5.2 Numeric and categorical covariate

Now: two covariates Z, W

Z: categorical $\qquad$ W: numeric

Example: $\quad$ = drive $\qquad$ = weight

Cars2004nh $\quad$ front, rear, 4x4

$\in \{1, 2, 3\}$

$\in \{1, 2, .., G\}$ $\qquad$ Y = consumption

Z: parameterized by chosen (pseudo)contrasts

$$\mathbb{C} = \underbrace{\begin{pmatrix} c_1^T \\ \vdots \\ c_G^T \end{pmatrix}}_{G-1 \text{ columns}} \qquad \text{i.e. } s_Z(z) = \mathbb{C}_z$$

W: parameterized by simple transformation (to start)

i.e. $s_W(w) = s_W(w) \in \mathbb{R}^1$

Example: $s_W(w) = \log(w)$

5.2.1 Additivity

MODEL: $E(Y \mid Z=z, W=w) = m(z,w) =$

$$= \beta_0 + c_z^T \underbrace{\beta^Z}_{= (\beta^Z_1, \ldots, \beta^Z_{G-1})^T} + \beta^W \cdot s_W(w)$$

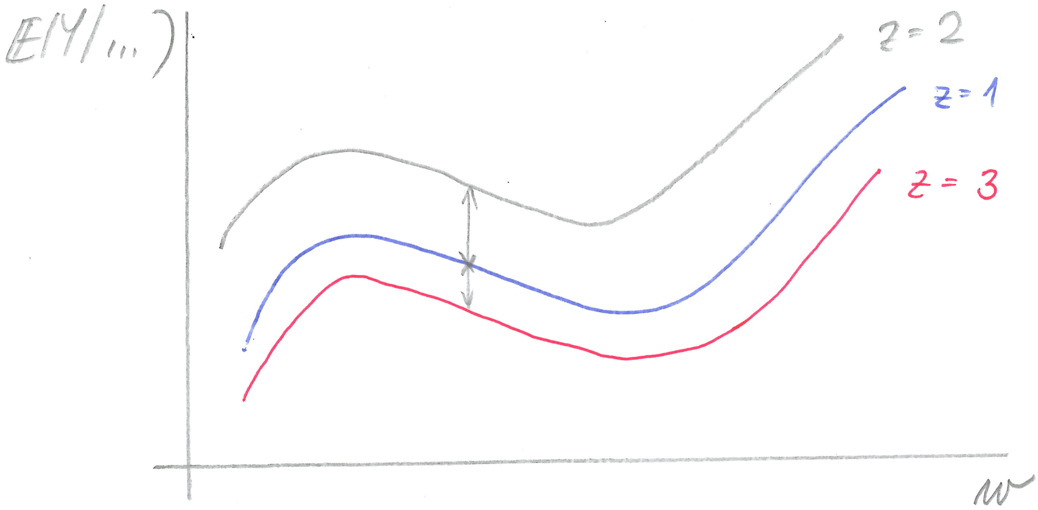
e.g. $\log(w)$

Interpretation:

$$E(Y \mid Z=z, W=w) = \underbrace{\beta_0 + c_z^T \beta^Z} + \beta^W \cdot s_W(w)$$

slope common to all groups

intercept which depends on z and is parameterized by chosen (pseudo) contrasts



plot Cars2004nh 6

plot Cars2004nh 7

Reference group (pseudo) contrasts used for Z

$$\mathbb{C} = \begin{pmatrix} 0 & \cdots & 0 \\ 1 & \cdots & 0 \\ 0 & \cdots & 1 \end{pmatrix}$$

Remember from section 4.4:

$$\beta_0 = m_1$$
$$\beta_1^Z = m_2 - m_1$$
$$\beta_{G-1}^Z = m_G - m_1$$

$m_1, \ldots, m_G$: group means

Here: $E(Y \mid Z = z, W = w) = \beta_0 + \mathbb{C}_z^T \beta^Z + \beta^W \cdot s_W(w)$

$$\beta^Z = (\beta_1^Z, \ldots, \beta_{G-1}^Z)^T$$

β_0 = intercept for group $Z=1$

β_1^Z = intcpt group $Z=2$ $\underset{\text{minus}}{-}$ intcpt group $Z=1$

$\vdots$

β_{G-1}^Z = intcpt group $Z=G$ $\underset{\text{minus}}{-}$ intcpt group $Z=1$

Routput (also slope can be mentioned)

Also: For any $w \in \mathbb{R}$:

$$\beta_1^Z = E(Y \mid Z=2, W=w) - E(Y \mid Z=1, W=w)$$
$$\vdots$$
$$\beta_{G-1}^Z = E(Y \mid Z=G, W=w) - E(Y \mid Z=1, W=w)$$

$$\beta_j^Z - \beta_l^Z = E(Y \mid Z = j+1, W=w) - E(Y \mid Z = l+1, W=w)$$

$j, l = 1, \ldots, G-1$

set of (partial) effects of Z given $W=w$

Sum contrasts used for Z

$$C = \begin{pmatrix} 1 & \cdots & 0 \\ & \ddots & \\ 0 & & 1 \\ -1 & \cdots & -1 \end{pmatrix}$$

Remember from section 4.4:

$$\beta_0 = \overline{m} = \frac{1}{G}\sum_{g=1}^{G} m_g$$

$$\alpha_1^Z := \beta_1^Z = m_1 - \overline{m}$$

$$\vdots$$

$$\alpha_{G-1}^Z := \beta_{G-1}^Z = m_{G-1} - \overline{m}$$

$$\alpha_G^Z = -\sum_{g=1}^{G-1} \beta_g^Z = m_G - \overline{m}$$

$m_1, \ldots, m_G$: group means

Here: $\mathbb{E}(Y \mid Z = z, W = w) = \beta_0 + C_z^T \beta^Z + \beta^W s_W(w)$

$$\beta^Z = (\beta_1^Z, \ldots, \beta_{G-1}^Z)^T$$

Let $\gamma_z = \beta_0 + C_z^T \beta^Z$ = intercept for group $Z = z$, $z = 1, \ldots, G$

$$\beta_0 = \frac{1}{G}\sum_{g=1}^{G} \gamma_g =: \overline{\gamma}$$

$$\alpha_1^Z := \beta_1^Z = \gamma_1 - \overline{\gamma}$$

$$\vdots$$

$$\alpha_{G-1}^Z := \beta_{G-1}^Z = \gamma_{G-1} - \overline{\gamma}$$

$$\alpha_G^Z := -\sum_{g=1}^{G} \beta_g^Z = \gamma_G - \overline{\gamma}$$

10

Also: for any $w \in \mathbb{R}$

$$\alpha_j^Z - \alpha_\ell^Z = \mathbb{E}(Y \mid Z = j, W = w) - \mathbb{E}(Y \mid Z = \ell, W = w)$$

14 plot 11

5.2.2 Partial effects

12

Additivity: Effect of Z (or W) on the response expectation does not depend on a value of W (or Z).

MODEL: $\mathbb{E}(Y \mid Z=z, W=w) = \beta_0 + c_z^T \beta^Z + \beta^W s_W(w)$

plot 13

Partial effect of a numeric covariate:

$$\mathbb{E}(Y \mid Z=z, W=w+1) - \mathbb{E}(Y \mid Z=z, W=w) =$$

$$= \underline{\beta^W} \left(s_W(w+1) - s_W(w) \right)$$

No effect $\equiv$ $H_0: \beta^W = 0$ 14

$\rightarrow$ t-test on regression coef.

Partial effect of a categorical covariate 15

$$\mathbb{E}(Y \mid Z=j, W=w) - \mathbb{E}(Y \mid Z=\ell, W=w)$$

$$= (c_j^T - c_\ell^T)\beta^Z \qquad j, \ell = 1, \dots, G$$

No effect $\equiv$ $H_0: \beta^Z = 0$

$\rightarrow$ Wald F-test on a (subvector) of regression coefs.

Equivalent way of testing no effect of Z given W:

If $\beta^Z = \mathbb{0}$, a (sub)model is obtained ~~with~~ which assumes $\mathbb{E}(Y \mid Z = z, W = w) =$

$$= \beta_0 + \beta^W s_W(w)$$

$$H_0: \mathbb{E}(Y \mid Z, W) \in \mathcal{M}(\mathbb{X}^0)$$

$$\mathbb{X}^0 = \begin{pmatrix} 1 & s_W(W_1) \\ \vdots & \vdots \\ 1 & s_W(W_n) \end{pmatrix}$$

$$H_1: \mathbb{E}(Y \mid Z, W) \in \mathcal{M}(\mathbb{X}) \setminus \mathcal{M}(\mathbb{X}^0)$$

$$\mathbb{X} = \begin{pmatrix} 1 & \mathbb{C}^T_{Z_1} & s_W(W_1) \\ \vdots & \vdots & \vdots \\ 1 & \mathbb{C}^T_{Z_n} & s_W(W_n) \end{pmatrix}$$

→ F-test on submodel

16

- in this context also called as <u>Analysis of covariance</u> (ANCOVA)

5.2.3 Interactions

$$\mathbb{S}_W^T(w) \otimes \mathbb{S}_Z^T(z) = S_W(w) \cdot \mathbb{C}_Z^T =: \underbrace{\mathbb{S}_{ZW}^T(z,w)}_{\text{vector of length } G-1}$$

$$\underset{1}{S_W(w)} \otimes \underset{G-1}{\mathbb{C}_Z^T}$$

Interaction model in general

$$E(Y \mid Z=z, W=w) = \beta_0 + \mathbb{S}_Z^T(z)\beta^Z + \mathbb{S}_W^T(w)\beta^W + \mathbb{S}_{ZW}^T(z,w)\beta^{ZW}$$

$$\overset{\text{NOW}}{=} \beta_0 + \mathbb{C}_Z^T\beta^Z + S_W(w)\cdot\beta^W + S_W(w)\,\mathbb{C}_Z^T\beta^{ZW}$$

$$\beta^Z = (\beta_1^Z, \ldots, \beta_{G-1}^Z)^T \qquad \beta^{ZW} = (\beta_1^{ZW}, \ldots, \beta_{G-1}^{ZW})^T$$

$$= \underbrace{(\beta_0 + \mathbb{C}_Z^T\beta^Z)}_{\gamma_{z,0}} + \underbrace{(\beta^W + \mathbb{C}_Z^T\beta^{ZW})}_{\gamma_{z,1}} S_W(w)$$

$\gamma_{z,0}$ intercept (depends on z)

$\gamma_{z,1}$ slope (depends on z)

plot Cars2004nh ~~#~~ 18

plot Cars2004nh ~~#~~ 19

Reference group (pseudo) contrasts used for Z

Remember from Section 4.4

$\beta_0 = m_1$ $m_1, \ldots, m_G$

$\beta_1^Z = m_2 - m_1$ group means

$\vdots$

$\beta_{G-1}^Z = m_G - m_1$

Now:

$$(\beta_1^Z, \ldots, \beta_{G-1}^Z)^T \qquad (\beta_1^{ZW}, \ldots, \beta_{G-1}^{ZW})^T$$

$$E(Y \mid Z = z, W = w) = \underbrace{(\beta_0 + \mathbb{C}_z^T \beta^Z)}_{\substack{\text{intercepts} \\ =: \gamma_{z,0}}} + \underbrace{(\beta^W + \mathbb{C}_z^T \beta^{ZW})}_{\substack{\text{slopes} \\ =: \gamma_{z,1}}} S_w(w)$$

$\beta_0 = \gamma_{1,0}$ (intercept for group $Z = 1$)

$\beta_1^Z = \gamma_{2,0} - \gamma_{1,0}$ (intcpt group $Z=2$ — minus intcpt group $Z=1$)

$\vdots$

$\beta_{G-1}^Z = \gamma_{G,0} - \gamma_{1,0}$

$\beta^W = \gamma_{1,1}$ (slope for group $Z=1$)

$\beta_1^{ZW} = \gamma_{2,1} - \gamma_{1,1}$ (slope group $Z=2$ — minus slope group $Z=1$)

$\vdots$

$\beta_{G-1}^{ZW} = \gamma_{G,1} - \gamma_{1,1}$

Sum contrasts used for Z

Remember from section 4.4

$$\beta_0 = \bar{m} = \frac{1}{G}\sum_{g=1}^{G} m_g \qquad m_1, \dots, m_G: \text{ group means}$$

$$\alpha_1^Z = \beta_1^Z = m_1 - \bar{m}$$

$$\alpha_{G-1}^Z = \beta_{G-1}^Z = m_{G-1} - \bar{m}$$

$$\alpha_G^Z = -\sum_{g=1}^{G-1} \beta_g^Z = m_G - \bar{m}$$

NOW:

$$E(Y \mid Z=z, W=w) = (\underbrace{\beta_0 + c_z^T \beta^Z}_{\substack{\text{intercepts} \\ =: \gamma_{z,0}}}) + (\underbrace{\beta^W + \cancel{\beta}\, c_z^T \beta^{ZW}}_{\substack{\text{slopes} \\ =: \gamma_{z,1}}}) \quad S_W(w)$$

$$\beta_0 = \frac{1}{G}\sum_{g=1}^{G} \gamma_{g,0} =: \bar{\gamma}_0 \quad \text{(mean intercept)}$$

$$\alpha_{1,0}^Z := \beta_1^Z = \gamma_{1,0} - \bar{\gamma}_0 \quad \text{(intcpt group } Z=1 \text{ — mean intcpt)}$$

$$\vdots$$

$$\alpha_{G-1,0}^Z := \beta_{G-1}^Z = \gamma_{G-1,0} - \bar{\gamma}_0$$

$$\alpha_{G,0}^Z := -\sum_{g=1}^{G-1} \beta_g^Z = \gamma_{G,0} - \bar{\gamma}_0$$

$$\beta^W = \frac{1}{G}\sum_{g=1}^{G} \gamma_{g,1} =: \bar{\gamma}_1 \quad \text{(mean slope)}$$

$$\alpha_{1,1}^Z := \beta_1^{ZW} = \gamma_{1,1} - \bar{\gamma}_1 \quad \text{(slope group } Z=1 \text{ — mean slope)}$$

$$\alpha_{G-1,1}^Z := \beta_{G-1}^{ZW} = \gamma_{G-1,1} - \bar{\gamma}_1$$

$$\alpha_{G,1}^Z := \cancel{\beta_G} - \sum_{g=1}^{G-1} \beta_g^{ZW} = \gamma_{G,1} - \bar{\gamma}_1$$

5.2.4 Additivity or interactions? 22

Can additivity be assumed?

23

H_0: model which assumes

$$\mathbb{E}(Y \mid Z = z, W = w) = \beta_0 + c_z^T \beta^Z + \beta^W s_W(w)$$

holds

H_1: interactions are needed

H_0: $\beta^{ZW} = \mathbb{0}$ in interaction model

→ (Wald) F-test on a subvector of regression coefficients

or equivalently submodel F-test

24

! If additivity rejected, it does not make much sense to interprete and estimate partial effects!

5.2.5 More complex parameterizations of a numeric covariate

- numeric covariate W can (indeed) be parameterized in a more complex way using $\mathbf{s}_W = (s_W^1, \dots, s_W^{k-1})^T$

  → polynomials

  → splines

Additivity model

$$E(Y|Z=z, W=w) = \underbrace{\beta_0 + \mathbf{c}_z^T \beta^Z}_{\text{intercept which depends on } z} + \mathbf{s}_W^T(w)\, \beta^W$$

Interaction model

$$E(Y|Z=z, W=w) =$$

$$= \underbrace{\beta_0 + \mathbf{c}_z^T \beta^Z}_{\text{intercept which depends on } z} + \mathbf{s}_W^T(w)\, \beta^W + \overbrace{(\mathbf{s}_W^T \otimes \mathbf{c}_z^T)}^{(s_W^1(w)\, \mathbf{c}_z^T, \dots, s_W^{k-1}(w)\, \mathbf{c}_z^T)} \beta^{ZW}$$

$$\beta^W = (\beta_1^W, \dots, \beta_{k-1}^W)^T$$

$$\beta^{ZW} = (\underbrace{\beta_{1,1}^{ZW}, \dots, \beta_{G-1,1}^{ZW}}_{\beta_{\bullet 1}^{ZW}}, \dots, \underbrace{\beta_{1,k-1}^{ZW}, \dots, \beta_{G-1,k-1}^{ZW}}_{\beta_{\bullet k-1}^{ZW}})^T$$

$$=$$

$$= \beta_0 + \mathbb{C}_z^T \beta^z + \sum_{j=1}^{k-1} \beta_j^w s_w^j(w) + \sum_{j=1}^{k-1} \mathbb{C}_z^T \beta_{\bullet j}^{zw} s_w^j(w)$$

$$= \underbrace{(\beta_0 + \mathbb{C}_z^T \beta^z)}_{\text{intercepts which depend on } z} + \sum_{j=1}^{k-1} \underbrace{(\beta_j^w + \mathbb{C}_z^T \beta_{\bullet j}^{zw})}_{\text{slopes which depend on } z} s_w^j(w)$$

---

e.g. $s_w(w) = (w, w^2, \dots, w^{k-1})^T$ (polynomial of degree $k-1$)

$$E(Y \mid Z=z, W=w) = (\beta_0 + \mathbb{C}_z^T \beta^z) + \sum_{j=1}^{k-1} \underbrace{(\beta_j^w + \mathbb{C}_z^T \beta_{\bullet j}^{zw})}_{\text{polynomial coefficients which depend on } z} w^j$$