ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Lineare Regression: Tests

Statistik (Biol./Pharm./HST) – FS 2015

Ersatz: Cooper & Shuttle

- 12-Minuten Test nach Cooper (1968)
- 20m-Shuttle-Test nach Leger (1982)

Eur J Appl Physiol (1982) 49: 1–12

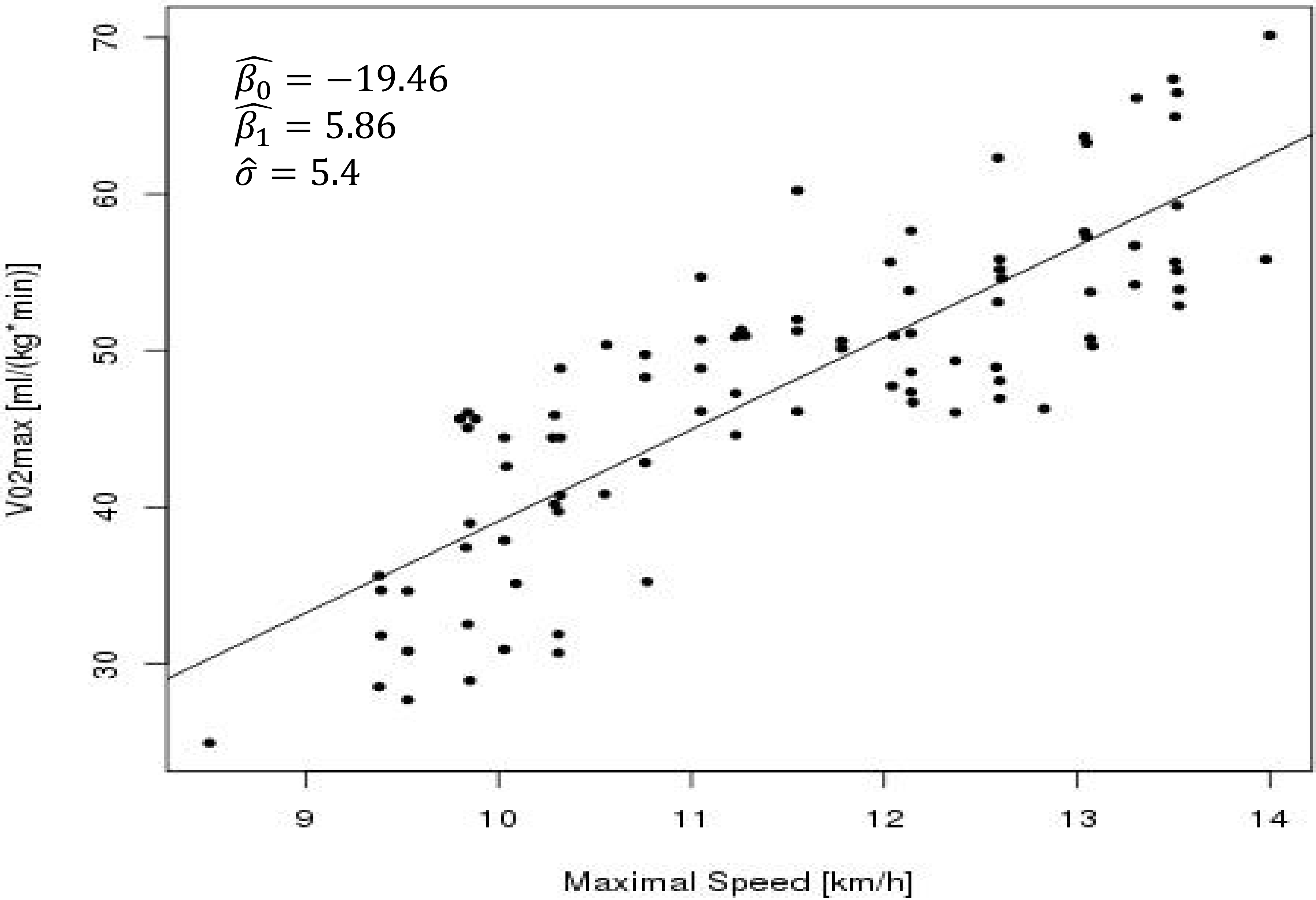
European Journal of
Applied Physiology
and Occupational Physiology

A Maximal Multistage 20-m Shuttle Run Test to Predict $\dot{V}O_2$ max*

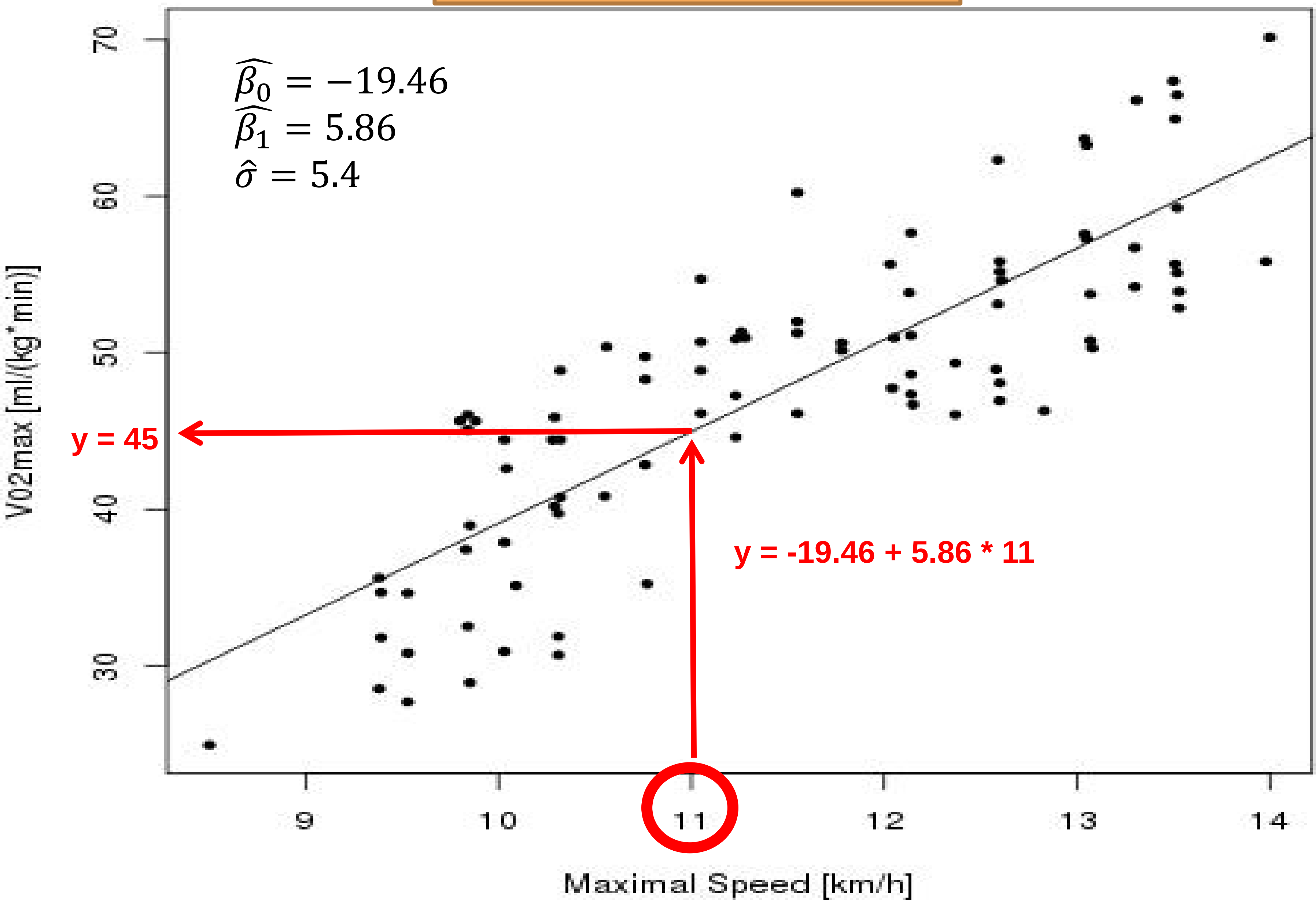
Luc A. Léger[1] and J. Lambert[2]

[1] Département d'éducation physique, Université de Montréal,
CEPSUM, C.P. 6128, Succ. "A", Montréal (Québec), Canada, H3C 3J7
[2] Département de Médecine sociale et préventive, Université de Montréal, Canada

Methode der kleinsten Quadrate



- Wie genau stimmen Parameter?
- Wie genau stimmt Vorhersage?



t-Test in der Linearen Regression

1. Modell: $Y_i = \beta_0 + \beta_1 x_i + E_i, E_1, \ldots, E_n \;\; iid \;\; N(0, \sigma^2)$
2. Nullhypothese: $H_0: \beta_1 = 0$
   Alternative: $H_A: \beta_1 \neq 0$ (es wird normalerweise ein zweiseitiger Test durchgeführt)
3. Teststatistik:
   $$T = \frac{beobachtet - erwartet}{geschätzter\ Standardfehler} = \frac{\widehat{\beta_1} - 0}{\widehat{s.e.}\left(\widehat{\beta_1}\right)}$$
   Dabei ist $s.e.\left(\widehat{\beta_1}\right) = \sqrt{Var\left(\widehat{\beta_1}\right)}$ der “Standard Error” von $\widehat{\beta_1}$
   Verteilung der Teststatistik unter $H_0: T \sim t_{n-2}$
4. Signifikanzniveau: α
5. Verwerfungsbereich der Teststatistik:
   $$K = \left(-\infty, -t_{n-2;1-\frac{\alpha}{2}}\right] \cup \left[t_{n-2;1-\frac{\alpha}{2}}, \infty\right)$$
6. Testentscheid: Überprüfe, ob der beobachtete Wert der Teststatistik im Verwerfungsbereich liegt.

Lineare Regression in R

Modell: $Y_i = \beta_0 + \beta_1 x_i + E_i,\ E_i \sim N(0, \sigma^2)\ i.i.d$

Modell: $Y_i = -19.46 + 5.86 x_i + E_i,\ E_i \sim N(0, 5.43^2)\ i.i.d$

Standardfehler von $\widehat{\beta_1}$ $(= \hat{\sigma}_{\widehat{\beta_1}})$

Approx. 95%-VI:

$5.86 \pm 2 * 0.41$

Exaktes 95%-VI:

$t_{89;0.975}$

Beobachtete Teststatistik

im Test $H_0: \beta_1 = 0$ vs. $H_A: \beta_1 \neq 0$

P-Wert:

Angenommen $\beta_1 = 0$; wie wa. ist Beobachtung oder etwas extremeres?

```
> fitShuttle <- lm(vo2max ~ vmax, data = dat)
> summary(fitShuttle)

Call:
lm(formula = vo2max ~ vmax, data = dat)

Residuals:
     Min       1Q   Median       3Q      Max
-10.2230  -4.3976  -0.2016   4.7026  12.0348

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -19.4582     4.7239  -4.119  8.5e-05 ***
vmax          5.8566     0.4082  14.347  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.433 on 89 degrees of freedom
Multiple R-squared: 0.6981, Adjusted R-squared: 0.6948
F-statistic: 205.8 on 1 and 89 DF,  p-value: < 2.2e-16
```

Freiheitsgrade: n – (Anz. β's) = 91 – 2 = 89

Bsp Prüfungsfrage

Kann $H_0: \beta_1 = 0$ auf dem 5%-Signifikanzniveau verworfen werden?

```
Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.9552 -1.3273 -0.0089  1.2986  3.5242 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)   
(Intercept)   1.0289     0.3302   3.116  0.00385 **
x             1.8859     0.2777   [redacted]
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.925 on 32 degrees of freedom
Multiple R-squared: 0.5904, Adjusted R-squared: 0.5776 
F-statistic: 46.13 on 1 and 32 DF,  p-value: 1.119e-07
```

Bsp Prüfungsfrage

Kann $H_0: \beta_1 = 0$ auf dem 5%-Signifikanzniveau verworfen werden?

```
Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.9552 -1.3273 -0.0089  1.2986  3.5242 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   1.0289     0.3302   3.116  0.00385 ** 
x             1.8859     0.2777   6.792 1.12e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.925 on 32 degrees of freedom
Multiple R-squared: 0.5904, Adjusted R-squared: 0.5776 
F-statistic: 46.13 on 1 and 32 DF,  p-value: 1.119e-07
```

Ja:

- t value: 1.88/0.277 = 6.79
- Verwerfungsbereich:

$\left[-\infty, -t_{32;0.975}\right] \cup \left[t_{32;0.975}, \infty\right]$

$t_{32;0.975} = 2.037$

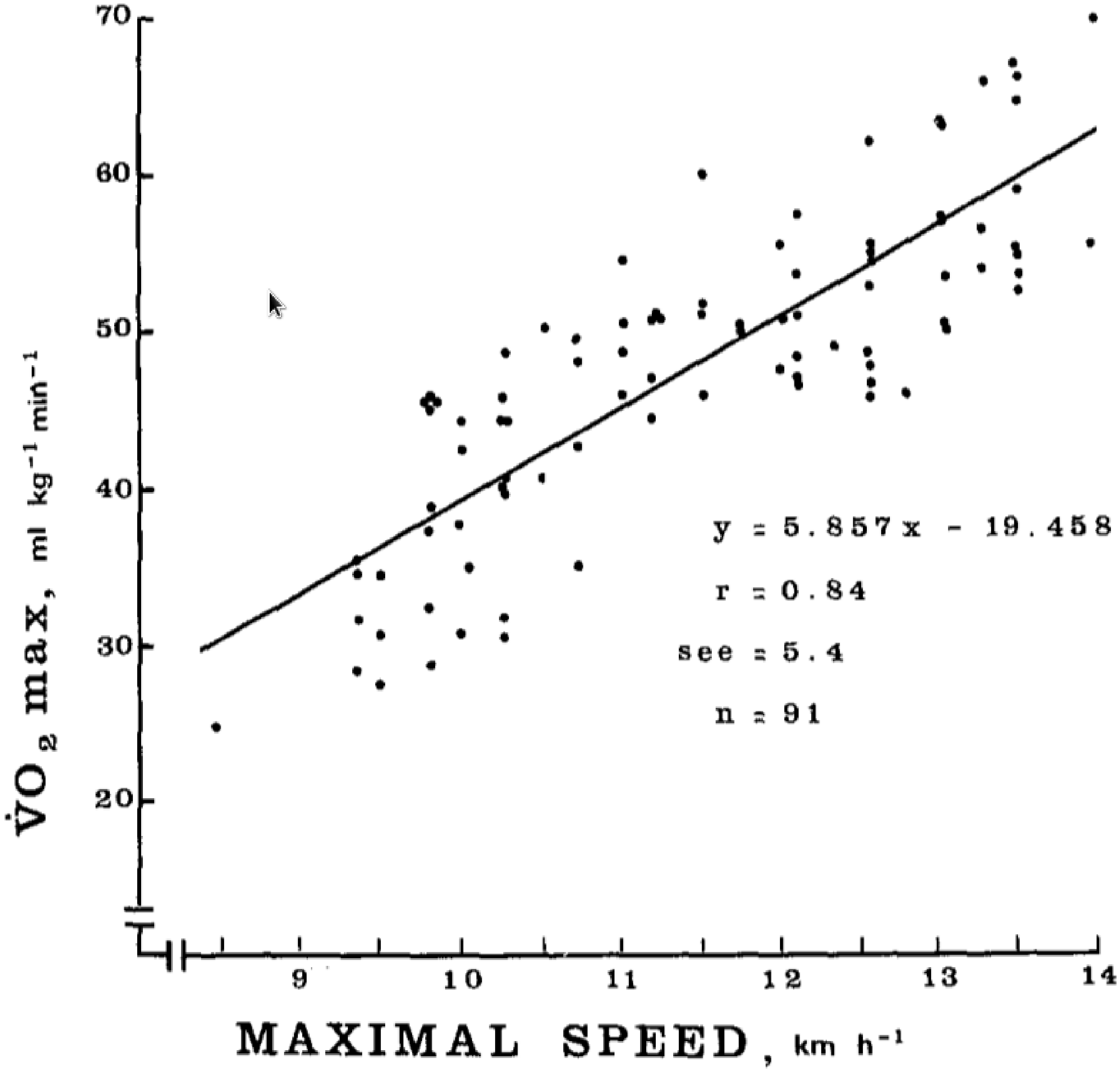


Fig. 2. $\dot{V}O_2$ max as a function of the maximal speed achieved in the 20-m shuttle run test for a total sample of 91 adult subjects. Each point in this figure represents maximal effort

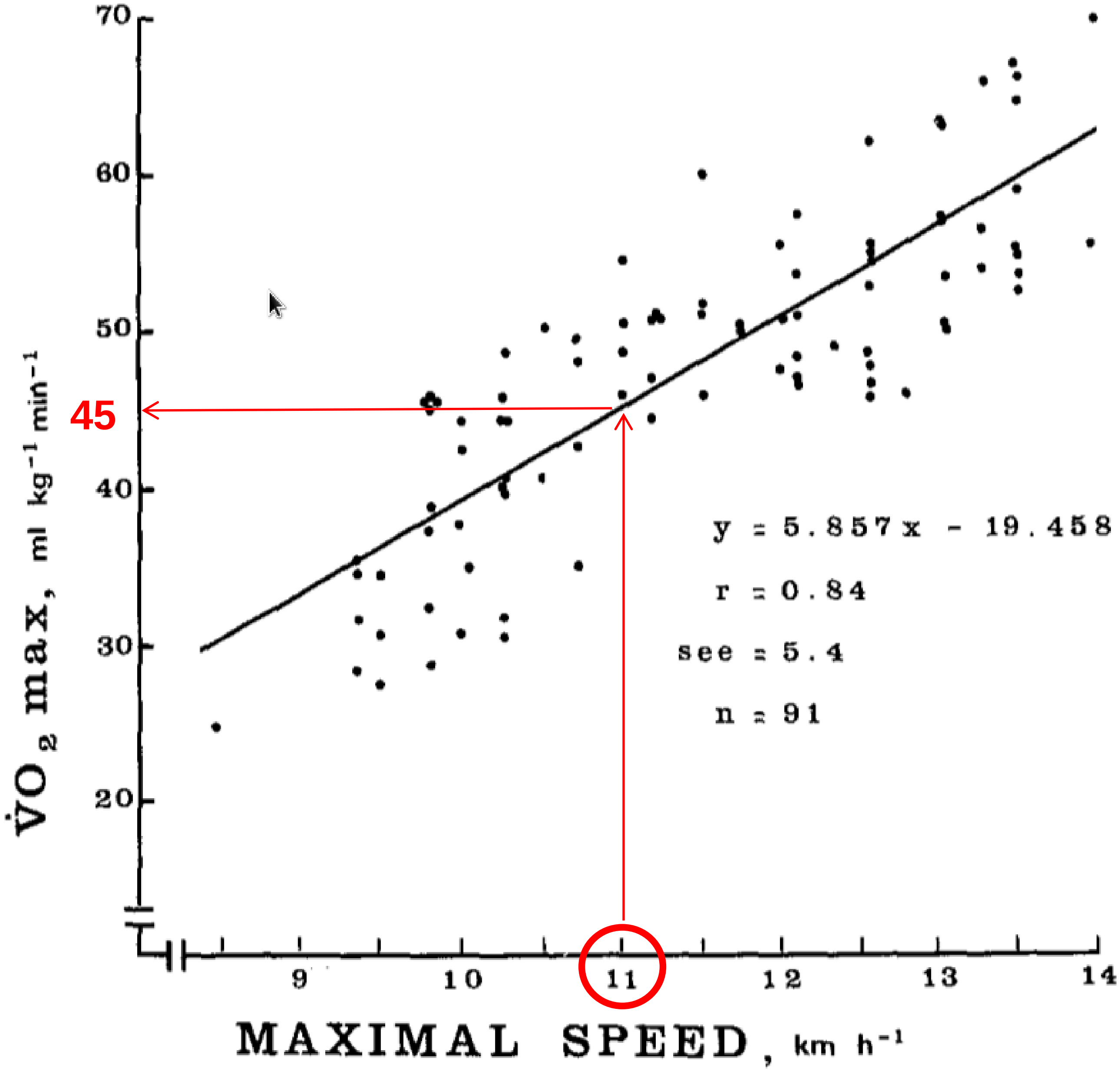


Fig. 2. $\dot{V}O_2$ max as a function of the maximal speed achieved in the 20-m shuttle run test for a total sample of 91 adult subjects. Each point in this figure represents maximal effort

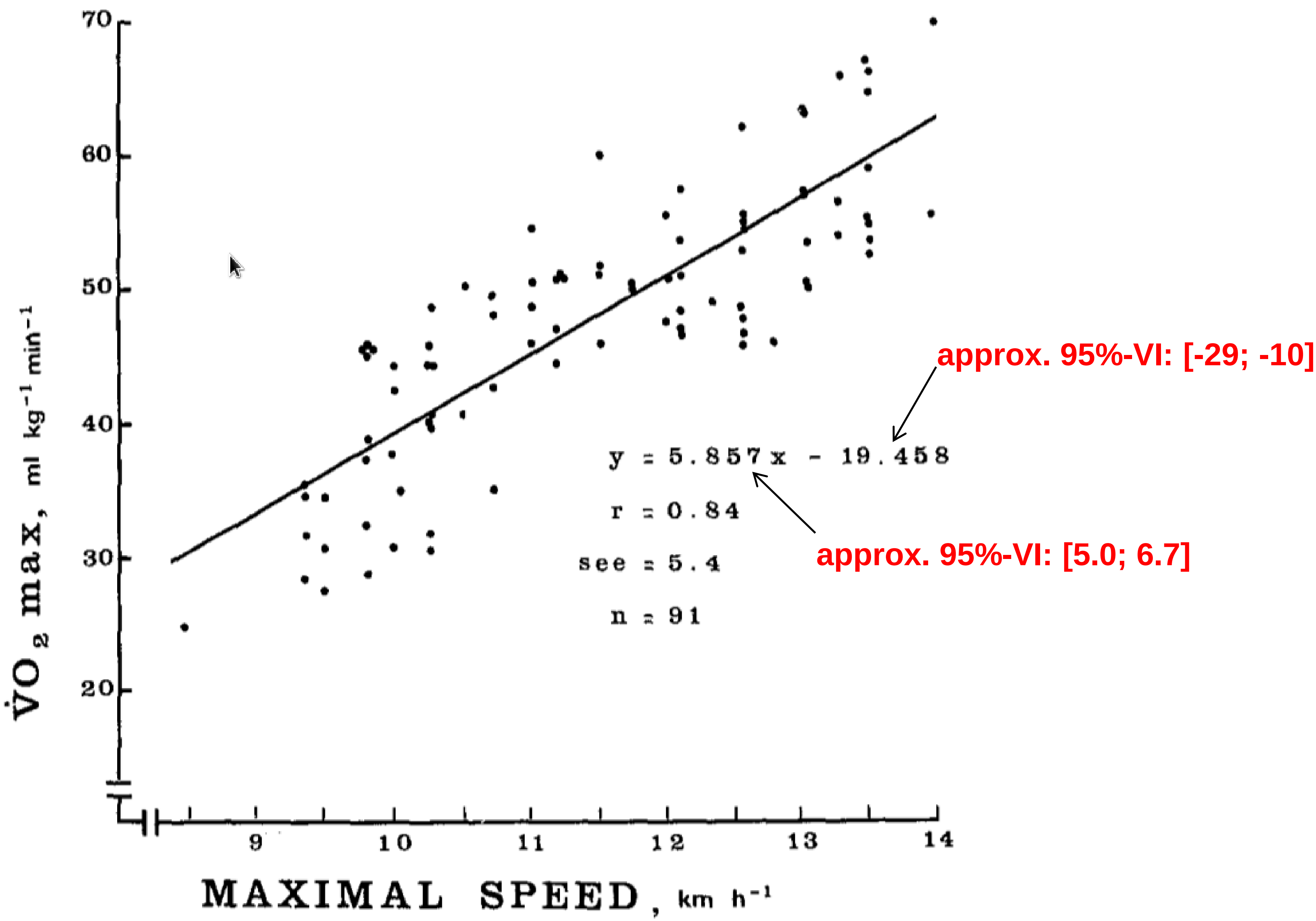


Fig. 2. $\dot{V}O_2$ max as a function of the maximal speed achieved in the 20-m shuttle run test for a total sample of 91 adult subjects. Each point in this figure represents maximal effort

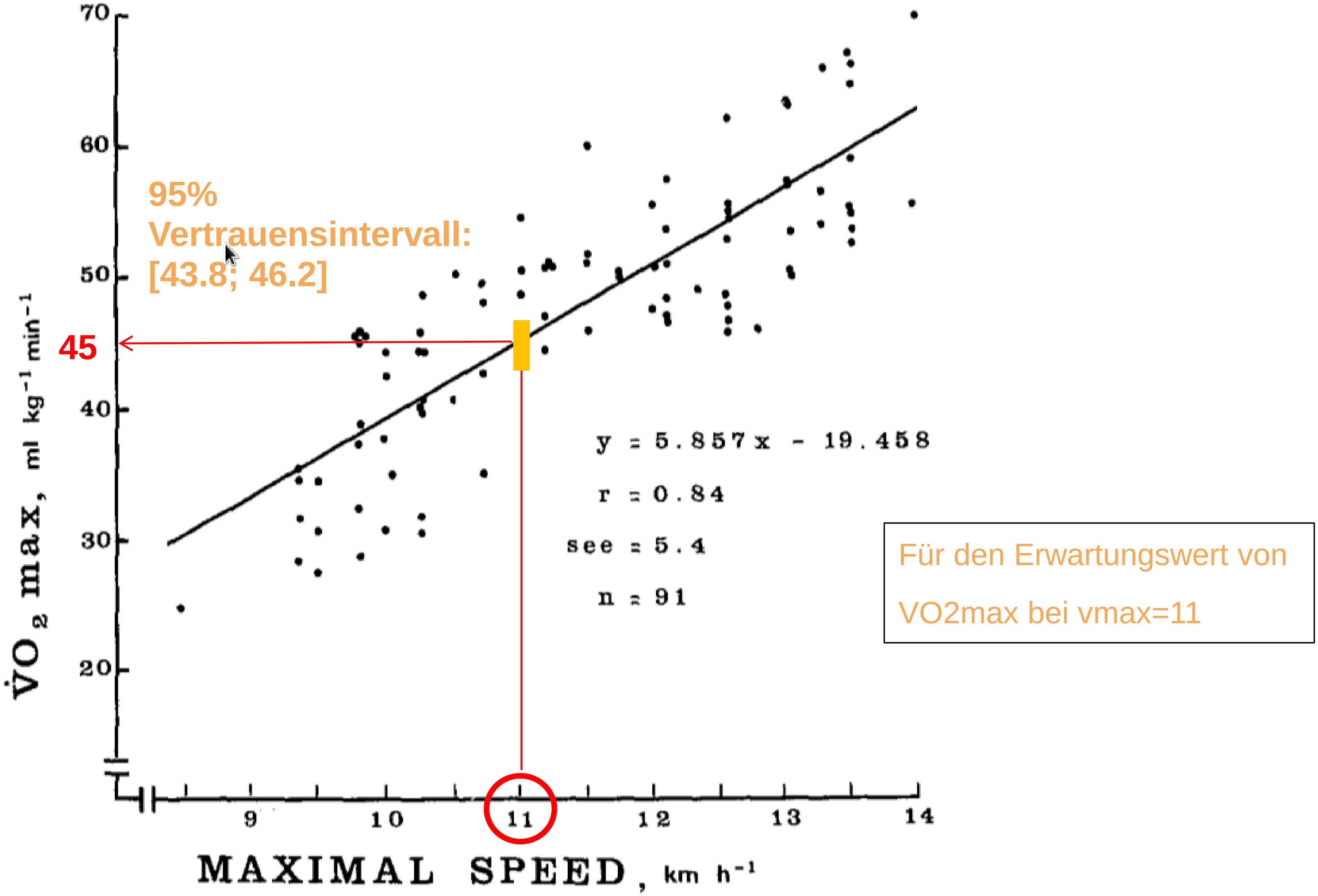


Fig. 2. $\dot{V}O_2$ max as a function of the maximal speed achieved in the 20-m shuttle run test for a total sample of 91 adult subjects. Each point in this figure represents maximal effort

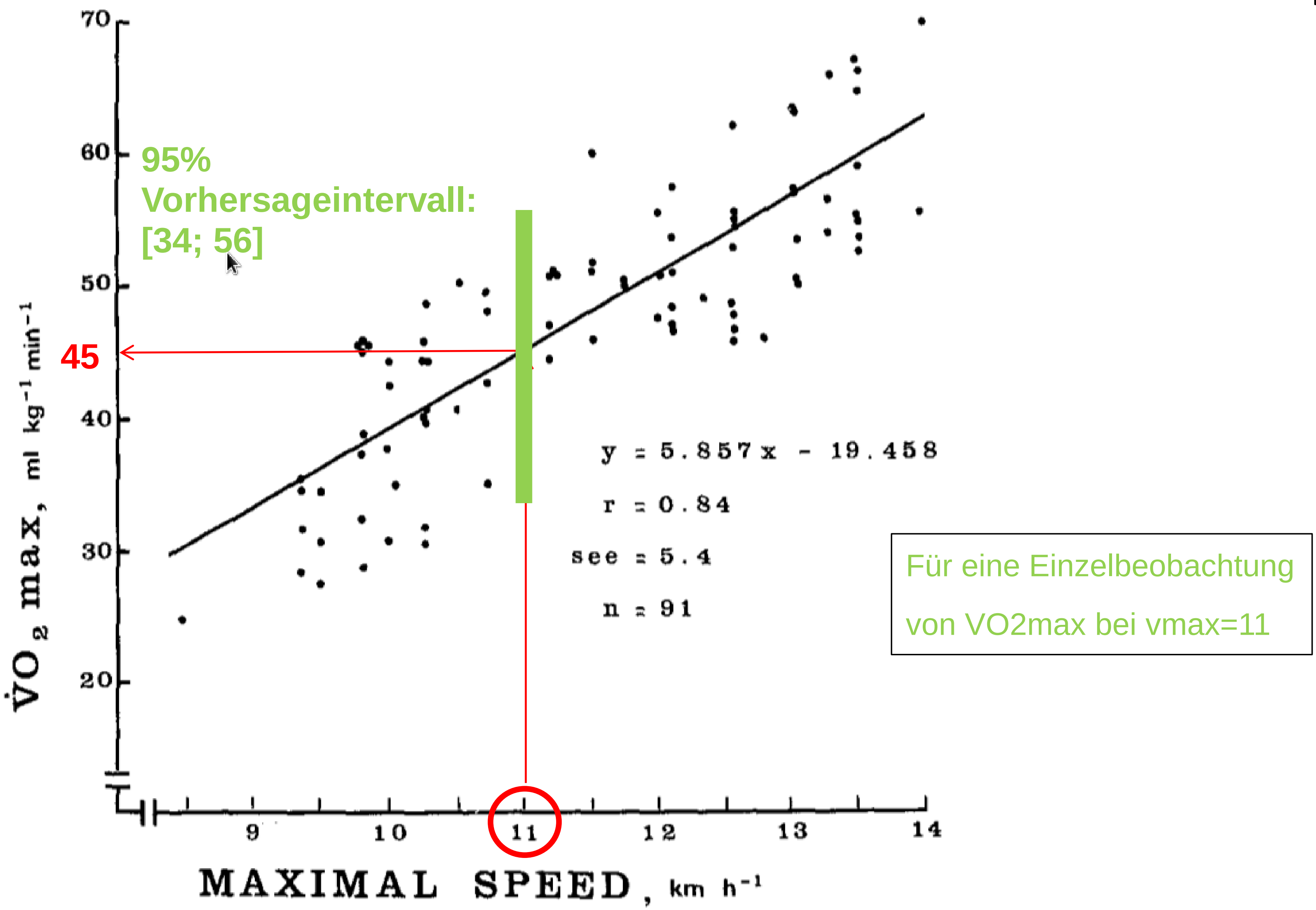
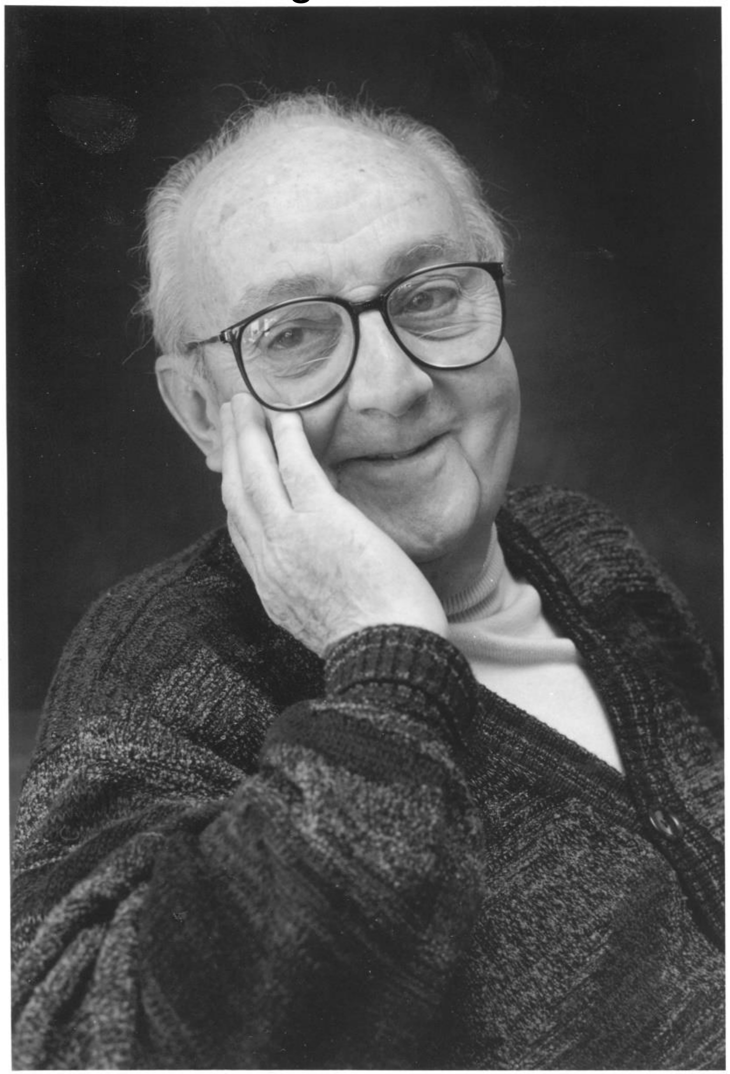


Fig. 2. $\dot{V}O_2$ max as a function of the maximal speed achieved in the 20-m shuttle run test for a total sample of 91 adult subjects. Each point in this figure represents maximal effort

George E.P. Box



“Essentially,
all models are
wrong,
but some are
useful.“

Residuenanalyse: Wie gut stimmt das Modell ?

$$\underline{Y_i = \beta_0 + \beta_1 x_i} + \varepsilon_i \;\; ; \varepsilon_i \sim \underline{N}(0, \underline{\sigma^2}) \quad iid$$

- Form des funktionellen Zusammenhangs
- Varianz der Fehler ist konstant
- Fehler sind normalverteilt

Einfache Regression:

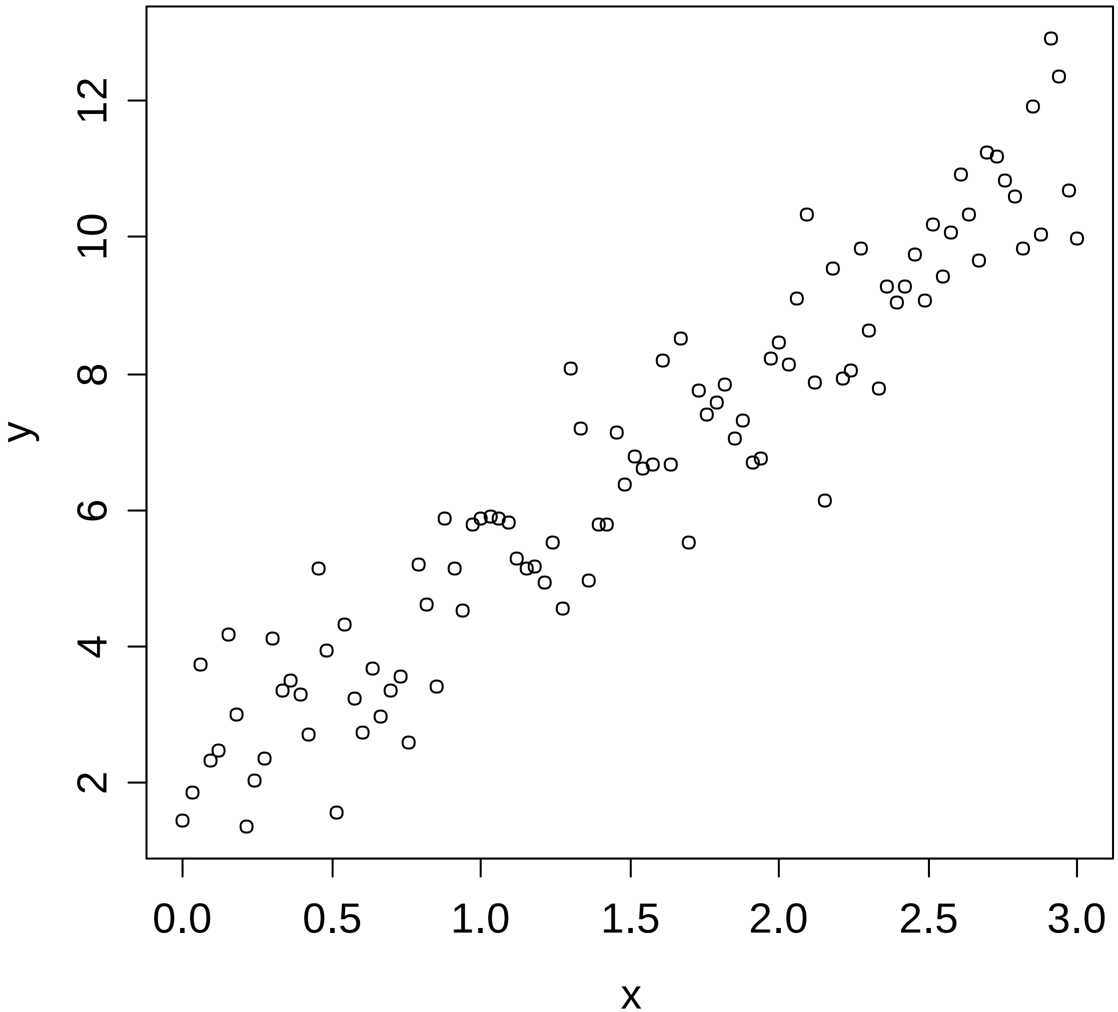
Streudiagramm

Multiple Regression:

Tukey-Anscombe Plot

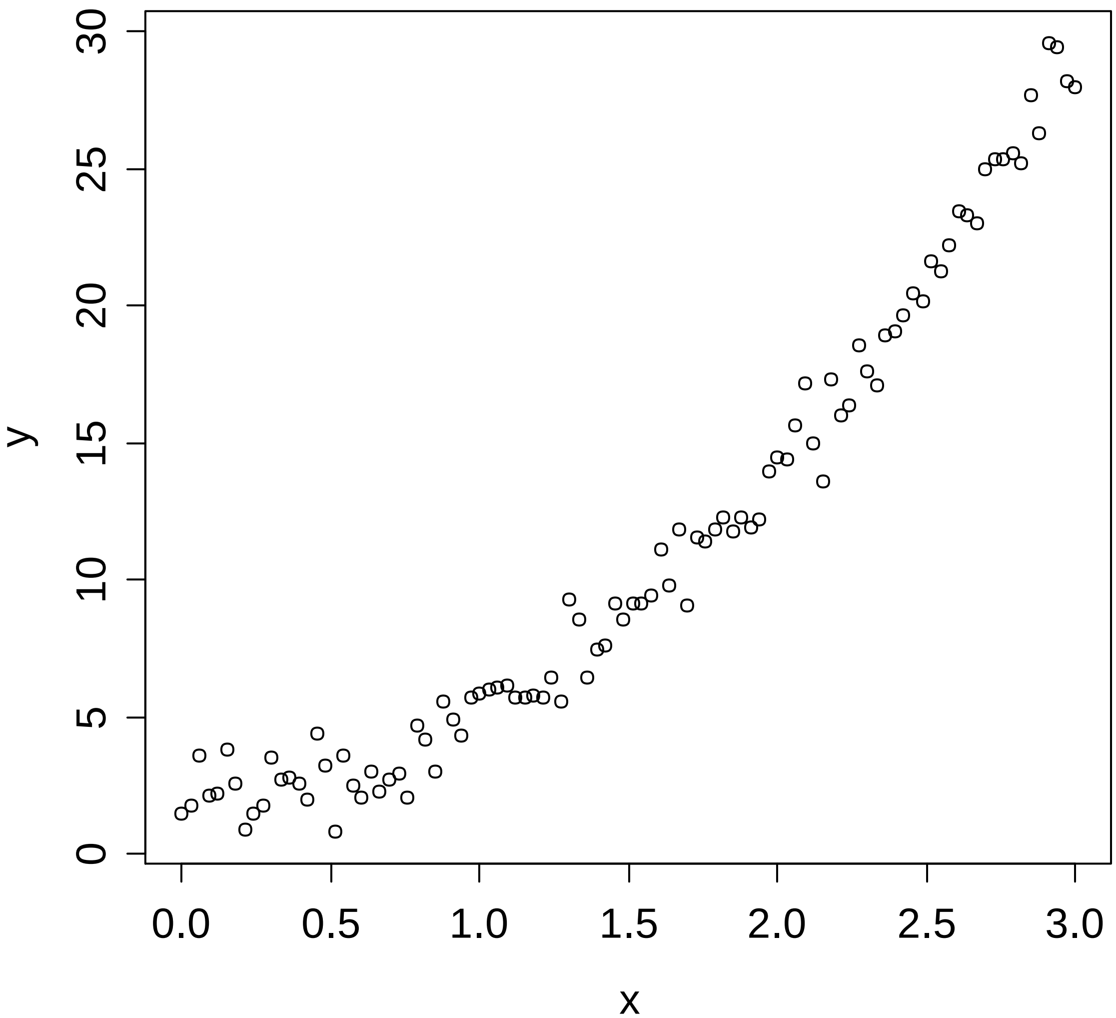
QQ-Plot der Residuen

Streudiagramm bei einfacher linearer Regression



OK

Streudiagramm bei einfacher linearer Regression

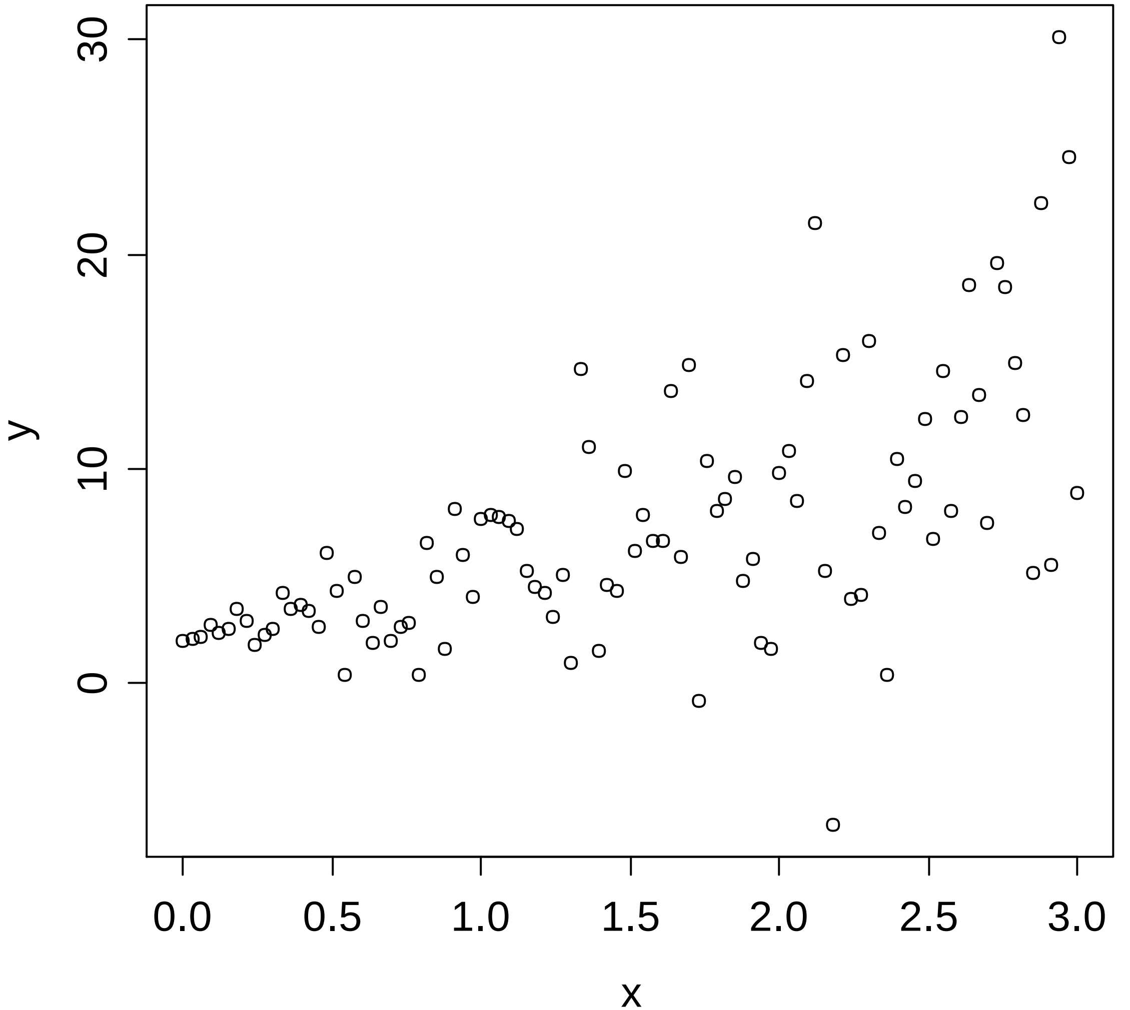


Systematischer Fehler

Krümmung:

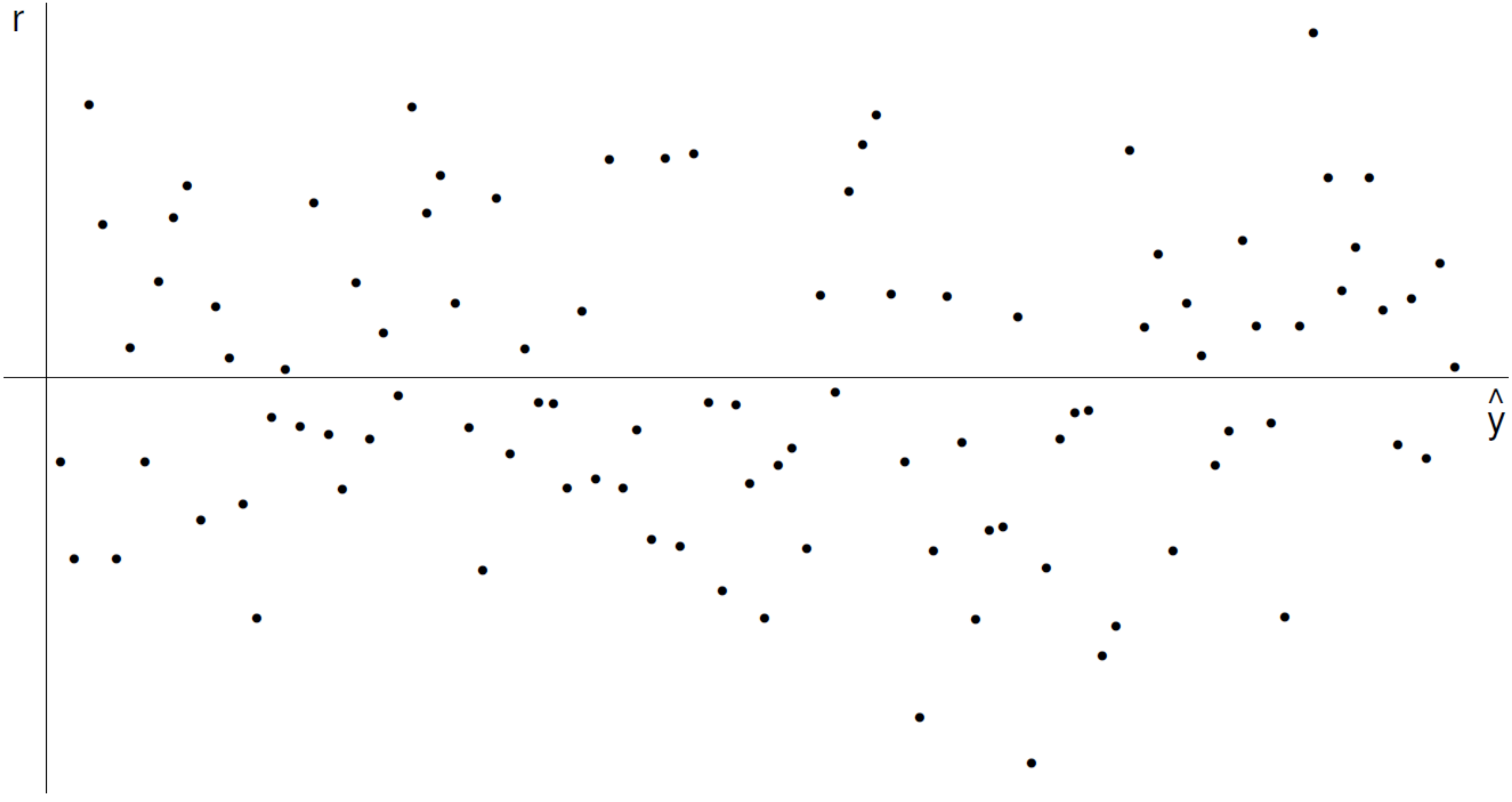
$$y = b_0 + b_1 x + b_2 x^2$$

Streudiagramm bei einfacher linearer Regression

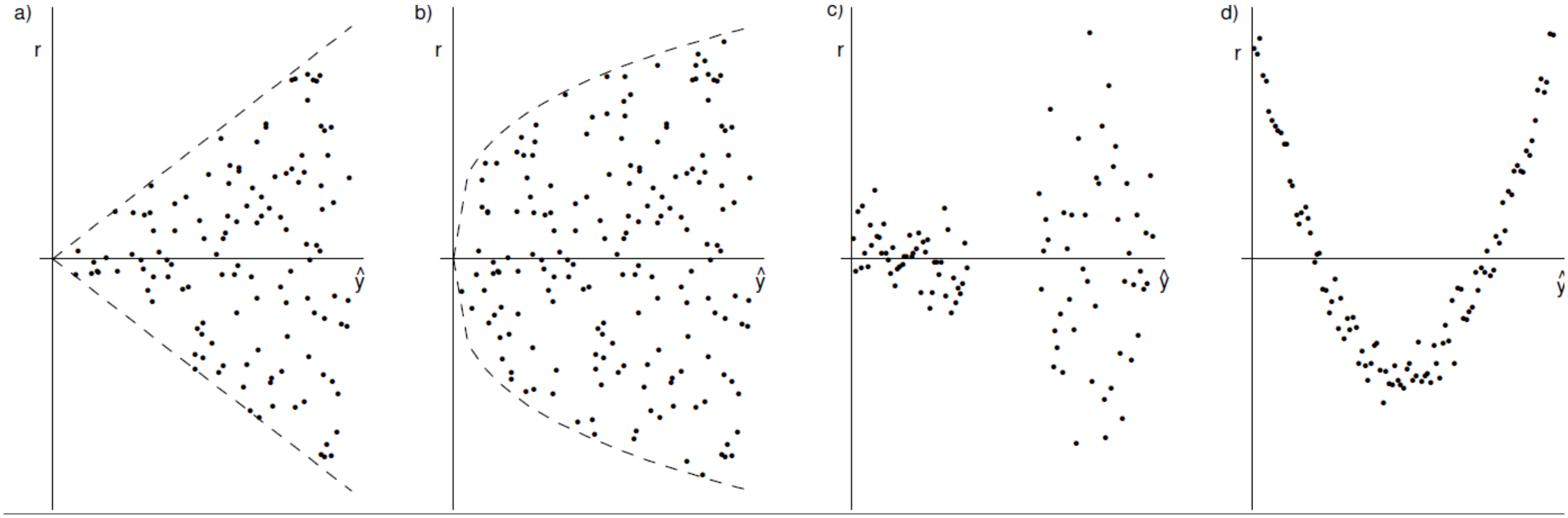


Fehlervarianz nicht konstant

Beispiel für guten Tukey-Anscombe Plot



Beispiele für schlechte Tukey-Anscombe Plots



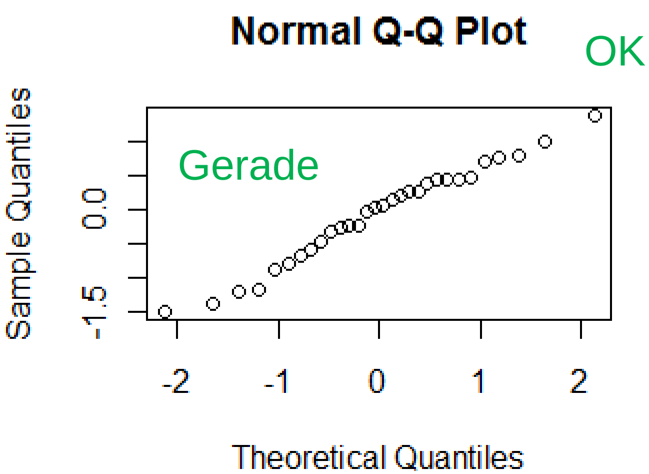
Fehlervarianz nicht konstant

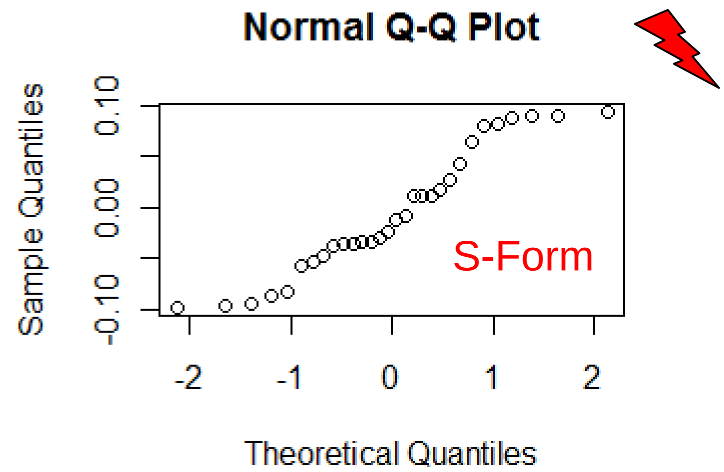
Systematischer Fehler

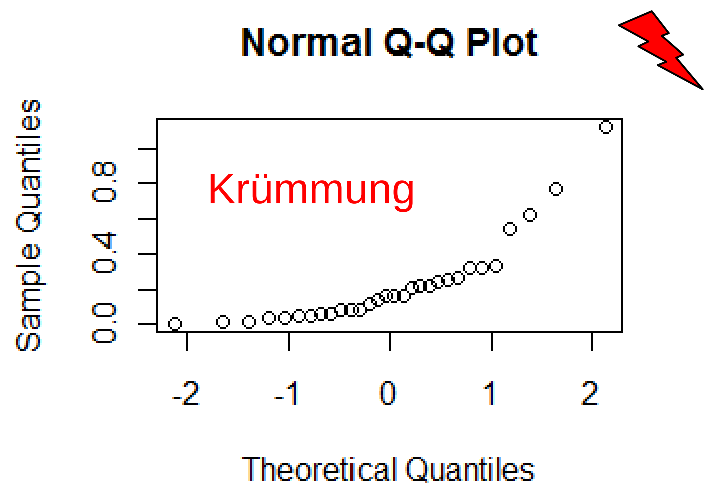
Residuenanalyse: QQ-Plot

Gerade = “gut”

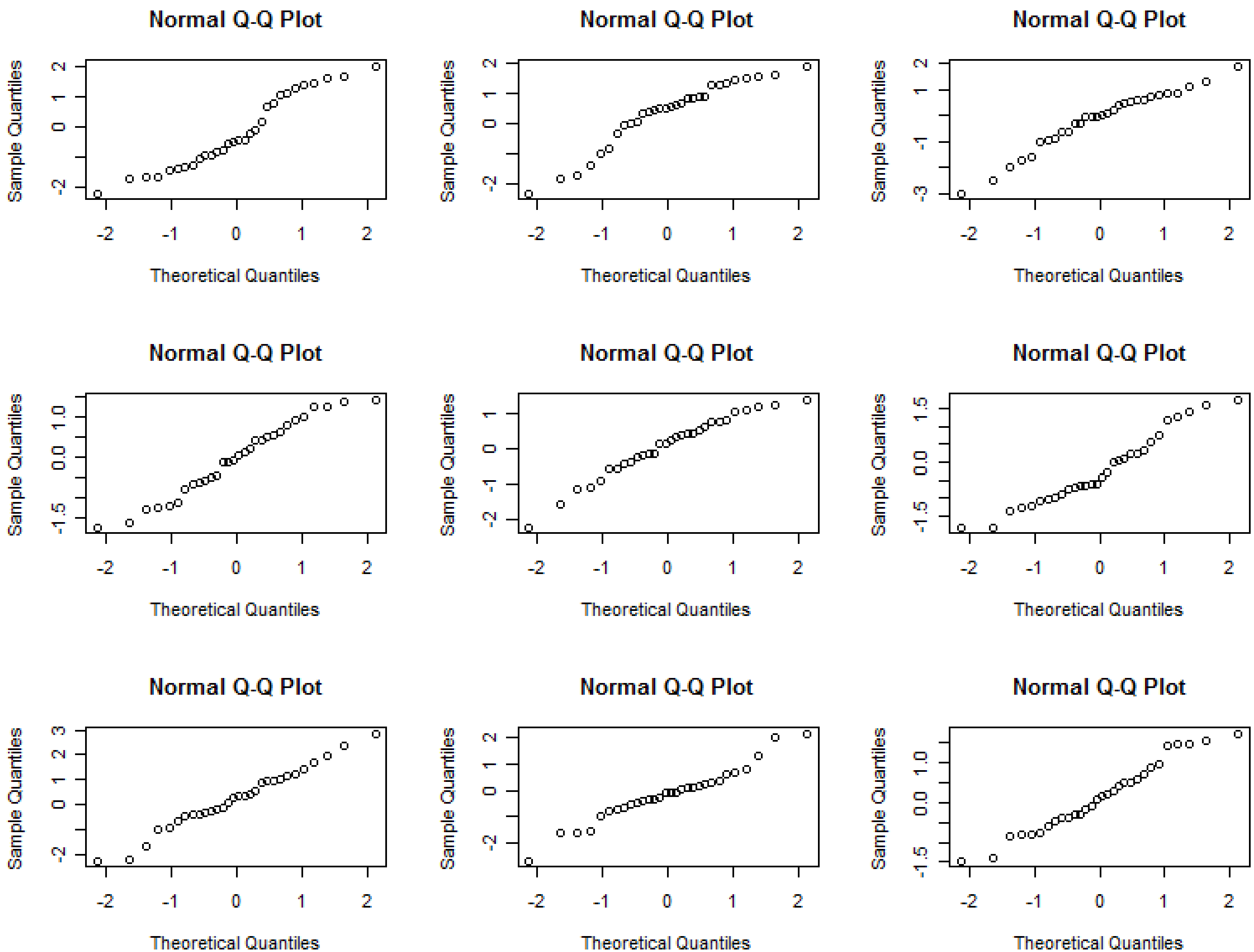
Krümmung = “schlecht”







QQ-Plots: Streuung von “guten” QQ-Plots ($n = 30, R_i \sim N(0,1)$)



Falls Residuenplots schlecht

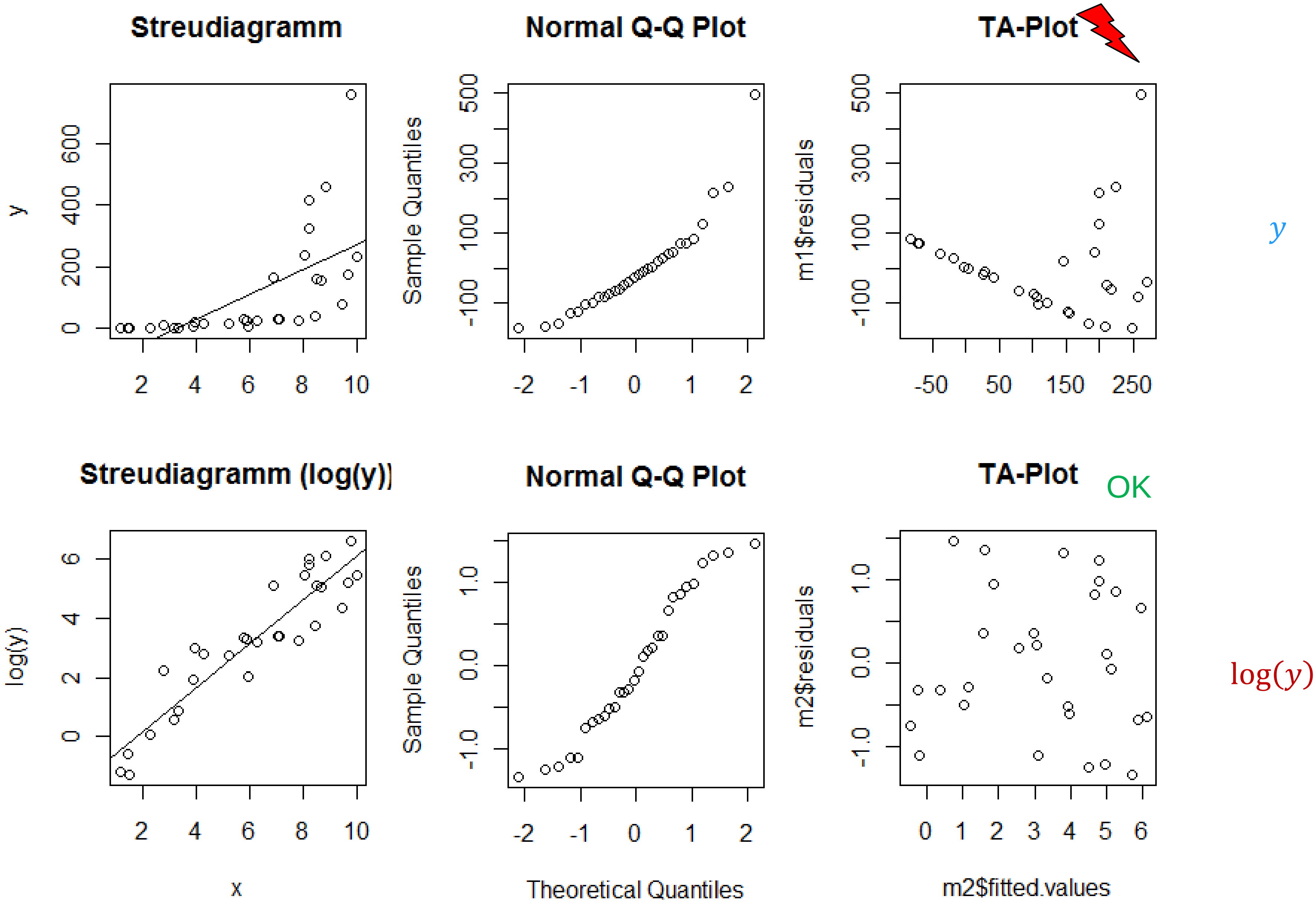
- Oft helfen Transformationen von x oder y
- Achtung: Vorsicht beim Interpretieren der neuen Parameter

- Bsp: $\log(y)$ statt y

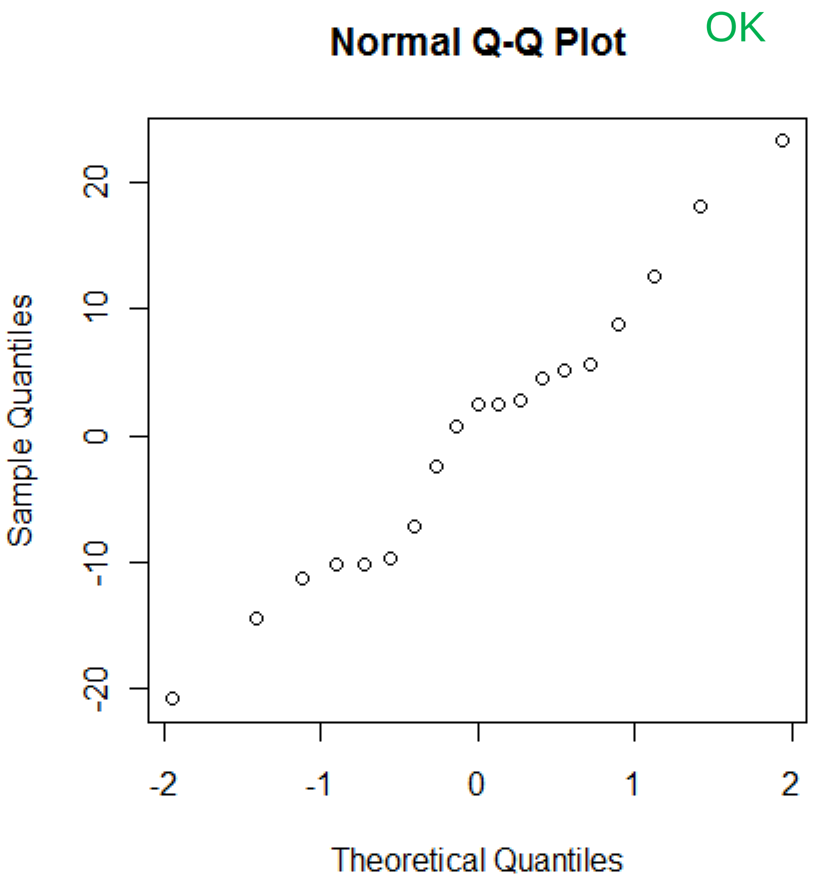
  Vorher: $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$
  Wenn x durch x+1 ersetzt wird, ändert sich Y im Mittel zu $Y + \beta_1$

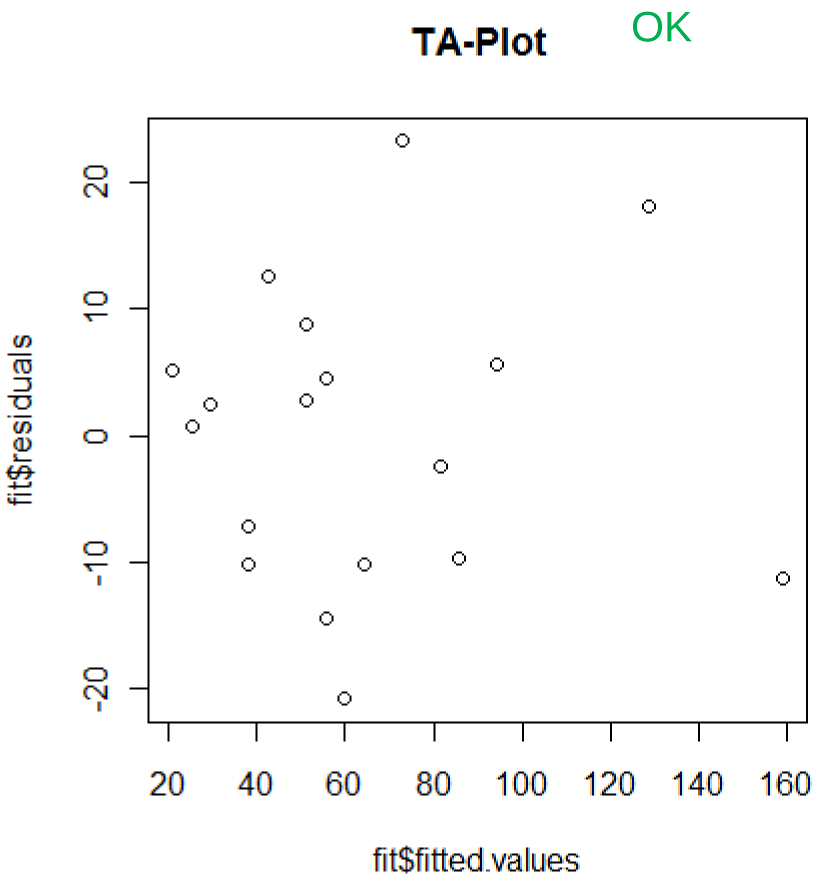
  Nachher:
  $\log(Y_i) = \beta_0 + \beta_1 x_i + \varepsilon_i \leftrightarrow Y_i = \exp(\beta_0 + \beta_1 x_i + \varepsilon_i)$
  Wenn x durch x+1 erstetzt wird, ändert sich Y “im Mittel” zu $Y * \exp(\beta_1)$

Bsp: Ohne Log-Transformation



Residuenanalyse: Supermarkt





Residuenanalyse: Beep-Test