Assignment 9 Solutions: Other Outcomes

Part 1: Ordinal, Multinomial, and Count Outcomes

Applied Quantitative Methods II, UC3M

Spring 2026

Contents


```
library(carData)
library(MASS)
library(nnet)
library(pscl)
library(AER)
library(marginaleffects)
library(ggplot2)
data(BEPS)
```

1. Ordered logit: perceptions of the national economy

a) Explore the outcome and convert to an ordered factor:

```
table(BEPS$economic.cond.national)
```

```
##
##   1   2   3   4   5
##  37 257 607 542  82
```

```
BEPS$econ_ord = factor(BEPS$economic.cond.national, ordered = TRUE)
```

The distribution is concentrated in the middle categories (2, 3, and 4), with category 3 ("stayed about the same") being the modal response. Very few respondents chose the extreme ends (1 = got much worse, 5 = got much better). OLS treats the numeric values 1–5 as equally spaced, implying that the difference between "got much worse" and "got a little worse" is identical to the difference between "stayed the same" and "got a little better." This is almost certainly wrong for Likert-type survey items, where psychological distances between adjacent categories need not be equal. Ordered logit avoids this assumption by estimating threshold parameters that let the data determine the spacing of the latent scale.

b) Fit the ordered logit model:

```
m_ologit = polr(econ_ord ~ age + gender + Europe + political.knowledge,
                data = BEPS, Hess = TRUE)
summary(m_ologit)
```

```
## Call:
## polr(formula = econ_ord ~ age + gender + Europe + political.knowledge,
##     data = BEPS, Hess = TRUE)
##
## Coefficients:
##                         Value Std. Error t value
## age                  0.003999   0.003044   1.314
## gendermale           0.192062   0.096475   1.991
## Europe              -0.122693   0.014978  -8.192
## political.knowledge -0.120719   0.045154  -2.674
##
## Intercepts:
##     Value    Std. Error t value
## 1|2  -4.4796   0.2727   -16.4247
## 2|3  -2.1843   0.2225    -9.8153
## 3|4  -0.3183   0.2141    -1.4867
## 4|5   2.2397   0.2348     9.5391
##
## Residual Deviance: 3833.171
## AIC: 3849.171
```

Note that `polr()` uses a reversed sign convention: the raw coefficient reported in the output is negated relative to the standard parameterization. A positive raw coefficient in the `polr()` output corresponds to a *negative* association with higher ordinal categories. The raw coefficient on `Europe` is negative, which — after applying the sign reversal — implies a positive association: respondents with stronger pro-EU attitudes tend to perceive national economic conditions as having improved. This is plausible: Blair's government was broadly pro-European and pro-single-market, so EU supporters may have held more favorable views of its economic stewardship. For reliable interpretation of magnitudes, we use average marginal effects below.

c) Compute average marginal effects across all response categories:

```
avg_slopes(m_ologit)
```

```
##
##                  Term Group      Contrast   Estimate Std. Error     z Pr(>|z|)
##   Europe              1 dY/dX          0.0028957  0.0005836  4.96  < 0.001
##   Europe              2 dY/dX          0.0157837  0.0019914  7.93  < 0.001
##   Europe              3 dY/dX          0.0097485  0.0013067  7.46  < 0.001
##   Europe              4 dY/dX         -0.0222323  0.0025214 -8.82  < 0.001
##   Europe              5 dY/dX         -0.0061956  0.0009877 -6.27  < 0.001
##   age                 1 dY/dX         -0.0000944  0.0000734 -1.29  0.19832
##   age                 2 dY/dX         -0.0005145  0.0003918 -1.31  0.18910
##   age                 3 dY/dX         -0.0003178  0.0002429 -1.31  0.19072
##   age                 4 dY/dX          0.0007247  0.0005506  1.32  0.18810
```

```
## age                   5 dY/dX          0.0002020  0.0001553  1.30  0.19338
## gender                1 male - female -0.0044922  0.0023459 -1.91  0.05550
## gender                2 male - female -0.0246336  0.0123540 -1.99  0.04615
## gender                3 male - female -0.0154980  0.0079927 -1.94  0.05250
## gender                4 male - female  0.0348983  0.0175541  1.99  0.04681
## gender                5 male - female  0.0097255  0.0049872  1.95  0.05116
## political.knowledge   1 dY/dX          0.0028491  0.0011560  2.46  0.01372
## political.knowledge   2 dY/dX          0.0155298  0.0058173  2.67  0.00759
## political.knowledge   3 dY/dX          0.0095917  0.0036514  2.63  0.00862
## political.knowledge   4 dY/dX         -0.0218747  0.0081279 -2.69  0.00712
## political.knowledge   5 dY/dX         -0.0060960  0.0023669 -2.58  0.01001
##     S      2.5 %     97.5 %
## 20.4  0.0017518  0.0040396
## 48.6  0.0118806  0.0196869
## 43.4  0.0071873  0.0123097
## 59.6 -0.0271741 -0.0172906
## 31.4 -0.0081315 -0.0042598
##  2.3 -0.0002382  0.0000494
##  2.4 -0.0012824  0.0002534
##  2.4 -0.0007938  0.0001582
##  2.4 -0.0003545  0.0018039
##  2.4 -0.0001024  0.0005063
##  4.2 -0.0090900  0.0001056
##  4.4 -0.0488471 -0.0004202
##  4.3 -0.0311634  0.0001675
##  4.4  0.0004929  0.0693037
##  4.3 -0.0000492  0.0195003
##  6.2  0.0005834  0.0051149
##  7.0  0.0041282  0.0269314
##  6.9  0.0024350  0.0167484
##  7.1 -0.0378051 -0.0059442
##  6.6 -0.0107350 -0.0014569
##
## Type: probs
```

The AMEs show the average change in the probability of each response category associated with a one-unit increase in each predictor. For `Europe`, the AMEs on the lower categories (1 and 2) are negative, while the AMEs on the higher categories (4 and 5) are positive, consistent with a positive association between pro-EU sentiment and more optimistic economic assessments. As a sanity check, the AMEs for any given predictor must sum to zero across the five categories because probabilities are constrained to sum to one.

d) Predicted probabilities by gender, at the mean of all other covariates:

```r
predictions(m_ologit, newdata = datagrid(gender = c("female", "male")))
```

```
##
##  Group gender Estimate Std. Error    z Pr(>|z|)    S  2.5 % 97.5 %
##      1 female   0.0267    0.00451 5.92   <0.001 28.2 0.0179 0.0356
##      1 male     0.0222    0.00380 5.84   <0.001 27.5 0.0147 0.0296
```

```
##       2 female   0.1874    0.01265 14.82   <0.001 162.6 0.1627 0.2122
##       2 male     0.1615    0.01160 13.92   <0.001 143.8 0.1387 0.1842
##       3 female   0.4237    0.01368 30.98   <0.001 697.6 0.3969 0.4505
##       3 male     0.4088    0.01381 29.61   <0.001 637.7 0.3817 0.4359
##       4 female   0.3200    0.01547 20.69   <0.001 313.5 0.2897 0.3503
##       4 male     0.3570    0.01621 22.02   <0.001 354.5 0.3252 0.3888
##       5 female   0.0421    0.00518  8.14   <0.001  51.2 0.0320 0.0523
##       5 male     0.0506    0.00604  8.37   <0.001  54.0 0.0388 0.0624
##
## Type: probs
```

The predicted probabilities for the most pessimistic category (1 = got much worse) and the most optimistic category (5 = got much better) are shown separately for female and male respondents, with all other covariates held at their sample means. Any gender differences in these predicted probabilities should be modest, given that `gender` does not appear to be a strong driver of economic perceptions relative to the other covariates in the model. The overlapping confidence intervals for male and female respondents across most categories suggest that the gender gap in economic optimism is not large in this dataset.

2. Multinomial logit: vote choice

a) Set the reference category and fit the multinomial logit:

```
BEPS$vote = relevel(BEPS$vote, ref = "Conservative")
m_mlogit = multinom(vote ~ economic.cond.national + Blair + Hague +
                           Kennedy + Europe, data = BEPS, trace = FALSE)
summary(m_mlogit)
```

```
## Call:
## multinom(formula = vote ~ economic.cond.national + Blair + Hague +
##     Kennedy + Europe, data = BEPS, trace = FALSE)
##
## Coefficients:
##                  (Intercept) economic.cond.national      Blair      Hague
## Labour            -0.7685077              0.6228794 0.8157754 -0.8805887
## Liberal Democrat   0.2627294              0.1840322 0.2847314 -0.8228982
##                    Kennedy     Europe
## Labour           0.2396046 -0.2102419
## Liberal Democrat 0.6763287 -0.2107956
##
## Std. Errors:
##                  (Intercept) economic.cond.national      Blair      Hague
## Labour             0.4973142             0.09708813 0.07381306 0.07112439
## Liberal Democrat   0.5276718             0.10184961 0.07486476 0.07588714
##                     Kennedy     Europe
## Labour           0.07532422 0.02688425
## Liberal Democrat 0.08381384 0.02849781
##
## Residual Deviance: 2364.339
```

```
## AIC: 2388.339
```

The model produces two sets of log-odds coefficients: Labour vs. Conservative and Liberal Democrat vs. Conservative. The coefficient on `Blair` in the Labour vs. Conservative equation is strongly positive: higher approval of Tony Blair is associated with substantially greater log-odds of voting Labour rather than Conservative. This makes intuitive sense — Blair was the Labour leader, so voters who rated him favorably were much more likely to have voted for his party. By contrast, the `Blair` coefficient in the Liberal Democrat vs. Conservative equation is expected to be smaller or near zero, since Blair approval does not strongly differentiate Liberal Democrat voters from Conservatives.

b) Compute average marginal effects across all predictors and outcome categories:

```
avg_slopes(m_mlogit)
```

```
##
##                     Term            Group Estimate Std. Error      z Pr(>|z|)
##  Blair                   Conservative      -0.0728    0.00703 -10.36  < 0.001
##  Blair                   Labour             0.1156    0.00920  12.57  < 0.001
##  Blair                   Liberal Democrat  -0.0428    0.00875  -4.89  < 0.001
##  Europe                  Conservative       0.0256    0.00283   9.06  < 0.001
##  Europe                  Labour            -0.0151    0.00341  -4.41  < 0.001
##  Europe                  Liberal Democrat  -0.0106    0.00317  -3.33  < 0.001
##  Hague                   Conservative       0.1043    0.00636  16.39  < 0.001
##  Hague                   Labour            -0.0695    0.00861  -8.07  < 0.001
##  Hague                   Liberal Democrat  -0.0348    0.00814  -4.27  < 0.001
##  Kennedy                 Conservative      -0.0509    0.00808  -6.30  < 0.001
##  Kennedy                 Labour            -0.0296    0.01046  -2.83  0.00462
##  Kennedy                 Liberal Democrat   0.0805    0.00998   8.07  < 0.001
##  economic.cond.national Conservative      -0.0539    0.01046  -5.15  < 0.001
##  economic.cond.national Labour             0.0918    0.01303   7.05  < 0.001
##  economic.cond.national Liberal Democrat  -0.0379    0.01224  -3.10  0.00196
##      S   2.5 %   97.5 %
##   81.2 -0.0866 -0.05904
##  117.9  0.0976  0.13362
##   19.9 -0.0599 -0.02563
##   62.8  0.0201  0.03114
##   16.6 -0.0217 -0.00837
##   10.2 -0.0168 -0.00435
##  198.2  0.0918  0.11674
##   50.3 -0.0864 -0.05260
##   15.7 -0.0507 -0.01883
##   31.7 -0.0667 -0.03506
##    7.8 -0.0501 -0.00912
##   50.3  0.0610  0.10008
##   21.9 -0.0744 -0.03343
##   39.0  0.0663  0.11738
##    9.0 -0.0619 -0.01391
##
## Type: probs
```

```
## Comparison: dY/dX
```

The AME of `Blair` on the probability of voting Labour is positive and substantial. A one-unit increase in Blair approval (on the 1–5 scale) is associated with a meaningful increase in the average probability of voting Labour, holding all other variables constant. This reflects the strong personalization of vote choice in 1997: feelings toward the party leader were a major driver of vote intention, and Blair in particular was unusually popular relative to his Conservative counterpart.

c) The multinomial logit assumes Independence of Irrelevant Alternatives (IIA): the odds ratio between any two alternatives (e.g., Labour vs. Conservative) is unaffected by the presence or characteristics of the third alternative (Liberal Democrats). In the red bus / blue bus analogy, IIA fails because two alternatives are near-perfect substitutes and removing one simply shifts its probability to the other rather than distributing it proportionally. For British party choice, IIA is a moderate concern: Labour and the Liberal Democrats are both centre-left parties, sharing some ideological space, so some voters may treat them as partial substitutes in a way IIA cannot accommodate. The Conservatives, however, occupy a clearly distinct ideological position (right-wing), so the three-party menu is not as degenerate as two buses of different colours. Overall, IIA is plausible for Conservative vs. the others but is a more legitimate worry for the Labour/Liberal Democrat distinction.

```r
data(bioChemists)
```

3. Poisson regression: publication counts

a) Explore the outcome variable:

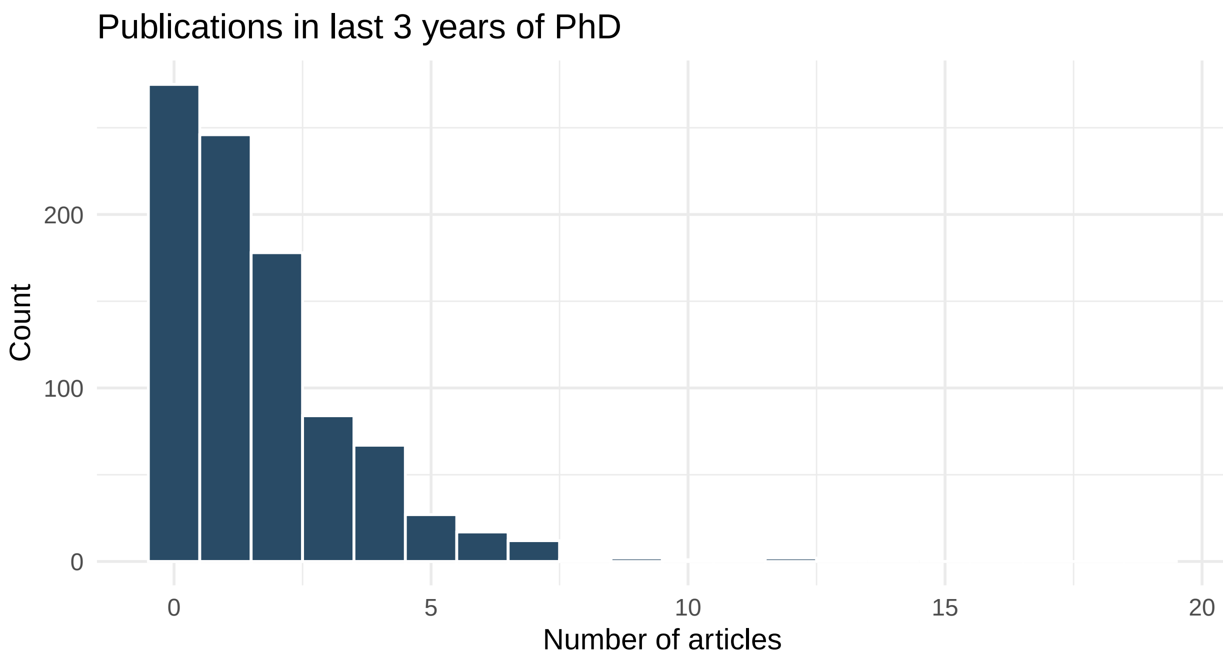
```r
summary(bioChemists$art)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.000   1.000   1.693   2.000  19.000
```

```r
var(bioChemists$art)
```

```
## [1] 3.709742
```

```r
ggplot(bioChemists, aes(x = art)) +
  geom_histogram(binwidth = 1, fill = "#294b66", color = "white") +
  theme_minimal() +
  labs(title = "Publications in last 3 years of PhD",
       x = "Number of articles", y = "Count")
```



The distribution of `art` is right-skewed, with a mode at zero and a long upper tail. The mean is around 1.69 while the variance is approximately 3.71 — roughly twice the mean. Under the Poisson assumption, the variance should equal the mean; a ratio substantially above 1 indicates overdispersion. This pattern is a first signal that a standard Poisson model may underestimate uncertainty and produce anti-conservative standard errors.

b) Fit the Poisson regression:

```
m_pois = glm(art ~ fem + mar + kid5 + phd + ment,
             data = bioChemists, family = poisson)
summary(m_pois)
```

```
##
## Call:
## glm(formula = art ~ fem + mar + kid5 + phd + ment, family = poisson,
##     data = bioChemists)
##
## Coefficients:
##              Estimate Std. Error z value             Pr(>|z|)
## (Intercept)  0.304617   0.102981   2.958               0.0031 **
## femWomen    -0.224594   0.054613  -4.112           0.00003915 ***
## marMarried   0.155243   0.061374   2.529               0.0114 *
## kid5        -0.184883   0.040127  -4.607           0.00000408 ***
## phd          0.012823   0.026397   0.486               0.6271
## ment         0.025543   0.002006  12.733 < 0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1817.4  on 914  degrees of freedom
## Residual deviance: 1634.4  on 909  degrees of freedom
```

```
## AIC: 3314.1
##
## Number of Fisher Scoring iterations: 5
```

```
exp(coef(m_pois)["ment"])
```

```
##     ment
## 1.025872
```

The incidence rate ratio (IRR) for `ment` is 1.026: each additional article published by the mentor is associated with a multiplicative increase in expected student articles by that factor, holding all else constant. The effect is modest but positive, suggesting that more productive mentors slightly boost student output. The residual deviance is substantially larger than the residual degrees of freedom (their ratio is well above 2), which is another clear diagnostic signal of overdispersion — the Poisson model does not adequately capture the variation in publication counts.

c) Formal overdispersion test:

```
dispersiontest(m_pois)
```

```
##
##  Overdispersion test
##
## data:  m_pois
## z = 5.7825, p-value = 0.000000003681
## alternative hypothesis: true dispersion is greater than 1
## sample estimates:
## dispersion
##    1.82454
```

The dispersion test strongly rejects the null hypothesis of equidispersion ($p < 0.001$). The estimated dispersion parameter is well above 1, confirming that the variance in `art` substantially exceeds its mean. This means the Poisson standard errors are too small: the model underestimates uncertainty, inflates test statistics, and produces p-values that are misleadingly small. A model that explicitly accounts for overdispersion — such as the negative binomial — is needed.

4. Negative binomial regression

a) Fit the negative binomial model:

```
m_nb = glm.nb(art ~ fem + mar + kid5 + phd + ment,
              data = bioChemists)
summary(m_nb)
```

```
##
## Call:
## glm.nb(formula = art ~ fem + mar + kid5 + phd + ment, data = bioChemists,
##     init.theta = 2.264387695, link = log)
##
## Coefficients:
##                Estimate Std. Error z value             Pr(>|z|)
```

```
## (Intercept)  0.256144   0.137348   1.865             0.062191 .
## femWomen    -0.216418   0.072636  -2.979             0.002887 **
## marMarried   0.150489   0.082097   1.833             0.066791 .
## kid5        -0.176415   0.052813  -3.340             0.000837 ***
## phd          0.015271   0.035873   0.426             0.670326
## ment         0.029082   0.003214   9.048 < 0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(2.2644) family taken to be 1)
##
##     Null deviance: 1109.0  on 914  degrees of freedom
## Residual deviance: 1004.3  on 909  degrees of freedom
## AIC: 3135.9
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  2.264
##           Std. Err.:  0.271
##
##  2 x log-likelihood:  -3121.917
```

The coefficient on `ment` is similar to the Poisson estimate, indicating that the point estimate is reasonably stable. The key difference is in the standard errors: the negative binomial model produces larger, more honest uncertainty estimates. The estimated overdispersion parameter `theta` (shown in the summary) quantifies how much the variance exceeds the Poisson prediction; a smaller `theta` means more severe overdispersion. Here `theta` is moderate, indicating meaningful but not extreme extra-Poisson variation.

b) Compare model fit by AIC:

```
AIC(m_pois)
```

```
## [1] 3314.113
```

```
AIC(m_nb)
```

```
## [1] 3135.917
```

The negative binomial AIC is substantially lower than the Poisson AIC, despite the NB model having one additional parameter (`theta`). Under AIC, the improvement in fit more than compensates for the added complexity. This confirms that overdispersion is a genuine feature of the data, not noise, and that the negative binomial is the more appropriate model for these publication counts.

c) Predicted article counts by gender, holding other variables at sample means:

```
predictions(m_nb, newdata = datagrid(fem = c("Men", "Women")))
```

```
##
##    fem Estimate Pr(>|z|)    S 2.5 % 97.5 %
##  Men       2.05   <0.001 93.7  1.80   2.32
##  Women     1.65   <0.001 42.2  1.44   1.88
```

```
##
## Type: invlink(link)
```

The predicted number of articles for men exceeds that for women, holding marital status, number of young children, PhD prestige, and mentor productivity constant at their sample means. The confidence intervals provide information on whether this gender gap is statistically distinguishable: if the intervals do not overlap, the difference is significant at conventional levels. The gap reflects a persistent within-group gender difference in publication productivity that is not simply an artefact of other observable characteristics.

d) Summary of findings:

The Poisson model is not adequate for this dataset. The variance-to-mean ratio of `art` is roughly double, the residual deviance far exceeds the degrees of freedom, and the formal `dispersiontest()` rejects equidispersion with a p-value well below 0.001. The negative binomial model, which adds a dispersion parameter to accommodate this extra variation, achieves a substantially lower AIC and produces more reliable (wider) standard errors. On substantive findings: the mentor's productivity (`ment`) has a positive and statistically significant effect, with an IRR slightly above 1 — each additional mentor article is associated with a modest multiplicative increase in expected student articles, suggesting that working with a productive mentor confers a real, if small, boost. Gender (`fem`) and number of young children (`kid5`) are both negative and statistically significant: women publish fewer articles on average, and each additional child under age 5 is associated with reduced output. PhD program prestige (`phd`) and marital status (`mar`) are not statistically significant in the negative binomial model. Together, the results point to early-career productivity being shaped by mentor environment, gender, and family demands — patterns consistent with broader literature on PhD student outcomes in STEM fields.