



Wahrscheinlichkeiten und statistische Modelle für Pferderennen

Gehalten von Clara Wendelborn

Im Rahmen des Bachelor Seminars Statistik im Sport

Bei Prof. Dr. Friedrich Leisch, Manuel Eugster, Sebastian Kaiser



Inhaltsverzeichnis

- Einleitung
 - Pferderennen
 - Wetten
 - Methodik
 - Diskussion
 - Ausblick
 - Fazit
- ☛ Einleitung
 - ☛ Der Pferderennsport
 - ☛ Das Pferderennen
 - ☛ Die Wetten
 - ☛ Methodik
 - ☛ Der Favorite-Longshot Bias
 - ☛ Platzierungswahrscheinlichkeiten
 - ☛ Diskussion
 - ☛ Ausblick
 - ☛ Fazit



Der Pferderennsport I

- Einleitung
 - Pferderennen
 - Wetten
 - Methodik
 - Diskussion
 - Ausblick
 - Fazit
- ☛ Ursprünglich zur Selektion von Leistungsmerkmalen
 - ☛ Ziel: als erstes das Ziel zu erreichen
 - ☛ Pro Renntag 8-13 Rennen mit je \approx 13 Pferden
 - ☛ Pferde rennen zwischen 60 und 70 km/h
 - ☛ Jockeys dürfen maximal 55kg wiegen
 - ☛ Rennbahnen nicht standadisiert



Der Pferderennsport II

- Einleitung
 - Pferderennen
 - Wetten
 - Methodik
 - Diskussion
 - Ausblick
 - Fazit
- ☛ Diverse Arten von Rennen:
 - ☛ Galopp-, Trab-, Kutschenrennen etc.
 - ☛ Ausschreibung: Ort, Zeit, Distanz, Dotierung und Bestimmung der zur Teilnahme berechtigten Pferde.
 - ☛ Galopprennpferde:
 - Flieger: 1000m-1400m,
 - Meiler: 1600m-2000m
 - Steher: 2200-4800m.





Die Wetten

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

- Pro Renntag Programm zu kaufen mit wichtigen Informationen für Wettende zu Startern
- Einige Wettarten:
 - WIN: das Pferd ist der Sieger des Rennens
 - PLACE: das Pferd muss Erster oder Zweiter werden
 - SHOW: das Pferd muss Erster, Zweiter oder Dritter werden
 - EXACTA: die ersten beiden Plätze müssen in der richtigen Reihenfolge vorhergesagt werden
 - TRIFECTA: die ersten drei Plätze müssen in richtiger Reihenfolge vorhergesagt werden
 - SUPERFECTA: die ersten vier Plätze müssen in richtiger Reihenfolge vorhergesagt werden

18.12.2009

Clara- Astrid Wendelborn

5



Die Quoten I

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

- Quoten über Totalisatorprinzip → alle 20s aktueller Stand
- Auch bei Buchmachern ,tw. feste Quoten
- Minimaler Einsatz 2€
- 75% aller Einsätze stehen zur Gewinnausschüttung zur Verfügung
- Vortages-, Tages-, Startquoten → Endquoten erst nach dem Start bekannt

18.12.2009

Clara- Astrid Wendelborn

6



Die Quoten II

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

- Beispiel ohne Steuern:
 - Insgesamteinsatz: 100000€
 - Bei Gewinn der jeweiligen Pferde:

Pferd	Einsatz	Quote	Gewinn
Step Dancer	20000€	50:10	5-fache des Einsatzes
Alando	30000€	33:10	3,3-fache des Einsatzes
Nymphe	50000€	20:10	Doppelte des Einsatzes

18.12.2009

Clara- Astrid Wendelborn

7



Pferderennen in Hong Kong

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

- Das folgende Paper bezieht sich hauptsächlich auf Daten aus Hong Kong
- Zwei Rennbahnen: Sha Tin und Happy Valley
- Bis zu 13 Rennen mit je ca. acht Pferden pro Renntag
- Einzig zugelassenes Glücksspiel: Pferderennen
- Umsatz an einem Renntag so hoch wie der Jahresumsatz aller deutschen Rennbahnen zusammen
 - 10000000000€
- 81% der Umsätze werden als Gewinne wieder ausgeschüttet
 - pro Renntag 100000000€ Wettumsätze
 - sehr hoher Gewinnpool

18.12.2009

Clara- Astrid Wendelborn

8



Der Favorite- Longshot Bias I

- Einleitung
- Pferderennen
- Wetten
- Methodik
- Diskussion
- Ausblick
- Fazit

- ☞ Übersetzt Favoriten- und Außenseiterverzerrung → nicht adäquat
- ☞ Wetter tendieren dazu Favoriten zu unterwetten, Longshots dagegen zu überwetten

	Step Dancer	Alando	Nymphe
Wahre Gewinnwahrscheinlichkeit	1/9	2/9	2/3
Wettanteil	20%	30%	50%
Gewinnwahrscheinlichkeit nach Wettenden	1/5	3/10	1/2

- ☞ Favorit Nymphe unterwettet; „Longshots“ Alando und Step Dancer überwettet



Der Favorite- Longshot Bias II

- Einleitung
- Pferderennen
- Wetten
- Methodik
- Diskussion
- Ausblick
- Fazit

- ☞ Erklärung: Gewinn höher bei Sieg von Pferden auf die weniger Menschen setzen; logisch auf Favoriten zu setzen
- ☞ In Hong Kong tritt das Phänomen nicht auf → nicht universal
 - ☞ Eventuell wegen der Höhe des dortigen Wettpools
 - ☞ Intensivere Vorbereitung auf Wette
 - ☞ Vorsichtigeres Wetten
- ☞ **Wie berechnet sich der Bias und wo tritt er auf?**



Umsetzung im Paper I

- Einleitung
- Pferderennen
- Wetten
- Methodik
- Diskussion
- Ausblick
- Fazit

- ☞ Risikoparameter β soll geschätzt werden: gibt an, ob Risiko präferiert wird oder nicht
- ☞ Gewinnwahrscheinlichkeit von Pferd $i = \pi_i$

$$= \frac{(\text{Wettanteil von Pferd } i)^\beta}{\sum (\text{Wettanteile aller anderen Pferde})^\beta} = \frac{P_i^\beta}{\sum_{j=1}^n P_j^\beta} \quad (1)$$

☞ Wettanteil $P_i = \frac{1 - t}{1 + O_i}$ also $= \frac{1 - \text{Steuern}}{1 + \text{Gewinnchance}}$

- ☞ Wettanteil bekannt
- ☞ $\beta > 1 \rightarrow$ Risiko wird präferiert
- ☞ $\beta = 1 \rightarrow$ Risiko neutral
- ☞ $\beta < 1 \rightarrow$ Risikoavers



Umsetzung im Paper II

- Einleitung
- Pferderennen
- Wetten
- Methodik
- Diskussion
- Ausblick
- Fazit

Rennen	Geschätztes β	p-Wert für $\beta \neq 1$	Gewinnpool
Atlantic City	1.10	0.08	unbekannt
Meadlowlands	1.12	0.02	52000\$
Saratoga	1.16	~ 0	25000\$
Roosevelt	1.13	~ 0	218000\$
Yonkers	1.13	~ 0	228000\$
Japan	1.07	0.01	168000\$
Happpy Valley	1.04	0.25	1100000\$
Sha Tin	0.94	0.04	1100000\$
Shanghai	1.03	0.38	unbekannt

- ☞ Risikopräferenz in China und Hong Kong neutral → hoher Gewinnpool



Die ökonomische Nutzentheorie I

- Einleitung
- Pferderennen
- Wetten
- Methodik
- Diskussion
- Ausblick
- Fazit

- ☞ Instrument zur Modellierung menschlichen Verhaltens
- ☞ Ziel: individuellen Nutzen bei Entscheidungen zu maximieren
- ☞ Nutzen aufgrund von Unsicherheit nicht exakt zu quantifizieren: Erwartungsnutzen
- ☞ Wettender fällt seine Entscheidung unter Risiko und Unsicherheit



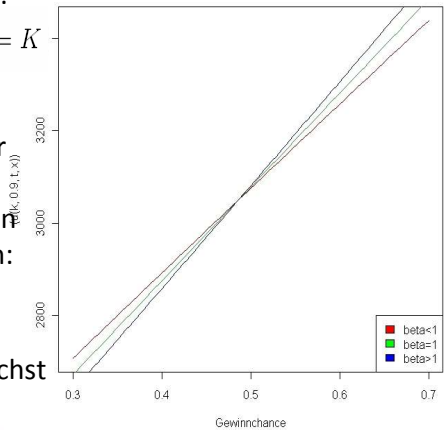
Die ökonomische Nutzentheorie II

- Einleitung
- Pferderennen
- Wetten
- Methodik
- Diskussion
- Ausblick
- Fazit

- ☞ Erwarteter Nutzen:

$$E(U_i) = \pi_i U(1 + O_i) = K$$

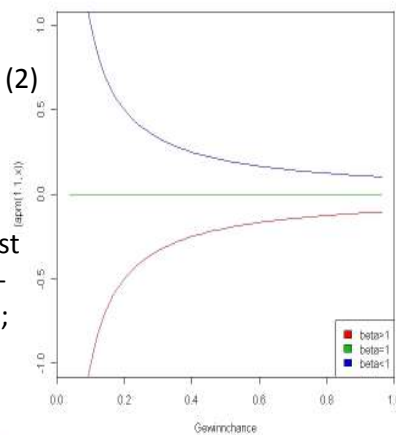
$$\Leftrightarrow U(1 + O_i) = \frac{K}{\pi_i}$$
- ☞ Nutzenfunktion für die Gewinnchance ergibt sich für einen erwarteten Nutzen: $K = 30$
- ☞ Mit steigender Gewinnchance wächst Nutzen



Das Arrow-Pratt Maß

- Einleitung
- Pferderennen
- Wetten
- Methodik
- Diskussion
- Ausblick
- Fazit

- ☞ Maß für die Intensität der Risikoablehnung
- $$\frac{U''(1 + O_i)}{U'(1 + O_i)} = -\frac{\beta - 1}{O_i + 1} \quad (2)$$
- ☞ Stetig wachsend und kleiner 0 für $\beta > 0$
- ☞ Risikoabneigung wächst für steigende Gewinnchancen eines Pferdes; risikoaversion fällt dagegen



Diskussion der Methodik des Bias

- Einleitung
- Pferderennen
- Wetten
- Methodik
- Diskussion
- Ausblick
- Fazit

- ☞ Verbessert Modell
- ☞ Bei allen Rennarten geeignet?
 → differenzieren
 - ☞ Für Aufgewichtsrennen nicht geeignet
- ☞ Bias weiter überprüfen
 - ☞ Wie wetten Europäer?
 - ☞ Nationale Unterschiede?



Platzierungswahrscheinlichkeiten I

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

Inwieweit können Platzierungen von Pferden vorhergesagt werden?

1. Wissen über Gewinnwahrscheinlichkeiten erlangen
 2. Mit deren Hilfe mittlere Rennzeiten schätzen
 3. Mit deren Hilfe Platzierungswahrscheinlichkeiten schätzen
- Verbessert die Aufnahme von Varianzen und Korrelationen das Modell?

18.12.2009

Clara- Astrid Wendelborn

17



Platzierungswahrscheinlichkeiten II

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

Modell von Harville:

- π_i wie in Formel (1) über die Wettanteile bestimmen
- Pferd i wird erster und Pferd j zweiter (Exacta Wette)

$$\pi_{ij} = \frac{\pi_i \pi_j}{1 - \pi_i} \quad (3)$$

- Pferd i wird erster, Pferd j zweiter, Pferd k dritter (Trifecta Wette)

$$\pi_{ijk} = \frac{\pi_i \pi_j \pi_k}{(1 - \pi_i)(1 - \pi_i - \pi_j)} \quad (4)$$

18.12.2009

Clara- Astrid Wendelborn

18



Platzierungswahrscheinlichkeiten III

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

Modell von Henery:

1. π_i wie in Formel (1) über die Wettanteile bestimmen
2. T_i : Rennzeit von Pferd i; θ_i : erwartete Rennzeit von Pferd i

$$\begin{aligned} \pi_i &= P(T_i < \text{MIN}\{T_r\}) \\ &= \int_{-\infty}^{\infty} \prod_{r \neq i} [1 - F(t_i | \theta_r)] f(t_i | \theta_i) dt_i \end{aligned} \quad (5)$$

3. Hierüber θ_i bestimmen

18.12.2009

Clara- Astrid Wendelborn

19



Platzierungswahrscheinlichkeiten VI

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

- Mithilfe des vorher geschätzten θ_i s:

$$\begin{aligned} \pi_{ij} &= P(T_i < T_j \mid \text{MIN}\{T_r\}) \\ &= \int_{-\infty}^{\infty} F(t_j | \theta_i) \prod_{r \neq i, j} [1 - F(t_j | \theta_r)] f(t_j | \theta_j) dt_j \end{aligned} \quad (6)$$

- Wahrscheinlichkeit, dass Pferd i erster und Pferd j zweiter wird

18.12.2009

Clara- Astrid Wendelborn

20



Platzierungswahrscheinlichkeiten V

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

- Henry nahm an: $T_i \sim N(\theta_i, 1)$ mit T_i unabhängig
- Formel sehr kompliziert für weitere Platzierungen:
→ Einfache Approximation von Lo und Bacon-Shone:
Weiterentwicklung von Formel (4):

$$\pi_{ijk} = \pi_i \frac{\pi_j^\lambda}{\sum_{s \neq i} \pi_s^\lambda} \frac{\pi_k^\tau}{\sum_{t \neq i, j} \pi_t^\tau} \quad (7)$$

18.12.2009

Clara- Astrid Wendelborn

21



Platzierungswahrscheinlichkeiten VI

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

Erweiterung des Henry Modells:

- Aufnahme von Korrelationen und Varianzen
- Nicht mehr $\text{Corr}(T_i, T_j) = \rho \quad \forall i, j$,
sondern $\text{Corr}(T_i, T_j) = \rho_{ij} = \psi_i \psi_j \quad \forall i \neq j \quad (8)$

Dabei ist $\log\left(\frac{\psi_i}{1 - \psi_i}\right) = -\delta - \gamma(\theta_i - \bar{\theta})$ mit $\bar{\theta} = \frac{1}{n} \sum_i \theta_i$, (9)

- und $\sigma_i = \exp(\kappa(\theta_i - \bar{\theta})) \quad (10)$
→ für κ und γ gleich 0 ergibt sich das Henry Modell

18.12.2009

Clara- Astrid Wendelborn

22



Maximum- Likelihood Schätzung

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

- Schätzfunktion um einen unbekannt Parameter zu schätzen
- auch bei komplexen Schätzsituationen anwendbar ist
- Verteilungsfunktion der Variablen muss bekannt sein
- Als Schätzwerte die Werte mit maximaler Likelihoodfunktion für die Parameter ermitteln, vorausgesetzt diese Parameter existieren.

$\ln L(\theta) = \sum_{i=1}^n \ln f(x_i | \theta)$ man maximiert die Log-Likelihoodfunktion durch Ableiten, wenn man die Ableitung =0 setzt kann man den unbekannt Parameter bestimmen

18.12.2009

Clara- Astrid Wendelborn

23



Umsetzung im Paper I

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

$\theta_i \approx \frac{\phi(z_0)(n-1)(z_i - z_0)}{M_1} \quad (11)$

$z_0 = \Phi^{-1}(1/n), \quad z_i = \Phi^{-1}(P_i)$

- M setzt sich aus dem Erwartungswert und der Varianz der Standardnormalverteilung zusammen, sowie über die Parameter κ, δ und γ

- Log-Likelihoodfunktion über Daten aus mehreren Rennen:

$$\ln L = \sum_{l=1}^{\#\text{Rennen}} \log \pi_{[12345]l}$$

18.12.2009

Clara- Astrid Wendelborn

24



Umsetzung im Paper II

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

- Modell wurde an 400 Rennen aus Hong Kong mit je acht Pferden angepasst; Platzierungswahrscheinlichkeiten wurde für die ersten fünf Pferde berechnet:

Modell	Geschätzte Parameter	P-Wert
a1) Inkonstante Korrelation (nur γ)	$\gamma = 0.58$	0.06
a2) Inkonstante Korrelation (γ und δ)	$\gamma = 0.60$ und $\delta = 0.05$	0.18
a3) Inkonstante Varianz	$\kappa = 0.08$	0.06

- p-Wert signifikante Unterschiede zwischen Henery Modell und dem mit Korrelation und Varianz

18.12.2009

Clara- Astrid Wendelborn

25



Diskussion der Methodik I

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

- Wenige betrachtete Variablen: Wettanteile und Rennzeiten
- Rennzeiten pro Pferd schwer einschätzbar, da unterschiedlich verteilt
- Weitere interessante Variablen: Startposition, Einfluss des Jockeys, etc.



18.12.2009

Clara- Astrid Wendelborn

26



Diskussion der Methodik II

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

- Wettanteile = subjektiv, tw. falsch, wie durch Favorite-Longshot Bias gezeigt
- Endgültige Quoten erst nach dem Schließen der Wettbüros bekannt → zu spät zum selber nutzen
- relativ leicht verständlich
- Mit Programm einfach anwendbar

18.12.2009

Clara- Astrid Wendelborn

27



Ausblick

Einleitung
Pferderennen
Wetten
Methodik
Diskussion
Ausblick
Fazit

- Generell: Anwendung der Methoden zur Überprüfung in Praxis
- Methoden auf viele Sportarten anwendbar:
- Methodik des Favorite-Longshot bias auf alle Sportarten auf die gewettet werden kann und wo es Favoriten gibt anwendbar, z.B. Fußball, Handball
- Methodik der Platzierungswahrscheinlichkeiten auf alle Rennen anwendbar, auf die gewettet wird, z.B. Autorennen, Hunderennen, Bootsrennen
- Problematik bleibt bestehen

18.12.2009

Clara- Astrid Wendelborn

28



Fazit

- Einleitung
- Pferderennen
- Wetten
- Methodik
- Diskussion
- Ausblick
- Fazit

- ☞ Erster Teil: Favorite-Longshot Bias ist kein universales Phänomen → nicht in allen Datensätzen
 - ☞ In den USA scheint Bias konsistent aufzutreten.
 - ☞ Hypothese: Phänomen hängt von der Größe des Wettpools ab
- ☞ Zweiter Teil: Modelle um Platzierungswahrscheinlichkeiten zu bestimmen
 - ☞ Es zeigte sich, dass es sich empfiehlt Korrelationen und Varianzen mit in das Modell aufzunehmen.