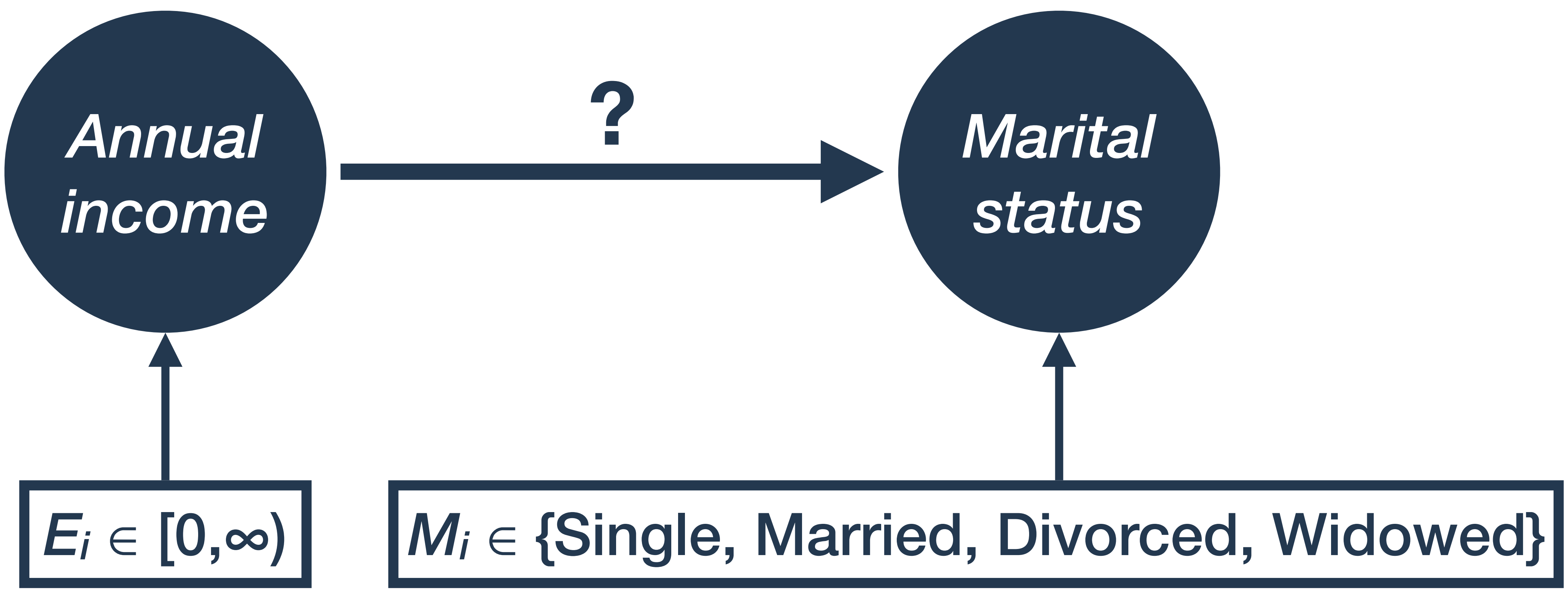
Agenda

Agenda

1. Introducing the multinomial distribution
2. Categorical outcome variables
3. Softmax link function
4. Interpreting coefficients
5. Multinomial logistic in R using ‘brms’

Income and marital status



The problem | Outcome variable has multiple (>2) categories. Binomial and Poisson models won’t work.

The solution | Use a multinomial outcome distribution (and a new link function) to account for the data.

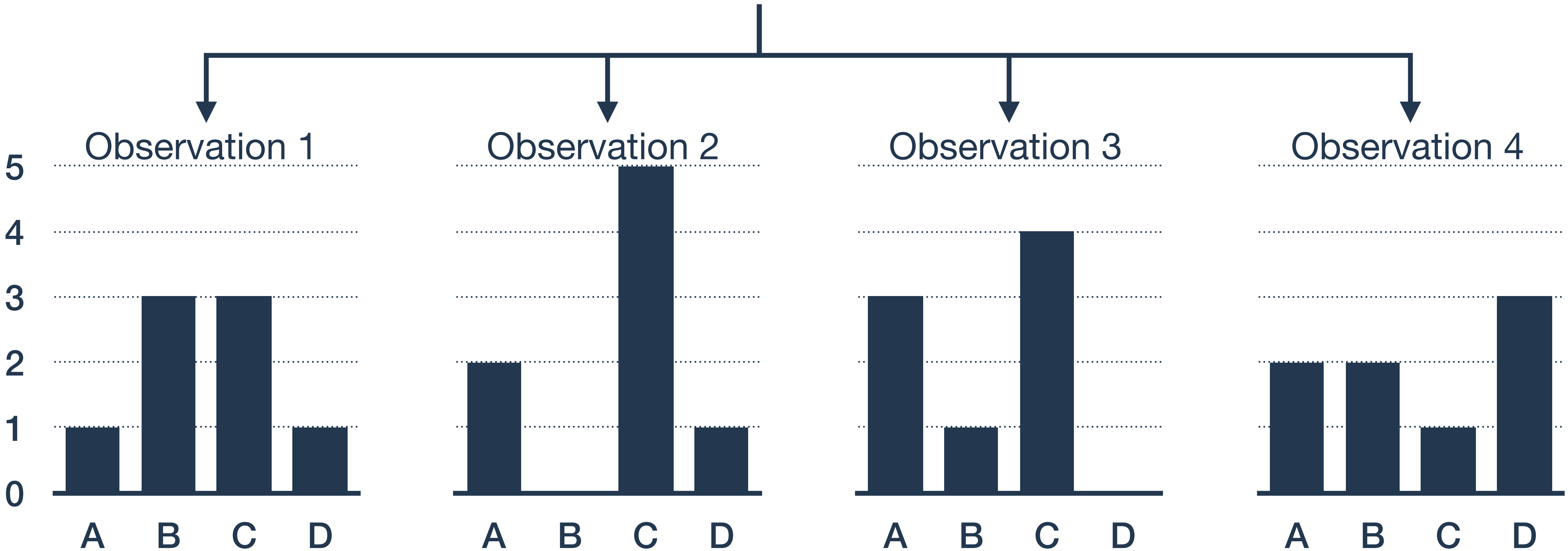
Multinomial distribution

$$\text{Multinom}\left(n, (p_1, \ldots, p_k)\right)$$

Multinomial distribution | Result of n trials, each of which can result in one of k outcomes with probability p_1, p_2, …, p_k.

Each 'observation' describes outcome of *n* trials:

$$\text{Multinom}\left(8, (0.20, 0.10, 0.45, 0.25)\right)$$



Multinomial distribution

Binomial, Bernoulli, and categorical distributions are special cases of the multinomial.

Binomial distribution	$\mathrm{Bin}(n, p) = \mathrm{Multinom}(n, (1-p, p))$
Bernoulli distribution	$\mathrm{Bernoulli}(p) = \mathrm{Multinom}(1, (1-p, p))$
Categorical distribution	$\mathrm{Cat}(p_1, p_2, \ldots, p_k) = \mathrm{Multinom}(1, (p_1, p_2, \ldots, p_k))$

	1 trial	>1 trials
2 categories	Bernoulli	Binomial
>2 categories	Categorical	Multinomial

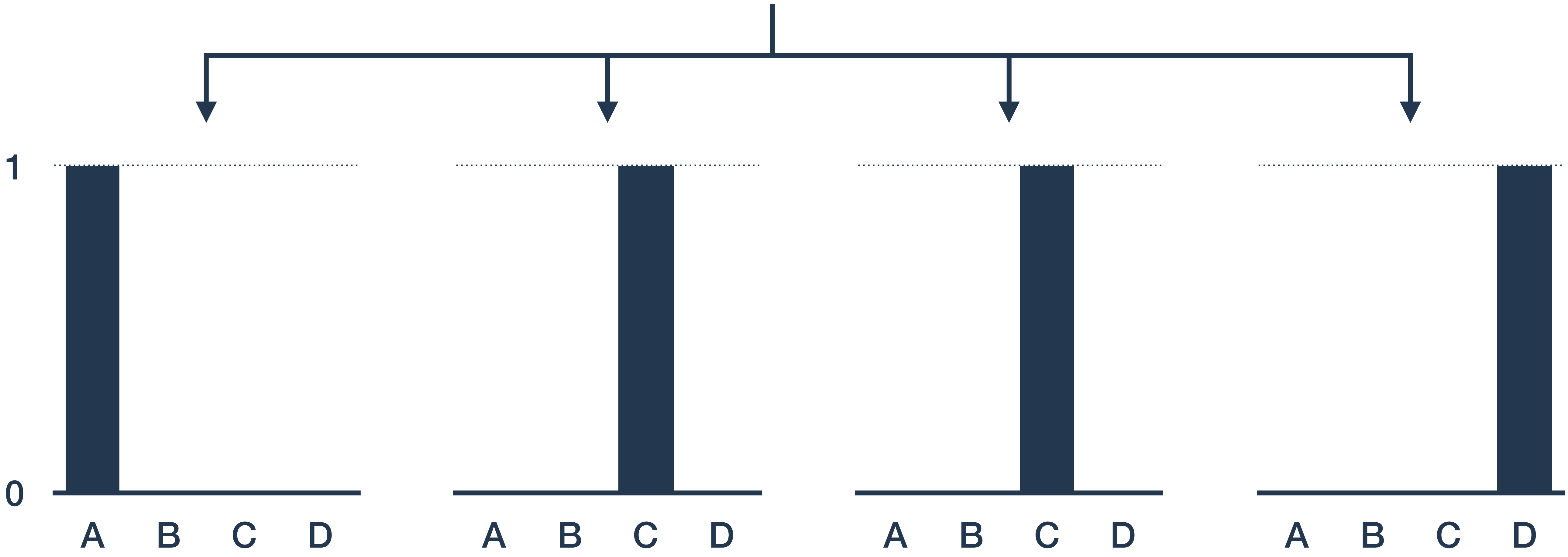
Categorical distribution

$$\text{Cat}\left(p_1, \ldots, p_k\right)$$

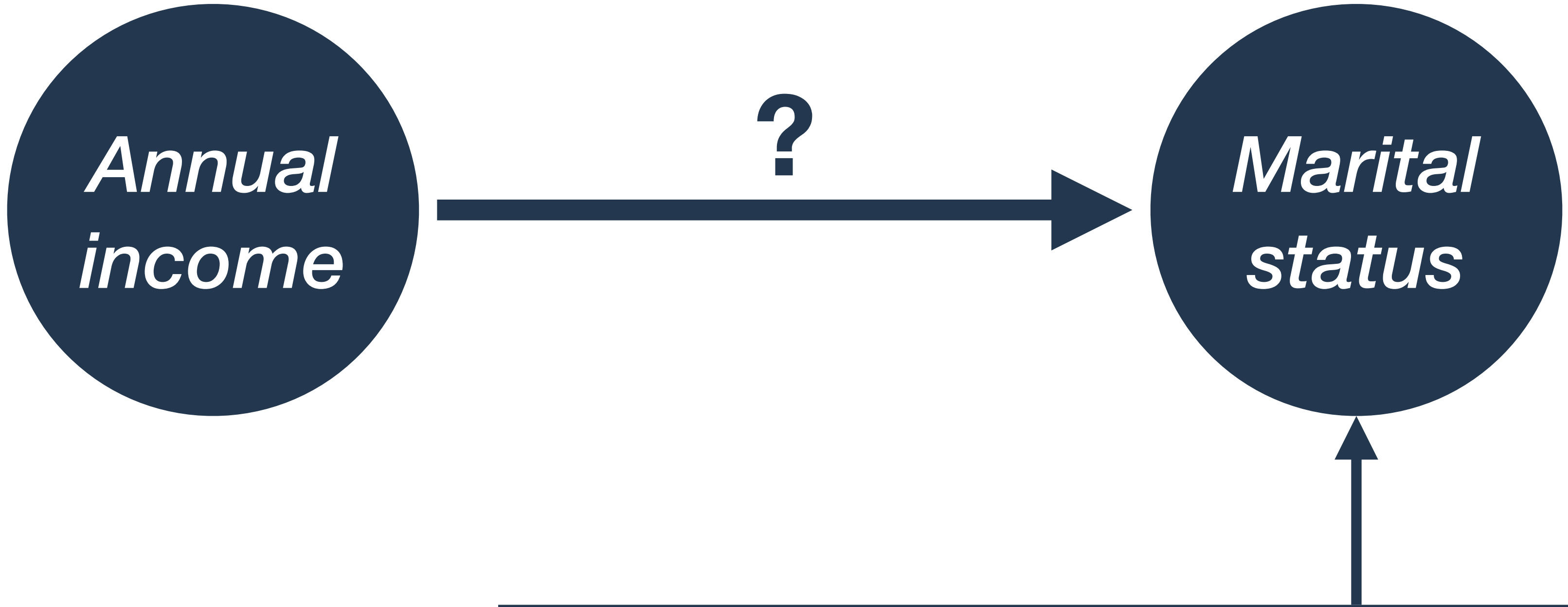
Categorical distribution | Multinomial distribution with just one trial

---

$$\text{Cat}\left(0.20, 0.10, 0.45, 0.25\right)$$



Categorical outcome



$M_i \in$ {Single, Married, Divorced, Widowed}

One category has to be the *reference* category.

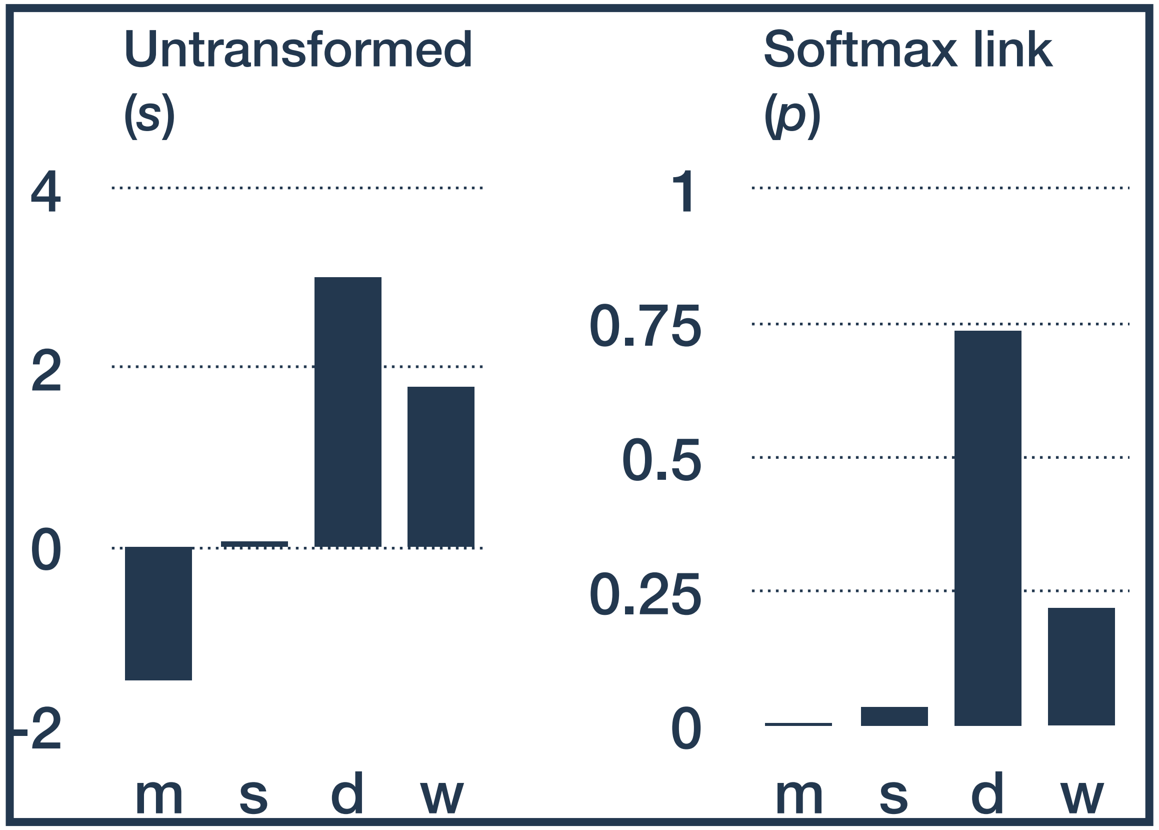
$$s_s = 0$$
$$s_m = a_m + \beta_m E_i$$
$$s_d = a_d + \beta_d E_i$$
$$s_w = a_w + \beta_w E_i$$

Each other category gets its own coefficient.

Softmax link function

$M_i \in$ {Single, Married, Divorced, Widowed}

$$M_i \sim \text{Cat}\left(\text{softmax}(s_s, m_s, s_d, s_w)\right)$$
$$s_s = 0$$
$$s_m = a_m + \beta_m E_i$$
$$s_d = a_d + \beta_d E_i$$
$$s_w = a_w + \beta_w E_i$$



Softmax is a multivariate generalization of inverse logit.

$$p_s = \text{softmax}(s_s) = \frac{\exp(s_s)}{\exp(s_s) + \exp(s_m) + \exp(s_d) + \exp(s_w)}$$

Multinomial logistic regression

Multinomial logistic (or categorical) regression model.

$$
\begin{aligned}
M_i &\sim \text{Cat}\left(\text{softmax}\left(s_{si}, s_{mi}, s_{di}, s_{wi}\right)\right) \\
s_{si} &= 0 \\
s_{mi} &= a_m + \beta_m E_i \\
s_{di} &= a_d + \beta_d E_i \\
s_{wi} &= a_w + \beta_w E_i \\
a_s, a_d, a_w &\sim \text{Norm}(0, 2) \\
\beta_s, \beta_d, \beta_w &\sim \text{Norm}(0, 3)
\end{aligned}
$$

Multinomial logistic regression

With two categories, the multinomial logistic model is the standard (binomial) logistic model.

$$M_i \sim \text{Cat}\left(\text{softmax}(s_{1i}, s_{2i})\right)$$
$$s_{1i} = 0$$
$$s_{2i} = \alpha + \beta E_i$$
$$\alpha \sim \text{Norm}(0, 1)$$
$$\beta \sim \text{Norm}(0, 3)$$

$$p_{2i} = \frac{\exp(s_{2i})}{1 + \exp(s_{2i})} = \text{logit}^{-1}(s_{2i})$$

Interpreting estimates

$$
\begin{aligned}
M_i &\sim \text{Cat}\left(\text{softmax}(s_{mi}, s_{si}, s_{di}, s_{wi})\right) \\
s_{mi} &= 0 \\
s_{si} &= \alpha_s + \beta_s E_i \\
s_{di} &= \alpha_d + \beta_d E_i \\
s_{wi} &= \alpha_w + \beta_w E_i \\
\alpha_s, \alpha_d, \alpha_w &\sim \text{Norm}(0, 2) \\
\beta_s, \beta_d, \beta_w &\sim \text{Norm}(0, 3)
\end{aligned}
$$

	Mean	***90% credible interval***	
$\boldsymbol{\alpha_s}$	5.35	4.73	5.98
$\boldsymbol{\beta_s}$	-0.59	-0.65	-0.53
$\boldsymbol{\alpha_d}$	0.57	-0.24	1.37
$\boldsymbol{\beta_d}$	-0.18	-0.25	-0.10
$\boldsymbol{\alpha_w}$	1.94	0.89	2.98
$\boldsymbol{\beta_w}$	-0.40	-0.50	-0.30

Interpreting estimates

	Mean	*90% credible interval*	
α_s	5.35	4.73	5.98
β_s	-0.59	-0.65	-0.53
α_d	0.57	-0.24	1.37
β_d	-0.18	-0.25	-0.10
α_w	1.94	0.89	2.98
β_w	-0.40	-0.50	-0.30

Estimation in R with brms

Estimating with brms

The `brm()` function allows you to use the model syntax from `lm()` and `glm()`

```
model <- marital_status ~ log_income

fit <- brm(model, data=d)
```

Priors are set using the `prior()` function

```
model <- marital_status ~ log_income
pr <- c(
  prior(normal(0,2), class='b'), # coefficients
  prior(normal(0,3), class='Intercept') # interc.
)
fit <- brm(model, data=d, prior=pr)
```