Data analysis with R software

Data-analyysi R-ohjelmistolla

Tommi Härkänen

National Institute for Health and Welfare (THL), Helsinki
E-mail: tommi.harkanen@helsinki.fi

University of Helsinki, February 14, 2012

Contents

Interaction of a categorical and a continuous covariate

Interaction of two categorical covariates

Regression coefficients

Interaction of continuous and categorical covariates

Imaginary example in R: $lm(y \sim age + gender + age*gender)$

```
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)             2.0              ...
age                     0.1              ...
genderFemale            3.0              ...
age:genderFemale       -0.2              ...
```

Age	Gender	Linear predictor		Prediction
0	Male	$2.0 + 0{\times}0.1 + 0{\times}3.0 + (\text{-}0.2){\times}0{\times}0$	$=$	2.0
0	Female	$2.0 + 0{\times}0.1 + 1{\times}3.0 + (\text{-}0.2){\times}0{\times}0$	$=$	5.0
40	Male	$2.0 + 40{\times}0.1 + 0{\times}3.0 + (\text{-}0.2){\times}40{\times}0$	$=$	6.0
40	Female	$2.0 + 40{\times}0.1 + 1{\times}3.0 + (\text{-}0.2){\times}40{\times}1$	$=$	1.0

Regression coefficients

Interaction of continuous and categorical covariates

$$\mathbb{E}[Y_i] = \beta_0 + \beta_{\text{gender=Female}} + \beta_{\text{Age}} \times \text{Age}_i + \beta_{\text{Age,Gender=Female}} \times \text{Age}_i$$
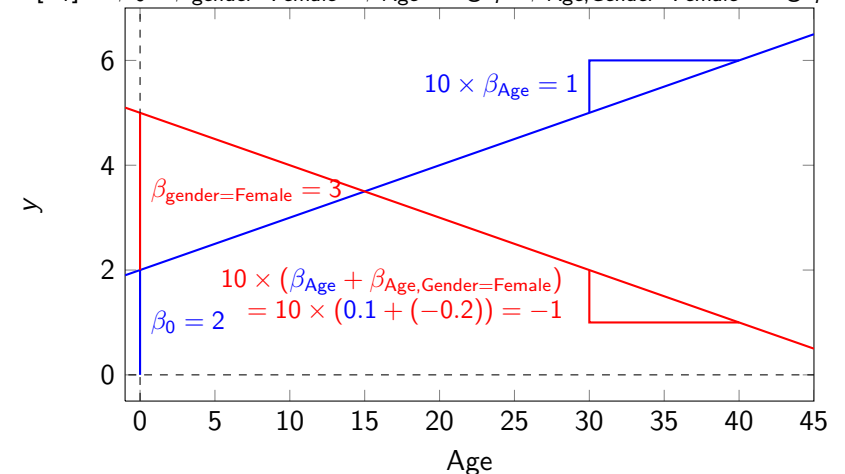
Example of interaction of two categorical covariates

Using Nhanes data. Regress weight on gender, smoking (`har1`,
"Have you smoked 100+ cigarettes in life") and their interaction.

```
> fit1 <- with(nhanes, lm(ham6s_kg ~ hssex + har1 + hssex*har1))
> round(summary(fit1)$coefficients, d=2)
```

```
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                  79.27       0.64  124.74     0.00
hssexFemale                  -8.05       1.10   -7.35     0.00
har1No (HAR14)               -1.94       1.06   -1.83     0.07
hssexFemale:har1No (HAR14)   -1.62       1.51   -1.07     0.28
```

Gender	Smoking	Linear predictor	Prediction
Male	Yes	79.27 + 0×-8.05 + 0×-1.94 + 0×-1.62 = 79.27	
Female	Yes	79.27 + 1×-8.05 + 0×-1.94 + 0×-1.62 = 71.22	
Male	No (HAR14)	79.27 + 0×-8.05 + 1×-1.94 + 0×-1.62 = 77.33	
Female	No (HAR14)	79.27 + 1×-8.05 + 1×-1.94 + 1×-1.62 = 67.65	