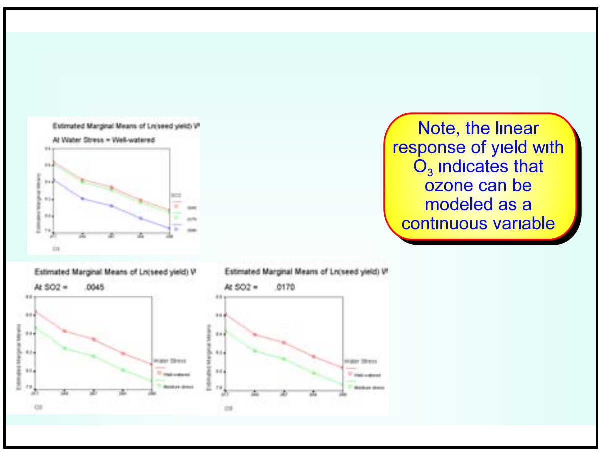
Slide 55 Log (Yield) vs. Ozone

NOTES:



Slide 56 Scatterplots of expected effects

NOTES:



Slide 57 Williams cultivar

NOTES:

ANOVA tables for Soybean yield

Analysis of variance tables for screening effects on log soybean seed yield

***Forrest* Cultivar**

Source	df	Sum of Squares	Mean Square	F-stat	p-value
ozone	1	.7208	.7208	29.4	.00003
SULPHUR	2	.0635	.0317	1.30	.30
water	1	.0080	.0080	.328	.57
ozone×SULPHUR	2	.0173	.0087	.353	.71
ozone×water	1	.0136	.0136	.556	.46
SULPHUR×water	2	.0285	.0143	.583	.57
ozone×SULPHUR×water	2	.0683	.0342	1.46	.26
residuals	18	.4211	.0234		

***Williams* Cultivar**

Source	df	Sum of Squares	Mean Square	F-stat	p-value
ozone	1	1.150	1.150	86.8	<.00001
SULPHUR	2	.2780	.1390	10.5	.001
water	1	.2376	.2376	17.9	.0005
ozone×SULPHUR	2	.0037	.0019	.140	.87
ozone×water	1	.0128	.0128	.964	.34
SULPHUR×water	2	.0263	.0131	.004	.39
ozone×SULPHUR×water	2	.0093	.0047	.352	.71
residuals	18	.2384	.0132		

No evidence for a sulphur or water effect on Forrest, but strong evidence on Williams

Slide 58 ANOVA tables for Soybean yield

NOTES:

Conclusion about main effects

Sulfur dioxide to be handled as a continuous variable

Display 14.16

Coefficient estimates and standard errors for the linear soybean models, with Y = log(soybean seed yield)

Variable	Forrest Coefficient	Forrest St. Error	Forrest 2-Sided p-Value	Williams Coefficient	Williams St. Error	Williams 2-Sided p-Value
CONSTANT	8.608	0.080		8.825	0.058	
ozone	-5.397	0.929	<.0001	-6.806	0.679	<.0001
sulphur	-1.566	0.989	.1252	-3.512	0.723	<.0001
water	0.094	0.153	.5453	0.507	0.112	.0001

Page 425 Ozone treated as a continuous variable to free df for interaction F tests No evidence for a quadratic effect for sulfur, so SO_2 also modeled as a continuous variable

Slide 59 Conclusion about main effects

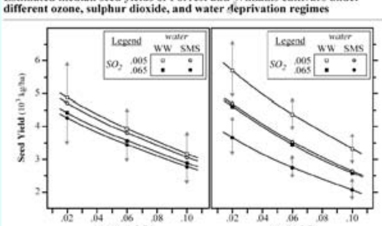
NOTES:

Final results for Soybean yield

To produce in SPSS: Run regression with data as extra rows and Hotelling-Working-Scheffe's mulitplier, p. 266-267

Display 14.17

Estimated median seed yields of Forrest and Williams cultivars under different ozone, sulphur dioxide, and water deprivation regimes



Slide 60 Final results for Soybean yield

NOTES:

Is there really no interaction?

Descriptive summary of interactions, p. 413-414
Low power for detecting interaction effects

- $\mu\{logyield\}=\beta_o+\beta_1 ozone+\beta_2 sulfur+\beta_3 water+\beta_4(ozone \times water)$, $\beta_4=-29.5 \pm 34.5$
- Forrest cultivar
  - Sulfur effect
    - Estimate at high sulfur dioxide, the rate of decline is 14% of the rate of decline at low sulfur dioxide
    - The 95% CI is 0.2% to 1450 %
  - Water effect: water stress produces a 430% decline vs. Ozone [9.2% to 20,100%]
- Williams cultivar
  - Sulfur effect:
    - High sulfur, rate of decline with ozone is 41% rate of decline at low sulfur
    - 95% CI is 1.4% to 1197%
  - Ozone effect: 23.8% under stress (1.5% to 390%)

Slide 61 Is there really no interaction?

NOTES:

Case Study 14.2 was based on a Split-plot design. These designs are common in agriculture and industrial applications, but less common in environmental science. The following slides present an example of a split-plot design to assess the effects of trawling on benthic communities

Slide 62 Case Study 14.2 was based on a Split-plot design. These designs are common in agriculture and industrial applications, but less common in environmental science. The following slides present an example of a split-plot design to assess the effects of trawling on benthic communities

NOTES:

Split-plot designs

Multiple treatment levels are nested within a larger treatment level, from my statistical terms appendix

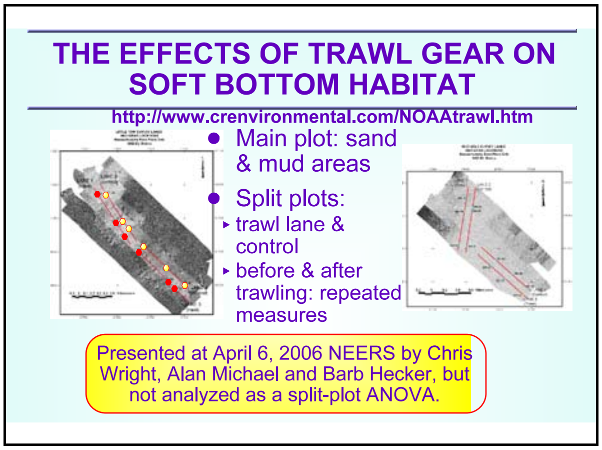
For example, an entire field could receive a given level of fertilizer, and different watering levels could be used on different portions of the field. Or, different greenhouses could be used to control temperature for a large number of trays of plants, and then different watering levels and fertilizer levels could be used within different areas or blocks of each greenhouse. The ANOVA table is often split, with tests of the main plot being based on a partition of the degrees of freedom of the main plots (*e.g.,* fields or greenhouses), whereas the factors being assessed in the subplots (*e.g.,* water or fertilizer level) can be assessed with error terms incorporating a much larger number of degrees of freedom. Cochran & Cox (1957, p. 296-297) compare split plot and randomized blocks design with **A** being the main factor and **B** being the split-plot factor:

1) **B** and **AB** effects estimated more precisely than **A** effects in the split-plot design
2) Overall experimental error is the same between designs: increased precision on **B** and **AB** effects are at the expense of precision for tests of **A** effects,
3) The chief advantage of the split plot over the factorial is combining factors that are expensive to create (the **A** or main plot factors) with relatively inexpensive subplot factors.

Consider the use of a split plot design when **B** and **AB** effects of more interest than **A**, or if the **A** effects can not be fully replicated with small amounts of resources.

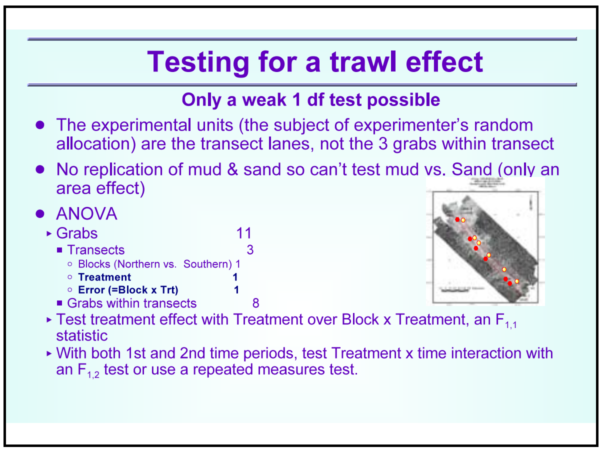
Slide 63 Split-plot designs

NOTES:



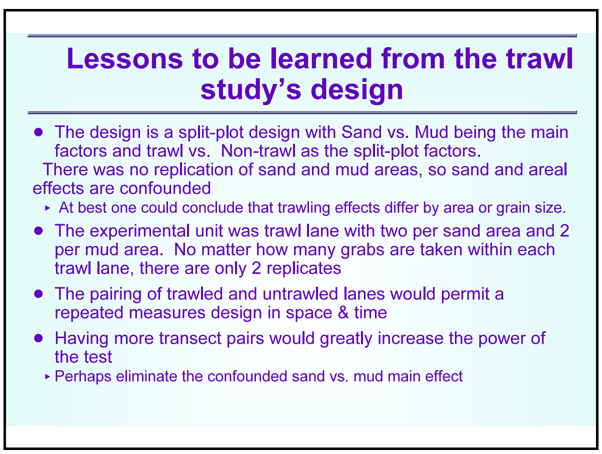
Slide 64 THE EFFECTS OF TRAWL GEAR ON SOFT BOTTOM HABITAT

NOTES:



Slide 65 Testing for a trawl effect

NOTES:



Slide 66 Lessons to be learned from the trawl study's design

NOTES: