

Chapter 14 : Multifactor studies without replication

Linear Models – $Y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij}$ This is a randomized block design with no replication beyond the treatment by block interaction. The treatment arrangement can be a single factor or factorial. In this case the treatments are fixed. The EMS below a generic example with t levels of the treatment and b levels of the block.

Source	d.f.	EMS (random)
Block	$b - 1$	$\sigma^2 + t\sigma^2_{\beta}$
Treatment	$t - 1$	$\sigma^2 + b\sigma^2_{\tau}$
Error (interaction)	$(t - 1)(b - 1)$	σ^2

If the two sources are both considered “treatments of interest” $Y_{ij} = \mu + \tau_{i1} + \tau_{i2} + \varepsilon_{ij}$ the nature of the error term is in doubt. If one of the treatments is random, then the interaction is an appropriate error term, but cannot itself be tested. If the effects are fixed, then the interaction would be fixed. In this case the test is suspect unless we know that there are no interactions. This basically requires that the model be additive.

Source	d.f.	EMS (fixed)
Treatment 1	$t_1 - 1$	$\sigma^2 + Q_{\tau_1}$
Treatment 2	$t_2 - 1$	$\sigma^2 + Q_{\tau_2}$
Error (interaction)	$(t_1 - 1)(t_2 - 1)$	$\sigma^2 + Q_{\tau_1\tau_2}$

Chimpanzee example

In this study a researcher taught 10 signs of American Sign Language (ASL) to four chimpanzees. (Data from R. S. Fouts, "Acquisition and Testing of Gestural Signs in Four Young Chimpanzees," Science 180 (1973): 978-80.) The goals were to determine whether some signs were more easily acquired than others and whether some chimps tended to learn signs more quickly than other chimps. The subjects were four young chimpanzees—males Bruno and Booe, and females Cindy and Thelma. The ASL signs—hat, shoe, fruit, drink, more, look, key, listen, string, and food—covered a wide range of objects, actions, and concepts. Chimpanzees were taught individually, using a system of rewards (and in the case of a reluctant Bruno, threats), until they could successfully produce unprompted responses on five consecutive occasions. Display 14.1 shows the times, in minutes, required to teach each sign to each subject, according to that criterion. The factors are arranged according to increasing average times.

Chimp	Listen	Drink	Shoe	Key	More	Food	Fruit	Hat	Look	String	Mean
Booe	12	15	14	10	10	80	80	78	115	129	54.3
Cindy	10	25	18	25	15	55	20	99	54	476	79.7
Bruno	2	36	60	40	225	14	177	178	345	287	136.4
Thelma	15	18	20	40	24	190	195	297	420	372	159.1
Mean	9.75	23.5	28	28.75	68.5	84.75	118	163	233.5	316	107.38

```

1 *****;
2 *** Chimpanzee language study ***;
3 *****;
4
5 dm'log;clear;output;clear';
6 options nodate nocenter nonumber ps=512 ls=132 nolabel;
7 ODS HTML style=minimal rs=none
body='C:\Geaghan\Current\EXST3201\Fall2005\SAS\Chimpanzee01.html' ;
NOTE: Writing HTML Body file: C:\Geaghan\Current\EXST3201\Fall2005\SAS\Chimpanzee01.html
8
9 Title1 'Chapter 14 : Chimpanzee sign language study';
10 filename input1 'C:\Geaghan\Current\EXST3201\Datasets\ASCII\case1401.csv';
11
12 data Chimp; infile input1 misover DSD dlm="," firstobs=2;
13 input MINUTES CHIMP $ SIGN $;
14 label Minutes = 'Time needed to learn the sign'
15 Chimp = 'Chimpanzee name'
16 Sign = 'Sign learned';
17 logMinutes = log(minutes);
18 ID = chimp; if chimp = 'BOOEE' then ID = 'A';
19 datalines;
NOTE: The infile INPUT1 is:
File Name=C:\Geaghan\Current\EXST3201\Datasets\ASCII\case1401.csv,
RECFM=V,LRECL=256
NOTE: 40 records were read from the infile INPUT1.
The minimum record length was 16.
The maximum record length was 21.
NOTE: The data set WORK.CHIMP has 40 observations and 5 variables.
NOTE: DATA statement used (Total process time):
real time 0.03 seconds
cpu time 0.03 seconds
20 run;
21
22 PROC PRINT DATA=Chimp; TITLE2 'Data Listing';
23 RUN;

```

NOTE: There were 40 observations read from the data set WORK.CHIMP.

NOTE: The PROCEDURE PRINT printed page 1.

NOTE: PROCEDURE PRINT used (Total process time):

```

real time 0.25 seconds
cpu time 0.06 seconds

```

Chapter 14 : Chimpanzee sign language study
Data Listing

Obs	MINUTES	CHIMP	SIGN	log Minutes	ID						
1	12	BOOEE	listen	2.48491	A	19	54	CINDY	look	3.98898	CINDY
2	15	BOOEE	drink	2.70805	A	20	476	CINDY	string	6.16542	CINDY
3	14	BOOEE	shoe	2.63906	A	21	2	BRUNO	listen	0.69315	BRUNO
4	10	BOOEE	key	2.30259	A	22	36	BRUNO	drink	3.58352	BRUNO
5	10	BOOEE	more	2.30259	A	23	60	BRUNO	shoe	4.09434	BRUNO
6	80	BOOEE	food	4.38203	A	24	40	BRUNO	key	3.68888	BRUNO
7	80	BOOEE	fruit	4.38203	A	25	225	BRUNO	more	5.41610	BRUNO
8	78	BOOEE	hat	4.35671	A	26	14	BRUNO	food	2.63906	BRUNO
9	115	BOOEE	look	4.74493	A	27	177	BRUNO	fruit	5.17615	BRUNO
10	129	BOOEE	string	4.85981	A	28	178	BRUNO	hat	5.18178	BRUNO
11	10	CINDY	listen	2.30259	CINDY	29	345	BRUNO	look	5.84354	BRUNO
12	25	CINDY	drink	3.21888	CINDY	30	287	BRUNO	string	5.65948	BRUNO
13	18	CINDY	shoe	2.89037	CINDY	31	15	THELMA	listen	2.70805	THELMA
14	25	CINDY	key	3.21888	CINDY	32	18	THELMA	drink	2.89037	THELMA
15	15	CINDY	more	2.70805	CINDY	33	20	THELMA	shoe	2.99573	THELMA
16	55	CINDY	food	4.00733	CINDY	34	40	THELMA	key	3.68888	THELMA
17	20	CINDY	fruit	2.99573	CINDY	35	24	THELMA	more	3.17805	THELMA
18	99	CINDY	hat	4.59512	CINDY	36	190	THELMA	food	5.24702	THELMA
						37	195	THELMA	fruit	5.27300	THELMA
						38	297	THELMA	hat	5.69373	THELMA
						39	420	THELMA	look	6.04025	THELMA
						40	372	THELMA	string	5.91889	THELMA

The variable “ID” is created as an identifier for graphics. Only the first letter will be used, and it was needed because two chimp names started with “B”.

The PROC MIXED below is run with “CHIMP” and “SIGN” included as fixed effects. The interaction is omitted and becomes the default error term which would be a random effect.

```

25     PROC mixed DATA=chimp cl covtest; class chimp sign;
26         Title2 'Unreplicated multifactor study';
27         MODEL minutes = chimp sign / outp=resids;
28         lsmeans chimp sign / pdiff adjust=tukey;
29         ods output diffs=ppp lsmeans=mmm;
30         ods listing exclude diffs;* lsmeans; *this is now just a comment;
31     run;
NOTE: The data set WORK.MMM has 14 observations and 8 variables.
NOTE: The data set WORK.PPP has 51 observations and 12 variables.
NOTE: The data set WORK.RESIDS has 40 observations and 12 variables.
NOTE: The PROCEDURE MIXED printed page 2.
NOTE: PROCEDURE MIXED used (Total process time):
      real time          0.61 seconds
      cpu time           0.52 seconds
32     TITLE3 'Post hoc adjustment with macro by Arnold Saxton';
33     %include 'C:\Geaghan\Current\EXST3201\Fall2005\SAS\pdmix800.sas';
706    %pdmix800(ppp,mmm,alpha=.05,sort=yes);
PDMIX800 08.08.2003 processing
3.8700871961
Tukey values for CHIMP are 102.414 (avg) 102.414 (min) 102.414 (max).
4.8644540923
Tukey values for SIGN are 203.537 (avg) 203.537 (min) 203.537 (max).
707    RUN;
708    QUIT;

```

Chapter 14 : Chimpanzee sign language study
Unreplicated multifactor study

The Mixed Procedure

Model Information

Data Set	WORK.CHIMP
Dependent Variable	MINUTES
Covariance Structure	Diagonal
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Residual

Class Level Information

Class	Levels	Values
CHIMP	4	BOOEE BRUNO CINDY THELMA
SIGN	10	drink food fruit hat key listen look more shoe string

Dimensions

Covariance Parameters	1
Columns in X	15
Columns in Z	0
Subjects	1
Max Obs Per Subject	40

Number of Observations	
Number of Observations Read	40
Number of Observations Used	40
Number of Observations Not Used	0

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr Z	Alpha	Lower	Upper
Residual	7002.94	1905.96	3.67	0.0001	0.05	4377.39	12974

This residual is the variation that remains after fitting the main effects. The only variation remaining for this analysis is the “chimp*sign” interaction.

Fit Statistics

-2 Res Log Likelihood	335.1
AIC (smaller is better)	337.1
AICC (smaller is better)	337.2
BIC (smaller is better)	338.4

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
CHIMP	3	27	3.38	0.0326
SIGN	9	27	5.95	0.0001

Note that there is not test of interaction. This term was used as the error term.

Least Squares Means

Effect	CHIMP	SIGN	Estimate	Standard Error	DF	t Value	Pr > t
CHIMP	BOOEE		54.3000	26.4631	27	2.05	0.0500
CHIMP	BRUNO		136.40	26.4631	27	5.15	<.0001
CHIMP	CINDY		79.7000	26.4631	27	3.01	0.0056
CHIMP	THELMA		159.10	26.4631	27	6.01	<.0001
SIGN		drink	23.5000	41.8418	27	0.56	0.5790
SIGN		food	84.7500	41.8418	27	2.03	0.0528
SIGN		fruit	118.00	41.8418	27	2.82	0.0089
SIGN		hat	163.00	41.8418	27	3.90	0.0006
SIGN		key	28.7500	41.8418	27	0.69	0.4979
SIGN		listen	9.7500	41.8418	27	0.23	0.8175
SIGN		look	233.50	41.8418	27	5.58	<.0001
SIGN		more	68.5000	41.8418	27	1.64	0.1132
SIGN		shoe	28.0000	41.8418	27	0.67	0.5091
SIGN		string	316.00	41.8418	27	7.55	<.0001

Effect=CHIMP ADJUSTMENT=Tukey(P<.01) bygroup=1

Obs	CHIMP	SIGN	Estimate	StdErr	MSGROUP
1	THELMA		159.10	26.4631	A
2	BRUNO		136.40	26.4631	A
3	CINDY		79.7000	26.4631	A
4	BOOEE		54.3000	26.4631	A

Effect=SIGN ADJUSTMENT=Tukey(P<.01) bygroup=2

Obs	CHIMP	SIGN	Estimate	StdErr	MSGROUP
5		string	316.00	41.8418	A
6		look	233.50	41.8418	AB
7		hat	163.00	41.8418	AB
8		fruit	118.00	41.8418	AB
9		food	84.7500	41.8418	AB
10		more	68.5000	41.8418	B
11		key	28.7500	41.8418	B
12		shoe	28.0000	41.8418	B
13		drink	23.5000	41.8418	B
14		listen	9.7500	41.8418	B

710 proc univariate data=resids plot normal; var resid; run;

NOTE: The PROCEDURE UNIVARIATE printed page 4.

NOTE: PROCEDURE UNIVARIATE used (Total process time):

real time 0.13 seconds
cpu time 0.04 seconds

Chapter 14 : Chimpanzee sign language study
Unreplicated multifactor study
Post hoc adjustment with macro by Arnold Saxton

The UNIVARIATE Procedure

Variable: Resid

Moments

N	40	Sum Weights	40
Mean	0	Sum Observations	0
Std Deviation	69.6289387	Variance	4848.1891
Skewness	0.29164063	Kurtosis	0.68447068
Uncorrected SS	189079.375	Corrected SS	189079.375
Coeff Variation	.	Std Error Mean	11.0093019

Basic Statistical Measures

Location		Variability	
Mean	0.000000	Std Deviation	69.62894
Median	0.450000	Variance	4848
Mode	.	Range	339.50000
		Interquartile Range	80.17500

Tests for Location: Mu0=0

Test	-Statistic-		-----p Value-----
Student's t	t	0	Pr > t 1.0000
Sign	M	0	Pr >= M 1.0000
Signed Rank	S	-15	Pr >= S 0.8432

Tests for Normality

Test	--Statistic--		-----p Value-----
Shapiro-Wilk	W	0.983199	Pr < W 0.8055
Kolmogorov-Smirnov	D	0.088432	Pr > D >0.1500
Cramer-von Mises	W-Sq	0.034289	Pr > W-Sq >0.2500
Anderson-Darling	A-Sq	0.255314	Pr > A-Sq >0.2500

Quantiles (Definition 5)

Quantile	Estimate
100% Max	187.675
99%	187.675
95%	131.125
90%	82.375
75% Q3	36.700
50% Median	0.450
25% Q1	-43.475
10%	-83.275
5%	-116.850
1%	-151.825
0% Min	-151.825

Extreme Observations

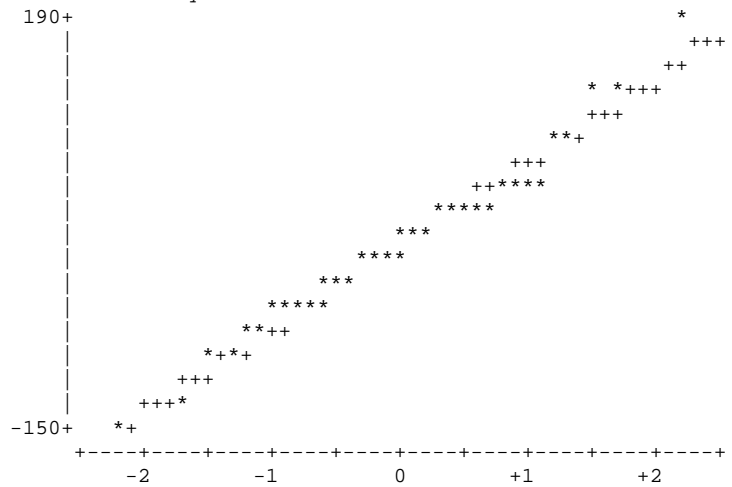
-----Lowest-----		-----Highest-----	
Value	Obs	Value	Obs
-151.825	19	82.275	38
-133.925	10	82.475	29
-99.775	26	127.475	25
-96.225	35	134.775	39
-70.325	17	187.675	20

Stem Leaf Boxplot

18	8	1	0
16			
14			
12	75	2	
10			
8	22	2	
6			
4	5845	4	
2	4589049	7	+-+--+-+
0	3458	4	*-+--*+
-0	87452	5	
-2	7626	4	
-4	8760	4	+-+--+-+
-6	050	3	
-8	6	1	
-10	0	1	
-12	4	1	
-14	2	1	

-----+-----+-----+
 Multiply Stem.Leaf by 10**+1

Normal Probability Plot



```
711 options ps=52 ls=111;
712 proc plot data=resids; plot resid*pred; run;
713 options ps=512 ls=132;
```

NOTE: There were 40 observations read from the data set WORK.RESIDS.

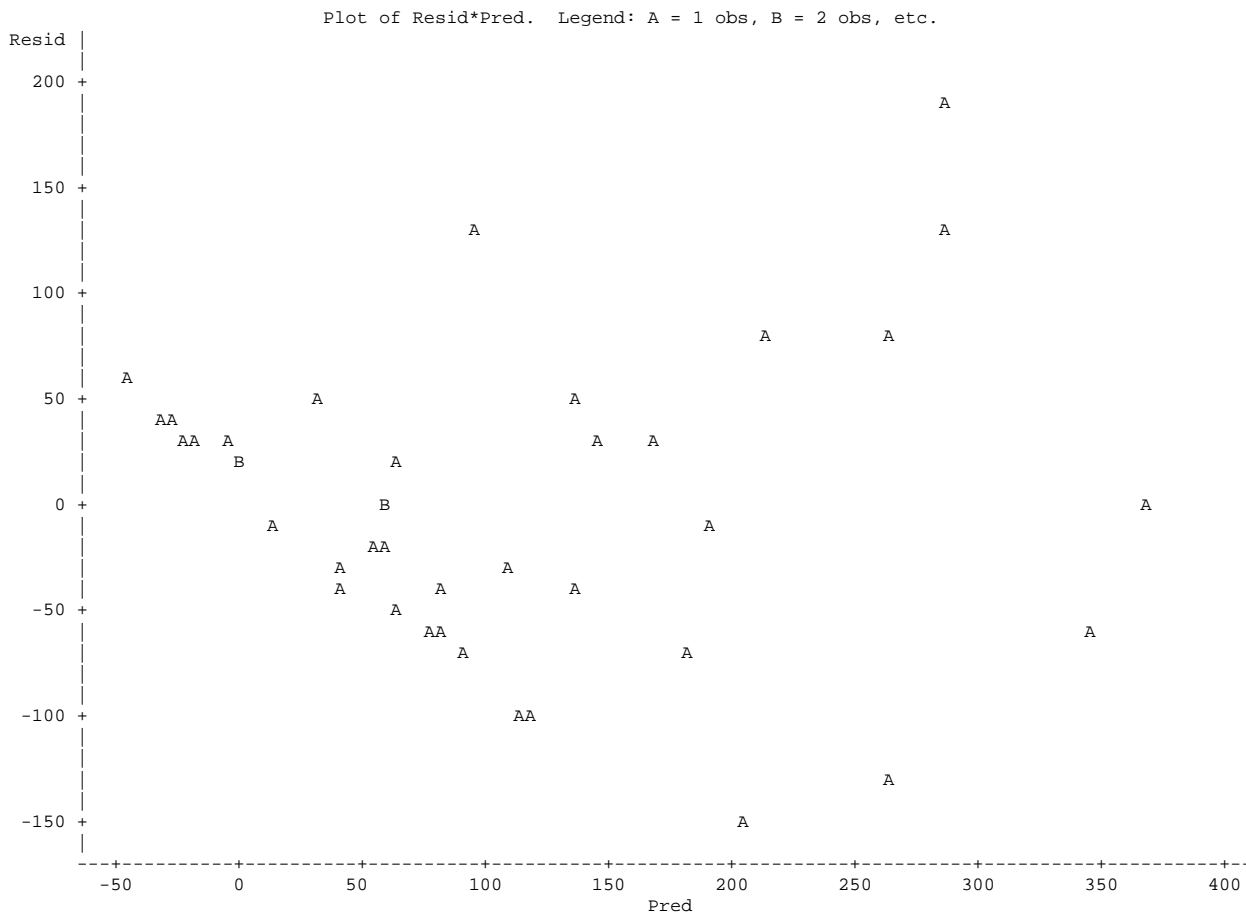
NOTE: The PROCEDURE PLOT printed page 5.

NOTE: PROCEDURE PLOT used (Total process time):

real time 0.07 seconds

cpu time 0.02 seconds

Chapter 14 : Chimpanzee sign language study
 Unreplicated multifactor study
 Post hoc adjustment with macro by Arnold Saxton



Based on this residual plot, which shows a fairly clear case of non-homogeneous variance, the textbook switches to a logarithmic transformation.

```

715     PROC mixed DATA=chimp cl covtest; class chimp sign;
716         Title2 'Unreplicated multifactor study';
717         MODEL logminutes = chimp sign / outp=resids;
718         lsmeans chimp sign / pdiff adjust=tukey;
719         ods output diffs=ppp lsmeans=mmm;
720         * ods listing exclude diffs lsmeans; *this is now just a comment;
721     run;
NOTE: The data set WORK.MMM has 14 observations and 8 variables.
NOTE: The data set WORK.PPP has 51 observations and 12 variables.
NOTE: The data set WORK.RESIDS has 40 observations and 12 variables.
NOTE: The PROCEDURE MIXED printed page 6.
NOTE: PROCEDURE MIXED used (Total process time):
      real time      0.52 seconds
      cpu time       0.48 seconds
722     TITLE3 'Post hoc adjustment with macro by Arnold Saxton';
723     %include 'C:\Geaghan\Current\EXST3201\Fall2005\SAS\pdmix800.sas';
1396     %pdmix800(ppp,mmm,alpha=.05,sort=yes);
PDMIX800 08.08.2003 processing
3.8700871961
Tukey values for CHIMP are 0.98956 (avg) 0.98956 (min) 0.98956 (max).
4.8644540923
Tukey values for SIGN are 1.96665 (avg) 1.96665 (min) 1.96665 (max).
1397     RUN;
1398     QUIT;

```

Chapter 14 : Chimpanzee sign language study
Unreplicated multifactor study

The Mixed Procedure

Model Information

Data Set	WORK.CHIMP
Dependent Variable	logMinutes
Covariance Structure	Diagonal
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Residual

Class Level Information

Class	Levels	Values
CHIMP	4	BOOEE BRUNO CINDY THELMA
SIGN	10	drink food fruit hat key listen look more shoe string

Dimensions

Covariance Parameters	1
Columns in X	15
Columns in Z	0
Subjects	1
Max Obs Per Subject	40

Number of Observations

Number of Observations Read	40
Number of Observations Used	40
Number of Observations Not Used	0

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr > Z	Alpha	Lower	Upper
Residual	0.6538	0.1779	3.67	0.0001	0.05	0.4087	1.2113

Fit Statistics

-2 Res Log Likelihood	84.5
AIC (smaller is better)	86.5
AICC (smaller is better)	86.7
BIC (smaller is better)	87.8

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
CHIMP	3	27	2.72	0.0642
SIGN	9	27	7.76	<.0001

Least Squares Means

Effect	CHIMP	SIGN	Estimate	Standard Error	DF	t Value	Pr > t
CHIMP	BOOEE		3.5163	0.2557	27	13.75	<.0001
CHIMP	BRUNO		4.1976	0.2557	27	16.42	<.0001
CHIMP	CINDY		3.6091	0.2557	27	14.11	<.0001
CHIMP	THELMA		4.3634	0.2557	27	17.06	<.0001
SIGN		drink	3.1002	0.4043	27	7.67	<.0001
SIGN		food	4.0689	0.4043	27	10.06	<.0001
SIGN		fruit	4.4567	0.4043	27	11.02	<.0001
SIGN		hat	4.9568	0.4043	27	12.26	<.0001
SIGN		key	3.2248	0.4043	27	7.98	<.0001
SIGN		listen	2.0472	0.4043	27	5.06	<.0001
SIGN		look	5.1544	0.4043	27	12.75	<.0001
SIGN		more	3.4012	0.4043	27	8.41	<.0001
SIGN		shoe	3.1549	0.4043	27	7.80	<.0001
SIGN		string	5.6509	0.4043	27	13.98	<.0001

Differences of Least Squares Means

Effect	CHIMP	SIGN	_CHIMP	_SIGN	Estimate	Standard Error	DF	t Value	Pr > t	Adjustment	Adj P
CHIMP	BOOEE		BRUNO		-0.6813	0.3616	27	-1.88	0.0704	Tukey	0.2584
CHIMP	BOOEE		CINDY		-0.09287	0.3616	27	-0.26	0.7993	Tukey	0.9939
CHIMP	BOOEE		THELMA		-0.8471	0.3616	27	-2.34	0.0268	Tukey	0.1133
CHIMP	BRUNO		CINDY		0.5885	0.3616	27	1.63	0.1153	Tukey	0.3808
CHIMP	BRUNO		THELMA		-0.1658	0.3616	27	-0.46	0.6503	Tukey	0.9674
CHIMP	CINDY		THELMA		-0.7543	0.3616	27	-2.09	0.0466	Tukey	0.1833
SIGN		drink		food	-0.9687	0.5718	27	-1.69	0.1017	Tukey	0.7892
SIGN		drink		fruit	-1.3565	0.5718	27	-2.37	0.0250	Tukey	0.3791
SIGN		drink		hat	-1.8566	0.5718	27	-3.25	0.0031	Tukey	0.0762
SIGN		drink		key	-0.1246	0.5718	27	-0.22	0.8291	Tukey	1.0000
SIGN		drink		listen	1.0530	0.5718	27	1.84	0.0765	Tukey	0.7047
SIGN		drink		look	-2.0542	0.5718	27	-3.59	0.0013	Tukey	0.0353
SIGN		drink		more	-0.3010	0.5718	27	-0.53	0.6029	Tukey	0.9999
SIGN		drink		shoe	-0.05467	0.5718	27	-0.10	0.9245	Tukey	1.0000
SIGN		drink		string	-2.5507	0.5718	27	-4.46	0.0001	Tukey	0.0043
SIGN		food		fruit	-0.3879	0.5718	27	-0.68	0.5033	Tukey	0.9994
SIGN		food		hat	-0.8880	0.5718	27	-1.55	0.1320	Tukey	0.8583
SIGN		food		key	0.8441	0.5718	27	1.48	0.1514	Tukey	0.8899
SIGN		food		listen	2.0217	0.5718	27	3.54	0.0015	Tukey	0.0402
SIGN		food		look	-1.0856	0.5718	27	-1.90	0.0683	Tukey	0.6699
SIGN		food		more	0.6677	0.5718	27	1.17	0.2531	Tukey	0.9714
SIGN		food		shoe	0.9140	0.5718	27	1.60	0.1216	Tukey	0.8375
SIGN		food		string	-1.5820	0.5718	27	-2.77	0.0101	Tukey	0.1979
SIGN		fruit		hat	-0.5001	0.5718	27	-0.87	0.3895	Tukey	0.9961
SIGN		fruit		key	1.2319	0.5718	27	2.15	0.0403	Tukey	0.5083
SIGN		fruit		listen	2.4096	0.5718	27	4.21	0.0003	Tukey	0.0079
SIGN		fruit		look	-0.6977	0.5718	27	-1.22	0.2329	Tukey	0.9624
SIGN		fruit		more	1.0555	0.5718	27	1.85	0.0759	Tukey	0.7021
SIGN		fruit		shoe	1.3019	0.5718	27	2.28	0.0309	Tukey	0.4339
SIGN		fruit		string	-1.1942	0.5718	27	-2.09	0.0463	Tukey	0.5499
SIGN		hat		key	1.7320	0.5718	27	3.03	0.0053	Tukey	0.1197
SIGN		hat		listen	2.9097	0.5718	27	5.09	<.0001	Tukey	0.0009
SIGN		hat		look	-0.1976	0.5718	27	-0.35	0.7323	Tukey	1.0000
SIGN		hat		more	1.5556	0.5718	27	2.72	0.0113	Tukey	0.2150
SIGN		hat		shoe	1.8020	0.5718	27	3.15	0.0039	Tukey	0.0932
SIGN		hat		string	-0.6941	0.5718	27	-1.21	0.2353	Tukey	0.9635
SIGN		key		listen	1.1776	0.5718	27	2.06	0.0492	Tukey	0.5682
SIGN		key		look	-1.9296	0.5718	27	-3.37	0.0022	Tukey	0.0577
SIGN		key		more	-0.1764	0.5718	27	-0.31	0.7601	Tukey	1.0000
SIGN		key		shoe	0.06993	0.5718	27	0.12	0.9036	Tukey	1.0000
SIGN		key		string	-2.4261	0.5718	27	-4.24	0.0002	Tukey	0.0074
SIGN		listen		look	-3.1073	0.5718	27	-5.43	<.0001	Tukey	0.0003
SIGN		listen		more	-1.3540	0.5718	27	-2.37	0.0253	Tukey	0.3815
SIGN		listen		shoe	-1.1077	0.5718	27	-1.94	0.0632	Tukey	0.6457
SIGN		listen		string	-3.6037	0.5718	27	-6.30	<.0001	Tukey	<.0001
SIGN		look		more	1.7532	0.5718	27	3.07	0.0049	Tukey	0.1111
SIGN		look		shoe	1.9996	0.5718	27	3.50	0.0016	Tukey	0.0439
SIGN		look		string	-0.4965	0.5718	27	-0.87	0.3929	Tukey	0.9963
SIGN		more		shoe	0.2463	0.5718	27	0.43	0.6700	Tukey	1.0000
SIGN		more		string	-2.2497	0.5718	27	-3.93	0.0005	Tukey	0.0158
SIGN		shoe		string	-2.4960	0.5718	27	-4.37	0.0002	Tukey	0.0054

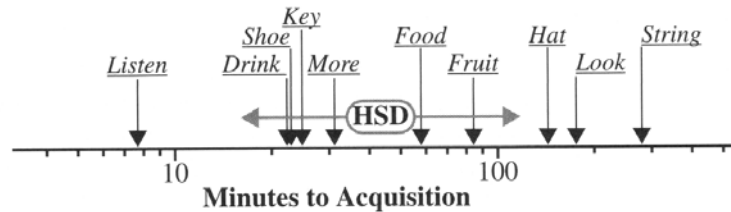
Chapter 14 : Chimpanzee sign language study
 Unreplicated multifactor study
 Post hoc adjustment with macro by Arnold Saxton

Effect=CHIMP ADJUSTMENT=Tukey(P<.01) bygroup=1

Obs	CHIMP	Estimate	StdErr	MSGROUP	Antilog(Estimate)
1	THELMA	4.3634	0.2557	A	78.524
2	BRUNO	4.1976	0.2557	A	66.526
3	CINDY	3.6091	0.2557	A	36.933
4	BOOEE	3.5163	0.2557	A	33.660

Effect=SIGN ADJUSTMENT=Tukey(P<.01) bygroup=2

Obs	SIGN	Estimate	StdErr	MSGROUP	Antilog(Estimate)
5	string	5.6509	0.4043	A	284.547
6	look	5.1544	0.4043	AB	173.192
7	hat	4.9568	0.4043	ABC	142.138
8	fruit	4.4567	0.4043	ABC	86.203
9	food	4.0689	0.4043	ABC	58.493
10	more	3.4012	0.4043	BCD	30.000
11	key	3.2248	0.4043	BCD	25.149
12	shoe	3.1549	0.4043	CD	23.451
13	drink	3.1002	0.4043	CD	22.202
14	listen	2.0472	0.4043	D	7.746

Display 14.2 Multiple comparisons of sign means on the log scale

The graphic in Display 14.2 is only precisely possible for balanced experiments. When unbalanced the variance between treatment levels is given by $s_{\bar{y}_1 - \bar{y}_2}^2 = MSE \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$, where each pair of treatment levels potentially has a different value of $s_{\bar{y}_1 - \bar{y}_2}^2$ and therefore a different value of the “HSD” in the graphic above. The HSD (as opposed to the LSD) is the Tukey interval, sometimes referred to as the “honest significant difference”.

Note the log scale on the graphic above (I used natural logs). The estimates and confidence intervals should be done on the transformed data. To recreate the interval drawn on the graphic we note that the HSD centers a little over 40, say 41. We then note that Saxton’s macro reports the following:

Tukey values for SIGN are 1.96665 (avg) 1.96665 (min) 1.96665 (max).

Using half of this interval we can calculate $\exp(41 \pm 1.96665/2)$ which gives lower and upper bounds of 15.337 and 109.607, a fair approximation of the table above.

```
1400      proc univariate data=resids plot normal; var resid; run;
NOTE: The PROCEDURE UNIVARIATE printed page 8.
NOTE: PROCEDURE UNIVARIATE used (Total process time):
      real time          0.08 seconds
      cpu time           0.04 seconds
```

Chapter 14 : Chimpanzee sign language study
Unreplicated multifactor study
Post hoc adjustment with macro by Arnold Saxton

The UNIVARIATE Procedure
Variable: Resid

Moments			
N	40	Sum Weights	40
Mean	0	Sum Observations	0
Std Deviation	0.67277851	Variance	0.45263093
Skewness	-0.4126989	Kurtosis	1.10905879
Uncorrected SS	17.6526062	Corrected SS	17.6526062
Coeff Variation	.	Std Error Mean	0.10637562

Basic Statistical Measures			
Location		Variability	
Mean	0.000000	Std Deviation	0.67278
Median	0.035119	Variance	0.45263
Mode	.	Range	3.44471
		Interquartile Range	0.80535

Tests for Location: Mu0=0			
Test	-Statistic-	-----p Value-----	
Student's t	t	0	Pr > t 1.0000
Sign	M	2	Pr >= M 0.6358
Signed Rank	S	29	Pr >= S 0.7019

Tests for Normality

Test	--Statistic--	-----p Value-----
Shapiro-Wilk	W 0.963814	Pr < W 0.2257
Kolmogorov-Smirnov	D 0.094756	Pr > D >0.1500
Cramer-von Mises	W-Sq 0.074443	Pr > W-Sq 0.2419
Anderson-Darling	A-Sq 0.509372	Pr > A-Sq 0.1955

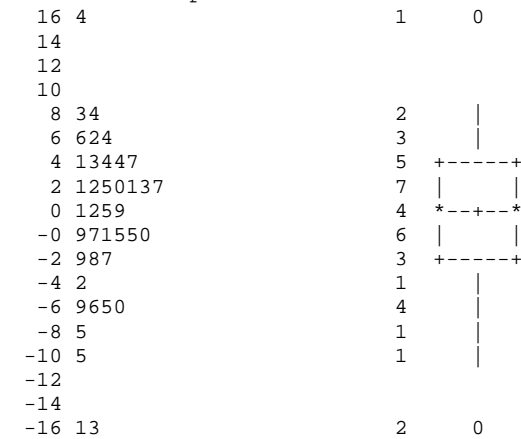
Quantiles (Definition 5)

Quantile	Estimate
100% Max	1.7389032
99%	1.7389032
95%	0.8350244
90%	0.7274318
75% Q3	0.4221268
50% Median	0.0351189
25% Q1	-0.3832191
10%	-0.7731295
5%	-1.3892767
1%	-1.7058028
0% Min	-1.7058028

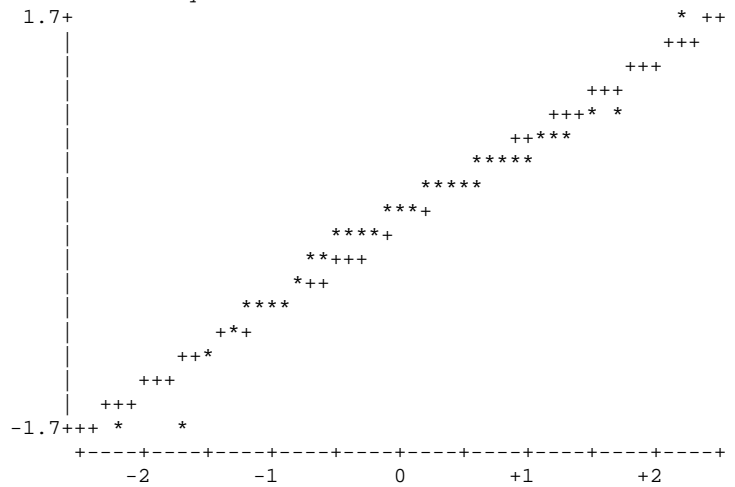
Extreme Observations

-----Lowest-----		-----Highest-----	
Value	Obs	Value	Obs
-1.705803	26	0.718498	6
-1.630025	21	0.736365	36
-1.148528	17	0.826983	20
-0.852978	19	0.843066	1
-0.693280	5	1.738903	25

Stem Leaf Boxplot



Normal Probability Plot



Multiply Stem.Leaf by 10**-1

```

1401 options ps=52 ls=111;
1402 proc plot data=resids; plot resid*pred; run;
1403 options ps=512 ls=132;
1404

```

NOTE: There were 40 observations read from the data set WORK.RESIDS.

NOTE: The PROCEDURE PLOT printed page 9.

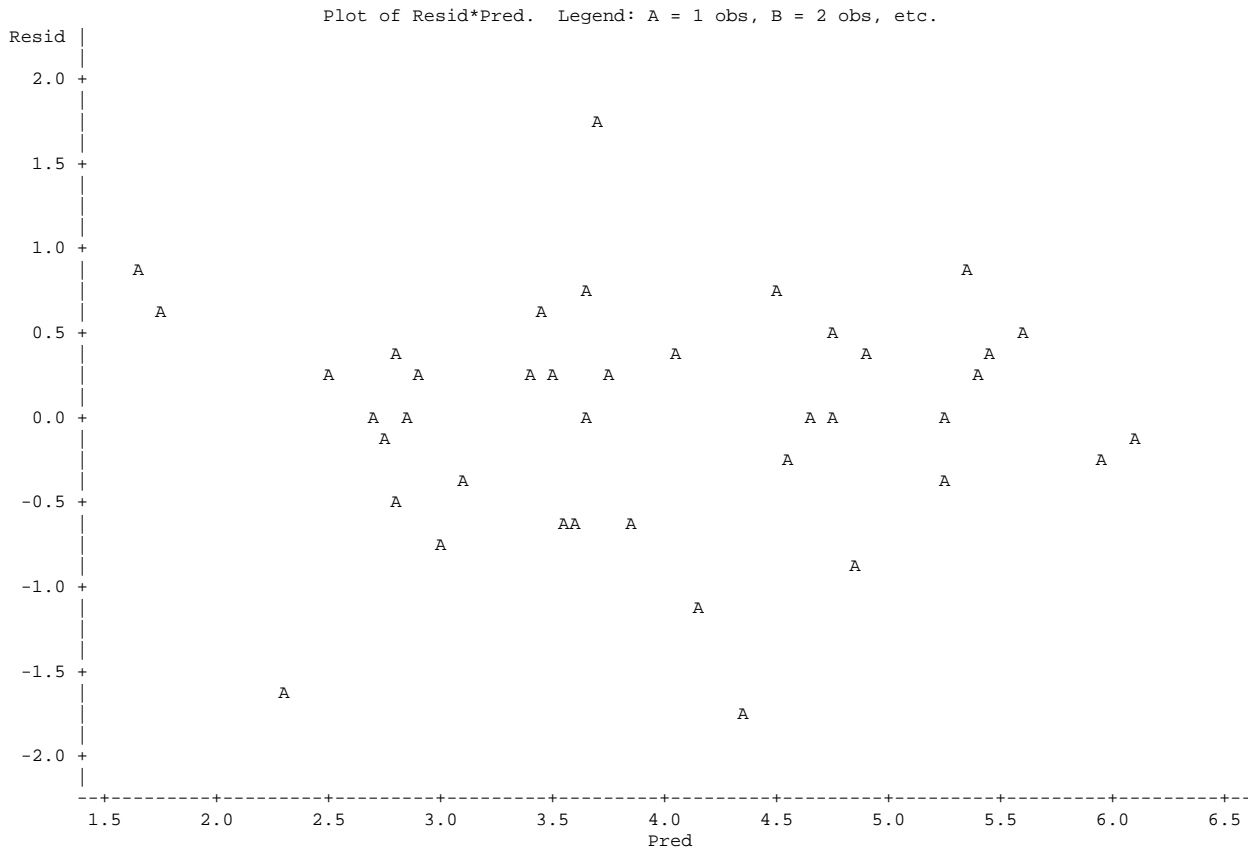
NOTE: PROCEDURE PLOT used (Total process time):

```

real time      0.06 seconds
cpu time       0.01 seconds

```

Chapter 14 : Chimpanzee sign language study
 Unreplicated multifactor study
 Post hoc adjustment with macro by Arnold Saxton



The residual plot with logarithms appears to be much improved, and variance now appears to be homogeneous.

```

1405      proc print data=mmm;
1406          TITLE3 'LSMeans output dataset of estimates';
1407      run;
NOTE: There were 14 observations read from the data set WORK.MMM.
NOTE: The PROCEDURE PRINT printed page 10.
NOTE: PROCEDURE PRINT used (Total process time):
      real time          0.06 seconds
      cpu time           0.02 seconds
    
```

Chapter 14 : Chimpanzee sign language study
 Unreplicated multifactor study
 LSMeans output dataset of Estimates

Obs	Effect	CHIMP	SIGN	Estimate	StdErr	DF	tValue	Probt
1	CHIMP	BOOEE		3.5163	0.2557	27	13.75	<.0001
2	CHIMP	BRUNO		4.1976	0.2557	27	16.42	<.0001
3	CHIMP	CINDY		3.6091	0.2557	27	14.11	<.0001
4	CHIMP	THELMA		4.3634	0.2557	27	17.06	<.0001
5	SIGN		drink	3.1002	0.4043	27	7.67	<.0001
6	SIGN		food	4.0689	0.4043	27	10.06	<.0001
7	SIGN		fruit	4.4567	0.4043	27	11.02	<.0001
8	SIGN		hat	4.9568	0.4043	27	12.26	<.0001
9	SIGN		key	3.2248	0.4043	27	7.98	<.0001
10	SIGN		listen	2.0472	0.4043	27	5.06	<.0001
11	SIGN		look	5.1544	0.4043	27	12.75	<.0001
12	SIGN		more	3.4012	0.4043	27	8.41	<.0001
13	SIGN		shoe	3.1549	0.4043	27	7.80	<.0001
14	SIGN		string	5.6509	0.4043	27	13.98	<.0001

```

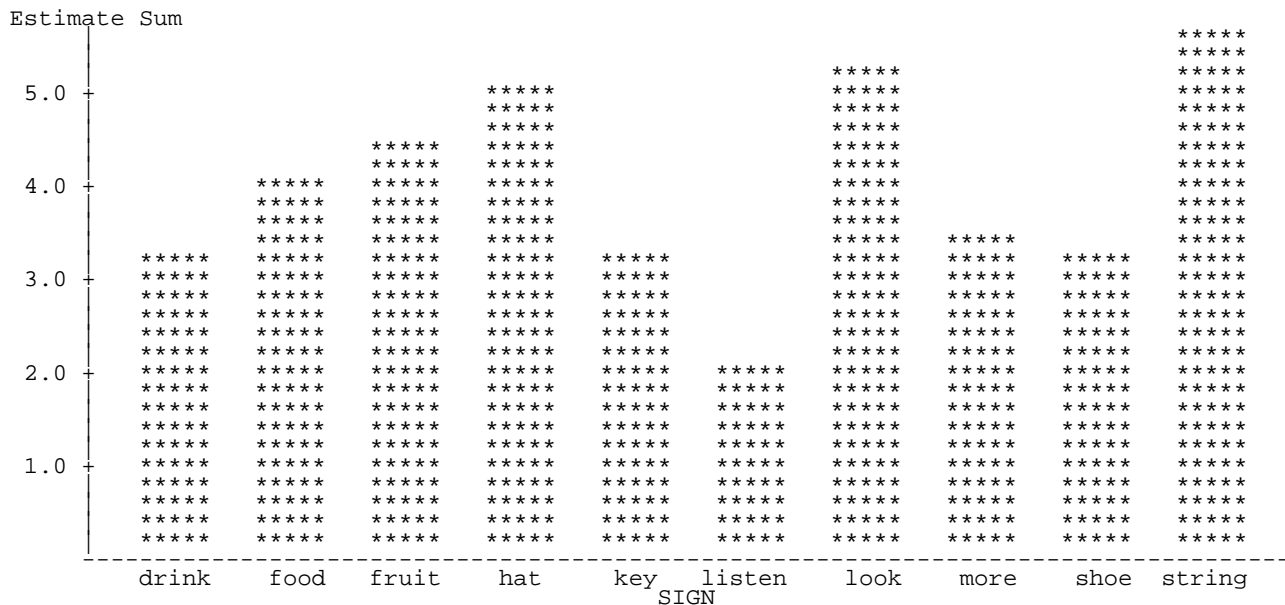
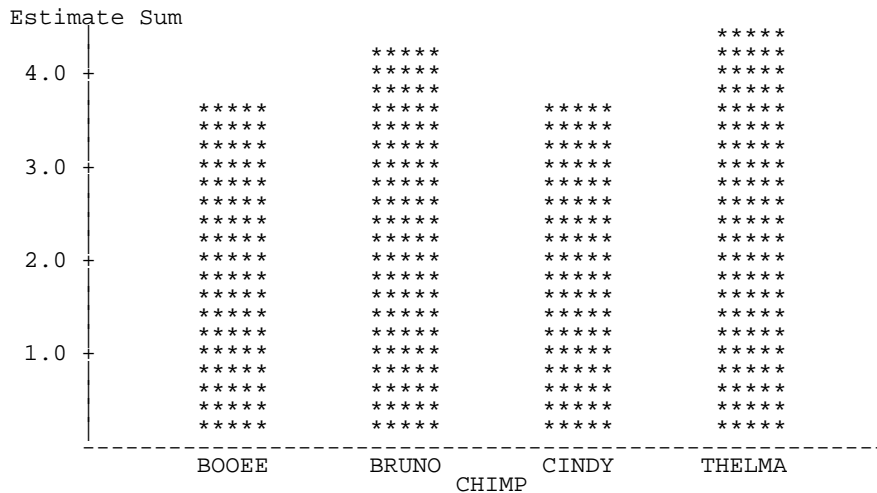
1409      data chimps; set mmm; format estimate f5.1; if effect='CHIMP'; run;
NOTE: There were 14 observations read from the data set WORK.MMM.
NOTE: The data set WORK.CHIMPS has 4 observations and 8 variables.
NOTE: DATA statement used (Total process time):
      real time           0.02 seconds
      cpu time            0.03 seconds
1410      data sign; set mmm; format estimate f5.1; if effect='SIGN'; run;
NOTE: There were 14 observations read from the data set WORK.MMM.
NOTE: The data set WORK.SIGN has 10 observations and 8 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.02 seconds

1411
1412      options ps=52 ls=99 nolabel;
1413      TITLE3 'Graphics';
1414      proc chart data=chimps; vbar chimp / sumvar=estimate; run;
NOTE: The PROCEDURE CHART printed page 11.
NOTE: PROCEDURE CHART used (Total process time):
      real time           0.05 seconds
      cpu time            0.02 seconds
1415      proc chart data=sign; vbar sign / sumvar=estimate; run;
NOTE: The PROCEDURE CHART printed page 12.
NOTE: PROCEDURE CHART used (Total process time):
      real time           0.07 seconds
      cpu time            0.02 seconds

1416

```

Chapter 14 : Chimpanzee sign language study
Unreplicated multifactor study
Graphics



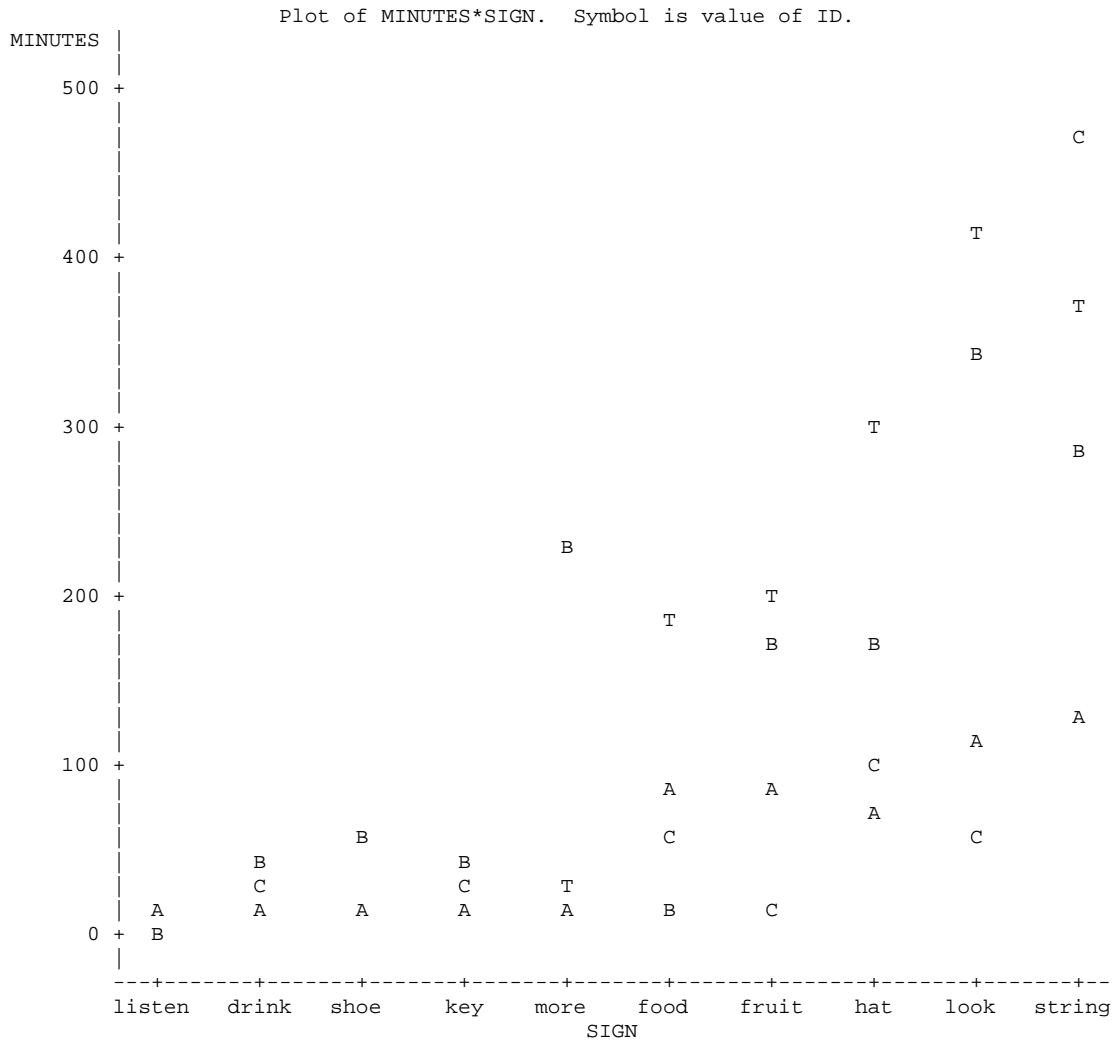
SIGN

```

1417      proc plot data=chimp; plot minutes * sign = ID /
1418          haxis = 'listen' 'drink' 'shoe' 'key' 'more' 'food' 'fruit' 'hat' 'look' 'string';
1419      run;
NOTE: There were 40 observations read from the data set WORK.CHIMP.
NOTE: The PROCEDURE PLOT printed page 13.
NOTE: PROCEDURE PLOT used (Total process time):
      real time      0.06 seconds
      cpu time       0.01 seconds

```

Chapter 14 : Chimpanzee sign language study
 Unreplicated multifactor study
 Graphics



NOTE: 7 obs hidden.

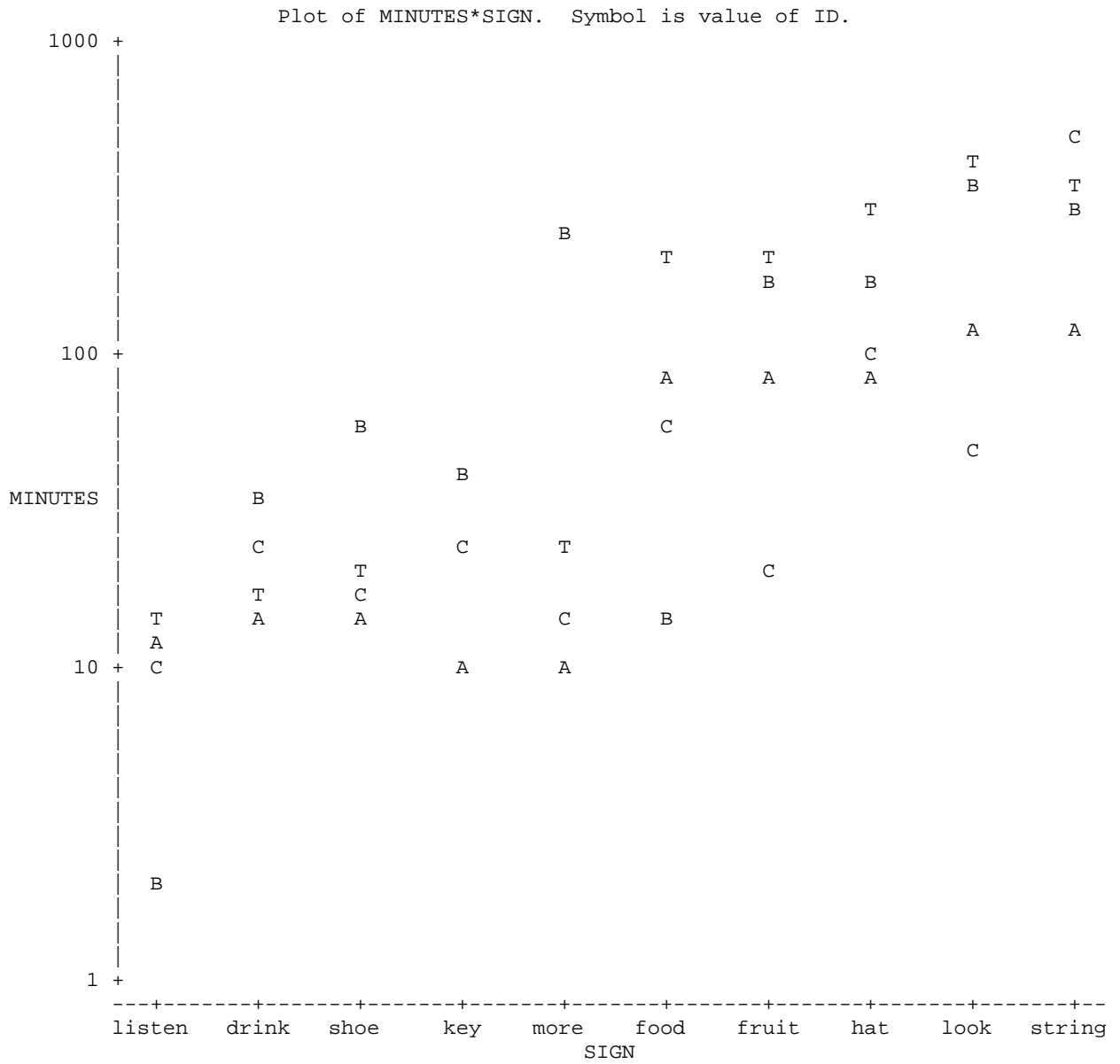
```

1420      proc plot data=chimp; plot minutes * sign = ID / vaxis=1 10 100 1000
1421          haxis = 'listen' 'drink' 'shoe' 'key' 'more' 'food' 'fruit' 'hat' 'look' 'string';
1422      run;
1423      ; run;
1424
NOTE: There were 40 observations read from the data set WORK.CHIMP.
NOTE: The PROCEDURE PLOT printed page 14.
NOTE: PROCEDURE PLOT used (Total process time):
      real time      0.08 seconds
      cpu time       0.01 seconds

```

Note the request for VAXIS that creates a log scale and the HAXIS that orders the signs in the desired order, lowest to highest in this case.

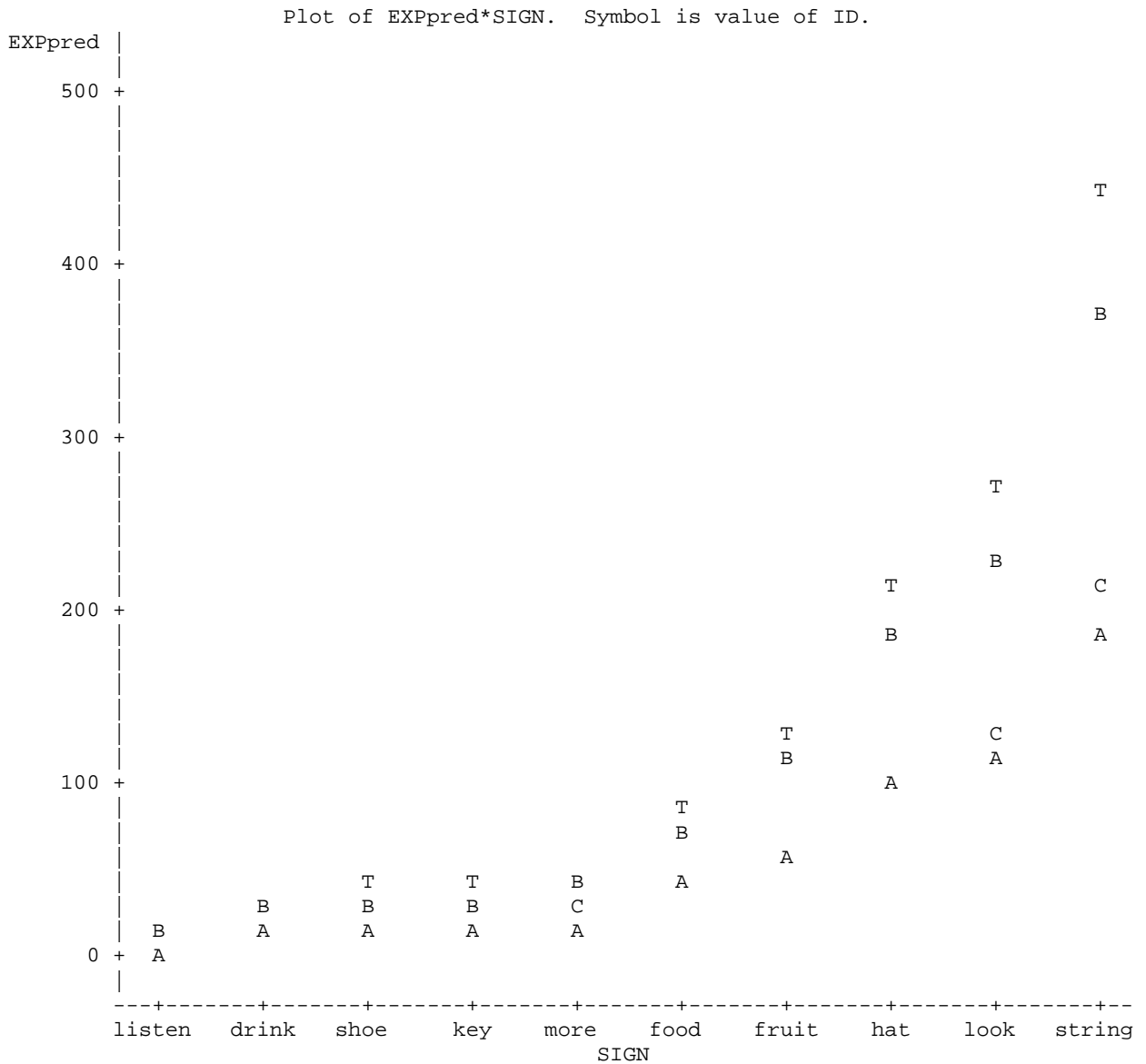
Chapter 14 : Chimpanzee sign language study
 Unreplicated multifactor study
 Graphics



```

1425      data resids; set resids; EXPpred = exp(pred); format estimate f5.1; run;
NOTE: Variable estimate is uninitialized.
NOTE: There were 40 observations read from the data set WORK.RESIDS.
NOTE: The data set WORK.RESIDS has 40 observations and 14 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.02 seconds
1426      proc plot data=resids; plot EXPpred * sign = ID /
1427          haxis = 'listen' 'drink' 'shoe' 'key' 'more' 'food' 'fruit' 'hat' 'look' 'string';
1428      run;
1429
1430      run;
NOTE: There were 40 observations read from the data set WORK.RESIDS.
NOTE: The PROCEDURE PLOT printed page 15.
NOTE: PROCEDURE PLOT used (Total process time):
      real time           0.06 seconds
      cpu time            0.01 seconds
    
```

Chapter 14 : Chimpanzee sign language study
 Unreplicated multifactor study
 Graphics



NOTE: 10 obs hidden.