Ordinal Examples

The examples shown below come from a study conducted by Karen Seccombe that examined health coverage among low income families.[1] I illustrate ordinal analyses with two variables—the extent to which the participant reported cutting meals (`cutmeals`) and a three-category education measure (`ed`).

SPSS
A simple way to get a linear-by-linear association model is through the crosstabs procedure which we used for nominal contingency chi-square tests. The linear-by-linear (i.e., ordinal loglinear) test is found in the default output. The linear-by-linear association reported is the same as the loglinear model test.

```
crosstabs /tables=ed by cutmeal
  /cells=count row column
 /statistics=chisq phi.
```

ed * cutmeal how often cut meal size Crosstabulation

			cutmeal how often cut meal size			
			0 never or rarely	1 some months but not every month	2 almost every month	Total
ed	1.00 less than hs	Count	85	16	13	114
		% within ed	74.6%	14.0%	11.4%	100.0%
		% within cutmeal how often cut meal size	19.8%	28.6%	19.7%	20.7%
	2.00 hs	Count	183	17	27	227
		% within ed	80.6%	7.5%	11.9%	100.0%
		% within cutmeal how often cut meal size	42.7%	30.4%	40.9%	41.2%
	3.00 college	Count	161	23	26	210
		% within ed	76.7%	11.0%	12.4%	100.0%
		% within cutmeal how often cut meal size	37.5%	41.1%	39.4%	38.1%
Total		Count	429	56	66	551
		% within ed	77.9%	10.2%	12.0%	100.0%
		% within cutmeal how often cut meal size	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	3.888 [a]	4	.421
Likelihood Ratio	3.867	4	.424
Linear-by-Linear Association	.000	1	.984
N of Valid Cases	551		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 11.59.

Or, the loglinear procedure `genlog` can be used if you compute an association variable and add that to the model. The `genlog` approach is easily expanded to include additional variables and effects. Adding keyword `freq` to the `/print` subcommand prints the frequencies (omitted below).

```
compute assoc = cutmeal*ed.
genlog ed cutmeal with assoc
/print=est freq
/plot=none
/design = ed cutmeal assoc.
```

[1] Seccombe, K., Newsom, J.T., & Hoffman, K. (2006). Access to healthcare after welfare reform. *Inquiry, 43*, 167-179.

Although the values are similar, the likelihood ratio and Pearson chi-square tests do not provide a test of the association term by itself. A comparison of the fit to a model without the association term, which I will not illustrate here, would be needed for that. (If this was a binary loglinear model, the fit would be perfect—a saturated model).

Goodness-of-Fit Tests [a,b]

	Value	df	Sig.
Likelihood Ratio	3.866	3	.276
Pearson Chi-Square	3.890	3	.274

a. Model: Poisson

b. Design: Constant + ed + cutmeal + assoc

The test of the association term is found in the last row (highlighted) of the parameter estimates (*b* values) table, however. Notice that this sig. value (p-value) matches the linear-by-linear sig. value in the crosstabs output above.

Parameter Estimates [b,c]

					95% Confidence Interval	
Parameter	Estimate	Std. Error	Z	Sig.	Lower Bound	Upper Bound
Constant	3.212	.653	4.918	<.001	1.932	4.492
[ed = 1.00]	-.606	.253	-2.397	.017	-1.102	-.111
[ed = 2.00]	.080	.148	.543	.587	-.209	.369
[ed = 3.00]	0[a]	.	.	.	.	.
[cutmeal = 1]	1.879	.388	4.849	<.001	1.120	2.639
[cutmeal = 2]	-.161	.257	-.624	.533	-.665	.344
[cutmeal = 3]	0[a]	.	.	.	.	.
assoc	.002	.084	.021	.984	-.162	.166

a. This parameter is set to zero because it is redundant.

b. Model: Poisson

c. Design: Constant + ed + cutmeal + assoc

Retest the ordinal model without the association variable to compare to the saturated model. The difference in likelihood ratio tests is a likelihood ratio chi-square that can be tested to significance to determine of there is a relationship between education and cutting back on meals.

```
genlog ed cutmeal with assoc
/model=poisson
/print=est freq
/plot=none
/design = ed cutmeal.
```

Goodness-of-Fit Tests [a,b]

	Value	df	Sig.
Likelihood Ratio	3.867	4	.424
Pearson Chi-Square	3.888	4	.421

a. Model: Poisson

b. Design: Constant + ed + cutmeal

Likelihood ratio from saturated model is equal to 3.866. Likelihood ratio from model without `assoc` is 3.867. Likelihood ratio difference is 3.867 – 3.866 = .001. This can be compared to the critical chi-square value with df = 1, which is 3.84, so the difference is not significant, indicating no significant relationship between the two variables.

The Kendall's tau-b, Gamma, Somer's d provide other ways to test association among ordinal variables but are not equivalent to the loglinear test.

```
* get Kendall's tau-b (btau), Gamma, Somer's d (d).

crosstabs /tables=ed by cutmeal
  /cells=count row column
 /statistics=btau gamma d.
```

Directional Measures

			Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Ordinal by Ordinal	Somers' d	Symmetric	-.001	.039	-.013	.989
		ed Dependent	-.001	.053	-.013	.989
		cutmeal how often cut meal size Dependent	.000	.030	-.013	.989

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Symmetric Measures

		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Ordinal by Ordinal	Kendall's tau-b	-.001	.040	-.013	.989
	Gamma	-.001	.081	-.013	.989
N of Valid Cases		551			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

R
Note: The glm function in R does not quite match SPSS or SAS, so may not be correct.

```
#clear console
> cat("\014")
> rm(mymodel)
> rm(mytable)
>
#Loglinear models
> #two-way loglinear
> library(MASS)
> #make summary frequency table for two variables
> mytable = table(d$ed, d$cutmeal)
> tbl = table(d$ed, d$cutmeal)
>
> #View(tbl)
>
> #create data frame from the matrix tbl (need both below)
> mytable = as.data.frame(tbl)
>
> #conversion leads to new column names Var1 and Var2, so rename them
> colnames(mytable)[colnames(mytable)=="Var1"] <- "ed"
> colnames(mytable)[colnames(mytable)=="Var2"] <- "cutmeal"
>
> #linear-by-linear association model needs numeric variables for the association term
> mytable$ed = as.numeric(mytable$ed)
> mytable$cutmeal = as.numeric(mytable$cutmeal)
> mytable$assoc = mytable$ed*mytable$cutmeal
>
> #convert variables back to factors so that comparisons are used by glm
> mytable$ed = as.factor(mytable$ed)
> mytable$cutmeal = as.factor(mytable$cutmeal)
>
> #use the codingMatrices package to get an exact match with SPSS
> #contr.SAS options needed to get the correct contrasts and Type III SS
> library(codingMatrices)
> mymodel = glm(Freq ~ ed + cutmeal + assoc, family=poisson(link="log"),data=mytable,
>     contrasts=list(ed=contr.SAS,cutmeal=contr.SAS))
> coef(summary(mymodel))
               Estimate Std. Error     z value        Pr(>|z|)
(Intercept)  3.211919805 0.65315131  4.91757388 0.0000008762339
ed1         -0.606302396 0.25291124 -2.39729318 0.0165167021564
ed2          0.080146515 0.14761352  0.54294835 0.5871653660056
cutmeal1     1.879273063 0.38758486  4.84867515 0.0000012428876
cutmeal2    -0.160566786 0.25733781 -0.62395334 0.5326582299075
```

```
assoc          0.001717556 0.08373613   0.02051153 0.9836353158063
```

```
> #DescTools can be used for association statistics that gives confidence intervals
> library("DescTools")
> Assocs(tbl,conf.level = .95)
                        estimate  lwr.ci  upr.ci
Contingency Coeff.        0.0837      -       -
Cramer V                  0.0594  0.0000  0.1005
Kendall Tau-b            -0.0005 -0.0789  0.0779
Goodman Kruskal Gamma    -0.0011 -0.1604  0.1583
Stuart Tau-c             -0.0004 -0.0577  0.0569
Somers D C|R             -0.0004 -0.0598  0.0590
Somers D R|C             -0.0007 -0.1041  0.1027
Pearson Correlation       0.0009 -0.0827  0.0844
Spearman Correlation     -0.0005 -0.0840  0.0830
Lambda C|R                0.0000  0.0000  0.0000
Lambda R|C                0.0185  0.0000  0.0564
Lambda sym                0.0135  0.0000  0.0410
Uncertainty Coeff. C|R    0.0051 -0.0051  0.0154
Uncertainty Coeff. R|C    0.0033 -0.0033  0.0099
Uncertainty Coeff. sym    0.0040 -0.0040  0.0120
Mutual Information        0.0051      -       -

> #vcd assocstats gives chi-square and likelihood ratio significance tests
> library(vcd)
> assocstats(tbl)
                 X^2 df P(> X^2)
Likelihood Ratio 3.8668  4  0.42434
Pearson          3.8880  4  0.42137

Phi-Coefficient    : NA
Contingency Coeff.: 0.084
Cramer's V         : 0.059
```