

Bayesian data analysis 4

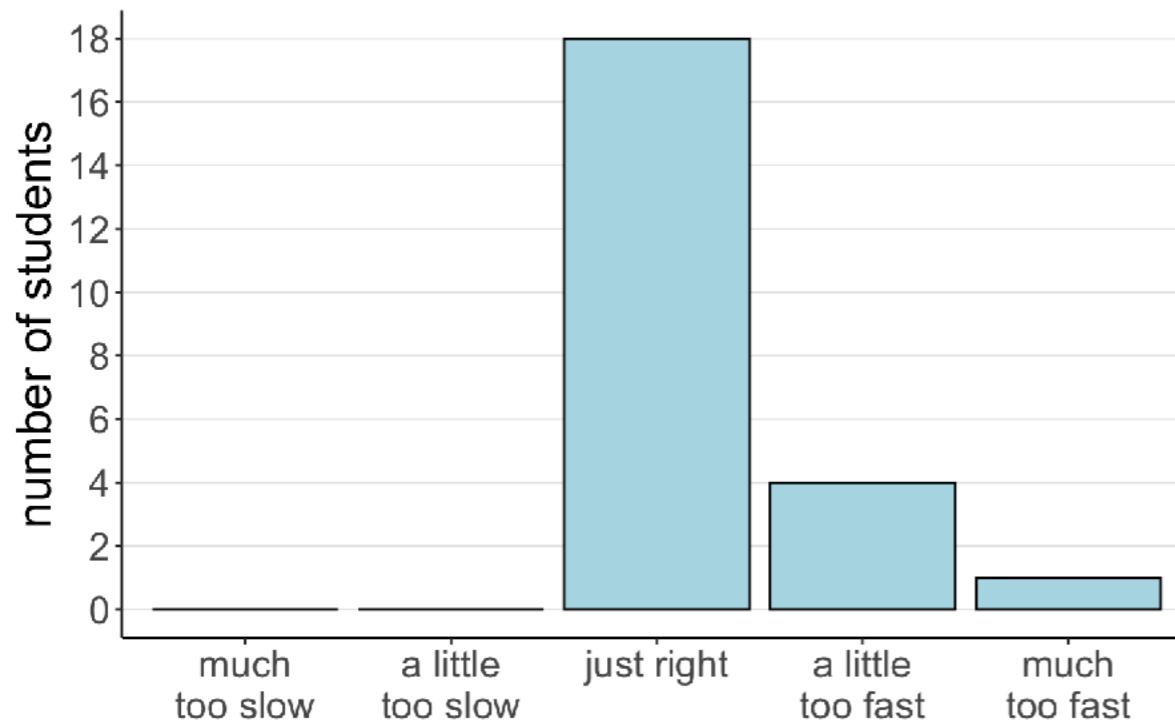
<u>P-VALUE</u>	<u>INTERPRETATION</u>
0.001	HIGHLY SIGNIFICANT
0.01	HIGHLY SIGNIFICANT
0.02	HIGHLY SIGNIFICANT
0.03	HIGHLY SIGNIFICANT
0.04	SIGNIFICANT
0.049	SIGNIFICANT
0.050	OH CRAP. REDO CALCULATIONS.
0.051	ON THE EDGE OF SIGNIFICANCE
0.06	ON THE EDGE OF SIGNIFICANCE
0.07	HIGHLY SUGGESTIVE,
0.08	SIGNIFICANT AT THE P<0.10 LEVEL
0.09	SIGNIFICANT AT THE P<0.10 LEVEL
0.099	HEY, LOOK AT THIS INTERESTING
≥0.1	SUBGROUP ANALYSIS

Logistics

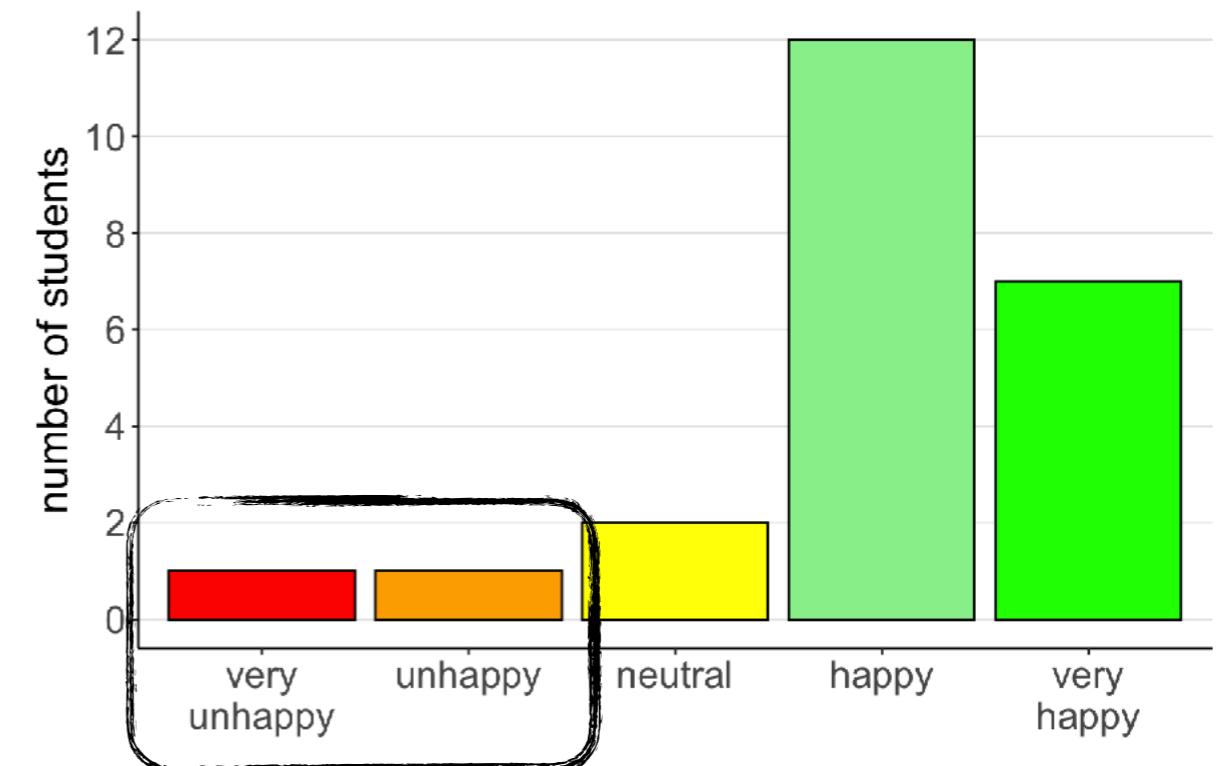
Feedback

Feedback

How was the pace of today's class?



How happy were you with today's class overall?



maybe we can check
in at office hour?

Feedback

I found this lecture to be quite clear and helpful. The visualizations were especially helpful for solidifying the quite abstract concepts covered.

I think the session was too ambitious. Maybe the class in general. We try to cover too much material in little time. This class I couldn't follow anything.

Homework

Homework



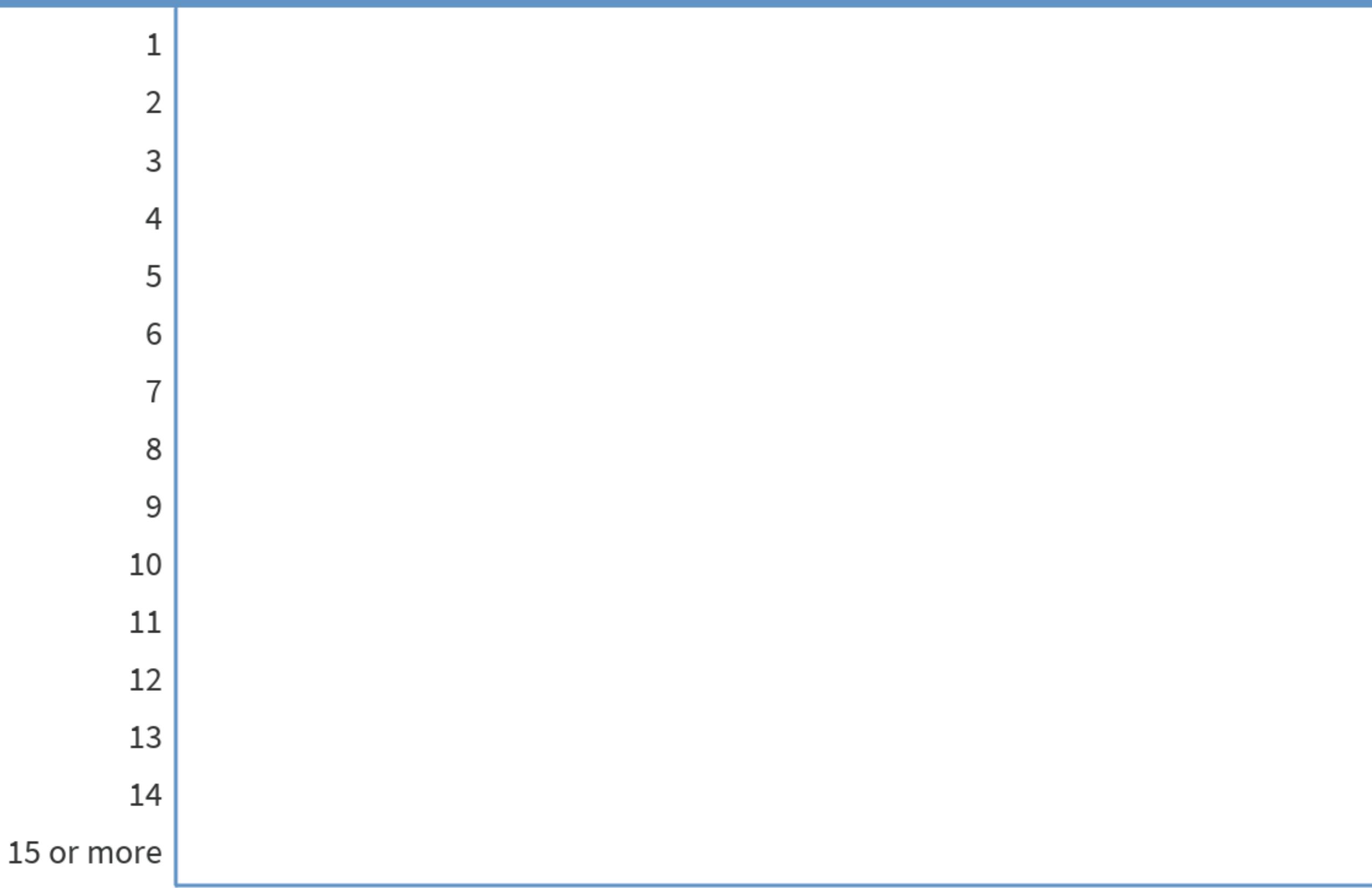
Jae-Young Son
@psychNerdJae

the problem with upper-level stats classes is that when a homework question seems unsolvable, it's never clear whether it's intentional/pedagogical, whether it's an error in how the assignment was written, or whether you as a student are failing to get something right

3:33 PM · Feb 29, 2020 · [Twitter Web App](#)

**homework 7 will be released at the end
of today (definitely before midnight)**

How many hours did it take you to complete Homework 6?

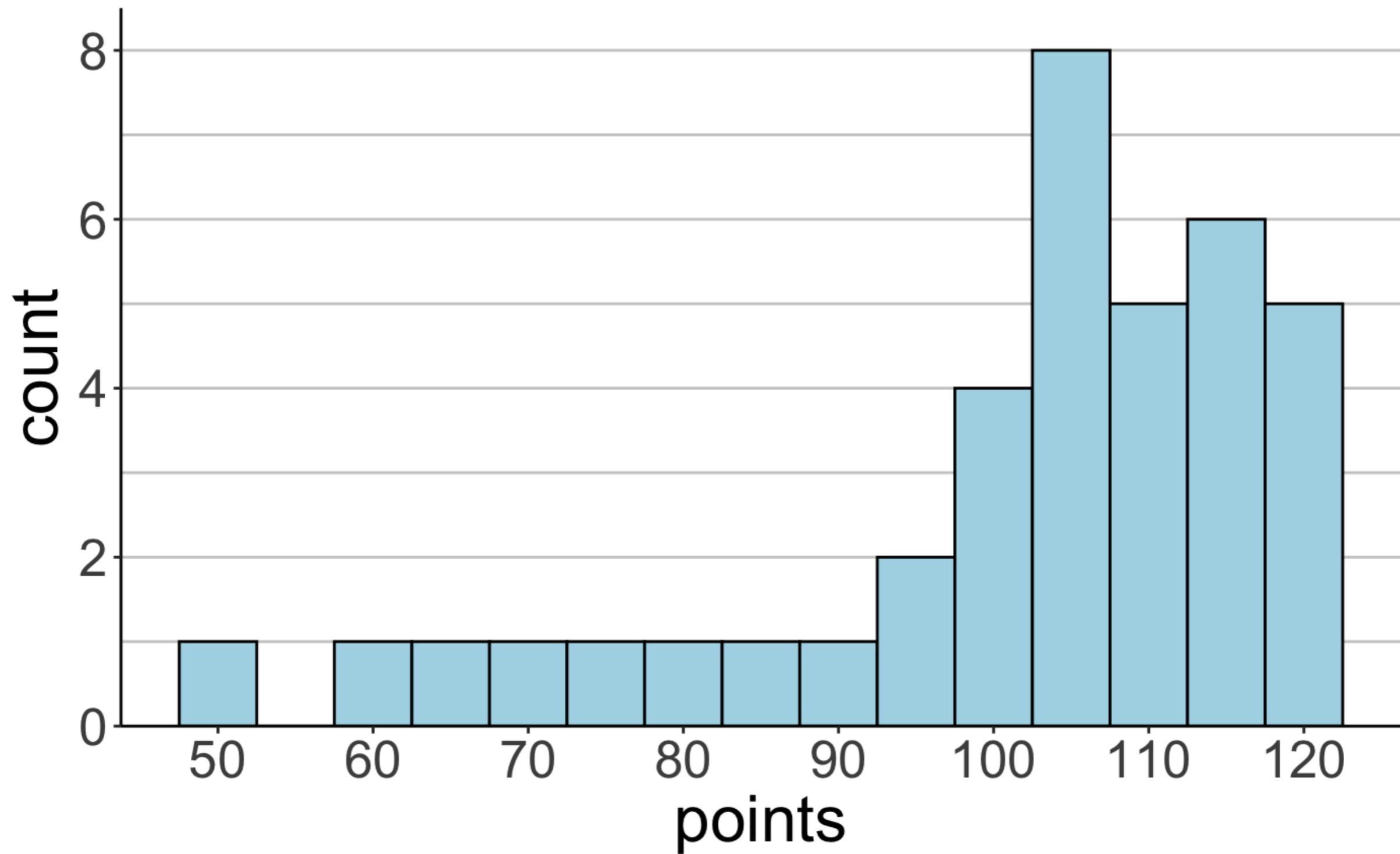


Homework

- we won't release the grades for HW5 to HW7 until after HW7 is due
- we will still release the solutions once we've graded the homeworks

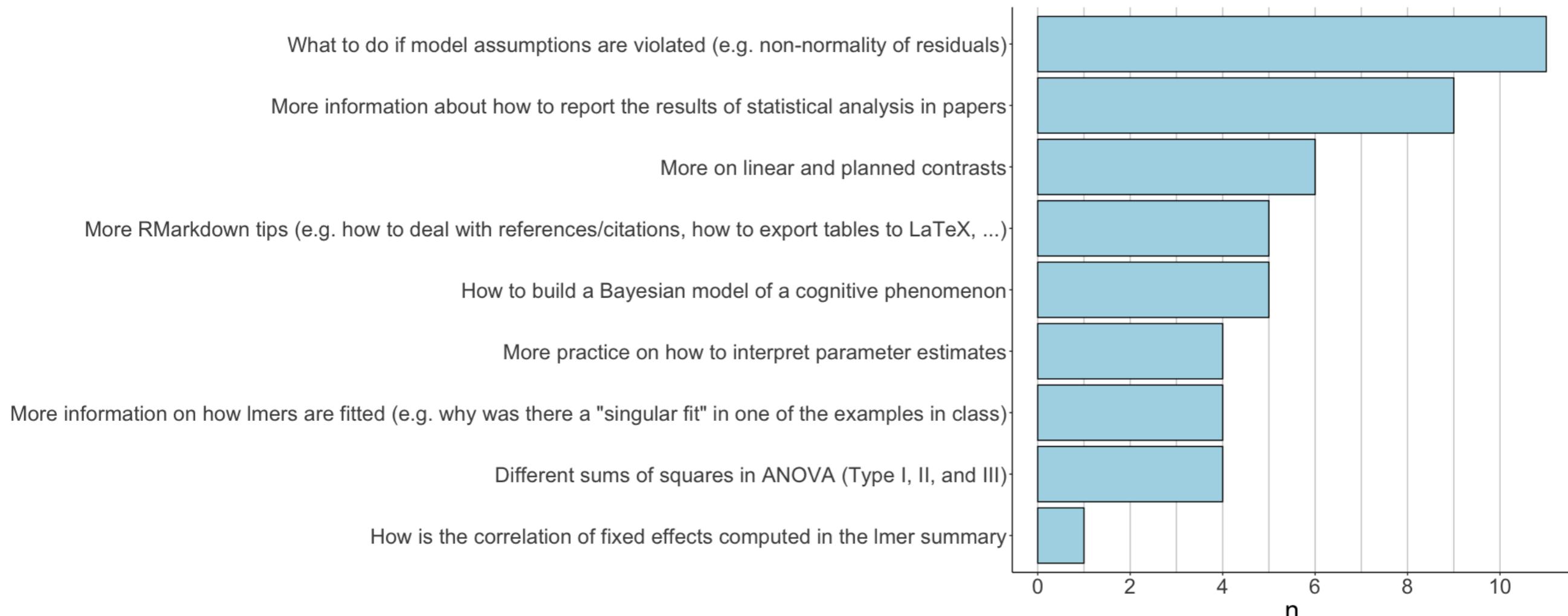
Midterm

Midterm point distribution

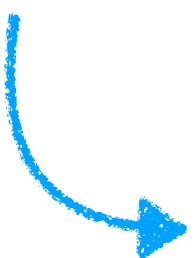


Class on Monday

Class on Monday



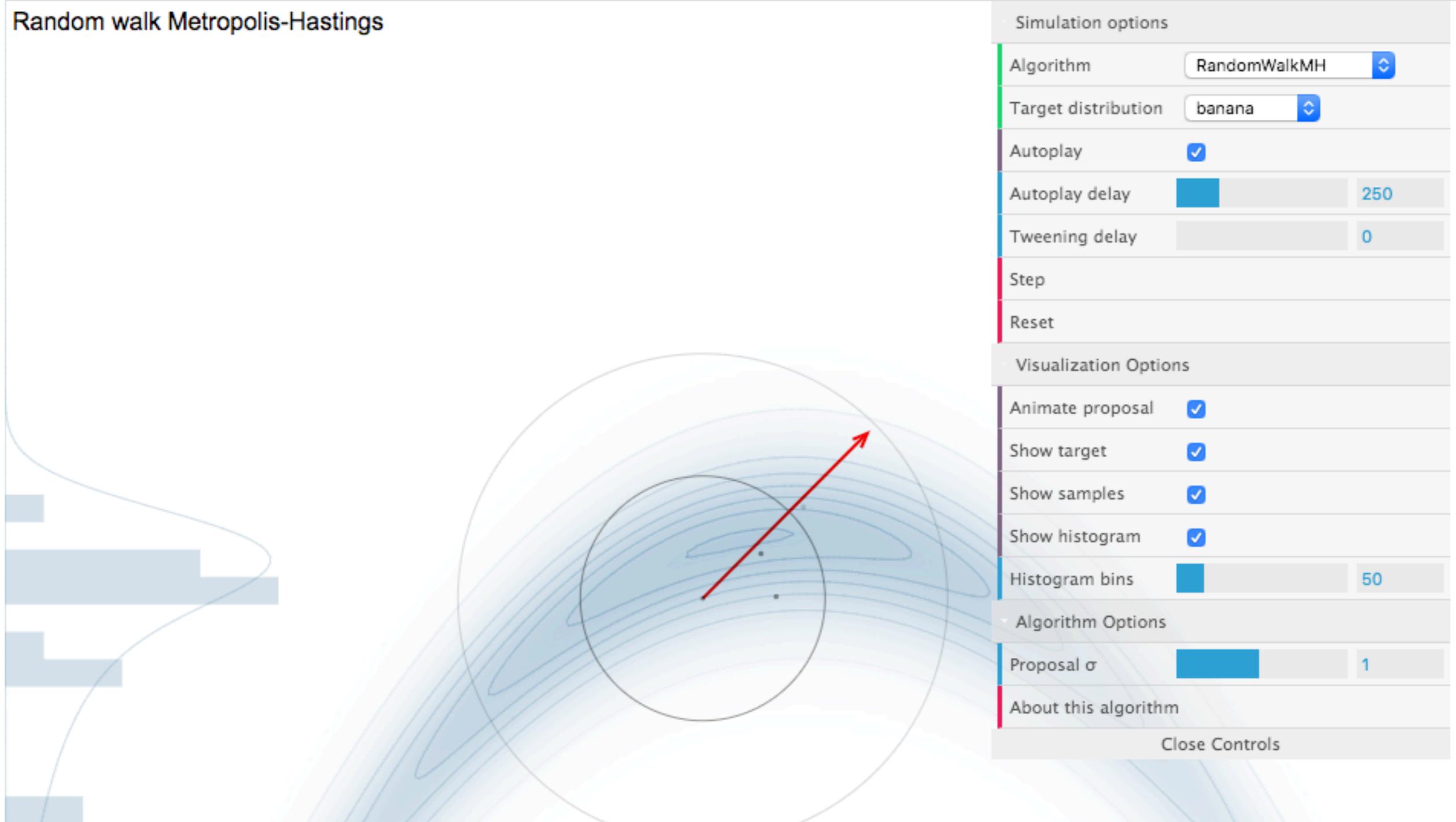
decide how Tobi
spends his weekend :)



<https://forms.gle/gKekvt62y9c8X7SQ6>

Things that came up

MCMC visualizations



<https://chi-feng.github.io/mcmc-demo/>

Questions

question @81 stop following 13 views Actions ▾

Brm vs. lmer

When should we use brm over lmer? Is there an advantage to using both? Thank you.

rstudio

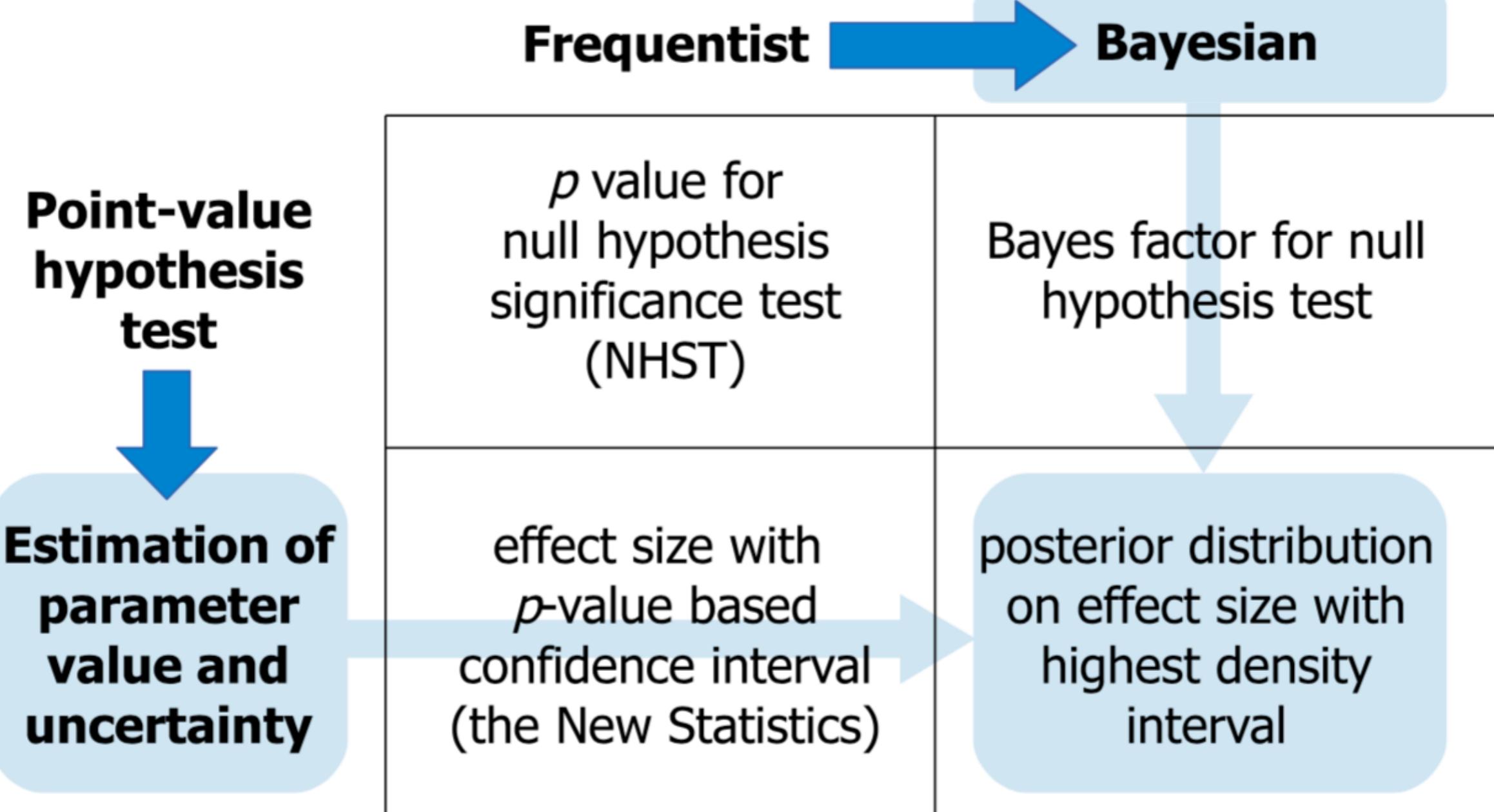
edit · good question | 0 Updated 1 day ago by Anonymous Scale

Bayesian vs. Frequentist

In classical frequentist statistics, parameter estimates are obtained by finding those parameter values that maximize the likelihood. In contrast, **Bayesian statistics estimate the full (joint) posterior distribution of the parameters.** Estimating the full posterior distribution not only is fully consistent with probability theory, but also is much more informative than estimating a single point (with an approximate measure of uncertainty commonly known as standard error).

Obtaining the posterior distribution analytically is rarely possible, and thus **Bayesian statistics relies on Markov-Chain Monte Carlo methods to obtain samples (i.e., random values) from the posterior distribution.** Such sampling algorithms are computationally very intensive, and thus **fitting models using Bayesian statistics is usually much slower than fitting models using frequentist statistics.** However, the advantages of Bayesian statistics—such as **greater modeling flexibility**, inclusion of prior distributions, and **more informative results**—are often worth the increased computational cost.

Frequentist vs. Bayes



very nice paper!

Plan for today

- Evidence for the "null hypothesis"
- Dealing with unequal variance
- Better modeling slider data
 - ZOIB: Zero-one inflated beta binomial model
- Better modeling Likert scale data
 - Ordinal logistic regression

Plan for today

- **Evidence for the "null hypothesis"**
- Dealing with unequal variance
- Better modeling slider data
 - ZOIB: Zero-one inflated beta binomial model
- Better modeling Likert scale data
 - Ordinal logistic regression

Evidence for the null hypothesis



[Front Psychol. 2014; 5: 781.](#)

Published online 2014 Jul 29. doi: [10.3389/fpsyg.2014.00781](https://doi.org/10.3389/fpsyg.2014.00781)

PMCID: PMC4114196

PMID: [25120503](#)

Using Bayes to get the most out of non-significant results

Zoltan Dienes*

► Author information ► Article notes ► Copyright and License information [Disclaimer](#)

[HTML] Using Bayes to get the most out of non-significant results

[Z Dienes - Frontiers in psychology, 2014 - frontiersin.org](#)

No scientific conclusion follows automatically from a statistically non-significant result, yet people routinely use non-significant results to guide conclusions about the status of theories (or the effectiveness of practices). To know whether a non-significant result counts against a theory, or if it just indicates data insensitivity, researchers must use one of: power, intervals (such as confidence or credibility intervals), or else an indicator of the relative evidence for one theory over another, such as a Bayes factor. I argue Bayes factors allow theory to be ...

☆ 99 Cited by 966 Related articles All 14 versions Web of Science: 583 Import into BibTeX »

- There is nothing special about H_0 compared to H_1 in Bayesian inference
- We can get evidence of H_0 over H_1 (e.g. using the Bayes factor approach)

Evidence for the null hypothesis

Data



#6 heads, #4 tails

Hypotheses

Is the coin's weight different from 0.5?

$$H_0 : \theta = 0.5$$

$$H_1 : \theta \neq 0.5$$



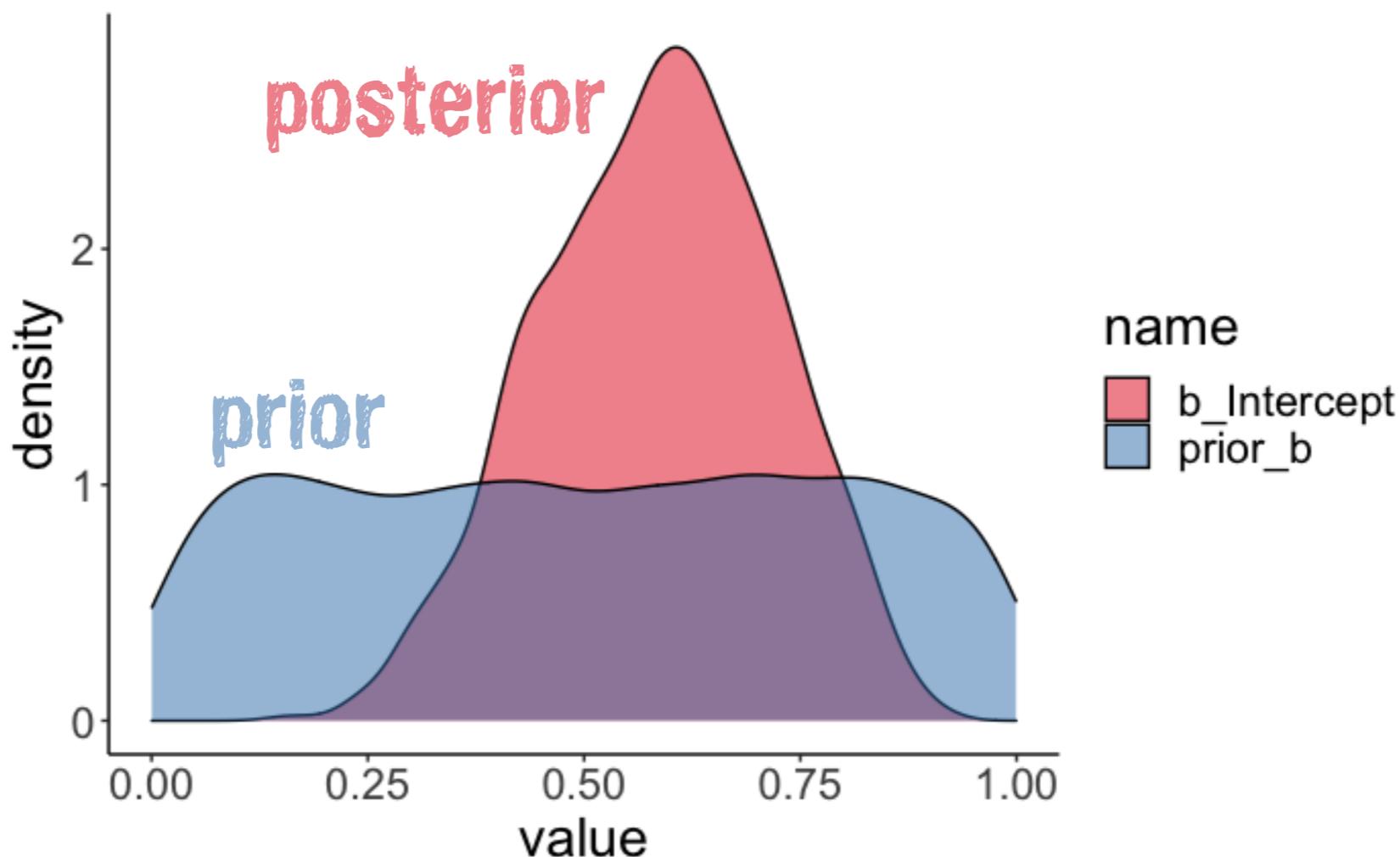
requires integrating over
all alternative values?!

Evidence for the null hypothesis

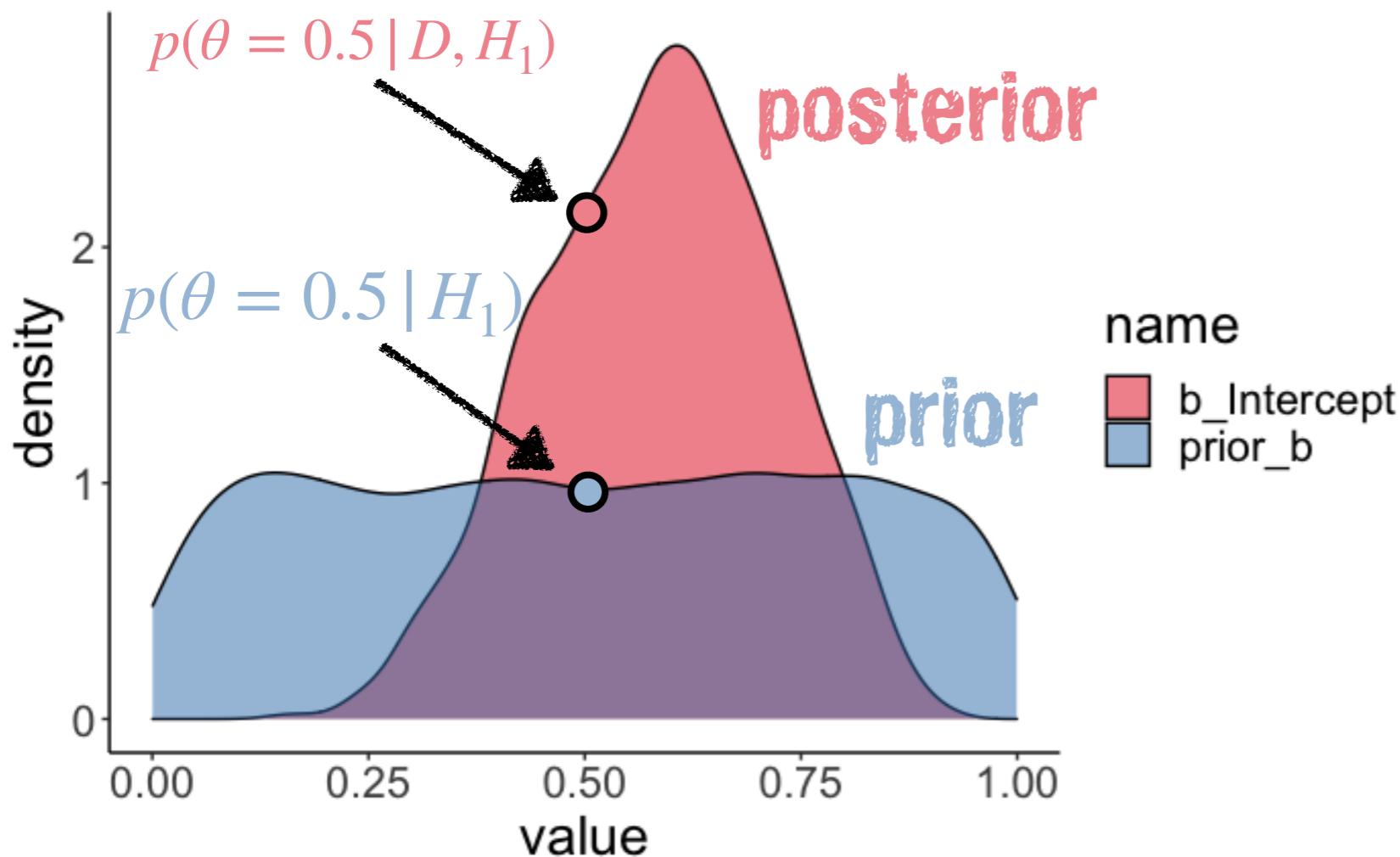
```
1 df.null = tibble(s = 6, k = 10)
2
3 fit.brm0 = brm(s | trials(k) ~ 0 + Intercept,
4                  family = binomial(link = "identity"),
5                  prior = set_prior("beta(1, 1)", class = "b", lb = 0, ub = 1),
6                  data = df.null,
7                  sample_prior = TRUE,
8                  cores = 4,
9                  file = "cache/brm0")
```

explicitly specify the Intercept
so we can set a prior

also get samples from the
prior (only works if priors have
been explicitly specified)



Evidence for the null hypothesis



The Savage–Dickey density ratio

$$BF_{01} = \frac{p(D | H_0)}{p(D | H_1)} = \frac{p(\theta = 0.5 | D, H_1)}{p(\theta = 0.5 | H_1)}$$

we don't have to integrate over
all possible values of theta

```
1 fit.brm0 %>%
2   hypothesis(hypothesis = "Intercept = 0.5")
```

$$BF_{01} = 2.22$$

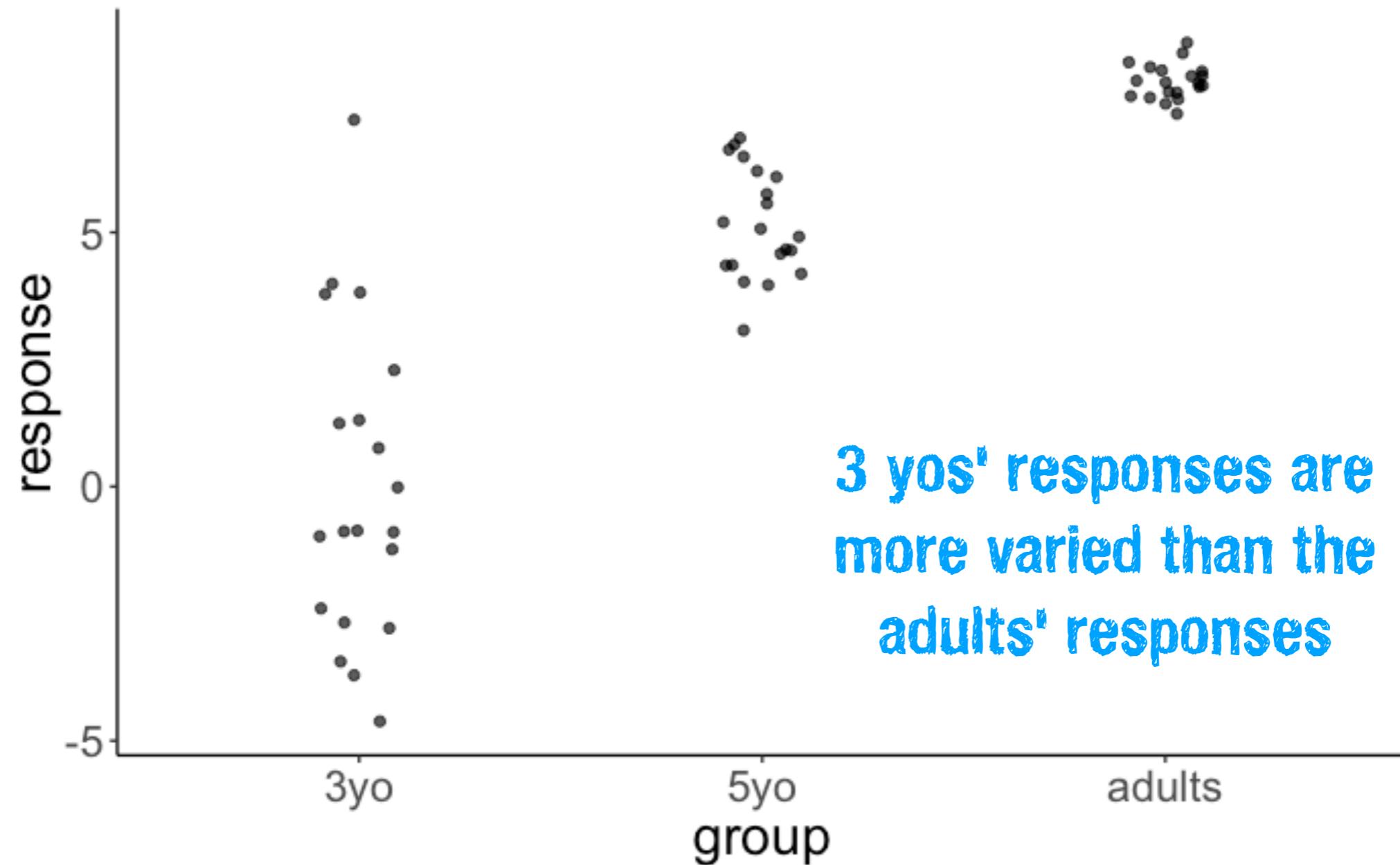
H_0 is more than twice as likely as H_1

Plan for today

- Evidence for the "null hypothesis"
- **Dealing with unequal variance**
- Better modeling slider data
 - ZOIB: Zero-one inflated beta binomial model
- Better modeling Likert scale data
 - Ordinal logistic regression

Unequal variance aka heteroscedasticity

```
1 df.variance = tibble(group = rep(c("3yo", "5yo", "adults"), each = 20),  
2                         response = rnorm(n = 60,  
3                                         mean = rep(c(0, 5, 8), each = 20),  
4                                         sd = rep(c(3, 1.5, 0.3), each = 20)))
```



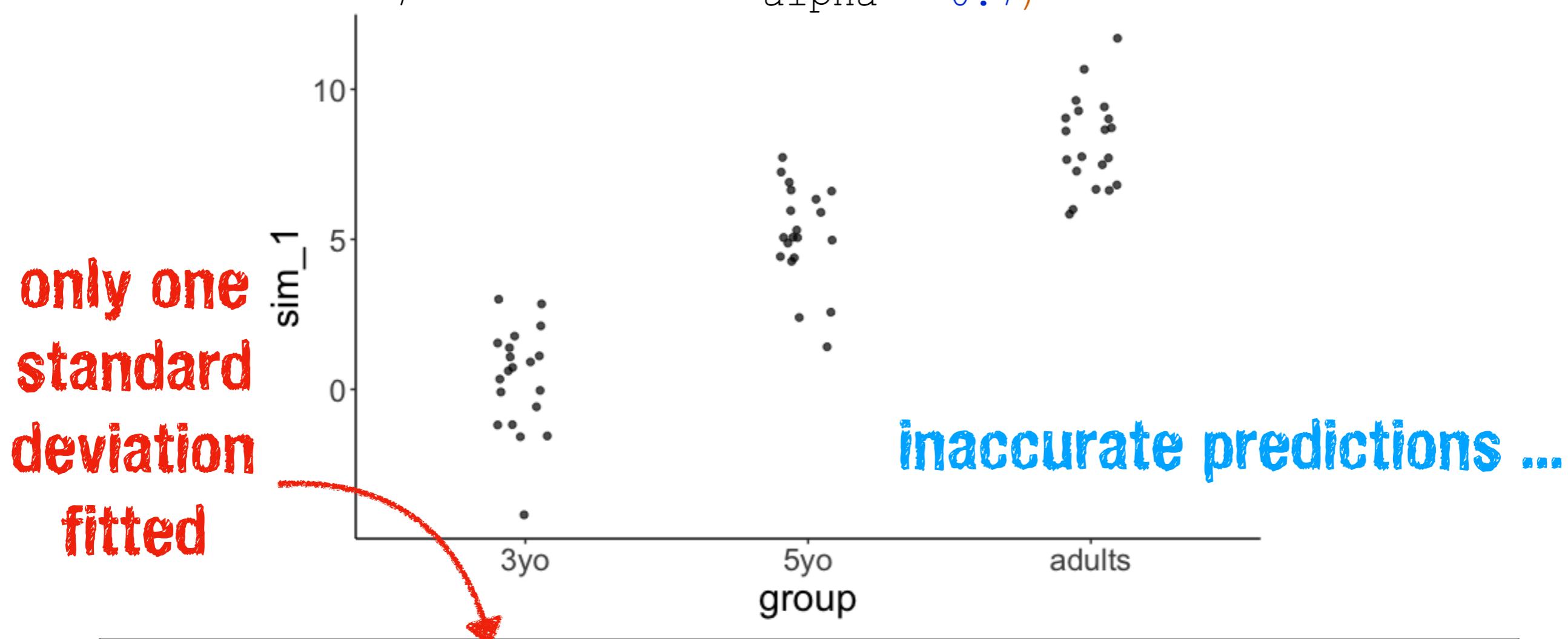
Unequal variance aka heteroscedasticity

```
1 fit.lm1 = lm(formula = response ~ 1 + group,  
2                         data = df.variance)  
3  
4 fit.lm1 %>%  
5   summary()
```

```
Call:  
lm(formula = response ~ 1 + group, data = df.variance)  
  
Residuals:  
    Min      1Q  Median      3Q     Max  
-4.6145 -0.8288 -0.0879  0.6315  7.2193  
  
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)  
(Intercept) -0.005336  0.421618 -0.013    0.99  
group5yo      5.172810  0.596258  8.675 5.25e-12 ***  
groupadults   7.970655  0.596258 13.368 < 2e-16 ***  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 1.886 on 57 degrees of freedom  
Multiple R-squared:  0.7635, Adjusted R-squared:  0.7552  
F-statistic: 91.99 on 2 and 57 DF,  p-value: < 2.2e-16
```

Unequal variance aka heteroscedasticity

```
1 fit.lm1 %>%
2   simulate() %>%
3   bind_cols(df.variance) %>%
4   ggplot(aes(x = group, y = sim_1)) +
5   geom_jitter(height = 0,
6                 width = 0.1,
7                 alpha = 0.7)
```



r.squared	adj.r.squared	sigma	statistic	p.value	df	logLik	AIC	BIC	deviance	df.residual
0.76	0.76	1.89	91.99	0	3	-121.65	251.3	259.68	202.65	57

Unequal variance aka heteroscedasticity

```
1 fit.brml = brm(formula = bf(response ~ group,  
2                   sigma ~ group),  
3                   data = df.variance,  
4                   file = "cache/brml",  
5                   seed = 1)
```

modeling both the
means and variances

```
Family: gaussian  
Links: mu = identity; sigma = log  
Formula: response ~ group  
         sigma ~ group  
Data: df.variance (Number of observations: 60)  
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;  
         total post-warmup samples = 4000
```

Population-Level Effects:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	-0.01	0.73	-1.41	1.51	1.01	1107	1072
sigma_Intercept	1.15	0.17	0.85	1.51	1.00	1991	1922
group5yo	5.18	0.77	3.60	6.65	1.00	1252	1327
groupadults	7.98	0.74	6.47	9.37	1.01	1110	1079
sigma_group5yo	-1.05	0.24	-1.51	-0.57	1.00	2249	2420
sigma_groupadults	-2.19	0.24	-2.66	-1.74	1.00	2171	2427

Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

Unequal variance aka heteroscedasticity

```
Family: gaussian  
Links: mu = identity; sigma = log ← on a log scale!  
Formula: response ~ group  
         sigma ~ group  
Data: df.variance (Number of observations: 60)  
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;  
         total post-warmup samples = 4000
```

Population-Level Effects:

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	-0.01	0.73	-1.41	1.51	1.01	1107	1072
sigma_Intercept	1.15	0.17	0.85	1.51	1.00	1991	1922
group5yo	5.18	0.77	3.60	6.65	1.00	1252	1327
groupadults	7.98	0.74	6.47	9.37	1.01	1110	1079
sigma_group5yo	-1.05	0.24	-1.51	-0.57	1.00	2249	2420
sigma_groupadults	-2.19	0.24	-2.66	-1.74	1.00	2171	2427

Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

mean = **c(0, 5, 8)**
sd = **c(3, 1.5, 0.3)**



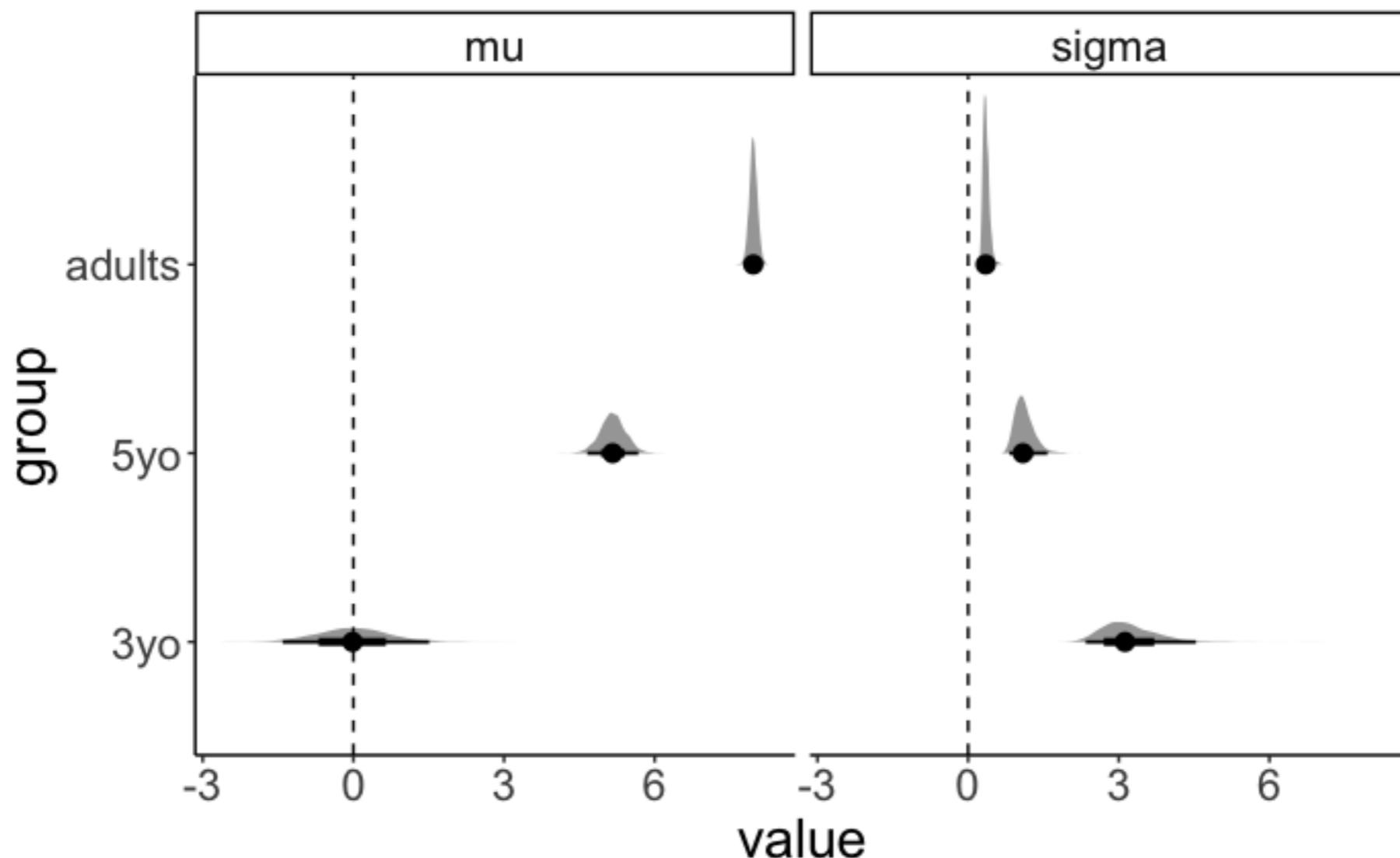
3 year olds $e^{1.15} = 3.16$

5 year olds $e^{1.15+(-1.05)} = 1.10$

adults $e^{1.15+(-2.19)} = 0.35$

Unequal variance aka heteroscedasticity

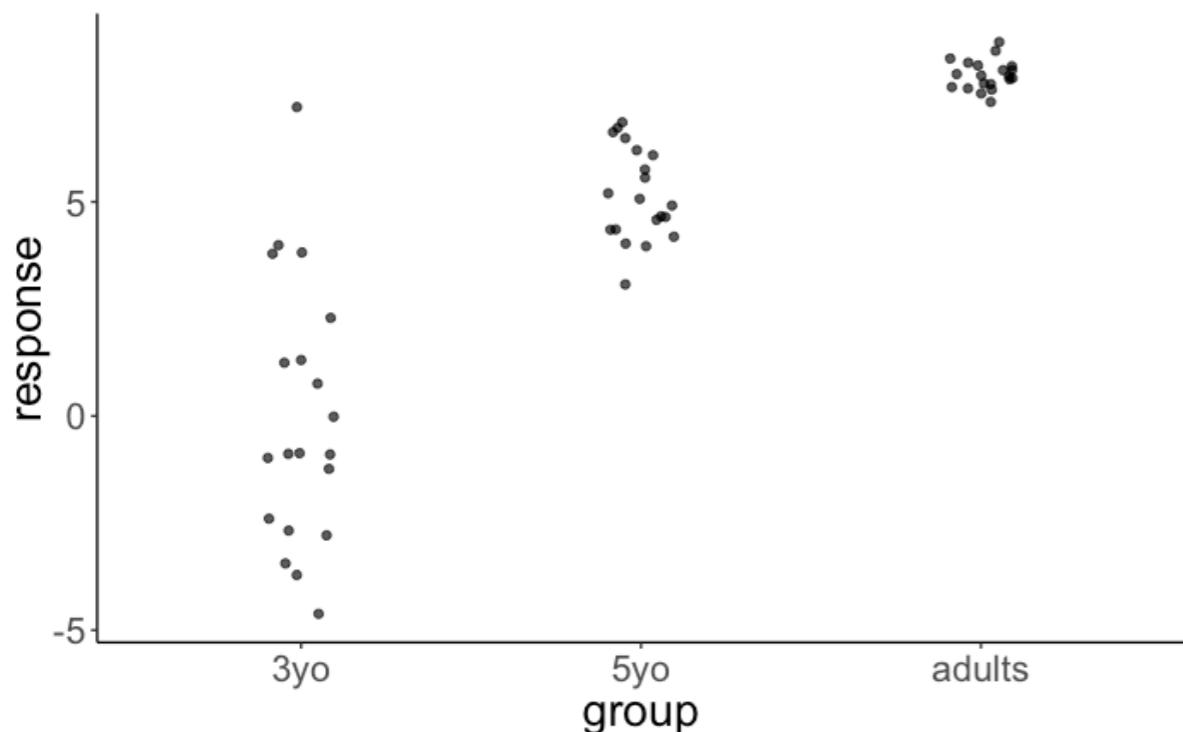
```
1 df.variance %>%
2   expand(group) %>%
3   add_fitted_draws(fit.brml, dpar = TRUE) %>%
4   select(group, .row, .draw, posterior = .value, mu, sigma) %>%
5   pivot_longer(cols = c(mu, sigma),
6                 names_to = "index",
7                 values_to = "value") %>%
8   ggplot(aes(x = value, y = group)) +
9   geom_halfeyeh() +
10  geom_vline(xintercept = 0, linetype = "dashed") +
11  facet_grid(cols = vars(index))
```



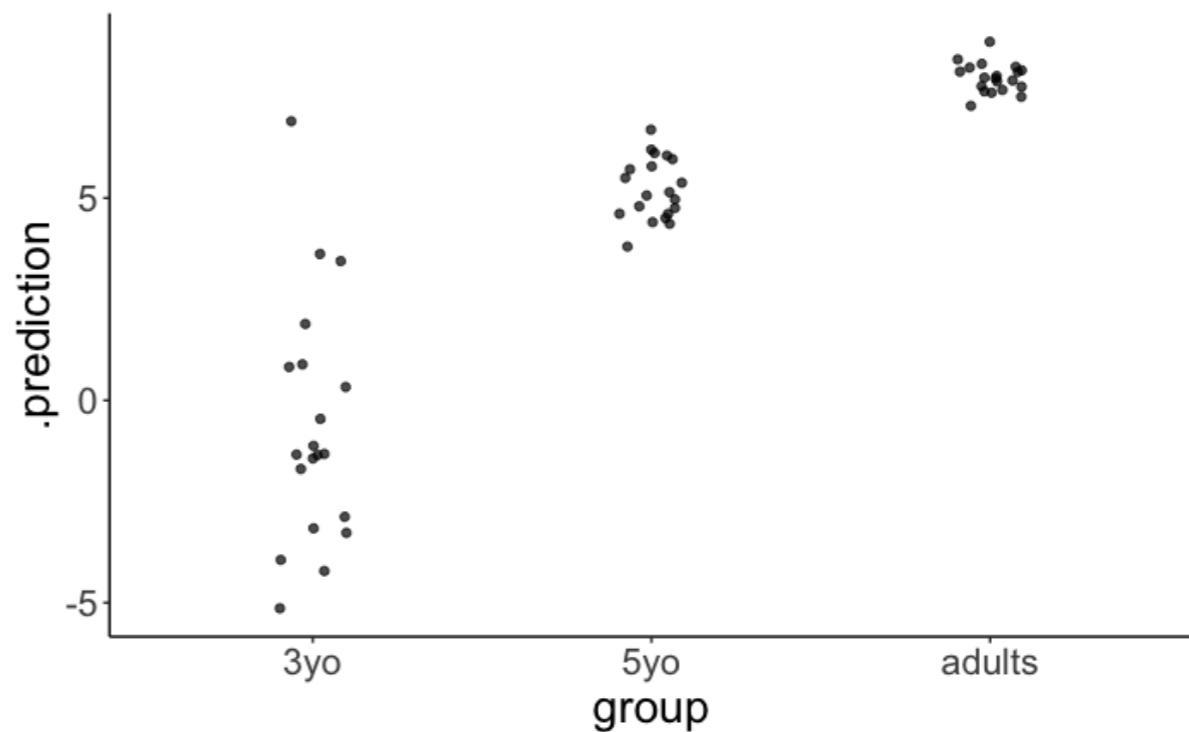
Unequal variance aka heteroscedasticity

```
1 df.variance %>%
2   add_predicted_draws(model = fit.brml,
3                       n = 1) %>%
4   ggplot(aes(x = group, y = .prediction)) +
5   geom_jitter(height = 0,
6                width = 0.1,
7                alpha = 0.7)
```

original data



predicted data



these predictions look good!

Plan for today

- Evidence for the "null hypothesis"
- Dealing with unequal variance
- **Better modeling slider data**
 - **ZOIB: Zero-one inflated beta binomial model**
- Better modeling Likert scale data
 - Ordinal logistic regression

Example taken from ...

How to analyze visual analog (slider) scale data?

A reasonable choice might be the zero-one-inflated beta model

Feb 18, 2019 · 25 min read · psychology, statistics

- [Introduction](#)
- [The zero-one-inflated beta model](#)
- [ZOIB regression](#)
- [Simulation: Compare ZOIB and t-test performances](#)
- [Discussion](#)
- [References](#)

Introduction

In psychological experiments, subjective responses are often collected using two types of response scales: ordinal and visual analog scales. These scales are unlikely to provide normally distributed data. However, researchers often analyze responses from these scales with models that assume normality of the data.¹

Ordinal scales, of which binary ratings are a special case, provide ordinal data and are thus better analyzed using ordinal models (Bürkner and Vuorre 2018; Liddell and Kruschke 2018).

<https://vuorre.netlify.com/post/2019/02/18/analyze-analog-scale-ratings-with-zero-one-inflated-beta-models/#zoib-regression>

In general, what is more important to you?

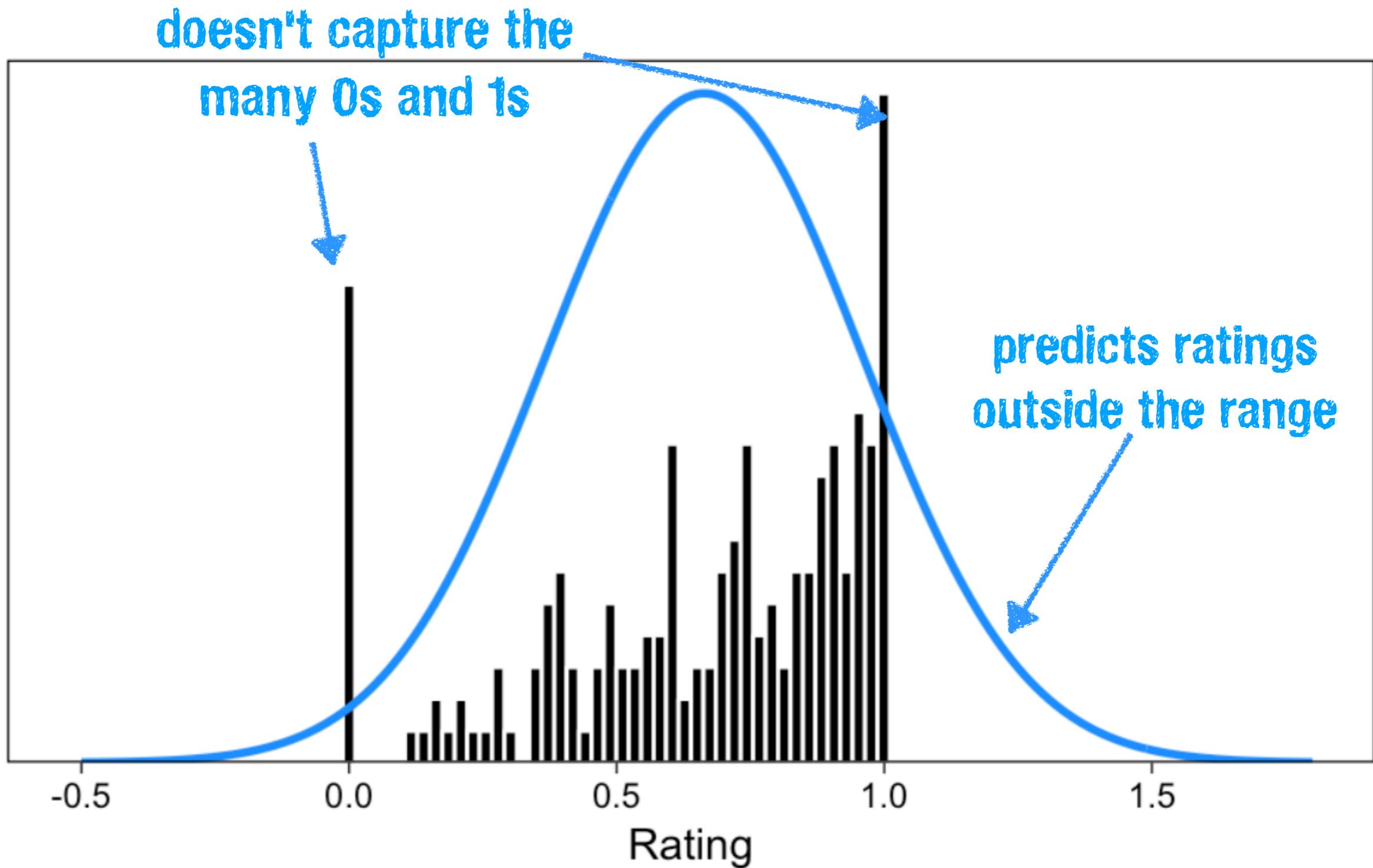
Leisure

Money



Next

Normality assumption is (almost always) violated



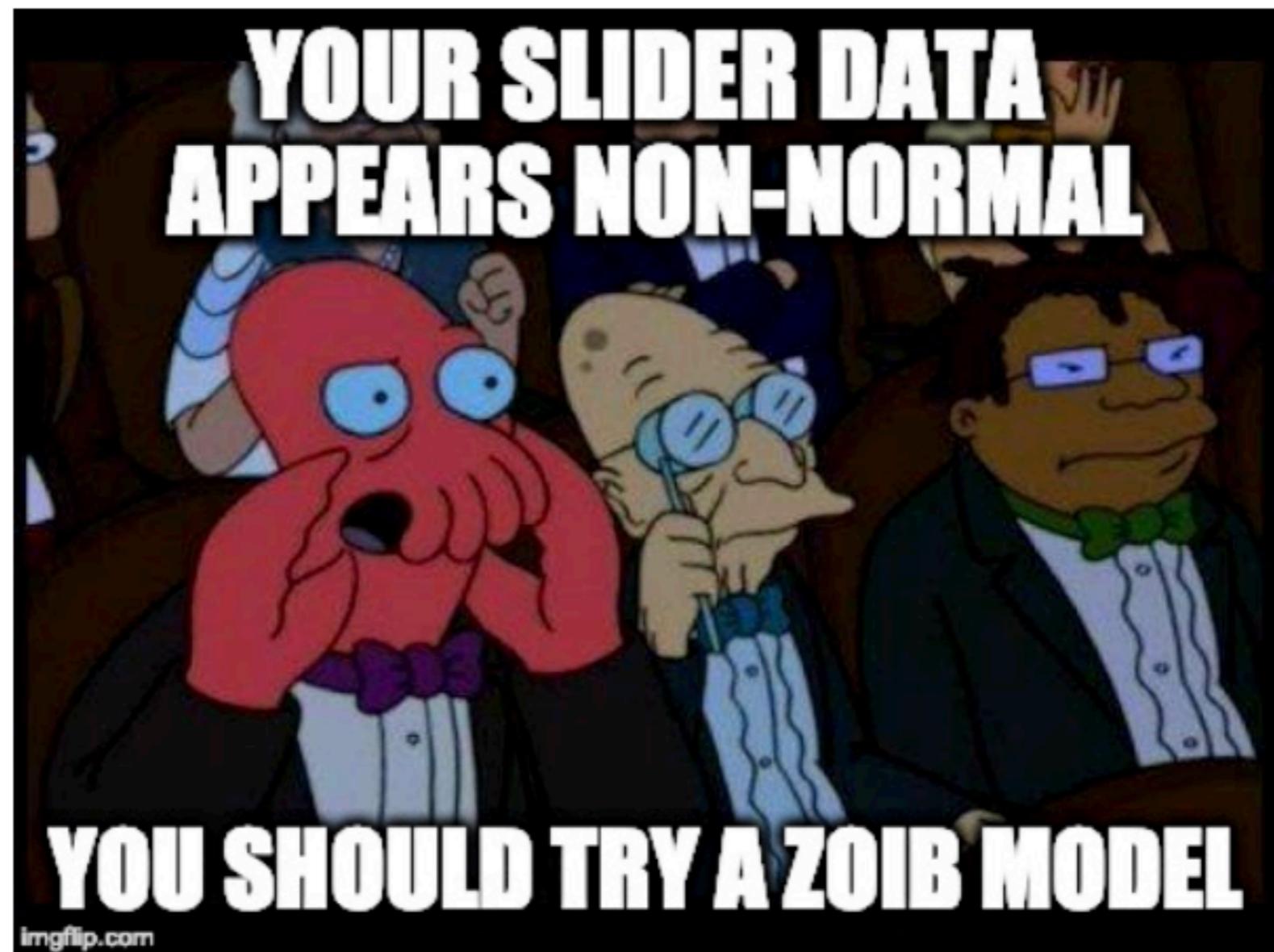
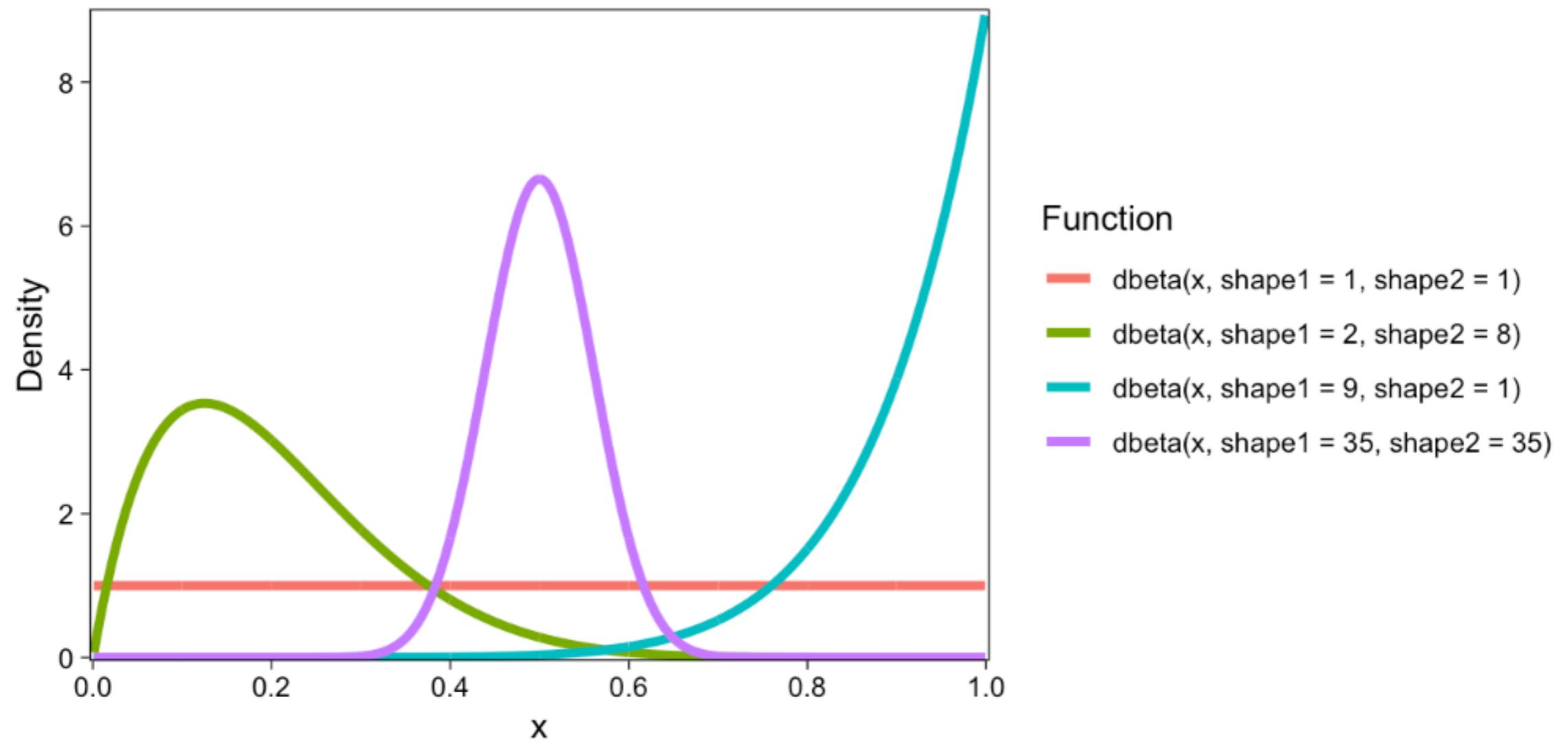


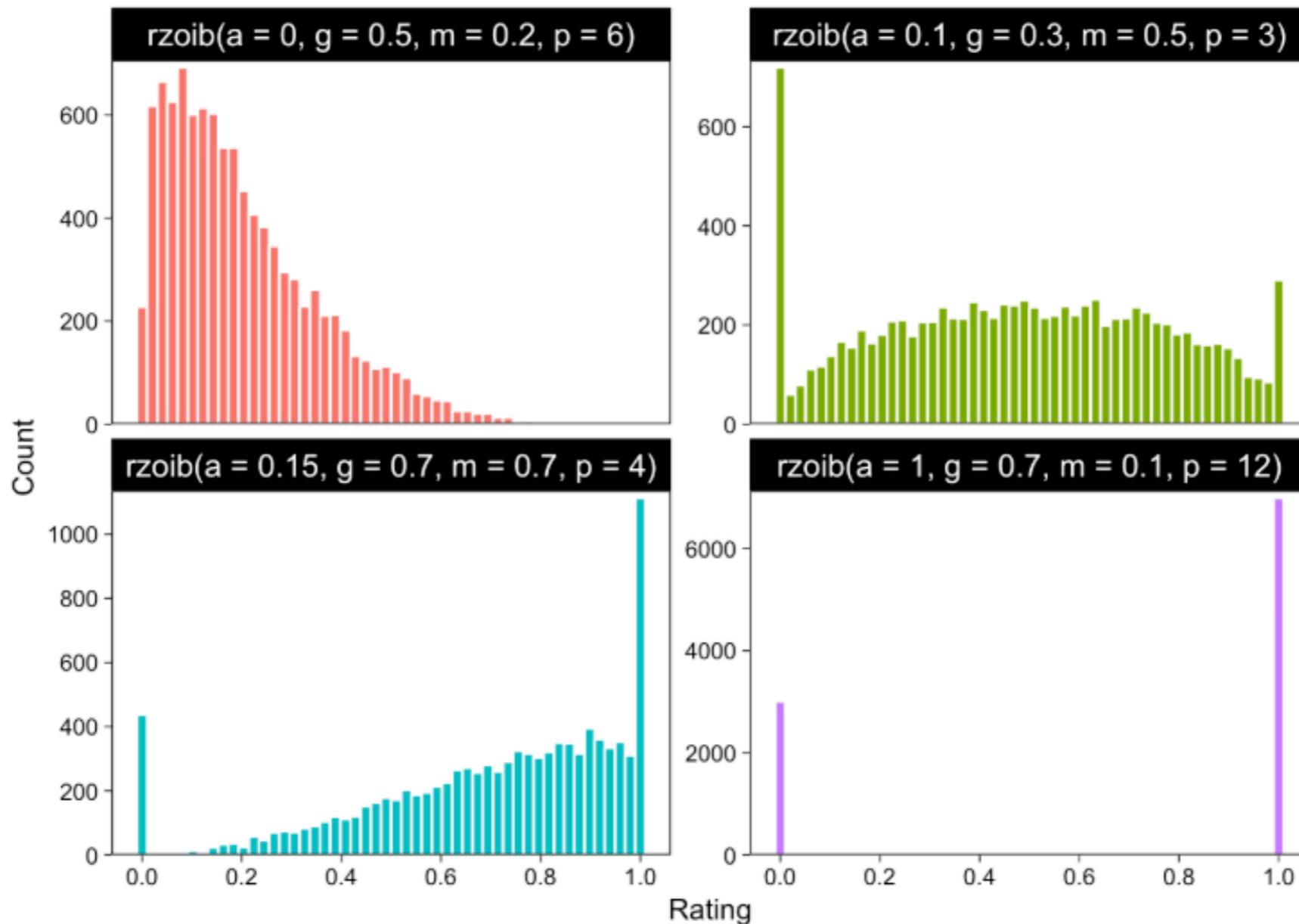
Figure 3: Dr. John A. Zoidberg thinks you should try a ZOIB model on your slider scale data.

Beta distribution



Zero-one inflated **beta** binomial model

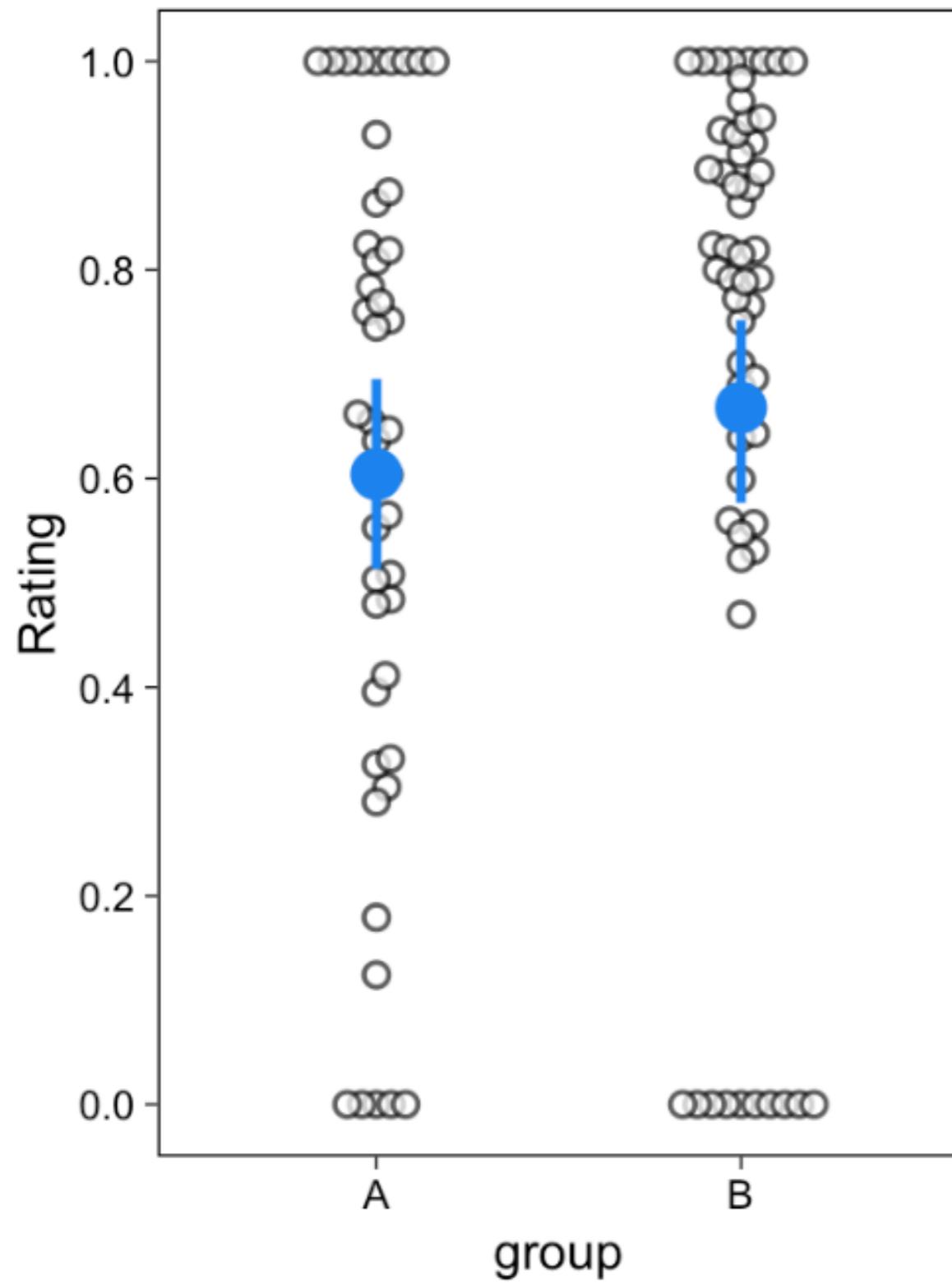
Zero-one inflated beta binomial model



Generative process

Some chance the a person will pick a 0 or 1, if not then response is determined by the beta distribution.

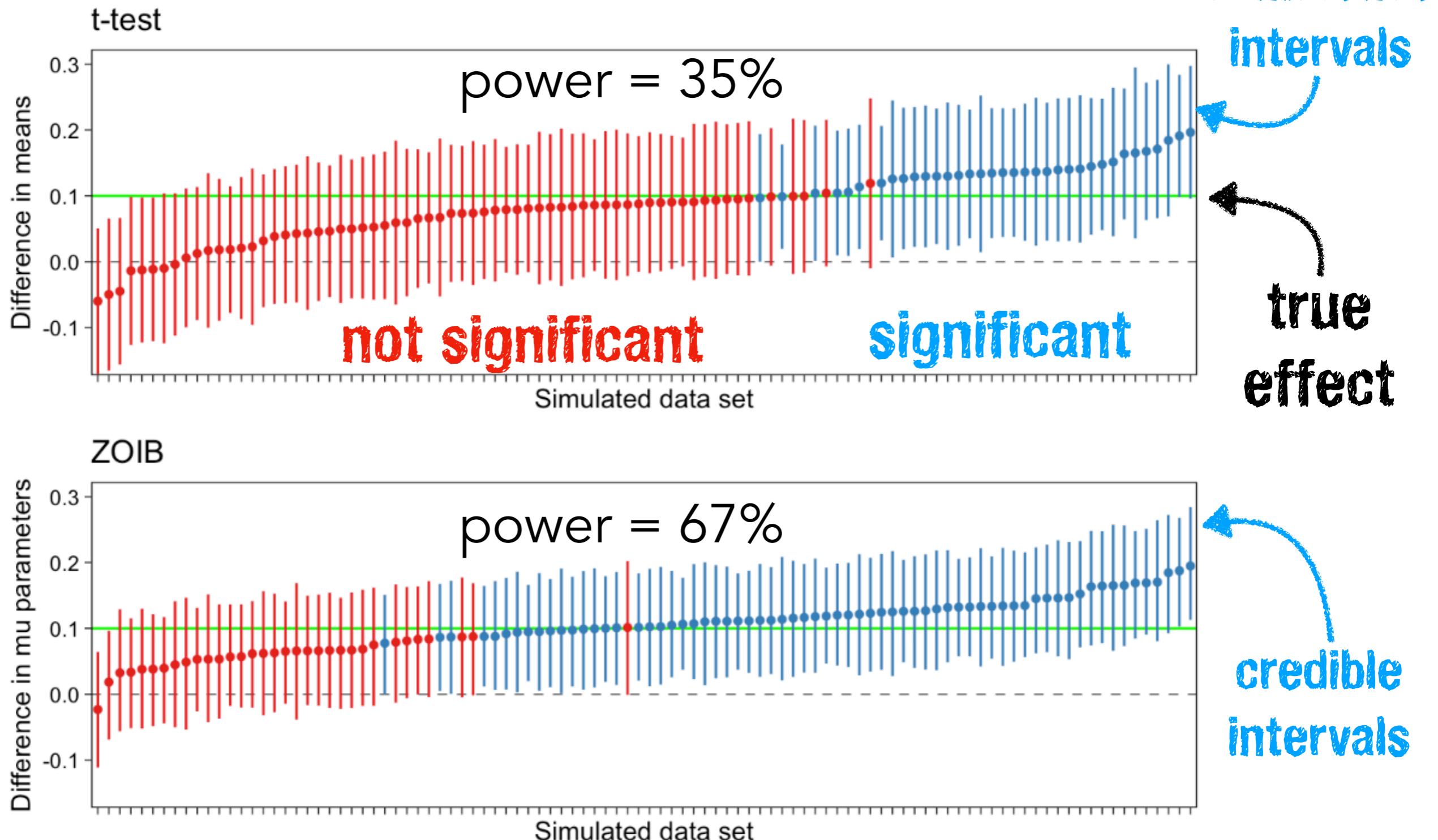
Example data



Fit the ZOIB

```
zoib_model = bf(  
    Rating ~ group,  
    phi ~ group,  
    zoi ~ group,  
    coi ~ group,  
    family = zero_one_inflated_beta()  
)  
  
fit = brm(  
    formula = zoib_model,  
    data = dat  
)
```

Capturing the data-generating process gives you power



Plan for today

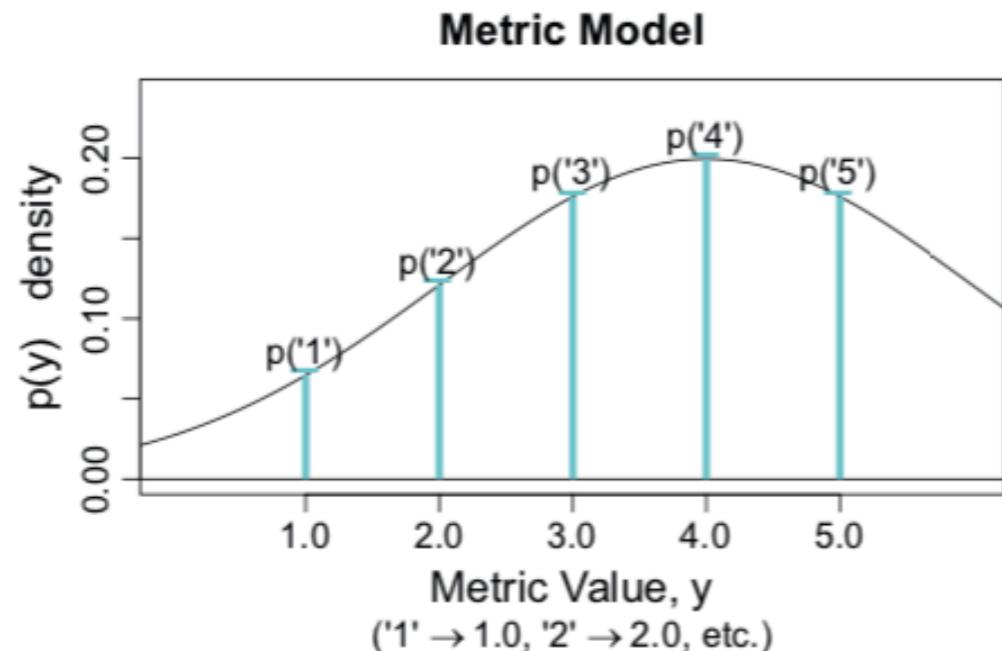
- Evidence for the "null hypothesis"
- Dealing with unequal variance
- Better modeling slider data
 - ZOIB: Zero-one inflated beta binomial model
- **Better modeling Likert scale data**
 - **Ordinal regression**

We surveyed all articles in the *Journal of Personality and Social Psychology* (*JPSP*), *Psychological Science* (*PS*), and the *Journal of Experimental Psychology: General* (*JEP:G*) that mentioned the term “Likert,” and found that **100% of the articles** that analyzed ordinal data did so using a metric model.

great paper!

Liddell & Kruschke (2018) Analyzing ordinal data with metric models: What could possibly go wrong?. *Journal of Experimental Social Psychology*

Ordinal regression



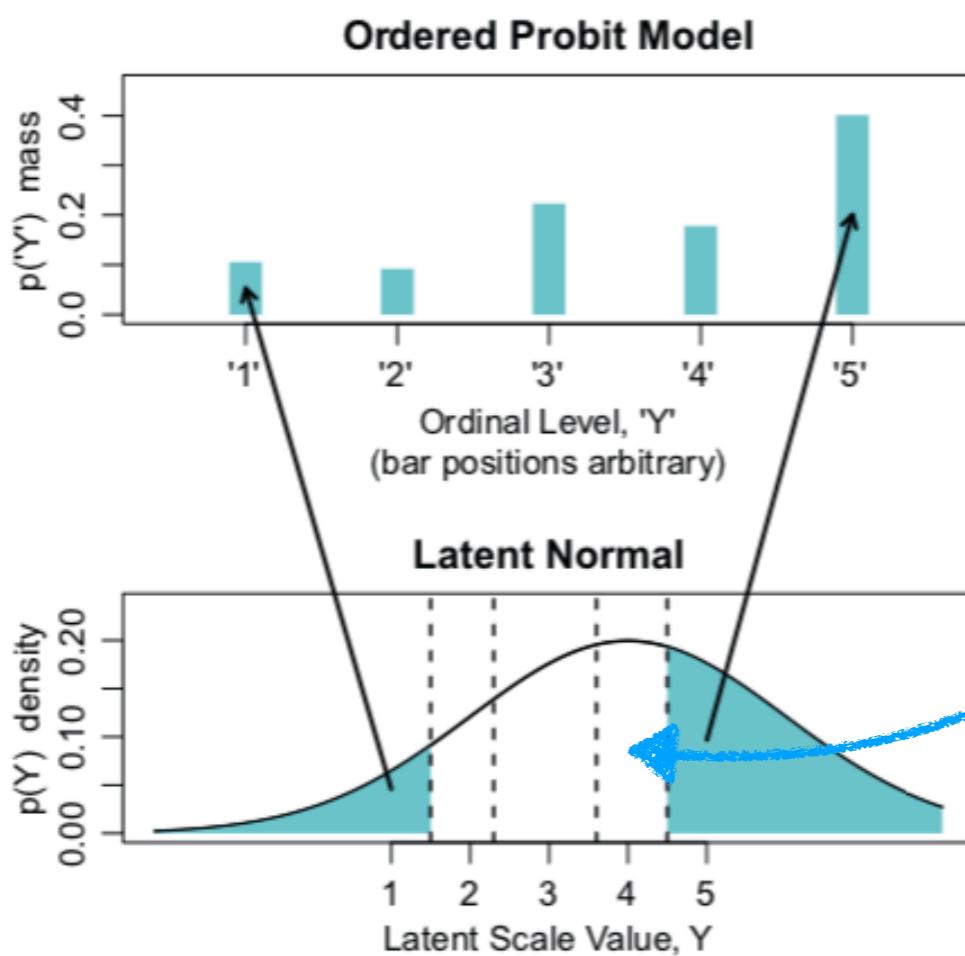
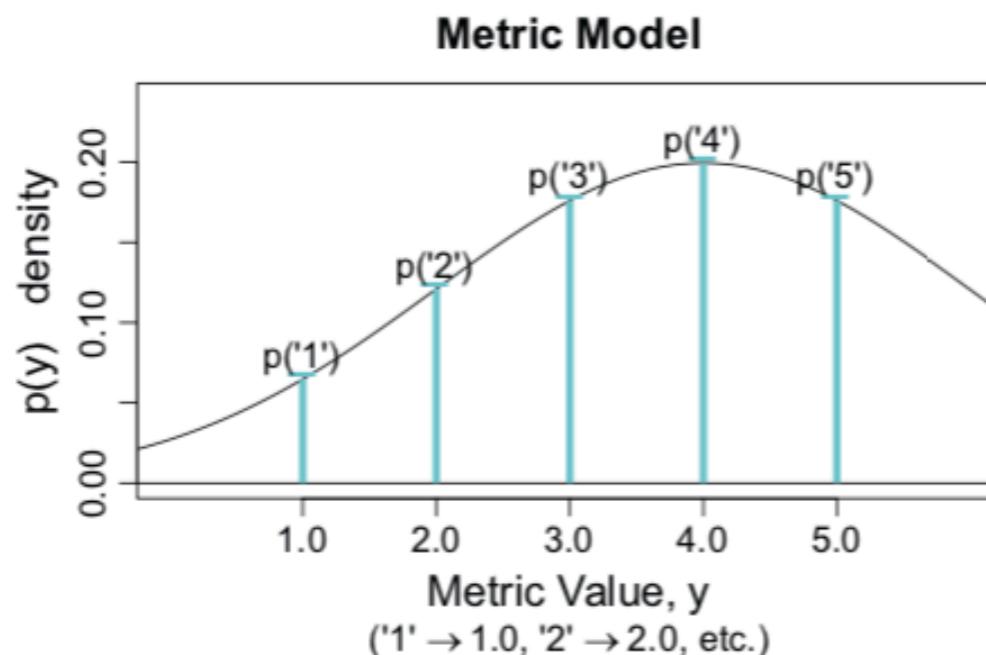
metric model

- fits a Gaussian distribution with mean and standard deviation
- makes the assumption that categories are **equidistant**

I loooove Bayesian statistics!

- completely disagree (1)
- moderately disagree (2)
- neither disagree nor agree (3)
- moderately agree (4)
- completely agree (5)

Ordinal regression

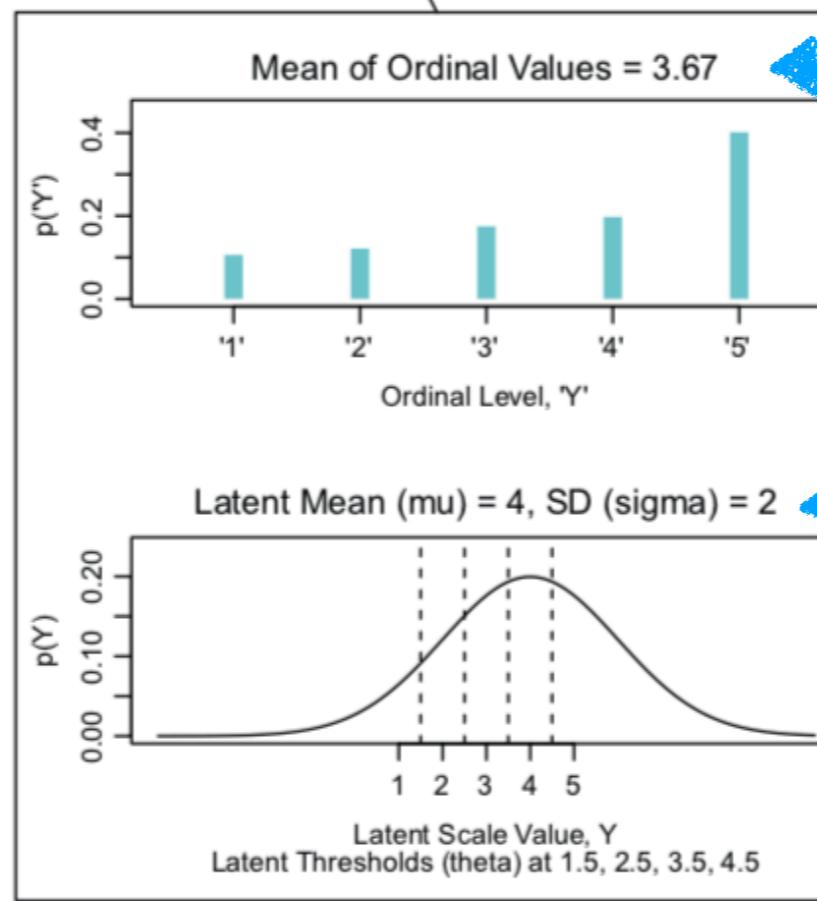
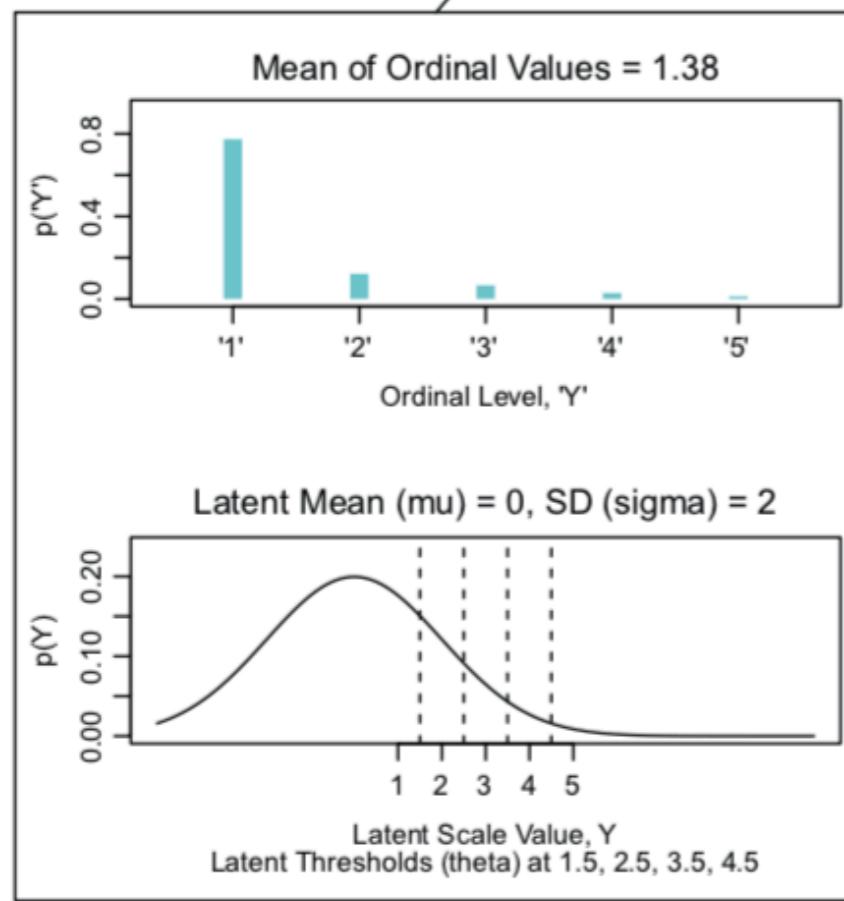
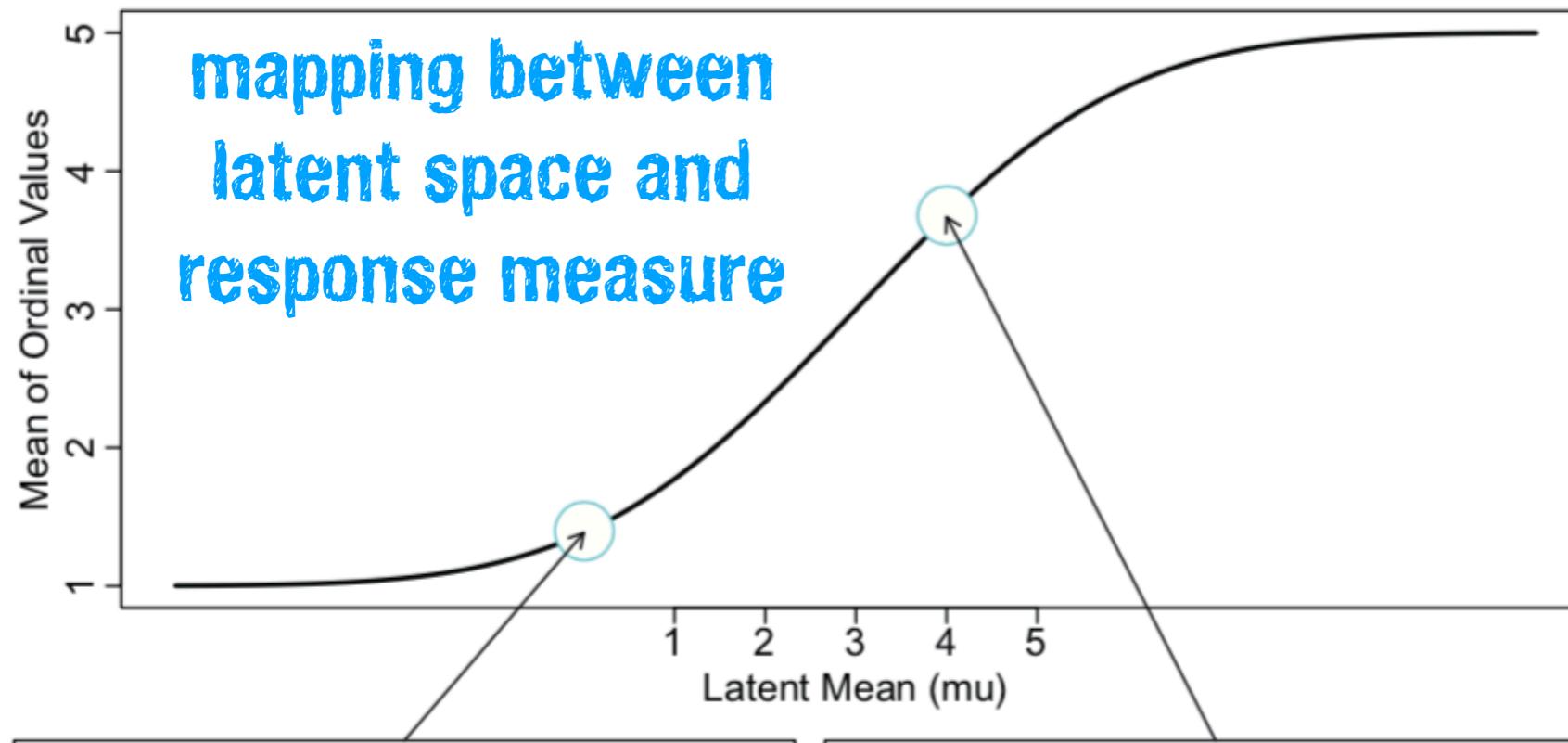


ordered probit model

- assumes an ordering of the categories
- assumes a normal distribution in latent space
- finds a good mapping from that normal distribution in latent space to the ordered responses

finds thresholds that separate the categories

Ordinal regression



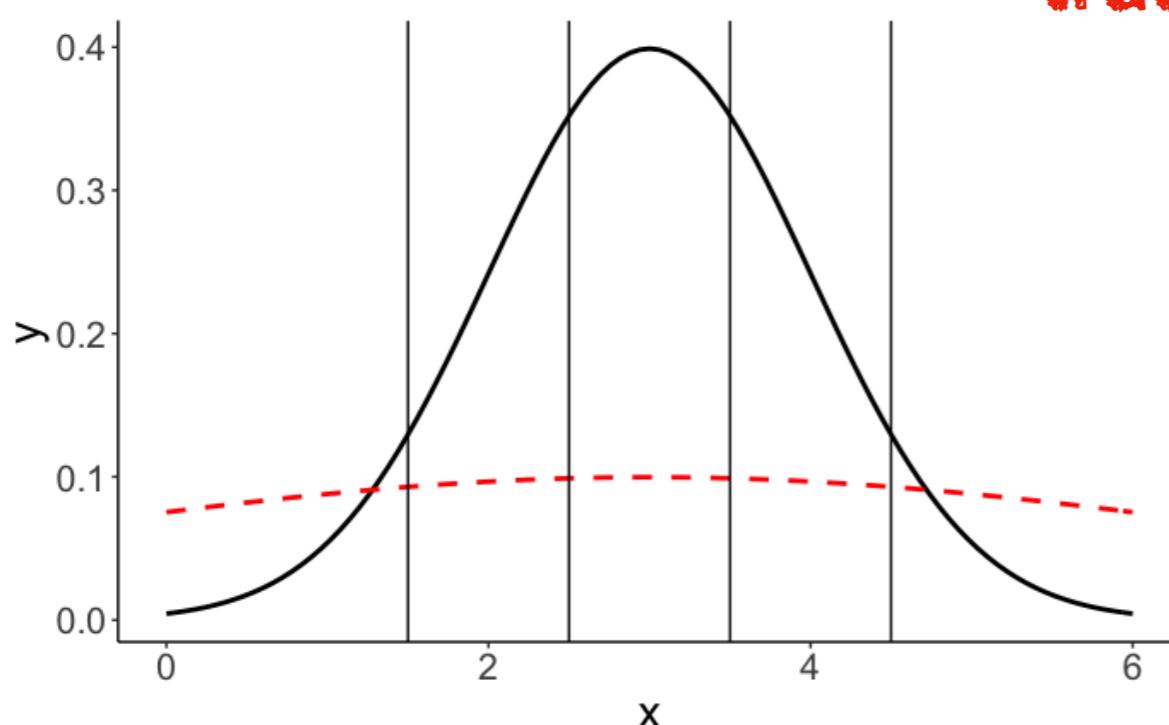
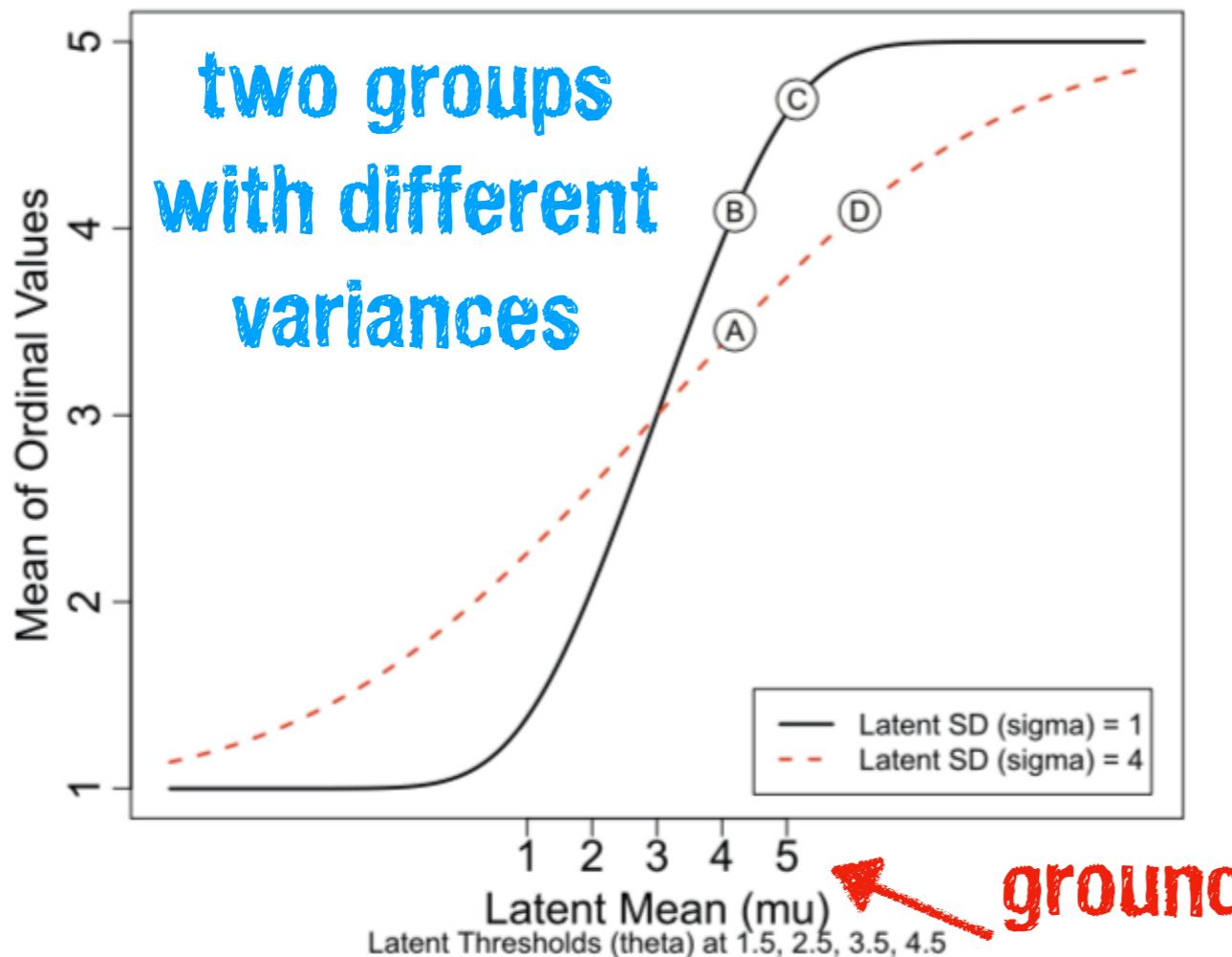
mean of the
ordinal values

different mean
in latent space

Does it matter?

- treating ordinal data as metric can lead to:
 - low rates of correct detection (Type II error)
 - distorted effect size estimates
 - inflated false alarms (Type I error)
 - inversions of differences between groups
- main reasons for why this happens:
 - response categories may not be equidistant
 - response distribution may be non-normal
 - variances of unobserved variables may differ between groups, conditions, time points, ...

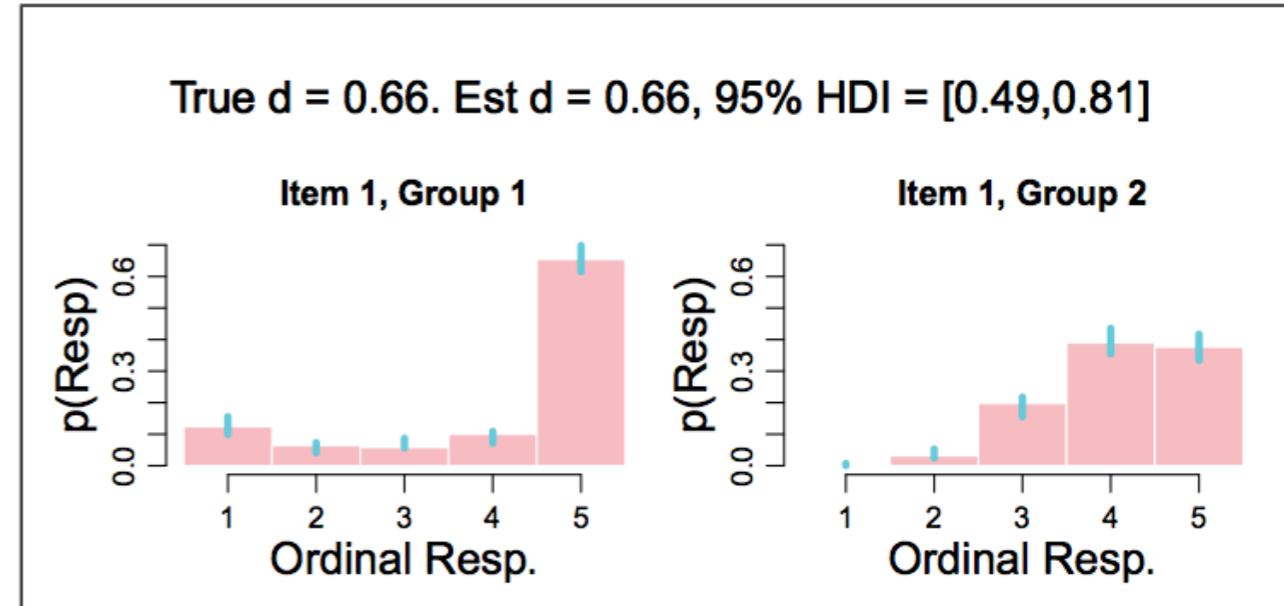
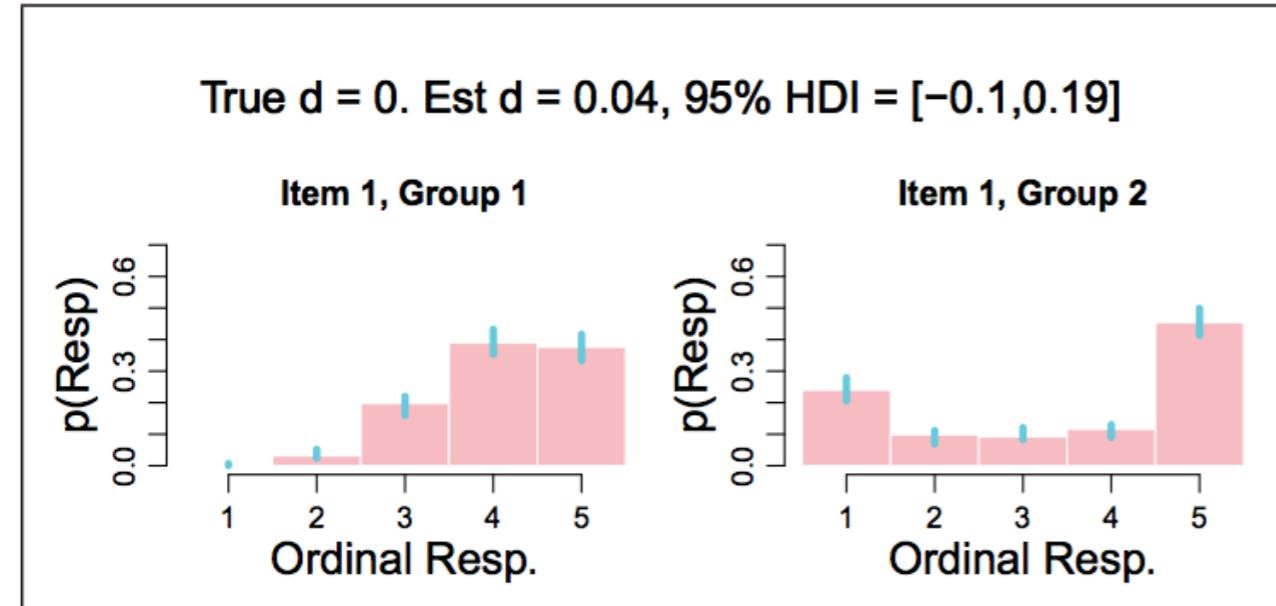
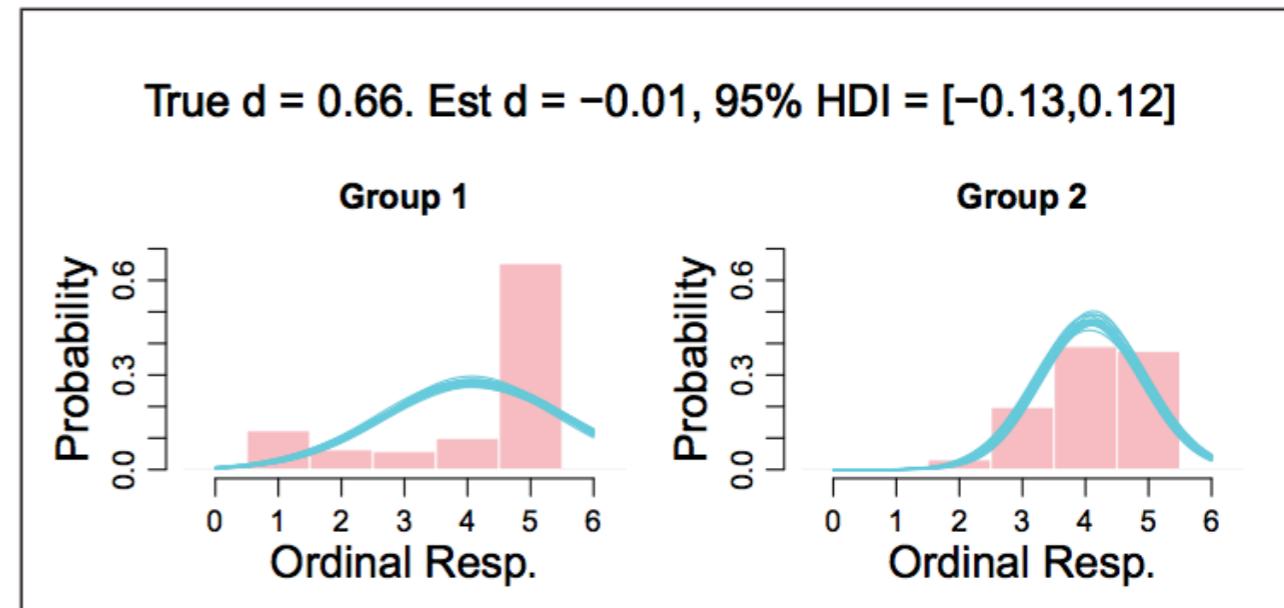
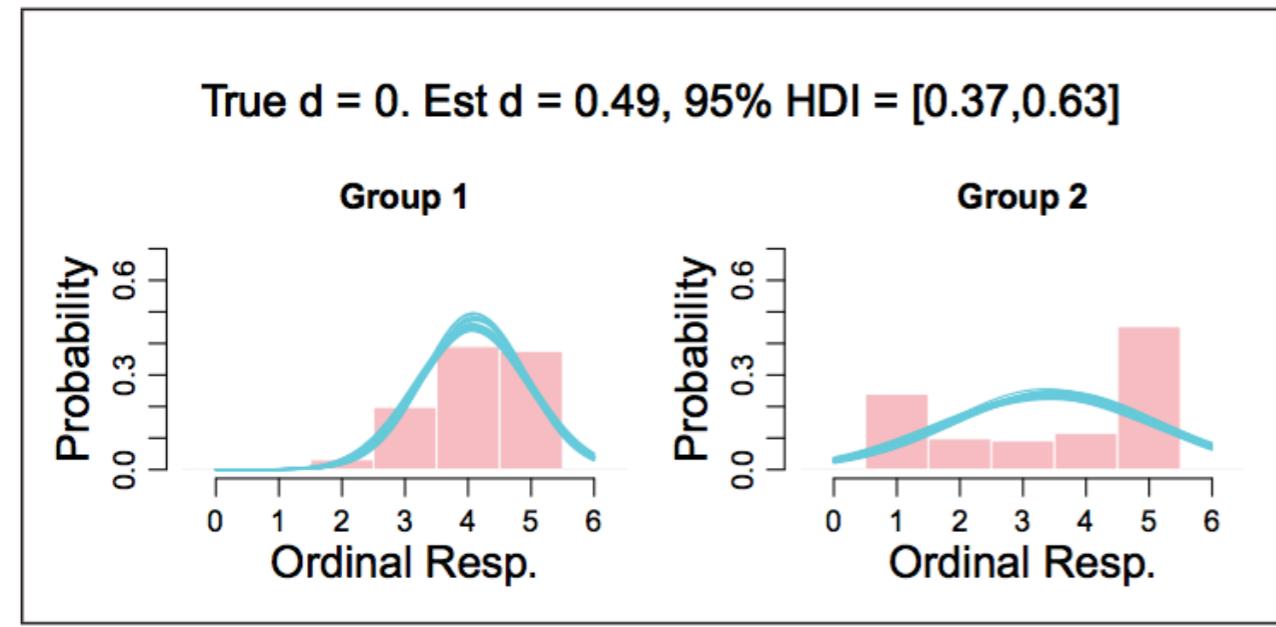
What could possibly go wrong?



- A vs. B: false positive (Type I error)
→ **no difference in latent space**
- B vs. D: false negative (Type II error)
→ **difference in latent space**
- C vs. D: reversal
→ **C is greater than D in metric space, but D is greater than C in latent space**

Does it matter?

false positive



Liddell & Kruschke (2018) Analyzing ordinal data with metric models: What could possibly go wrong?. Journal of Experimental Social Psychology

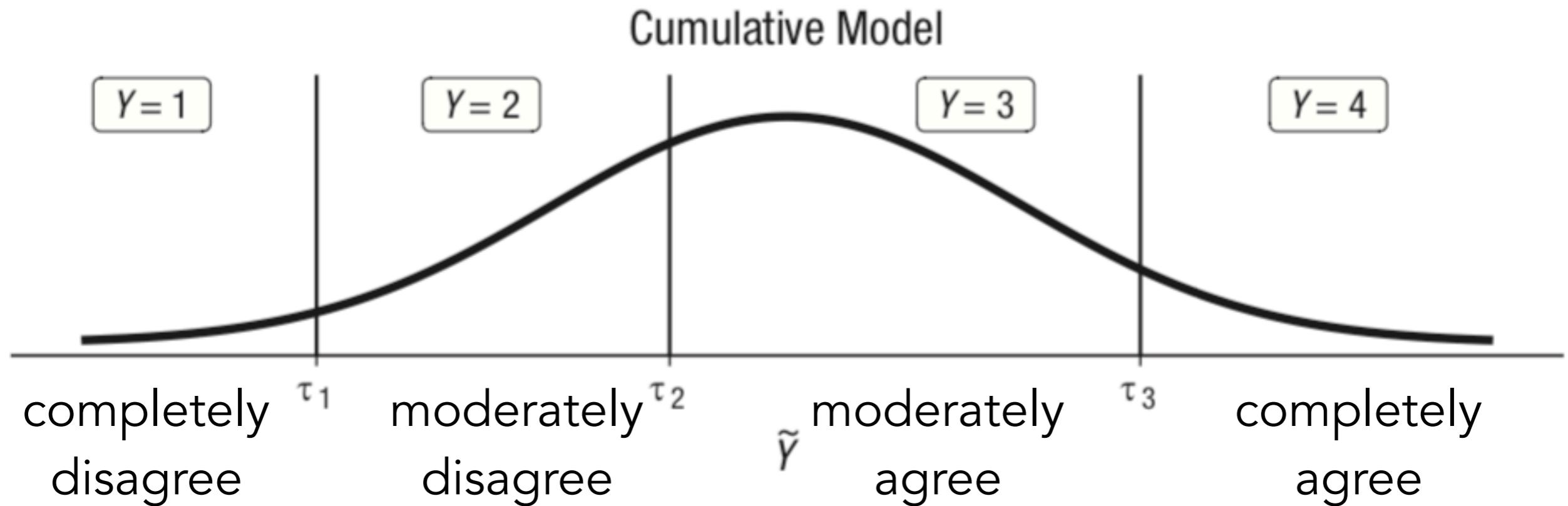
What to do?

- take averages across a number of items and hope that these are normally distributed
 - doesn't help
- use non-parametric approaches
 - only allows for answering very simple questions (e.g. difference between two groups)
- we want to use regression-like methods ...



to the rescue!

What shall we do?



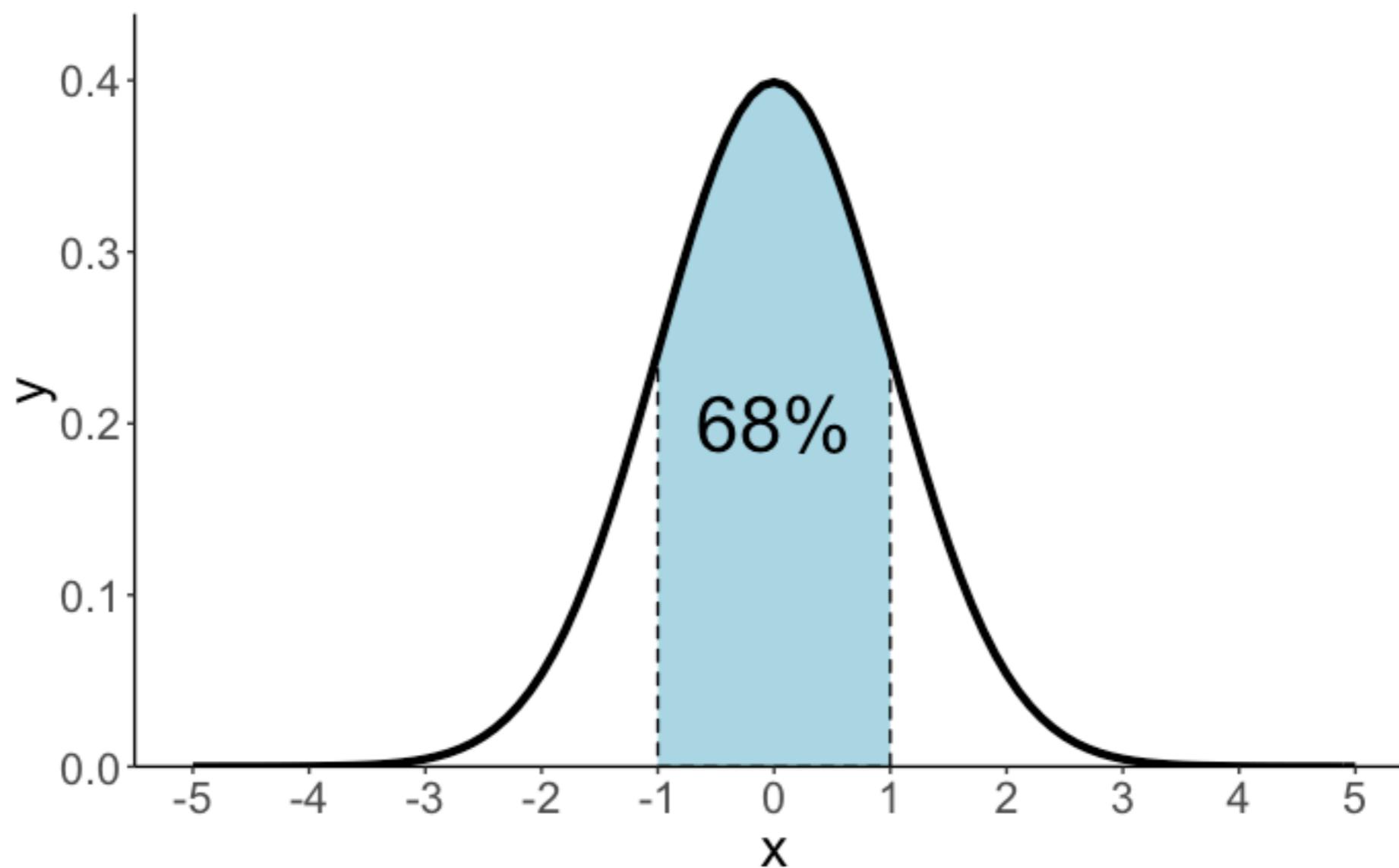
$$\Pr(Y = k) = \Phi(\tau_k - (b_1 x_1 + b_2 x_2)) - \Phi(\tau_{k-1} - (b_1 x_1 + b_2 x_2))$$

probability of k is the cumulative probability of k minus the cumulative probability of $k-1$

Computing probabilities

Find the probability between two values of interest:

$$\text{pnorm}(1) - \text{pnorm}(-1) = \\ 0.84 - 0.16 = 0.68$$



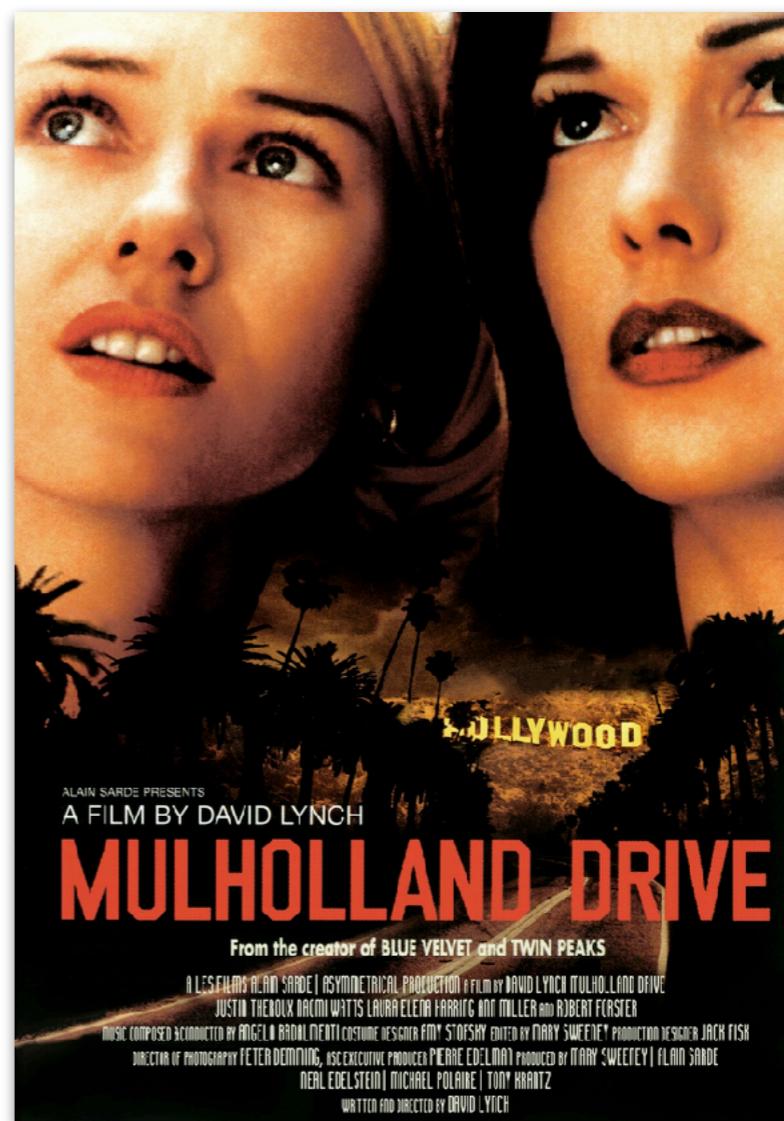
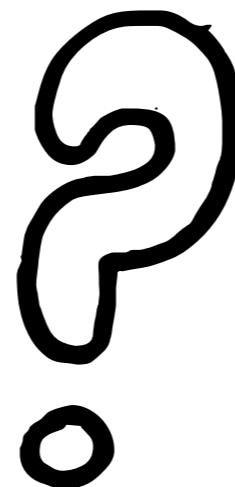
Which movie shall I watch?



Customer reviews

★★★★★ 4.5 out of 5

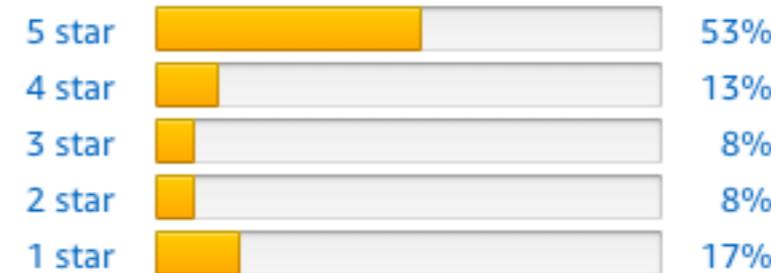
2,029 customer ratings



Customer reviews

★★★★☆ 3.8 out of 5

1,558 customer ratings



Amazon movie ratings

```
1 df.movies = read_csv(file = "data/MoviesData.csv")
2
3 df.movies = df.movies %>%
4   pivot_longer(cols = n1:n5,
5                 names_to = "stars",
6                 values_to = "rating") %>%
7   mutate(stars = str_remove(stars, "n"),
8           stars = as.numeric(stars))
```

ID	Descrip	stars	rating
1	The Whole Truth	1	49
1	The Whole Truth	2	70
1	The Whole Truth	3	119
1	The Whole Truth	4	217
1	The Whole Truth	5	245
2	Priceless	1	67
2	Priceless	2	22
2	Priceless	3	22
2	Priceless	4	60
2	Priceless	5	574

Amazon movie ratings

```
1 df.movies = read_csv(file = "data/MoviesData.csv")
2
3 df.movies = df.movies %>%
4   pivot_longer(cols = n1:n5,
5                 names_to = "stars",
6                 values_to = "rating") %>%
7   mutate(stars = str_remove(stars, "n"),
8           stars = as.numeric(stars)) %>%
9   uncount(weights = rating) %>%
10  mutate(id = as.factor(ID)) %>%
11  filter(ID <= 6)
```

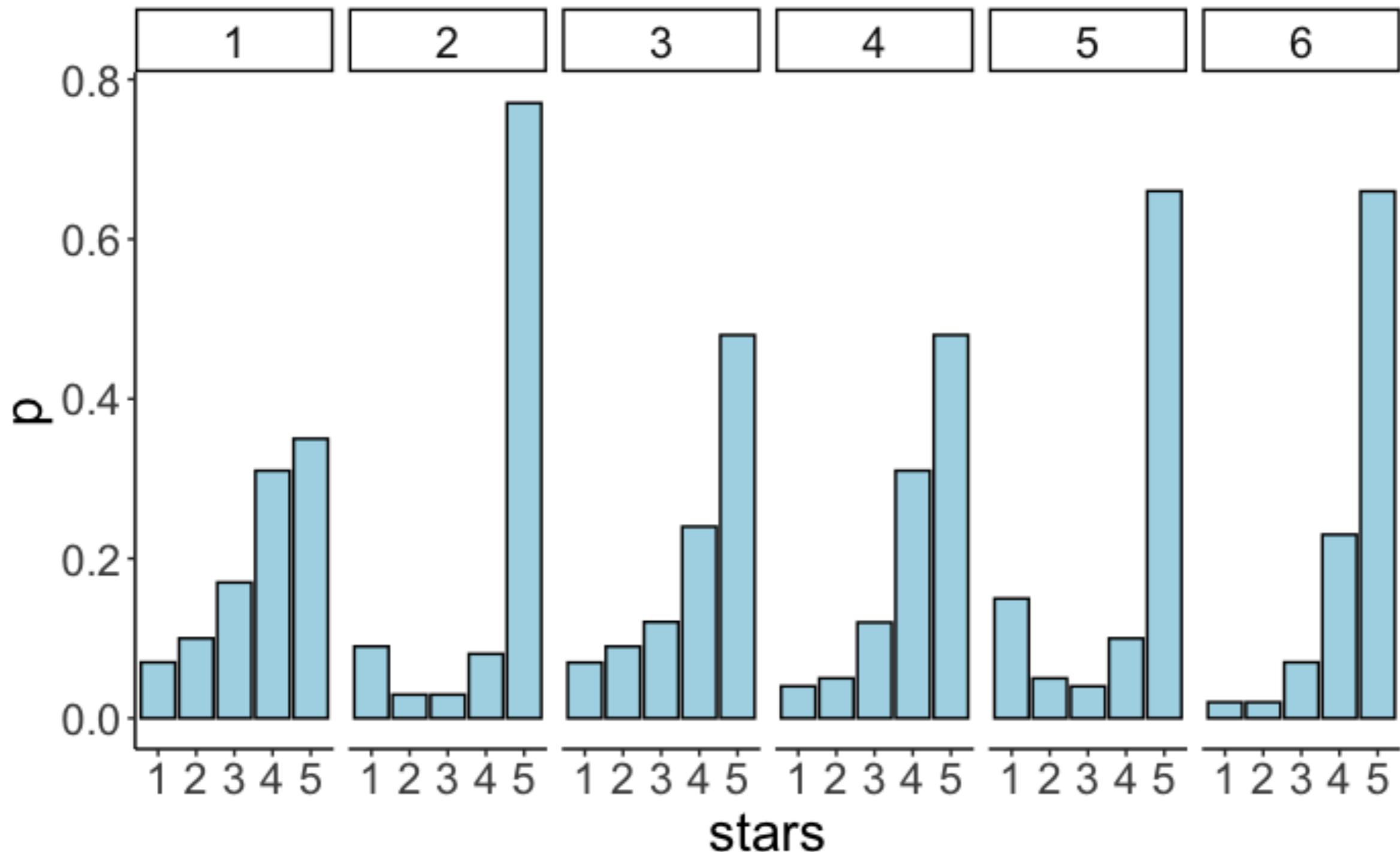
• very cool function!

ID	Descrip	stars	rating
1	The Whole Truth	1	49
1	The Whole Truth	2	70
1	The Whole Truth	3	119
1	The Whole Truth	4	217
1	The Whole Truth	5	245

`uncount()`



Amazon movie ratings



these aren't normally distributed ..

Fit the ordinal regression model

```
1 fit.brm5 = brm(formula = stars ~ 1 + id,  
2                   family = cumulative(link = "probit"),  
3                   data = df.movies,  
4                   file = "cache/brm5",  
5                   seed = 1)
```

linking function

thresholds

difference in
mean to
reference
category

```
Family: cumulative  
Links: mu = probit; disc = identity  
Formula: stars ~ 1 + id  
Data: df.movies (Number of observations: 21708)  
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;  
         total post-warmup samples = 4000
```

Population-Level Effects:

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-1.22	0.04	-1.31	-1.14	1.00	1877	2488
Intercept[2]	-0.90	0.04	-0.98	-0.82	1.00	1787	2419
Intercept[3]	-0.44	0.04	-0.52	-0.36	1.00	1692	2185
Intercept[4]	0.32	0.04	0.24	0.40	1.00	1634	2101
id2	0.84	0.06	0.71	0.96	1.00	2354	2553
id3	0.22	0.05	0.11	0.32	1.00	2146	2516
id4	0.33	0.04	0.24	0.41	1.00	1647	2315
id5	0.44	0.05	0.34	0.54	1.00	1982	2608
id6	0.75	0.04	0.67	0.83	1.00	1659	2158

Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

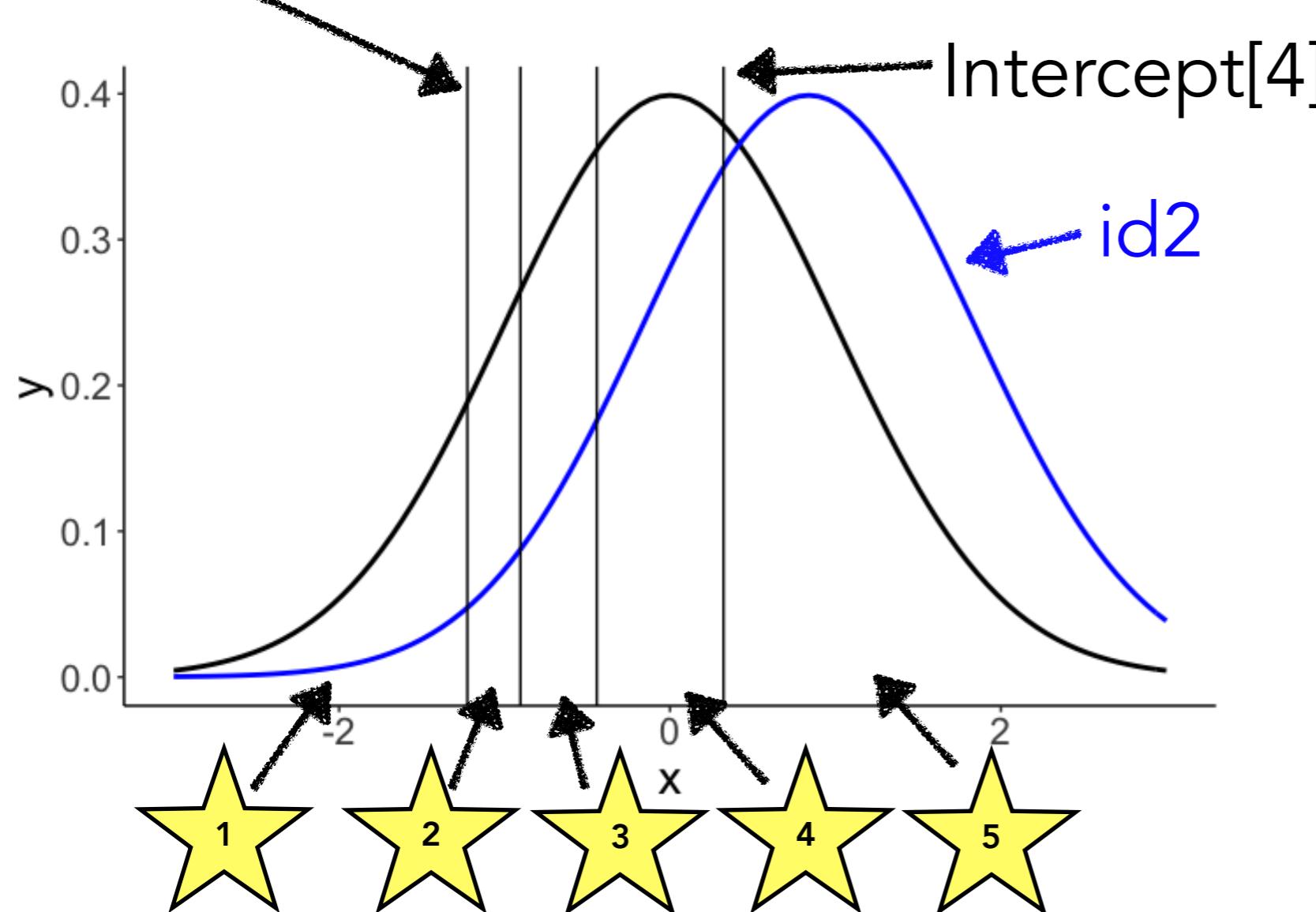
Interpretation

```
Family: cumulative  
Links: mu = probit; disc = identity  
Formula: stars ~ 1 + id  
Data: df.movies (Number of observations: 21708)  
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;  
total post-warmup samples = 4000
```

Population-Level Effects:

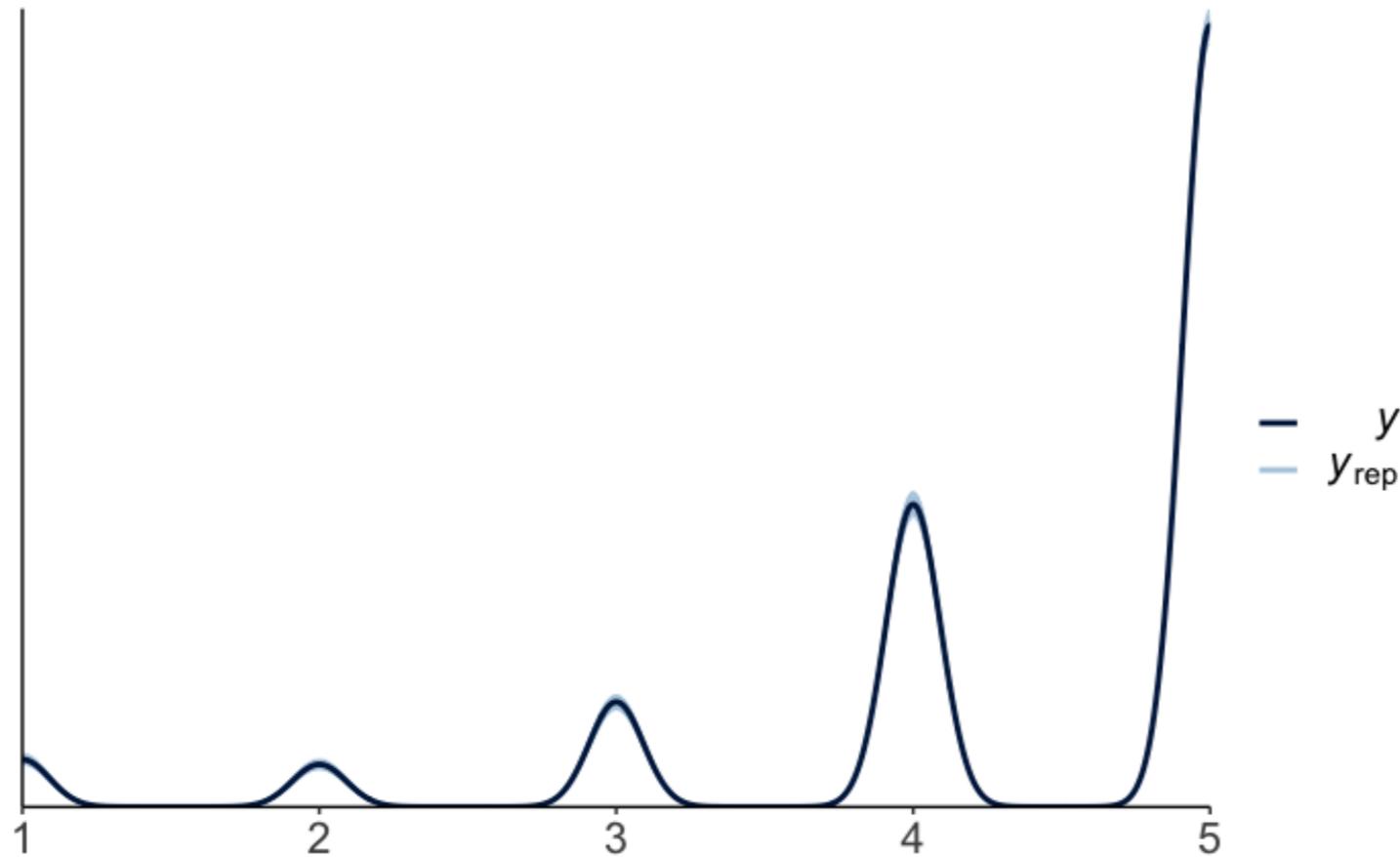
	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-1.22	0.04	-1.31	-1.14	1.00	1877	2488
Intercept[2]	-0.90	0.04	-0.98	-0.82	1.00	1787	2419
Intercept[3]	-0.44	0.04	-0.52	-0.36	1.00	1692	2185
Intercept[4]	0.32	0.04	0.24	0.40	1.00	1634	2101
id2	0.84	0.06	0.71	0.96	1.00	2354	2553
id3	0.22	0.05	0.11	0.32	1.00	2146	2516
id4	0.33	0.04	0.24	0.41	1.00	1647	2315
id5	0.44	0.05	0.34	0.54	1.00	1982	2608
id6	0.75	0.04	0.67	0.83	1.00	1659	2158

Intercept[1]



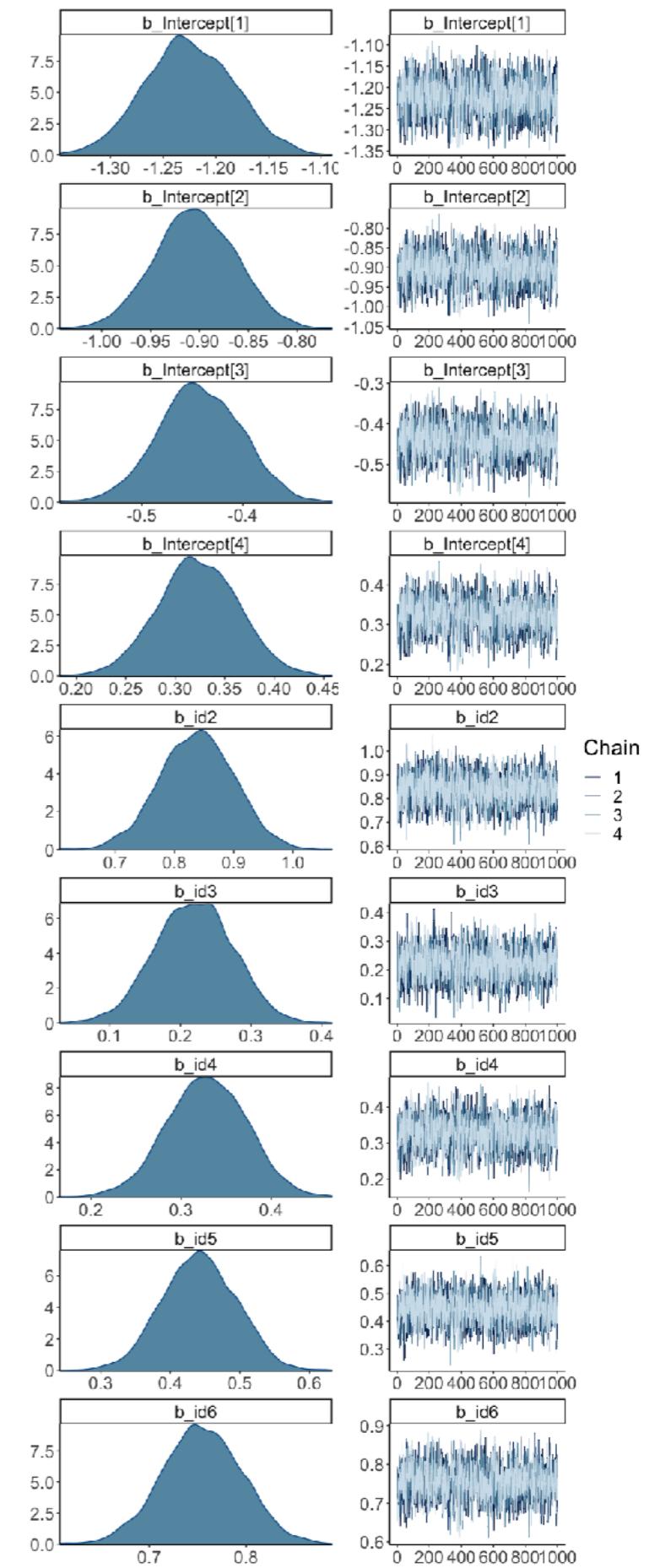
Check the model

```
1 fit.brms %>%  
2   plot(N = 9)
```

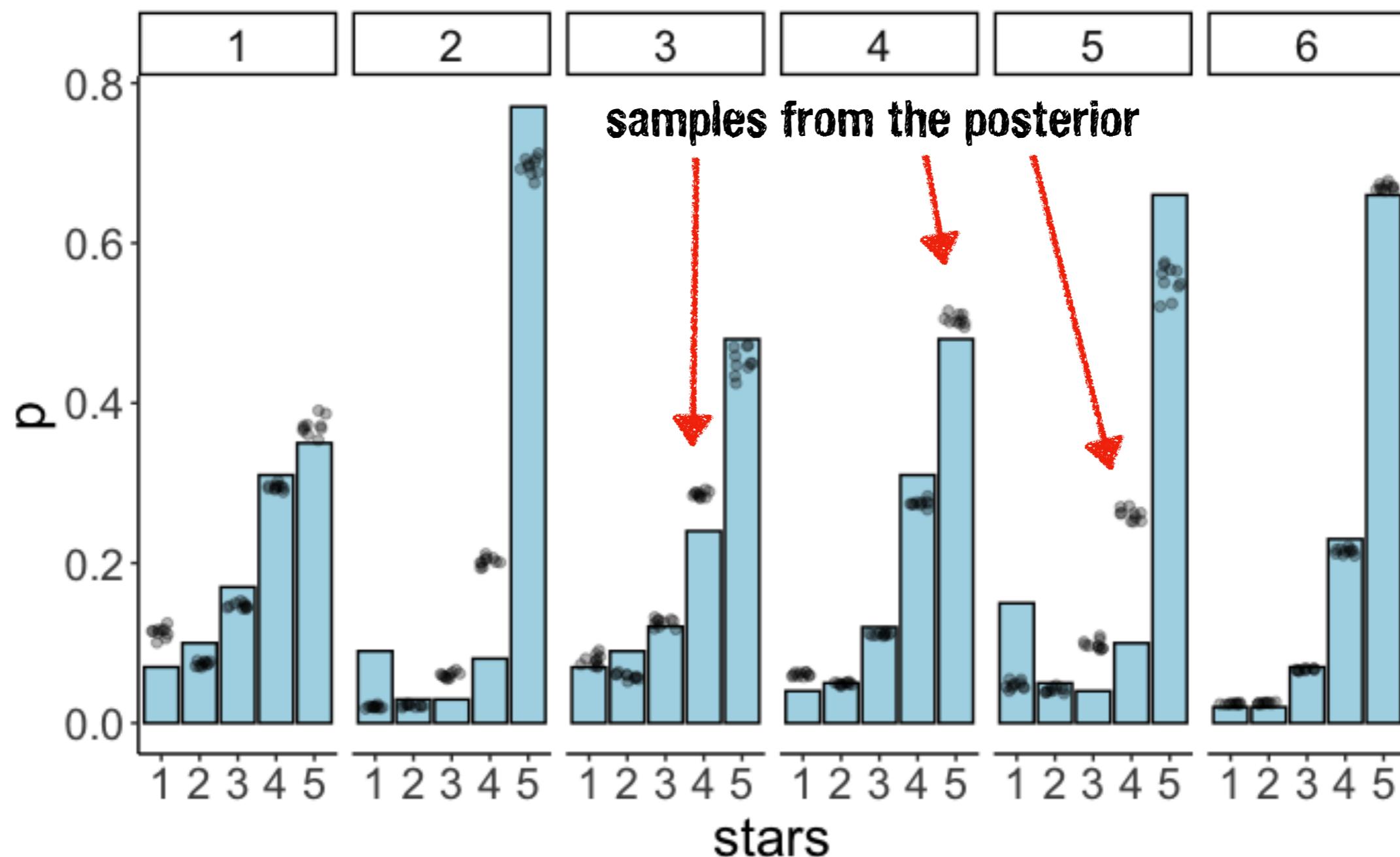


looking good!

```
1 fit.brms %>%  
2   plot(N = 9)
```



Illustrate the predictions



**predictions look pretty good but
maybe we can do better?**

Before doing better, let's do worse!

```
1 fit.brm6 = brm(formula = stars ~ 1 + id,  
2                   data = df.movies,  
3                   file = "cache/brm6",  
4                   seed = 1)
```

Family: gaussian
Links: mu = identity; sigma = identity
Formula: stars ~ 1 + id
Data: df.movies (Number of observations: 21708)
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
total post-warmup samples = 4000

Population-Level Effects:

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	3.77	0.04	3.70	3.84	1.00	1203	1621
id2	0.64	0.05	0.54	0.75	1.00	1605	2335
id3	0.20	0.05	0.10	0.30	1.00	1558	2147
id4	0.37	0.04	0.29	0.45	1.00	1267	1862
id5	0.30	0.05	0.21	0.40	1.00	1441	2154
id6	0.72	0.04	0.65	0.79	1.00	1205	1720

Family Specific Parameters:

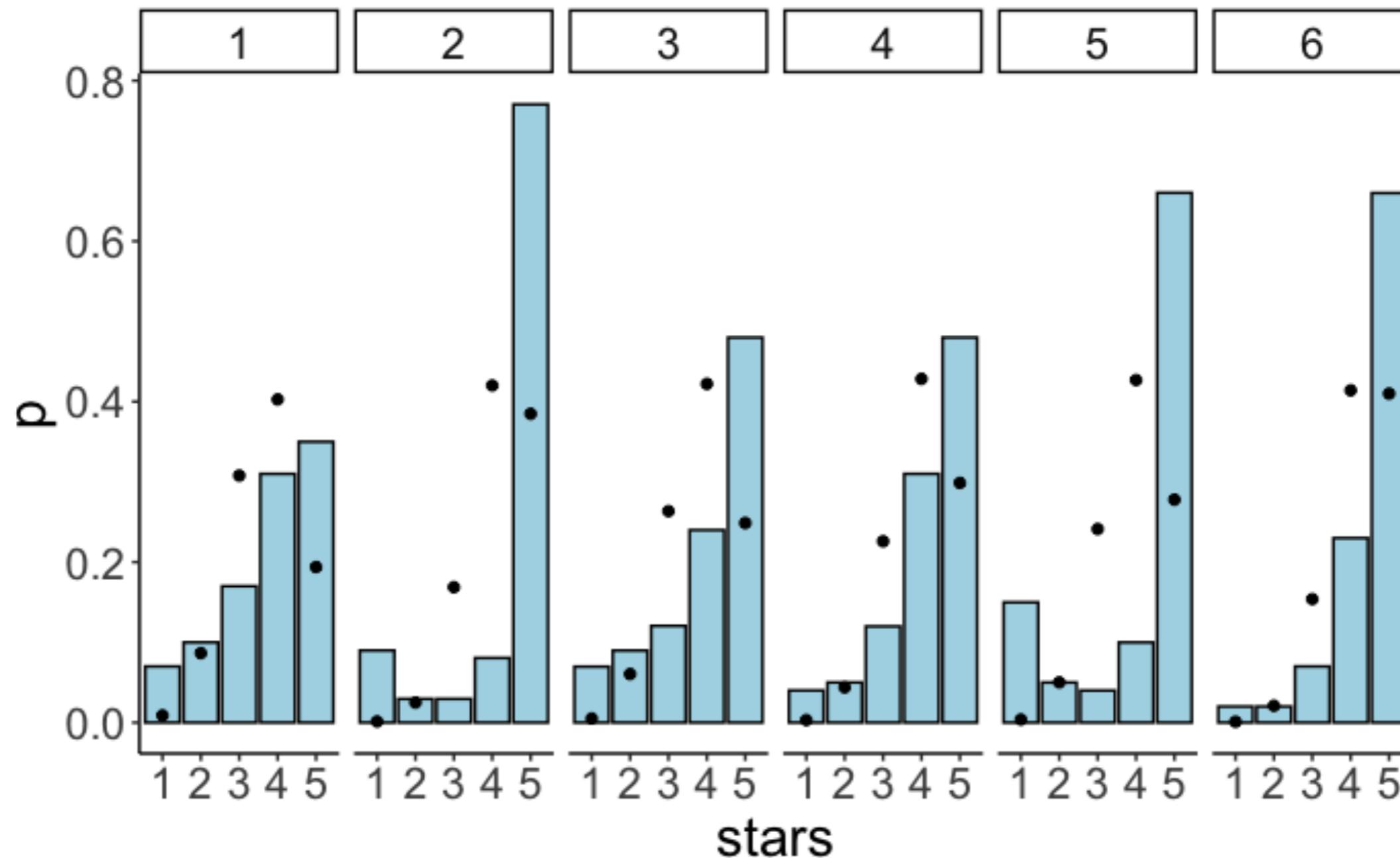
	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	1.00	0.00	0.99	1.01	1.00	3300	2723

Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

mean for the reference category

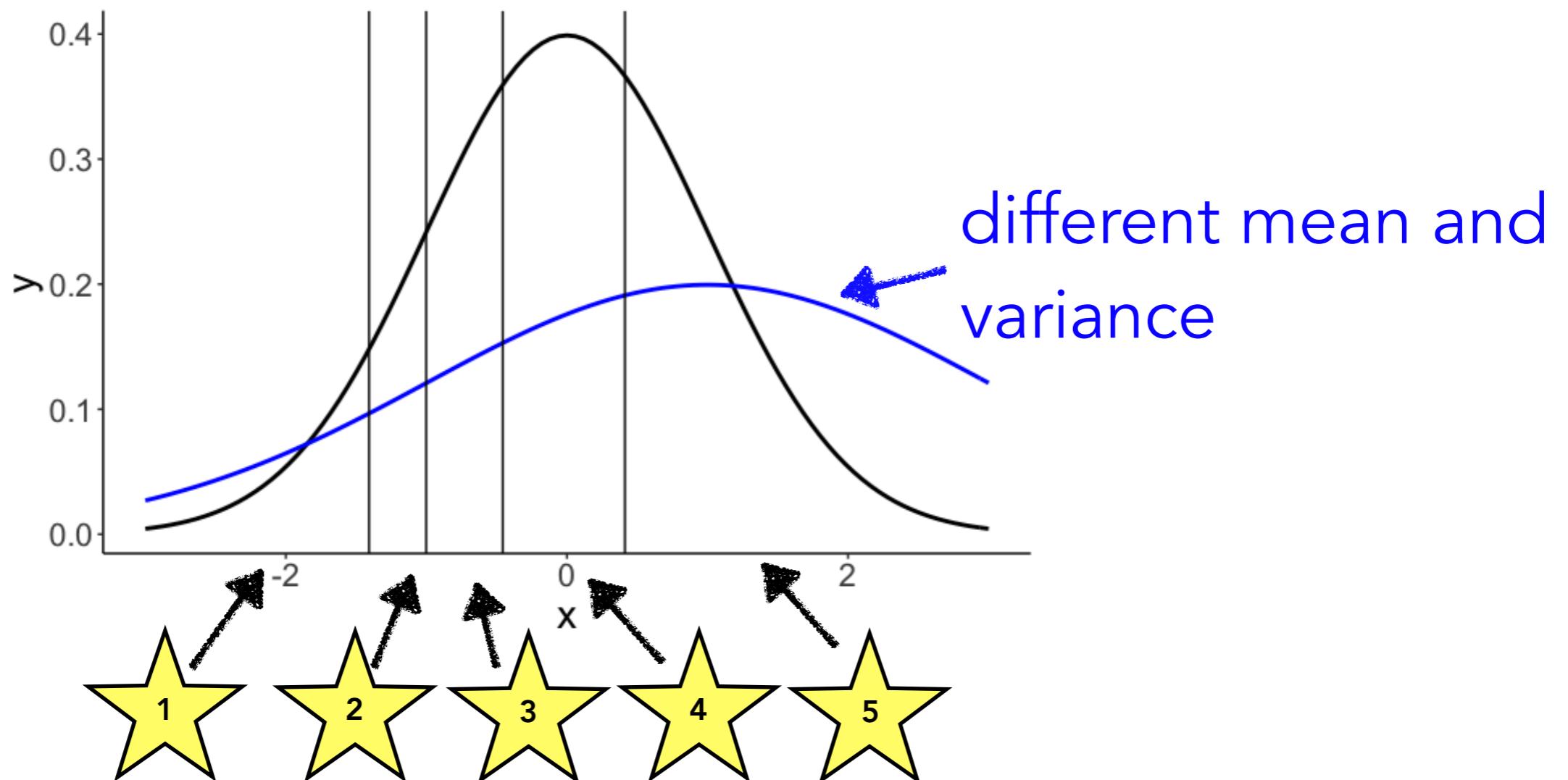
difference in mean to reference category

Predictions from the metric model



that's not looking good ...

Let's relax the assumption of equal variances



Fit a model that doesn't assume equal variance

```
1 fit.brm7 = brm(formula = bf(stars ~ 1 + id) + lf(disc ~ 0 + id, cmc = FALSE),  
2   family = cumulative(link = "probit"),  
3   data = df.movies,  
4   file = "cache/brm7",  
5   seed = 1)
```

tricky formula

thresholds

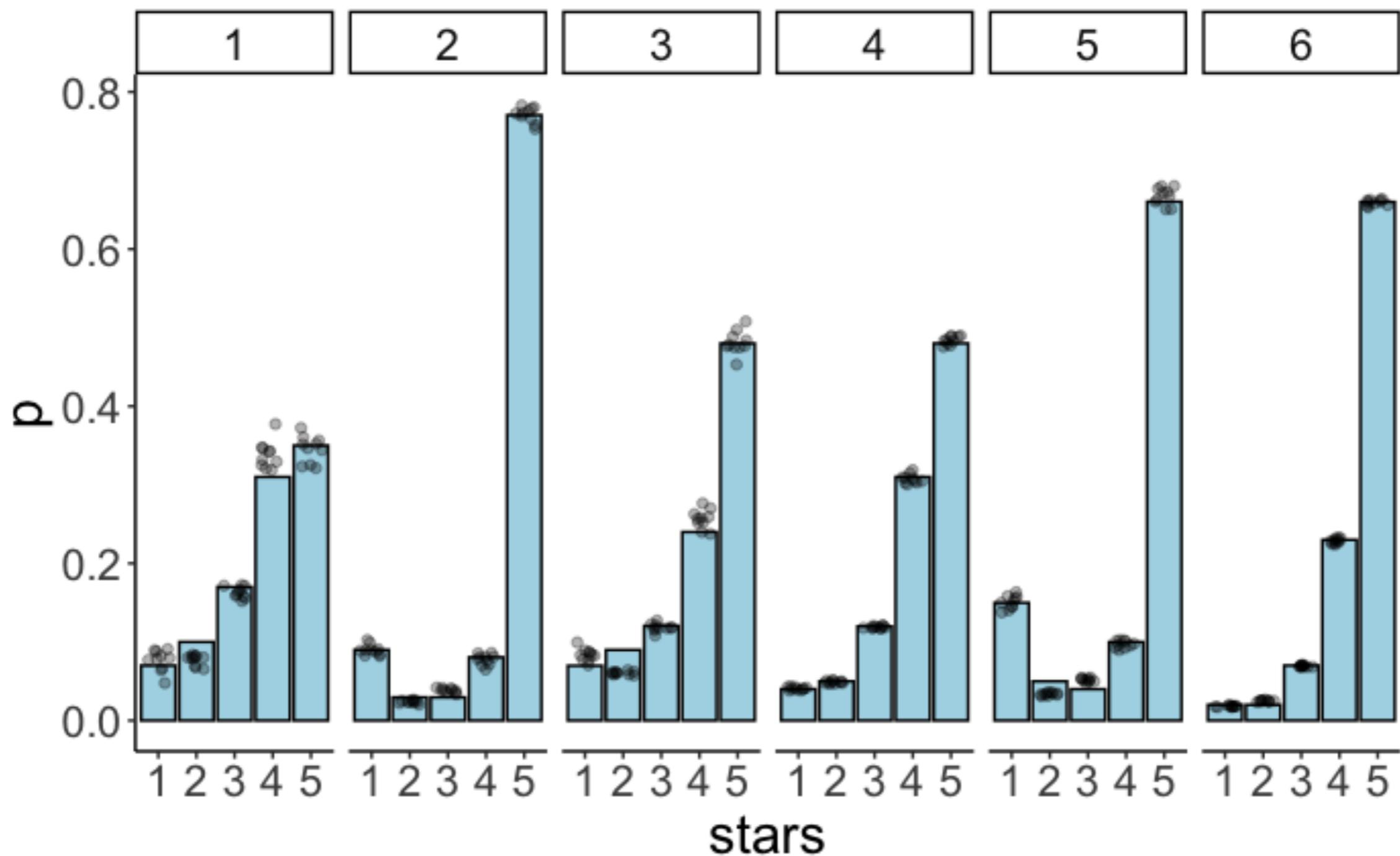
difference in
mean

difference in
variance

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept[1]	-1.41	0.06	-1.53	-1.29	1.00	1484	2421
Intercept[2]	-1.00	0.05	-1.10	-0.90	1.00	1852	2561
Intercept[3]	-0.46	0.04	-0.54	-0.37	1.00	2405	2684
Intercept[4]	0.41	0.05	0.32	0.51	1.00	1336	2161
id2	2.71	0.33	2.14	3.44	1.00	1681	1865
id3	0.33	0.07	0.20	0.47	1.00	1961	2618
id4	0.36	0.05	0.26	0.46	1.00	1525	2753
id5	1.65	0.17	1.34	2.00	1.00	1929	2281
id6	0.86	0.06	0.74	0.98	1.00	1112	1769
disc_id2	-1.12	0.10	-1.33	-0.94	1.00	1672	1943
disc_id3	-0.23	0.06	-0.34	-0.11	1.00	1342	1955
disc_id4	-0.01	0.04	-0.09	0.07	1.00	1043	1747
disc_id5	-1.09	0.07	-1.23	-0.95	1.00	1681	1996
disc_id6	-0.08	0.04	-0.15	0.00	1.00	941	1489

see for
details

Illustrate the predictions

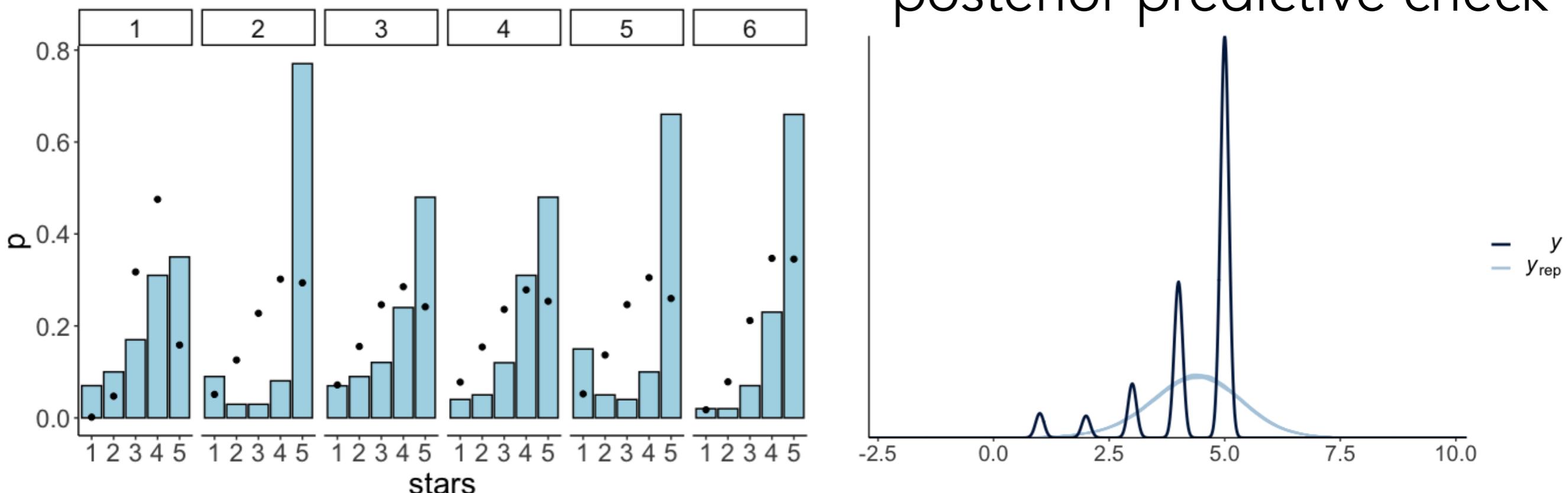


this looks excellent!

Metric model with different variances still doesn't work well ...

```
1 fit.brms = brm(formula = bf(stars ~ 1 + id,  
2                               sigma ~ 1 + id),  
3 data = df.movies,  
4 file = "cache/brms",  
5 seed = 1)
```

posterior predictive check



still no good ...

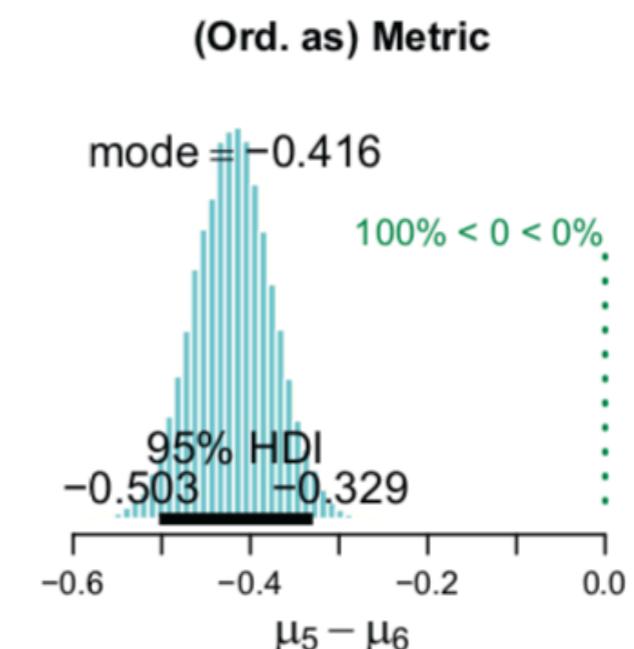
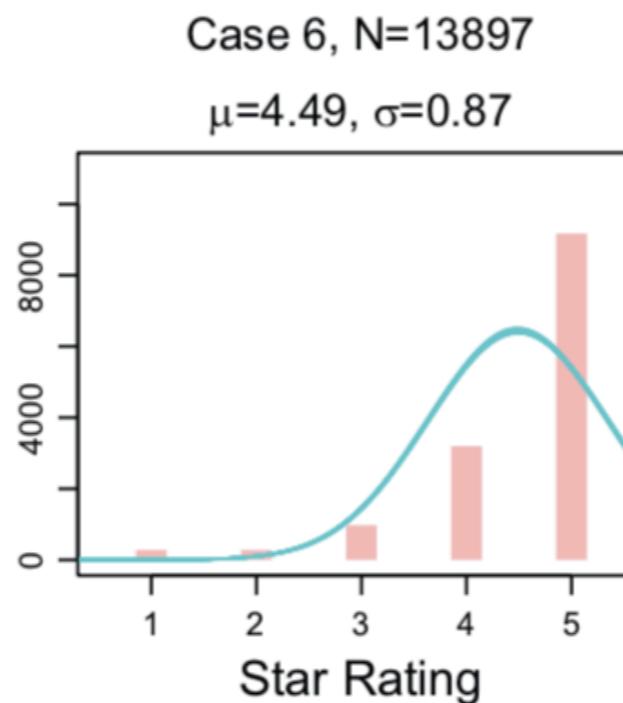
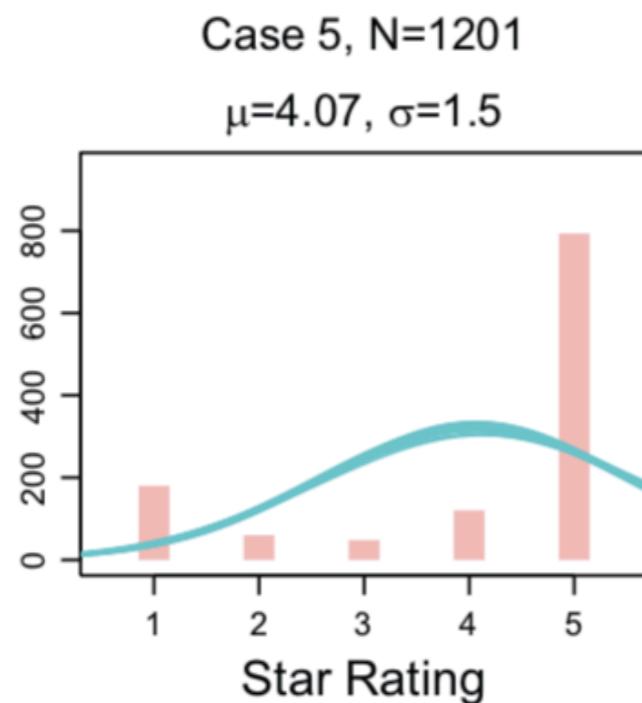
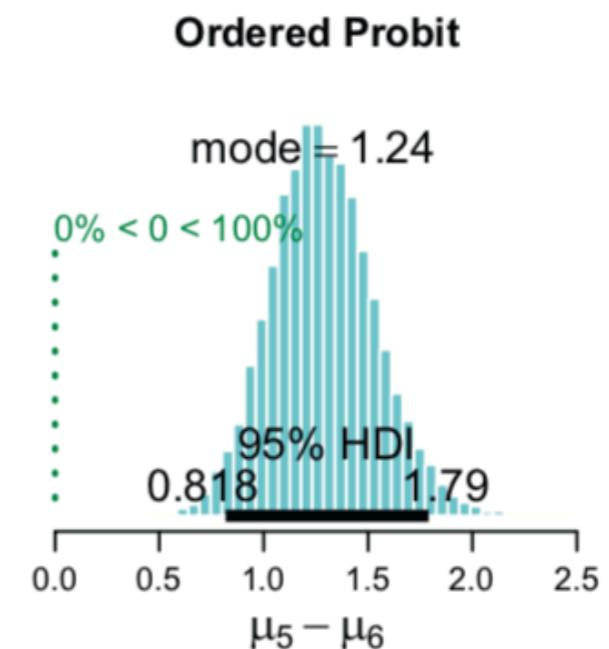
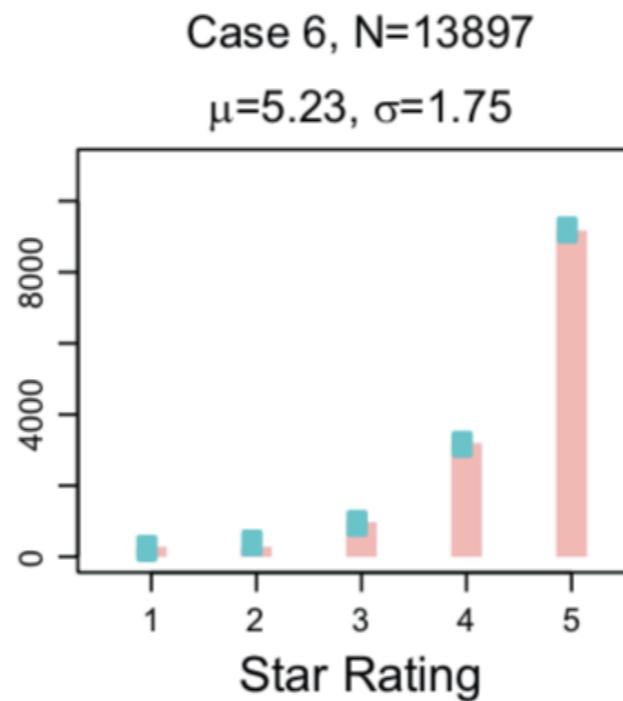
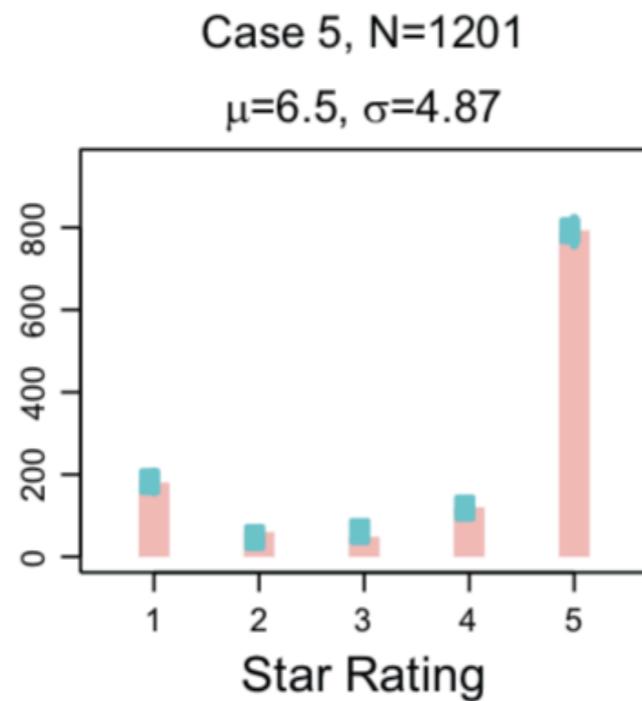
Model comparison

```
1 fit.brm5 = add_criterion(fit.brm5,  
2                             criterion = "loo")  
3  
4 fit.brm6 = add_criterion(fit.brm6,  
5                             criterion = "loo")  
6  
7 loo_compare(fit.brm5, fit.brm6)
```

	elpd_diff	se_diff
fit.brm5	0.0	0.0
fit.brm6	-7657.1	109.2

the ordinal regression model is muuuuuuch better!

Which movie shall I watch?



run an ordinal regression to find out!

Plan for today

- Evidence for the "null hypothesis"
- Dealing with unequal variance
- Better modeling slider data
 - ZOIB: Zero-one inflated beta binomial model
- Better modeling Likert scale data
 - Ordinal regression

Feedback

How was the pace of today's class?

much a little just a little much
too too right too too
slow slow

How happy were you with today's class overall?



What did you like about today's class? What could be improved next time?

Thank you!