

Let's take a stab at our first note on a topic that pre-establishing the definitions of [probability model homotopy](#) makes much easier to write.

In this note we will discuss *tailored probability models*. There are models deliberately fit to training data that has an outcome prevalence equal to the expected outcome prevalence on the data they are to be applied on. This is a very typical modeling case, it is achieved for free when the training data is thought to be statistically [exchangeable](#) with the future application data, which is a good experimental design (in our formal notation, this is the  $\mathcal{O}$ -model-homotopy, in the limited case where it is a correct procedure). Tailored models can be simulated by re-weighting or re-sampling the training data to have the same prevalence as expected in the future application data (in our formal notation, this is the  $\mathcal{T}$ -model-homotopy).

Informally, tailored models are very careful models that have been built to anticipate how they are going to be applied in the future. Our claim is: the model tailoring process is not monotone. That is, some predictions reverse order under the model tailoring process. This implies model tailoring is not always as simple as adjusting the predictions in any monotone manner. So, assuming the tailored models are correct, such simple statistical adjustments may in fact be insufficient.

Let us make the above precise and work through an example using logistic regression (a model one might expect to have monotone tailoring properties, but does not).

Using our [probability model homotopy](#) notation and definitions what we were saying above can be refined and condensed into the following technical claim.

Even in the case of logistic regression models, [the tailored probability model homotopy  \$\mathcal{T}\$](#)  can not always be factored into  $\mathcal{T}(x, p) = f_p(m(x))$ , where  $m(x)$  is a probability model.

This statement, once unwound using the definitions, contains all of the content of the earlier claims. The earlier claims are of use, as they help point out why we should care. The discussion emphasizes that if  $\mathcal{T}$  did factor in this way, then a number of simple statistical corrections would be shown to be sufficient, though it turns out they are not.

It only remains to exhibit a simple logistic regression example proving the claim. That is quite easy using [R](#).

```
# attach our packages
library(wrapr)

# build our example data
# modeling y as a function of x1 and x2 (plus intercept)
d <- wrapr::build_frame(
  "x1"  , "x2", "y", "w2" |
    0    , 0  , 0  , 2    |
    0    , 0  , 0  , 2    |
    0    , 1  , 1  , 5    |
    1    , 0  , 0  , 2    |
    1    , 0  , 0  , 2    |
    1    , 0  , 1  , 5    |
    1    , 1  , 0  , 2    )
```

```
# fit a model at prevalence 0.2857143
m_0.29 <- glm(
  y ~ x1 + x2,
  data = d,
  family = binomial())
# add in predictions
d$pred_m_0.29 <- predict(
  m_0.29, newdata = d, type = 'response')
```

```
# fit a model at prevalence 0.5
m_0.50 <- glm(
  y ~ x1 + x2,
  data = d,
  weights = w2,
  family = binomial())
# add in predictions
d$pred_m_0.50 <- predict(
  m_0.50, newdata = d, type = 'response')
```

Now notice the relative order of the predictions in rows 1 and 5 are reversed in model `m_0.50` relative to the order given by model `m_0.29`.

```
interesting_rows <- c(1, 5)
d$pred_m_0.29[interesting_rows]

## [1] 0.2304816 0.1796789

d$pred_m_0.50[interesting_rows]

## [1] 0.3655679 0.3930810
```

This means no monotone correction that looks only at the predictions can make the same adaptations as these two prevalence tailored models. And that is our demonstration.