

1. Ist Fussball ein Glückspiel? Nicht ganz. Der Zufall wird auch bei der anstehenden Weltmeisterschaft eine so überragende Rolle spielen, dass man mit einer statistischen Analyse ziemlich weit kommt. Zum Beispiel lassen sich die Anzahl der Tore mit einer Poissonverteilung ziemlich genau beschreiben.

Gegeben seien die Daten der Ausgänge (reguläre Spielzeit) aller 575 Fussballweltmeisterschaftsspiele von 1990 bis 2002. Die empirische Verteilung der erzielten Tore ist



Goals

Goals	0	1	2	3	4	5	6	7	8
Frequency	19	49	60	47	32	18	3	3	1

Testen Sie, ob die Anzahl der Tore poissonverteilt mit  $\lambda = 2$  ist (Signifikanzniveau=0.05). (Lösungsblatt: Wert der Teststatistik) (3)

2. Ein begeisterter Fussballfan gibt bei der Fussballweltmeisterschaft in Südafrika Tipps ab, wobei er die Ziffern 0 (Unentschieden), 1 (erstgenannte Mannschaft gewinnt), 2 (letztgenannte gewinnt) unter Zuhilfenahme der Wahrscheinlichkeitsfunktion

$$P(X = k) = \begin{cases} 1/4 + ak + bk^2 & k = 0, 1, 2 \\ 0 & \text{sonst} \end{cases}$$

verteilt. Die Größen  $a$  und  $b$  hält er geheim. Es ist aber bekannt, dass für seine Tipps ausserdem  $P(X = 1) = \frac{1}{3}$  gilt. Bestimmen Sie  $a$  und  $b$  sowie die zugehörige Verteilungsfunktion (+ graphische Darstellung!). (Lösungsblatt:  $a$  und  $b$ ) (4)

3. Die Überwachung der Lauffreudigkeit von Fussballspielern stellt Heutzutage kein technisches Problem mehr dar und es ist state-of-the-art diese Daten für wichtige Spieler zu sammeln. Man schätzt, dass während eines Weltmeisterschafts-Fussballspieles ein Mittelfeldspieler im Durchschnitt neun bis elf Kilometer zurücklegt.

Aus taktischen Überlegungen erwägt der Trainer der Mannschaft von Uruguay einen defensiven zentralen und möglichst lauffreudigen Mittelfeldspieler (einen Ausputzer) einzusetzen um die Angriffslinien der Mannschaft von Südafrika zu stören.



Von zwei in Frage kommenden Spielern liegen die Ergebnisse (Anzahl an gelaufenen Kilometern) für Fussballspiele in der Vergangenheit vor. Das Resultat lautet wie folgend:

Two Sample t-test

```
data km by Spieler
t = 0.431, df = 108, p-value = 0.6673
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-2.46 3.82
sample estimates:
mean in group Spieler1 mean in group Spieler2
10.42 9.73
```

Welche der folgenden Aussagen sind korrekt? (Signifikanzniveau 5%)

- (a) Der Absolutwert der Teststatistik ist 1.96.
- + (b) Die Nullhypothese lautet, dass die Stichprobenmittel gleich sind.
- + (c) Das angegebene Konfidenzintervall beschreibt einen Intervallschätzer für ein Mittel=0.
- (d) Das Testresultat zeigt, dass die Kilometerleistung von Spieler 1 grösser ist als von Spieler 2.

(2)

(Lösungsblatt ankreuzen(!), zB ein Plus wenn korrekt und ein Minus wenn nicht korrekt)

4. Vuvuzelas, die berühmt berüchtigten Trompeten der Fussballfans in Südafrika stellen eine hohe Lärmbelastung dar (bis 143 Dezibel) die möglicherweise zu Gehörschäden bei Stadionbesuchern während der Weltmeisterschaft in Südafrika führen können.

Vuvuzelas sind aus Polyethylen gefertigt. Bei sachgemässrer Verwendung kann damit ein Dauerschallpegel von bis ca. 105 dB mit einer einzigen Vuvuzela erreicht werden. Zwar werden in Südafrika Ohrenstöpsel - sogenannte Vuvu-Stopper oder Tulazela - angeboten, die aber aufgrund einer hohen Nachfrage größtenteils schon ausverkauft sind. In der Medizin geht man davon aus, dass ab 102 dB Dauergehörschäden bei den Stadionbesuchern auftreten können.



Eine gemittelte Lautstärkemessung an 24 Zeitpunkten während eines Weltmeisterschaftsspiels in Südafrika (hypothetische Werte da noch keine aktuellen Messergebnisse verfügbar gemacht wurden) ergab folgende Abweichungen (in dB) vom zumutbaren Durchschnittswert 102dB

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Zeit	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Abweichung	-9	-2	0	-13	-4	1	14	0	11	9	15	13	8	0	7

	16	17	18	19	20	21	22	23	24
Zeit	16	17	18	19	20	21	22	23	24
Abweichung	-4	-2	2	28	20	18	10	24	11

Es kann angenommen werden, dass die Daten normalverteilt sind.

Wird der zumutbare Wert im Schnitt überschritten? (Signifikanzniveau 0.05)  
(Lösungsblatt: Wert der Teststatistik)

(2)

5. Eine Stichprobe unter den bestbezahlten Fussballspielern wurde erhoben und unter anderem das Jahreseinkommen (in Mio. Pfund) erfragt.



- Überprüfen Sie grafisch mit einem Wahrscheinlichkeitsnetz, ob die Jahreseinkommen der Spieler log-normalverteilt sind. Schätzen Sie das Mittel und die Standardabweichung der Jahreseinkommen unter Verwendung des Wahrscheinlichkeitsnetzes (3)
- Überprüfen Sie die logarithmierten Jahreseinkommen (unter Normalverteilungsannahme), ob für die amerikanischen und afrikanischen Fussballer die Varianzen übereinstimmen (Signifikanzniveau  $\alpha = 0.05$ ). (2)
- Testen Sie (unter Normalverteilungsannahme) für die Einkommen der europäischen Spieler, ob das mittlere Jahreseinkommen signifikant kleiner als 8 Mio Pfund ist (Signifikanzniveau  $\alpha = 0.05$ ) (2)

(Lösungsblatt: Wert der Teststatistik aus c, Wert der Teststatistik aus d)



	Name	Einkommen	Herkunft
1	Drogba	4.8	Afrika
2	Adebayor	<del>7.1</del>	Afrika
3	Eto-o	<del>4.1</del>	Afrika
4	Kanoute	<del>5.2</del>	Afrika
5	Kaka	<del>8.7</del>	Amerika
6	Ronaldinho	<del>6.5</del>	Amerika
7	Robinho	<del>5.2</del>	Amerika
8	Messi	<del>9.1</del>	Amerika
9	Tevez	<del>7.8</del>	Amerika
10	Casillas	5.2	Europa
11	Gerrard	6.5	Europa
12	Raul	<del>5.6</del>	Europa
13	Ibrahimovic	<del>10.1</del>	Europa
14	Bezena	<del>7.7</del>	Europa
15	Ferdinand	<del>5.6</del>	Europa
16	Alves	6.1	Europa
17	Deco	5.2	Europa
18	Terry	6.5	Europa
19	Lampard	6.5	Europa
20	Henry	6.5	Europa
21	Xavi	6.5	Europa
22	Rooney	5.2	Europa
23	Ronaldo	<del>11.3</del>	Europa
24	Valdez	5.2	Europa
25	Toure	<del>5.6</del>	Europa
26	Ballack	<del>5.6</del>	Europa

 $n = 4$  $n = 5$  $n = 12$ 

6. Betrachten Sie die Daten aus dem vorigen Beispiel.

Nehmen Sie nun an, dass irrtümlich auch ein Nicht-Spitzenverdiener im weltweiten Fussballgewerbe gefragt wurde. Noch schlimmer, nehmen Sie an, dass ein österreichischer Fussballer aus der österreichischen Liga wurde befragt.

Welche Schätzungen werden beeinflusst?

- (a) Schätzung der Regressionskoeffizienten (laut Vorlesungsskriptum) bei Regression von Einkommen (abh. Variable) gegen andere (nicht gelistete Kovariablen)
- (b) Interquartilsabstand
- (c) F-Test bei Varianzanalyse bzgl. der Gruppen
- (d) Arithmetisches Mittel
- (e) Medium Absolute Deviation

(2)

(In Lösungsblatt eintragen: Minus 1 Punkt pro falscher Antwort. Wenn TRUE dann zB ein Plus im Lösungsblatt zur Frage eintragen, wenn FALSCH dann ein Minus eintragen.)  
(Lösungsblatt ankreuzen)