



- Rückenschwimmen

- Rückenschwimmen

- Schmetterling (Delfin)

- Wettkampfbestimmungen
- Strecken
- Einzelwettkampf und Staffelwettkampf
- Ablauf des Wettkampfs

- Entwicklung der Schwimgeschwindigkeit und Vergleich zwischen weiblichen und männlichen Eliteschwimmern über einen längeren Zeitraum

- Die Aufzeichnungen von männlichen und weiblichen Rekorden von 1962 bis 2007 stammen von der Website *USA Swimming*
- Schwimmdisziplinen:
  - 50 m Freistil
  - 100 m Freistil
  - 200 m Freistil
  - 100 m Rückenstil
  - 100 m Bruststil
  - 100 m Schmetterlingstil
- Untersuchungseinheit ist die Durchschnittsgeschwindigkeit
  - ⇒ schnellere Streckenzeit resultiert also in einer höheren Durchschnittsgeschwindigkeit

- Betrachtung der Entwicklung der männlichen und weiblichen Durchschnittsgeschwindigkeit einzeln und im Vergleich zueinander
- die abhängige Variable  $Y$  „Geschwindigkeit“ wird auf die unabhängige Variable  $X$  „Jahreszahl“ regressiert
- für geschlechtsbedingte Unterschiede bezüglich der Durchschnittsgeschwindigkeit wird eine weitere Regressionsanalyse mit der Differenz aus männlichen und weiblichen Durchschnittsgeschwindigkeiten als neue abhängige Variable durchgeführt

- Regressionsanalyse

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi} + \varepsilon_i \quad i = 1, \dots, n$$

$$E(\varepsilon_i) = 0$$

$$V(\varepsilon_i) = \sigma^2$$

$\{\varepsilon_i \mid i = 1, \dots, n\}$  stoch. unabhängig

$$\varepsilon_i \sim N(0, \sigma^2)$$

$y_i$  : Zielgröße (Zufallsgröße), abhängige Variable  
 $x_{1i}, \dots, x_{pi}$  : feste bekannte Einflussgrößen, unabhängige Variablen  
 $\varepsilon_i$  : Zufallsfehler mit unbekannter Varianz  $\sigma^2$   
 $\beta_0, \beta_1, \dots, \beta_p$  : Regressionsparameter  
 $n$  : Anzahl der Beobachtungen

- Schätzung der Regressionsparameter  $\beta_k$  mit Kleinsten-Quadrate-Methode (KQ-Methode)

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

- Modellgüte mittels:

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad R^2 \in [0; 1]$$

- Überprüfung, ob die Regression den Sachverhalt sinnvoll modelliert, mittels *Overall-F-Test*:

$$H_0 : \beta_1 = \dots = \beta_p = 0 \quad \text{vs.} \quad H_1 : \beta_k \neq 0 \text{ für mindestens ein } k$$

$$F = \frac{n - p - 1}{p} \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \sim F_{p, n-p-1}$$

Dabei wird  $H_0$  verworfen, falls gilt:

$$F > F_{1-\alpha, p, n-p-1}$$

- Auswahl des besten Modells mit:
  - $\Delta R^2 = R_2^2 - R_1^2$ , wobei  $R_i^2$  das Bestimmtheitsmaß von Modell (i) darstellt
  - Entscheidung für das kompliziertere Modell bei  $\Delta R^2 \geq 0.02$

- Verwendete Regressionsfunktionen:

- lineare Einfachregression

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i; \quad i = 1, \dots, n \quad (1)$$

- Polynom zweiten Grades als Regressionsfunktion

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i; \quad i = 1, \dots, n \quad (2)$$

- Polynom dritten Grades als Regressionsfunktion

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \varepsilon_i; \quad i = 1, \dots, n \quad (3)$$

- Geschlechtsspezifischer Vergleich

- Verwendung von Regressionsmodellen der Form (1) - (3) mit neuer abhängiger Variable  $\check{Y}$  als Differenz der männlichen Durchschnittsgeschwindigkeiten  $Y^m$  und der weiblichen Durchschnittsgeschwindigkeiten  $Y^w$

- $\check{y}_i := y_i^m - y_i^w, i = 1, \dots, n$

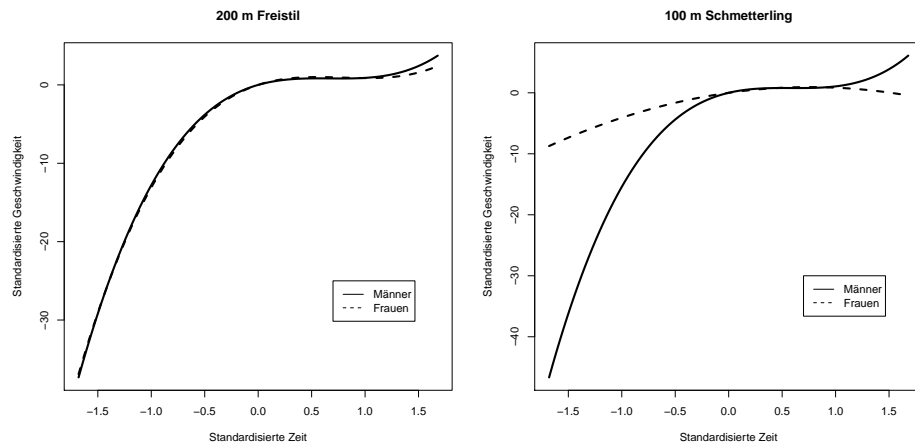
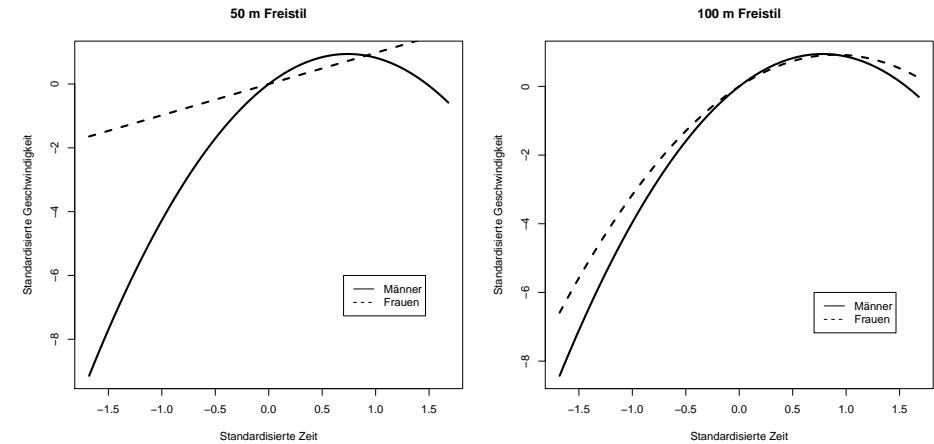
- Modellentscheidung analog zu den Modellen (1) - (3) mittels  $\Delta R^2$

Disziplin	Standardisierte Regressionskoeffizienten			F-Test	$R^2$	$\Delta R^2$
	$\beta_1$	$\beta_2$	$\beta_3$			
<b>50 m Freistil:</b>						
Männer	2.548	-1.723		$F(2, 25) = 198.64^*$	0.94	0.17
Frauen	0.978			$F(1, 26) = 573.83^*$	0.96	
<b>100 m Freistil:</b>						
Männer	2.417	-1.547		$F(2, 25) = 1391.82^*$	0.99	0.14
Frauen	2.035	-1.123		$F(2, 43) = 710.08^*$	0.97	0.08
<b>200 m Freistil:</b>						
Männer	3.895	-5.950	2.949	$F(3, 42) = 1391.82^*$	0.99	0.03
Frauen	4.392	-6.111	2.589	$F(3, 42) = 687.91^*$	0.98	0.03
<b>100 m Rücken:</b>						
Männer	2.087	-1.179		$F(2, 43) = 805.38^*$	0.97	0.03
Frauen	1.983	-1.061		$F(2, 43) = 981.38^*$	0.98	0.07
<b>100 m Brust:</b>						
Männer	1.810	-0.881		$F(2, 43) = 511.78^*$	0.96	0.05
Frauen	2.036	-1.138		$F(2, 43) = 389.59^*$	0.95	0.06
<b>100 m Schmetterling</b>						
Männer	4.141	-7.187	4.096	$F(3, 42) = 468.87^*$	0.97	0.07
Frauen	2.465	-1.626		$F(2, 43) = 383.85^*$	0.95	0.16

\*  $p \leq 0.01$

Tabelle: Regressionskoeffizienten, F-Test,  $R^2$ ,  $\Delta R^2$  und  $p$ -Werte für den Verlauf der 12 amerikanischen Schwimmrekorde (Geschlecht · 6 Disziplinen) von 1962 bis 2007.

Vergleich der standardisierten Regressionsfunktionen beider Geschlechter für die Disziplinen 50, 100 und 200 m Freistil sowie 100 m Schmetterling

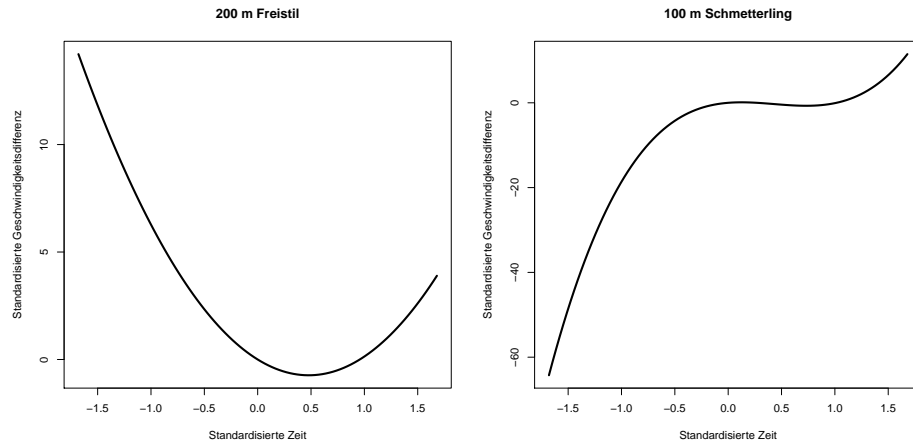


- Geschlechtlicher Vergleich der Entwicklung amerikanischer Schwimmrekorde

Disziplin	Standardisierte Regressionskoeffizienten			F-Test	$R^2$	$\Delta R^2$	$p$ -Wert
	$\beta_1$	$\beta_2$	$\beta_3$				
<b>50 m Freistil:</b>	4.668	-8.560	3.619	$F(3, 24) = 11.69$	0.59	0.05	$\leq 0.01$
<b>100 m Freistil:</b>	0.745	1.832	-2.720	$F(3, 42) = 9.49$	0.40	0.03	$\leq 0.01$
<b>200 m Freistil:</b>	-3.069	3.207		$F(2, 43) = 34.79$	0.62	0.62	$\leq 0.01$
<b>100 m Rücken:</b>	1.062	-0.947		$F(2, 43) = 1.73$	0.07	0.05	0.19
<b>100 m Brust:</b>	-1.114	0.894		$F(2, 43) = 2.62$	0.11	0.05	0.08
<b>100 m Schmett.:</b>	1.985	-9.351	7.284	$F(3, 42) = 26.91$	0.66	0.23	$\leq 0.01$

Tabelle: Regressionskoeffizienten, F-Test,  $R^2$ ,  $\Delta R^2$  und  $p$ -Werte für die Differenzen der Geschwindigkeitsrekorde zwischen Männern und Frauen für den Zeitraum 1962 bis 2007.

Standardisierte Regressionsfunktionen der Geschwindigkeitsdifferenzen von Mann und Frau für die Disziplinen 200 m Freistil sowie 100 m Schmetterling



- Bestzeiten männlicher und weiblicher heranwachsender Eliteschwimmer
- Die Aufzeichnungen aus den Jahren 1962, 1969, 1982, 1992 und 2002 stammen von der Website *USA Swimming*
- Schwimmdisziplinen:
  - 50 m Freistil
  - 100 m Freistil
  - 200 m Freistil
  - 100 m Rückenstil
  - 100 m Bruststil
  - 100 m Schmetterlingstil
- Betrachtung von fünf Altersgruppen
  - 10 Jahre & jünger
  - 11-12 Jahre
  - 13-14 Jahre
  - 15-16 Jahre
  - 17-18 Jahre

- Betrachtung der Entwicklung der männlichen und weiblichen Durchschnittsgeschwindigkeit einzeln und im Vergleich zueinander
- die abhängige Variable Y „Geschwindigkeit“ wird auf die unabhängige Variable X „Jahreszahl“ regressiert
- für geschlechtsbedingte Unterschiede bezüglich der Durchschnittsgeschwindigkeit wird eine weitere Regressionsanalyse mit der Differenz aus männlichen und weiblichen Durchschnittsgeschwindigkeiten als neue abhängige Variable durchgeführt
- Analyse der Stärke von geschlechtsspezifischen Effekten

- analog zur 1. Studie werden die Regressionsmodelle (1) - (3) für die allgemeine und zwischengeschlechtliche Geschwindigkeitsentwicklung verwendet
- Modellentscheidung basiert wieder auf den Differenzen  $\Delta R^2$
- Untersuchung geschlechtsspezifischer Effekte mittels *Cohens d*

$$d = \frac{\bar{y}_m - \bar{y}_w}{s_m}$$

mit den Hypothesen

$$H_0 : \mu_m = \mu_w \quad \text{vs.} \quad H_1 : \mu_m \neq \mu_w$$

- Regressionsanalyse
  - Polynome mindestens zweiten Grades
  - bester Fit bei 52 der 60 Kombinationen mit einem Polynom dritten Grades
  - Werte des Bestimmtheitsmaßes  $R^2$  jetzt nur noch zwischen 0.141 und 0.744
  - Overall-F-Tests erzielten in 57 Fällen ein signifikantes Ergebnis
- Geschlechtlicher Vergleich
  - lineares Regressionsmodell nur in drei von 30 Fällen angemessen
  - ansonsten Polynome zweiten oder dritten Grades

- keine eindeutige Interpretation der Geschwindigkeitsentwicklung mittels der Regressionsanalyse möglich
  - **weder:** männliche Rekorde werden immer über dem der Frauen liegen
  - **noch:** geschlechtsspezifische Leistungsunterschiede werden zukünftig verschwinden
- Effektgrößenanalyse ermöglicht bessere Interpretation
  - normierte männliche Durchschnittsgeschwindigkeiten sind höher als die der Frauen
  - Verstärkung der Unterschiede mit wachsendem Alter der Schwimmer
  - Hypothese  $\mu_m = \mu_w$  wird in fast allen Fällen signifikant verworfen  
 $\Rightarrow \mu_m \neq \mu_w$

- Geschlechtsspezifische Effekte

Disziplin	Altersgruppen				
	10 & jünger	11-12	13-14	15-16	17-18
50 m Freistil:	0.63**	2.44**	5.97**	5.33**	8.14**
100 m Freistil:	1.00**	2.00**	5.08**	5.51**	7.36**
200 m Freistil:	0.81**	1.64**	3.60**	6.49**	6.21**
100 m Rücken:	0.39	1.30**	3.29**	4.81**	5.36**
100 m Brust:	0.18	1.14**	2.37**	3.94**	5.01**
100 m Schmetterling:	0.60**	2.08**	3.09**	4.99**	5.22**

**Tabelle:** Differenzen der Effektgrößen und zugehörige  $p$ -Werte (\* für  $p < 0.05$ , \*\* für  $p < 0.01$ ) für den Geschlechtsvergleich bei amerikanischen Elitejugendschwimmern nach Alter, Schwimmdisziplin und Bahnlänge für den Zeitraum 1962 bis 2007.

- Datensatzwahl der ersten Studie ist nicht gut für den Sachverhalt geeignet
- Voraussetzungen der Regressionsanalyse sind verletzt
- einfache Regressionsfunktionen sind nicht ausreichend
- Auswirkung von Trainer, Herkunft, psychischer Verfassung, sozi-kultureller Umwelt sowie physiologische oder biomechanische Faktoren werden nicht berücksichtigt

- Extremwerttheorie
- Flexiblere Regressionsmodelle z.B. allgemeine *additive* Form:

$$y_i = f_1(z_{i1}) + \dots + f_q(z_{iq}) + \beta_0 + \beta_1 x_{1i} \dots + \beta_p x_{pi} + \varepsilon_i,$$

- Analyse der Geschwindigkeit in Anbetracht neuer Regeln, Bekleidung oder Technik

Vielen Dank!