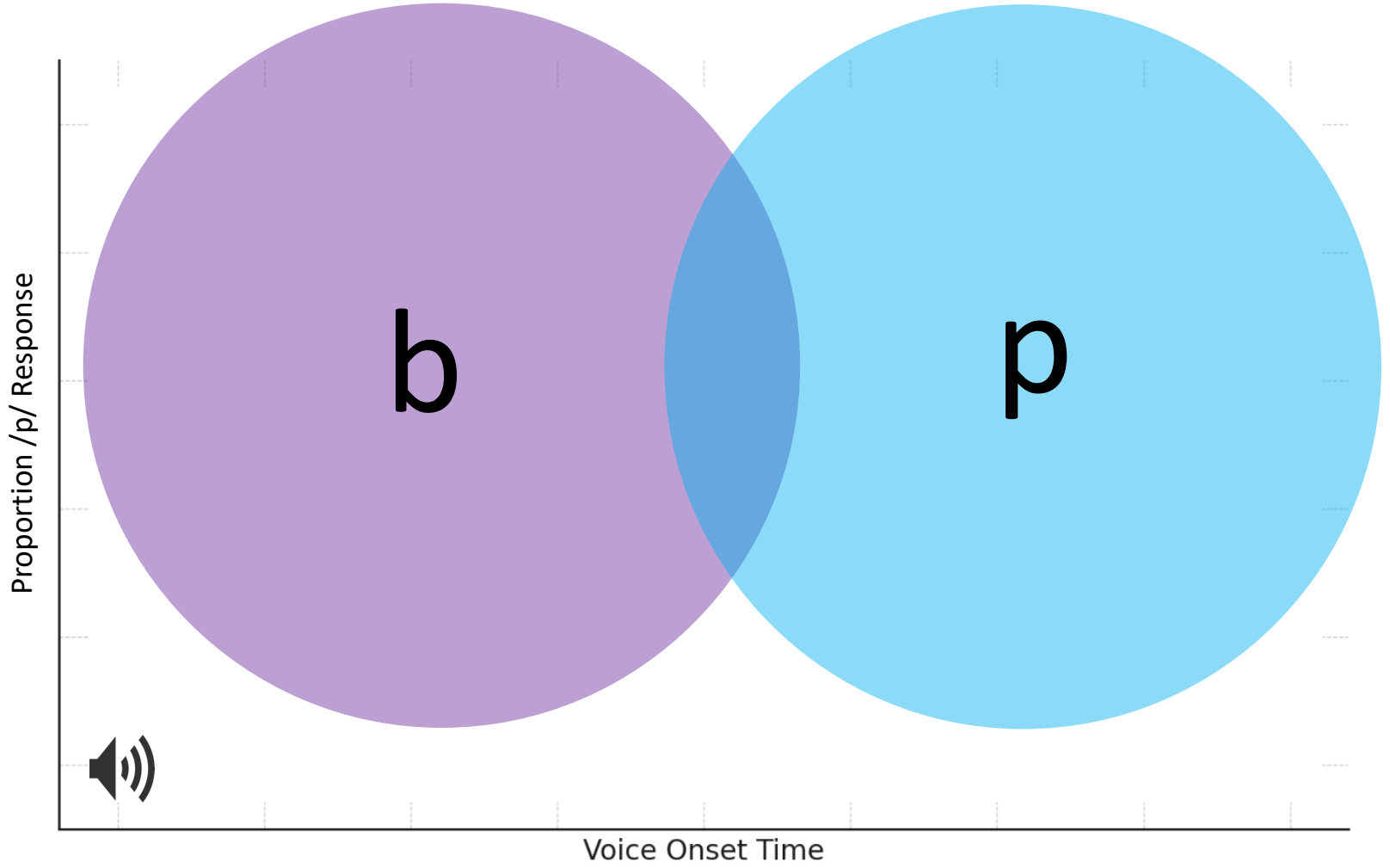
Logistic Regression

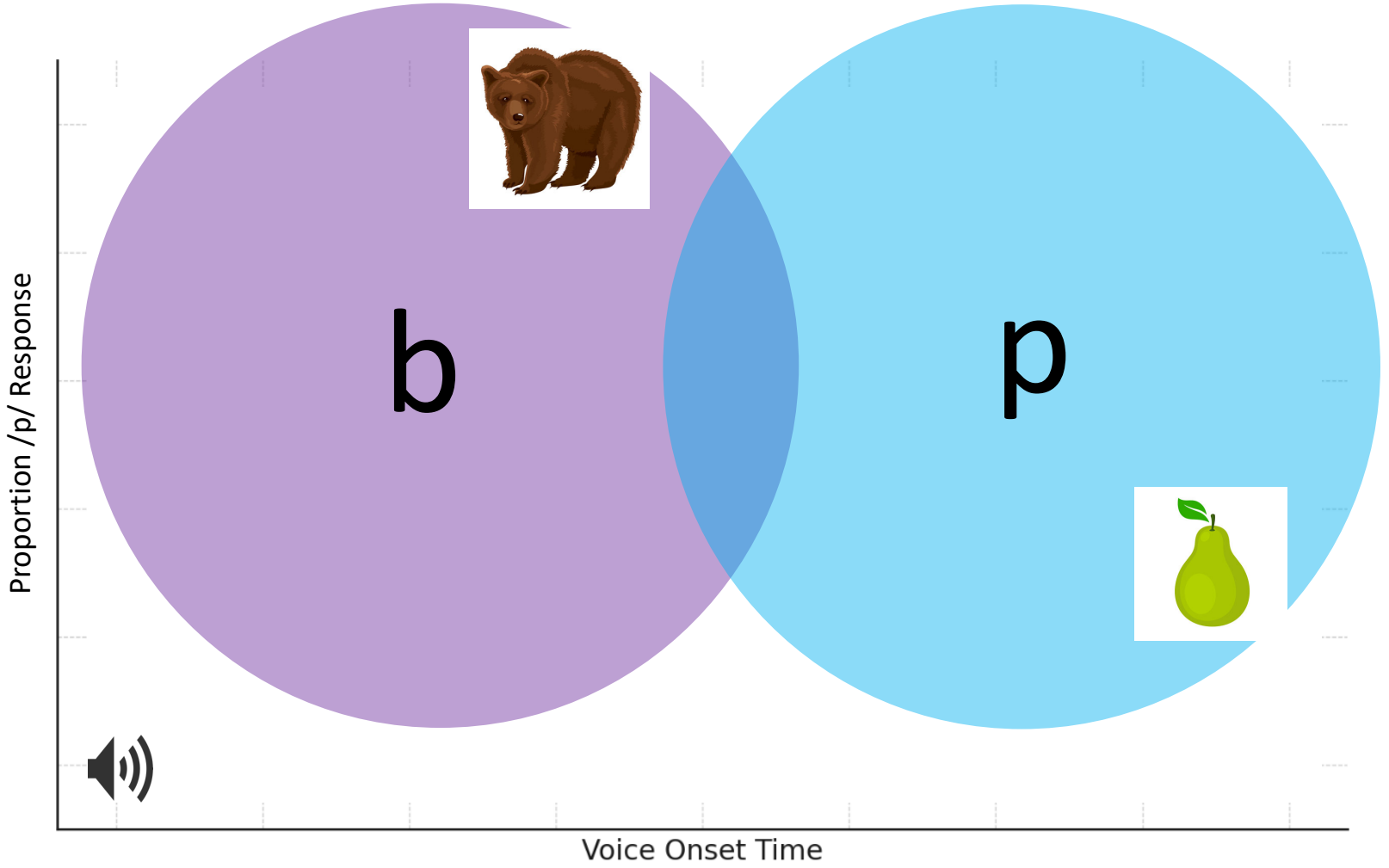
Abby Fergus

Why logistic regression?

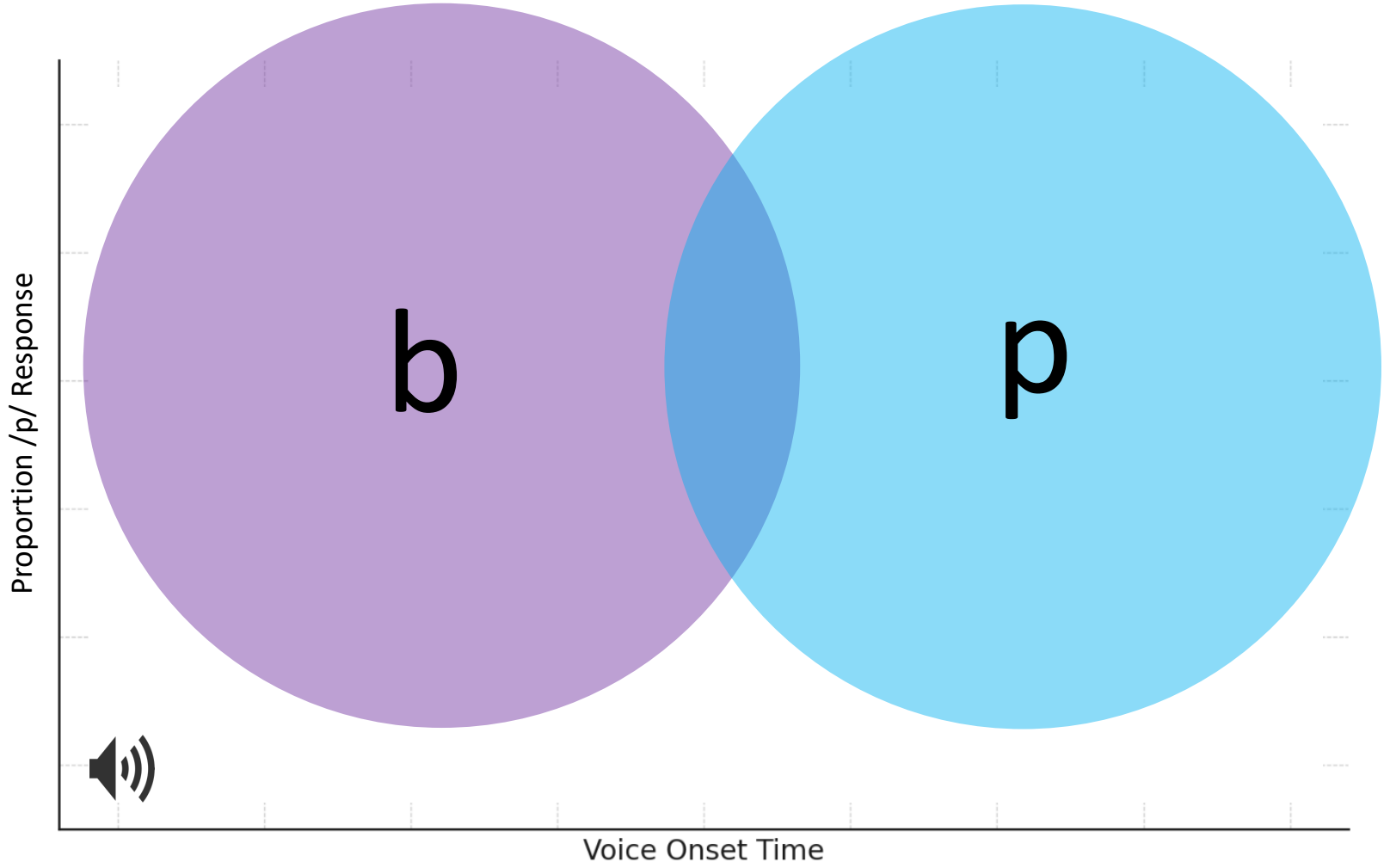
Why logistic regression?



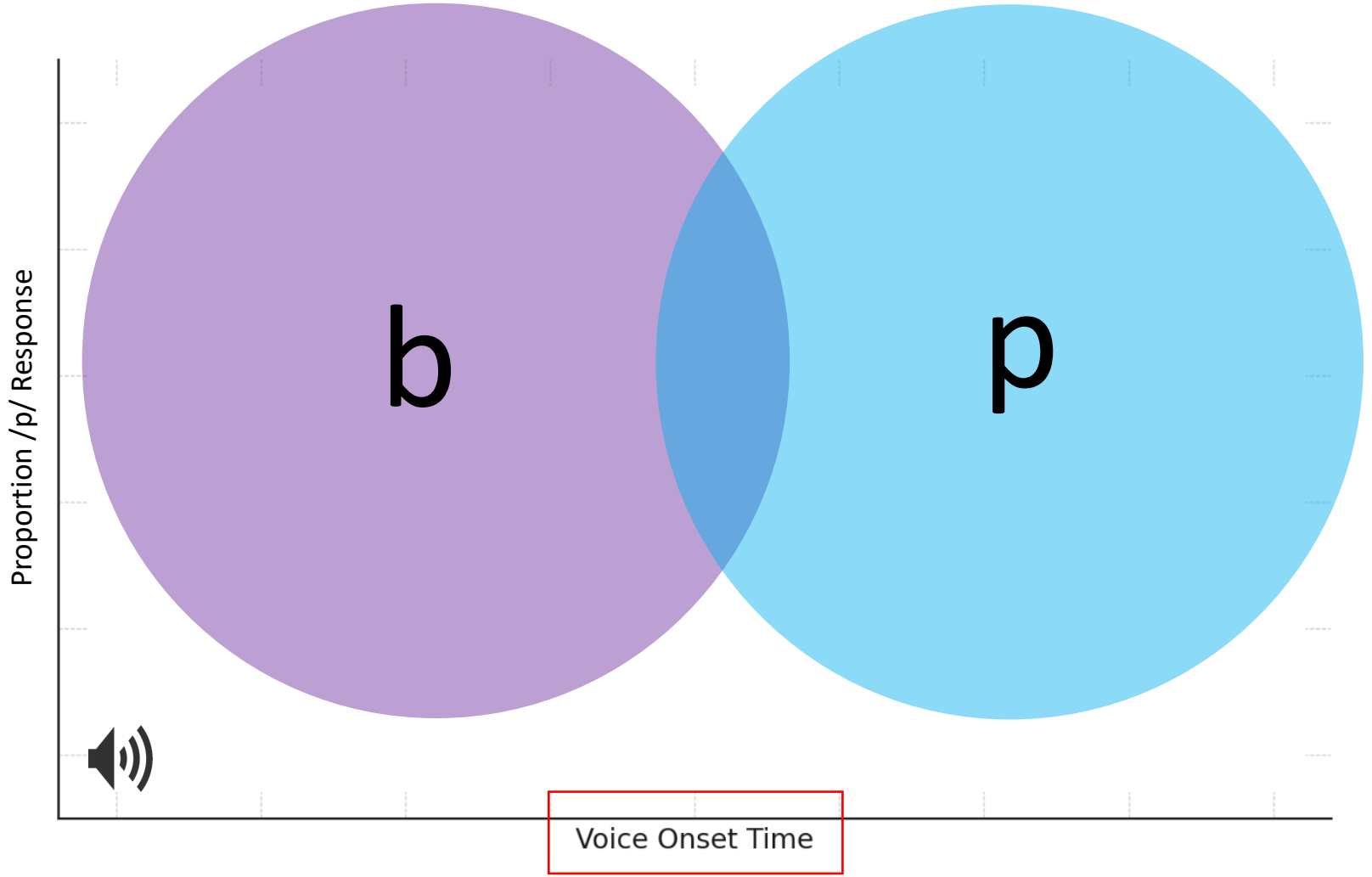
Why logistic regression?



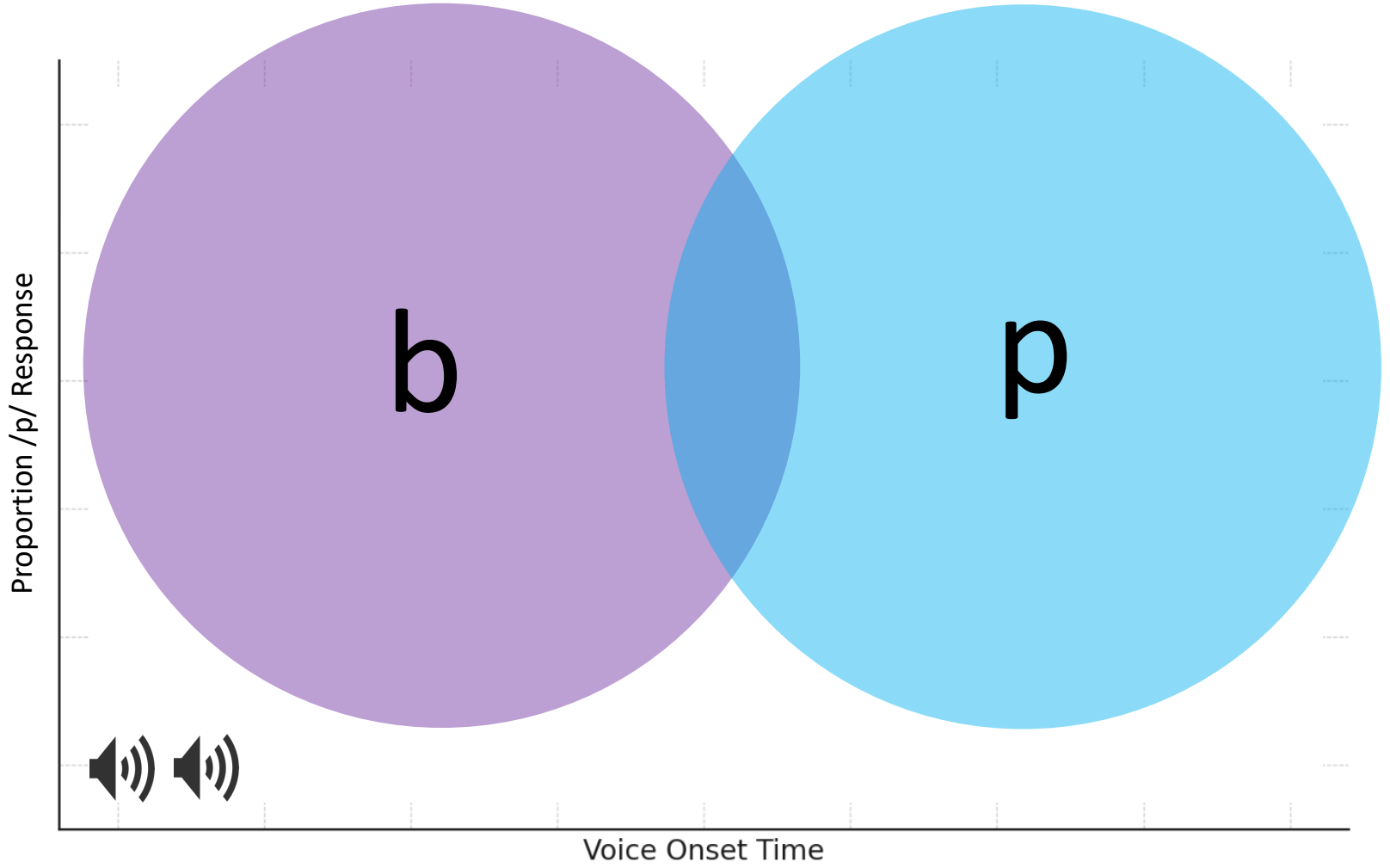
Why logistic regression?



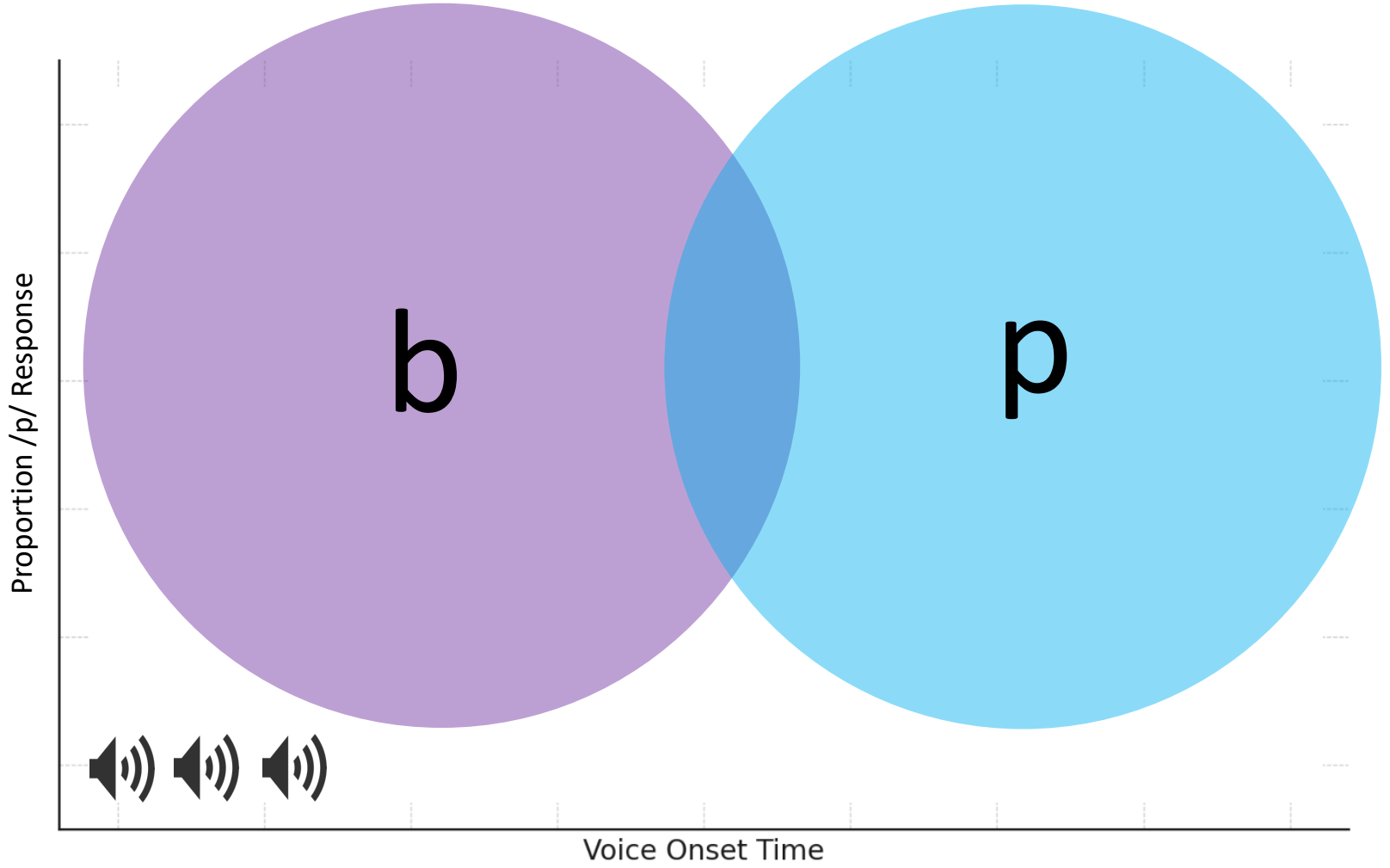
Why logistic regression?



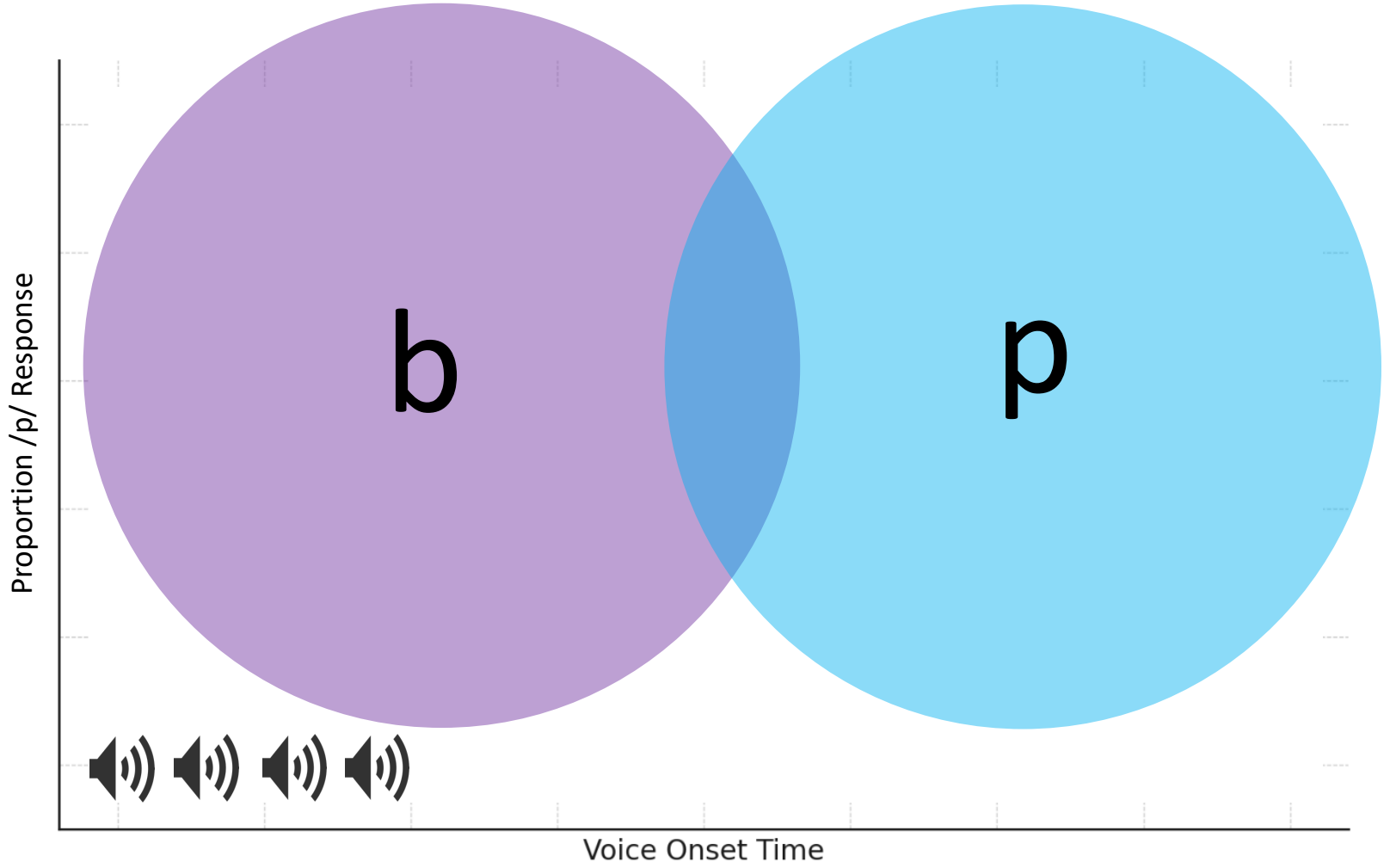
Why logistic regression?



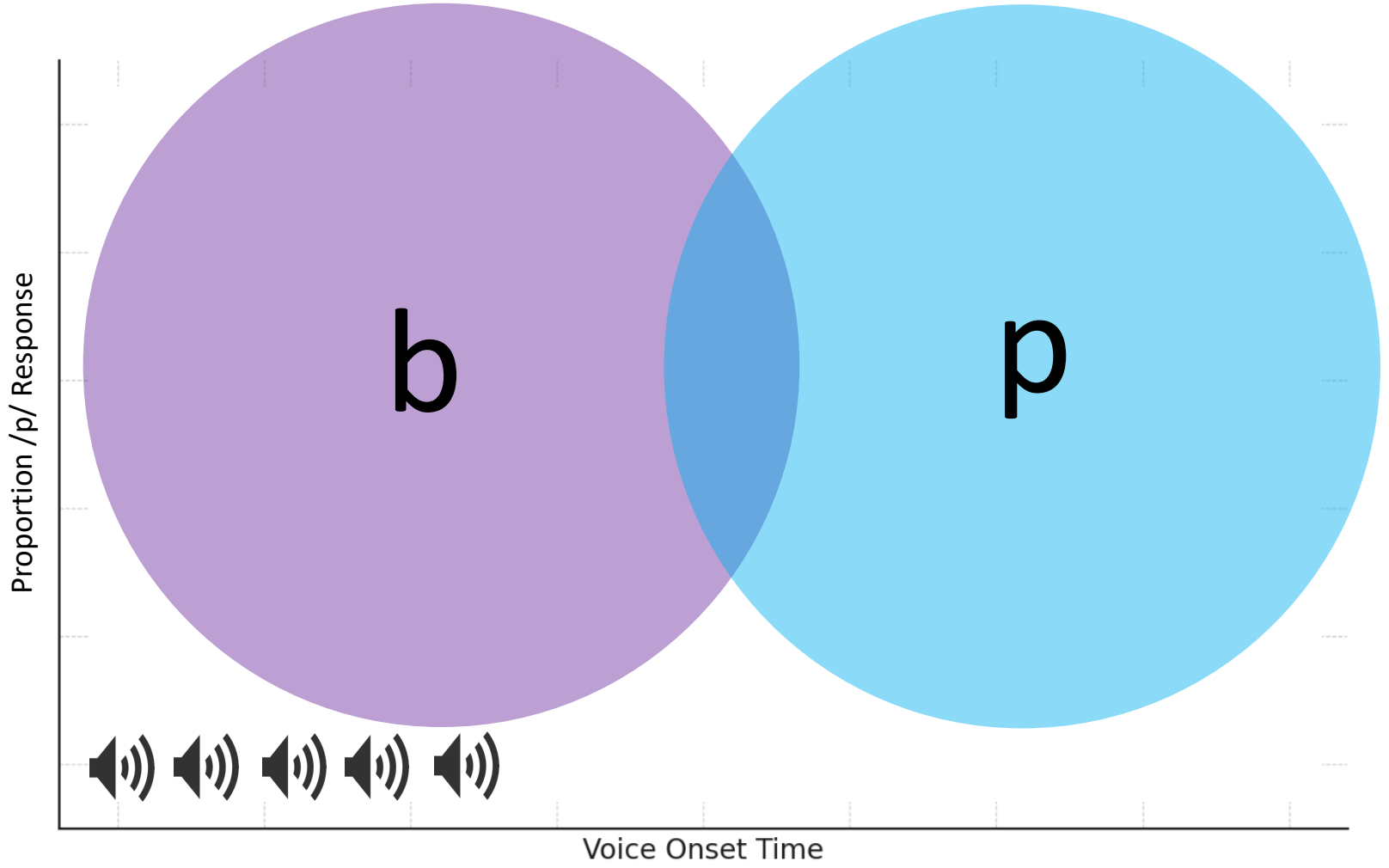
Why logistic regression?



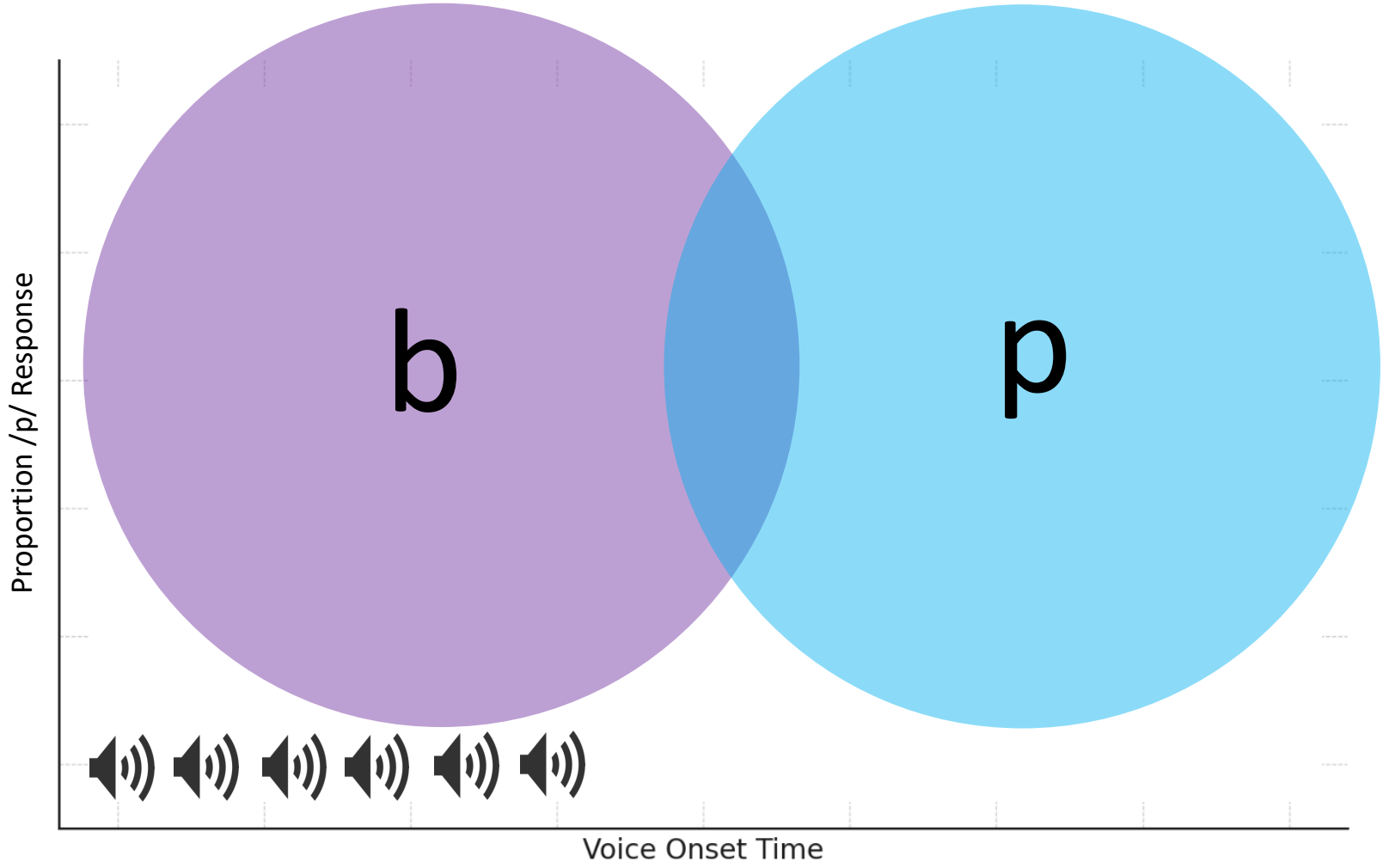
Why logistic regression?



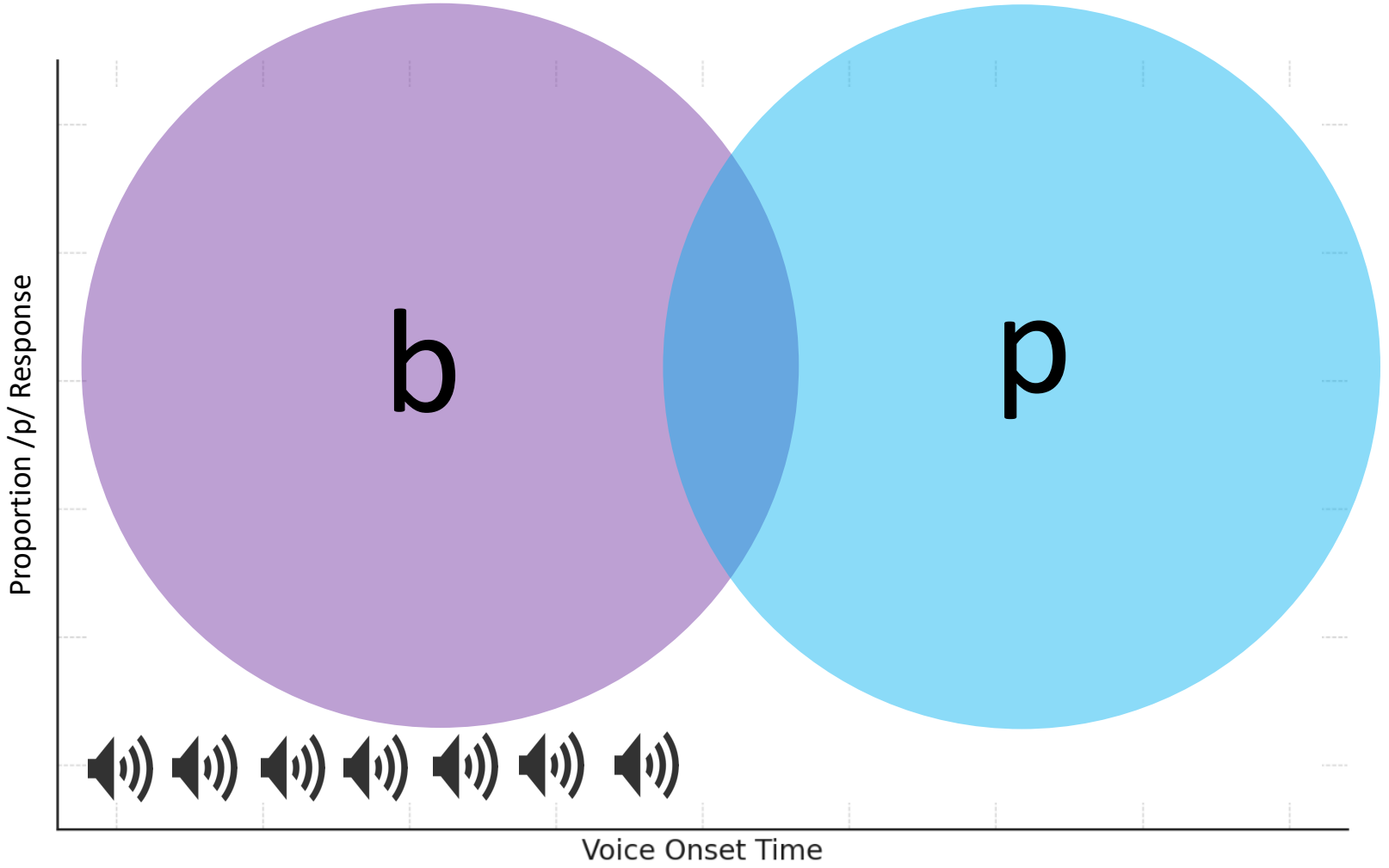
Why logistic regression?



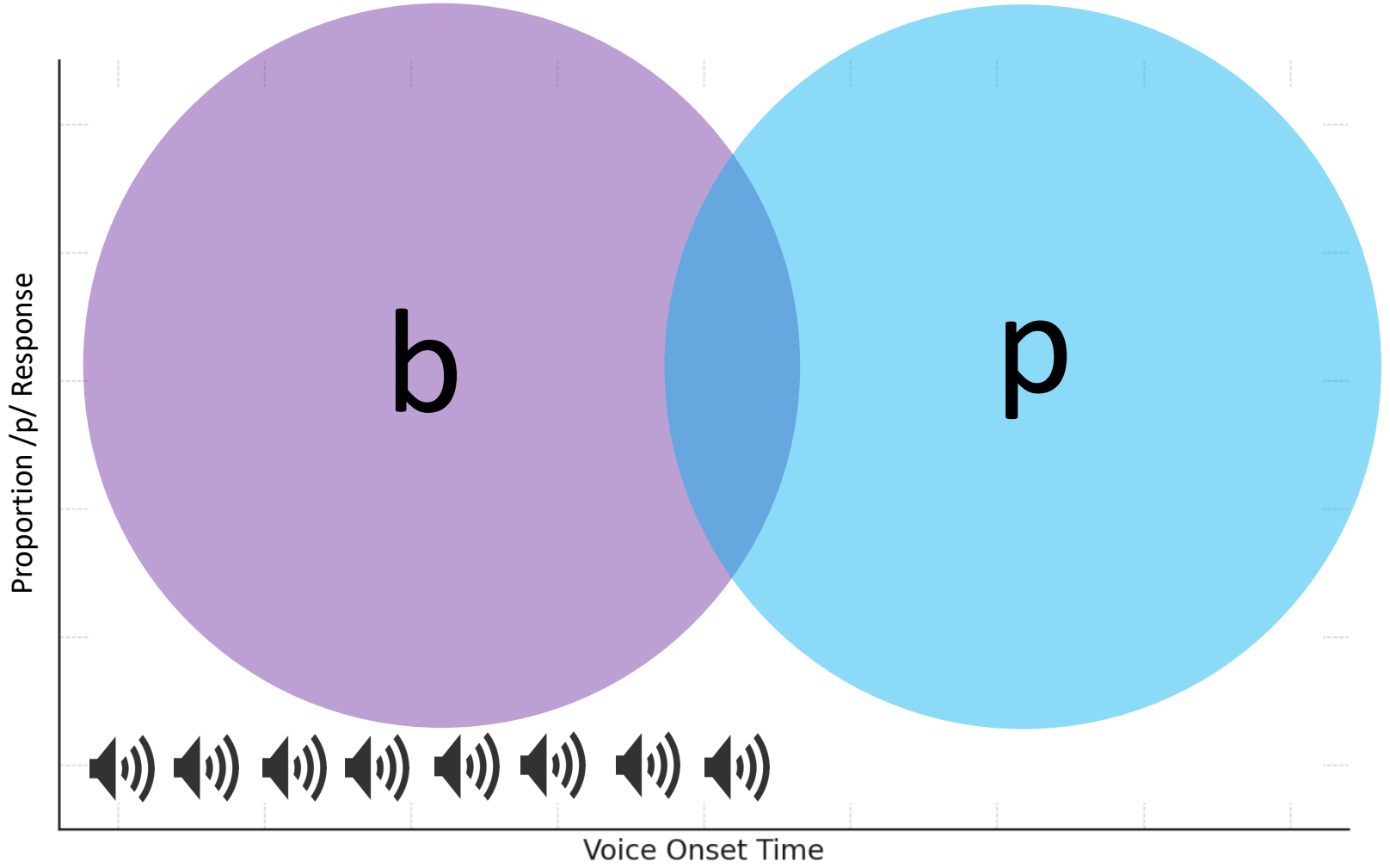
Why logistic regression?



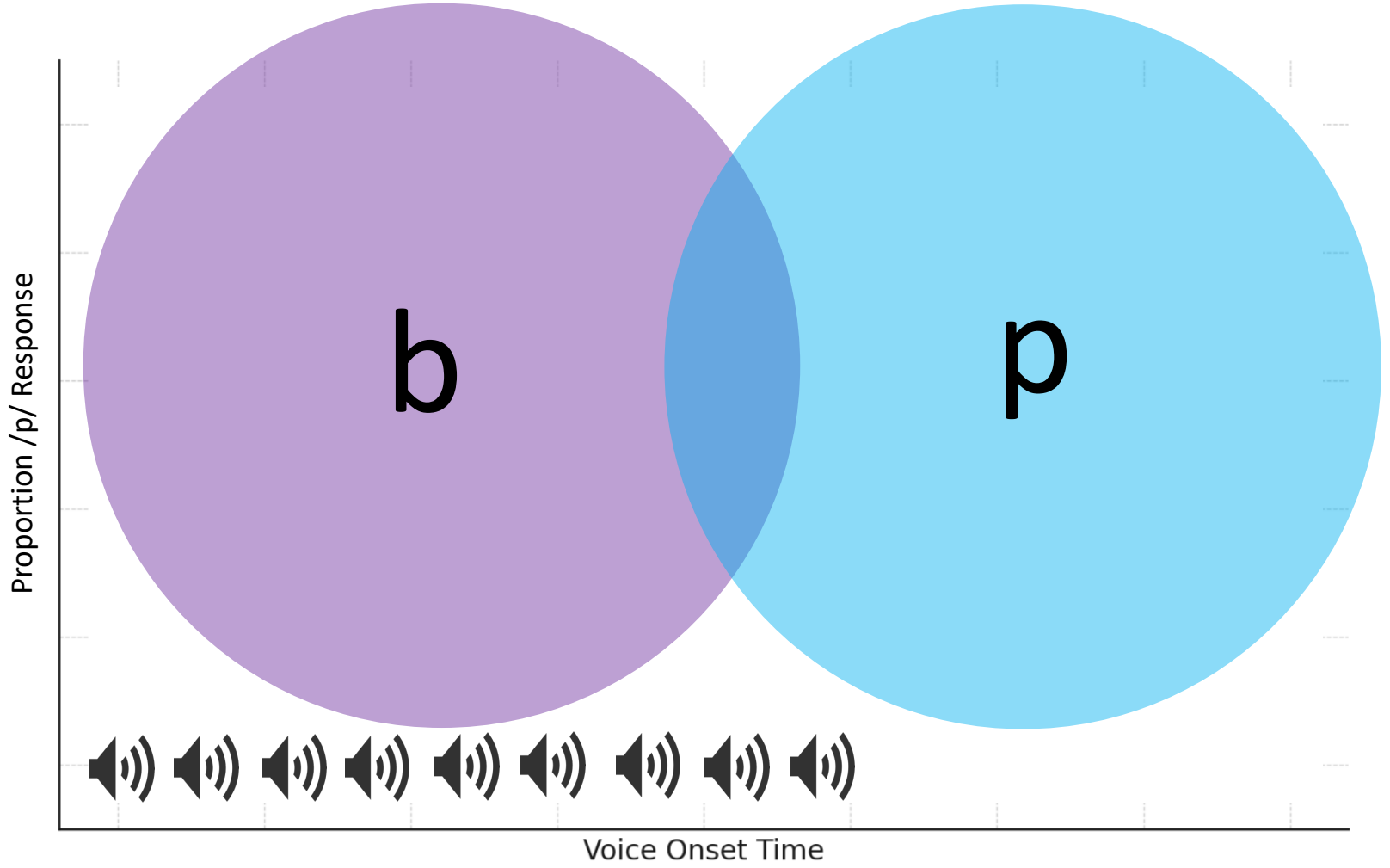
Why logistic regression?



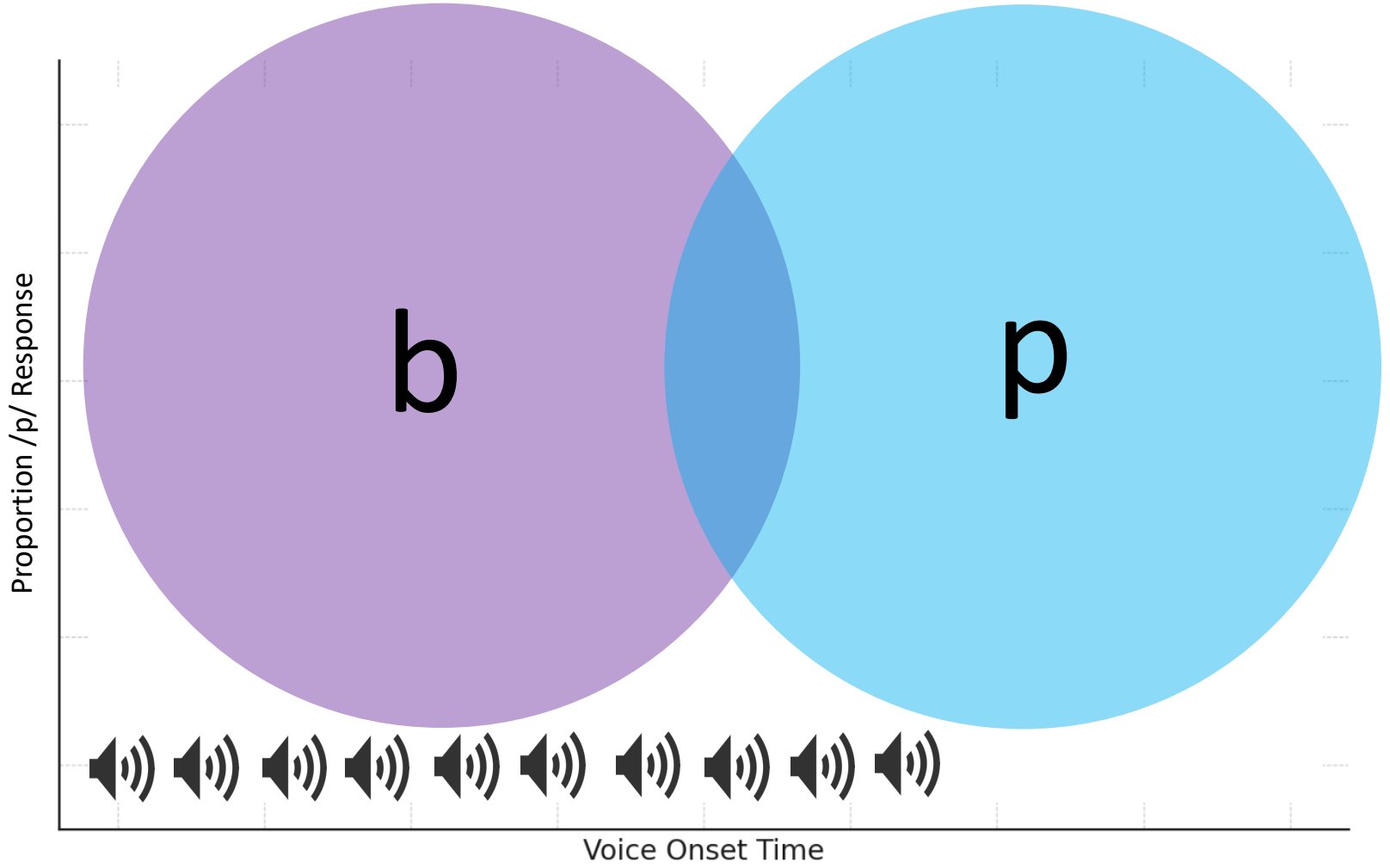
Why logistic regression?



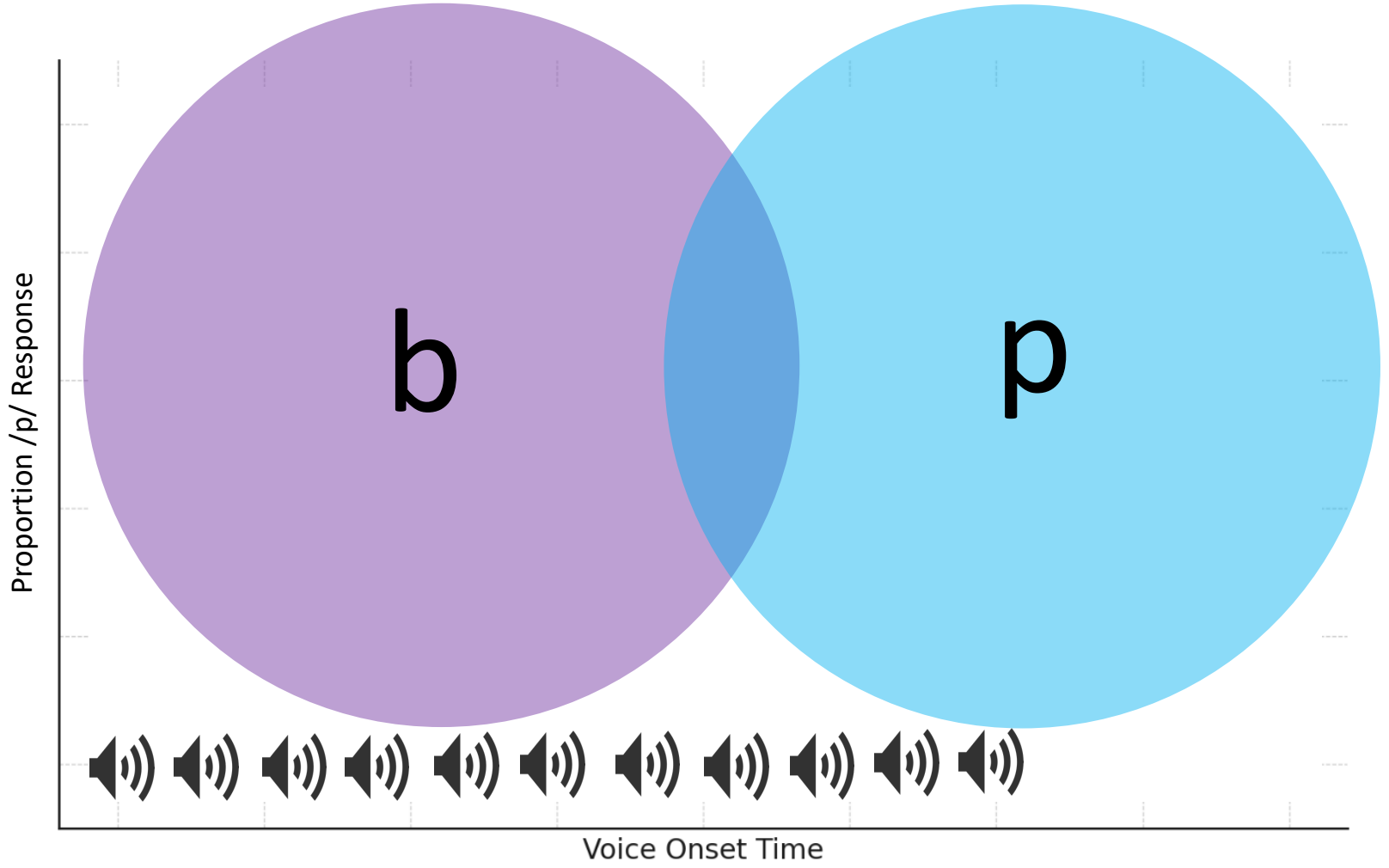
Why logistic regression?



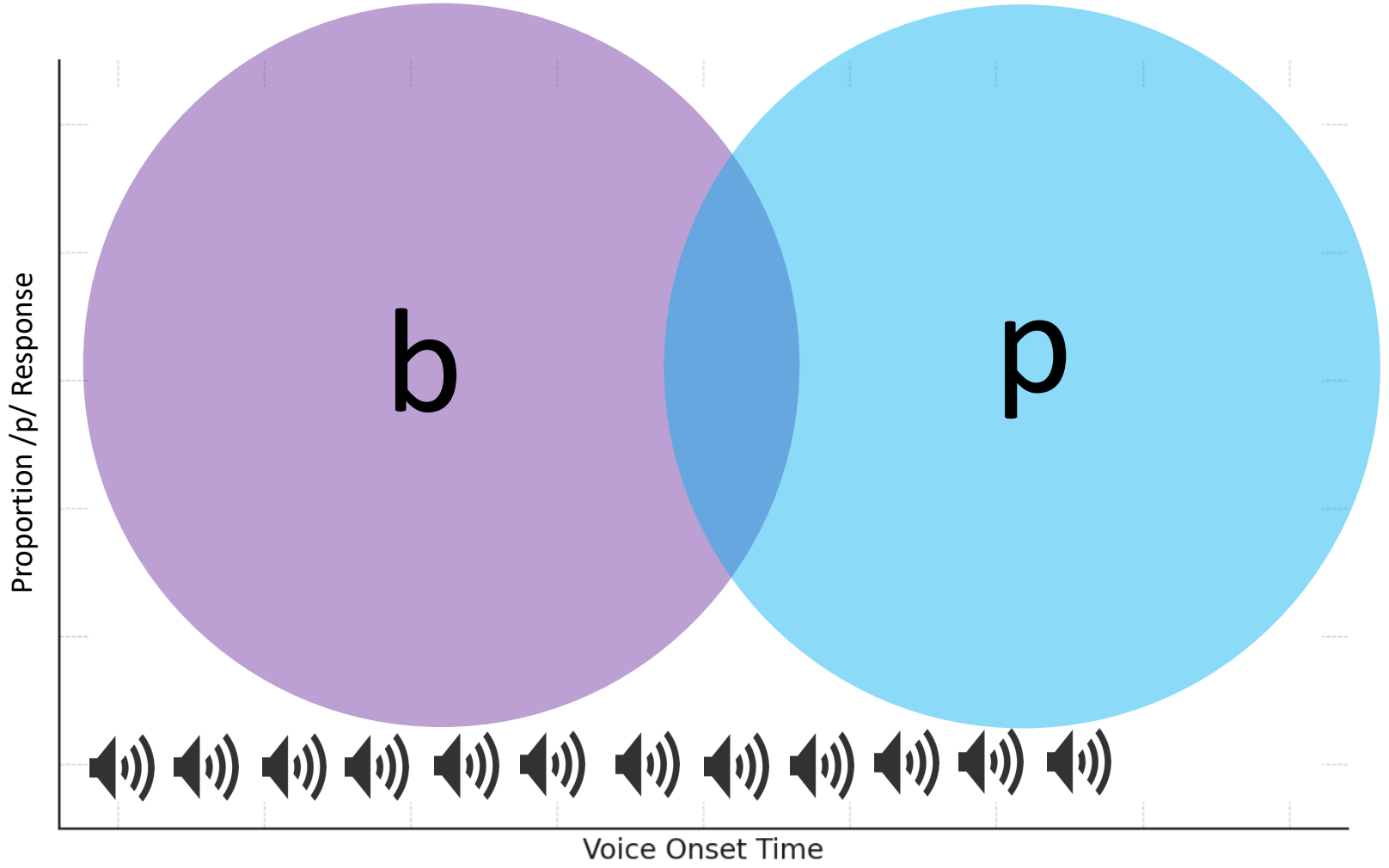
Why logistic regression?



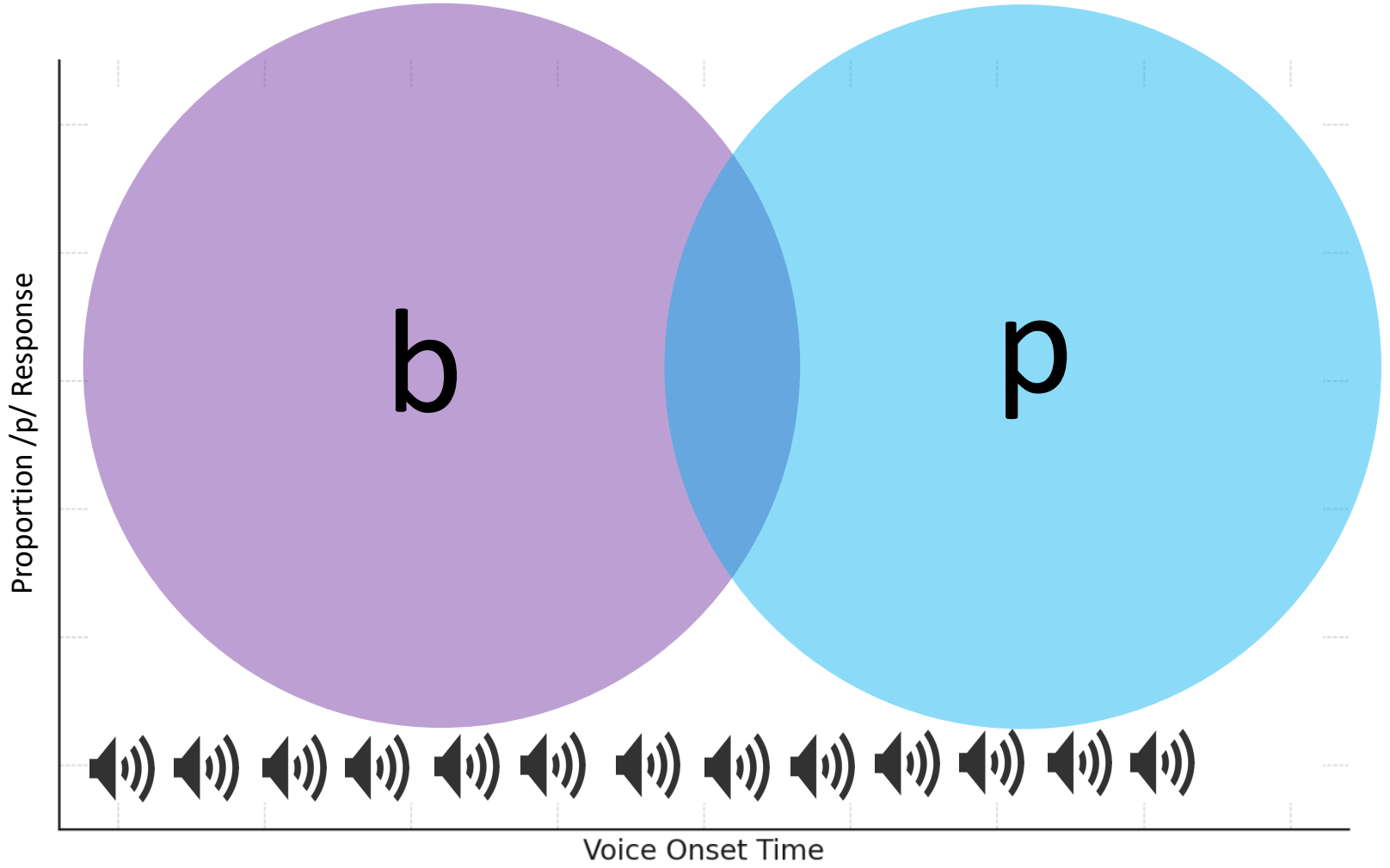
Why logistic regression?



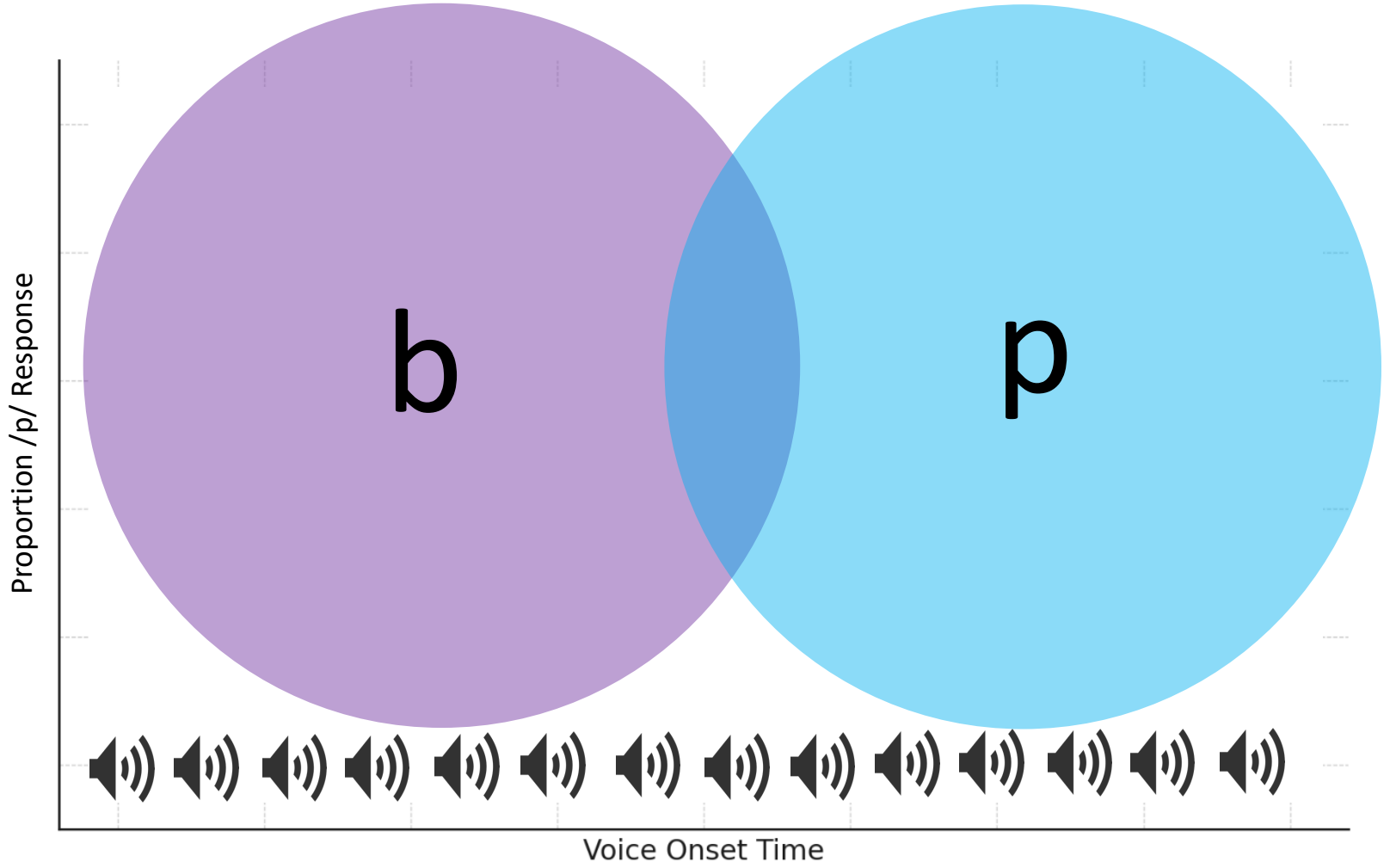
Why logistic regression?



Why logistic regression?



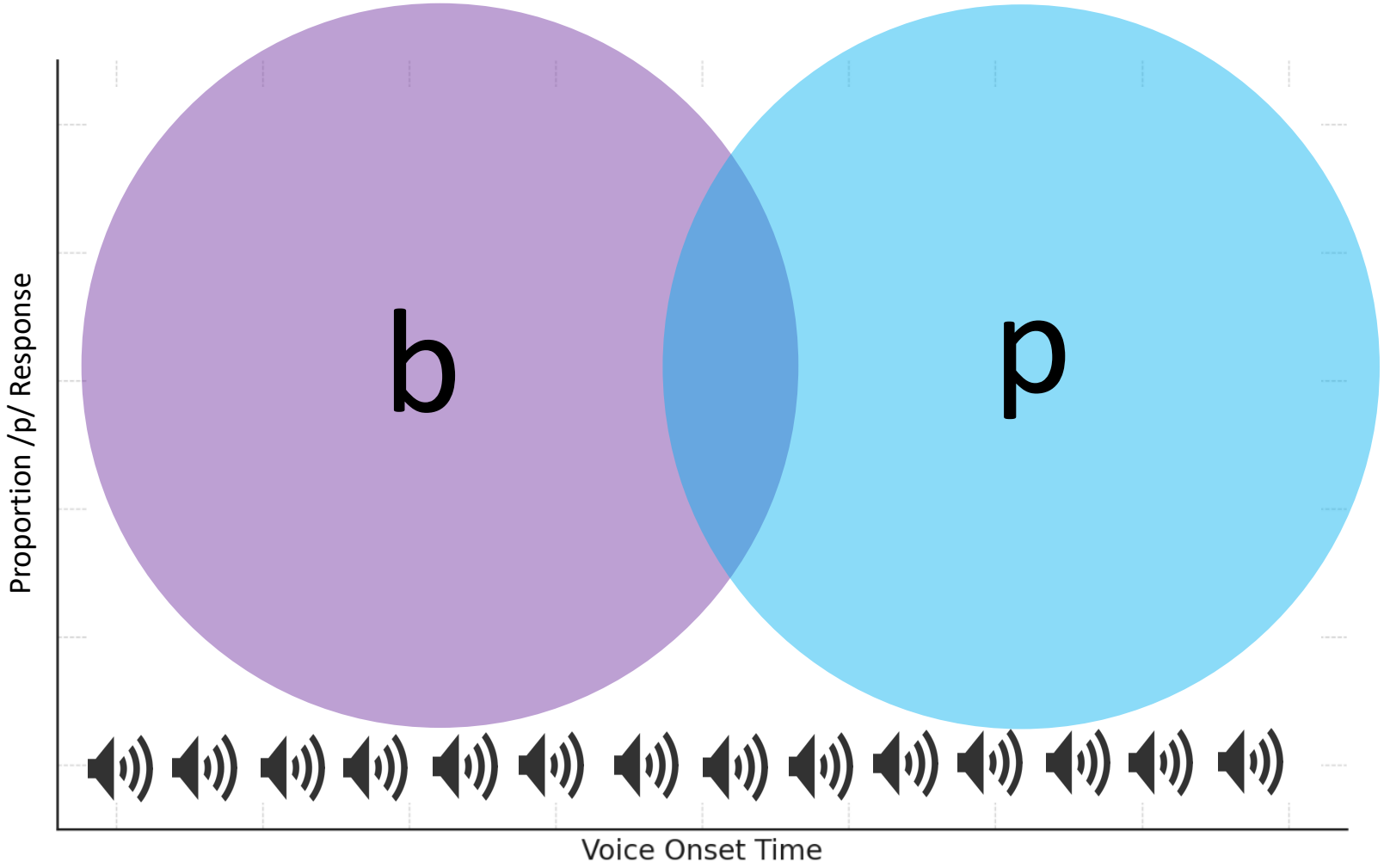
Why logistic regression?



Why logistic regression?

What sound did you hear?

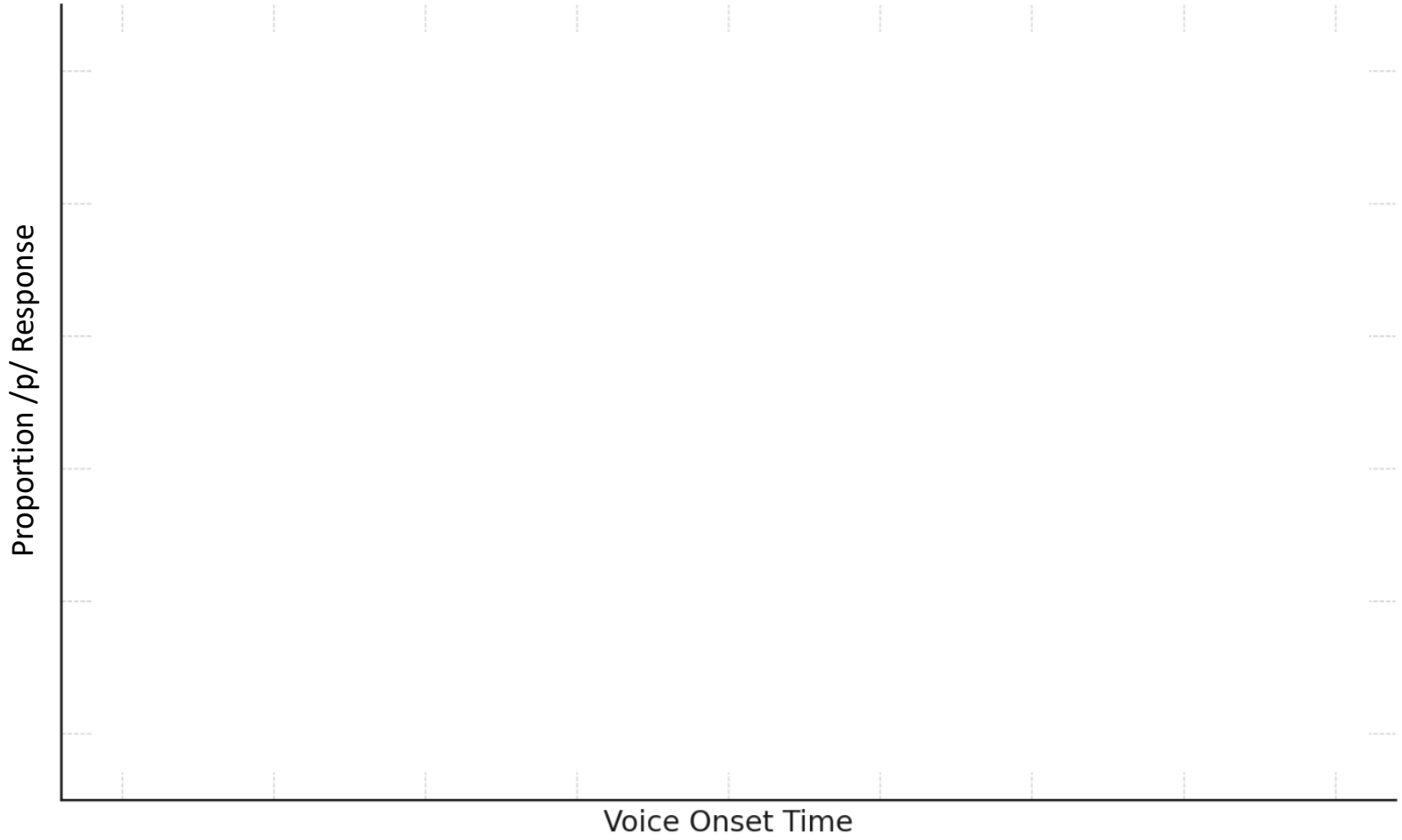
- b
- p



Why logistic regression?

What sound did you hear?

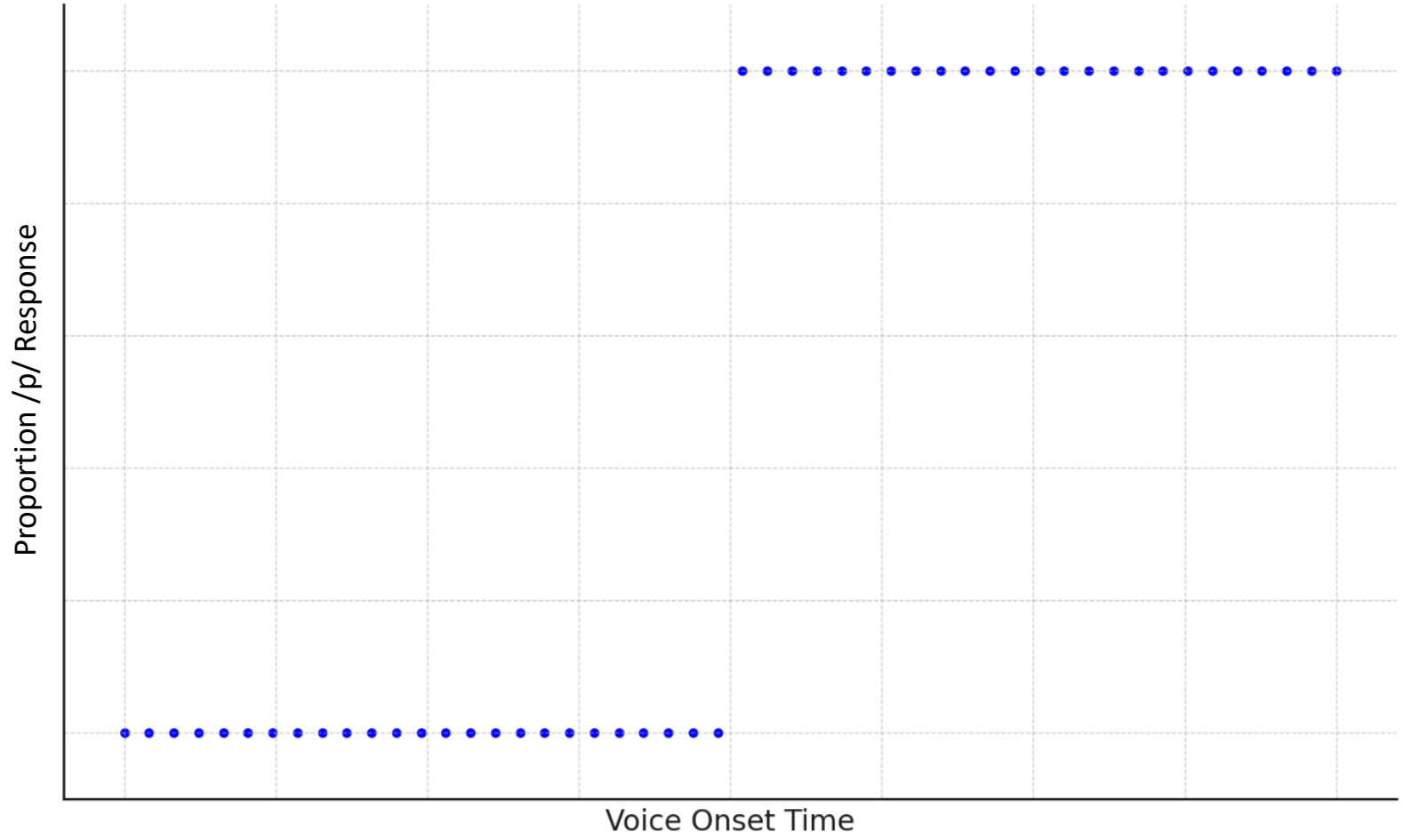
- b
- p



Why logistic regression?

What sound did you hear?

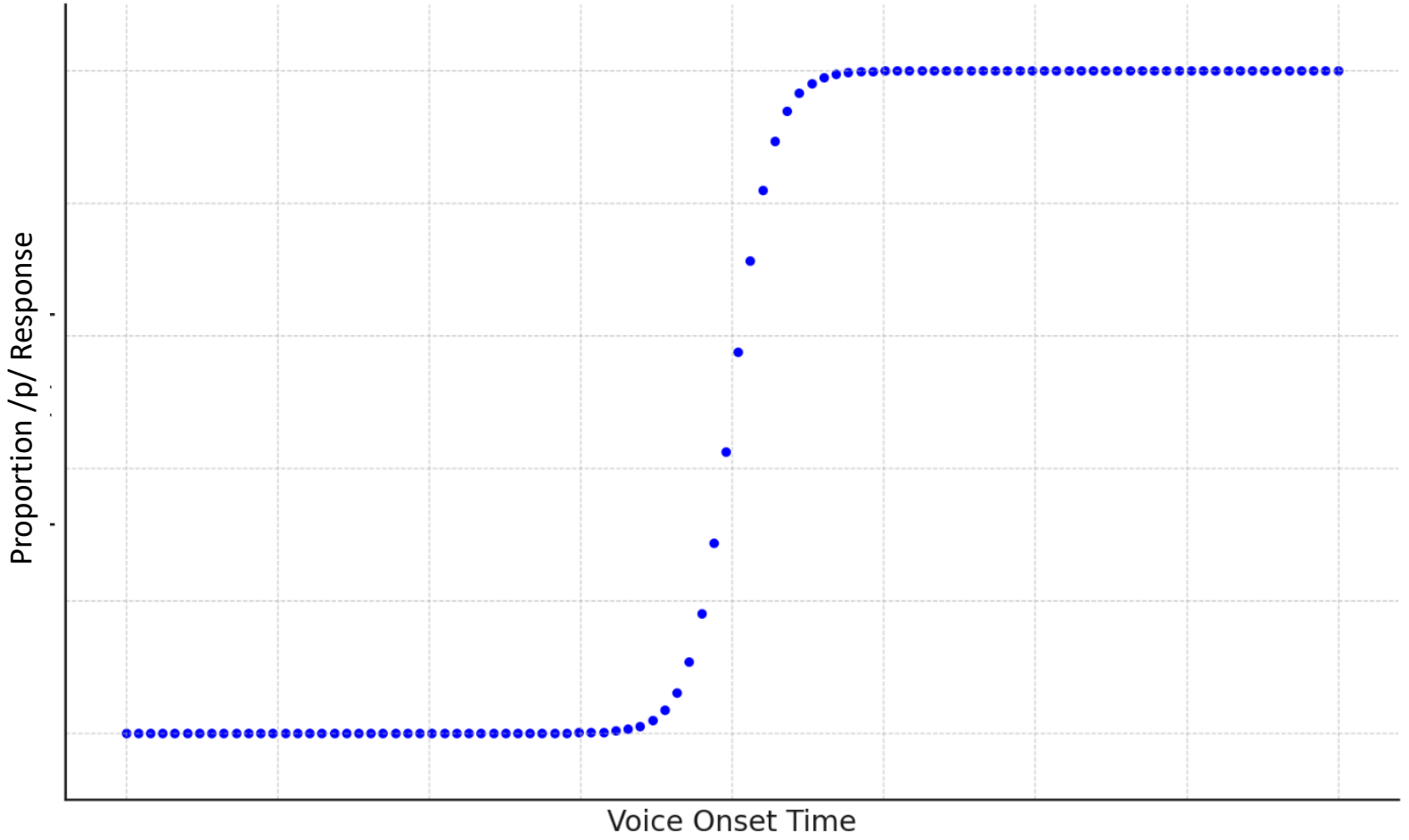
- b
- p



Why logistic regression?

What sound did you hear?

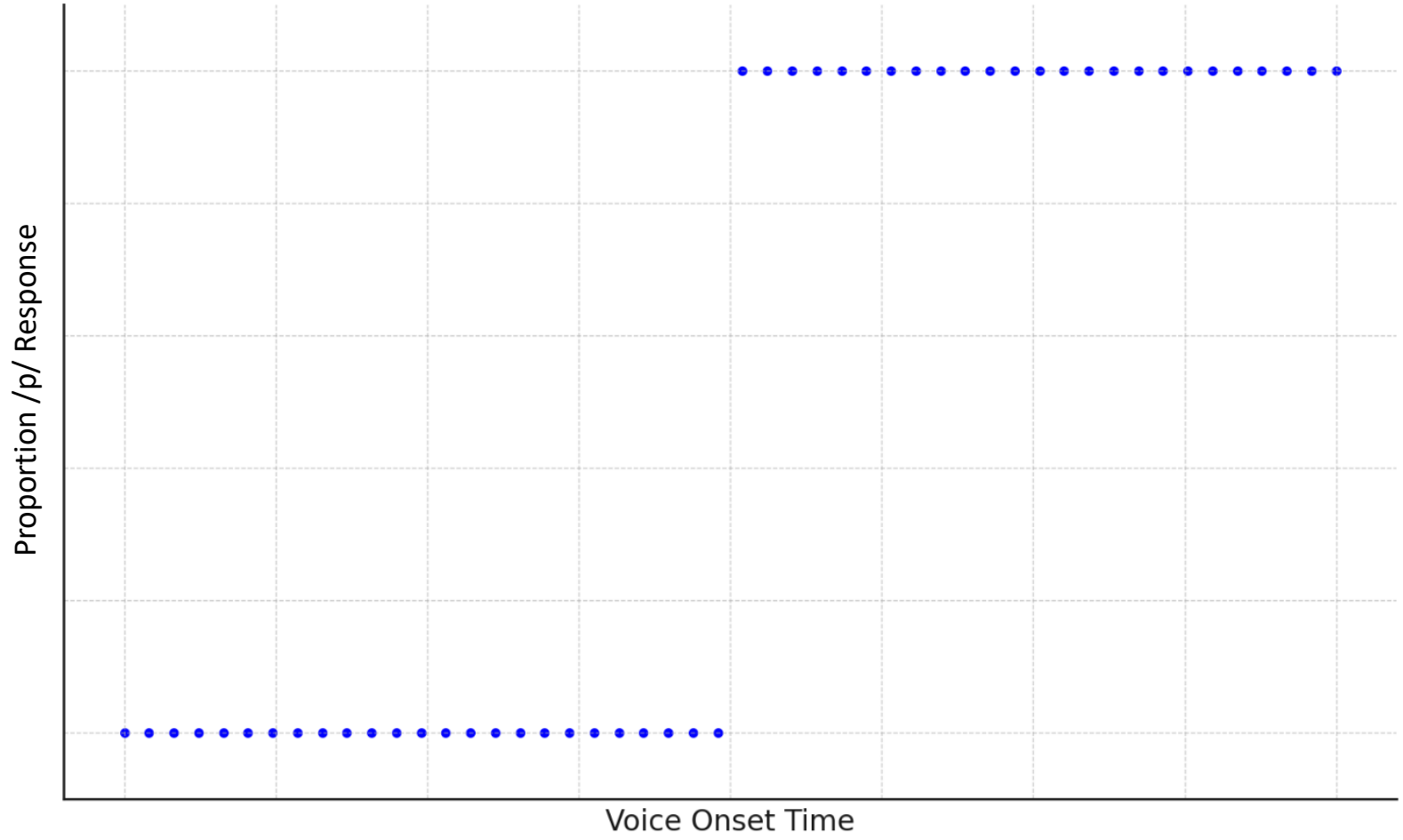
b ---------- p



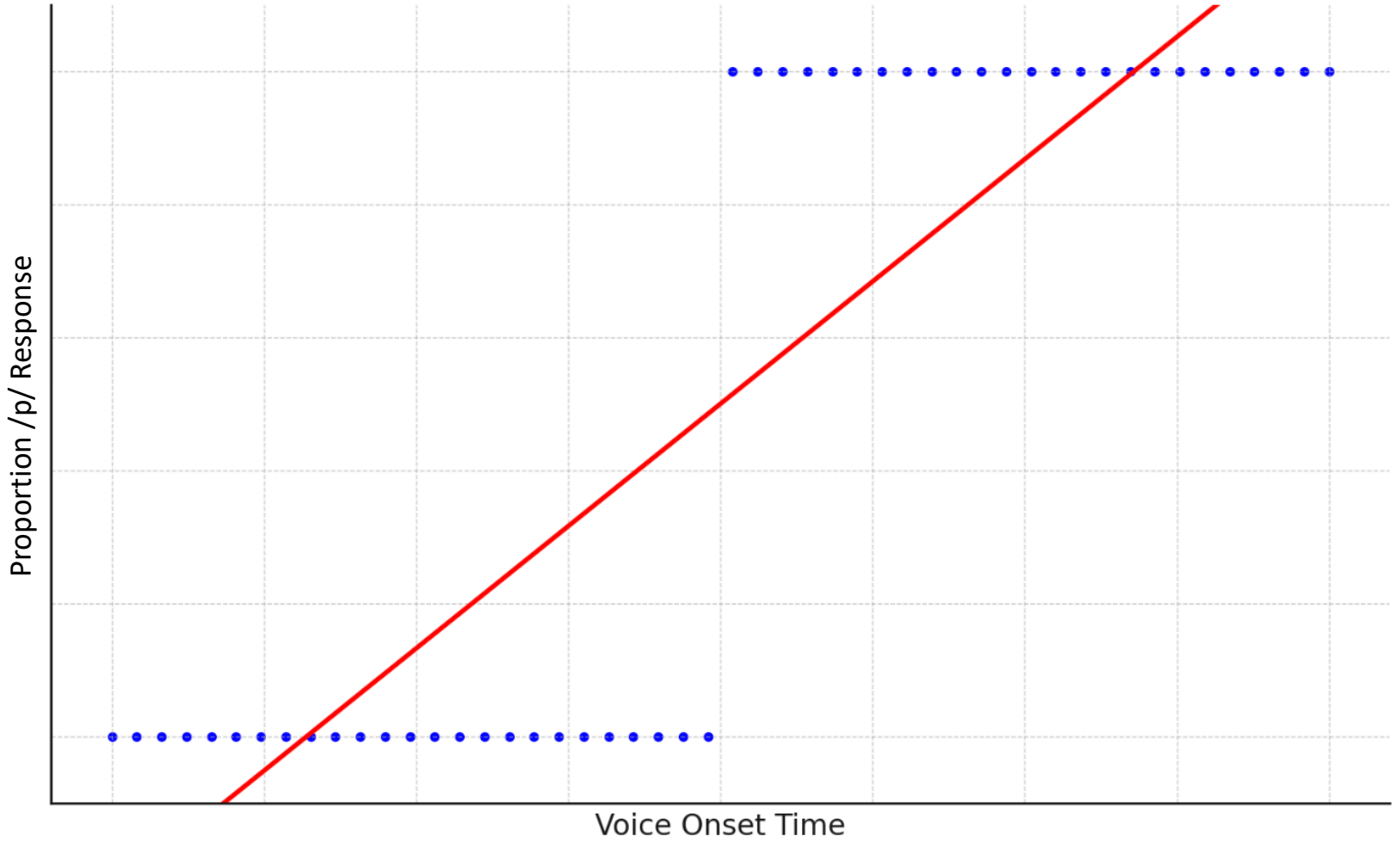
Why logistic regression?

What sound did you hear?

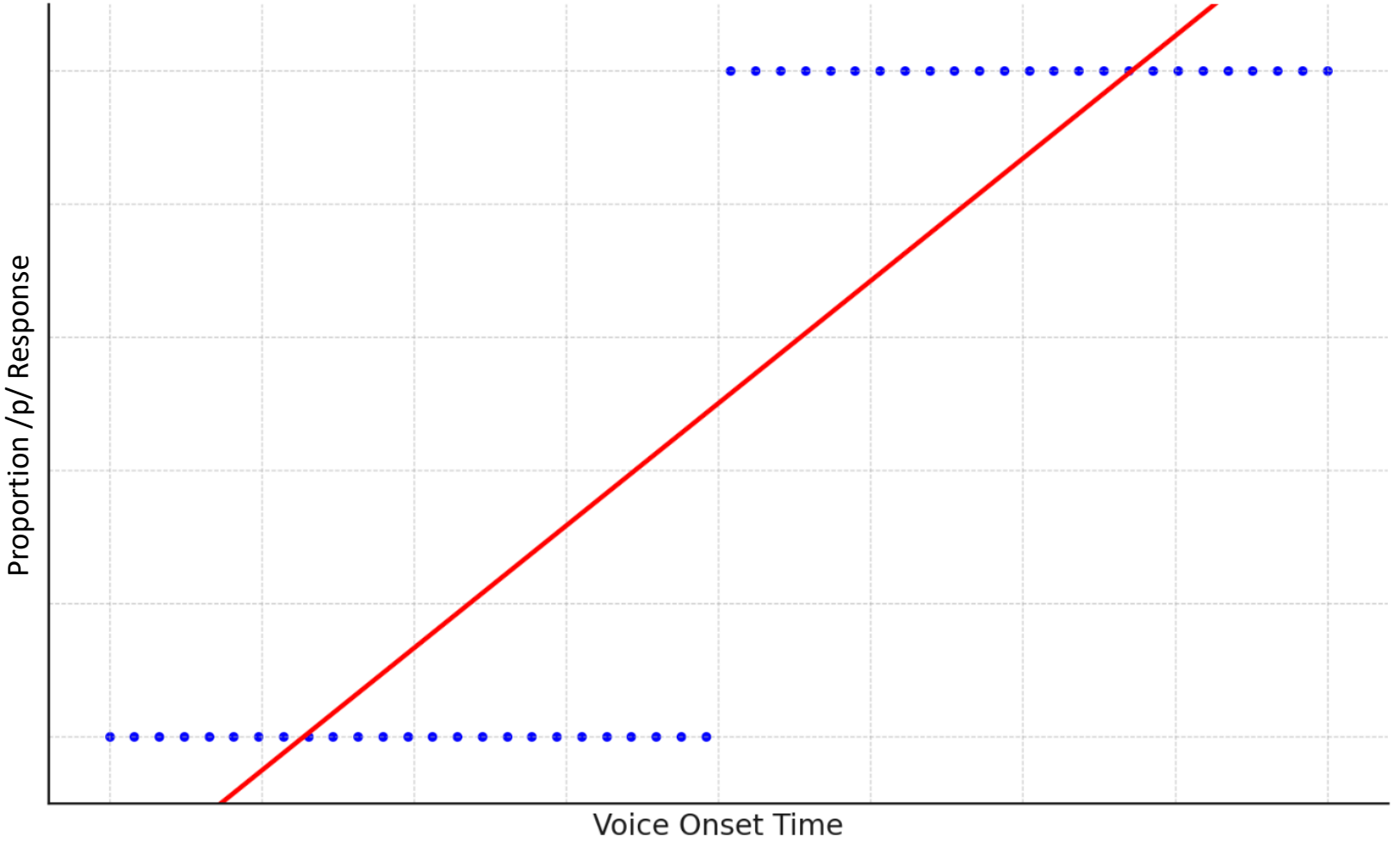
- b
- p



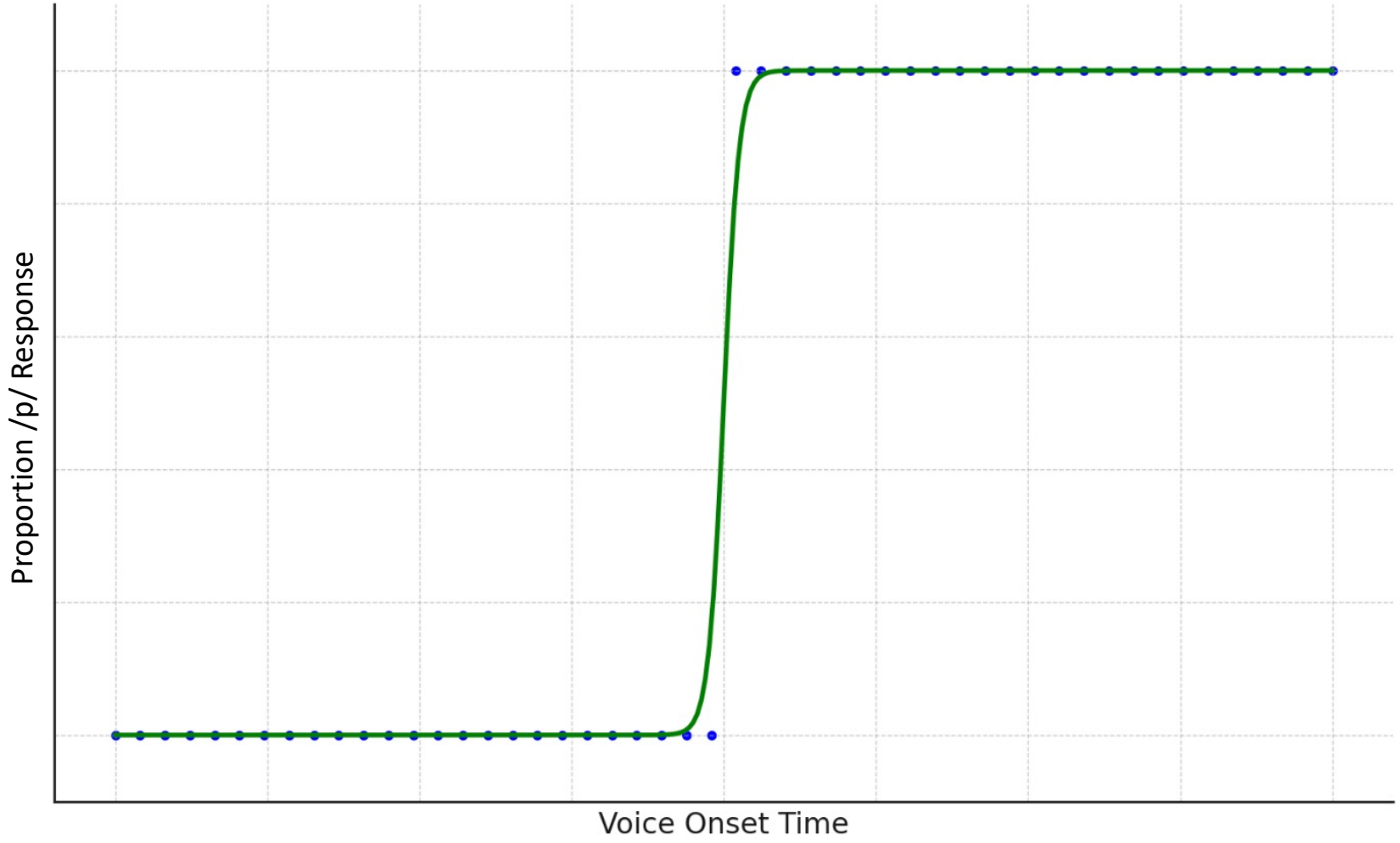
Why logistic regression?



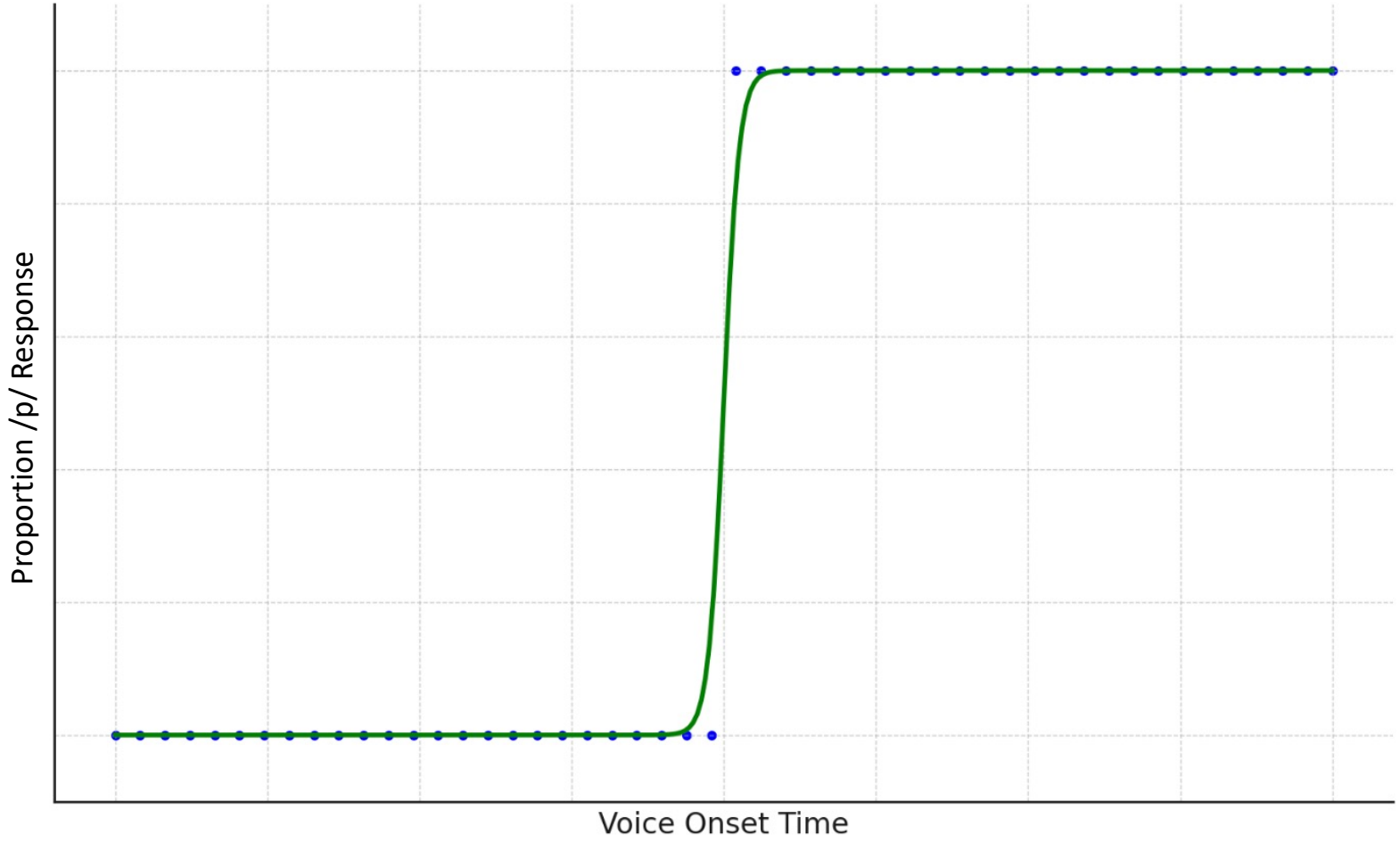
Why logistic regression?



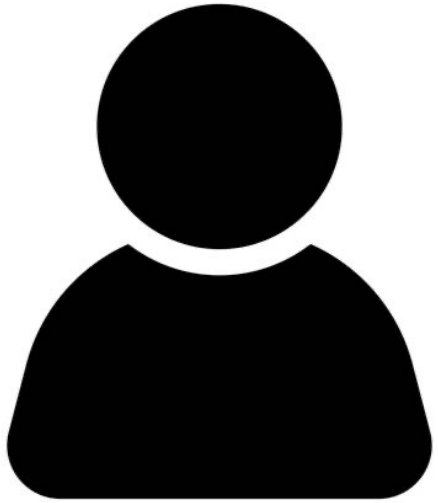
Why logistic regression?



Why logistic regression?



Another example



Another example

They cooked a casserole for him.

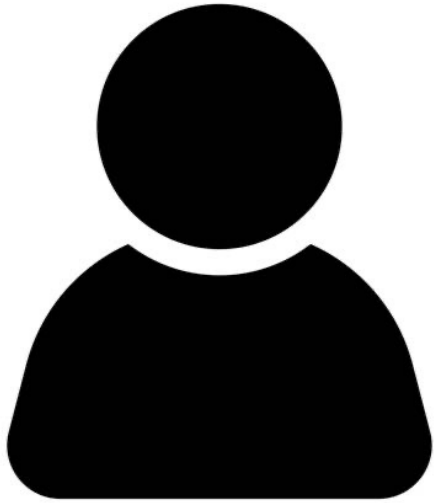
The parents grilled a chicken breast for Maria.

Maya baked home-made cookies for them.

Sarah made a burger for the kid.

The chef boiled a few carrots for her.

Amir sliced some vegetables for Luca.



Another example

They cooked a casserole for him.

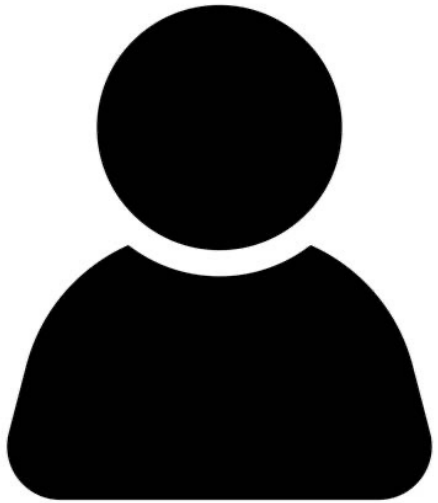
The parents grilled a chicken breast for Maria.

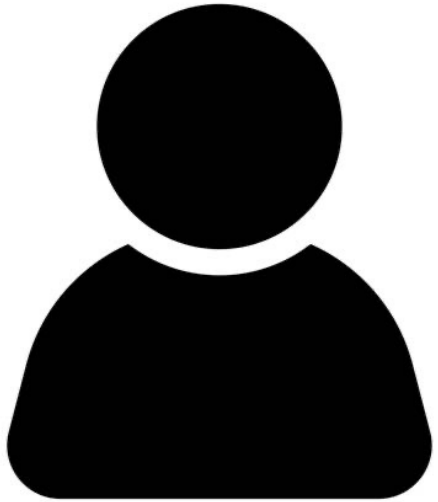
Maya baked home-made cookies for them.

Sarah made a burger for the kid.

The chef boiled a few carrots for her.

Amir sliced some vegetables for Luca.





Another example

They cooked a casserole for him.

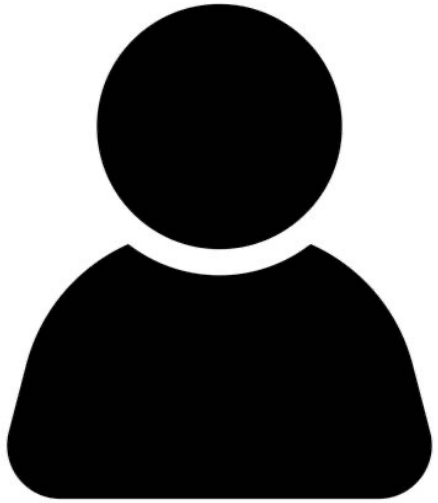
The parents grilled a chicken breast for Maria.

Maya baked home-made cookies for them.

Sarah made a burger for the kid.

The chef boiled a few carrots for her.

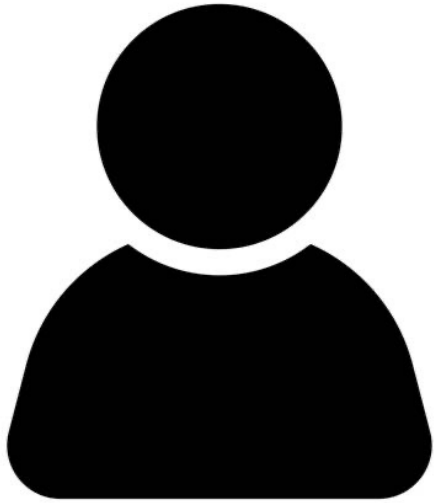
Amir sliced some vegetables for Luca.



Did you read this exact sentence?

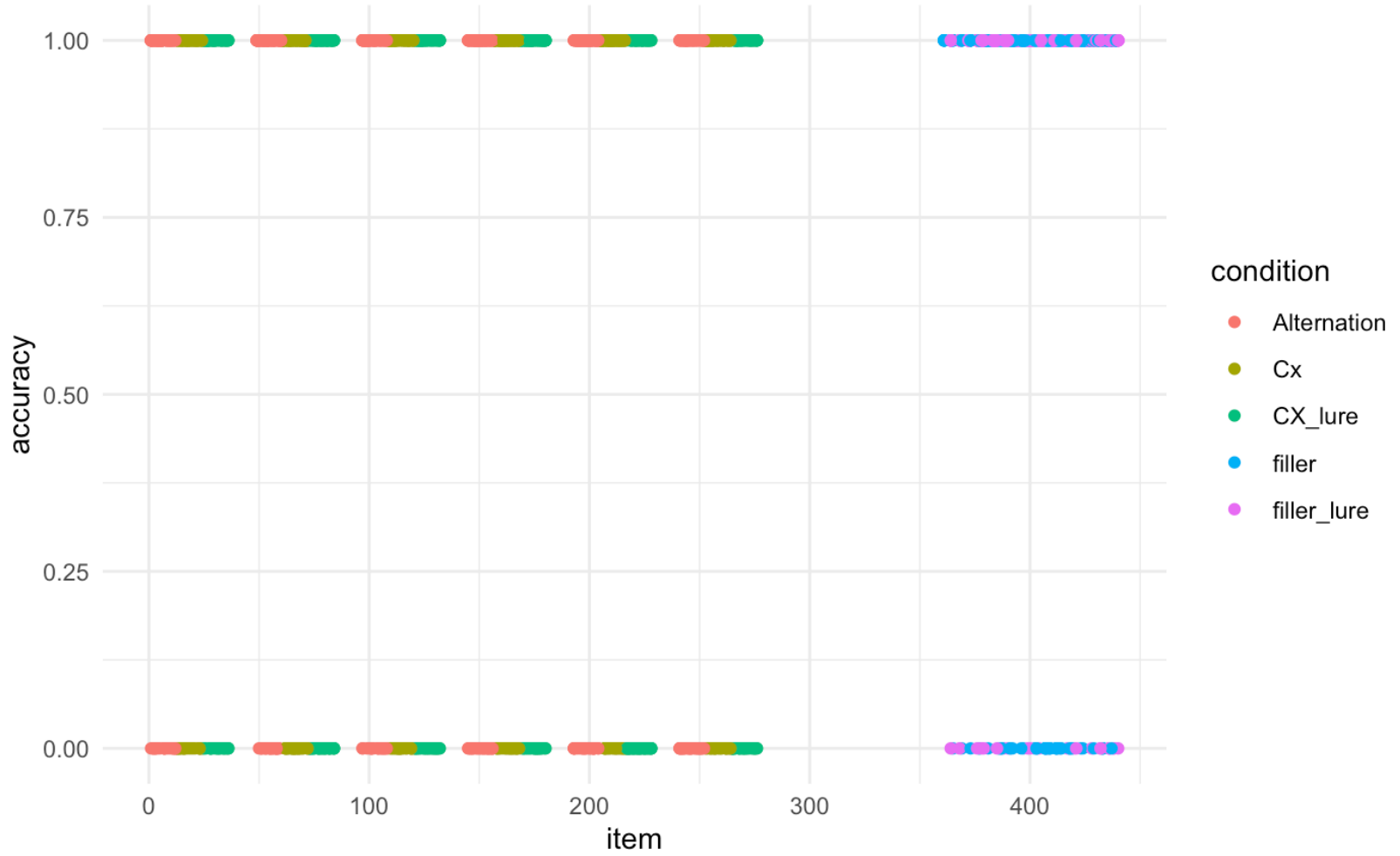
They made Maria home-made cookies.

- ◯ Yes
- ◯ No



Another example

Another example



Why logistic regression?

Why logistic regression?

```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```

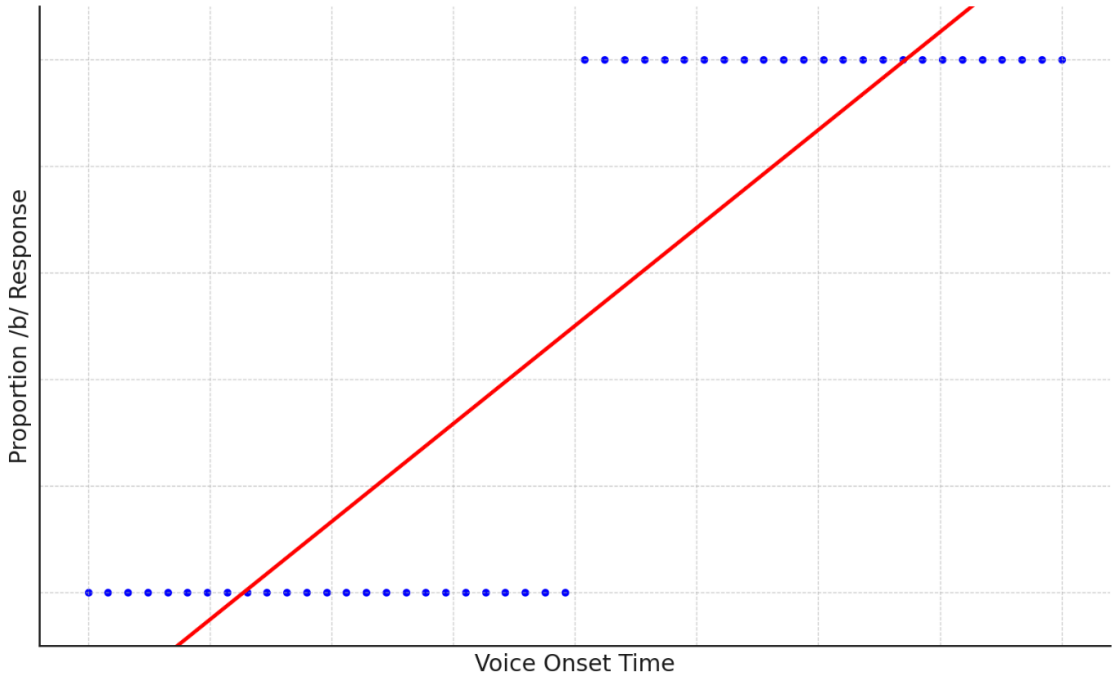
Why logistic regression?

```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```

linearity: there is a linear relationship between the IV and the DV

Why logistic regression?

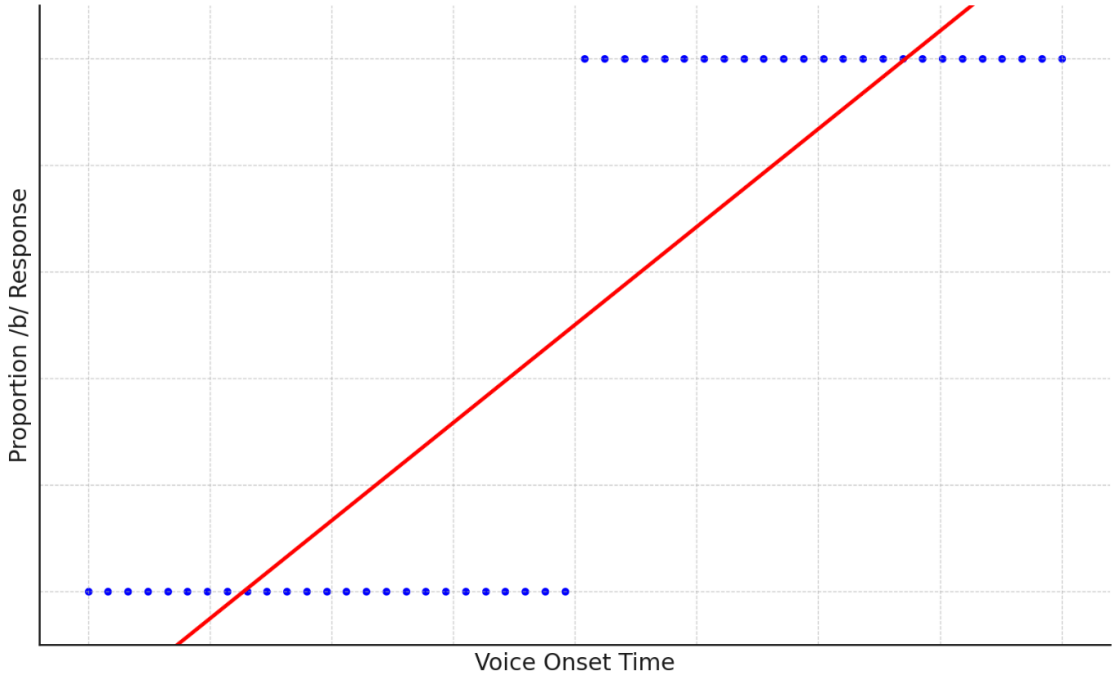
```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```



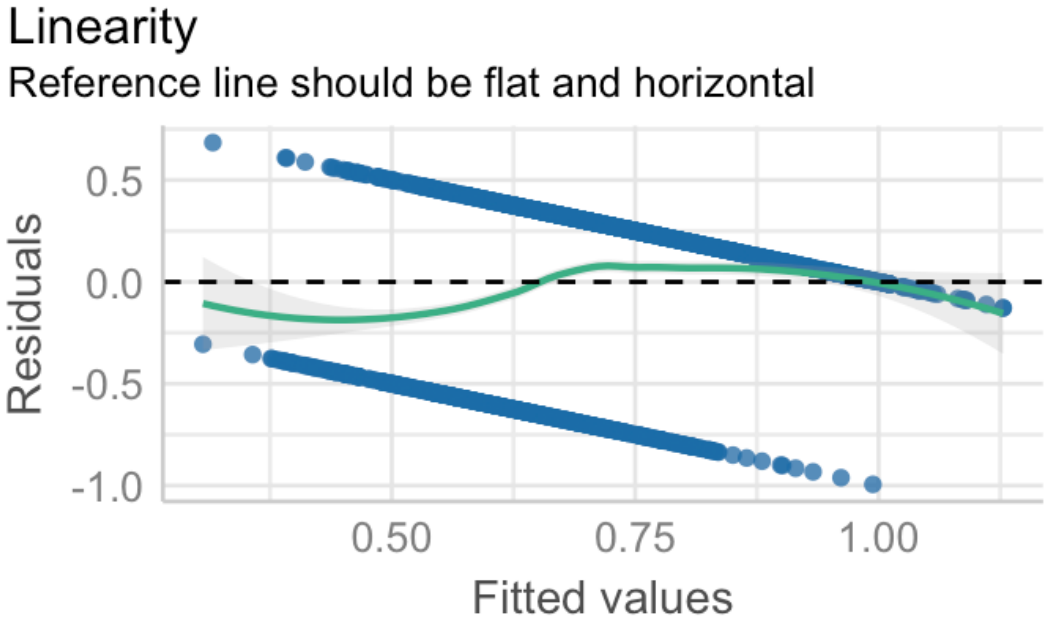
linearity: there is a linear relationship between the IV and the DV

Why logistic regression?

```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```

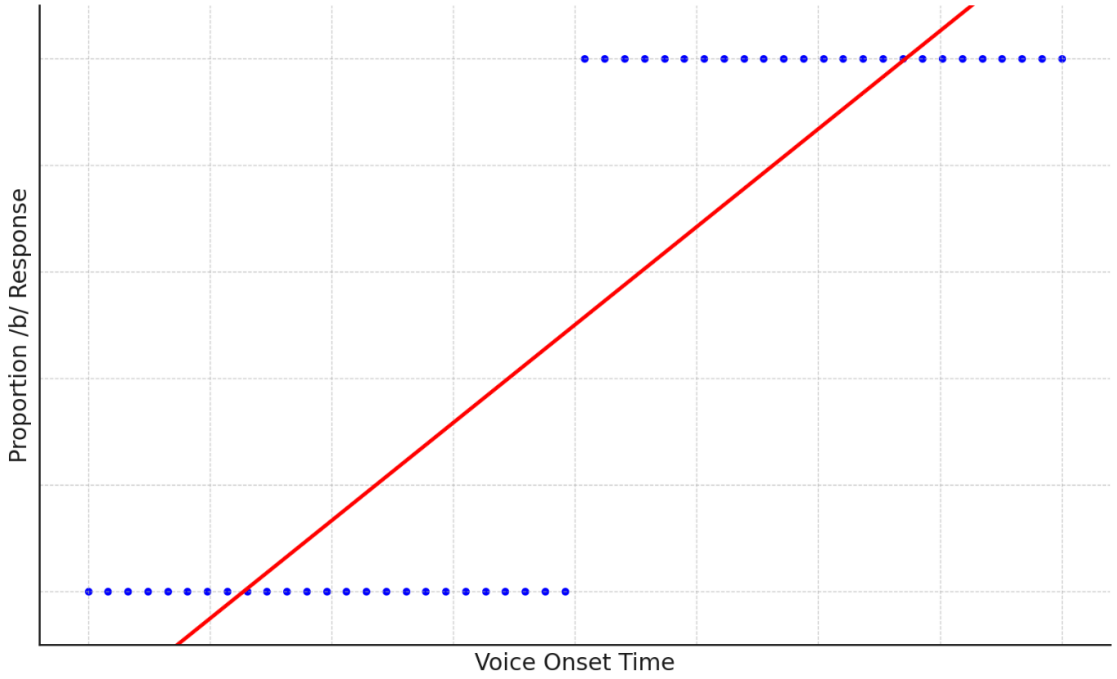


linearity: there is a linear relationship between the IV and the DV

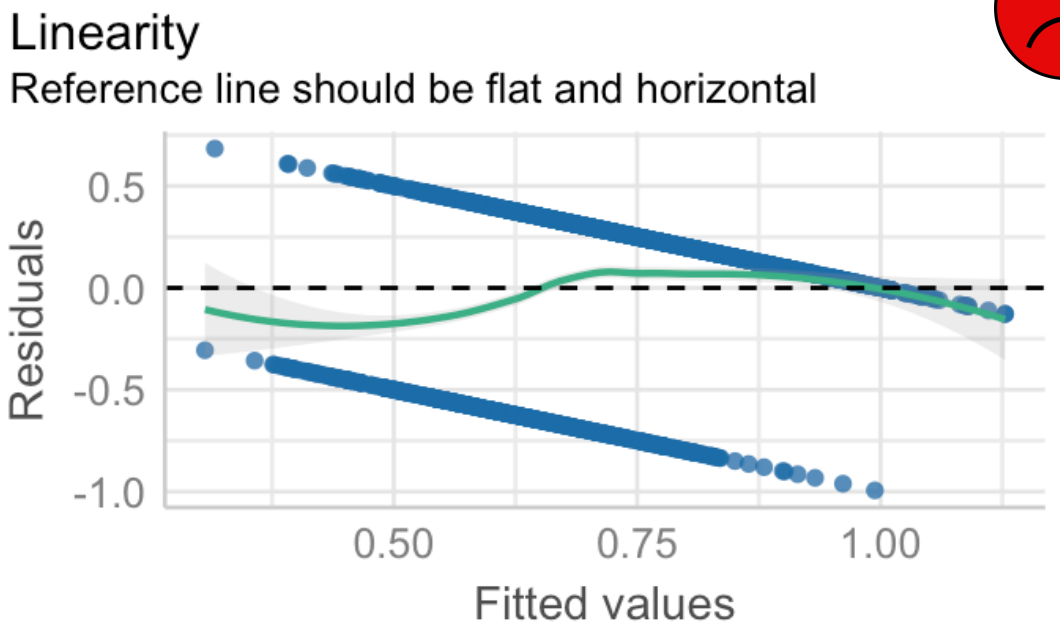


Why logistic regression?

```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```



linearity: there is a linear relationship between the IV and the DV



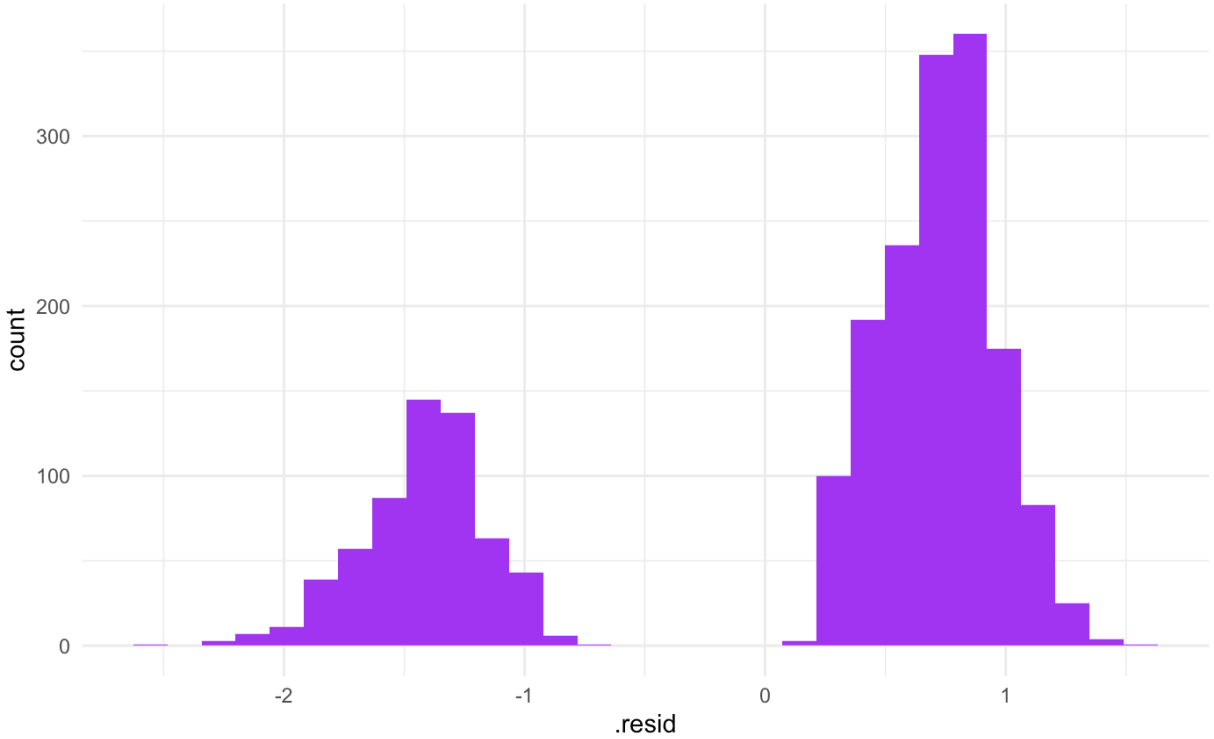
Why logistic regression?

```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```

normality: residuals have a normal distribution

Why logistic regression?

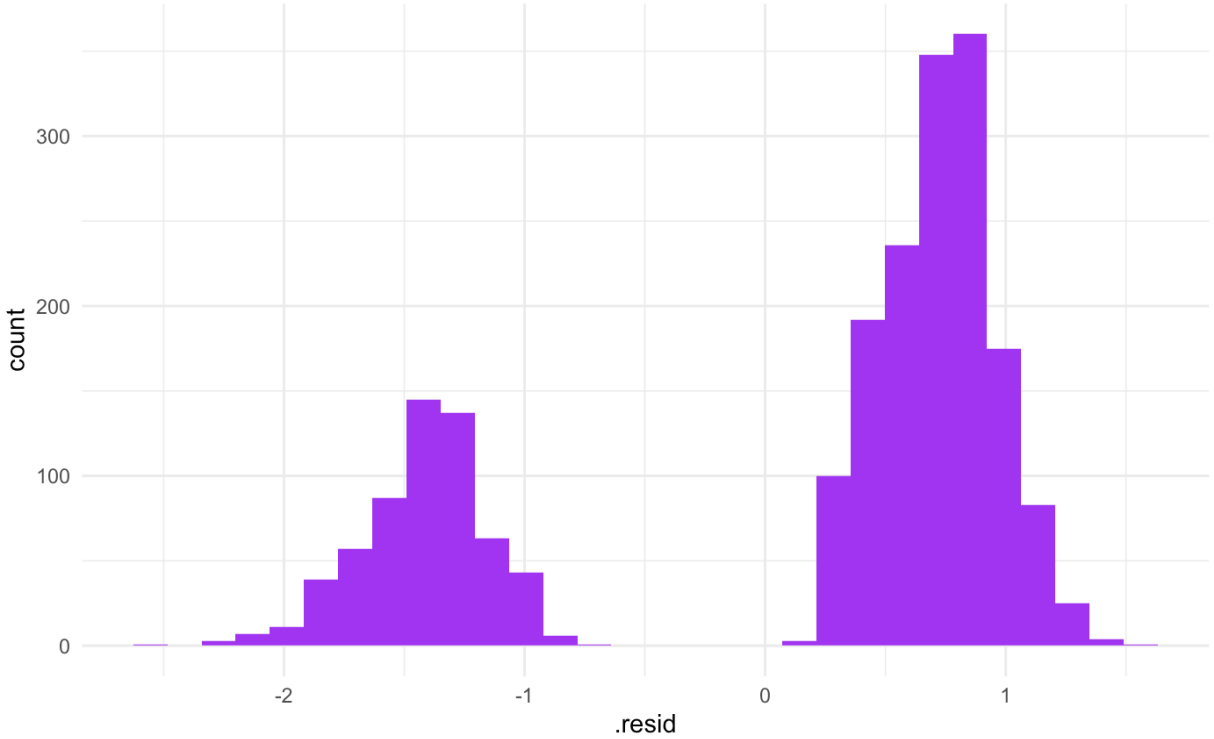
```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```



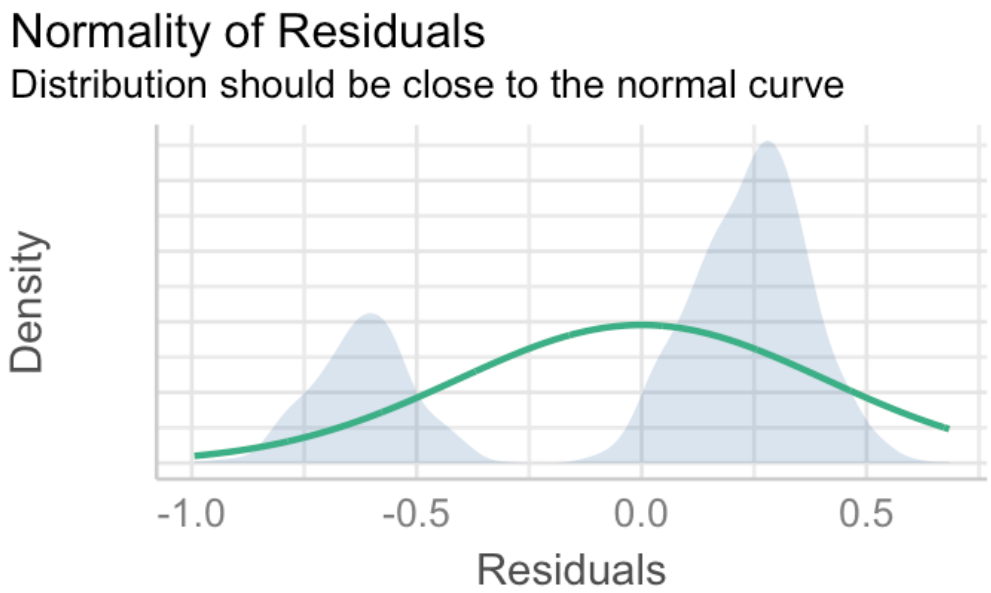
normality: residuals have a normal distribution

Why logistic regression?

```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```

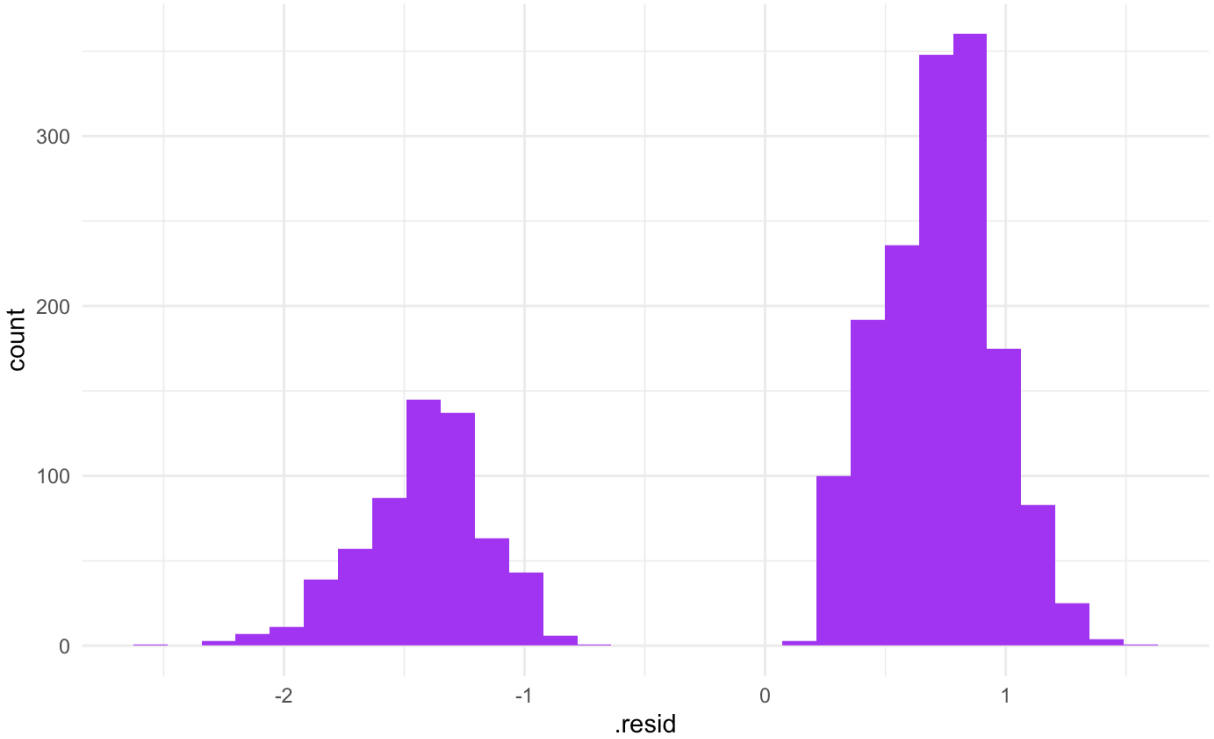


normality: residuals have a normal distribution

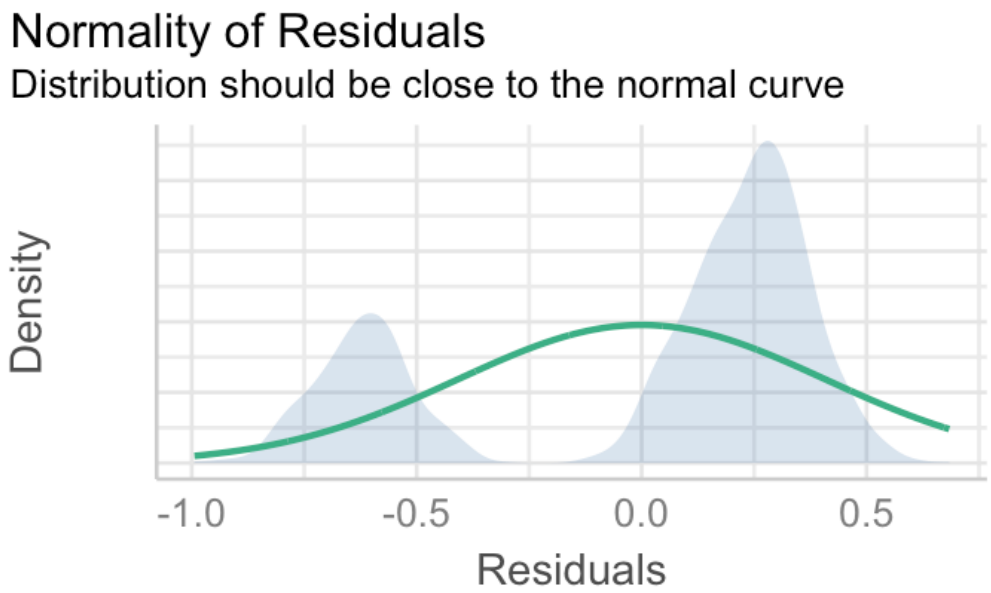


Why logistic regression?

```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```



normality: residuals have a normal distribution



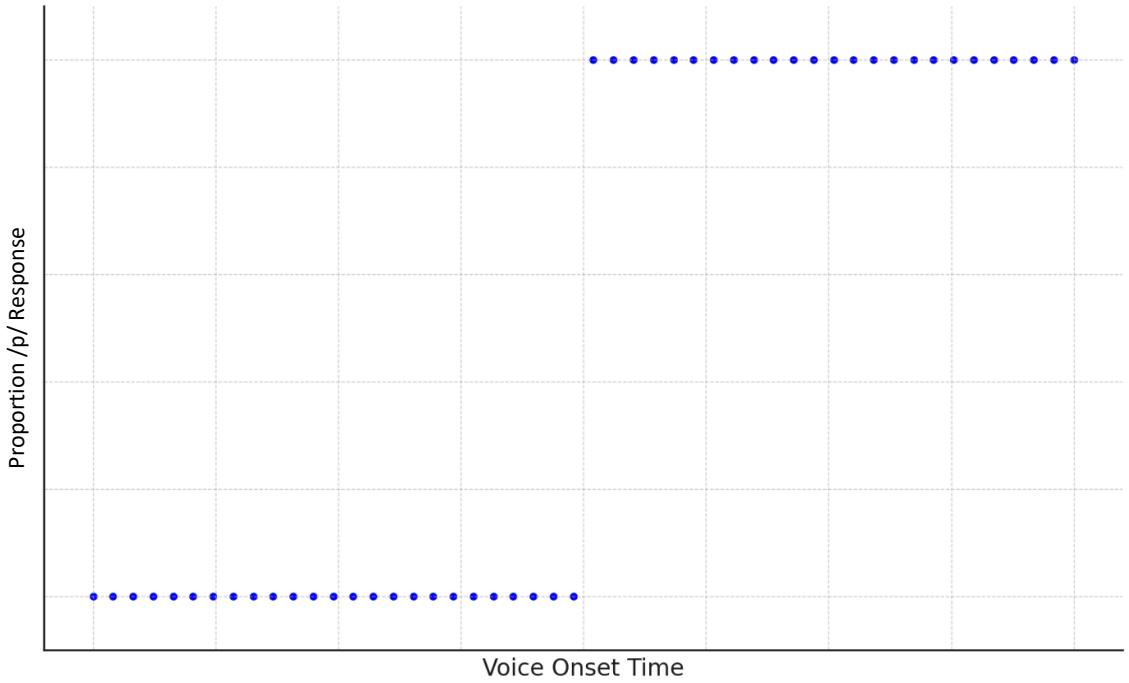
Why logistic regression?

```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```

homogeneity: error variance is the same across different levels of the IV

Why logistic regression?

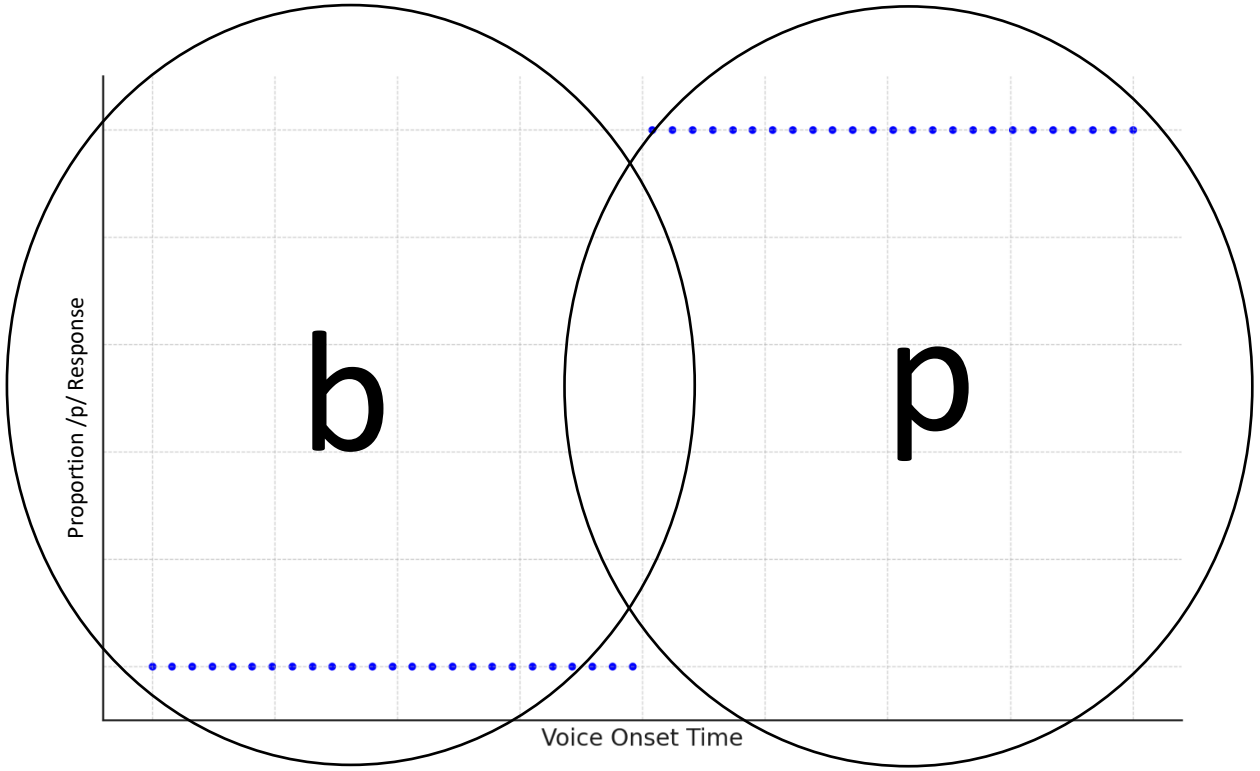
```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```



homogeneity: error variance is the same across different levels of the IV

Why logistic regression?

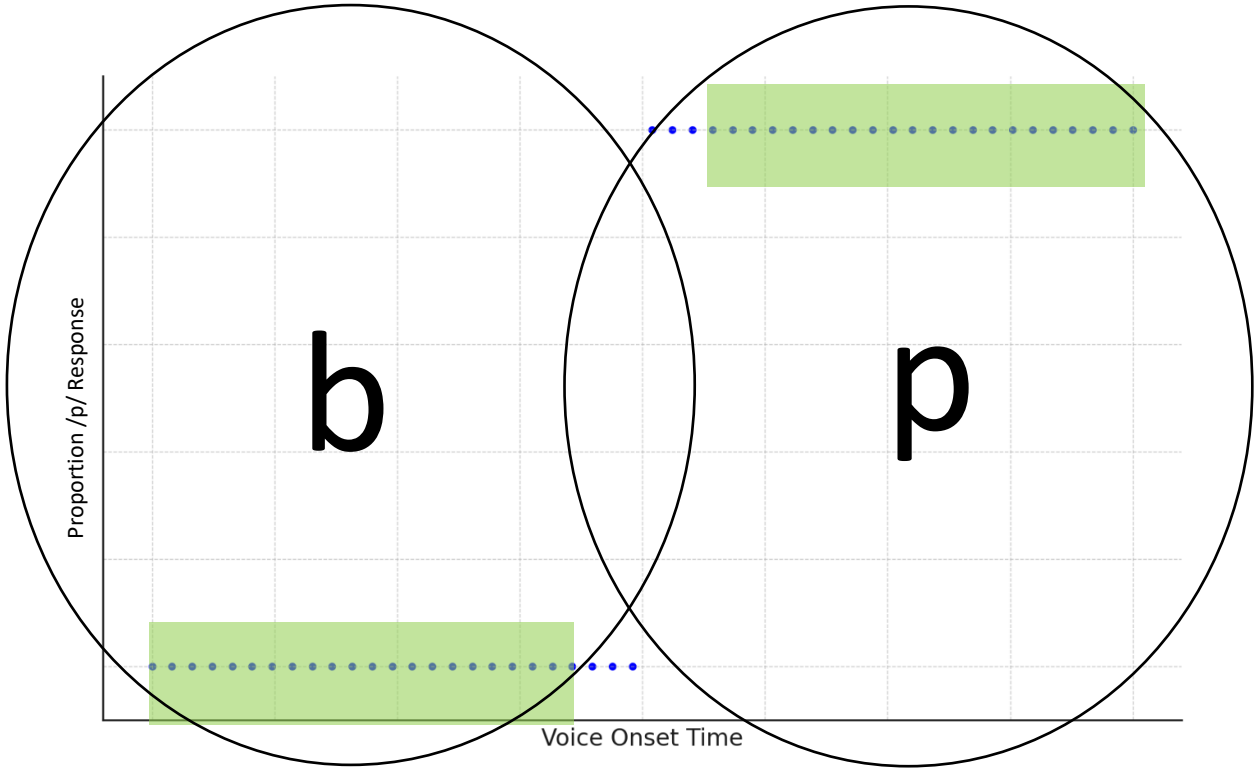
```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```



homogeneity: error variance is the same across different levels of the IV

Why logistic regression?

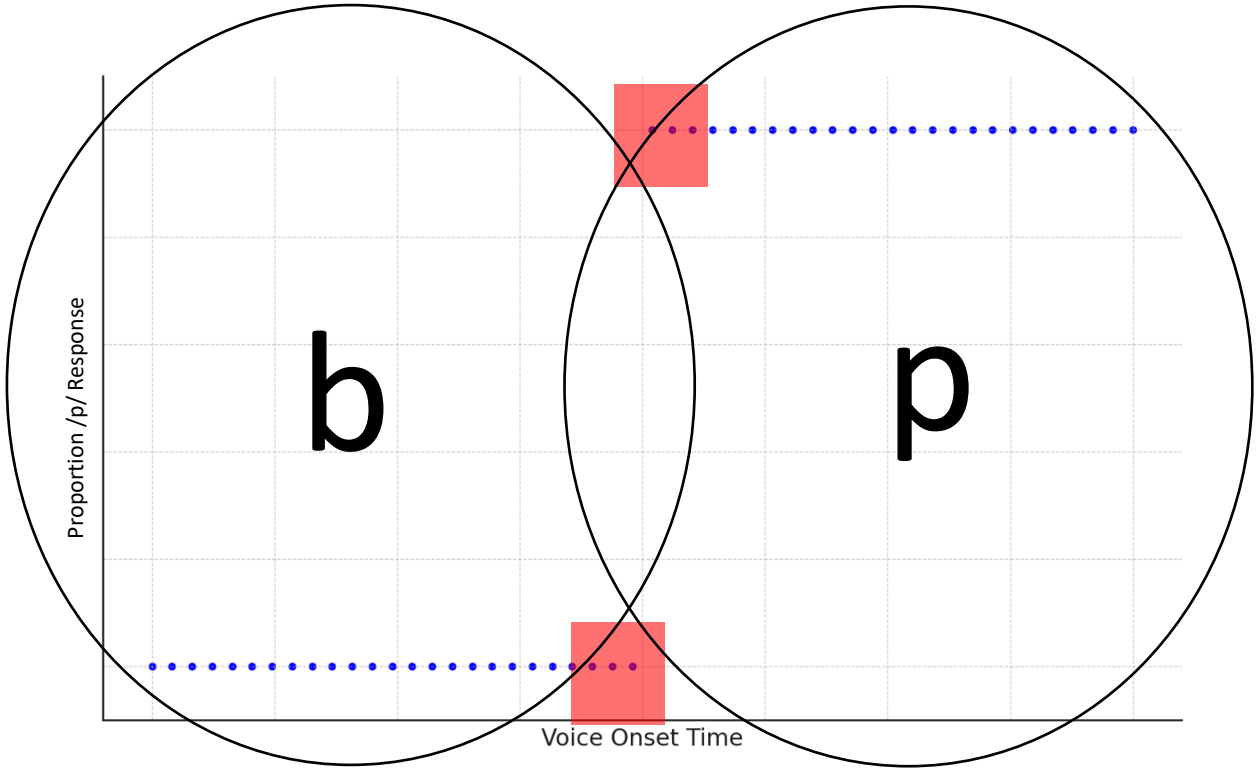
```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```



homogeneity: error variance is the same across different levels of the IV

Why logistic regression?

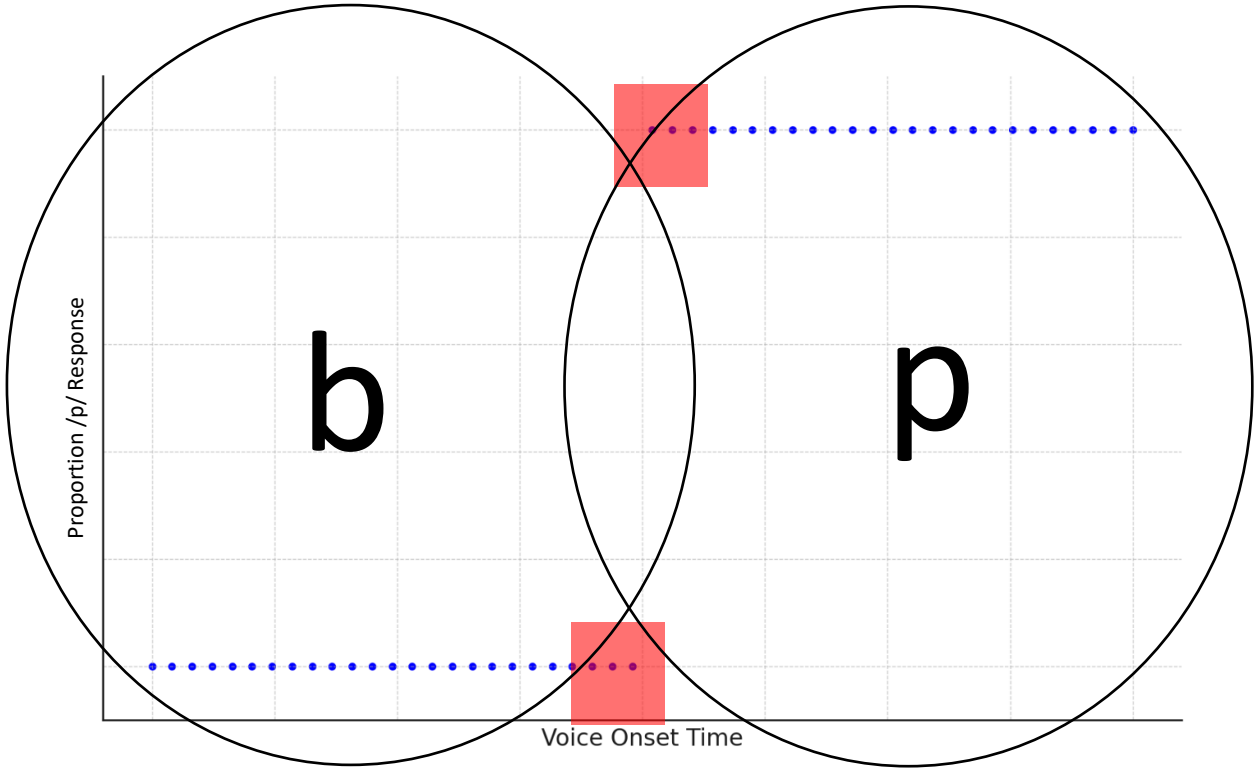
```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```



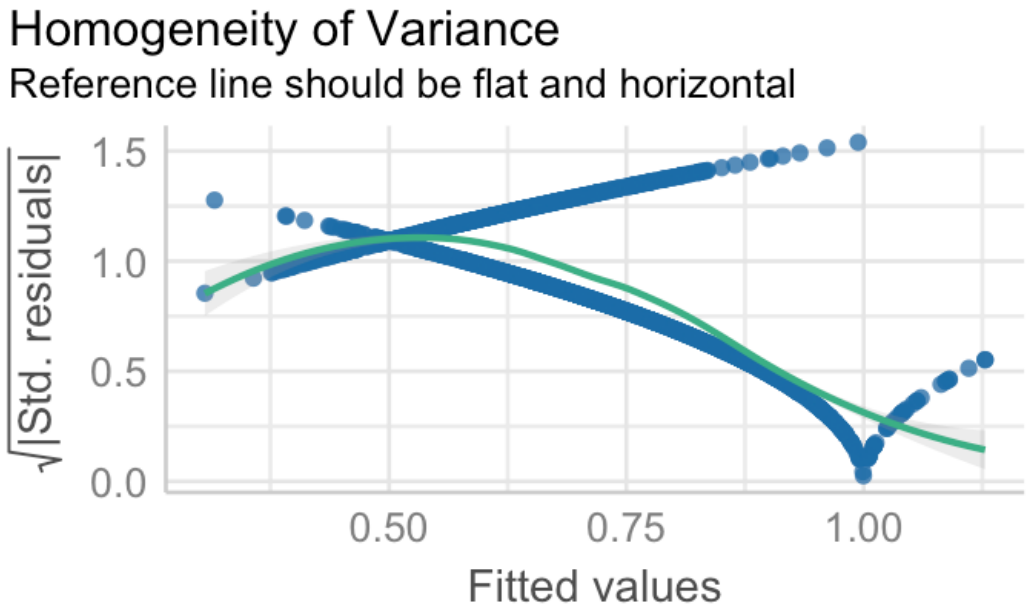
homogeneity: error variance is the same across different levels of the IV

Why logistic regression?

```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```

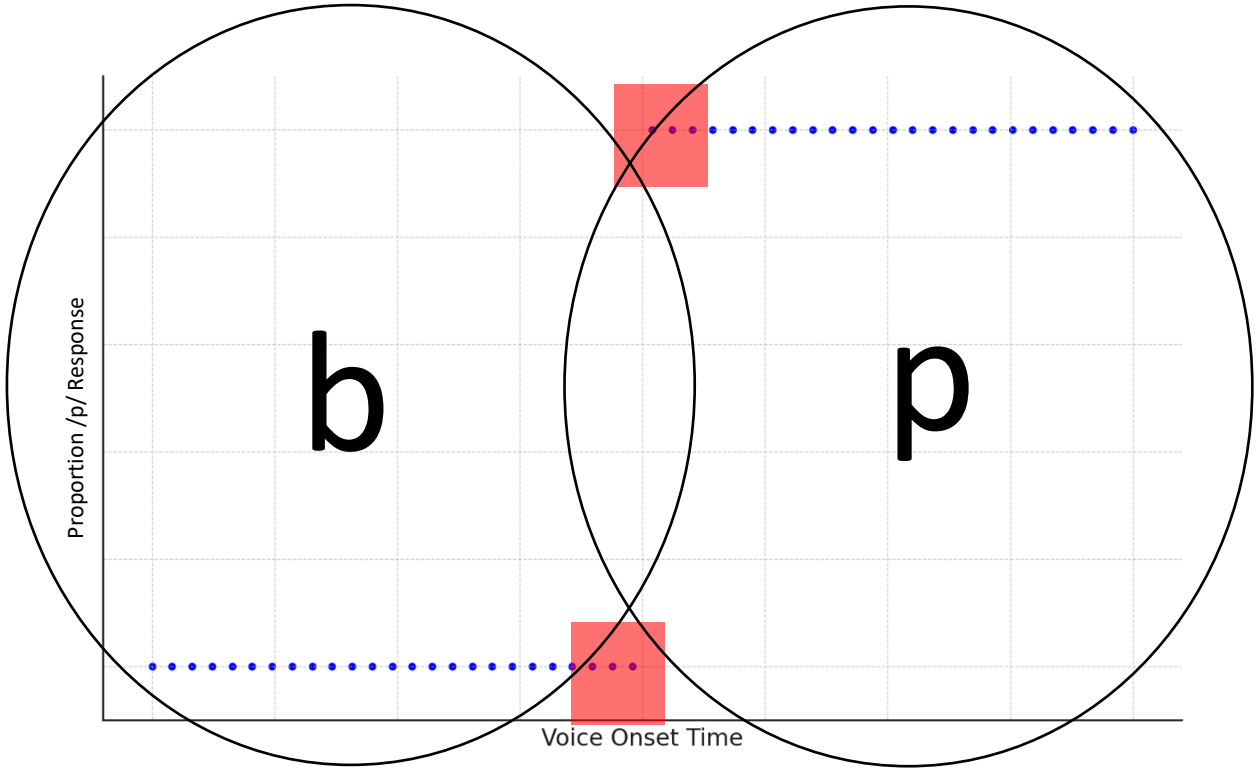


homogeneity: error variance is the same across different levels of the IV

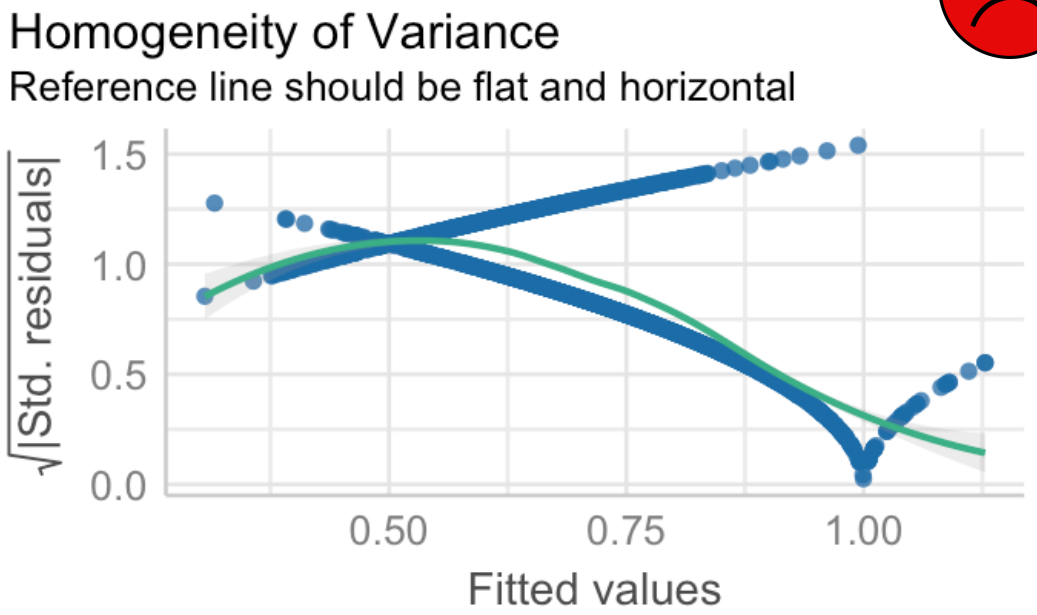


Why logistic regression?

```
performance::check_model(model_all, check=c("normality", "linearity", "homogeneity"))
```



homogeneity: error variance is the same across different levels of the IV



When use logistic regression?

When use logistic regression?

Binary outcome!

When use logistic regression?

Ex: yes vs. no

Binary outcome!

When use logistic regression?

Ex: yes vs. no

Binary outcome!

Ex: b vs. p

When use logistic regression?

Ex: yes vs. no

Ex: accuracy (1 vs. 0)

Binary outcome!

Ex: b vs. p

When use logistic regression?

Ex: yes vs. no

Ex: accuracy (1 vs. 0)

Binary outcome!

Ex: good vs. bad

Ex: b vs. p

When use logistic regression?

Ex: yes vs. no

Ex: accuracy (1 vs. 0)

Binary outcome!

Ex: good vs. bad

Anything with two choices

Ex: b vs. p

Logistic regression

$$Y = \beta_0 + \beta_1 X$$

Logistic regression

$$Y = \beta_0 + \beta_1 X$$

$$\text{logit}(p) = \beta_0 + \beta_1 X$$

Logistic regression

Linear regression: predicting Y value

$$Y = \beta_0 + \beta_1 X$$

$$\text{logit}(p) = \beta_0 + \beta_1 X$$

Logistic regression

Linear regression: predicting Y value

$$Y = \beta_0 + \beta_1 X$$

$$\text{logit}(p) = \beta_0 + \beta_1 X$$

Logistic regression: predicting log odds

Logistic regression

$$\text{logit}(p) = \beta_0 + \beta_1 \text{X}$$

p = the probability of a “success” or the dependent variable being 1

Logistic regression

$$\text{logit}(p) = \beta_0 + \beta_1 \text{X}$$

p = the probability of a “success” or the dependent variable being 1

logit*(p)* = logistic transformation of the probability of “success”

Logistic regression

$$\text{logit}(p) = \beta_0 + \beta_1 X$$

p = the probability of a "success" or the dependent variable being 1

logit(p) = logistic transformation of the probability of "success"

β_0 = intercept

Logistic regression

$$\text{logit}(p) = \beta_0 + \beta_1 \text{X}$$

p = the probability of a “success” or the dependent variable being 1

logit(p) = logistic transformation of the probability of “success”

β_0 = intercept

β_1 = logistic regression coefficient

Logistic regression

$$\text{logit}(p) = \beta_0 + \beta_1 X$$

p = the probability of a "success" or the dependent variable being 1

logit(p) = logistic transformation of the probability of "success"

β_0 = intercept

β_1 = logistic regression coefficient

X = predictor variable

Log odd, odds, probability?

$$\frac{p}{1-p}$$

Log odd, odds, probability?

probability of an event (p)

$$\frac{p}{1-p}$$

Log odd, odds, probability?

probability of an event (p)

$$\frac{p}{1-p}$$

probability of no event (1- p)

Log odd, odds, probability?

probability of an event (p)

Odd: ratio between probability of an event and probability of no event

$$\frac{p}{1-p}$$

probability of no event (1- p)

In our example

They cooked a casserole for him.

The parents grilled a chicken breast for Maria.

Maya baked home-made cookies for them.

Sarah made a burger for the kid.

The chef boiled a few carrots for her.

Amir sliced some vegetables for Luca.

Did you read this exact sentence?

They made Maria home-made cookies.

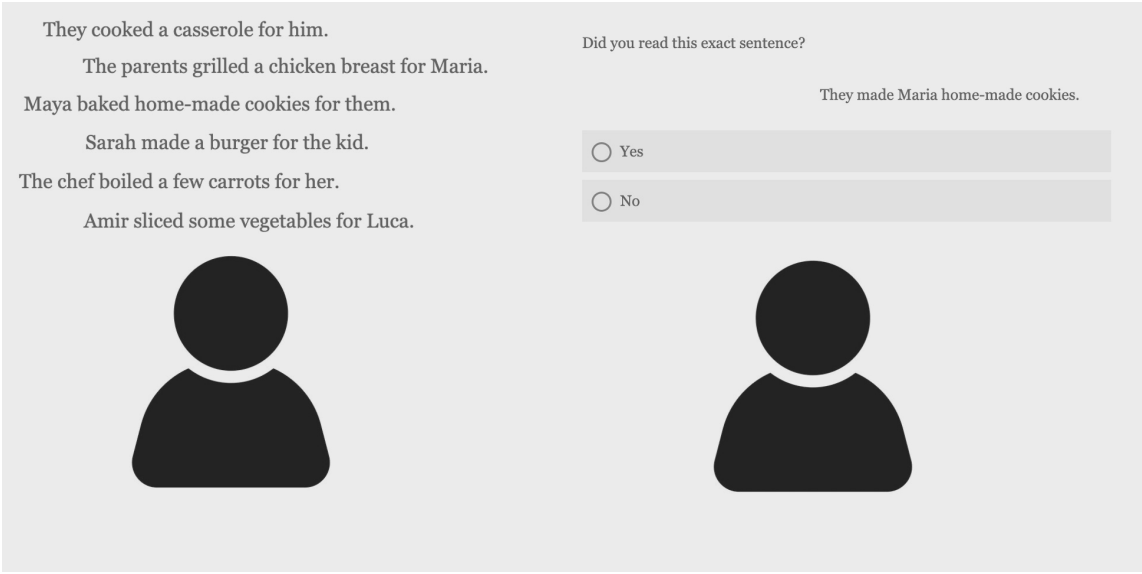
- ◯ Yes
- ◯ No

$$\frac{p}{1-p}$$

In our example

probability of getting it right (hits or correct rejects)

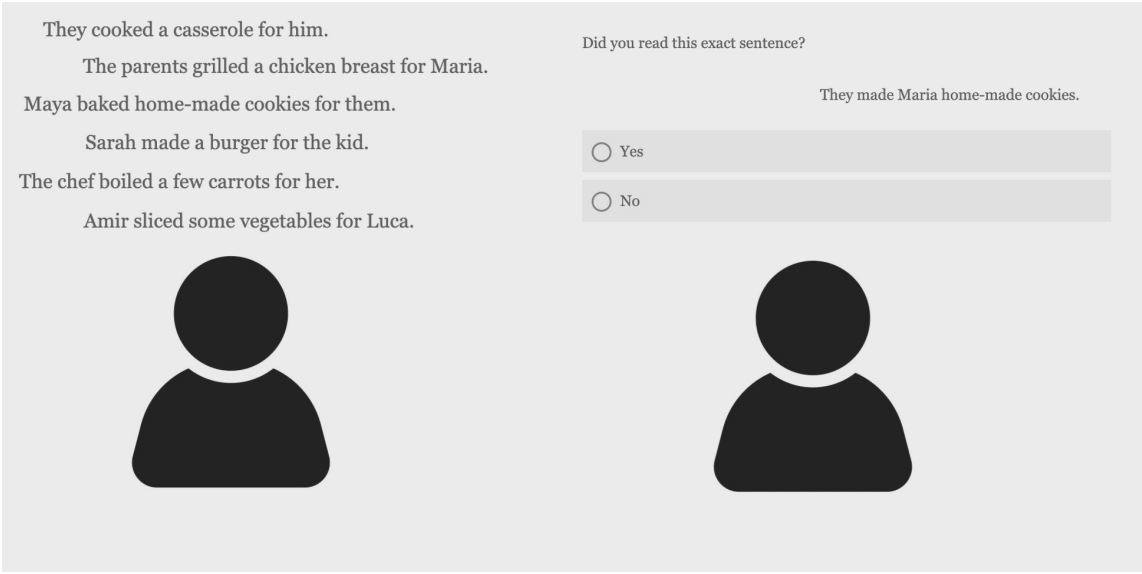
$$\frac{p}{1-p}$$



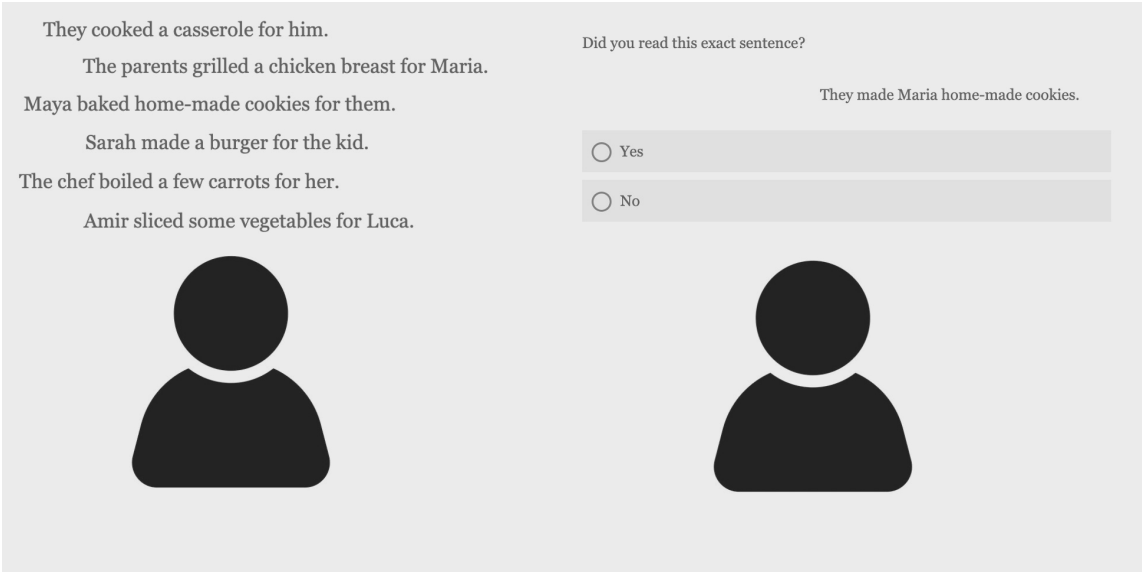
In our example

probability of getting it right (hits or correct rejects)

$$\frac{.4}{1-p}$$



In our example

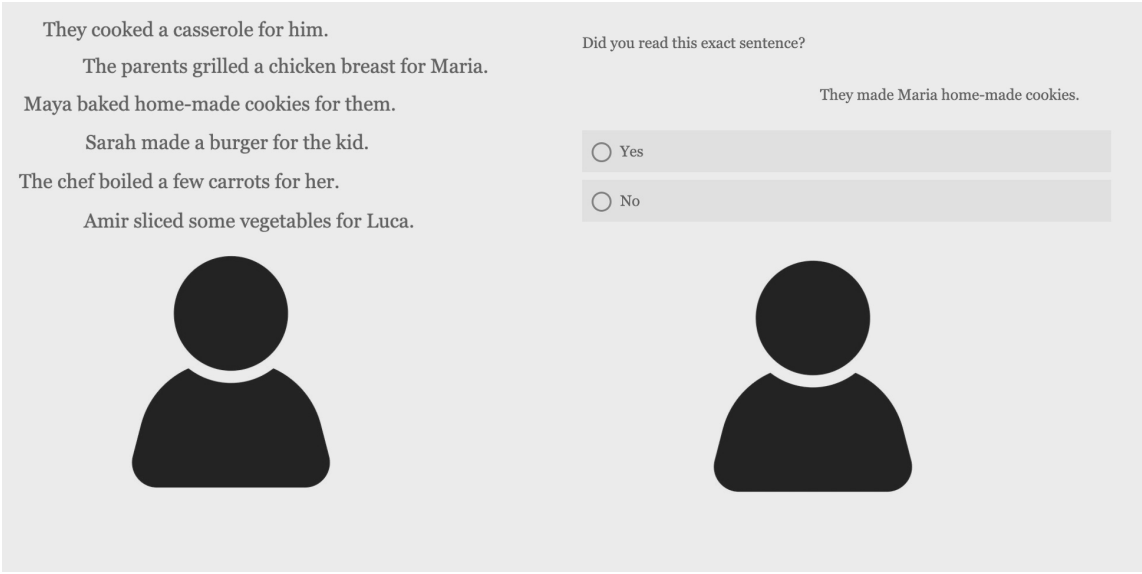


probability of getting it right (hits or correct rejects)

$$\frac{p}{1 - .4}$$

probability of getting it wrong (misses or false alarms)

In our example



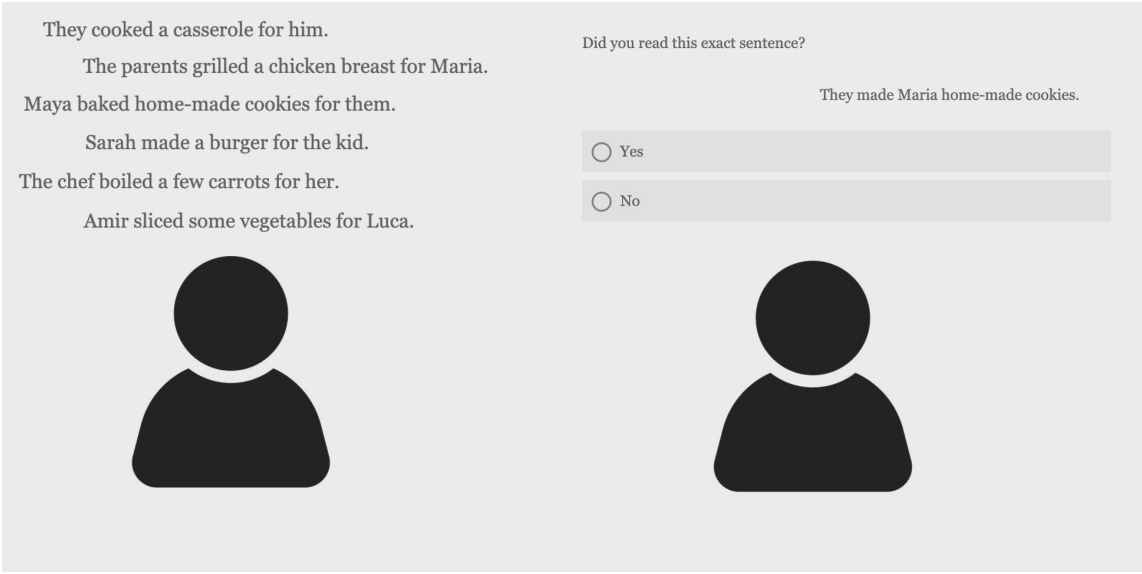
probability of getting it right (hits or correct rejects)

probability of getting it wrong (misses or false alarms)

$$\frac{p}{1-p}$$

Odd: ratio between correct and incorrect probabilities

In our example



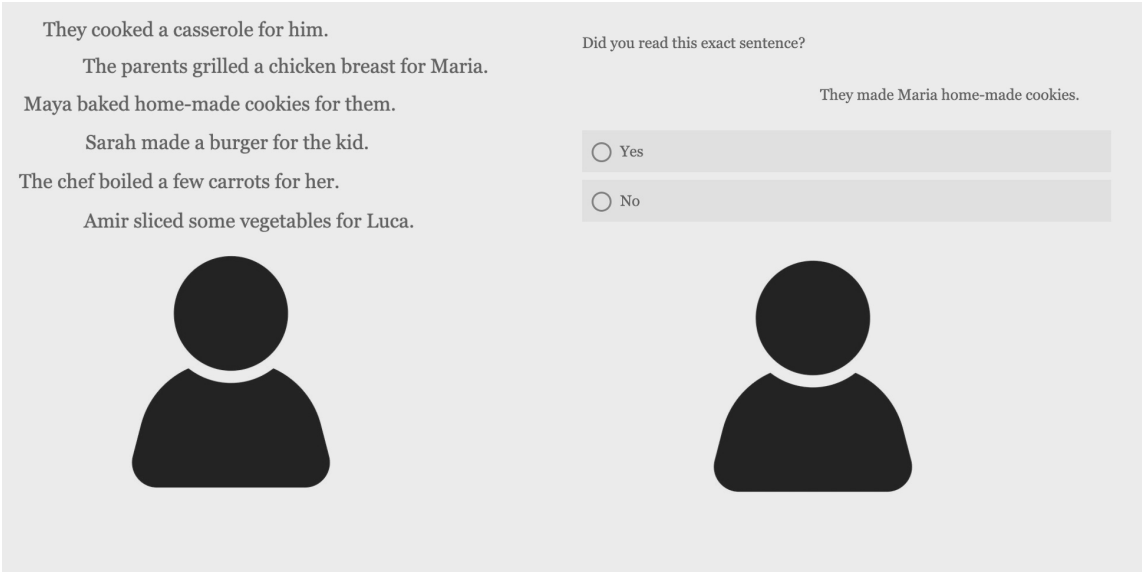
probability of getting it right (hits or correct rejects)

probability of getting it wrong (misses or false alarms)

$$\frac{.4}{1 - .4}$$

Odd: ratio between correct and incorrect probabilities

In our example



probability of getting it right (hits or correct rejects)

.67

probability of getting it wrong (misses or false alarms)

Odd: ratio between correct and incorrect probabilities

Logistic regression

$$\text{logit}(p) = \beta_0 + \beta_1 \text{X}$$

Logistic regression

$$\text{logit}(p) = \beta_0 + \beta_1 \text{X}$$

logit of probability

Logistic regression

$$ln(\frac{p}{1-p}) = \beta_0 + \beta_1 X$$

Logistic regression

$$ln(\frac{p}{1-p}) = \beta_0 + \beta_1 X$$

natural log of odds

Log odd, odds, probability?

$$ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X$$

logit/log odd

Log odd, odds, probability?

$$ln(\frac{p}{1-p}) = \beta_0 + \beta_1 X$$

Log odd, odds, probability?

$$ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X$$

odds

Log odd, odds, probability?

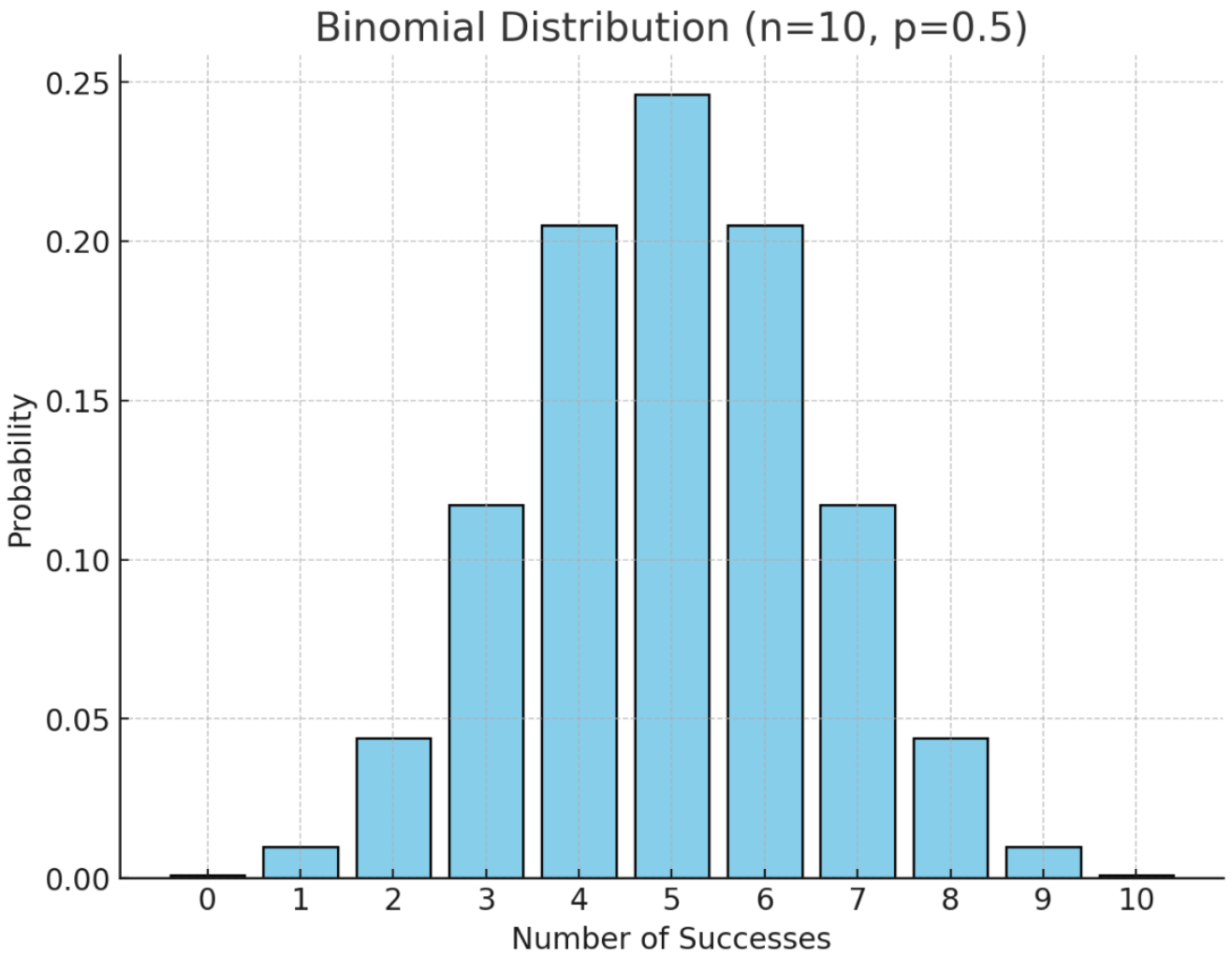
$$ln(\frac{p}{1-p}) = \beta_0 + \beta_1 X$$

Log odd, odds, probability?

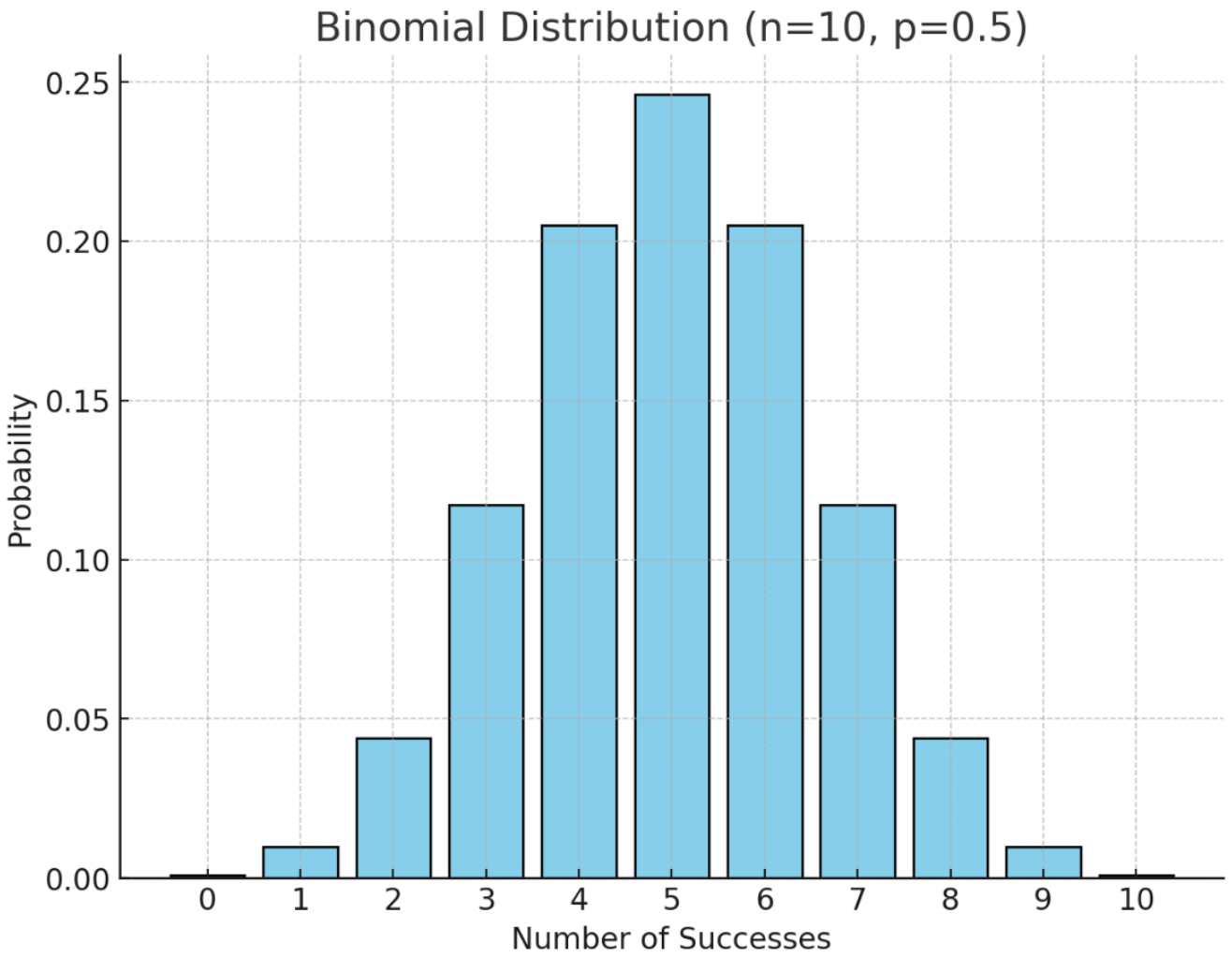
$$ln(\frac{p}{1-p}) = \beta_0 + \beta_1 X$$

probability

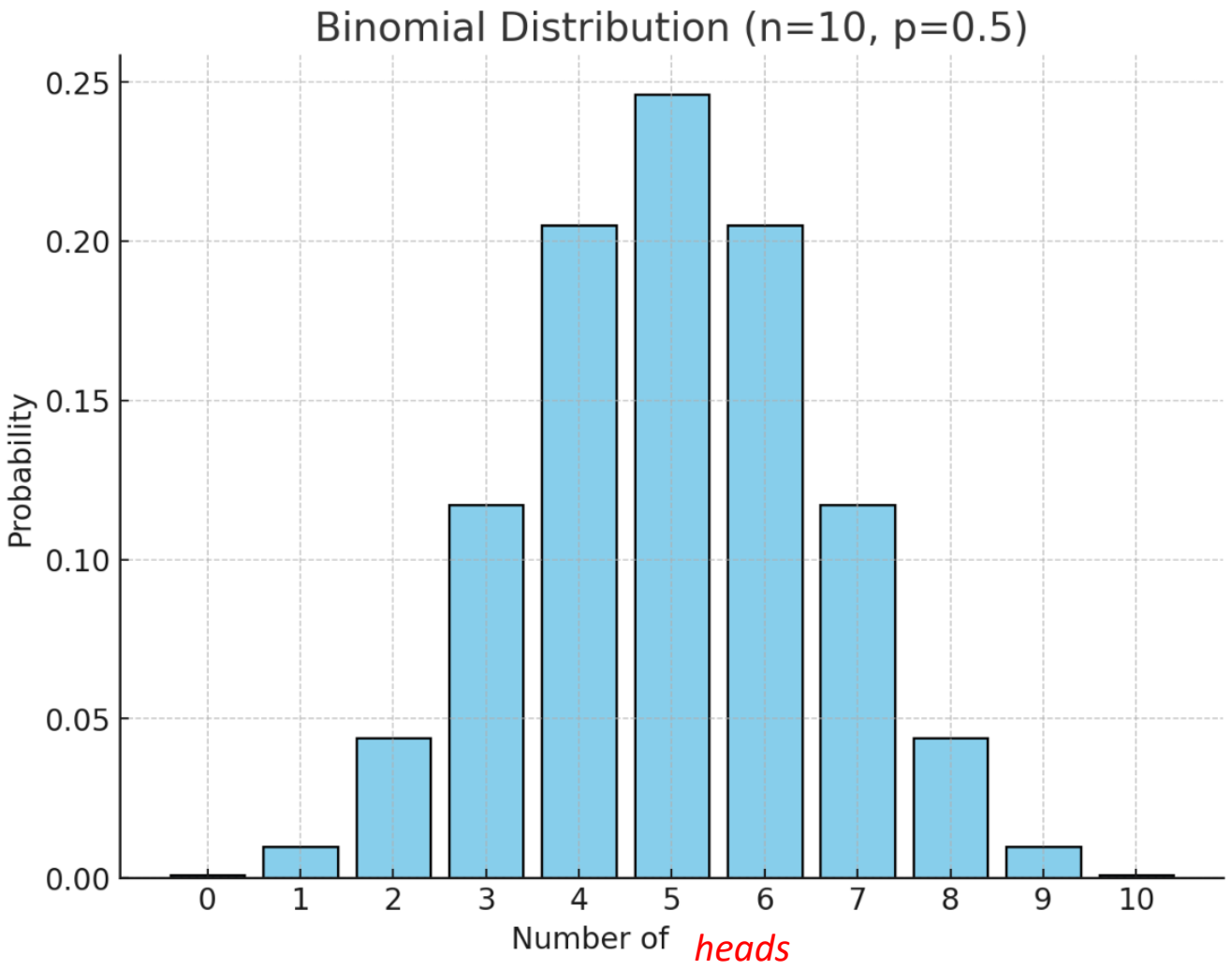
Binomial distribution



Binomial distribution

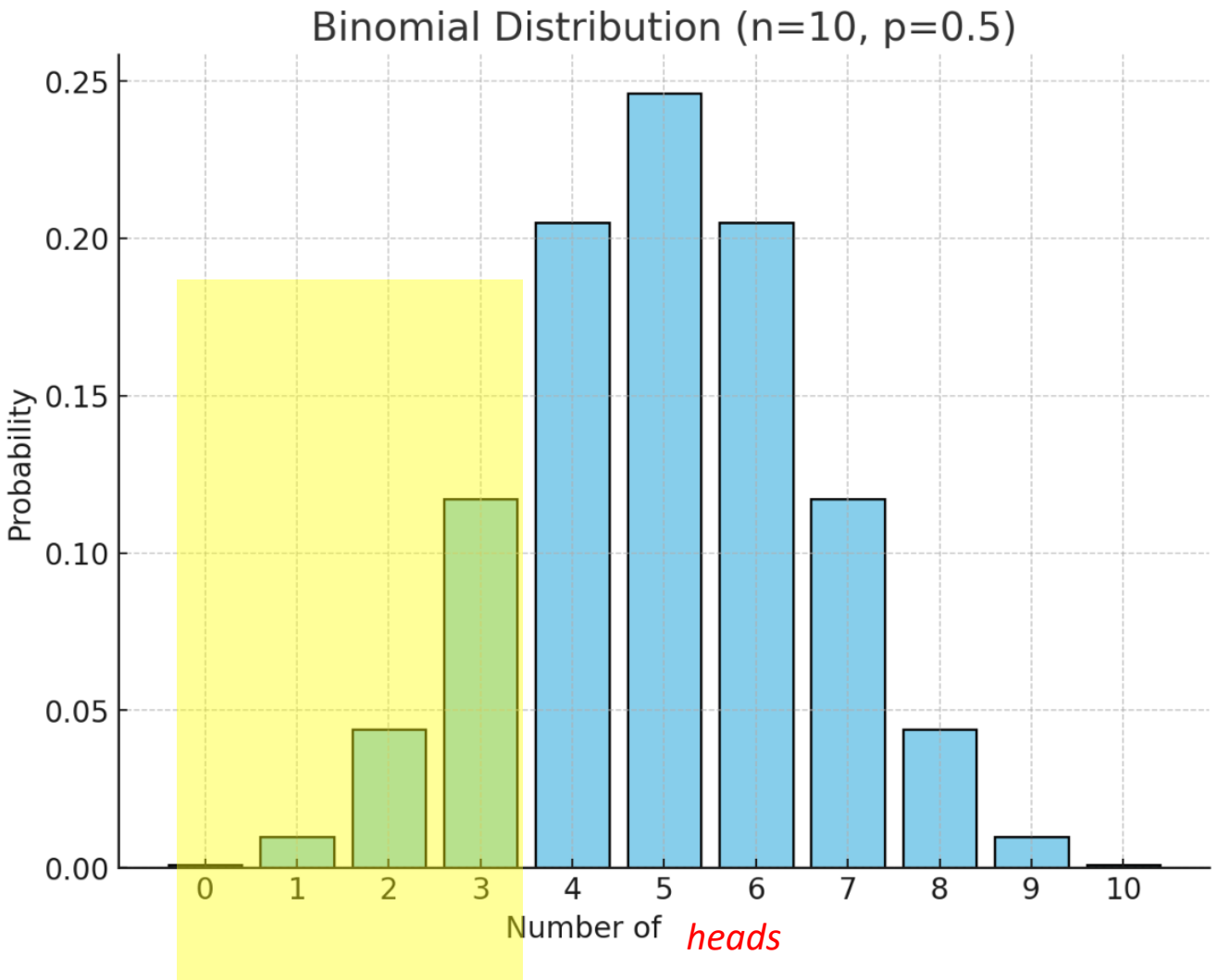


Binomial distribution

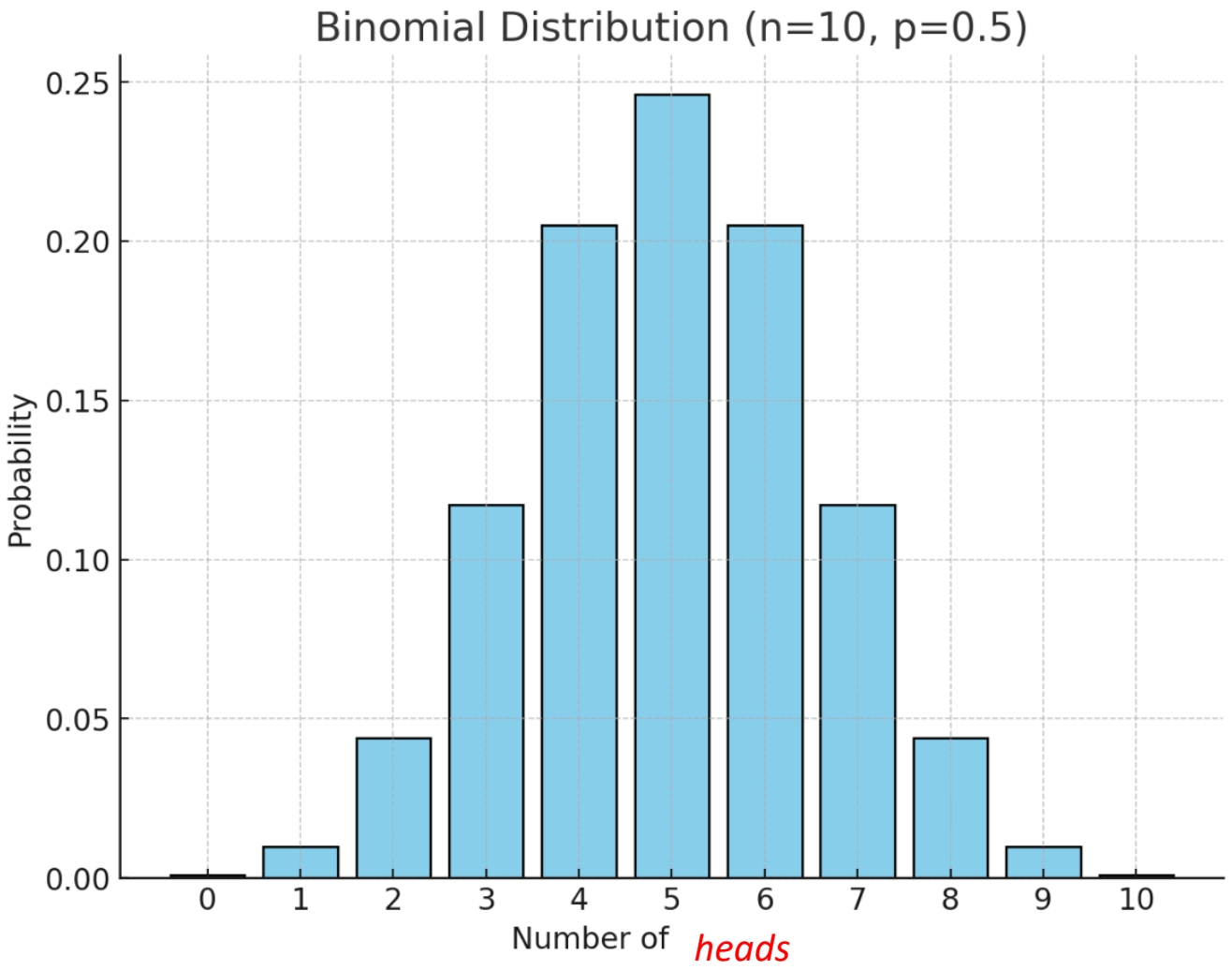


Binomial distribution

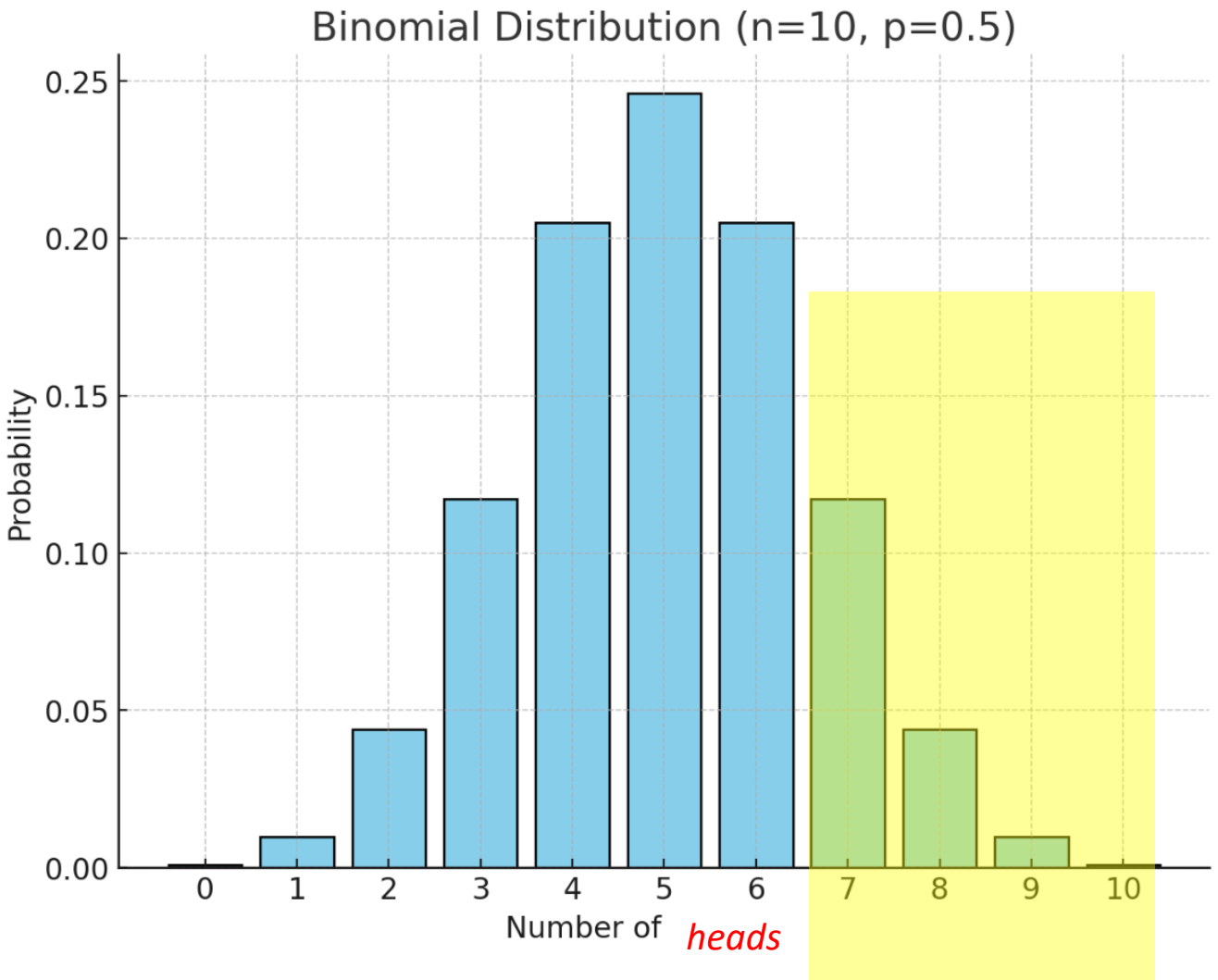
unlikely to get all tails



Binomial distribution

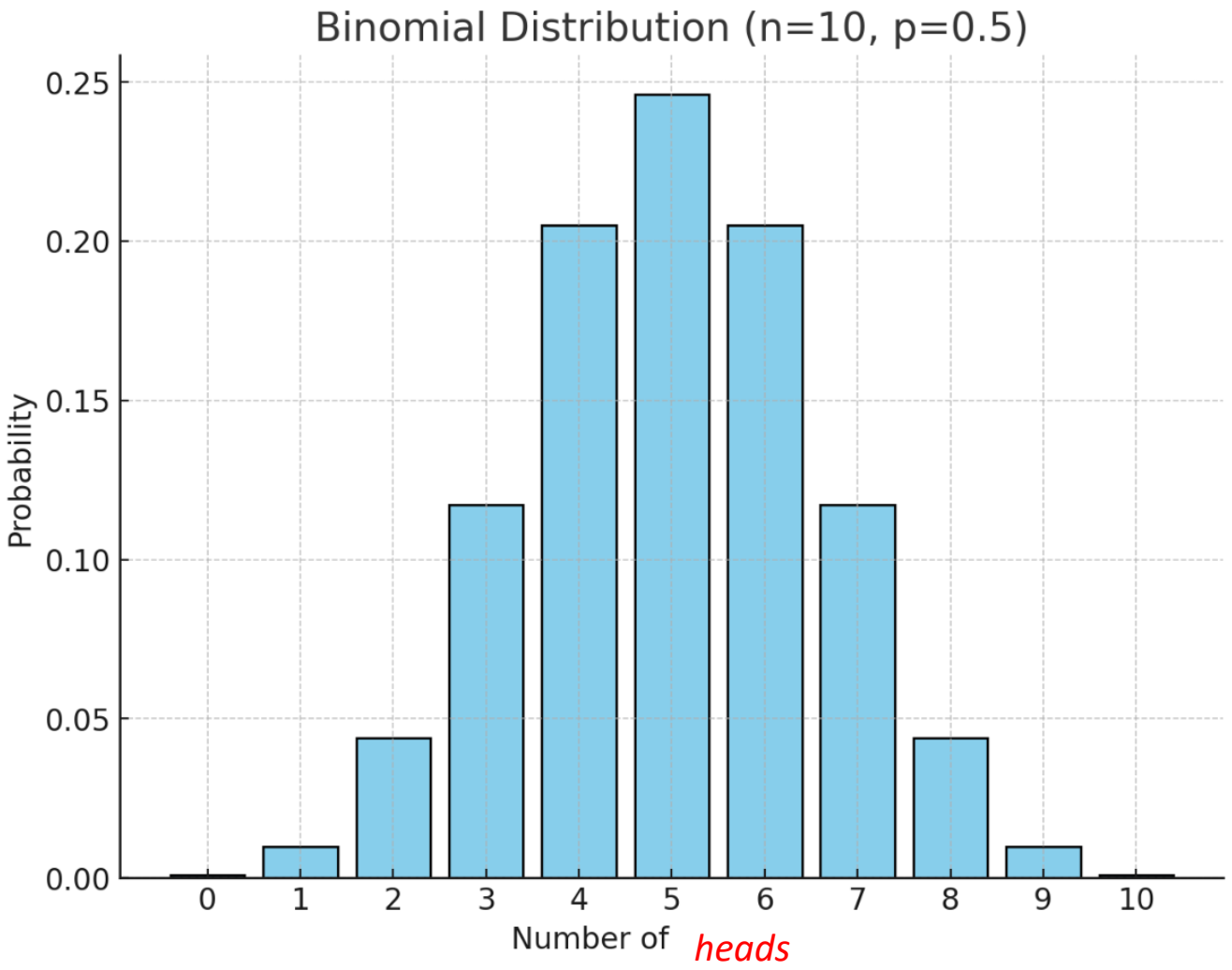


Binomial distribution



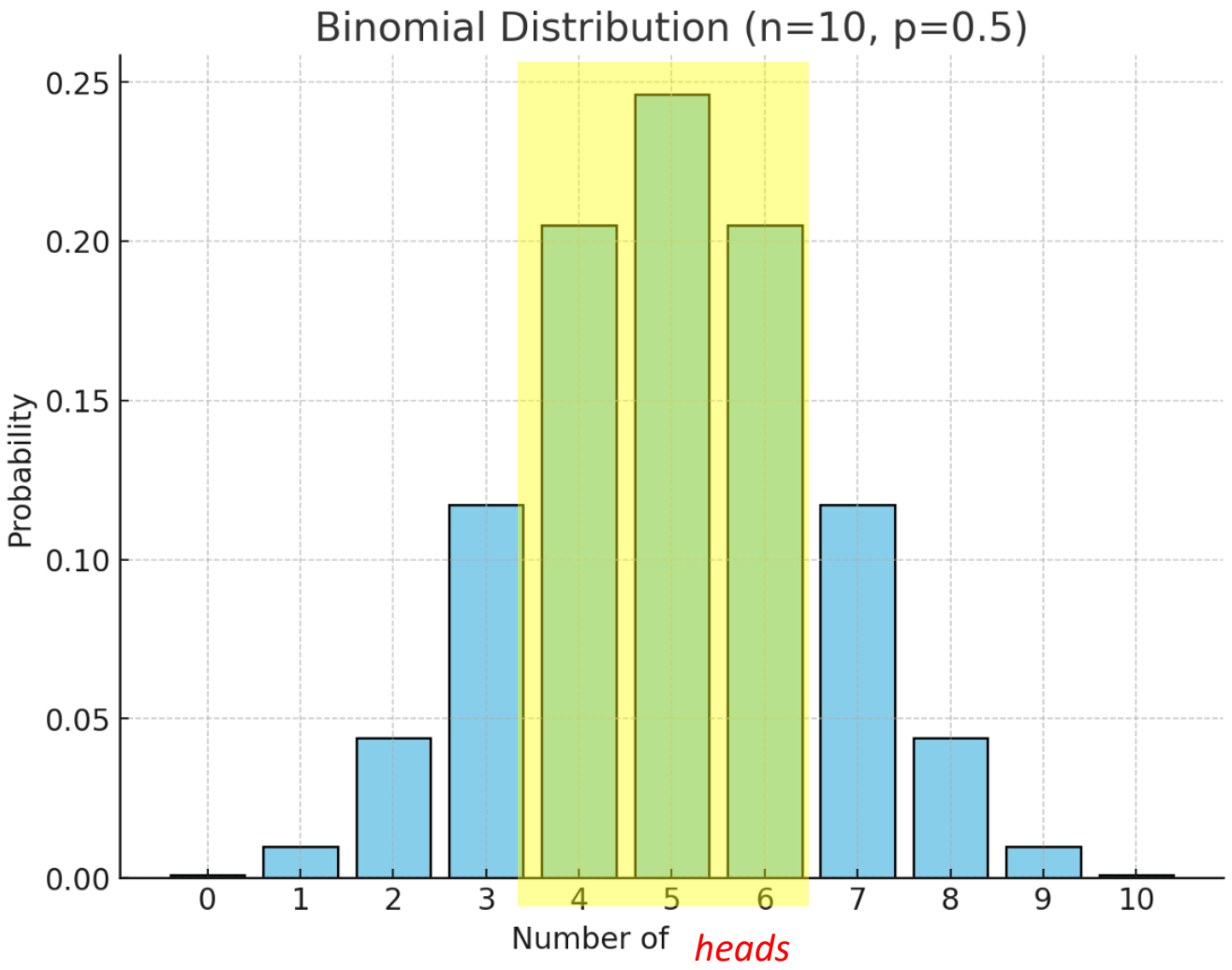
unlikely to get all heads

Binomial distribution

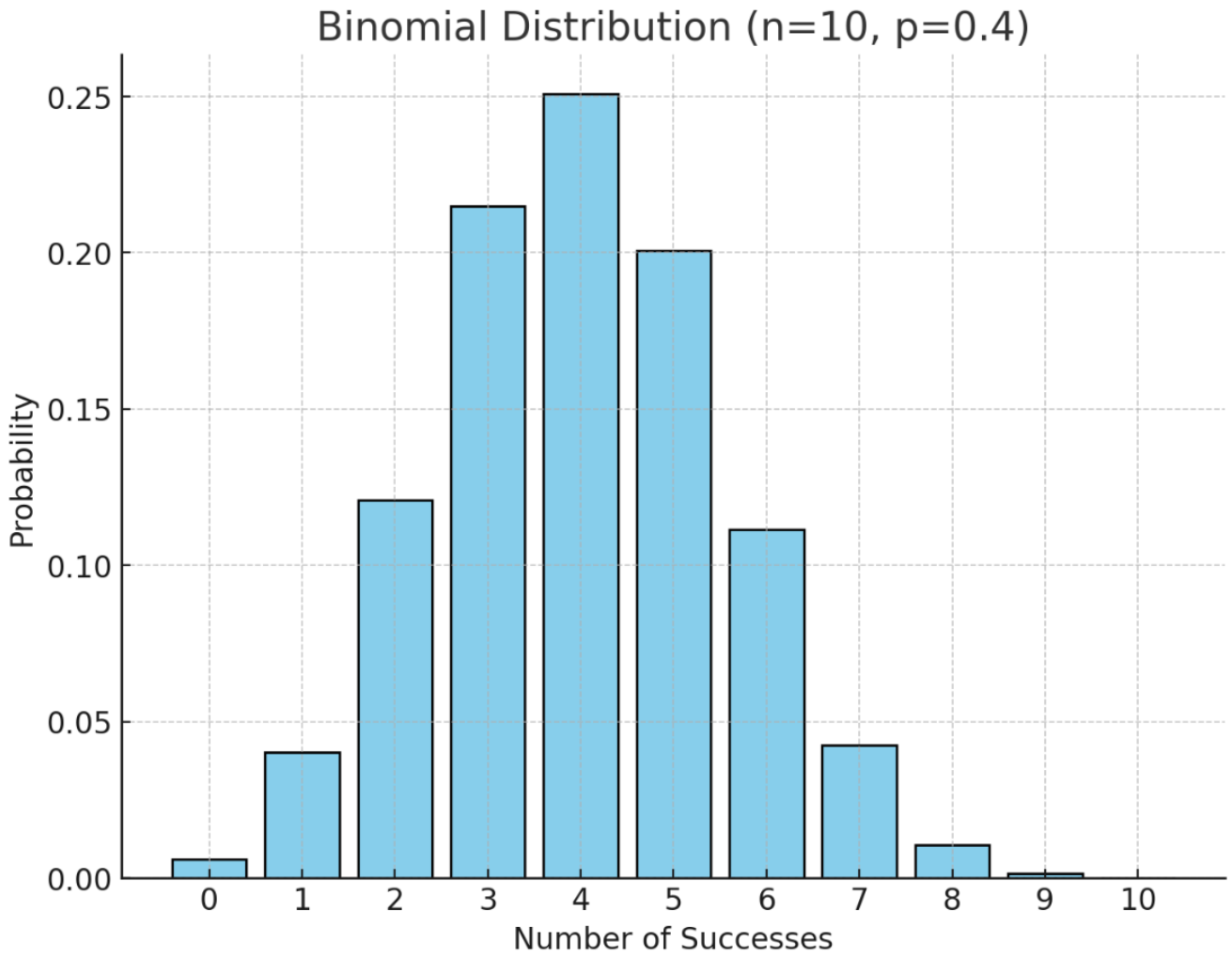


Binomial distribution

most likely to get even split



Our example



They cooked a casserole for him.

The parents grilled a chicken breast for Maria.

Maya baked home-made cookies for them.

Sarah made a burger for the kid.

The chef boiled a few carrots for her.

Amir sliced some vegetables for Luca.

Did you read this exact sentence?

They made Maria home-made cookies.

- ◯ Yes
- ◯ No

They cooked a casserole for him.

The parents grilled a chicken breast for Maria.

Maya baked home-made cookies for them.

Sarah made a burger for the kid.

The chef boiled a few carrots for her.

Amir sliced some vegetables for Luca.

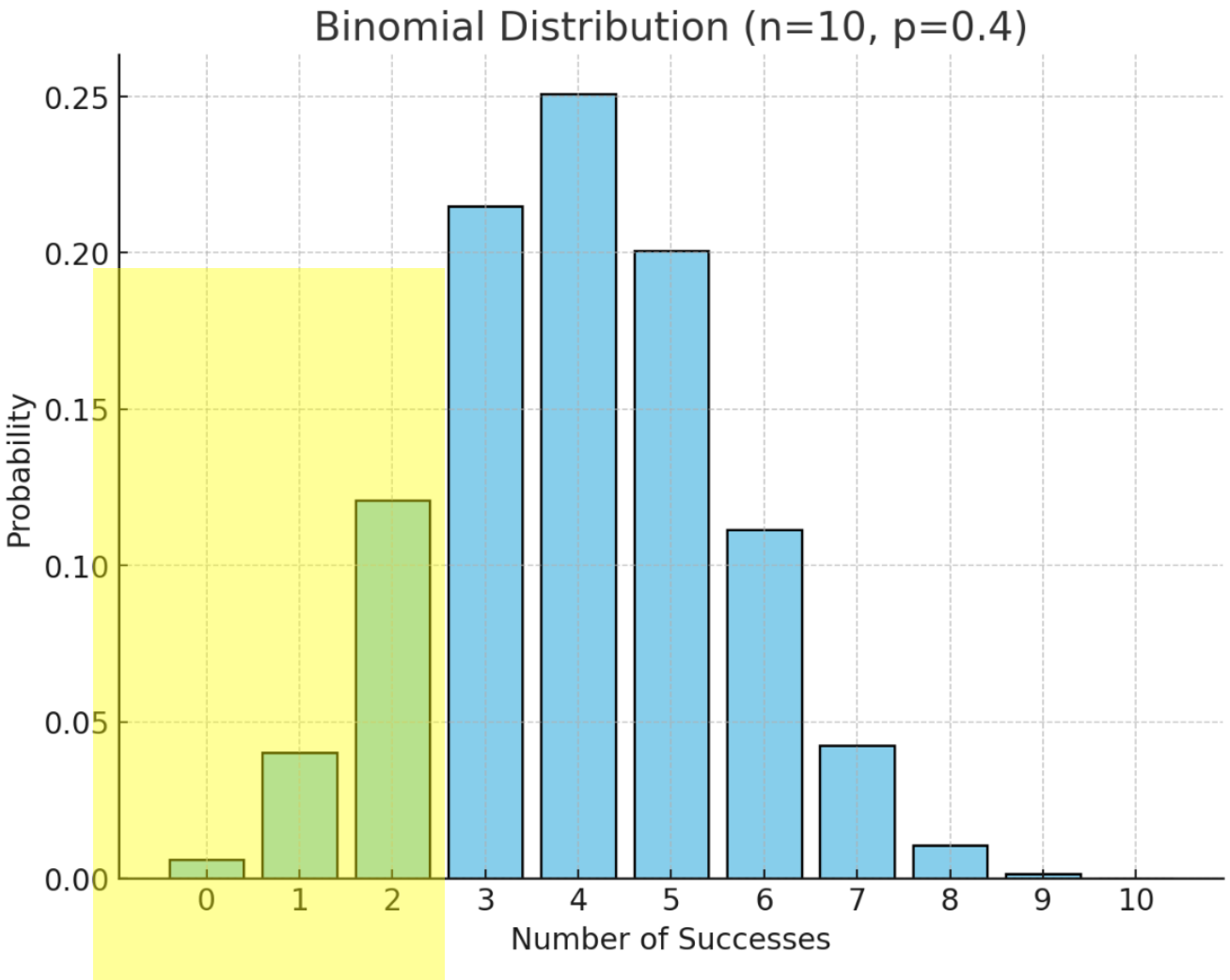
Did you read this exact sentence?

They made Maria home-made cookies.

Yes

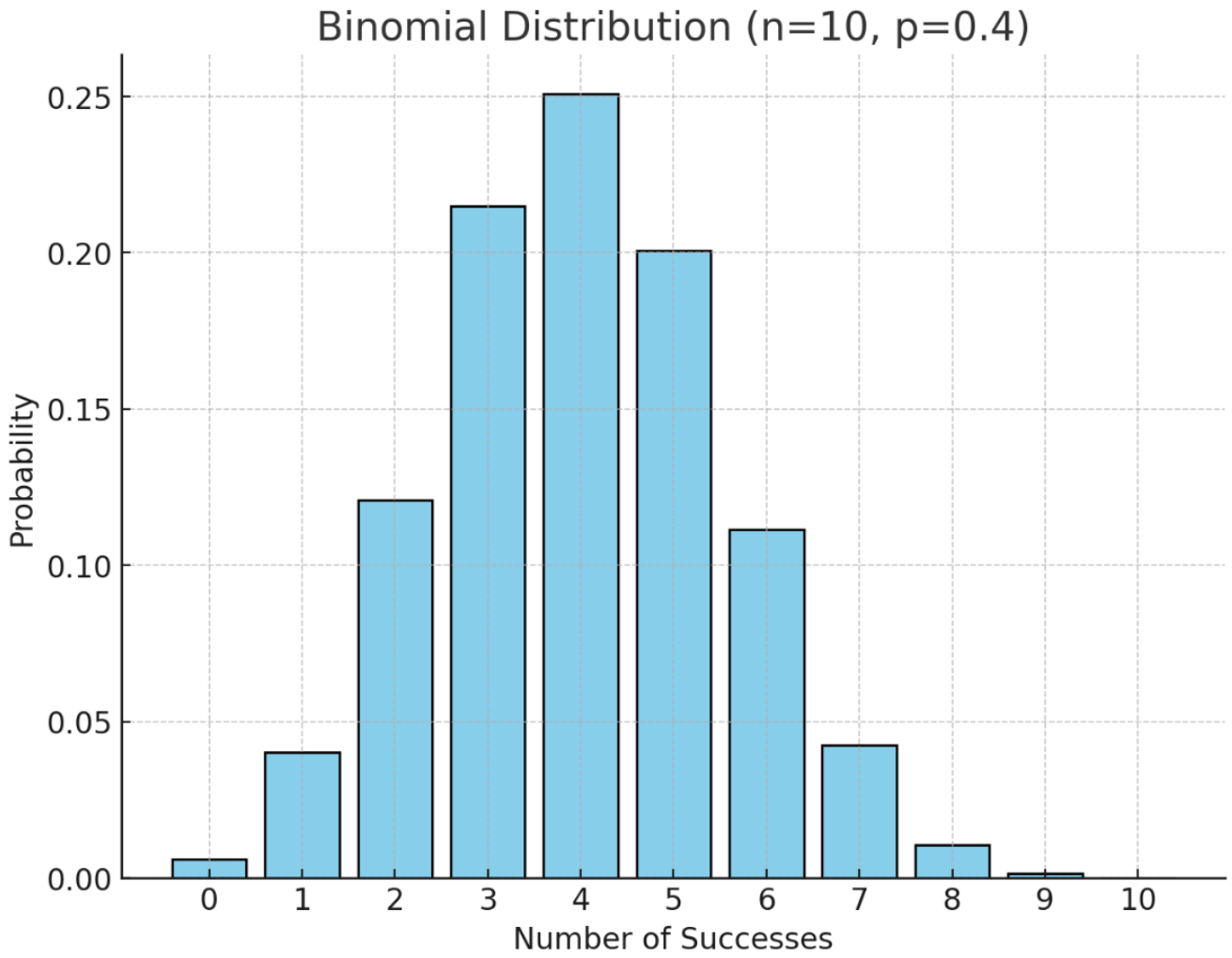
No

Our example



unlikely to only get them all wrong

Our example



They cooked a casserole for him.

The parents grilled a chicken breast for Maria.

Maya baked home-made cookies for them.

Sarah made a burger for the kid.

The chef boiled a few carrots for her.

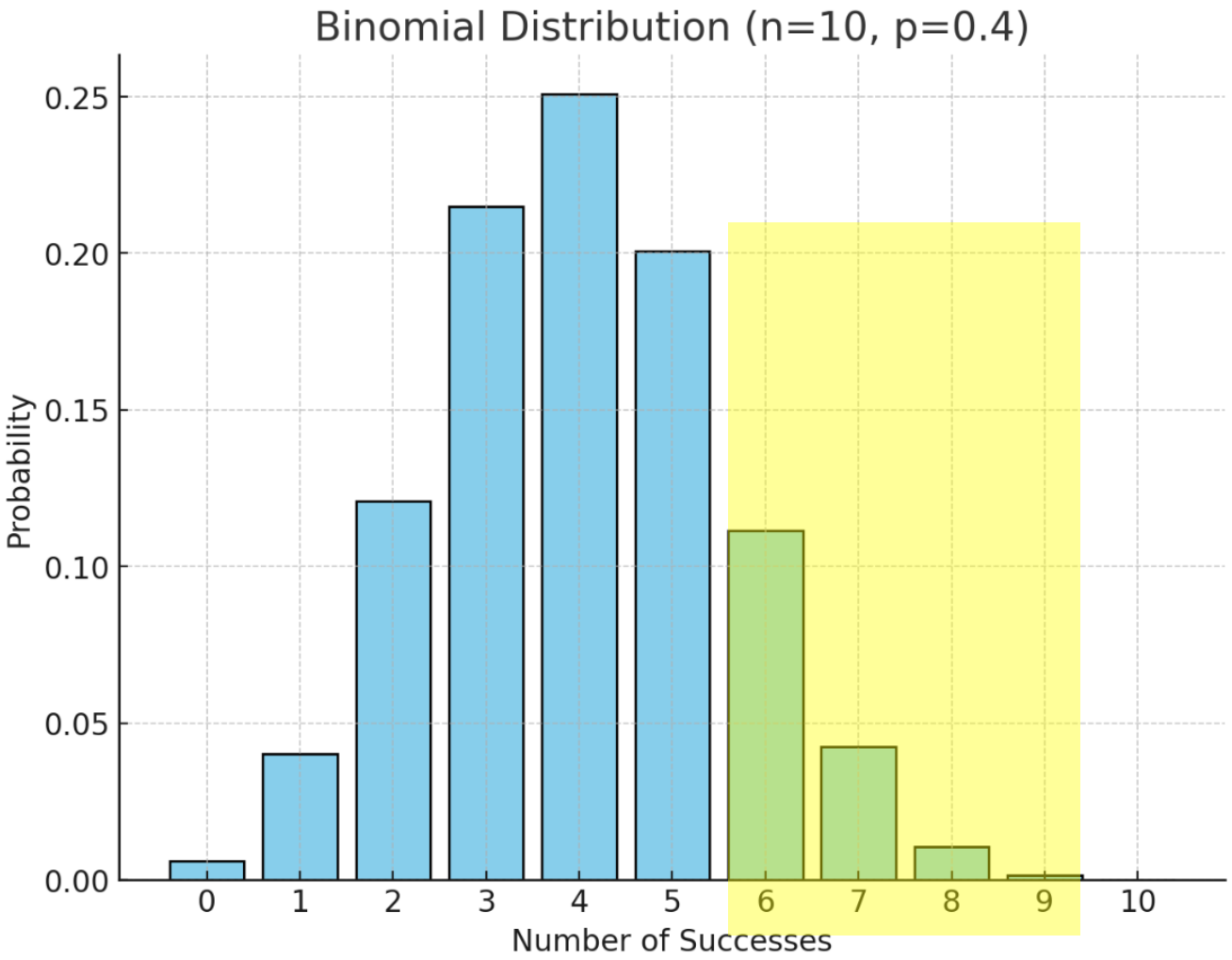
Amir sliced some vegetables for Luca.

Did you read this exact sentence?

They made Maria home-made cookies.

- ◯ Yes
- ◯ No

Our example



They cooked a casserole for him.

The parents grilled a chicken breast for Maria.

Maya baked home-made cookies for them.

Sarah made a burger for the kid.

The chef boiled a few carrots for her.

Amir sliced some vegetables for Luca.

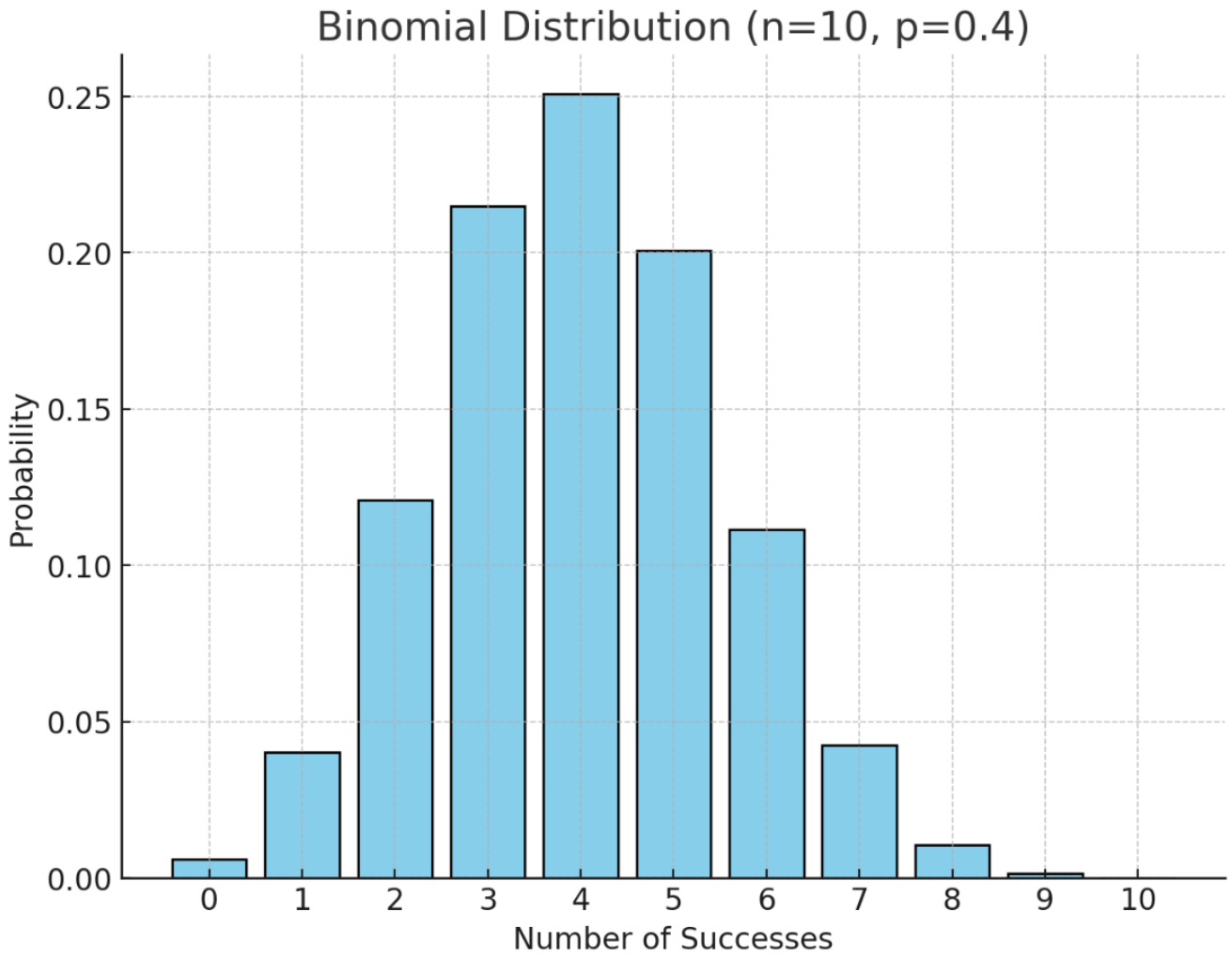
Did you read this exact sentence?

They made Maria home-made cookies.

- ◯ Yes
- ◯ No

unlikely to only get them all right

Our example



They cooked a casserole for him.

The parents grilled a chicken breast for Maria.

Maya baked home-made cookies for them.

Sarah made a burger for the kid.

The chef boiled a few carrots for her.

Amir sliced some vegetables for Luca.

Did you read this exact sentence?

They made Maria home-made cookies.

- ◯ Yes
- ◯ No

They cooked a casserole for him.

The parents grilled a chicken breast for Maria.

Maya baked home-made cookies for them.

Sarah made a burger for the kid.

The chef boiled a few carrots for her.

Amir sliced some vegetables for Luca.

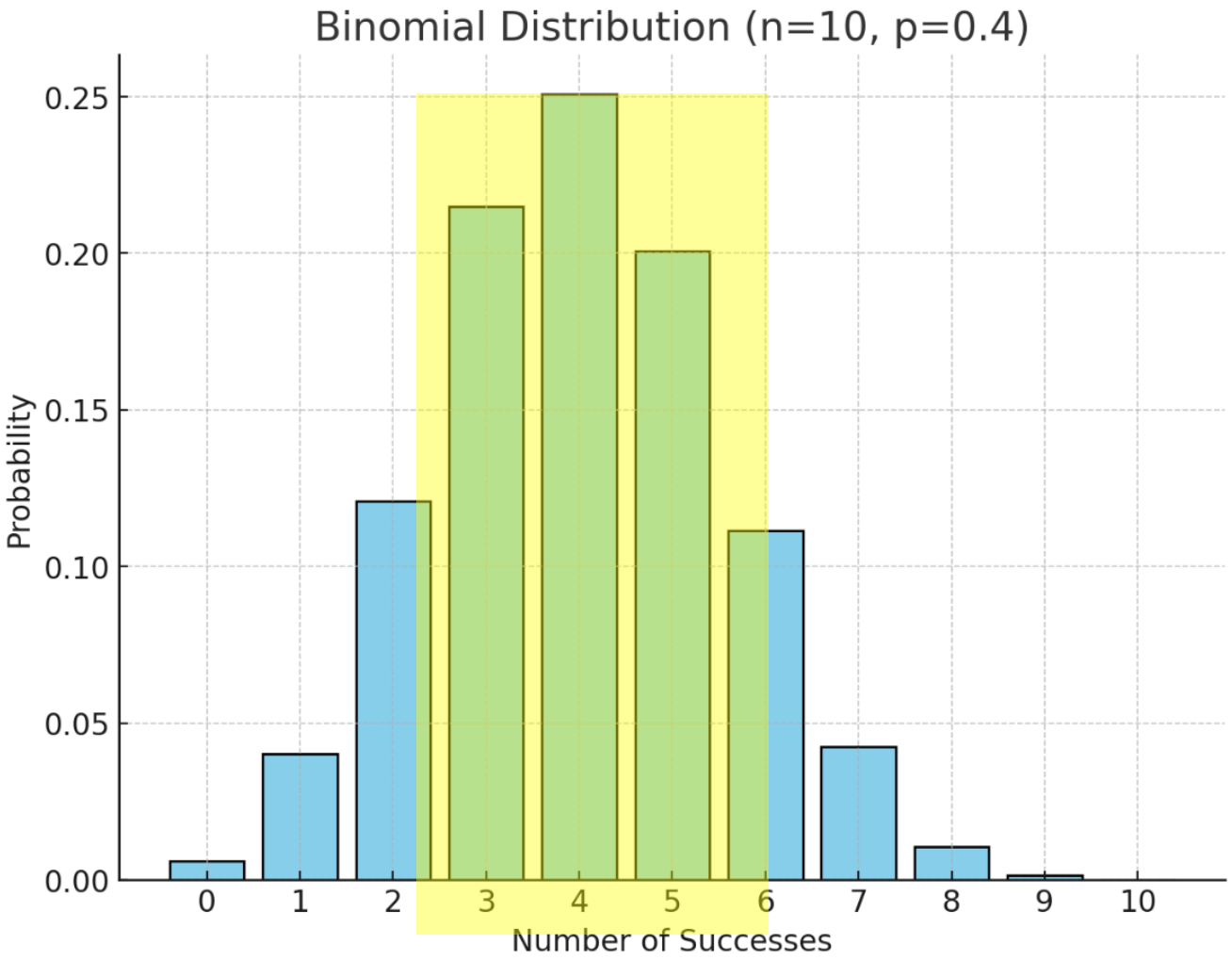
Did you read this exact sentence?

They made Maria home-made cookies.

Yes

No

Our example



most likely to be in between

In R!

lm -> glm

```
model_all <- glm(accuracy ~  condition, family = binomial, data)
```

In R!

lm -> glm

```
model_all <- glm(accuracy ~  condition, family = binomial, data)

model_all <- glmer(accuracy ~  condition + (1 | RandomID) +  (1| sentence/item),family = binomial, data)
```

lmer -> glmer

In R!

lm -> glm

```
model_all <- glm(accuracy ~  condition, family = binomial, data)

model_all <- glmer(accuracy ~  condition + (1 | RandomID) +  (1| sentence/item),family = binomial, data)
```

lmer -> glmer

In R!

also can specify the ”link”

```
model_all <- glmer(accuracy ~  condition + (1 | RandomID) +  (1| sentence/item),family = binomial (link = "logit"), data)

model_all <- glmer(accuracy ~  condition + (1 | RandomID) +  (1| sentence/item),family = binomial (link = "probit"), data)
```

In R!

also can specify the ”link”

```
model_all <- glmer(accuracy ~  condition + (1 | RandomID) +  (1| sentence/item),family = binomial (link = "logit"), data)

model_all <- glmer(accuracy ~  condition + (1 | RandomID) +  (1| sentence/item),family = binomial (link = "probit"), data)
```

In R!

also can specify the ”link”

```
model_all <- glmer(accuracy ~  condition + (1 | RandomID) +  (1| sentence/item),family = binomial (link = "logit"), data)

model_all <- glmer(accuracy ~  condition + (1 | RandomID) +  (1| sentence/item),family = binomial (link = "probit")  data)
```

Thank you!