Crash course on GLMs

HIERARCHICAL AND MIXED EFFECTS MODELS IN R

Richard Erickson
Data Scientist

datacamp

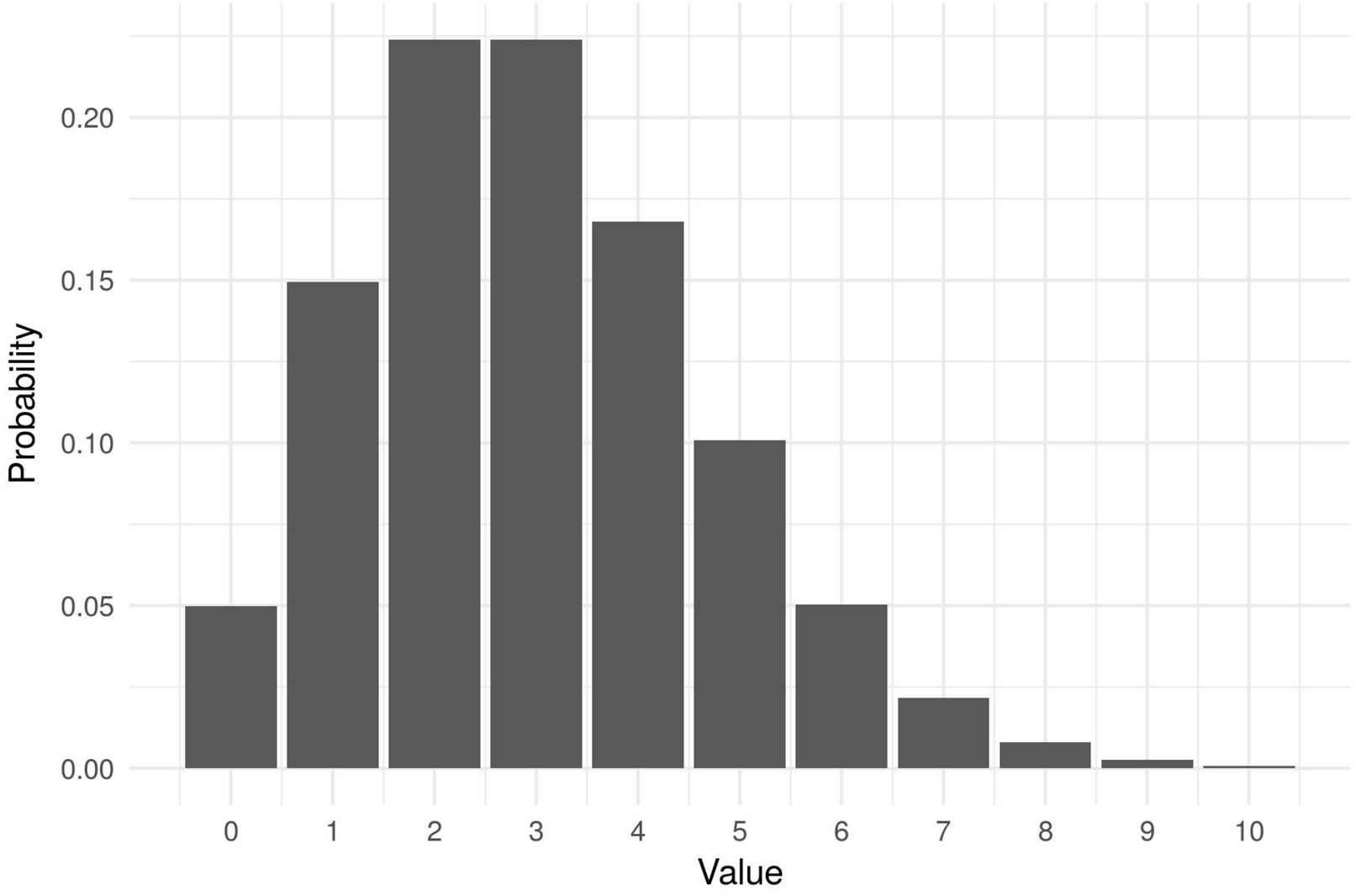
Assumption of normality

- Residuals are "normal"
- Transformation
- "The arcsine is asinine:..."
- Alternative distributions

R syntax for GLM

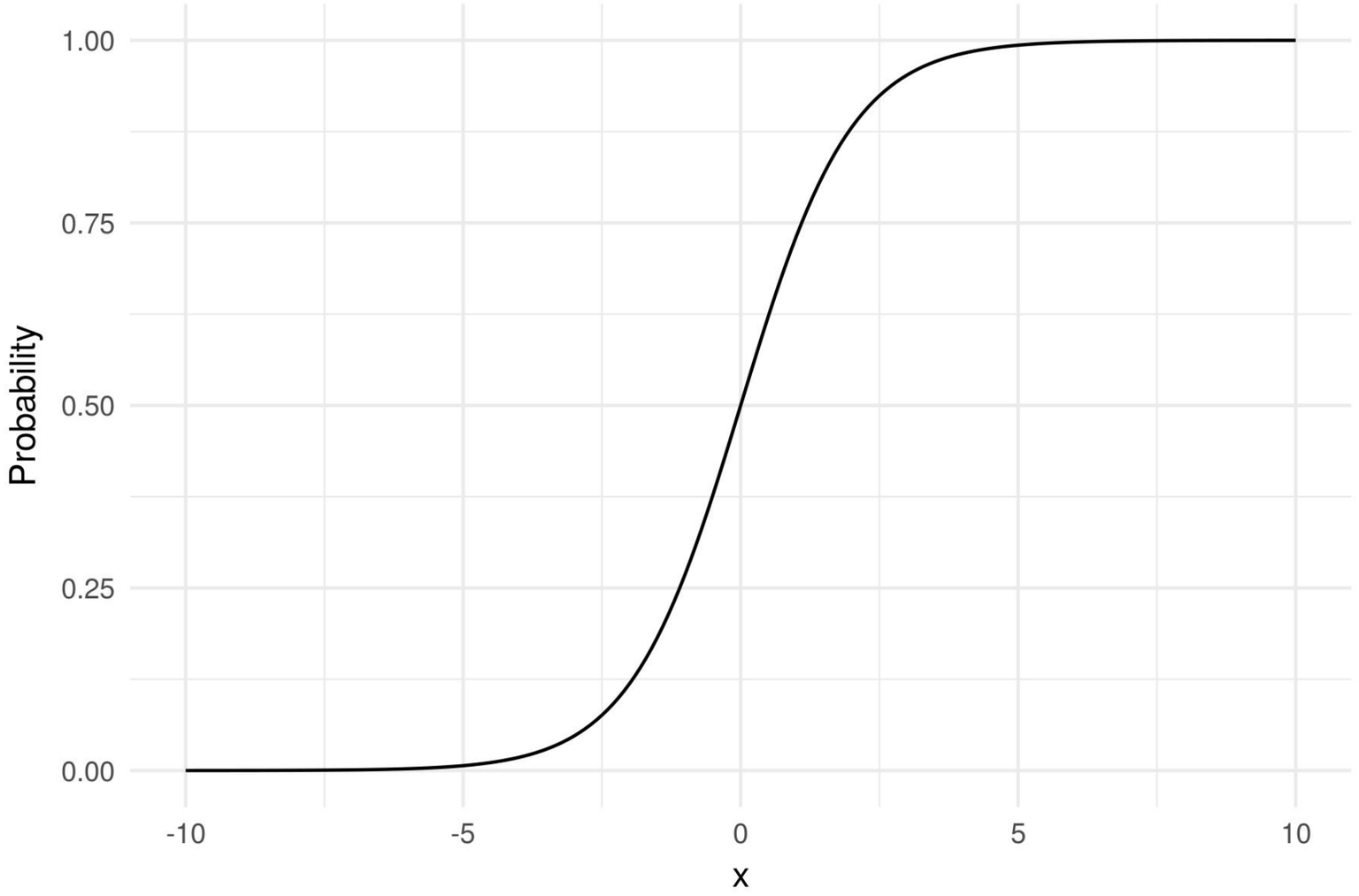
```
glm(y ~ x, family = "gaussian")
```

```
?family
```


0.20
0.15
0.10
0.05
0.00
Probability
0
1
2
3
4
5
6
7
8
9
10
Value

Example with Poisson regression

```
glm(y ~ x, family = "poisson")
```


1.00
0.75
0.50
0.25
0.00
Probability
-10
-5
0
5
10
x

Example with logistic regression

```
# binary: y = 0 or 1
glm(y ~ x, family = "binomial")
# Wilkinson-Rogers: cbind(success, failure)
glm(cbind(success, failure) ~ x, family = "binomial")
# Weighted format: y = 0.3, weights = 10
glm(y ~ x, weights = weights, family = "binomial")
```

Let's practice!

HIERARCHICAL AND MIXED EFFECTS MODELS IN R

Binomial data

HIERARCHICAL AND MIXED EFFECTS MODELS IN R

Richard Erickson
Data Scientist

datacamp

Examples of binomial data

- Coin/toss
- Yes/No
- Dead/alive
- Behavior
- Choice
- Study result

Binomial data with glmer

```
glmer(y ~ x + (1 | group), family = "error term")
```

Dose-response case study

- Study had increasing dose
- Repeated in triplicate
- Requires `glmer()`, not `glm()`

Internet purchase

- "Purchases" or "Pass"
- "ranking" of product 0 to 20
- "friend recommendation"
- Tracked sales with 4 focal groups in different "cities"
- Do friend recommendations help?

Odds-ratio

- Regression coefficients hard to explain
- Odds-ratios sometimes easier
- If Group A has odds-ratio of 2.0 $\times$, 2:1 A will do something compared to the other group
- Extract by exponentiation logistic coefficients
- `exp(coef(modelOut))`

Let's practice!

HIERARCHICAL AND MIXED EFFECTS MODELS IN R

Count data

HIERARCHICAL AND MIXED EFFECTS MODELS IN R

Richard Erickson
Data Scientist

datacamp

Examples of count data

- Events at a rate of time
- Events per area
- Differs from binomial because no explicit upper limit

Alternative to Chi-square test

- Chi-square test used to compare binned counts
- Poisson glm can be an alternative

R Syntax for Poisson regression with `glmer`

```
glm(y ~ x, family = 'poisson')


glmer(y ~ x + (1 | group), family = 'poisson')
```

Marketing click through case study

- Redesigned website
- Clicks on different webpages
- Focus groups looked at old and new
- Marketer was using a linear mixed model
- Demonstrate how a generalized model

Chlamydia by age-group and county data

- **State of IL**
- By county and age group
- Important for public health
- Public policy application
- Marketing/drug research

Let's apply Poisson regression!

HIERARCHICAL AND MIXED EFFECTS MODELS IN R