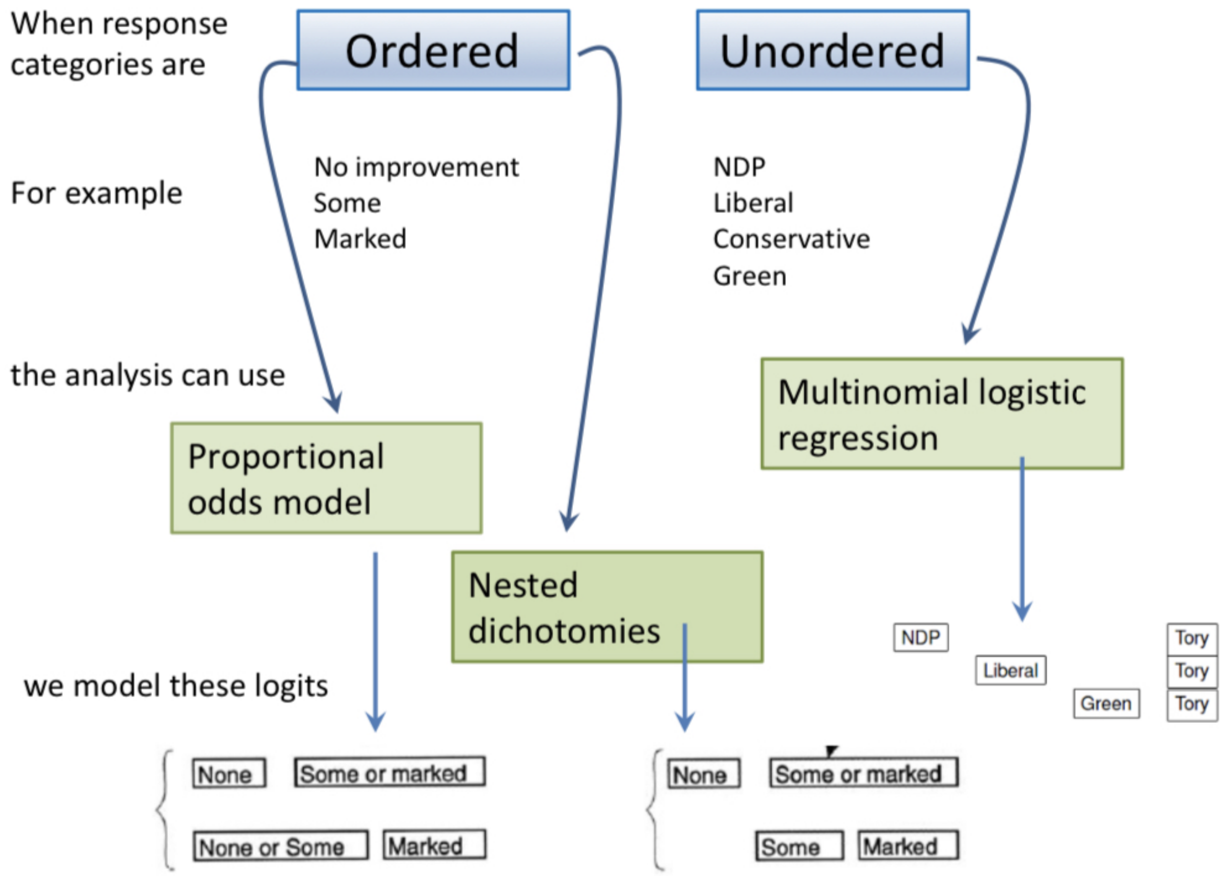
PSY 504: Advanced Statistics

Multinomial Regression

Jason Geller, Ph.D. (he/him/his)

Princeton University

Updated:2023-02-25


When response categories are
Ordered
Unordered
For example
No improvement
Some
Marked
NDP
Liberal
Conservative
Green
the analysis can use
Proportional odds model
Nested dichotomies
Multinomial logistic regression
we model these logits
None
Some or marked
None or Some
Marked
None
Some or marked
Some
Marked
NDP
Liberal
Green
Tory
Tory
Tory

Multinomial Logistic Regression

- In ordinal regression:

$$
\begin{aligned}
L_1 &= \alpha_1 - \beta_1 x_1 + \cdots + \beta_p X_p \\
L_2 &= \alpha_2 - \beta_1 x_1 + \cdots + \beta_p X_p \\
L_{J-1} &= \alpha_{J-1} - \beta_1 x_1 + \cdots + \beta_p X_p
\end{aligned}
$$

- In the multinomial logistic model:

$$
\begin{aligned}
L_1 &= \alpha_1 + \beta_1 x_1 + \cdots + \beta_p X_p \\
L_2 &= \alpha_2 + \beta_2 x_1 + \cdots + \beta_p X_p \\
L_{J-1} &= \alpha_{J-1} + \beta_j x_1 + \cdots + \beta_p X_p
\end{aligned}
$$

Multinomial Logistic Regression

- Choose a baseline category. Let's choose $y = 0$. Then,

$$P(y_i = 0|x_i;) = P_{i0} and P(y_i = 1|x_i) = P_{i1}$$

$$\log\left(\frac{p_{i1}}{p_{i0}}\right) = \beta_{0k} + \beta_{1k}x_i$$

- Slope, β_1: when x increases by one unit, the odds of Y = 1 vs. baseline is expected to multiply by a factor or $exp(\beta)$

- Intercept, β_0, when x = 0 the odds of Y = 1 is expeted to be exp(β)

Multinomial Logistic Regression

- Which of the following best describes your pattern of study?
  - Light cram
  - Heavy cram
  - Space out
- Let "Space out" be the baseline category. Then

$$\log\left(\frac{\pi_{light}}{\pi_{space}}\right) = \beta_{0B} + \beta_{1B}x_i$$

$$\log\left(\frac{\pi_{heavy}}{\pi_{space}}\right) = \beta_{0C} + \beta_{1C}x_i$$

Summary

- Multinomial logistic regression models the probabilities of j response categories (j-1)
  - Typically these compare each of the first m-1 categories to the last (reference) category
  - 1 vs. m, 2 vs.m, 3 vs. m
- Logits for any pair of categories can be calculated from the m-1 fitted ones

NHANES Data

- National Health and Nutrition Examination Survey is conducted by the National Center for Health Statistics (NCHS)
- The goal is to *"assess the health and nutritional status of adults and children in the United States"*
- This survey includes an interview and a physical examination

NHANES Data

- We will use the data from the **`NHANES`** R package
- Contains 75 variables for the 2009 - 2010 and 2011 - 2012 sample years
- The data in this package is modified for educational purposes and should not be used for research
- Original data can be obtained from the [NCHS website]() for research purposes
- Type **`?NHANES`** in console to see list of variables and definitions

Health Rating vs. Age & Physical Activity

- Question: Can we use a person's age and whether they do regular physical activity to predict their self-reported health rating?

- We will analyze the following variables:

  - **`HealthGen`:** Self-reported rating of participant's health in general. Excellent, Vgood, Good, Fair, or Poor.

  - **`Age`:** Age at time of screening (in years). Participants 80 or older were recorded as 80.

  - **`PhysActive`:** Participant does moderate to vigorous-intensity sports, fitness or recreational activities

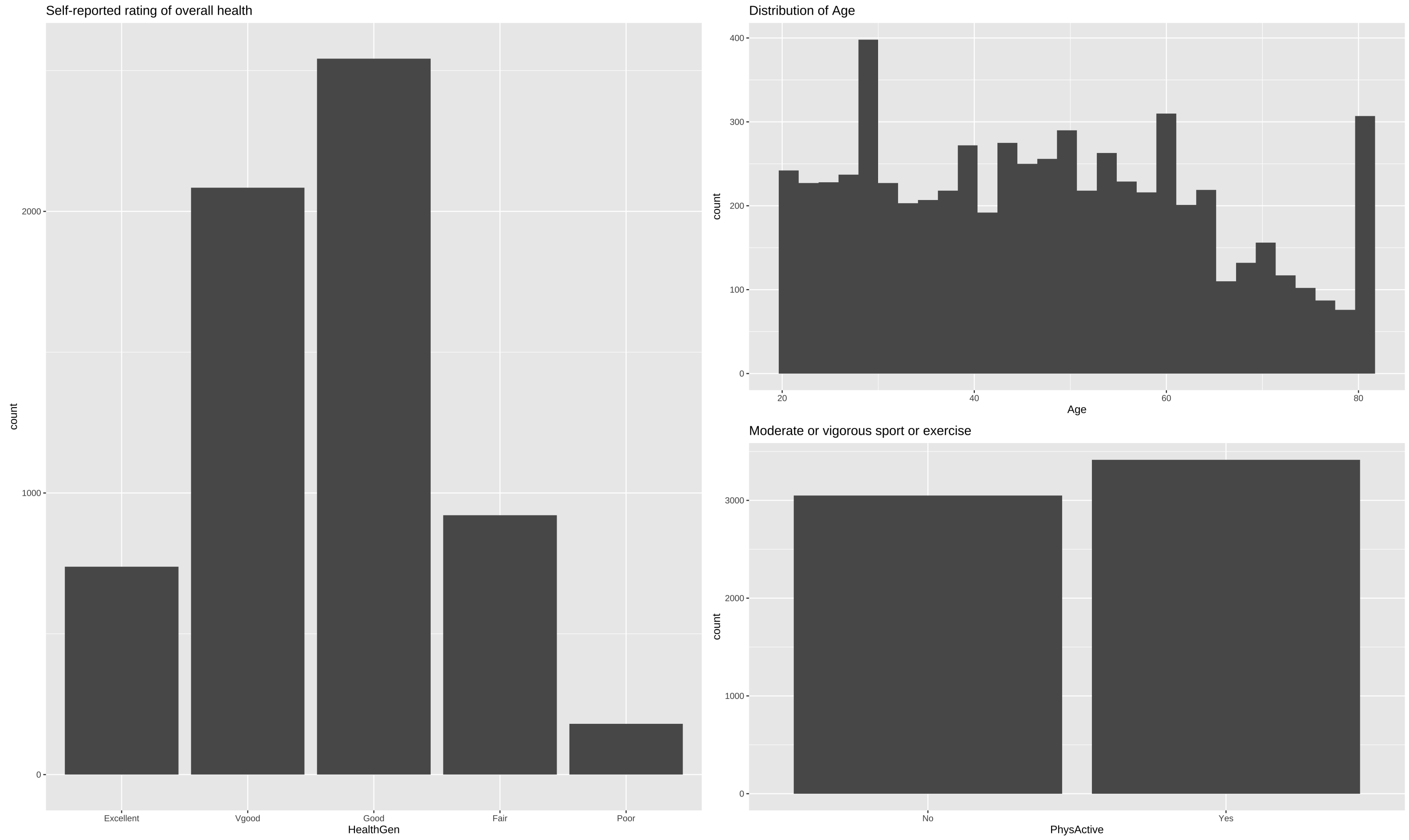
The data

```
library(NHANES)
nhanes_adult <- NHANES %>%
  #only use ages 18+
  filter(Age >= 18) %>%
  #select 4 vars from the full dataset
  dplyr::select(HealthGen, Education, Age, PhysActive) %>%
  # get rid of nas
  drop_na() %>%
  mutate(obs_num = 1:n())
```

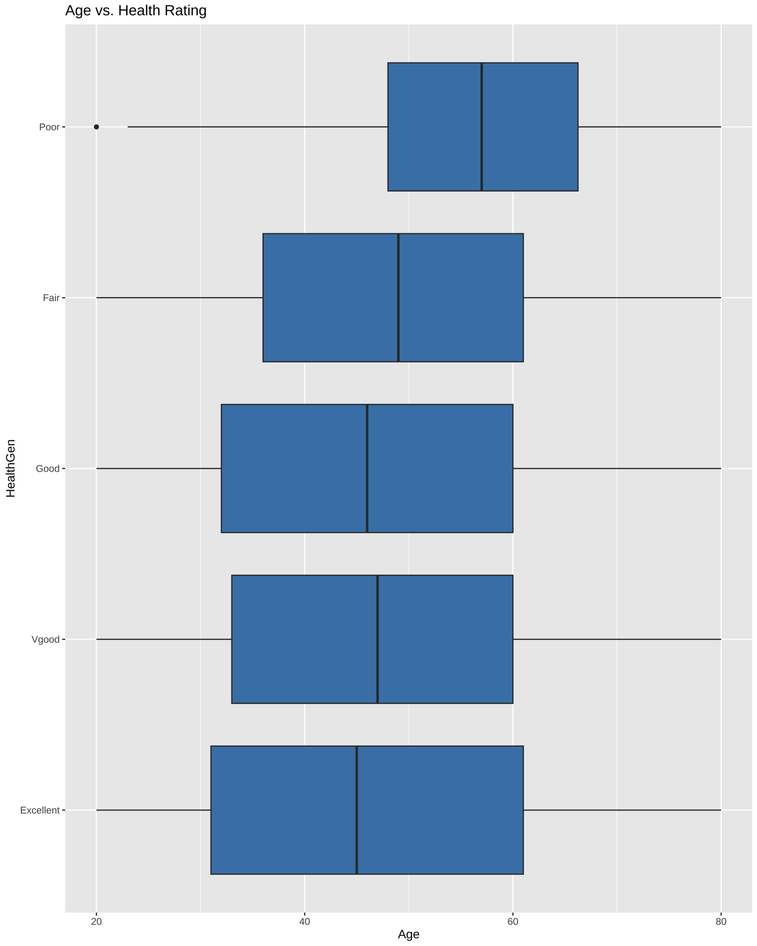
```
glimpse(nhanes_adult)
```

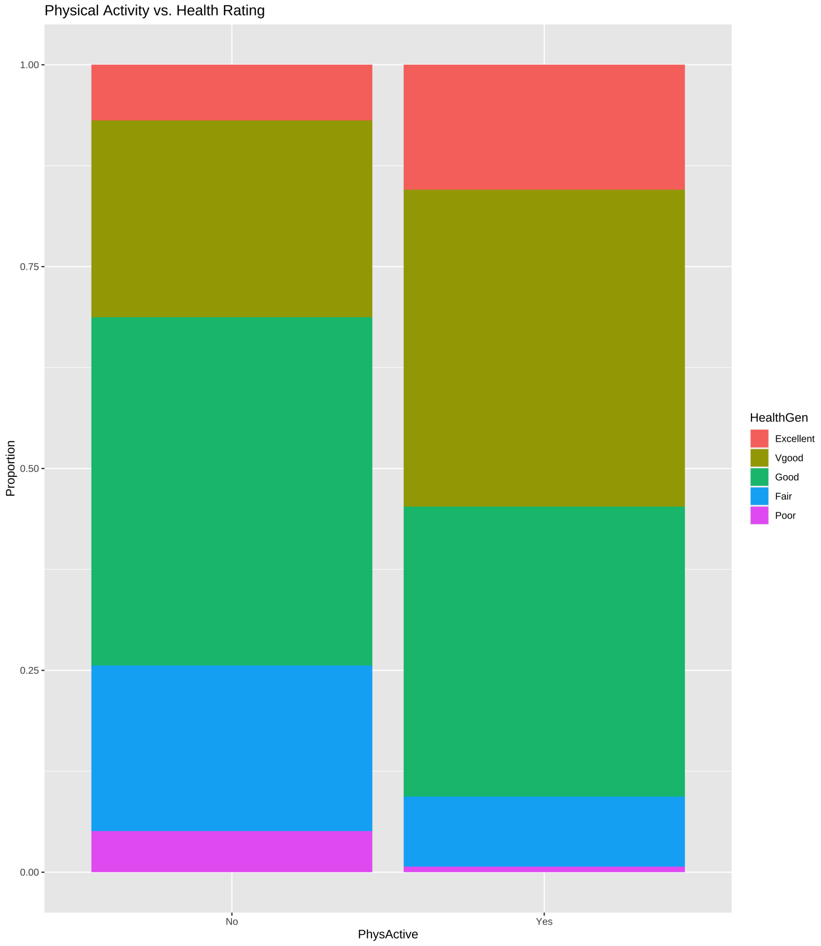
```
## Rows: 6,465
## Columns: 5
## $ HealthGen  <fct> Good, Good, Good, Good, Vgood, Vgood, Vgood, Vgood, Vgood, …
## $ Education  <fct> High School, High School, High School, Some College, Colleg…
## $ Age        <int> 34, 34, 34, 49, 45, 45, 45, 66, 58, 54, 50, 33, 60, 56, 56,…
## $ PhysActive <fct> No, No, No, No, Yes, Yes, Yes, Yes, Yes, Yes, Yes, No, No, …
## $ obs_num    <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, …
```

Exploratory data analysis



Exploratory data analysis





Multinomial Model in R

- Use the `multinom()` function in the **nnet** package

```
library(nnet)
health_m <- multinom(HealthGen ~ PhysActive + Age,
                     data = nhanes_adult)
```

- Put `results = "hide"` in the code chunk header to suppress convergence output

Output results

```
model_parameters(health_m, exponentiate = FALSE)%>%  filter(Response=="Fair")
  kable(digits = 3, format = "markdown")
```

Parameter	Coefficient	SE	CI	CI_low	CI_high	t	df_error	p	Response
(Intercept)	1.033	0.174	0.95	0.692	1.374	5.938	6453	0.000	Fair
PhysActiveYes	-1.662	0.109	0.95	-1.877	-1.448	-15.190	6453	0.000	Fair
Age	0.001	0.003	0.95	-0.005	0.007	0.373	6453	0.709	Fair

Output results

```r
model_parameters(health_m, exponentiate = TRUE) %>% filter(Response=="Fair")
```

Parameter	Coefficient	SE	CI	CI_low	CI_high	t	df_error	p	Response
(Intercept)	2.809	0.489	0.95	1.998	3.951	5.938	6453	0.000	Fair
PhysActiveYes	0.190	0.021	0.95	0.153	0.235	-15.190	6453	0.000	Fair
Age	1.001	0.003	0.95	0.995	1.007	0.373	6453	0.709	Fair

Fair vs. Excellent Health

The baseline category for the model is **`Excellent`**.

The model equation for the log-odds a person rates themselves as having "Fair" health vs. "Excellent" is

$$\log\left(\frac{\hat{\pi}_{Fair}}{\hat{\pi}_{Excellent}}\right) = 1.09 + 0.003 \text{ age} - 1.66 \text{ PhysActive}$$

Interpretations

$$\log\left(\frac{\hat{\pi}_{Fair}}{\hat{\pi}_{Excellent}}\right) = 1.09 + 0.003 \text{ age} - 1.66 \text{ PhysActive}$$

- For each additional year in age, the odds a person rates themselves as having fair health versus excellent health are expected to multiply by 1.003 (exp(0.003)), holding physical activity constant.
    - As Age ⬆️, more likely to report Fair vs. Excellent health

Interpretations

$$\log\left(\frac{\hat{\pi}_{Fair}}{\hat{\pi}_{Excellent}}\right) = 1.09 + 0.003 \text{ age} - 1.66 \text{ PhysActive}$$

- The odds a person who does physical activity will rate themselves as having fair health versus excellent health are expected to be 0.19 **`(exp(-1.66 ))`** times the odds for a person who doesn't do physical activity, holding age constant.
    - A person who does physical activity is more likely to rate themselves in Excellent vs. Fair health

Interpretations

$$\log\left(\frac{\hat{\pi}_{Fair}}{\hat{\pi}_{Excellent}}\right) = 1.09 + 0.003 \text{ age} - 1.66 \text{ PhysActive}$$

- The odds a 0 year old person who doesn't do physical activity rates themselves as having fair health vs. excellent health are 2.974 `(exp(1.09))`.
⚠️ Need to mean-center age for the intercept to have a meaningful interpretation!

Good vs. Excellent Health

```
model_parameters(health_m, exponentiate = FALSE)%>%  filter(Response=="Good")
  kable(digits = 3, format = "markdown")
```

Parameter	Coefficient	SE	CI	CI_low	CI_high	t	df_error	p	Response
(Intercept)	1.989	0.150	0.95	1.695	2.282	13.285	6453	0.000	Good
PhysActiveYes	-1.011	0.092	0.95	-1.192	-0.831	-10.979	6453	0.000	Good
Age	-0.003	0.003	0.95	-0.008	0.002	-1.187	6453	0.235	Good

```r
model_parameters(health_m, exponentiate = TRUE) %>%
  filter(Response=="Good") %>%
  kable(digits = 3, format = "markdown")
```

Parameter	Coefficient	SE	CI	CI_low	CI_high	t	df_error	p	Response
(Intercept)	7.306	1.094	0.95	5.448	9.797	13.285	6453	0.000	Good
PhysActiveYes	0.364	0.034	0.95	0.304	0.436	-10.979	6453	0.000	Good
Age	0.997	0.003	0.95	0.992	1.002	-1.187	6453	0.235	Good

Poor vs. Excellent Health

The baseline category for the model is **`Excellent`**.

The model equation for the log-odds a person rates themselves as having "Good" health vs. "Excellent" is

$$\log\left(\frac{\hat{\pi}_{Good}}{\hat{\pi}_{Excellent}}\right) = 1.844 - 0.003 \text{ age} - 1.011 \text{ PhysActive}$$

Interpretations

$$\log\left(\frac{\hat{\pi}_{Good}}{\hat{\pi}_{Excellent}}\right) = 1.844 - 0.003 \text{ age} - 1.011 \text{ PhysActive}$$

- For each additional year in age, the odds a person rates themselves as having poor health versus excellent health are expected to multiply by 1.003 (exp(0.02)), holding physical activity constant
    - As Age ⬆️, higher probability to report good health vs. excellent health

Interpretations

```
$$\log\Big(\frac{\hat{\pi}{Good}}{\hat{\pi}
{Excellent}}\Big) = 1.844 - 0.003 ~ \text{age} - 1.011
~ \text{PhysActive}$$
```

- The odds a person who does physical activity will rate themselves as having poor health versus excellent health are expected to be 0.364 **`(exp(-1.01))`** times the odds for a person who doesn't do physical activity, holding age constant
  - A person who does physical activity is more likely to rate themselves in Excellent vs. good health

Interpretations

$$\log\left(\frac{\hat{\pi}_{Good}}{\hat{\pi}_{Excellent}}\right) = 1.844 - 0.003 \text{ age} - 1.011 \text{ PhysActive}$$

- The odds a 0 year old person who doesn't do physical activity rates themselves as having poor health vs. excellent health are 6.297 **`(exp(1.84))`**.

- Those reporting no physical activity are more likely to report Good vs. Excellent health

⚠️ Need to mean-center age for the intercept to have a meaningful interpretation!

Change baseline

```
nhanes_adult %>%
  mutate(HealthGen = relevel(as.factor(HealthGen), ref= "Poor")) %>% dplyr::s
```

Hypothesis test for β_{jk}

The test of significance for the coefficient β_{jk} is

Hypotheses: $H_0 : \beta_{jk} = 0 \text{ vs } H_a : \beta_{jk} \neq 0$

Test Statistic:

$$z = \frac{\hat{\beta}_{jk} - 0}{SE(\hat{\beta}_{jk})}$$

P-value: $P(|Z| > |z|)$,

where $Z \sim N(0,1)$, the Standard Normal distribution

Confidence interval for β_{jk}

- We can calculate the **C% confidence interval** for β_{jk} using the following:

$$\hat{\beta}_{jk} \pm z^* SE(\hat{\beta}_{jk})$$

where z^* is calculated from the $N(0,1)$ distribution

> We are $C\%$ confident that for every one unit change in x_j, the odds of $y = k$ versus the baseline will multiply by a factor of $\exp\{\hat{\beta}_{jk} - z^* SE(\hat{\beta}_{jk})\}$ to $\exp\{\hat{\beta}_{jk} + z^* SE(\hat{\beta}_{jk})\}$, holding all else constant.

Interpreting confidence intervals for β_{jk}

```
tidy(health_m, conf.int = TRUE, exponentiate = FALSE) %>%
  filter(y.level == "Fair") %>%
  kable(digits = 3, format = "markdown")
```

y.level	term	estimate	std.error	statistic	p.value	conf.low	conf.high
Fair	(Intercept)	1.033	0.174	5.938	0.000	0.692	1.374
Fair	PhysActiveYes	-1.662	0.109	-15.190	0.000	-1.877	-1.448
Fair	Age	0.001	0.003	0.373	0.709	-0.005	0.007

We are 95% confident, that for each additional year in age, the odds a person rates themselves as having fair health versus excellent health will multiply by 0.997 (exp(-0.003)) to 1.009 (exp(0.009)) , holding physical activity constant.

Interpreting confidence intervals for β_{jk}

```
tidy(health_m, conf.int = TRUE, exponentiate = FALSE) %>%
  filter(y.level == "Fair") %>%
  kable(digits = 3, format = "markdown")
```

y.level	term	estimate	std.error	statistic	p.value	conf.low	conf.high
Fair	(Intercept)	1.033	0.174	5.938	0.000	0.692	1.374
Fair	PhysActiveYes	-1.662	0.109	-15.190	0.000	-1.877	-1.448
Fair	Age	0.001	0.003	0.373	0.709	-0.005	0.007

We are 95% confident that the odds a person who does physical activity will rate themselves as having fair health versus excellent health are 0.156 (exp(-1.856)) to 0.238 (exp(-1.435)) times the odds for a person

NHANES results

- **`emmeans`** approach

```
multi_an <- emmeans(health_m, ~ PhysActive|HealthGen)

# uses baseline as contast of interest
# can change this to get other baselines
# use trt.vs.ctrl" #ref = newbaseline

coefs = contrast(regrid(multi_an, "log"),"trt.vs.ctrl1", by="PhysActive")

update(coefs, by = "contrast") %>%
  kable(format = "markdown", digits = 3)
```

contrast	PhysActive	estimate	SE	df	t.ratio	p.value
Vgood - Excellent	No	1.264	0.079	12	16.069	0.000
Good - Excellent	No	1.844	0.075	12	24.663	0.000
Fair - Excellent	No	1.087	0.080	12	13.504	0.000

NHANES:

```
contrast(coefs, "revpairwise", by = "contrast") %>%
  kable(format = "markdown", digits = 3)
```

contrast1	contrast	estimate	SE	df	t.ratio	p.value
Yes - No	Vgood - Excellent	-0.332	0.095	12	-3.496	0.004
Yes - No	Good - Excellent	-1.011	0.092	12	-10.979	0.000
Yes - No	Fair - Excellent	-1.662	0.109	12	-15.190	0.000
Yes - No	Poor - Excellent	-2.670	0.236	12	-11.308	0.000

Calculating probabilities

For categories $2, \dots, K$, the probability that the i^{th} observation is in the j^{th} category is

$$\hat{\pi}_{ij} = \frac{\exp\{\hat{\beta}_{0j} + \hat{\beta}_{1j}x_{i1} + \dots + \hat{\beta}_{pj}x_{ip}\}}{1 + \sum\limits_{k=2}^{K}\exp\{\hat{\beta}_{0k} + \hat{\beta}_{1k}x_{i1} + \dots \hat{\beta}_{pk}x_{ip}\}}$$

For the baseline category, $k = 1$, we calculate the probability $\hat{\pi}_{i1}$ as

$$\hat{\pi}_{i1} = 1 - \sum_{k=2}^{K}\hat{\pi}_{ik}$$

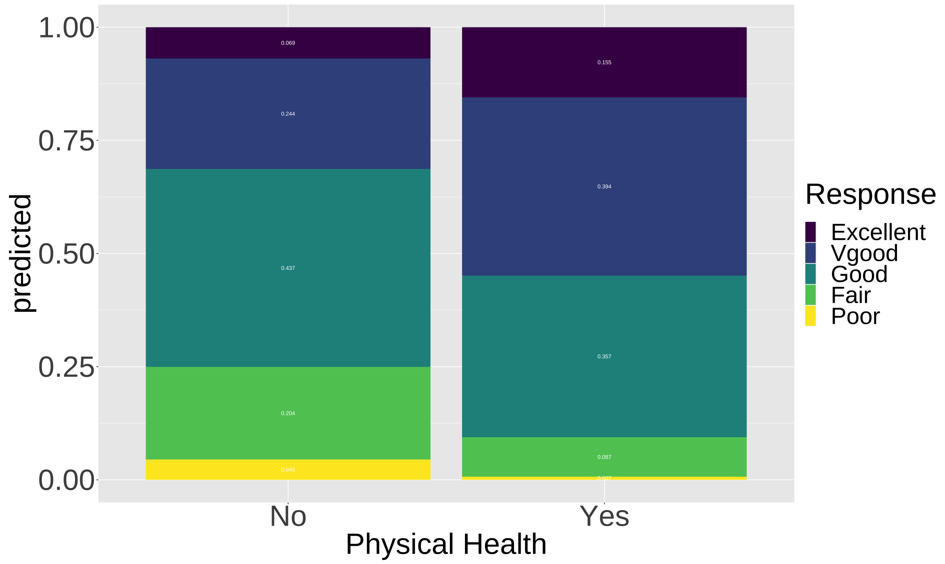
NHANES: Predicted probabilities

```r
#calculate predicted probabilities
ggemmeans(health_m, terms=c("PhysActive")) %>%
  kable(format = "markdown", digits = 3)
```

x	predicted	std.error	conf.low	conf.high	response.level	group
No	0.069	0.005	0.068	0.070	Excellent	1
No	0.244	0.008	0.241	0.247	Vgood	1
No	0.437	0.009	0.433	0.442	Good	1
No	0.204	0.007	0.202	0.207	Fair	1
No	0.045	0.004	0.045	0.045	Poor	1
Yes	0.155	0.006	0.153	0.157	Excellent	1
Yes	0.394	0.008	0.390	0.398	Vgood	1
Yes	0.357	0.008	0.353	0.361	Good	1

NHANES: Predicted probabilities

```r
pred_probs_plot <- ggemmeans(health_
  geom_bar(stat = "identity" ) +
    geom_text(aes(label = round(pred
  easy_add_legend_title("Response")
  labs(x="Physical Health") +
  easy_all_text_size(size=45) +
  scale_fill_viridis(discrete = TRUE
```



Model selection

Comparing nested models

- Suppose there are two models:
  - Reduced Model includes predictors $x_1, \ldots, x_q$
  - Full Model includes predictors $x_1, \ldots, x_q, x_{q+1}, \ldots, x_p$
- We want to test the hypotheses

$$
\begin{aligned}
H_0 &: \beta_{q+1} = \cdots = \beta_p = 0 \\
H_a &: \text{ at least 1 } \beta_j \text{ is not} 0
\end{aligned}
$$

- To do so, we will use the **Drop-in-Deviance test** (very similar to logistic regression)

Add `Education` to the model?

- We consider adding the participants' **`Education`** level to the model
    - Education takes values **`8thGrade`**, **`9-11thGrade`**, **`HighSchool`**, **`SomeCollege`**, and **`CollegeGrad`**
- Models we're testing:
    - Reduced Model: **`Age`**, **`PhysActive`**
    - Full Model: **`Age`**, **`PhysActive`**, **`Education`**

$$H_0 : \beta_{9-11thGrade} = \beta_{HighSchool} = \beta_{SomeCollege} = \beta_{CollegeGrad} = 0$$
$$H_a : \text{ at least one } \beta_j \text{ is not equal to } 0$$

Add `Education` to the model?

$$H_0: \beta_{9-11thGrade} = \beta_{HighSchool} = \beta_{SomeCollege} = \beta_{CollegeGrad} = 0$$

$$H_a: \text{ at least one } \beta_j \text{ is not equal to } 0$$

```
model_red <- multinom(HealthGen ~ Age + PhysActive,
               data = nhanes_adult)
model_full <- multinom(HealthGen ~ Age + PhysActive +
                         Education,
               data = nhanes_adult)
```

Add **Education** to the model?

```r
model_red <- multinom(HealthGen ~ Age + PhysActive,
                      data = nhanes_adult)
model_full <- multinom(HealthGen ~ Age + PhysActive + Education,
                       data = nhanes_adult)
```

```r
anova(model_red, model_full, test = "Chisq") %>%
  kable(format = "markdown")
```

Model	Resid. df	Resid. Dev	Test	Df	LR stat.	Pr(Chi)
Age + PhysActive	25848	16994.23		NA	NA	NA
Age + PhysActive + Education	25832	16505.10	1 vs 2	16	489.1319	0

At least one coefficient associated with **Education** is non-zero. Therefore, we will include **Education** in the model.

Model with `Education`

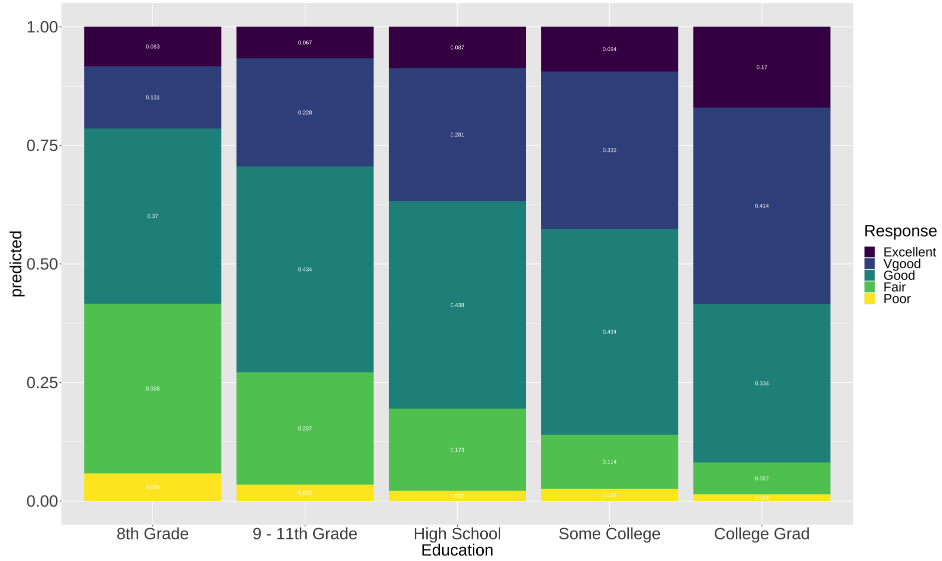
Parameter	Coefficient	SE	CI	CI_low	CI_high	t	df_error	p	Respon
(Intercept)	0.582	0.301	0.95	-0.009	1.173	1.930	6437	0.054	Vgood
Age	0.001	0.003	0.95	-0.004	0.006	0.419	6437	0.675	Vgood
PhysActiveYes	-0.264	0.099	0.95	-0.457	-0.071	-2.681	6437	0.007	Vgood
Education9 - 11th Grade	0.768	0.308	0.95	0.164	1.372	2.493	6437	0.013	Vgood
EducationHigh School	0.701	0.280	0.95	0.153	1.249	2.509	6437	0.012	Vgood
EducationSome College	0.788	0.271	0.95	0.255	1.320	2.901	6437	0.004	Vgood
EducationCollege Grad	0.408	0.268	0.95	-0.117	0.933	1.522	6437	0.128	Vgood
(Intercept)	2.041	0.272	0.95	1.508	2.573	7.513	6437	0.000	Good

Full model

```
car::Anova(model_full, type="II") %>%
  kable(format = "markdown", digits = 3)
```

	LR Chisq	Df	Pr(>Chisq)
Age	19.295	4	0.001
PhysActive	242.631	4	0.000
Education	489.132	16	0.000

```
plot_edu <- ggemmeans(model_full, te
  ggplot(aes(x = x, y = predicted, f
  geom_bar(stat = "identity" ) +
   geom_text(aes(label = round(predi
  easy_add_legend_title("Response")
  labs(x="Education", "Predicted Pro
  easy_all_text_size(size=25) +
   scale_fill_viridis(discrete = TRU
```



Emmeans

- Use **`emmeans`** to extract log odds coefs for compairsons of interest

```
multi_an <- emmeans(model_full, ~ Education|HealthGen)

coefs = contrast(regrid(multi_an, "log"),"trt.vs.ctrl1", by="Education")

update(coefs, by = "contrast") %>%
  kable(format = "markdown", digits = 3)
```

contrast	Education	estimate	SE	df	t.ratio	p.value
Vgood - Excellent	8th Grade	0.456	0.259	28	1.765	0.297
Good - Excellent	8th Grade	1.490	0.223	28	6.681	0.000
Fair - Excellent	8th Grade	1.458	0.221	28	6.609	0.000
Poor - Excellent	8th Grade	-0.344	0.269	28	-1.276	0.572
Vgood - Excellent	9 - 11th Grade	1.232	0.169	28	7.268	0.000

```r
contrast(coefs, "revpairwise", by = "contrast") %>%
  kable(format = "markdown", digits = 3)
```

contrast1	contrast	estimate	SE	df	t.ratio	p.value
(9 - 11th Grade) - 8th Grade	Vgood - Excellent	0.775	0.308	28	2.519	0.115
High School - 8th Grade	Vgood - Excellent	0.714	0.279	28	2.554	0.107
High School - (9 - 11th Grade)	Vgood - Excellent	-0.061	0.200	28	-0.307	0.998
Some College - 8th Grade	Vgood - Excellent	0.804	0.272	28	2.958	0.045
Some College - (9 - 11th Grade)	Vgood - Excellent	0.028	0.188	28	0.150	1.000
Some College - High	Vgood -	0.090	0.135	28	0.664	0.962

Assumptions for multinomial logistic regression

We want to check the following assumptions for the multinomial logistic regression model:

1. **Linearity**: Is there a linear relationship between the log-odds and the predictor variables?

2. **Randomness**: Was the sample randomly selected? Or can we reasonably treat it as random?

3. **Independence**: There is no obvious relationship between observations

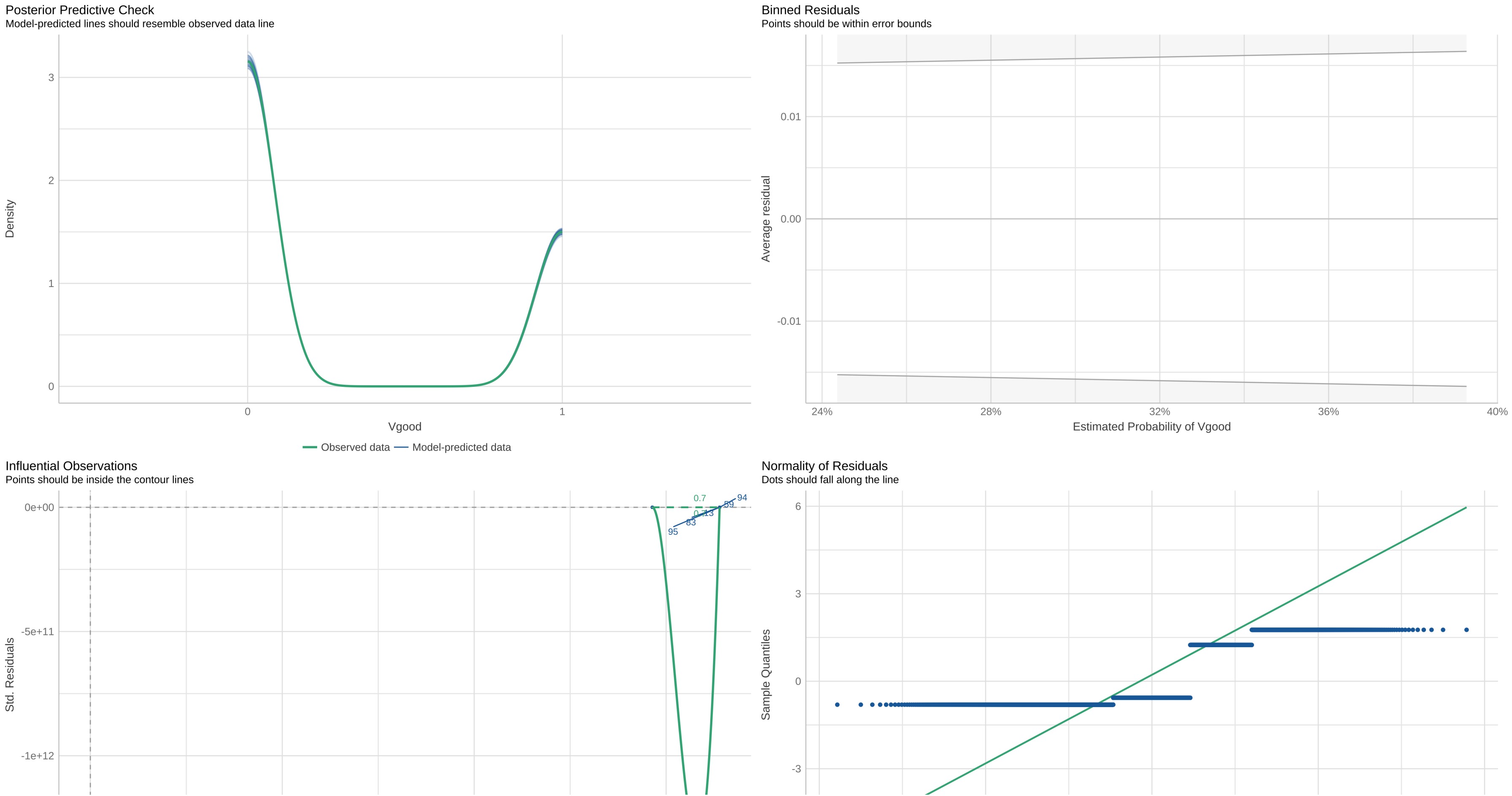
Assumptions for multinomial logistic regression

- Fit sep logistic regressions to check for linearity, outliers, and multicolinearity

```
nhanes_adult <- nhanes_adult %>%
  mutate(Excellent = factor(if_else(HealthGen == "Excellent", "1", "0")),
         Vgood = factor(if_else(HealthGen == "Vgood", "1", "0")),
         Good = factor(if_else(HealthGen == "Good", "1", "0")),
         Fair = factor(if_else(HealthGen == "Fair", "1", "0")),
         Poor = factor(if_else(HealthGen == "Poor", "1", "0"))
  )
```

```r
library(Stat2Data)

glm(Vgood~PhysActive, family=binomial, data=nhanes_adult) %>%
  performance::check_model(.)
```



Write-up

- Report full model results
    - χ^2 test
    - Age, $\chi^2(4)$ = 19.30, $p < .05$
    - PhysActive, $\chi^2(4)$ = 242.63, $p < .05$
    - Education,$\chi^2 (16)$ = 489.13, $p < .05$
- R^2
- Comparisons
    - Significant differences between all levels of self-reported health were compared against baseline of Excellence

Table

Table 2. Odds ratios for questions examining scheduling as a function of GPA and achievement goals.

Response	Frequency (%)	GPA	Performance Approach	Performance Avoidance	Mastery Approach	Mastery Avoidance
Question 2: How do you decide what to study next? (compared to "whatever's due soonest/overdue", 58% of responses)						
Longest time	2	0.71	1.04	0.95	1.08	1.09
Interesting	4	0.84	1.02	0.88**	0.94	1.00
Worst in	19	0.96	0.97	1.00	1.01	1.00
Plan ahead	17	1.24*	0.99	0.97	1.14***	0.97
Question 10: Which of the following best describes your pattern of study? (compared to "space out", 17% of responses)						
Light Cram	65	1.01	1.00	1.03	0.90**	1.03
Heavy Cram	17	0.96	0.99	1.11**	0.85***	1.01
Question 8: What time of day do you most often do your studying? (compared to "morning", 6% of responses)						
Afternoon	19	0.89	1.07*	1.10**	0.93	1.00
Evening	53	0.84	1.07*	1.07*	0.95	0.98
Late Night	22	0.79	1.03	1.11**	0.92	1.04
Question 9: During what time of day do you believe your studying is (or would be) most effective? (compared to "morning", 21% of responses)						
Afternoon	33	0.85	1.04	1.01	0.97	1.00
Evening	37	0.89	1.03	1.03	0.97	0.97
Late Night	8	0.84	1.03	1.01	1.01	1.01

Note: Frequency refers to the percentage of students who chose each response option.
*$p < .05$.
**$p < .01$.
***$p < .001$.