Models In Epidemiology And Biostatistics
Gordon Hilton Fick

Apgar Scores Example

Dohoo, Martin & Stryhn (2012) provide a detailed example with an ordinal outcome. The outcome is the Apgar score. See: https://en.wikipedia.org/wiki/Apgar_score
The Apgar score was designed to summarize the health of newborn children.
Apgar scores range from zero to 10. The five criteria are summarized using words chosen to form a backronym (Appearance, Pulse, Grimace, Activity, Respiration).
The five criteria of the Apgar score:

	Score of 0	**Score of 1**	**Score of 2**	**Component of backronym**
Skin color	blue or pale all over	blue at extremities, body pink (acrocyanosis)	no cyanosis body and extremities pink	**A**ppearance
Pulse rate	absent	<100 beats per minute	>100 beats per minute	**P**ulse
Reflex irritability grimace	no response to stimulation	grimace on suction or aggressive stimulation	cry on stimulation	**G**rimace
Activity	none	some flexion	flexed arms and legs that resist extension	**A**ctivity
Respiratory effort	absent	weak, irregular, gasping	strong, robust cry	**R**espiration

DMS obtained a random sample of 5000 births from http://www.cdc.gov/nchs

This dataset is bw5k.dta

```
. tab apgar

Apgar score |
  at 5 min. |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         10        0.20        0.20
          2 |          3        0.06        0.26
          3 |          5        0.10        0.36
          4 |          6        0.12        0.48
          5 |         19        0.38        0.86
          6 |         29        0.58        1.44
          7 |         99        1.98        3.42
          8 |        538       10.76       14.18
          9 |      4,103       82.06       96.24
         10 |        188        3.76      100.00
------------+-----------------------------------
      Total |      5,000      100.00
```

```
. tab previs

       # of |
   prenatal |
     visits |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |         47        0.94        0.94
          1 |         14        0.28        1.22
          2 |         42        0.84        2.06
          3 |         35        0.70        2.76
          4 |         68        1.36        4.12
          5 |         95        1.90        6.02
          6 |        126        2.52        8.54
          7 |        177        3.54       12.08
          8 |        284        5.68       17.76
          9 |        360        7.20       24.96
         10 |        734       14.68       39.64
         11 |        468        9.36       49.00
         12 |        946       18.92       67.92
         13 |        398        7.96       75.88
         14 |        381        7.62       83.50
         15 |        411        8.22       91.72
         16 |        160        3.20       94.92
         17 |         60        1.20       96.12
         18 |         49        0.98       97.10
         19 |         20        0.40       97.50
         20 |         60        1.20       98.70
         21 |         11        0.22       98.92
         22 |          9        0.18       99.10
         23 |          6        0.12       99.22
         24 |          9        0.18       99.40
         25 |         11        0.22       99.62
         26 |          3        0.06       99.68
         27 |          2        0.04       99.72
         29 |          3        0.06       99.78
         30 |          8        0.16       99.94
         31 |          1        0.02       99.96
         34 |          1        0.02       99.98
         40 |          1        0.02      100.00
------------+-----------------------------------
      Total |      5,000      100.00
```

DMS grouped apgar and previs so we will do the same.

```
. recode apgar (1/6=0) (7/8=1) (9/1=2), gen(apgar_c3)

. recode previs (1/5=0)(6/11=1)(12/1000=2), gen(previs_c3)

. gen white = (mrace_c4==2)
```

```
. mlogit apgar_c3 i.previs_c3 white gest male

Multinomial logistic regression                 Number of obs     =      5,000
                                                LR chi2(10)       =     108.02
                                                Prob > chi2       =     0.0000
Log likelihood = -2219.9695                     Pseudo R2         =     0.0238

------------------------------------------------------------------------------
    apgar_c3 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
0            |
   previs_c3 |
          1  |  -1.024911    .357607    -2.87   0.004    -1.725808   -.3240141
          2  |  -.9443598   .3614109    -2.61   0.009    -1.652712   -.2360076
             |
       white |   .4289323   .2530061     1.70   0.090    -.0669506    .9248152
        gest |  -.2156859   .0317664    -6.79   0.000     -.277947   -.1534249
        male |   .7148482   .2604891     2.74   0.006      .204299    1.225397
       _cons |   4.270922   1.142261     3.74   0.000     2.032132    6.509712
-------------+----------------------------------------------------------------
1            |
   previs_c3 |
          1  |  -.2946795   .1697455    -1.74   0.083    -.6273745    .0380154
          2  |  -.3356566   .1698413    -1.98   0.048    -.6685394   -.0027737
             |
       white |   .2794856   .0883631     3.16   0.002      .106297    .4526742
        gest |  -.1027939    .016823    -6.11   0.000    -.1357664   -.0698215
        male |   .0149094   .0855448     0.17   0.862    -.1527552    .1825741
       _cons |   2.166233   .6416813     3.38   0.001     .9085608    3.423905
-------------+----------------------------------------------------------------
2            |  (base outcome)
------------------------------------------------------------------------------

. ologit apgar_c3 i.previs_c3 white gest male

Ordered logistic regression                     Number of obs     =      5,000
                                                LR chi2(5)        =      90.90
                                                Prob > chi2       =     0.0000
Log likelihood = -2228.5279                     Pseudo R2         =     0.0200

------------------------------------------------------------------------------
    apgar_c3 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   previs_c3 |
          1  |   .4406487   .1565974     2.81   0.005     .1337234     .747574
          2  |   .4639134   .1566039     2.96   0.003     .1569753    .7708514
             |
       white |   -.296896   .0846716    -3.51   0.000    -.4628492   -.1309428
        gest |   .1284589   .0158105     8.12   0.000     .0974709    .1594469
        male |  -.0906923   .0820493    -1.11   0.269     -.251506    .0701215
-------------+----------------------------------------------------------------
       /cut1 |   .8684825   .6053449                     -.3179718    2.054937
       /cut2 |   3.328196   .6011476                      2.149969    4.506424
------------------------------------------------------------------------------
```

Interpret the coefficients from these 2 models.

```
. brant,detail

Estimated coefficients from binary logits

-------------------------------------
    Variable |  y_gt_0      y_gt_1
-------------+-----------------------
   previs_c3 |
          1  |    0.964       0.400
             |     2.70        2.54
          2  |    0.879       0.429
             |     2.44        2.72
             |
       white |   -0.382      -0.294
             |    -1.51       -3.47
        gest |    0.192       0.119
             |     6.17        7.59
        male |   -0.713      -0.079
             |    -2.74       -0.96
       _cons |   -3.190      -2.936
             |    -2.86       -4.93
-------------------------------------
                         legend: b/t

Brant test of parallel regression assumption

              |       chi2     p>chi2      df
 -------------+------------------------------
          All |      19.49      0.002       5
 -------------+------------------------------
  1.previs_c3 |       2.81      0.094       1
  2.previs_c3 |       1.75      0.186       1
        white |       0.13      0.717       1
         gest |       6.15      0.013       1
         male |       6.37      0.012       1

A significant test statistic provides evidence that the parallel
regression assumption has been violated.
```

There is evidence against the proportional odds assumption. In particular, gest and male have cut specific regression coefficients that are different.

There is a model that assumes proportional odds for some variables but does not make the assumption for others. These models are called 'partial proportional odds models'. For example :

$$\log\left(\frac{p_j}{1-p_j}\right)=\beta_{0j}+\beta_1*\text{previs}_1+\beta_2*\text{previs}_2+\beta_3*\text{white}+\beta_{4j}*\text{gest}+\beta_{5j}*\text{male}$$

These models use β_{0j} rather than $-\kappa_j$.

There are 'generalized ordinal logit models' as well. Such models can provide cut specific estimates for all cuts simultaneously. These models do not give the same estimates as the individual marginal models and these models are not the same as the multinomial models.

```
. ssc install gologit2
. gologit2 apgar_c3 i.previs_c3 white gest male,pl(white i.previs_c3)
```

Generalized Ordered Logit Estimates

	Number of obs	=	5,000
	LR chi2(7)	=	104.24
	Prob > chi2	=	0.0000
Log likelihood = -2221.8607	Pseudo R2	=	0.0229

(1) [0]white - [1]white = 0
(2) [0]1.previs_c3 - [1]1.previs_c3 = 0
(3) [0]2.previs_c3 - [1]2.previs_c3 = 0

apgar_c3	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
0						
previs_c3						
1	.4292983	.1563866	2.75	0.006	.1227861	.7358105
2	.4534526	.1565242	2.90	0.004	.1466709	.7602343
white	-.2956676	.0845814	-3.50	0.000	-.461444	-.1298911
gest	.1875419	.0274186	6.84	0.000	.1338023	.2412814
male	-.723602	.2579657	-2.81	0.005	-1.229205	-.2179985
_cons	-2.657577	1.026297	-2.59	0.010	-4.669082	-.6460712
1						
previs_c3						
1	.4292983	.1563866	2.75	0.006	.1227861	.7358105
2	.4534526	.1565242	2.90	0.004	.1466709	.7602343
white	-.2956676	.0845814	-3.50	0.000	-.461444	-.1298911
gest	.1211466	.015794	7.67	0.000	.090191	.1521022
male	-.0776052	.0821413	-0.94	0.345	-.2385992	.0833888
_cons	-3.043491	.6007176	-5.07	0.000	-4.220876	-1.866106