Error in library("abd"): there is no package called 'abd'

R01a - Simple linear regression:
Choosing explanatory variables

STAT 5870 (Engineering)
Iowa State University

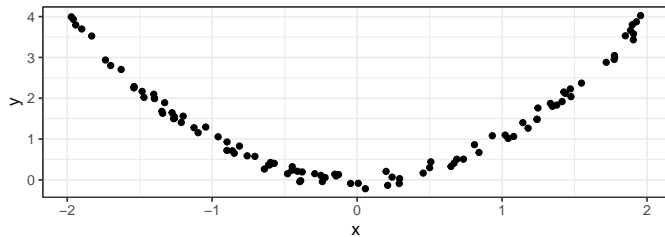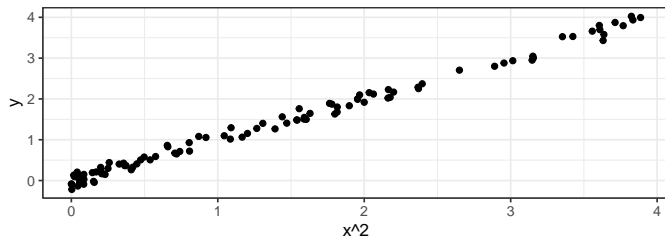August 28, 2024

Simple linear regression

Let

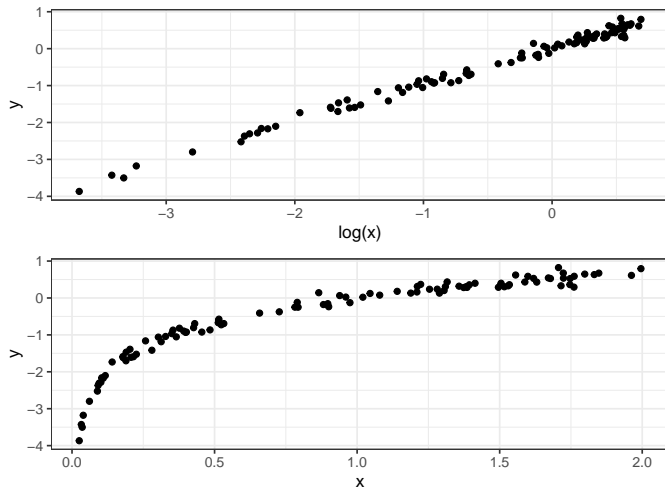$$Y_i \overset{ind}{\sim} N(\beta_0 + \beta_1 f(X_i), \sigma^2).$$

Possible choices for f:

- quadratic: $f(x) = x^2$
- logarithmic: $f(x) = \log(x)$
- centered: $f(x) = x - m$
- scaled: $f(x) = x/s$

Quadratic relationship

Logarithmic relationship

Shifting the intercept

The intercept is the expected response when the explanatory variable is zero. If we use

$$f(x) = x - m,$$

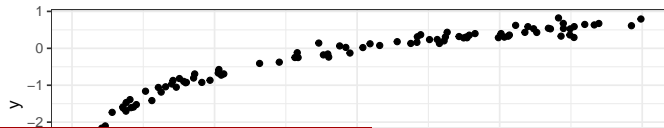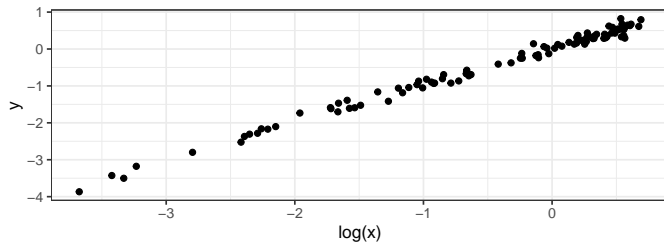then the new intercept is the expected response when the explanatory variable is m.

$$E[Y|X = x] = \beta_0 + \beta_1(x - m) = \tilde{\beta}_0 + \tilde{\beta}_1 x$$

so our new parameters for the mean are

- slope $\tilde{\beta}_1 = \beta_1$ (unchanged) but
- intercept $\tilde{\beta}_0 = (\beta_0 - m\beta_1)$.

Telomere data

```
Error in loadNamespace(x):  there is no package called 'abd'
Error in eval(expr, envir, enclos):  object 'Telomeres' not found
Error in eval(expr, envir, enclos):  object 'Telomeres' not found
```

Telomere data: shifting the intercept

```
m0 = lm(telomere.length ~   years    , abd::Telomeres)

Error in loadNamespace(x):  there is no package called 'abd'

m4 = lm(telomere.length ~ I(years-5), abd::Telomeres)

Error in loadNamespace(x):  there is no package called 'abd'

coef(m0)

Error in eval(expr, envir, enclos):  object 'm0' not found

coef(m4)

Error in eval(expr, envir, enclos):  object 'm4' not found

confint(m0)

Error in eval(expr, envir, enclos):  object 'm0' not found

confint(m4)

Error in eval(expr, envir, enclos):  object 'm4' not found
```

Rescaling the slope

The slope is the expected increase in the response when the explanatory variable increases by 1. If we use

$$f(x) = x/s,$$

then the new slope is the expected increase in the response when the explanatory variable increases by s.
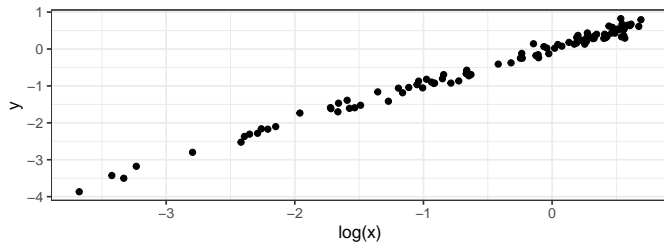
$$E[Y|X = x] = \beta_0 + \beta_1(x/s) = \tilde{\beta}_0 + \tilde{\beta}_1 x$$

so our new parameters are

- intercept $\tilde{\beta}_0 = \beta_0$ (unchanged) but
- slope $\tilde{\beta}_1 = \beta_1/s$.

Telomere data: rescaling the slope

```
Error in loadNamespace(x):  there is no package called 'abd'
Error in eval(expr, envir, enclos):  object 'Telomeres' not found
Error in eval(expr, envir, enclos):  object 'Telomeres' not found
```
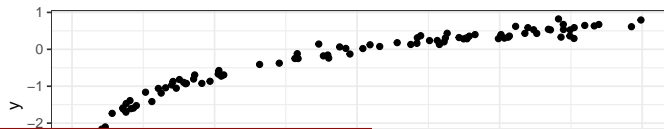
Telomere data: rescaling the slope

```
m0 = lm(telomere.length ~   years    , abd::Telomeres)

Error in loadNamespace(x):  there is no package called 'abd'

m4 = lm(telomere.length ~ I(years/2), abd::Telomeres)

Error in loadNamespace(x):  there is no package called 'abd'

coef(m0)

Error in eval(expr, envir, enclos):  object 'm0' not found

coef(m4)

Error in eval(expr, envir, enclos):  object 'm4' not found

confint(m0)

Error in eval(expr, envir, enclos):  object 'm0' not found

confint(m4)

Error in eval(expr, envir, enclos):  object 'm4' not found
```

Summary

Let

$$Y_i \overset{ind}{\sim} N(\beta_0 + \beta_1 f(X_i), \sigma^2).$$

Choose f based on

- Scientific understanding
- Interpretability
- Diagnostics