interaction_effects

November 17, 2020

GLMs with Interaction Effects in R

1 1. Interaction Effects

- Often called "moderation", or statistical moderation effects
- One variable "behaves" differently depending on the level another
- E.g. cigarette smoking is an *increasingly strong* predictor of lung cancer as age increases

2 2. Logistic Regression with Interaction Effect

- Using `titanic` data
- Predicting survival using age and sex

In [16]:
```
library(broom)
library(ggplot2)

titanic_df <- read.csv("data/titanic_subset.csv")

head(titanic_df)
```

A data.frame: 6 Œ 12

	passenger_id <int>	survived <int>	pclass <int>	name <chr>
1	1	0	3	Braund, Mr. Owen Harris
2	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs T
3	3	1	3	Heikkinen, Miss. Laina
4	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)
5	5	0	3	Allen, Mr. William Henry
6	6	0	3	Moran, Mr. James

2.1 2.1 Fitting Model with Interaction Term

In [18]:
```
fm3 <- glm(survived ~ sex + age + age*sex, titanic_df, family = binomial(link = "logi

tidy(fm3)
```

A tibble: 4 Œ 5

term <chr>	estimate <dbl>	std.error <dbl>	statistic <dbl>	p.value <dbl>
(Intercept)	0.59380093	0.31032443	1.913484	0.055686078
sexmale	-1.31775110	0.40842445	-3.226426	0.001253468
age	0.01970198	0.01057286	1.863449	0.062399191
sexmale:age	-0.04111845	0.01355105	-3.034336	0.002410658

2.1.1 2.1.1 Plotting the Effects

```
In [20]: ggplot(titanic_df, aes(x = age, y = as.numeric(survived), color = sex)) +
             stat_smooth(method = "glm", alpha = 0.2, size = 2, aes(fill = sex)) +
             geom_point(position=position_jitter(height = 0.03, width = 0)) +
             xlab("Age") +
             ylab("Pr(survived)")
```

```
`geom_smooth()` using formula 'y ~ x'

Warning message:
Removed 177 rows containing non-finite values (stat_smooth).
Warning message:
Removed 177 rows containing missing values (geom_point).
```

2.1.2 2.1.2 Predictions using Fitted Model

```
In [24]: fm2 <- glm(survived ~ age + sex + age*sex, titanic_df, family = binomial(link = "logi

         predict(fm2, newdata=data.frame(age = 50, sex = "female"), type = "response")
```

1: 0.829048659321465