# SOC542 Statistical Methods in Sociology II

## Interactions

Thomas Davidson

Rutgers University

March 3, 2025

# Plan

- Introducing interactions
- Types of interactions and their interpretations
- Marginal effects

# Updates

- Grading Homework 2
- Project proposals due Friday (3/7) at 5pm
  - See last week's slides for details
  - Recommend meeting to discuss plan
  - Submit via email as PDF

# Introducing interactions

**What is an statistical interaction?**

- Consider the following population model:

$$y = \beta_0 + \beta_1 x + \beta_2 z + u$$

- The coefficients $\beta_1$ and $\beta_2$ measure the relationship between $x$ and $y$ and $z$ and $y$, respectively.
    - The interpretation of either coefficient requires that we hold the other constant.
- *But what if we expect the effect of $x$ to vary as a function of $z$?*

# Introducing interactions

### What is an statistical interaction?

- If we expect there to be an **interaction** between $x$ and $z$, such that the effect of $x$ on $y$ varies according to the level of $z$, we can add an **interaction term** into our model formula.

$$y = \beta_0 + \beta_1 x + \beta_2 z + \beta_3 x \cdot z + u$$

- $\beta_1$ and $\beta_2$ are now considered as the **main effects**.
- $\beta_3$ is the coefficient for the interaction term, representing the effect of $x$ *times* $z$.

# Introducing interactions

### A simple population model

```
N <- 1000
x <- rnorm(N)
z <- rnorm(N)
y <- 3*x + 2*z + -5*(x*z) + rnorm(N, 10)
```

# Introducing interactions

## Comparing models

| | (1) | (2) |
|---|---|---|
| (Intercept) | 10.029*** | 10.010*** |
| | (0.153) | (0.032) |
| x | 2.935*** | 2.981*** |
| | (0.157) | (0.033) |
| z | 2.099*** | 2.016*** |
| | (0.151) | (0.031) |
| x × z | | -4.980*** |
| | | (0.034) |
| Num.Obs. | 1000 | 1000 |
| R2 | 0.351 | 0.972 |
| R2 Adj. | 0.350 | 0.972 |
| F | 269.689 | 11455.353 |

# Introducing interactions

### Why use interactions?

- We can use interaction terms as a way to encode theoretical knowledge about the relationship between variables, which is often important for answering theoretical questions.
- For example, if we expect there to be differences in income related to intersectional inequalities involving sex and race, we can add an interaction term to a model:

$$Income = \beta_0 + \beta_1 Sex + \beta_2 Race + \beta_3 Sex \cdot Race + u$$

# Introducing interactions

### Why use interactions?

- Block et al. 2023 make the case that interactional frameworks are a necessary condition for making claims about intersectionality:
  - If $\beta_3 = 0 \rightarrow$ "no interaction effect, no intersectionality" (p.801)
- If $\beta_3 \neq 0$ then there are intersectional differences
- These differences are symmetric:
  - The effect of sex depends on race
  - The effect of race depends on sex

# Types of interactions

### Dummy-dummy

$$y = \beta_0 + \beta_1 Male + \beta_2 Degree + \beta_3 Male \cdot Degree + u$$

# Types of interactions

## Dummy-dummy

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| (Intercept) | 19.962*** | 17.501*** | 12.128*** | 13.267*** |
| | (1.065) | (0.915) | (1.136) | (1.245) |
| Male | 10.600*** | | 11.113*** | 8.757*** |
| | (1.546) | | (1.443) | (1.791) |
| Degree | | 21.141*** | 21.431*** | 18.315*** |
| | | (1.538) | (1.506) | (2.060) |
| Male:Degree | | | | 6.678* |
| | | | | (3.015) |
| Num.Obs. | 1358 | 1358 | 1358 | 1358 |
| R2 | 0.034 | 0.122 | 0.159 | 0.162 |
| R2 Adj. | 0.033 | 0.122 | 0.158 | 0.160 |
| F | 47.019 | 188.947 | 128.195 | 87.344 |

# Types of interactions

## Dummy-dummy

$$y = \beta_0 + \beta_1 Male + \beta_2 Degree + \beta_3 Male \cdot Degree + u$$

- Female and people without a college degree are the reference categories.
- $\beta_1$ and $\beta_2$ represent the main effects of sex and degree on the outcome, but they only tell a partial story unless $\beta_3 = 0$.
- The coefficient $\beta_3$ represents the expected difference in the effect of degree for men versus women.[1]
- If $\beta_3 \neq 0$, the expected income for a male with a degree is $\beta_0 + \beta_1 + \beta_2 + \beta_3$. The same quantity for a female with a degree is $\beta_0 + \beta_2$.

[1] Note the symmetrical interpretation here: the difference in the effect of sex for college degree versus non-college degree. See McElreath 8.2 for further discussion.

# Types of interactions

## Dummy-dummy: Evaluating intersectional claims

$$y = \beta_0 + \beta_1 Female + \beta_2 Black + \beta_3 Female \cdot Black + u$$

| | (1) | (2) |
|---|---|---|
| (Intercept) | 32.950*** | 34.044*** |
| Female | -10.236*** | -12.357*** |
| Black | -8.764*** | -15.662*** |
| Female:Black | | 11.600** |
| Num.Obs. | 1191 | 1191 |
| R2 | 0.046 | 0.051 |
| R2 Adj. | 0.044 | 0.049 |
| F | 28.386 | 21.349 |

* p <0.05, ** p <0.01, *** p <0.001

# Types of interactions

### Dummy-dummy: Evaluating intersectional claims

- Since $\beta_3 \neq 0$ we can reject the null hypothesis of no interesectionality
- But this means that the other coefficients do not tell us the "separate, unconditional, independent, or average effects of gender and race" (Block et al. 2023)

# Types of interactions

### Dummy-dummy: Evaluating intersectional claims

The effect of gender depends on race:

$$\frac{\Delta Income}{\Delta Female} = \beta_1 + \beta_3 \cdot Black$$

The effect of race depends on gender:

$$\frac{\Delta Income}{\Delta Black} = \beta_2 + \beta_3 \cdot Female$$

# Types of interactions

### Reformulating the model

- Block et al. 2023 show how we could specify an equivalent, alternative model:
- Assuming White Female is the reference category, we could write this as:

$$y = \beta_0 + \gamma_1 BlackFemale + \gamma_2 BlackMale + \gamma_3 WhiteFemale + u$$

# Comparing frameworks

| | (1) | (2) |
|---|---|---|
| (Intercept) | 34.044*** | 34.044*** |
| Female | -12.357*** | |
| Black | -15.662*** | |
| Female:Black | 11.600** | |
| Black Female | | -16.420*** |
| Black Male | | -15.662*** |
| White Female | | -12.357*** |
| Num.Obs. | 1191 | 1191 |
| R2 | 0.051 | 0.051 |
| R2 Adj. | 0.049 | 0.049 |
| F | 21.349 | 21.349 |

* p <0.05, ** p <0.01, *** p <0.001

# Reformulating interactions

Table 1. Some Different Stories about the Impact of Gender and Race and the Predicted Model Parameters

| | Standard Interaction Model | Alternative Interaction Model |
|---|---|---|
| No intersectionality | $\beta_3 = 0$ | $\gamma_3 - \gamma_1 - \gamma_2 = 0$ |
| Only gender matters | $\beta_1 \neq 0, \beta_2 = 0$ | $\gamma_1 \neq 0, \gamma_2 = 0$ |
| Only race matters | $\beta_1 = 0, \beta_2 \neq 0$ | $\gamma_1 = 0, \gamma_2 \neq 0$ |
| Gender and race both have separate effects | $\beta_1 \neq 0, \beta_2 \neq 0$ | $\gamma_1 \neq 0, \gamma_2 \neq 0$ |
| Intersectionality | $\beta_3 \neq 0$ | $\gamma_3 - \gamma_1 - \gamma_2 \neq 0$ |
| Gender matters, but differently, for both White people and Black people | $\beta_1 \neq 0, \beta_1 + \beta_3 \neq 0,$ | $\gamma_1 \neq 0, \gamma_3 - \gamma_2 \neq 0$ |
| Race matters, but differently, for both men and women | $\beta_2 \neq 0, \beta_2 + \beta_3 \neq 0$ | $\gamma_2 \neq 0, \gamma_3 - \gamma_1 \neq 0$ |
| Gender matters for White people but not Black people | $\beta_1 \neq 0, \beta_1 + \beta_3 = 0,$ | $\gamma_1 \neq 0, \gamma_3 - \gamma_2 = 0$ |
| Race matters for men but not women | $\beta_2 \neq 0, \beta_2 + \beta_3 = 0$ | $\gamma_2 \neq 0, \gamma_3 - \gamma_1 = 0$ |
| Gender matters for Black people but not White people | $\beta_1 + \beta_3 \neq 0, \beta_1 = 0$ | $\gamma_3 - \gamma_2 \neq 0, \gamma_1 = 0$ |
| Race matters for women but not men | $\beta_2 + \beta_3 \neq 0, \beta_2 = 0$ | $\gamma_3 - \gamma_1 \neq 0, \gamma_2 = 0$ |

# Reformulating interactions

- Testing intersectional claims requires postestimation calculations to calculate different quantities:
  - Intersectional effect for gender and race
  - Effect of being female among White people
  - Effect of being female among Black people
  - Effect of being Black among men
  - Effect of being Black among women

# Types of interactions

### Continuous-dummy

$$y = \beta_0 + \beta_1 Age + \beta_2 Sex + \beta_3 Age \cdot Sex + u$$

# Types of interactions

## Continuous-dummy

| | (1) | (2) |
|---|---|---|
| (Intercept) | 4.489 | 7.431* |
| | (2.553) | (3.394) |
| Age | 0.353*** | 0.286*** |
| | (0.053) | (0.074) |
| Male | 10.158*** | 3.941 |
| | (1.523) | (4.967) |
| Age:Male | | 0.140 |
| | | (0.106) |
| Num.Obs. | 1358 | 1358 |
| R2 | 0.064 | 0.065 |
| R2 Adj. | 0.063 | 0.063 |
| F | 46.342 | 31.488 |

# Types of interactions

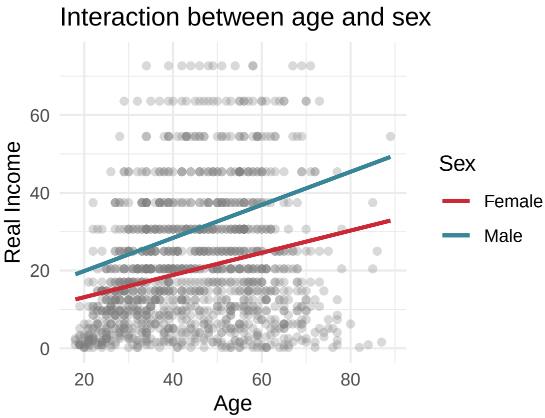
### Continuous-dummy

$$y = \beta_0 + \beta_1 Age + \beta_2 Sex + \beta_3 Age \cdot Sex + u$$

- The coefficients $\beta_1$ and $\beta_2$ represent the main effects of age and sex on income.
- For females, $\beta_1$ represents the relationship between age and income. For males, the relationship is $\beta_1 + \beta_3$.
  - Thus, the interaction term allows the *slope* to vary according to sex.

# Types of interactions

## Continuous-dummy




Income truncated to 75k to emphasize trends.

# Types of interactions

**Continuous-continuous**

$$y = \beta_0 + \beta_1 Age + \beta_2 HoursWorked + \beta_3 Age \cdot HoursWorked + u$$

# Types of interactions

## Continuous-continuous

| | (1) | (2) |
|---|---|---|
| (Intercept) | -13.689*** | 11.030 |
| | (3.799) | (7.796) |
| Age | 0.424*** | -0.106 |
| | (0.059) | (0.157) |
| Hours Worked | 0.498*** | -0.123 |
| | (0.057) | (0.180) |
| Age:Hours Worked | | 0.014*** |
| | | (0.004) |
| Num.Obs. | 1172 | 1172 |
| R2 | 0.091 | 0.101 |
| R2 Adj. | 0.090 | 0.099 |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

# Types of interactions

### Continuous-continuous

$$y = \beta_0 + \beta_1 Age + \beta_2 WorkHrs + \beta_3 Age \cdot WorkHrs + u$$

- The intercept no longer has a meaningful interpretation (income when age and work hours equal zero).
  - GHV 12.2 discuss standardization to make intercepts more interpretable in such contexts.
- $\beta_1$ and $\beta_2$ represent the main effects of age and work hours.
- The interaction term $\beta_3$ captures how the effect of work hours on income varies as a function of age.

## Types of interactions

### Continuous-continuous

- The effect of work hours on income is now also a function of age:

$$\frac{\Delta y}{\Delta_{WorkHrs}} = \beta_2 + \beta_3 Age$$

- Similarly,

$$\frac{\Delta y}{\Delta_{Age}} = \beta_1 + \beta_3 WorkHrs$$

# Types of interactions

### Continuous-continuous

- If Age changes by $\Delta$Age and WorkHrs by $\Delta$WorkHrs, the expected change in $y$ is:

$$\Delta y = (\beta_1 + \beta_3 WorkH)\Delta Age + (\beta_2 + \beta_3 Age)\Delta WorkH + \beta_3 \Delta Age \cdot \Delta WorkH$$

- The coefficient $\beta_3$ represents the effect of a unit increase in age *and* work hours, beyond the sum of the individual effects of unit increases alone.

# Types of interactions

## Dummy-categorical

| | (1) | (2) |
|---|---|---|
| (Intercept) | 22.657*** | 21.686*** |
| Male | 10.354*** | 12.357*** |
| Black | -8.753*** | -4.062 |
| Other race | -9.069*** | -8.545* |
| Male:Black | | -11.600** |
| Male:Other race | | -1.164 |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

# Types of interactions

### Dummy-categorical

$$y = \beta_0 + \beta_1 Male + \beta_2 Black + \beta_3 Other + \beta_4 Black \cdot Male + \beta_5 Other \cdot Male + u$$

- There is a separate coefficient for the interaction between the dummy variable and each of the categories, with the exception of the reference group.
- The interpretation is the same as the dummy-dummy model.

# Types of interactions

## Continuous-categorical

$$y = \beta_0 + \beta_1 Age + \beta_2 Black + \beta_3 Other + \beta_4 Black \cdot Age + \beta_5 Other \cdot Age + u$$

| | (1) | (2) |
|---|---|---|
| (Intercept) | 12.391*** | 11.405*** |
| Age | 0.334*** | 0.356*** |
| Black | -8.403*** | -1.744 |
| Other race | -6.901** | -8.009 |
| Age:Black | | -0.158 |
| Age:Other race | | 0.030 |

* p <0.05, ** p <0.01, *** p <0.001

# Types of interactions

## Categorical-categorical

| | (1) | (2) |
|---|---|---|
| (Intercept) | 23.151*** | 22.096*** |
| Black | -8.627*** | -4.916 |
| Other race | -8.231*** | -7.528 |
| Inspired Word | 4.485* | |
| Ancient Book | 8.583*** | |
| White:Inspired Word | | 5.815* |
| Black:Inspired Word | | 2.259 |
| Other race:Inspired Word | | 2.473 |
| White:Ancient Book | | 10.015*** |
| Black:Ancient Book | | -3.627 |
| Other race:Ancient Book | | 14.067* |

* p <0.05, ** p <0.01, *** p <0.001

# Types of interactions

### Three-way interactions

| | (1) | (2) |
|---|---|---|
| (Intercept) | 14.287*** | 14.980*** |
| Male | 11.310*** | 8.652*** |
| Black | -7.013*** | -4.260* |
| Other race | -6.281** | -6.289** |
| Degree | 20.656*** | 17.779*** |
| Male:White:Degree | | 9.407** |
| Male:Black:Degree | | -14.010* |
| Male:Other race:Degree | | 8.542 |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

# Types of interactions

### Interpreting interactions

- Interactions terms make models more challenging to interpret.
  - Like polynomial regression, the effect of a single predictor is represented by more than one coefficient (e.g. $y = \beta_0 + \beta_1 x + \beta_2 z + \beta_3 x \cdot z + u$).
- Three-way and more complex interactions are even more difficult to interpret and should be avoided unless there are strong theoretical reasons to use them.

# Marginal effects

### Definitions

- A **marginal effect** is the relationship between change in single predictor and the dependent variable while *holding other variables constant*.
- The **average marginal effect (AME)** is the *average* change in the outcome as a function of a unit change in $x$.
  - Coefficients in a standard OLS model represent average marginal effects.
- This quantity becomes more complicated to calculate when interaction terms are included, since the effect of a change in $x$ now depends on multiple parameters.

# Marginal effects

### Computing marginal effects

- Frequentist marginal effects computed by calculating *partial derivatives* and variance approximations are used to construct confidence intervals.

$$ME(x_i) = \frac{\Delta y}{\Delta x_i}$$

- We can use the `margins` package in R to do this.[2]
- Bayesian marginal effects can be calculated by sampling from the posterior distribution.

[2] See documentation for the margins package for further details.

# Marginal effects

## Marginal effects and OLS regression

| | (1) |
|---|---|
| (Intercept) | 8.213** |
| | (2.671) |
| Male | 9.981*** |
| | (1.519) |
| Age | 0.320*** |
| | (0.053) |
| Black | -7.495*** |
| | (2.068) |
| Other race | -7.456** |
| | (2.348) |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

# Marginal effects

### Marginal effects using `margins`

Note how the average marginal effects are equal to the OLS coefficients.

```r
library(margins)
me <- margins(m)
summary(me)
```

```
##     factor     AME     SE       z      p    lower   upper
##        age  0.3196 0.0533  6.0018 0.0000   0.2153  0.4240
##  raceBlack -7.4951 2.0683 -3.6238 0.0003 -11.5489 -3.4413
##  raceOther -7.4564 2.3484 -3.1751 0.0015 -12.0592 -2.8536
##        sex  9.9807 1.5185  6.5728 0.0000   7.0045 12.9569
```

# Marginal effects

### Marginal effects using `marginaleffects`

Note how the average marginal effects are equal to the OLS coefficients.

```
library(marginaleffects)
me <- avg_slopes(m)
print(me)
```

```
##
##  Term        Contrast Estimate Std. Error     z Pr(>|z|)    S   2.5 %
##  age   dY/dX              0.32     0.0533  6.00   <0.001 28.9   0.215
##  race Black - White      -7.50     2.0683 -3.62   <0.001 11.7 -11.549
##  race Other - White      -7.46     2.3484 -3.18   0.0015  9.4 -12.059
##  sex   1 - 0              9.98     1.5185  6.57   <0.001 34.2   7.005
##
## Columns: term, contrast, estimate, std.error, statistic, p.value, s.
## Type:  response
```

# Marginal effects

## Marginal effects with non-linear variables

| | (1) | (2) |
|---|---|---|
| (Intercept) | 8.213** | -36.429*** |
| Male | 9.981*** | 9.999*** |
| Age | 0.320*** | 2.460*** |
| Black | -7.495*** | -7.546*** |
| Other race | -7.456** | -7.802*** |
| Age^2 | | -0.023*** |

# Marginal effects

### Marginal effects with non-linear variables

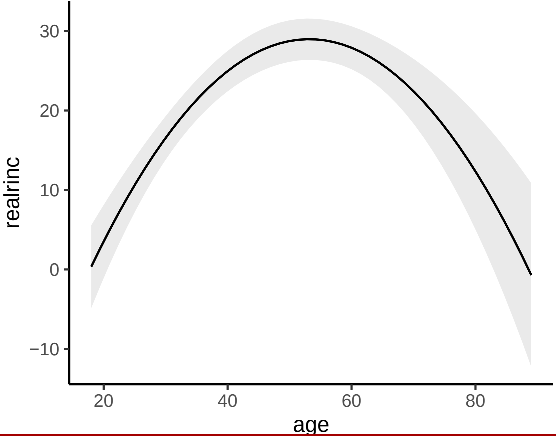
The margins commands are the same as above. Note how the AME now represents the total effect of age across the two parameters. There is no separate marginal effect for age squared.

```
##
##  Term       Contrast Estimate Std. Error     z Pr(>|z|)    S   2.5 %
##  age  dY/dX             0.403     0.0537  7.51   <0.001 44.0   0.298
##  race Black - White   -7.546     2.0322 -3.71   <0.001 12.3 -11.529
##  race Other - White   -7.802     2.3080 -3.38   <0.001 10.4 -12.325
##  sex  1 - 0             9.999     1.4920  6.70   <0.001 35.5   7.074
##
## Columns: term, contrast, estimate, std.error, statistic, p.value, s.
## Type:  response
```

# Marginal effects

## Marginal effects with non-linear variables

```
plot_predictions(m2, condition = "age") +
    theme_classic()
```



# Marginal effects

## Marginal effects with interactions

```r
m <- lm(realrinc ~ sex + age + I(age^2) + race + sex:race + sex:age,
        data = gss)
```

| | (1) |
|---|---|
| (Intercept) | -35.202*** |
| Male | 3.993 |
| Age | 2.441*** |
| Age^2 | -0.024*** |
| Black | -2.570 |
| Other race | -6.678* |
| Male:Black | -12.661** |
| Male:Other race | -2.375 |
| Male:Age | 0.187 |

# Marginal effects

## Marginal effects with interactions

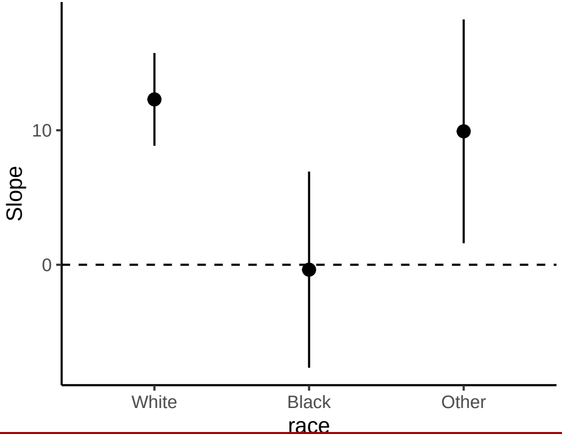
In this case, we can isolate the average marginal effect of each predictor.

```
##
##  Term      Contrast Estimate Std. Error     z Pr(>|z|)    S   2.5 %
##  age  dY/dX             0.41     0.0535  7.66   <0.001 45.6   0.305
##  race Black - White    -8.57     2.0453 -4.19   <0.001 15.1 -12.582
##  race Other - White    -7.80     2.3121 -3.38   <0.001 10.4 -12.336
##  sex  1 - 0             9.89     1.4865  6.65   <0.001 35.0   6.974
##
## Columns: term, contrast, estimate, std.error, statistic, p.value, s.
## Type:  response
```

# Marginal effects

## Conditional marginal effects (effect of sex by race)

```
plot_slopes(m, variables = "sex", condition = "race") +
    theme_classic() + geom_hline(yintercept=0, linetype = "dashed")
```



# Marginal effects

### Computing marginal effects for Bayesian models

- Unlike frequentist marginal effects, there is no need for additional calculus.
- Marginal effects can be computed directly from the posterior distribution.

# Marginal effects

| | OLS | Bayesian |
|---|---|---|
| Male | 3.993 | 3.892 |
| | [-6.005, 13.992] | [-6.085, 13.829] |
| Age | 2.441 | 2.437 |
| | [1.834, 3.049] | [1.812, 2.991] |
| Age^2 | -0.024 | -0.024 |
| | [-0.030, -0.017] | [-0.030, -0.017] |
| Black | -2.570 | -2.365 |
| | [-7.747, 2.607] | [-7.382, 2.643] |
| Other race | -6.678 | -6.436 |
| | [-13.239, -0.117] | [-12.844, -0.014] |
| Male:Black | -12.661 | -12.953 |
| | [-20.751, -4.570] | [-20.752, -4.496] |
| Male:Other race | -2.375 | -2.409 |
| | [-11.414, 6.664] | [-11.482, 6.153] |
| Male:Age | 0.187 | 0.187 |
| | [-0.019, 0.392] | [-0.020, 0.395] |

# Marginal effects

### Bayesian marginal effects

Fortunately for us, the `margins` and `marginaleffects` packages also work for Bayesian models.

```r
avg_slopes(m.b)
```

```
##
##  Term      Contrast Estimate   2.5 % 97.5 %
##  age  dY/dX            0.409   0.309  0.514
##  race Black - White   -8.481 -12.640 -4.734
##  race Other - White   -7.768 -11.816 -3.063
##  sex  1 - 0            9.843   7.152 12.841
##
## Columns: term, contrast, estimate, conf.low, conf.high
## Type:  response
```

# Marginal effects

### Bayesian marginal effects

To obtain the information used in these calculations, we can compute the *expected value* of the outcome at different levels of predictors using `epred_draws`.

```
data.range <- expand_grid(sex = c(0,1),
                          race = c("Black", "White", "Other"),
                          age = 18:75)

tidy_epred <- m.b %>% epred_draws(newdata = data.range)
```

# Marginal effects

### Bayesian marginal effects

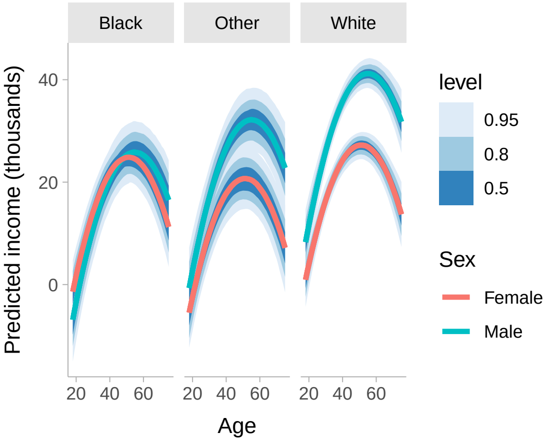
Like everything else we obtain from a Bayesian model, these marginal effects have a posterior distribution.

```
tail(tidy_epred %>% select(sex, race, age, .epred))
```

```
## # A tibble: 6 x 5
## # Groups:   sex, race, age, .row [1]
##    .row   sex race    age .epred
##   <int> <dbl> <chr> <int>  <dbl>
## 1   348     1 Other    75   14.8
## 2   348     1 Other    75   26.3
## 3   348     1 Other    75   15.6
## 4   348     1 Other    75   19.5
## 5   348     1 Other    75   40.1
## 6   348     1 Other    75   27.6
```

# Marginal effects

## Bayesian marginal effects



# Marginal effects

### Marginal effects and generalized linear models

- For generalized linear models (GLMs)—which will be our main focus after spring break—coefficients often do not have clear interpretations on the outcome scale so marginal effects even more important for interpretation.

## Next week

### Topic

- Missing data and imputation
- Model specification and robustness

## Lab

- Specifying and interpreting interaction terms