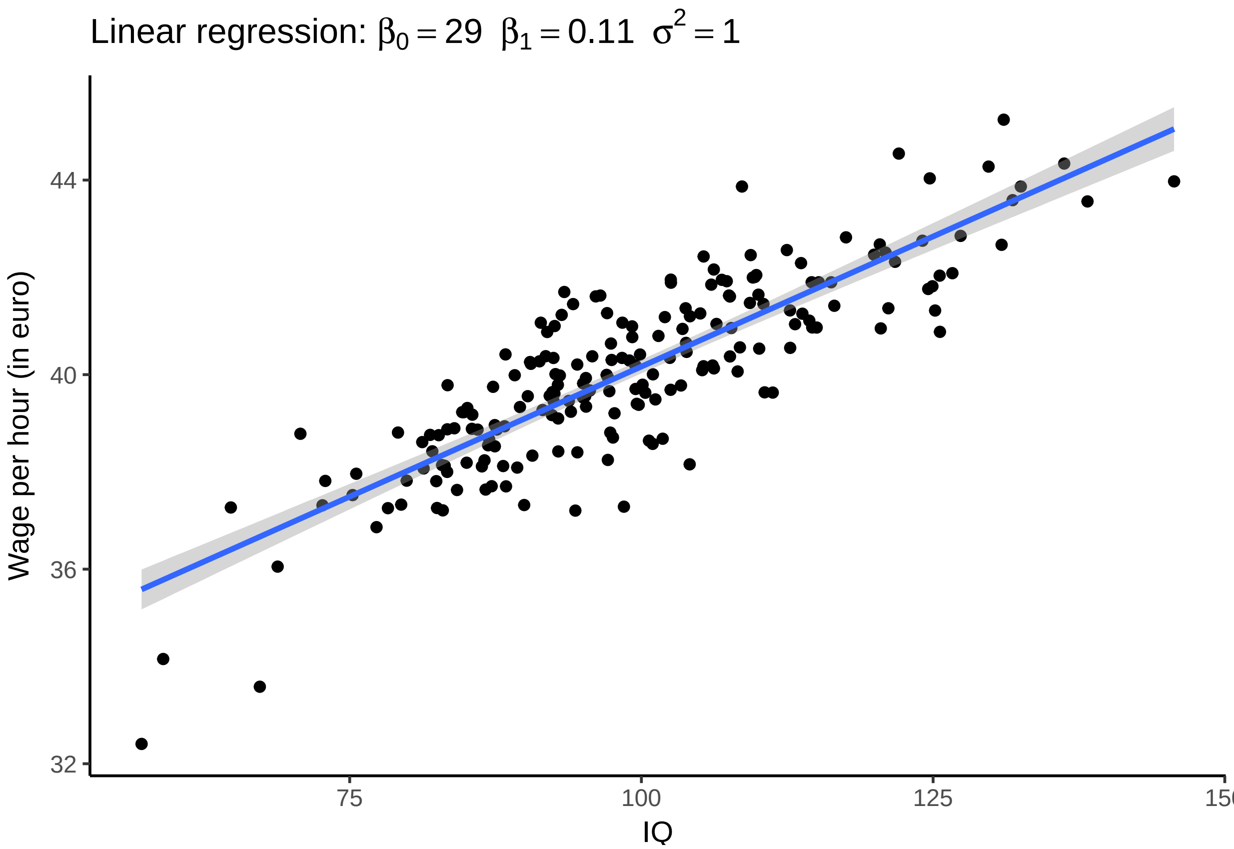
Lab 10

Wiktor Soral

December 11th 2017

Simple linear regression

Line of the best fit



Understanding parameters

- **intercept** (β_0) - place of the crossing of regression line and y-axis, e.g. what would be average wage for a person with 0 IQ
- because such intercept cannot be interpretted in a substantive way (imagine a person with 0 IQ), a popular approach is to center predictors on mean value (i.e. take mean as 0); as a result intercept can be interpretted, e.g. what would be average wage of a person with average IQ

Understanding parameters

- **slope** (β_0) - how much outcome variable will increase if we change a predictor by 1 unit, e.g. what is the difference in wages between people with IQ of 100 and 101

- commonly in psychology predictors have scales which are difficult to interpret, e.g. we don't really now what it means that a person obtained 16 points on a test, and how it is different from a person who obtained 17 points
- to operate on a more interpretable scales, we often standardize both predictor and outcome variables (this is also handy in the case of models with more than 1 predictor) - 1 unit becomes a 1 standard deviation
- because standardizing is also centering, there is no loss of interpretability of the intercept

Understanding parameters

- **variance of residuals (σ^2)** - variance of the differences between observed and predicted values
- this value explains how much unexplained variance of the outcome variable is still left (after we fitted regression line)
- by looking at the standard deviation (σ), we can say what is the average error we make by accepting the regression line as our model

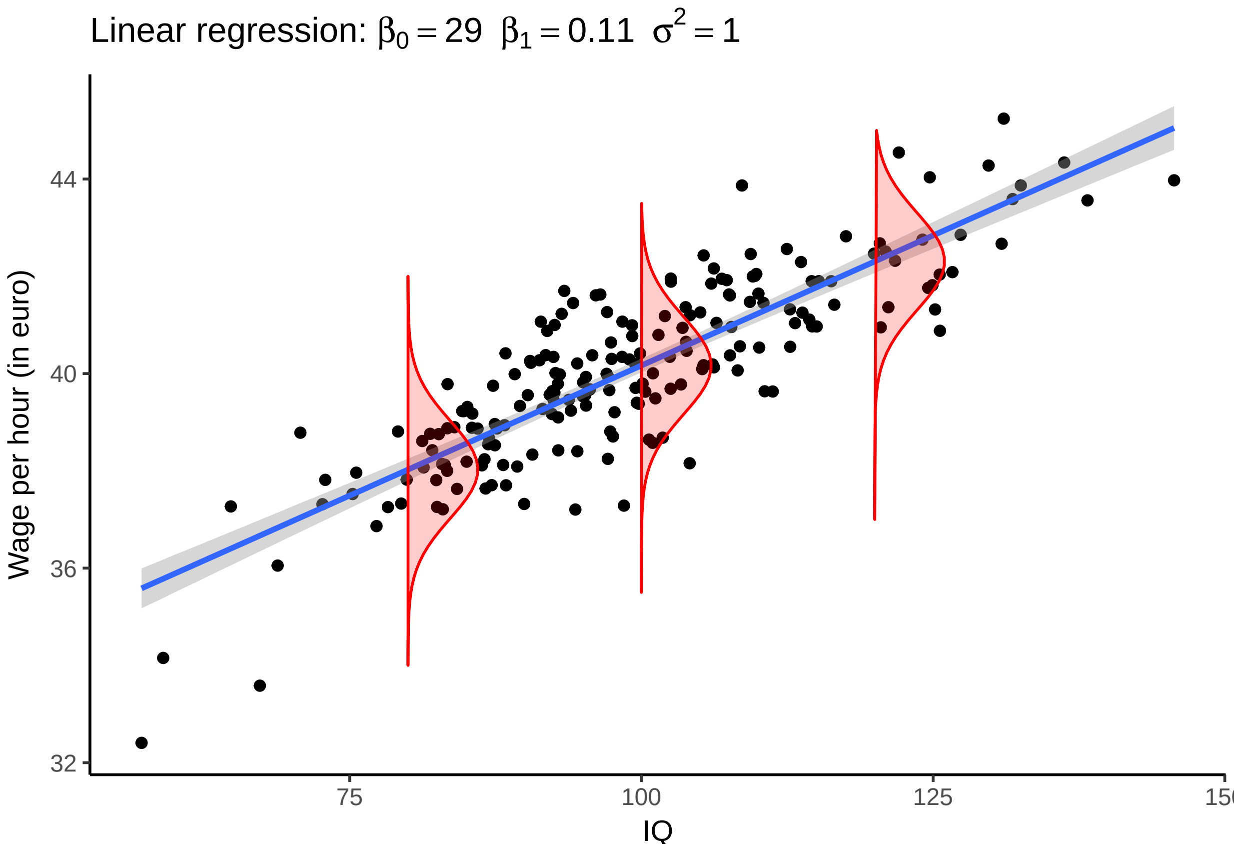
Understanding parameters

- **grey area** - explains uncertainty associated with out parameters' estimates - in Bayesian approach this is 95% credible interval
- in other words in Bayesian approach, we do not claim that the regression line is strictly the blue one, but rather that it lays somewhere over the gray region

Bayesian Linear Model

- $y \sim \mathcal{N}(\mu, \sigma^2)$
- $\mu = \beta_0 + \beta_1 * x$
- in other words y is distributed according to Normal distribution with mean μ and variance σ^2
- note that μ changes (is conditional) on the value of x

Bayesian Linear Model



- notice that we assume that the variance of errors is the same for all values of IQ - this is so called homoscedasticity assumption

Priors on the parameters

- $\beta_0 \sim \mathcal{N}(m_0, s_0)$
- $\beta_0 \sim \mathcal{N}(m_1, s_1)$
- $\sigma \sim \mathcal{U}(lower, upper)$
- values of the hyperparameters depend largely on the scale of x and y variables
- for weakly informative priors (standardized variables) it is common to set m_0 and m_1 to 0, s_0 and s_1 to 10, lower to $1e-3$ and upper to $1e+3$

Robust linear regression

- As with robust Bayesian t-test, we can use Student's t distribution, instead of Normal distribution to model our data, robustly against outliers
- $y \sim \mathcal{T}(\mu, \sigma^2, \nu)$
- $\mu = \beta_0 + \beta_1 * x$
- With prior on normality parameter ν, the same as previously:
- $(\nu - 1) \sim \mathcal{E}(1/29)$