Model design in R: Class 1

Alexandre Cremers

August 12th, 2019

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

- Class 1: introducing the problem and the solution
- Class 2: formalize what we'll do today
- Classes 3/4: Applications
- Class 5: A few more problems and a few more solutions. . .

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.
- Two candidate meanings $\rightsquigarrow$ 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 which readings are available in which contexts, 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: The truth-value judgment task.

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

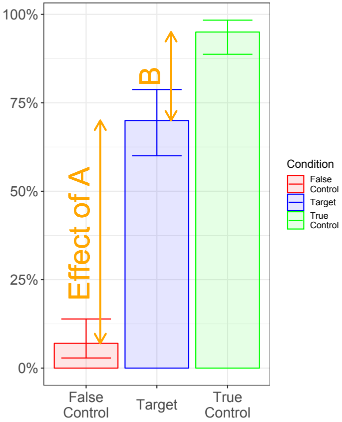
Bott&Noveck 2004

Predictions in Experimental SemPrag

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

- 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 less than an unambiguous true control

☞ We can simply compare our target to true and false controls.



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)
- 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
 a. Surface Scope: a > every
 b. Inverse Scope: every > a

} $SS \rightarrow IS$

(2) Every child rode a camel
 a. Surface Scope: every > a
 b. Inverse Scope: a > every

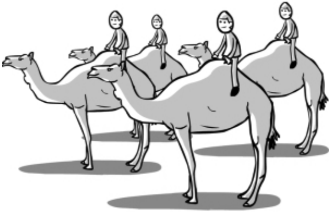
} $IS \rightarrow SS$

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

Experimental design

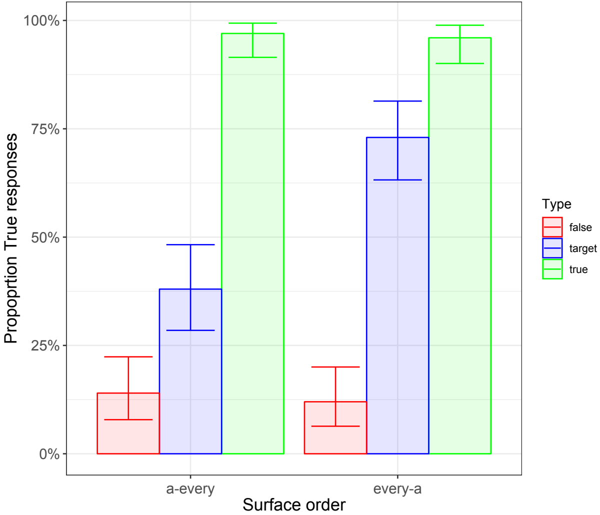
(1) A child rode every camel

(2) Every child rode a camel



- 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`

	β	z-value	p-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?

- 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`:

```
> contrasts(Type)
         true  false  (inter.)
 target     0      0        1
 true       1      0        1
 false      0      1        1
```

A bit of algebra. . .

`Answer ~ 1+Type`

		model predictors		
		(inter.)	true	false
experimental conditions	target	1	0	0
	true	1	1	0
	false	1	0	1
model parameters:		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:	$\beta_0 \times 1 + \beta_1 \times 0 + \beta_2 \times 0$	$= \beta_0$	(intercept)
true:	$\beta_0 \times 1 + \beta_1 \times 1 + \beta_2 \times 0$	$= \beta_0 + \beta_1$	(intercept)+Type:true
false:	$\beta_0 \times 1 + \beta_1 \times 0 + \beta_2 \times 1$	$= \beta_0 + \beta_2$	(intercept)+Type:false

Alternative parametrization (not particularly useful)

```
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: Rate of inverse scope readings

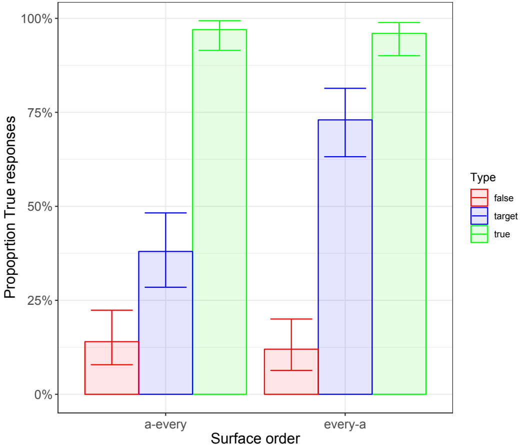
- The diagnosis for inverse scope differs in each condition.
- ☞ Our IS parameter will depend on both `Type` and `Order`.
- We need two more parameters to capture all 3 levels of `Type`.
- We'll leave `Order` as it is for now.

	a-every			every-a		
	β_{IS}	β_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

One possibility:

- T and F represent baseline for controls. The a-every target is indistinguishable from false control without IS. The every-a target is indistinguishable from true control without IS.
- NB: this model comes with its own assumptions (can you list them?) and may not always be the best.

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:

```
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
```

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.
- In this example, this was necessary to show that the rate of inverse scope readings doesn't depend on surface order.
- Next: introduce a bit of formalism, before moving to concrete examples in Class 3.

Questions?