Analysis of covariance

Session 9

MATH 80667A: Experimental Design and Statistical Methods
for Quantitative Research in Management
HEC Montréal

Outline

Analysis of covariance

Model assumptions

Analysis of covariance

What's in a model?

In experimental designs, the explanatories are

- experimental factors (categorical)
- continuous (dose-response)

Random assignment implies no systematic difference between groups.

ANCOVA = Analysis of covariance

- Analysis of variance with added continuous covariate(s) to reduce experimental error (similar to blocking).
- These continuous covariates are typically concomitant variables (measured alongside response).
- Including them in the mean response (as slopes) can help reduce the experimental error (residual error).

Control to gain power!

Identify external sources of variations

- enhance balance of design (randomization)
- reduce mean squared error of residuals to increase power

These steps should in principle increase power **if** the variables used as control are correlated with the response.

- Beware the kitchen sink approach
- Continuous variables are not used for assignment to treatment

Example

Abstract of van Stekelenburg et al. (2021)

> In three experiments with more than 1,500 U.S. adults who held false beliefs, participants first learned the value of scientific consensus and how to identify it. Subsequently, they read a news article with information about a scientific consensus opposing their beliefs. We found strong evidence that in the domain of genetically engineered food, this two-step communication strategy was more successful in correcting misperceptions than merely communicating scientific consensus.

Experiment 2: Genetically Engineered Food

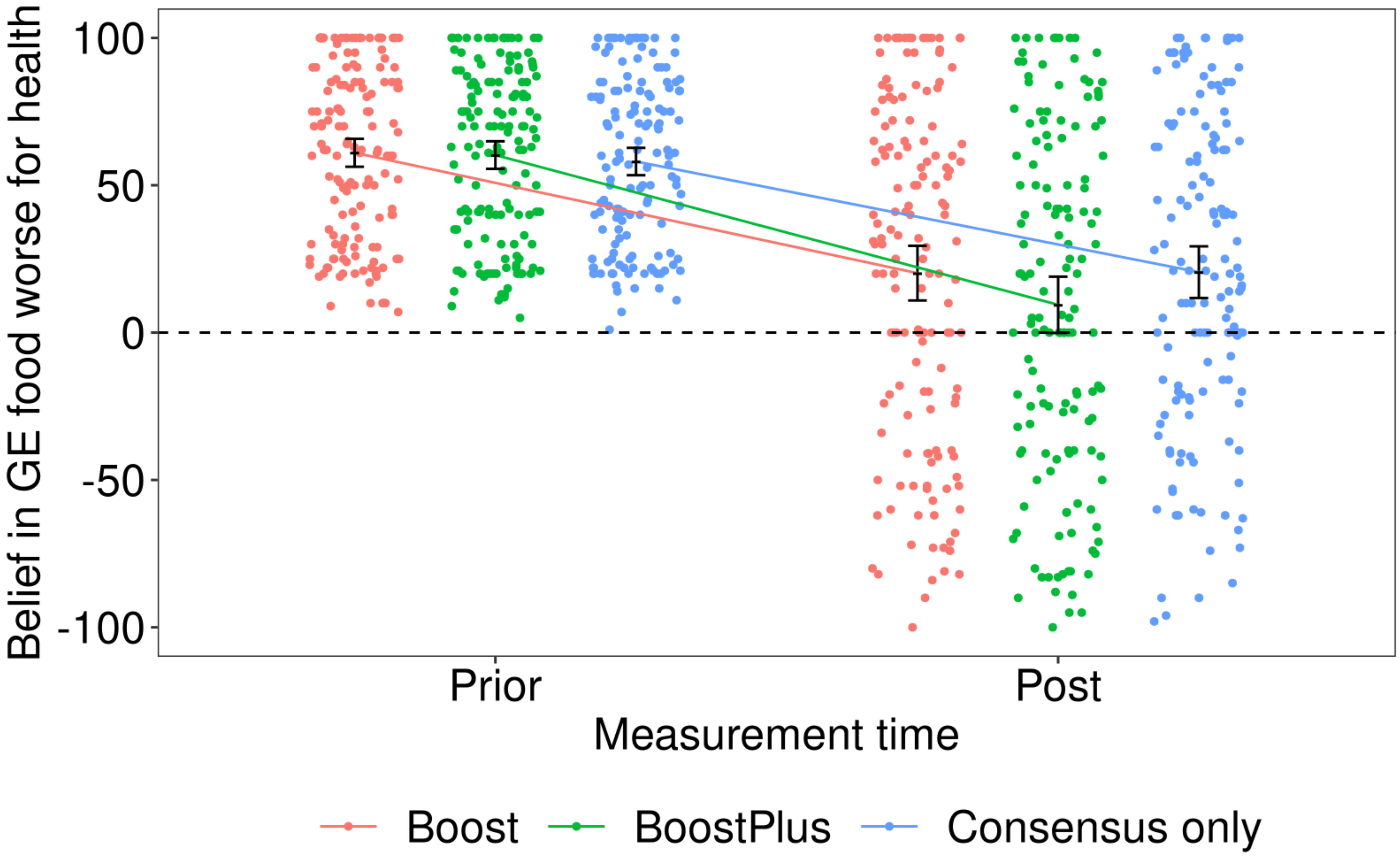
We focus on a single experiment; preregistered exclusion criteria led to $n = 442$ total sample size (unbalanced design).

Three experimental conditions:

`Boost` | `Boost Plus` | Consensus only (`consensus`)

Use `post` as response variable and `prior` beliefs as a control variable in the analysis of covariance.

> their response was measured on a visual analogue scale ranging from –100 (I am 100% certain this is false) to 100 (I am 100% certain this is true) with 0 (I don't know) in the middle.


Belief in GE food worse for health
100
50
0
-50
-100
Prior
Post
Measurement time
Boost
BoostPlus
Consensus only

Model formulation

Average for the rth replication of the ith experimental group is

$$\begin{aligned}\mathsf{E}(\texttt{post}_{ir}) &= \mu + \alpha_i \texttt{condition}_i + \beta \texttt{prior}_{ir}.\\ \mathsf{Va}(\texttt{post}_{ir}) &= \sigma^2\end{aligned}$$

We assume that there is no interaction between `condition` and `prior`

- the slopes for `prior` are the same for each `condition` group.
- the effect of prior is linear

Contrasts of interest

1. Difference between average boosts (`Boost` and `BoostPlus`) and control (`consensus`)
2. Difference between `Boost` and `BoostPlus` (pairwise)

Inclusion of the `prior` score leads to increased precision for the mean (reduces variability).

- The estimated marginal means will be based on detrended values $\neq$ group averages
- In the `emmeans` package, the average of the covariate is used as value.
- the difference between levels of `condition` are the same for any value of `prior` (parallel lines), but the uncertainty changes.

Data analysis

Loading data | Scatterplot | Model | ANOVA

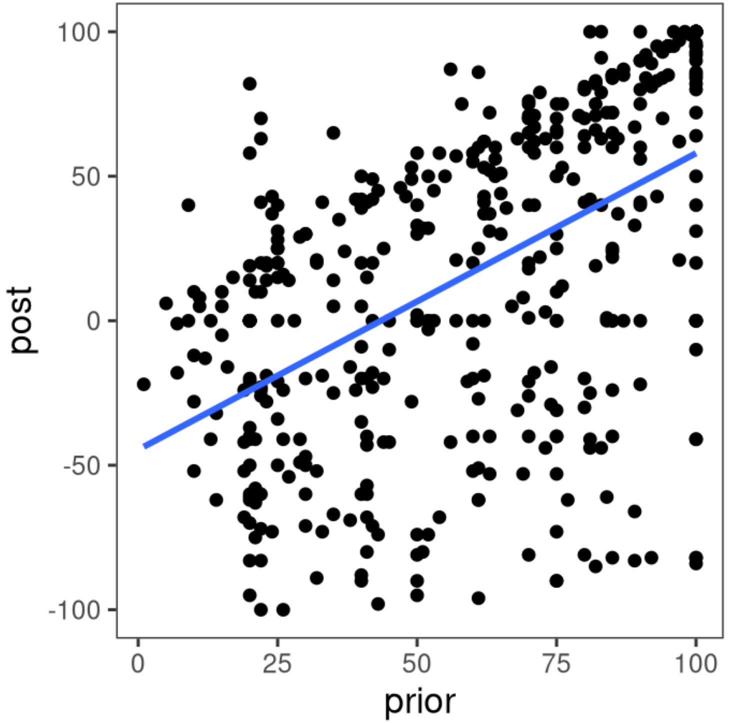
```r
library(tidyverse)
library(emmeans)
options(contrasts = c("contr.sum", "contr.poly"))
url <- "https://edsm.rbind.io/data/vanStekelenburg2021S2.csv"
exp2 <- read.csv(url, header = TRUE) %>%
  mutate(condition = factor(condition))
# Check balance
with(exp2, table(condition))
```

Data analysis

Loading data | **Scatterplot** | Model | ANOVA

```
ggplot(data = exp2,
       aes(x = prior, y = post)) +
       geom_point() +
       geom_smooth(method = "lm",
                   se = FALSE)
```

Strong correlation; note responses that achieve max of scale.



Data analysis

Loading data | Scatterplot | **Model** | ANOVA

```r
# Check that the data are well randomized
car::Anova(lm(prior ~ condition, data = exp2), type = 3)
# Fit linear model with continuous covariate
model1 <- lm(post ~ condition + prior, data = exp2)
# Fit model without for comparison
model2 <- lm(post ~ condition, data = exp2)
# Global test for differences - of NO INTEREST
car::Anova(model1, type = 3)
car::Anova(model2, type = 3)
```

Data analysis

Loading data | Scatterplot | Model | ANOVA

term	sum of squares	df	statistic	p-value
(Intercept)	166341	1	71.7	0.00
condition	**14107**	**2**	**3.0**	**0.05**
prior	385385	1	166.1	0.00
Residuals	1016461	438		

term	sum of squares	df	statistic	p-value
(Intercept)	123377	1	53.16	0.000
condition	**11680**	**2**	**2.52**	**0.082**
Residuals	1016461	438		

Data analysis

Contrasts | t-tests | Assumptions

```r
emm1 <- emmeans(model1, specs = "condition")
# Note order: Boost, BoostPlus, consensus
emm2 <- emmeans(model2, specs = "condition")
# Not comparable: since one is detrended and the other isn't
contrast_list <- list(
   "boost vs control" = c(0.5,  0.5, -1),
   #av. boosts vs consensus
   "Boost vs BoostPlus" = c(1, -1,  0))
contrast(emm1,
         method = contrast_list,
         p.adjust = "holm")
```

Data analysis

Contrasts | t-tests | Assumptions

contrast	estimate	se	df	t stat	p-value
boost vs control	-8.37	4.88	438	-1.72	0.09
Boost vs BoostPlus	9.95	5.60	438	1.78	0.08

Contrasts with ANCOVA with `prior` (Holm-Bonferroni adjustment with $k=2$ tests)

contrast	estimate	se	df	t stat	p-value
boost vs control	-5.71	5.71	439	-1.00	0.32
Boost vs BoostPlus	10.74	6.57	439	1.63	0.10

Contrasts for ANOVA (Holm-Bonferroni adjustment with $k=2$ tests)

Data analysis

Contrasts | t-tests | **Assumptions**

```r
# Test equality of variance
levene <- car::leveneTest(
   resid(model1) ~ condition,
   data = exp2,
   center = 'mean')
# Equality of slopes (interaction)
car::Anova(lm(post ~ condition * prior,
           data = exp2),
           model1, type = 3)
```

Levene's test of equality of variance: $F\,(2, 439) = 2.05$ with a p-value of 0.13.

term	sum of squares	df	statistic	p-value
(Intercept)	165573	1	71.3	0.0
condition	4245	2	0.9	0.4
prior	382596	1	164.9	0.0
condition:prior	**3257**	**2**	**0.7**	**0.5**
Residuals	1016461	438		

Fit the model with an interaction `condition*prior`. A significative interaction

The kitchen sink approach

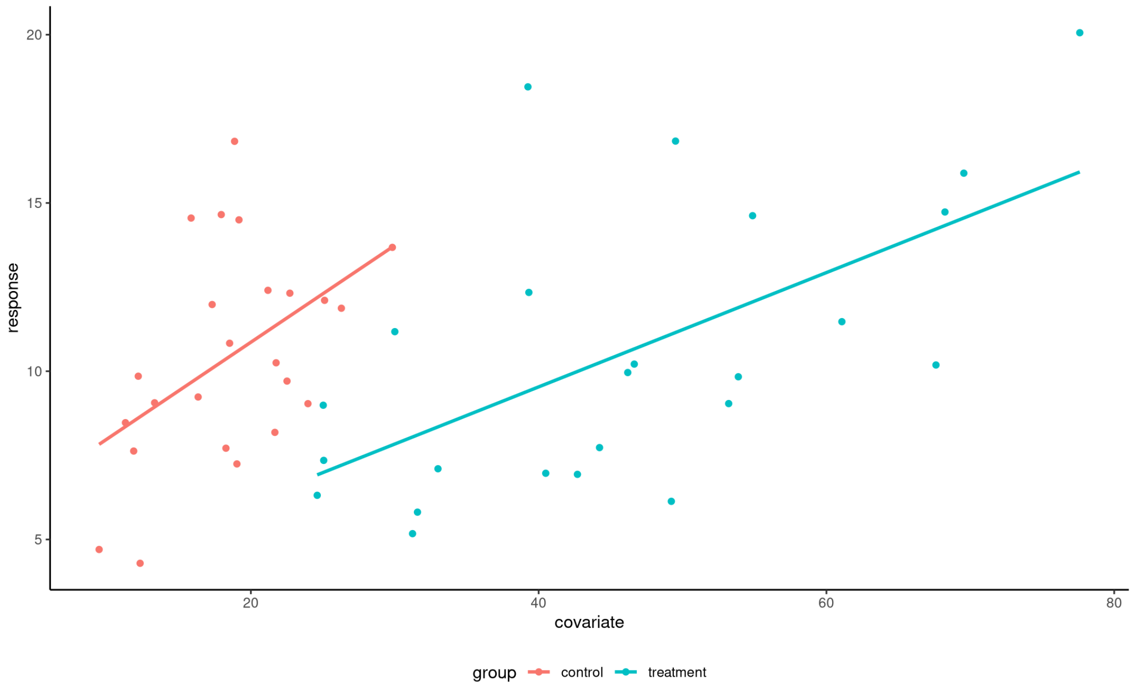
Should we control for more stuff?

NO! ANCOVA is a design to reduce error

- Randomization should ensure that there is no confounding
- No difference (on average) between group given a value of the covariate.
- If it isn't the case, adjustment matters

Equal trends

- If trends are different, meaningful comparisons (?)
- Differences between groups depend on value of the covariate



Due to lack of overlap, comparisons hazardous as they entail extrapolation one way or another.

Model assumptions

Are my conclusions valid?

Assumptions give us leverage for designing tests and models

Linear model assumptions

linearity **equal variance** **normality** **independence**