Models as prediction machines: How to convert confusing coefficients with marginaleffects

Julia M. Rohrer
juliarohrer.com
Leipzig University
🦋@dingdingpeng.the100.ci
www.the100.ci

Slides: Resources

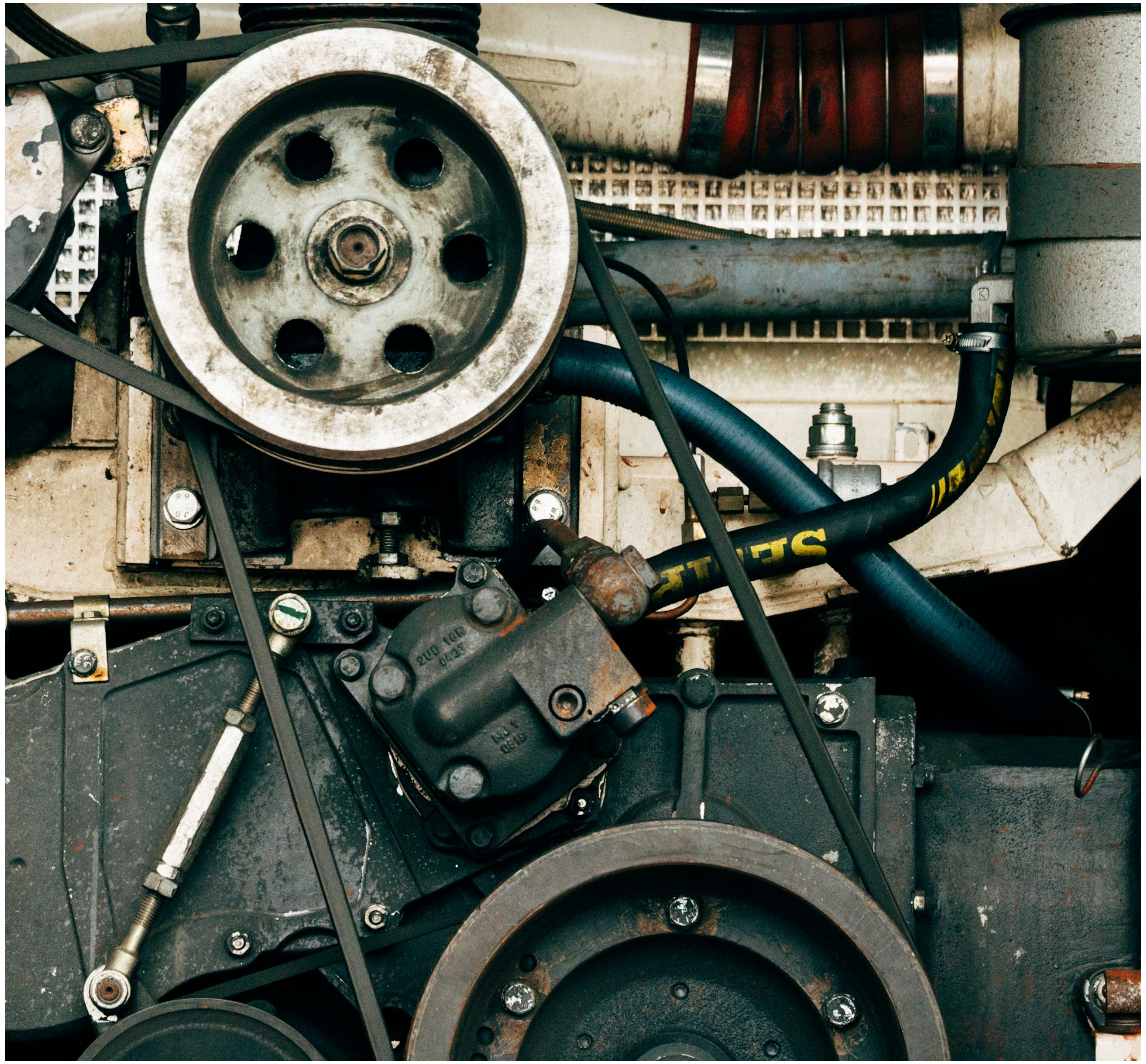


Foto von Elimende Inagella auf Unsplash

Standard statistics training

1. Take some model

```
lm()
glm()
lmer()
gam()
brm()
loess()
clm()
…
```

2. Make sure you code everything that goes into the model the right way

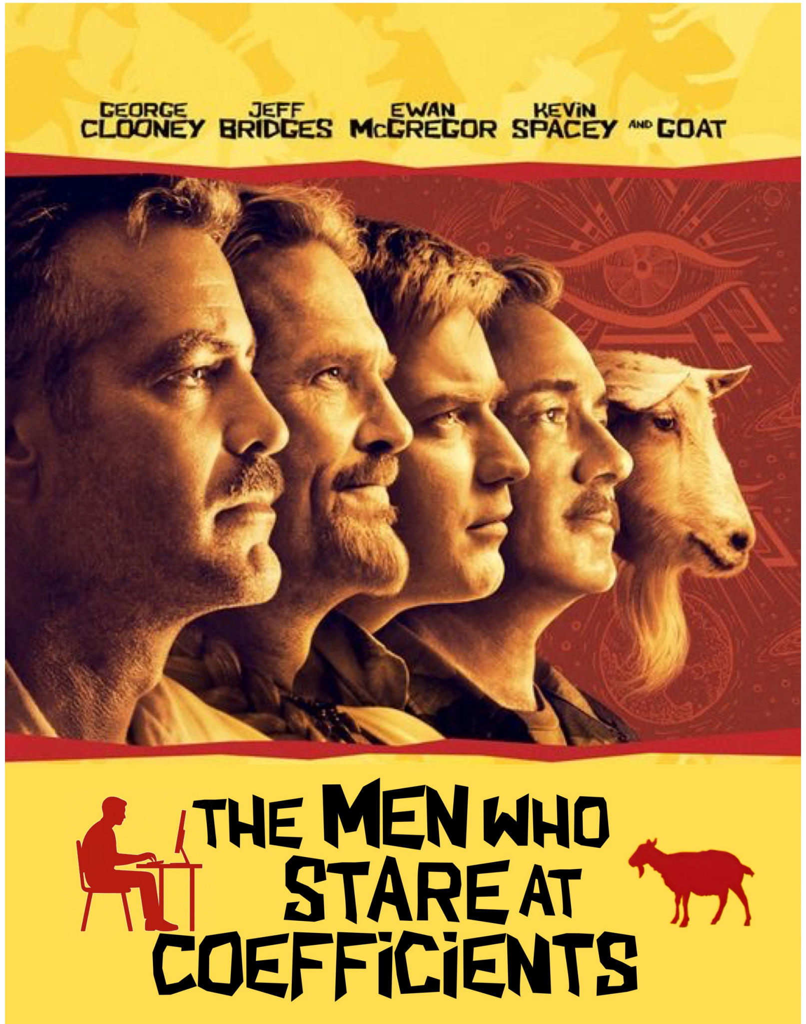
```
# assigning the deviation contrasts to race.f
contrasts(df$race.f) = contr.sum(4)

# center variables to make interaction interpretable
df$x1_cent <- scale(df$x1, center = TRUE, scale = FALSE)
df$x2_cent <- scale(df$x2, center = TRUE, scale = FALSE)
```

Standard statistics training

3. Fit model

```
estimator go brrrrrrr
```

4. ???

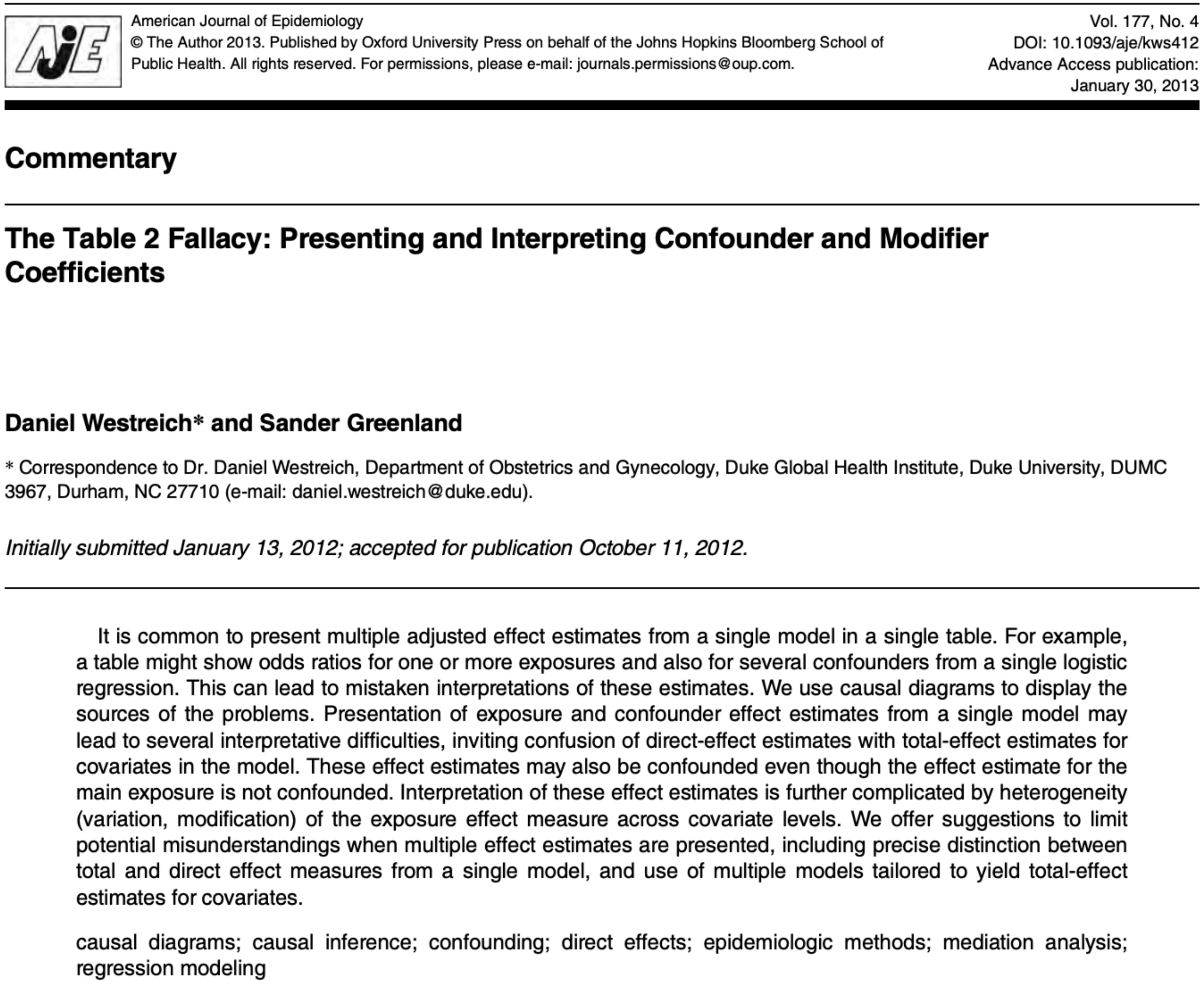



```
Call:
lm(formula = lebensz_org ~ sex * bs(alter, df = 3) * bs(einkommenj1,
    df = 3), data = soep[soep$alter < 60, ])

Residuals:
    Min      1Q  Median      3Q     Max
-7.8030 -0.6359  0.3642  1.0181  3.2115

Coefficients:
                                           Estimate Std. Error t value Pr(>|t|)
(Intercept)                                  7.8621     0.1032  76.148  < 2e-16 ***
sex                                         -0.3415     0.1380  -2.475  0.01332 *
bs(alter, df = 3)1                          -1.2071     0.3793  -3.183  0.00146 **
bs(alter, df = 3)2                          -1.1611     0.2675  -4.341 1.43e-05 ***
bs(alter, df = 3)3                          -0.9721     0.1881  -5.169 2.39e-07 ***
bs(einkommenj1, df = 3)1                    -2.9110     1.9740  -1.475  0.14031
bs(einkommenj1, df = 3)2                    19.8332    19.4858   1.018  0.30877
bs(einkommenj1, df = 3)3                    -0.8611   132.6094  -0.006  0.99482
sex:bs(alter, df = 3)1                       1.2525     0.4806   2.606  0.00917 **
sex:bs(alter, df = 3)2                       0.5992     0.3352   1.788  0.07380 .
sex:bs(alter, df = 3)3                       0.4095     0.2492   1.643  0.10035
sex:bs(einkommenj1, df = 3)1                -2.6841     3.2689  -0.821  0.41160
sex:bs(einkommenj1, df = 3)2                65.7747    40.8548   1.610  0.10743
sex:bs(einkommenj1, df = 3)3              -457.9446   394.2805  -1.161  0.24547
bs(alter, df = 3)1:bs(einkommenj1, df = 3)1  8.7901     3.9310   2.236  0.02536 *
bs(alter, df = 3)2:bs(einkommenj1, df = 3)1  3.1041     2.3496   1.321  0.18649
bs(alter, df = 3)3:bs(einkommenj1, df = 3)1  3.9785     2.2388   1.777  0.07558 .
bs(alter, df = 3)1:bs(einkommenj1, df = 3)2 -27.3680    28.4171  -0.963  0.33552
bs(alter, df = 3)2:bs(einkommenj1, df = 3)2 -12.4798    18.6404  -0.670  0.50319
bs(alter, df = 3)3:bs(einkommenj1, df = 3)2 -17.4873    20.0521  -0.872  0.38317
bs(alter, df = 3)1:bs(einkommenj1, df = 3)3 -21.9448   181.2849  -0.121  0.90365
bs(alter, df = 3)2:bs(einkommenj1, df = 3)3   8.6725   121.1156   0.072  0.94292
```

5. Profit

Some pitfalls of this approach

- it gets confusing quickly
  - interactions, nonlinearities, categorical predictors and their combinations
- it often needs to be relearned for different model classes and different model implementations
  - Ordinal model? Tough luck, now the coefficients mean something else
- sometimes it's not clear which coefficients answer your research question
- sometimes it's not clear which research question your coefficients answer


American Journal of Epidemiology
© The Author 2013. Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.
Vol. 177, No. 4
DOI: 10.1093/aje/kws412
Advance Access publication:
January 30, 2013

Commentary

The Table 2 Fallacy: Presenting and Interpreting Confounder and Modifier Coefficients

Daniel Westreich* and Sander Greenland

* Correspondence to Dr. Daniel Westreich, Department of Obstetrics and Gynecology, Duke Global Health Institute, Duke University, DUMC 3967, Durham, NC 27710 (e-mail: daniel.westreich@duke.edu).

Initially submitted January 13, 2012; accepted for publication October 11, 2012.

It is common to present multiple adjusted effect estimates from a single model in a single table. For example, a table might show odds ratios for one or more exposures and also for several confounders from a single logistic regression. This can lead to mistaken interpretations of these estimates. We use causal diagrams to display the sources of the problems. Presentation of exposure and confounder effect estimates from a single model may lead to several interpretative difficulties, inviting confusion of direct-effect estimates with total-effect estimates for covariates in the model. These effect estimates may also be confounded even though the effect estimate for the main exposure is not confounded. Interpretation of these effect estimates is further complicated by heterogeneity (variation, modification) of the exposure effect measure across covariate levels. We offer suggestions to limit potential misunderstandings when multiple effect estimates are presented, including precise distinction between total and direct effect measures from a single model, and use of multiple models tailored to yield total-effect estimates for covariates.

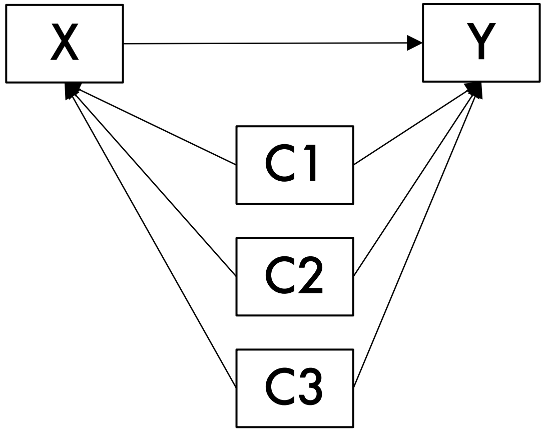
causal diagrams; causal inference; confounding; direct effects; epidemiologic methods; mediation analysis; regression modeling

More on the logic of third variable control:
Rohrer (2018), Wysocki et al. (2022)

Let‘s say you‘re interested in X → Y

You fit a model:
Y~ X + C1 + C2 + C3

This implies that you have the following causal model in your head:



If this model is correct, you should *not* interpret your model’s coefficients of C1, C2 and C3 substantively because the models controls for X – a *mediator* of their effects, which leads to

- overcontrol bias (removes part of the effect of interest) and
- potential collider bias (introduces new spurious associations)

Some pitfalls of this approach

- it gets confusing quickly
  - interactions, nonlinearities, categorical predictors and their combinations
- it often needs to be relearned for different model classes and different model implementations
  - Ordinal model? Tough luck, now the coefficients mean something else
- sometimes it's not clear which coefficients answer your research question
- sometimes it's not clear which research question your coefficients answer

But what's the alternative???

Models as prediction machines with `marginaleffects`

Fitted model

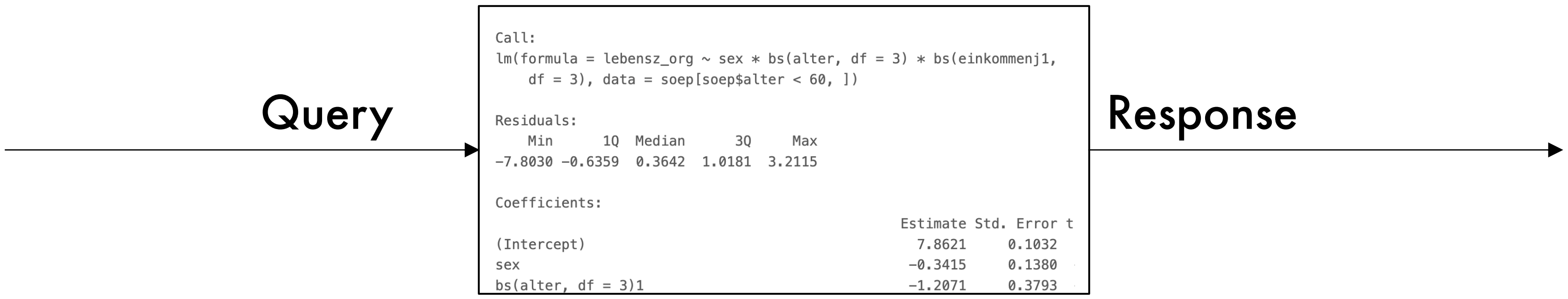
Feed in combinations of predictor values

Get out predicted outcomes

1. Decision: Which combinations to feed in to answer research question

2. Decision: How to combine the predicted outcomes to answer research question

Model predicting life satisfaction from gender, age and income



Query: What's the predicted life satisfaction for a woman who is 35 years old and earns 20,000 Euro?

```
predictions(mod, newdata = data.frame(sex = 1, alter = 35, einkommenj1 = 20000))
```

Response:

Estimate	Std. Error	z	Pr(>|z|)	S	2.5 %	97.5 %	sex	alter	einkommenj1
7.65	0.0489	156	<0.001	Inf	7.55	7.74	1	35	20000

Estimate	Std. Error	z	Pr(>|z|)	S	2.5 %	97.5 %	sex	alter	einkommenj1
7.65	0.0489	156	<0.001	Inf	7.55	7.74	1	35	20000

Query: What's the predicted life satisfaction for a woman who is 35 years old and earns twice as much, 40,000 Euro?

```
predictions(mod, newdata = data.frame(sex = 1, alter = 35, einkommenj1 = 40000))
```

Response:

Estimate	Std. Error	z	Pr(>|z|)	S	2.5 %	97.5 %	sex	alter	einkommenj1
7.77	0.0683	114	<0.001	Inf	7.64	7.91	1	35	40000

Query: What's the predicted difference in life satisfaction between that woman if she earns 20,000 Euro versus 40,000 Euro, *all else being equal?*

```
comparisons(mod,
            newdata = data.frame(sex = 1, alter = 35, einkommenj1 = 20000),
            variables = list("einkommenj1" = c(20000, 40000)))
```

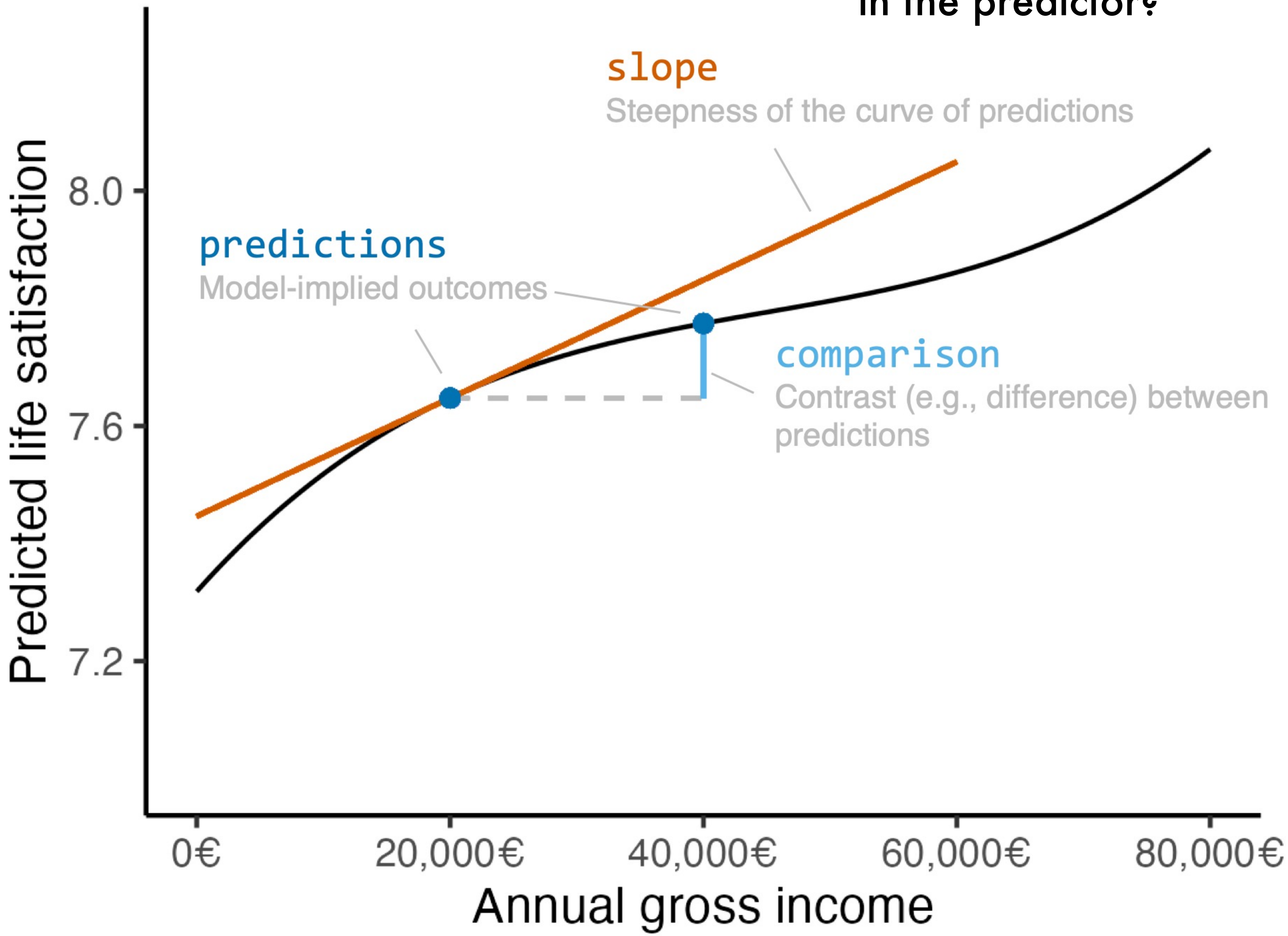
Response:

Estimate	Std. Error	z	Pr(>|z|)	S	2.5 %	97.5 %	sex	alter	einkommenj1
0.127	0.07	1.81	0.0697	3.8	-0.0102	0.264	1	35	20000

Counterfactual comparison: What if this one thing changed, but all else remained equal?

For a 35-year-old woman

Slope: partial derivative – what does the model imply for a tiny tiny change in the predictor?



Target quantities

- Essential building blocks:
  - `predictions()`
  - `comparisons()` – contrast various predictions
  - `slopes()` – partial derivatives of predictions
- We can generate these target quantities for arbitrary observations
  - any observation from the data
  - any made-up observation that may be of interest
    - for example: what does the model predict for a person who is perfectly average? ("marginal effect at the mean")

Target quantities

» Often, we are interested in target quantities *averaged over observations*

  » average for everybody in the data

    » `avg_predictions(), avg_comparisons(), avg_slopes()`

  » average for subgroups of the data

    » `avg_*(…, by = subgroupvariable)`

  » some target population that requires reweighting of the data

    » `avg_*(…, wts = weightingvariable)`

```
comparisons(mod, variables = "sex")
```

```
 Estimate Std. Error      z Pr(>|z|)    S   2.5 % 97.5 %
  -0.1330     0.1390 -0.957    0.338  1.6 -0.4054  0.139
   0.3483     0.0643  5.413   <0.001 23.9  0.2222  0.474
  -0.0589     0.1007 -0.584    0.559  0.8 -0.2562  0.139
  -0.1189     0.1136 -1.046    0.295  1.8 -0.3415  0.104
   0.0952     0.0640  1.489    0.137  2.9 -0.0301  0.221
--- 16282 rows omitted. See ?print.marginaleffects ---
   0.2512     0.0595  4.224   <0.001 15.3  0.1346  0.368
  -0.1049     0.0831 -1.262    0.207  2.3 -0.2677  0.058
  -0.0134     0.0627 -0.214    0.831  0.3 -0.1363  0.109
   0.0876     0.0657  1.334    0.182  2.5 -0.0411  0.216
   0.1094     0.0797  1.372    0.170  2.6 -0.0468  0.266
Term: sex
Comparison: 1 - 0
```

1. For everybody in the data, set sex to 1 and predict outcome
2. For everybody in the data, set sex to 0 and predict outcome
3. Take difference between the predictions → individual-level counterfactual comparison, "all else being equal"

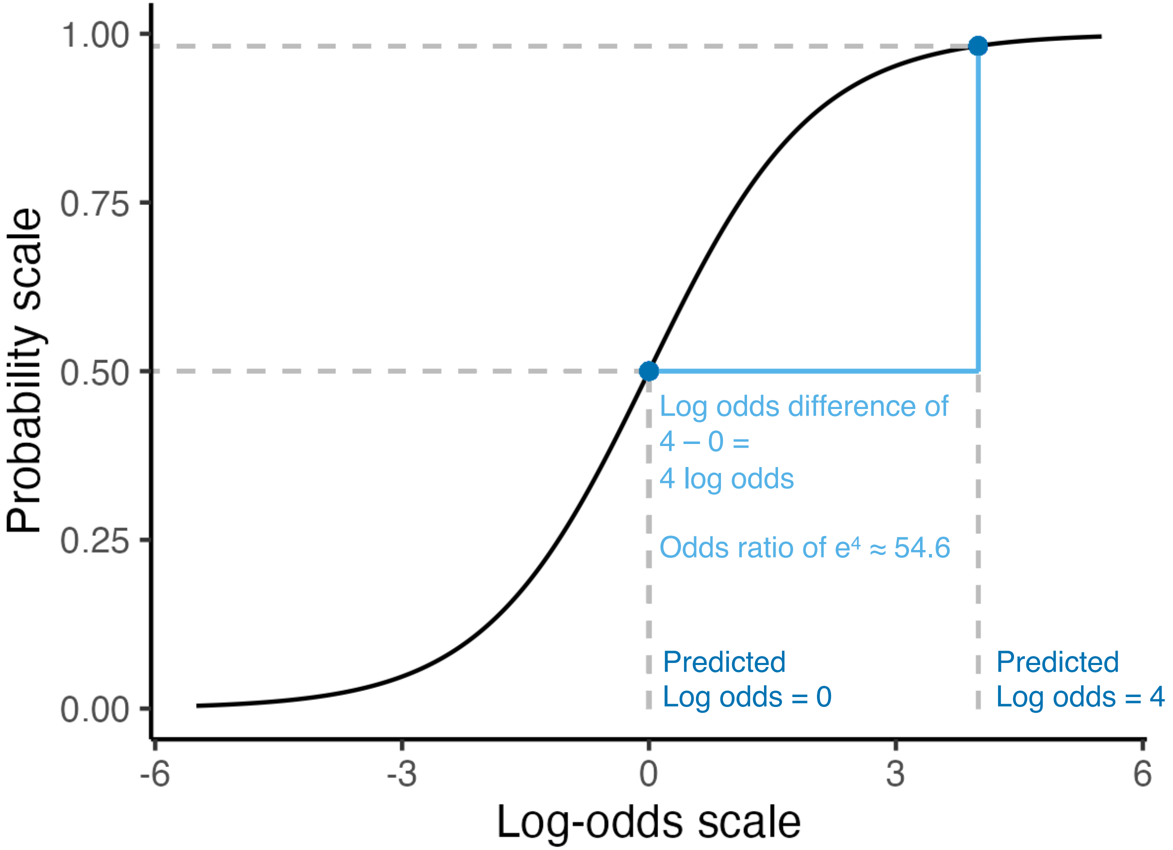
```
avg_comparisons(mod, variables = "sex")
```

```
 Estimate Std. Error    z Pr(>|z|)    S 2.5 % 97.5 %
    0.158     0.0287 5.51   <0.001 24.7 0.102  0.214

Term: sex
Comparison: 1 - 0
```

4. Average the individual-level counterfactual comparisons

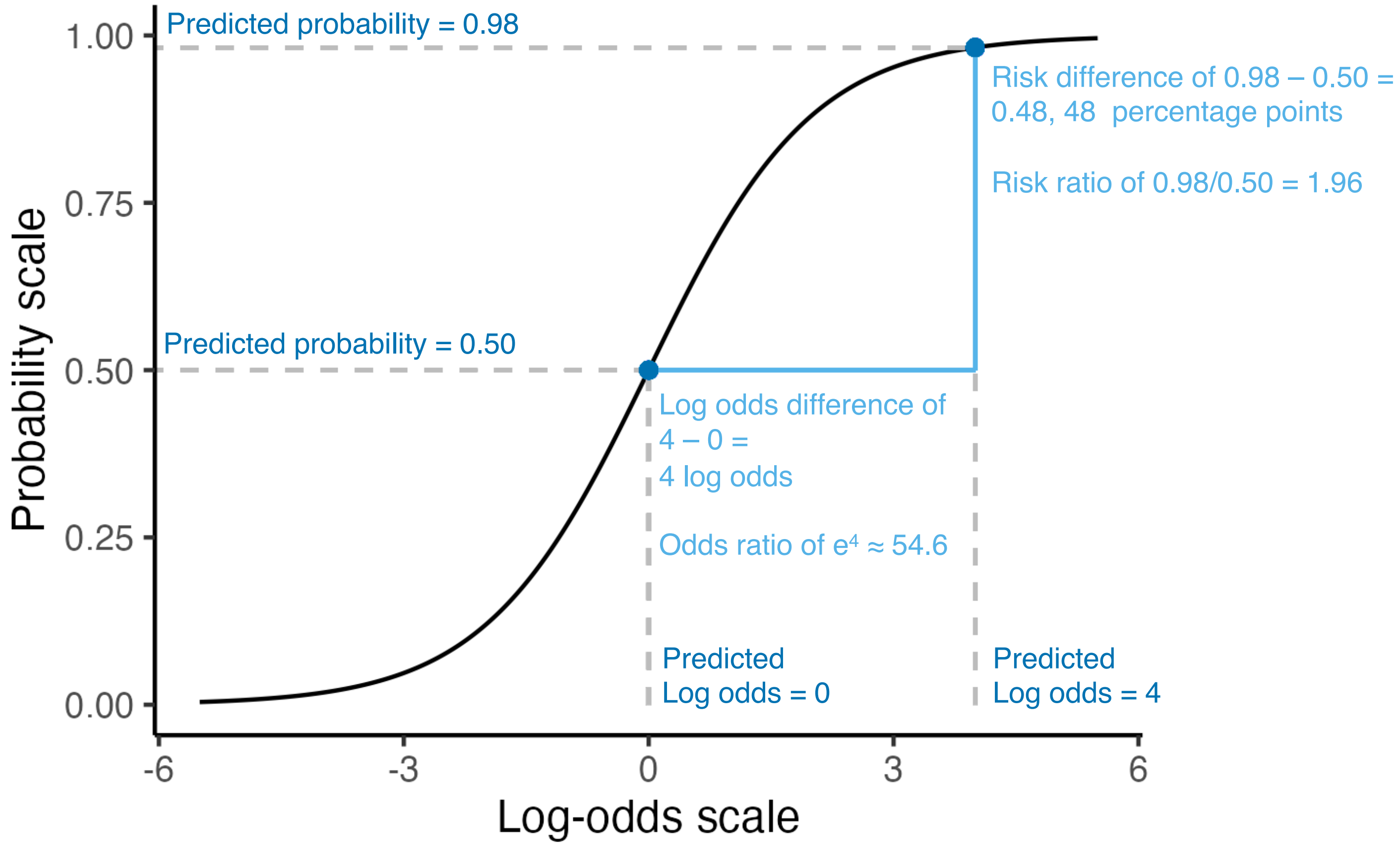
Target quantities

» target quantities can be
  » calculated on different scales (depending on the model type) and
  » contrasted in different manners



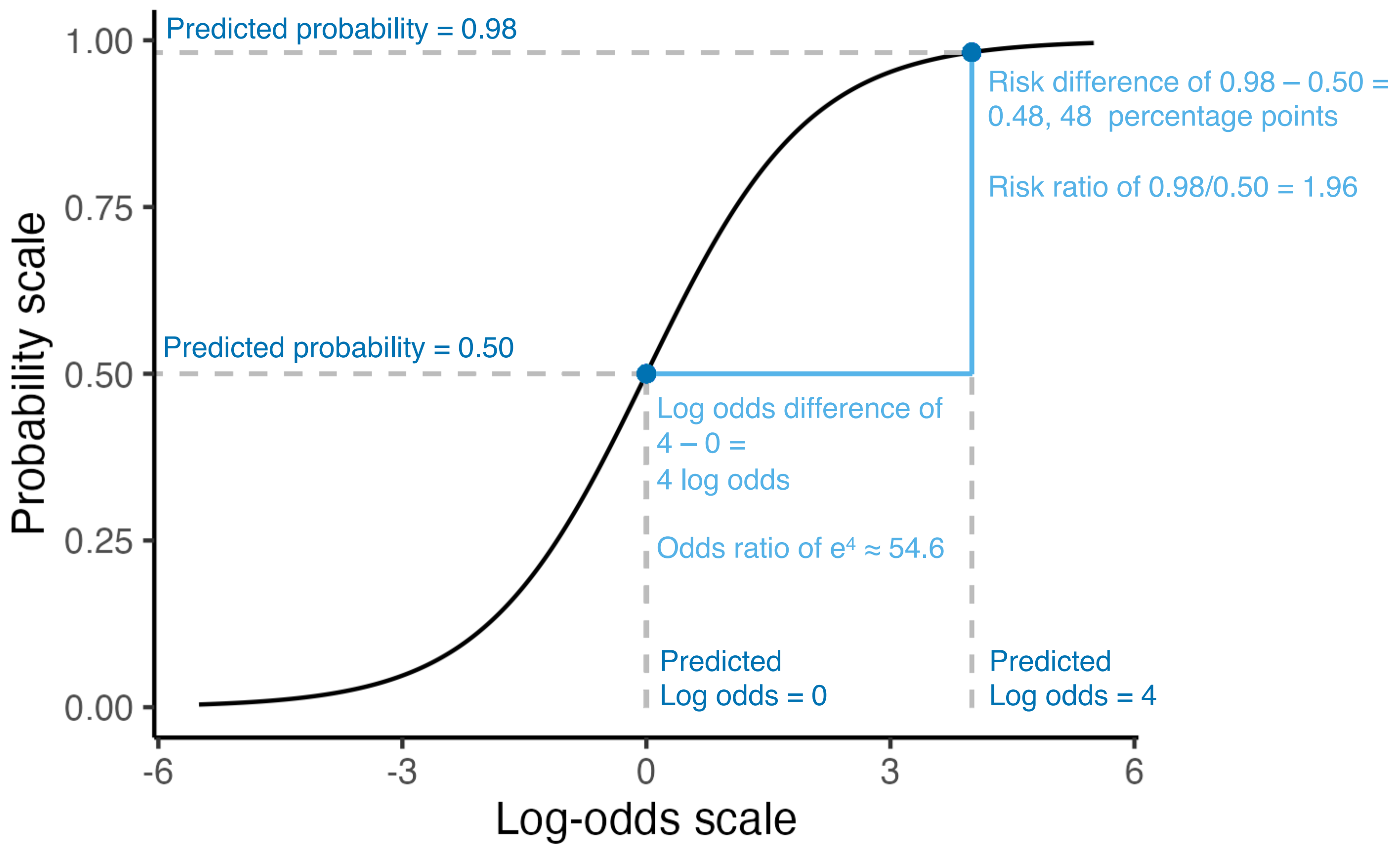
`predictions(…,type = “link”)`

```
predictions(…,
type = “response”)
```




```
predictions(…,type = “link”)
```

predictions(…,
type = “response”)



predictions(…,type = “link”)

Target quantities

- Basic decisions:
  - *What* to calculate
    - `predictions`, `comparisons`, `slopes`
  - for *whom*
    - `avg_*()`, `newdata =…`, `wts =…`
  - on *which scale*
    - `type =`
  - and, in the case of comparisons, how to *contrast*
    - `comparison =`

Uncertainty and significance

- And then you can do whatever you want with your target quantities
- slap confidence intervals on them
  - or credible intervals if you fitted a Bayesian model
  - `marginaleffects` will always automatically report intervals
- compare them in arbitrary ways
  - null hypothesis tests, equivalence tests
  - of any target quantity or combinations and transformations of them

Worked example

» https://j-rohrer.github.io/marginal-psych/

What else

- `marginaleffects` can do *a lot*
- supports over 100 different classes of models (in R)
- A wealth of case studies provided online
  - conjoint experiments, elasticity, interrupted time series, inverse probability weighting, matching, dealing with missing data, survival analysis, machine learning
- All objects produced are "tidy"
- Lots of plotting functionality (returns `ggplot2` objects)

Read more

- "Models as Prediction Machines" preprint by Vincent Arel-Bundock and me: https://osf.io/preprints/psyarxiv/g4s2a_v2
  - Preprint website with replication package: https://j-rohrer.github.io/marginal-psych/
- Website with documentation including full book on `marginaleffects` by Vincent: https://marginaleffects.com/

Thank you for your attention!

Julia M. Rohrer
juliarohrer.com
Leipzig University
@dingdingpeng.the100.ci
www.the100.ci

Slides: Resources