

# Ultimate 100 M World Records Through Extreme-Value Theory

Statistik im Sport

Florian Wickler

Institut für Statistik  
Ludwig-Maximilians-Universität München

05. Dezember 2009

## Gliederung

- 1 Einführung
  - Problemstellung
  - Der 100-Meter-Lauf
  - Die Daten
- 2 Extremwerttheorie
  - Grundlagen
  - Die Bestimmung der Schätzer
- 3 Berechnung des Weltrekords
  - Überprüfen der Extremwertbedingung
  - Der Vorabschätzer  $\hat{\gamma}_{ini}$
  - Bestimmung des Extremwertindex über  $\overline{AMSE}$
  - Berechnung des absoluten Weltrekords
- 4 Kritik und Vergleich

## Verbesserung des Weltrekords im Laufe der Geschichte

- erste registrierte Zeit: 11.0 s über 110 Yards (100.58 m), 1867
- bis 1887 wurde im Stehen gestartet
- seit 1937 sind Startblöcke erlaubt
- Bei über 2 m/s Rückenwind gibt es keine offiziellen Zeiten
- seit 1977 werden nur elektronisch gemessene Zeiten anerkannt
- Gras- und Aschenbahnen wurden durch Kunststoffbeläge ersetzt
- nur ein Olympiasieger seit 1988 wird nicht des Dopings verdächtigt
- aktueller Weltrekord: 9.58 s, 2009

## Erstellen der Datenliste

### Aufbau

- 1 Nur Spitzenathleten werden betrachtet
- 2 Beobachtungszeitraum: 01. Januar 1991 bis am 19. Juni 2008, da erst seit 1990 Dopingkontrollen im Wettkampf
- 3 von jedem Athleten wird nur die Bestzeit aufgenommen, um Unabhängigkeit zu gewährleisten
- 4 Glättung exakt gleicher Daten zur Auflösung von Clustern
- 5 Transformation der Zeiten in Geschwindigkeiten

# Die Datenliste

## Ergebnis

Disziplin	$n$	Bestzeit	langsamste Zeit
100m Männer	762	9.72s (37.04km/h)	10.30s (34.95km/h)
100m Frauen	479	10.65s (33.80km/h)	11.38s (31.63km/h)

- beide aktuellen Weltrekorde fehlen (F. Griffith-Joyner und U. Bolt)
- wahre Anzahl an „Athleten“ deutlich größer; wie groß spielt keine Rolle

# Ziel

## Ausgangslage

- 1 Zentraler Grenzwertsatz:  
Grenzwertverhalten von partiellen Summen  $X_1 + X_2 + \dots + X_n$
- 2  $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$  Geschwindigkeiten der Spitzenathleten

## Ziel

- Modellierung stochastischer Prozesse von Extremen
- Extremwerttheorie:  
Grenzwertverhalten vom  $\max(X_1, X_2, \dots, X_n)$
- absoluten Weltrekord  $\hat{=}$  rechter Endpunkt einer Verteilung  
 $x^* := \sup\{x \in \mathbb{R} : F(x) < 1\}$

## Extremwertbedingung

$$\lim_{n \rightarrow \infty} P\left(\frac{X_{n:n} - b_n}{a_n} \leq x\right) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G(x)$$

Einzig die Klasse der Extremwertverteilungen erfüllt den Grenzwert.

## Extremwertverteilung

Die Klasse der Extremwertverteilungen wird definiert als:

$$G_\gamma(x) = \begin{cases} \exp(-(1 + \gamma x)^{-1/\gamma}) & \text{mit } 1 + \gamma x > 0, \quad \gamma \in \mathbb{R} \setminus \{0\} \\ \exp(-e^{-x}) & \gamma = 0 \end{cases}$$

## Der Extremwertindex $\gamma$

- Der Gestaltungsparameter ( $\gamma$ ) der Extremwertverteilung wird Extremwertindex genannt
- mit Hilfe von  $\gamma$  kann die Verteilung exakt bestimmt werden  
→ der absolute Weltrekord  $x^*$  kann geschätzt werden
- der Weltrekord ist endlich (unendlich) falls  $\gamma < 0$  ( $\gamma > 0$ )

### Schätzung des Extremwertindex $\gamma$

$$\hat{\gamma} = M_n^{(1)} + 1 - \frac{1}{2} \left( 1 - \frac{(M_n^{(1)})^2}{M_n^{(2)}} \right)^{-1}$$

### Schätzung des absoluten Weltrekords

$$\hat{x}^* = \hat{b}_{n/k} - \frac{\hat{a}_{n/k}}{\hat{\gamma}}$$

### Schätzung der Folgen $a_{n/k}$ und $b_{n/k}$

$$\hat{a}_{n/k} = (1 - \min\{0, \hat{\gamma}\}) X_{n-k:n} M_n^{(1)}$$

$$\hat{b}_{n/k} = X_{n-k:n}$$

### Momentenschätzer

$$M_n^{(r)}(k) = \frac{1}{k} \sum_{i=0}^{k-1} (\log X_{n-i} - \log X_{n-k})^r \quad \text{für } r = 1, 2$$

### Beispiel

- Berechnung des ersten Moments
- für  $n = 10$
- und  $k = 4$

$$\diamond \leq \diamond \leq \diamond \leq \diamond \leq \diamond \leq \diamond \leq \diamond \leq \diamond \leq \diamond \leq \diamond \leq \diamond$$

### Der Schwellenwert $k$

Welche Athleten haben einen signifikanten Einfluss auf den absoluten Weltrekord?

- nur das Maximum enthält Informationen über das rechte Ende der Verteilung  
→ eher unrealistisch
- die  $k$  besten Läufer haben einen Einfluss; der Einfluss der Anderen ist vernachlässigbar klein

### die richtige Wahl von $k$

Wie weit muss und wie weit darf man vom Maximum weg?

- Für ein kleines  $k$  hat  $\hat{\gamma}$  eine hohe Varianz
- Für größere Werte von  $k$  nimmt der Bias von  $\hat{\gamma}$  zu

### Minimierung des AMSE zur Bestimmung von $k$

$$AVar(\hat{\gamma}) + (ABias(\hat{\gamma}))^2 = AMSE(\hat{\gamma})$$

### Zirkelschluss

$$AVar(\hat{\gamma}) = \frac{(1-\gamma)^2(1-2\gamma)(1-\gamma+6\gamma^2)}{(1-3\gamma)(1-4\gamma)} \frac{1}{k}$$

$$ABias(\hat{\gamma}) = \frac{1-2\gamma}{\gamma(1-\gamma)} c \left( \frac{n}{k} \right)$$

- $AMSE(\hat{\gamma})$  hängt von  $\gamma$  ab
- $AMSE(\hat{\gamma})$  hängt auch von  $k$  ab

# Die Schätzer

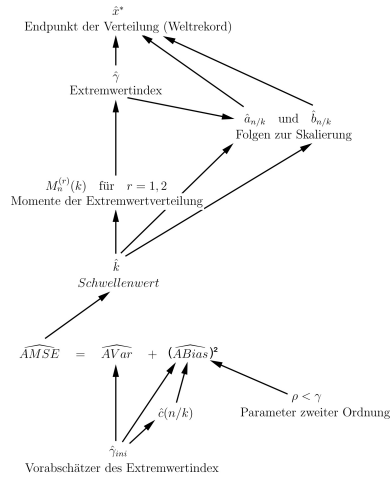
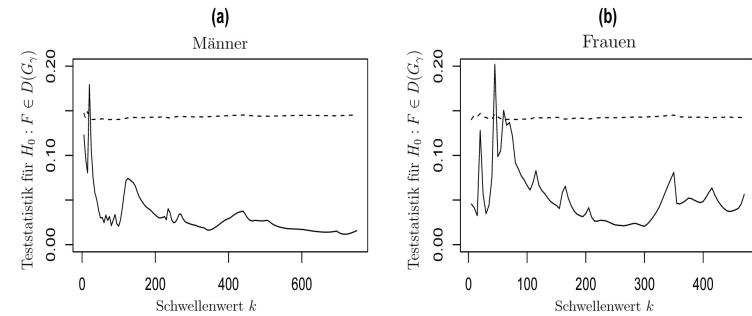
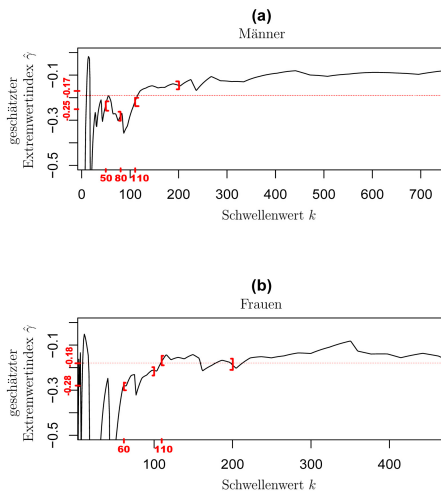


Abbildung: notwendige Schätzer zur Bestimmung des absoluten Weltrekords



Überprüfung der Extremwertbedingung für die 100 Meter der Männer und Frauen. Die durchgezogene Linie zeigt die Teststatistik, die gepunktete das 95%-Quantil.  
→ Für fast alle  $k$  gilt die Extremwertbedingung.

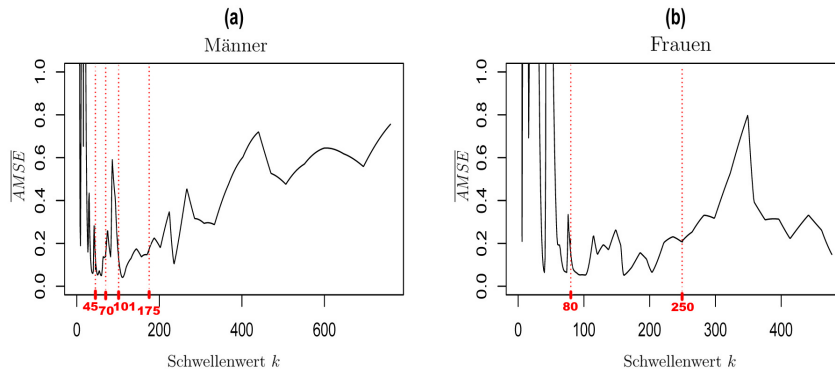


Der Momentenschätzer in Abhängigkeit von  $k$  für den 100-Meter-Lauf der Männer und der Frauen; die horizontale Gerade zeigt das endgültige  $\hat{\gamma}$ .

## Bestimmung von $k$

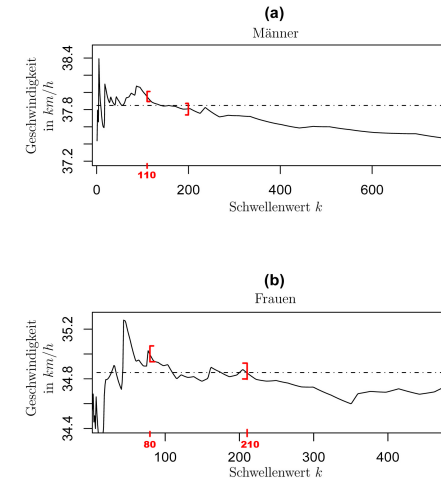
Für jeden Vorabschätzer  $\hat{\gamma}_{ini}$ : Berechnung eines  $\widehat{AMSE}$

- 9 Werte von  $\hat{\gamma}_{ini}$  bei den Männern  
→ 9 Graphiken des  $\widehat{AMSE}$
- 11 Werte von  $\hat{\gamma}_{ini}$  bei den Frauen  
→ 11 Graphiken des  $\widehat{AMSE}$
- grundsätzliche Gestalt der Graphen gleich
- Minimum der  $\widehat{AMSE}$ -Graphen liegt in unterschiedlichen  $k$ -Regionen
- Für jedem Graphen eine  $k$ -Region auswählen, in der der  $\widehat{AMSE}$  klein ist  
→  $\widehat{AMSE}$  bilden



Mittelwert der geschätzten AMSE der Momentenschätzer sowohl für Männer als auch für Frauen.

- für die Männer:  $\hat{\gamma} = -0.19$
- für die Frauen:  $\hat{\gamma} = -0.18$



Der geschätzte Endpunkt  $\hat{x}^*$  bei festgehaltenem  $\hat{\gamma}$  für den 100-Meter-Lauf der Männer und der Frauen; die gepunktete Linie zeigt die abschließende Wahl des Endwerts.

### Endergebnisse

	aktueller WR	absoluter WR	Spielraum
100m Männer	9.58s	9.51s (37.85km/h)	0.7%
100m Frauen	10.49s	10.33s (34.85km/h)	1.5%

### Achtung!

die aktuellen Weltrekorde wurden nicht berücksichtigt

### Kritik an der Methode

- Unabhängigkeit der Daten nicht gegeben
- sehr viele aufeinanderfolgende Schätzungen
- relativ willkürliche Schätzung anhand von Graphen
- aktueller Weltrekord der Männer fehlt

### Vergleich mit ähnlichen Berechnungen

Zeitspanne	Autor	100m Männer	100m Frauen
1991-2008	Einmahl	9.51s	10.33s
1991-2005	Einmahl	9.37s	10.36s
-2005	Einmahl	9.29s	10.11s
1896-2007	Berthelot	9.73s	10.43s
1900-2008	Denny	9.48s	10.39s