

Model design in R: Class 1

Alexandre Cremers

August 12th, 2019

What are we doing here?

What this course is not:

What are we doing here?

What this course is not:

- An introduction to R

What are we doing here?

What this course is not:

- An introduction to R
- A proper introduction to stats

What are we doing here?

What this course is not:

- An introduction to R
- A proper introduction to stats

So what is it?

What are we doing here?

What this course is not:

- An introduction to R
- A proper introduction to stats

So what is it?

- ☞ Identifying some issues with data analysis that are somewhat specific to Semantics & Pragmatics.

What are we doing here?

What this course is not:

- An introduction to R
- A proper introduction to stats

So what is it?

- ☞ Identifying some issues with data analysis that are somewhat specific to Semantics & Pragmatics.
- ☞ Discussing possible solutions and their implementation.

Plan

- Today: introducing the problem and the solution
- Tomorrow: formalize what we'll do today
- Wednesday/Thursday: Applications
(either my examples or yours!)
- Friday: A few more problems and a few more solutions. . .

Plan

- Today: introducing the problem and the solution
- Tomorrow: formalize what we'll do today
- Wednesday/Thursday: Applications
(either my examples or yours!)
- Friday: A few more problems and a few more solutions. . .

We may switch Thursday and Friday if necessary.

Experimental semantics and pragmatics

What's a typical empirical issue in SemPrag?

Experimental semantics and pragmatics

What's a typical empirical issue in SemPrag?

- We want to know what a sentence S means.

Experimental semantics and pragmatics

What's a typical empirical issue in SemPrag?

- We want to know what a sentence S means.
- We usually have at least two candidate meanings, A and B .

Experimental semantics and pragmatics

What's a typical empirical issue in SemPrag?

- We want to know what a sentence S means.
- We usually have at least two candidate meanings, A and B .
- This leaves at least three possible options:

Experimental semantics and pragmatics

What's a typical empirical issue in SemPrag?

- We want to know what a sentence S means.
- We usually have at least two candidate meanings, A and B .
- This leaves at least three possible options:
 - S only means A

Experimental semantics and pragmatics

What's a typical empirical issue in SemPrag?

- We want to know what a sentence S means.
- We usually have at least two candidate meanings, A and B .
- This leaves at least three possible options:
 - S only means A
 - S only means B

Experimental semantics and pragmatics

What's a typical empirical issue in SemPrag?

- We want to know what a sentence S means.
- We usually have at least two candidate meanings, A and B .
- This leaves at least three possible options:
 - S only means A
 - S only means B
 - S is ambiguous between A and B

Experimental semantics and pragmatics

What's a typical empirical issue in SemPrag?

- We want to know what a sentence S means.
- We usually have at least two candidate meanings, A and B .
- This leaves at least three possible options:
 - S only means A
 - S only means B
 - S is ambiguous between A and B
- We might also accept that S is ambiguous and be interested in figuring out in which contexts each reading is available,

Experimental semantics and pragmatics

What's a typical empirical issue in SemPrag?

- We want to know what a sentence S means.
- We usually have at least two candidate meanings, A and B .
- This leaves at least three possible options:
 - S only means A
 - S only means B
 - S is ambiguous between A and B
- We might also accept that S is ambiguous and be interested in figuring out in which contexts each reading is available,

Experimental semantics and pragmatics

What's a typical empirical issue in SemPrag?

- We want to know what a sentence S means.
- We usually have at least two candidate meanings, A and B .
- This leaves at least three possible options:
 - S only means A
 - S only means B
 - S is ambiguous between A and B
- We might also accept that S is ambiguous and be interested in figuring out in which contexts each reading is available, but this amounts to the same question relativized to a context (What does S mean in context C ?)

Experimental semantics and pragmatics

Most common method:

Experimental semantics and pragmatics

Most common method: The truth-value judgment task.

Experimental semantics and pragmatics

Most common method: The truth-value judgment task.

- Most often, one reading entails the other ($B \rightarrow A$).

Experimental semantics and pragmatics

Most common method: The truth-value judgment task.

- Most often, one reading entails the other ($B \rightarrow A$).
- We can construct a situation in which A is true and B is false.

Experimental semantics and pragmatics

Most common method: The truth-value judgment task.

- Most often, one reading entails the other ($B \rightarrow A$).
- We can construct a situation in which A is true and B is false.

Some elephants have trunks

False

True

Bott&Noveck 2004

Experimental semantics and pragmatics

Most common method: The truth-value judgment task.

- Most often, one reading entails the other ($B \rightarrow A$).
- We can construct a situation in which A is true and B is false.
- Asking participants whether they think that the sentence is true or false informs us on what their preferred reading is.

Some elephants have trunks

False

True

Bott&Noveck 2004

Predictions in Experimental SemPrag

S potentially ambiguous between readings A (true) and B (false).

Predictions in Experimental SemPrag

S potentially ambiguous between readings A (true) and B (false).

- If A is available, S should be judged true significantly more than an unambiguous false control

Predictions in Experimental SemPrag

S potentially ambiguous between readings A (true) and B (false).

- If A is available, S should be judged true significantly more than an unambiguous false control
- If B is available, S should be judged true significantly more than an unambiguous true control

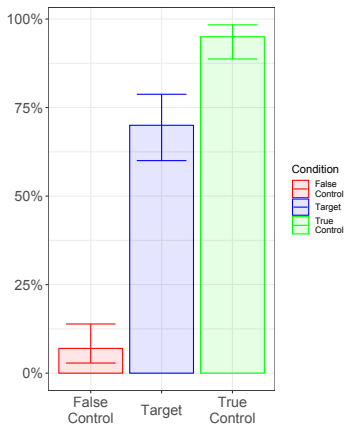
Predictions in Experimental SemPrag

S potentially ambiguous between readings A (true) and B (false).

- If A is available, S should be judged true significantly more than an unambiguous false control
 - If B is available, S should be judged true significantly more than an unambiguous true control
- ☞ We can simply compare our target to true and false controls.

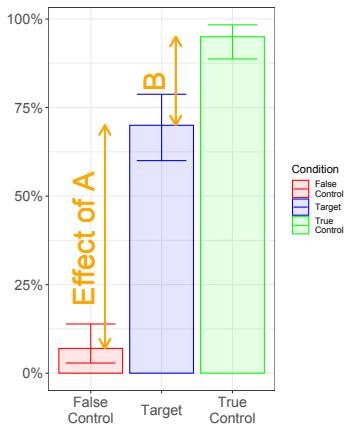
Predictions in Experimental SemPrag

S potentially ambiguous between readings A (true) and B (false).



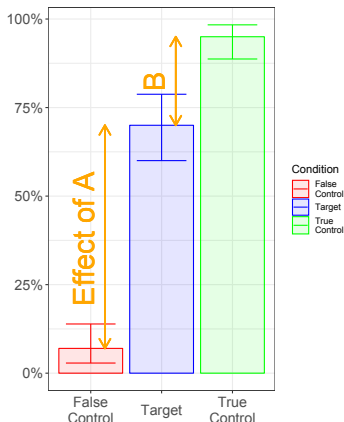
Predictions in Experimental SemPrag

S potentially ambiguous between readings A (true) and B (false).



Predictions in Experimental SemPrag

S potentially ambiguous between readings A (true) and B (false).



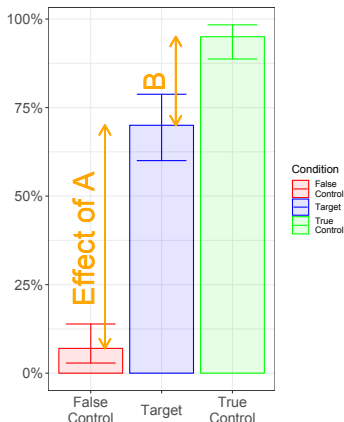
Possible statistical model:

```
Cond <- factor(Cond, levels=  
c("target", "false", "true"))
```

```
glm(Response~Cond)
```


Predictions in Experimental SemPrag

S potentially ambiguous between readings A (true) and B (false).



Possible statistical model:

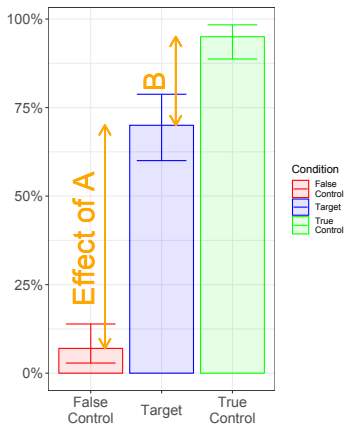
```
Cond <- factor(Cond, levels=
c("target", "false", "true"))
```

```
glm(Response~Cond)
```

-Cond:false gives an estimate of A ,
Cond:true gives an estimate of B .

Predictions in Experimental SemPrag

S potentially ambiguous between readings A (true) and B (false).



Possible statistical model:

```
Cond <- factor(Cond, levels=
c("target", "false", "true"))
```

```
glm(Response~Cond)
```

-Cond:false gives an estimate of A ,
Cond:true gives an estimate of B .

☞ We have a direct mapping between
Cond and our theoretical parameters.

Interim conclusion

- Very simple experimental design (Target/True/False)

Interim conclusion

- Very simple experimental design (Target/True/False)
- Direct mapping between theoretical parameters (availability of a reading) and experimental factors (differences between conditions / parameters in the statistical model)

Interim conclusion

- Very simple experimental design (Target/True/False)
- Direct mapping between theoretical parameters (availability of a reading) and experimental factors (differences between conditions / parameters in the statistical model)
- Relies on systematic entailment between the two readings of interest.

Breaking systematic entailment

- (1) A child rode every camel
- (2) Every child rode a camel

Breaking systematic entailment

- (1) A child rode every camel
 - a. Surface Scope: $a > \text{every}$
 - b. Inverse Scope: $\text{every} > a$

- (2) Every child rode a camel
 - a. Surface Scope: $\text{every} > a$
 - b. Inverse Scope: $a > \text{every}$

Breaking systematic entailment

- (1) A child rode every camel
- a. Surface Scope: $a > \text{every}$
 - b. Inverse Scope: $\text{every} > a$
- } $SS \rightarrow IS$
- (2) Every child rode a camel
- a. Surface Scope: $\text{every} > a$
 - b. Inverse Scope: $a > \text{every}$
- } $IS \rightarrow SS$

Breaking systematic entailment

- (1) A child rode every camel
- a. Surface Scope: $a > \text{every}$
 - b. Inverse Scope: $\text{every} > a$
- } $SS \rightarrow IS$
- (2) Every child rode a camel
- a. Surface Scope: $\text{every} > a$
 - b. Inverse Scope: $a > \text{every}$
- } $IS \rightarrow SS$

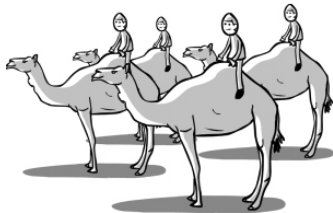
The entailment alternates from $A \rightarrow B$ to $B \rightarrow A$.

Experimental design

- (1) A child rode every camel
- (2) Every child rode a camel

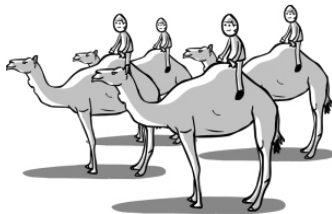
Experimental design

- (1) A child rode every camel
- (2) Every child rode a camel



Experimental design

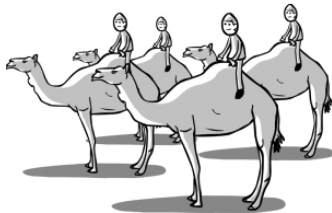
- (1) A child rode every camel
- (2) Every child rode a camel



- Impossible to systematically associate an answer to a reading

Experimental design

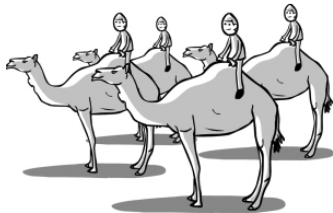
- (1) A child rode every camel
- (2) Every child rode a camel



- Impossible to systematically associate an answer to a reading
- We can however test each sentence and associate each answer to an interpretation:

Experimental design

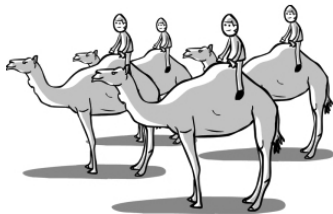
- (1) A child rode every camel
- (2) Every child rode a camel



- Impossible to systematically associate an answer to a reading
- We can however test each sentence and associate each answer to an interpretation:
 - T1: true under IS, false under SS

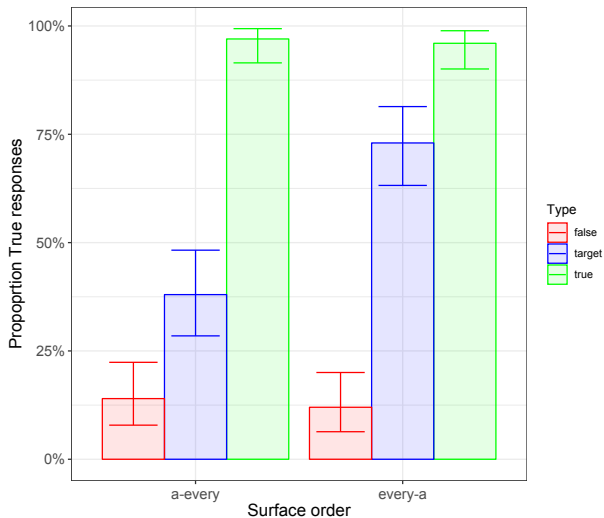
Experimental design

- (1) A child rode every camel
- (2) Every child rode a camel



- Impossible to systematically associate an answer to a reading
- We can however test each sentence and associate each answer to an interpretation:
 - T1: true under IS, false under SS
 - T2: true under SS, false under IS

Interpreting results



Direct comparison

```
Type: factor with levels: c('target', 'true', 'false')  
Order: sum coded ('a...every':-0.5, 'every...a':+0.5)  
Answer ~ Type * Order
```

Direct comparison

Type: factor with levels: c('target', 'true', 'false')

Order: sum coded ('a...every':-0.5, 'every...a':+0.5)

Answer ~ Type * Order

	β	<i>z</i> -value	<i>p</i> -value
(Intercept)	0.27	1.61	.107
Type:true	3.06	7.21	< .001***
Type:false	-2.49	-7.02	< .001***
Order	1.60	4.91	< .001***
[Type:true] × [Order]	-1.90	-2.25	.025*
[Type:false] × [Order]	-1.79	-3.26	.001**

What can we conclude from this model?

What can we conclude from this model?

- The “weak” reading ‘ $\forall > \exists$ ’ is more frequent with surface order ‘every... a’ (i.e. when it matches the surface order).

What can we conclude from this model?

- The “weak” reading ‘ $\forall > \exists$ ’ is more frequent with surface order ‘every... a’ (i.e. when it matches the surface order).
- This suggests that inverse scope is dispreferred.

What can we conclude from this model?

- The “weak” reading ‘ $\forall > \exists$ ’ is more frequent with surface order ‘every... a’ (i.e. when it matches the surface order).
- This suggests that inverse scope is dispreferred.
- But is inverse scope equally dispreferred in every configuration?

What can we conclude from this model?

- The “weak” reading ‘ $\forall > \exists$ ’ is more frequent with surface order ‘every... a’ (i.e. when it matches the surface order).
- This suggests that inverse scope is dispreferred.
- But is inverse scope equally dispreferred in every configuration?
- We could recode our dependent variable: instead of rates of True answers, look at rates of “Inverse scope” answers.

What can we conclude from this model?

- The “weak” reading ‘ $\forall > \exists$ ’ is more frequent with surface order ‘every... a’ (i.e. when it matches the surface order).
- This suggests that inverse scope is dispreferred.
- But is inverse scope equally dispreferred in every configuration?
- We could recode our dependent variable: instead of rates of True answers, look at rates of “Inverse scope” answers.

What can we conclude from this model?

- The “weak” reading ‘ $\forall > \exists$ ’ is more frequent with surface order ‘every... a’ (i.e. when it matches the surface order).
- This suggests that inverse scope is dispreferred.
- But is inverse scope equally dispreferred in every configuration?
- We could recode our dependent variable: instead of rates of True answers, look at rates of “Inverse scope” answers. But what does that mean for controls?

What does the Type factor really mean?

Let's forget about Order for a second and focus on 1+Type:

What does the Type factor really mean?

Let's forget about Order for a second and focus on 1+Type:
> contrasts(Type)

What does the Type factor really mean?

Let's forget about Order for a second and focus on 1+Type:

```
> contrasts(Type)
```

	true	false
target	0	0
true	1	0
false	0	1

What does the Type factor really mean?

Let's forget about Order for a second and focus on 1+Type:

	true	false	(inter.)
target	0	0	1
true	1	0	1
false	0	1	1

A bit of algebra. . .

Answer ~ 1+Type

	(inter.)	true	false
target	1	0	0
true	1	1	0
false	1	0	1

A bit of algebra...

Answer ~ 1+Type

		(inter.)	true	false
experimental conditions	{ target	1	0	0
	{ true	1	1	0
	{ false	1	0	1

A bit of algebra...

Answer ~ 1+Type

		model parameters		
		(inter.)	true	false
experimental conditions	{ target	1	0	0
	{ true	1	1	0
	{ false	1	0	1

A bit of algebra...

Answer $\sim 1 + \text{Type}$

		model parameters		
		(inter.)	true	false
experimental conditions	{ target	1	0	0
	{ true	1	1	0
	{ false	1	0	1
		β_0	β_1	β_2

A bit of algebra...

Answer ~ 1+Type

		model parameters		
		(inter.)	true	false
experimental conditions	target	1	0	0
	true	1	1	0
	false	1	0	1
		β_0	β_1	β_2

$$y_i = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$$

A bit of algebra...

Answer ~ 1+Type

		model parameters			
		(inter.)	true	false	
experimental conditions	{	target	1	0	0
		true	1	1	0
		false	1	0	1
			β_0	β_1	β_2

$$y_i = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$$

Predicted values for each experimental condition:

target: $\beta_0 \times 1 + \beta_1 \times 0 + \beta_2 \times 0$

A bit of algebra...

Answer ~ 1+Type

		model parameters			
		(inter.)	true	false	
experimental conditions	{	target	1	0	0
		true	1	1	0
		false	1	0	1
			β_0	β_1	β_2

$$y_i = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$$

Predicted values for each experimental condition:

target: β_0 (intercept)

A bit of algebra...

Answer ~ 1+Type

		model parameters			
		(inter.)	true	false	
experimental conditions	{	target	1	0	0
		true	1	1	0
		false	1	0	1
			β_0	β_1	β_2

$$y_i = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$$

Predicted values for each experimental condition:

target: β_0 (intercept)

true: $\beta_0 \times 1 + \beta_1 \times 1 + \beta_2 \times 0$

A bit of algebra...

Answer ~ 1+Type

		model parameters			
		(inter.)	true	false	
experimental conditions	{	target	1	0	0
		true	1	1	0
		false	1	0	1
			β_0	β_1	β_2

$$y_i = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$$

Predicted values for each experimental condition:

target: β_0 (intercept)
 true: $\beta_0 + \beta_1$ (intercept) + Type:true

A bit of algebra...

Answer ~ 1+Type

		model parameters			
		(inter.)	true	false	
experimental conditions	{	target	1	0	0
		true	1	1	0
		false	1	0	1
			β_0	β_1	β_2

$$y_i = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$$

Predicted values for each experimental condition:

- target: β_0 (intercept)
- true: $\beta_0 + \beta_1$ (intercept) + Type:true
- false: $\beta_0 \times 1 + \beta_1 \times 0 + \beta_2 \times 1$

A bit of algebra...

Answer ~ 1+Type

		model parameters			
		(inter.)	true	false	
experimental conditions	{	target	1	0	0
		true	1	1	0
		false	1	0	1
			β_0	β_1	β_2

$$y_i = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$$

Predicted values for each experimental condition:

target:	β_0	(intercept)
true:	$\beta_0 + \beta_1$	(intercept) + Type:true
false:	$\beta_0 + \beta_2$	(intercept) + Type:false

Alternative

		model parameters		
		target	true	false
experimental conditions	target	1	0	0
	true	0	1	0
	false	0	0	1
		β'_0	β'_1	β'_2

$$y_i = \beta'_0 x_{i0} + \beta'_1 x_{i1} + \beta'_2 x_{i2} + \varepsilon_i$$

Predicted values for each experimental condition:

target: β'_0

true: β'_1

false: β'_2

Alternative

		model parameters			
		target	true	false	
experimental conditions	{	target	1	0	0
		true	0	1	0
		false	0	0	1
			β'_0	β'_1	β'_2

$$y_i = \beta'_0 x_{i0} + \beta'_1 x_{i1} + \beta'_2 x_{i2} + \varepsilon_i$$

Predicted values for each experimental condition:

target:	β'_0	Type:target
true:	β'_1	Type:true
false:	β'_2	Type:false

Alternative

		model parameters		
		┌───────────────────┐		
		target	true	false
experimental conditions	{	target	1	0
		true	0	1
		false	0	0
			β'_0	β'_1
			β'_1	β'_2

$$y_i = \beta'_0 x_{i0} + \beta'_1 x_{i1} + \beta'_2 x_{i2} + \varepsilon_i$$

Predicted values for each experimental condition:

target:	$\beta'_0 = \beta_0$	Type:target
true:	$\beta'_1 = \beta_0 + \beta_1$	Type:true
false:	$\beta'_2 = \beta_0 + \beta_2$	Type:false

Alternative

Answer ~ 0+Type

		model parameters			
		target	true	false	
experimental conditions	{	target	1	0	0
		true	0	1	0
		false	0	0	1
			β'_0	β'_1	β'_2

$$y_i = \beta'_0 x_{i0} + \beta'_1 x_{i1} + \beta'_2 x_{i2} + \varepsilon_i$$

Predicted values for each experimental condition:

target:	$\beta'_0 = \beta_0$	Type:target
true:	$\beta'_1 = \beta_0 + \beta_1$	Type:true
false:	$\beta'_2 = \beta_0 + \beta_2$	Type:false

Goal for the week

Learn how to choose parameters for our statistical models that correspond to actual parameters in our theories.

Back to scope ambiguities

What we're interested in:

Back to scope ambiguities

What we're interested in: Rate of inverse scope readings

Back to scope ambiguities

What we're interested in: Rate of inverse scope readings

- The diagnosis for inverse scope differs in each condition.

Back to scope ambiguities

What we're interested in: Rate of inverse scope readings

- The diagnosis for inverse scope differs in each condition.
- ☞ Our IS parameter will depend on both `Type` and `Order`.

Back to scope ambiguities

What we're interested in: Rate of inverse scope readings

- The diagnosis for inverse scope differs in each condition.
- ☞ Our IS parameter will depend on both `Type` and `Order`.
- We need two more parameters to capture all 3 levels of `Type`.

Back to scope ambiguities

What we're interested in: Rate of inverse scope readings

- The diagnosis for inverse scope differs in each condition.
- ☞ Our IS parameter will depend on both `Type` and `Order`.
- We need two more parameters to capture all 3 levels of `Type`.
- We'll leave `Order` as it is for now.

	a-every	every-a
	β_{IS} β_1 β_2	β_{IS} β_1 β_2
target		
true		
false		

Your turn!

	a-every	every-a
	β_{IS} β_1 β_2	β_{IS} β_1 β_2
target	+1	-1
true		
false		

	a-every			every-a		
	β_{IS}	β_1	β_2	β_{IS}	β_1	β_2
target	+1			-1		
true	0			0		
false	0			0		

	a-every			every-a		
	β_{IS}	β_1	β_2	β_{IS}	β_1	β_2
target	+1			-1		
true	0	1		0	1	
false	0	0		0	0	

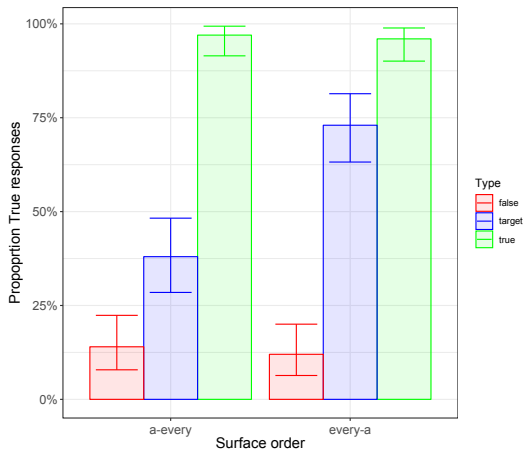
	a-every			every-a		
	β_{IS}	β_1	β_2	β_{IS}	β_1	β_2
target	+1			-1		
true	0	1	0	0	1	0
false	0	0	1	0	0	1

	a-every			every-a		
	β_{IS}	β_1	β_2	β_{IS}	β_1	β_2
target	+1	0	1	-1		
true	0	1	0	0	1	0
false	0	0	1	0	0	1

	a-every			every-a		
	β_{IS}	β_1	β_2	β_{IS}	β_1	β_2
target	+1	0	1	-1	1	0
true	0	1	0	0	1	0
false	0	0	1	0	0	1

	a-every			every-a		
	β_{IS}	β_T	β_F	β_{IS}	β_T	β_F
target	+1	0	1	-1	1	0
true	0	1	0	0	1	0
false	0	0	1	0	0	1

Visualizing model parameters



β_F

$\beta_F + \beta_{IS}$

β_T

β_F

$\beta_T - \beta_{IS}$

β_T

Reanalysing the data

We define our new factors in R:

Reanalysing the data

We define our new factors in R:

```
IS <- case_when(  
  Order=="a-every"&Type=="target" ~ 1,  
  Order=="every-a"&Type=="target" ~ -1,  
  T ~ 0  
)  
True <- case_when(  
  Type=="true" ~ 1,  
  Order=="every-a"&Type=="target" ~ 1,  
  T ~ 0  
)  
False <- case_when(  
  Type=="false" ~ 1,  
  Order=="a-every"&Type=="target" ~ 1,  
  T ~ 0  
)
```

Reanalysing the data

Model:

```
Answer~(0+True+False+Inverse)+(0+True+False+Inverse):Order
```

Reanalysing the data

Model:

```
Answer~(0+True+False+Inverse)+(0+True+False+Inverse):Order
```

Results:

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z)	
True	3.3575	0.3905	8.598	< 2e-16	***
False	-2.1487	0.2917	-7.367	1.74e-13	***
IS	1.8409	0.3415	5.390	7.03e-08	***
True:Order	-0.2984	0.7780	-0.384	0.701	
False:Order	-0.1894	0.4360	-0.434	0.664	
IS:Order	0.7148	0.6794	1.052	0.293	

Conclusion

- By re-parametrizing the statistical model we were able to get interpretable results.

Conclusion

- By re-parametrizing the statistical model we were able to get interpretable results.
- In this example, this was necessary to show that the rate of inverse scope readings doesn't depend on surface order.

Conclusion

- By re-parametrizing the statistical model we were able to get interpretable results.
- In this example, this was necessary to show that the rate of inverse scope readings doesn't depend on surface order.
- Tomorrow: introduce a bit of formalism, before moving to concrete examples on Wednesday.

Questions?