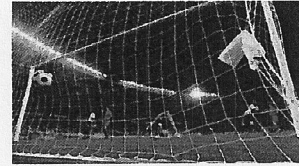


1. Ist Fußball ein Glückspiel? Nicht ganz. Der Zufall kann auch bei der anstehenden Europameisterschaft nicht machen was er will, sodass man mit einer statistischen Analyse oft einiges voraussagen kann. Zum Beispiel lassen sich die Anzahl der Tore mit einer Poissonverteilung ziemlich genau beschreiben.

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



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

Testen Sie, ob die Anzahl der Tore poissonverteilt ($P_\lambda(k) = \frac{\lambda^k}{k!} e^{-\lambda}$) mit $\lambda = 2$ ist (Signifikanzniveau=0.05). (Lösungsblatt: Wert der Teststatistik) (3)

2. Ein begeisterter Fußballfan gibt bei der Fußball Europameisterschaft in Polen/Ukraine 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 $P(X = 1) = \frac{1}{2}$ 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 Europameisterschafts-Fußballspieles ein Mittelfeldspieler im Durchschnitt neun bis elf Kilometer zurücklegt.

Aus taktischen Überlegungen erwägt der Trainer der Mannschaft von Irland einen defensiven zentralen und möglichst lauffreudigen Mittelfeldspieler (einen Ausputzer) einzusetzen um die Angriffslinien der Mannschaft von Spanien 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%)

905

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

(2)

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

- ✓ 4. Wer erinnert sich an die TrötKonzerte während der Weltmeisterschaft in Südafrika? Vuvuzelas, diese berühmt berüchtigten Trompeten der Fußballfans in Südafrika stellen eine hohe Lärmbelastung dar (bis 143 Dezibel), die möglicherweise zu Gehörschäden bei Stadionbesuchern führen.

Vuvuzelas sind aus Polyethylen gefertigt. Bei sachgemäßer 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) ergab folgende Abweichungen (in dB) vom zumutbaren Durchschnittswert 102dB

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

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.

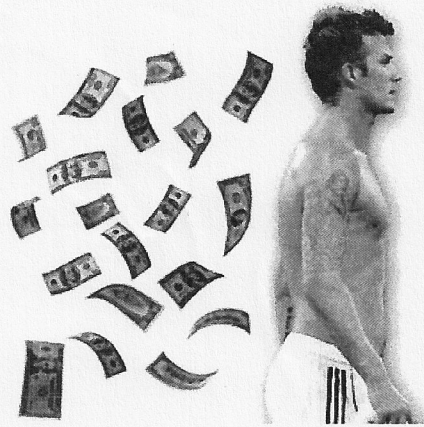
Wird der zumutbare Wert im Schnitt überschritten? (Signifikanzniveau 0.05)
(Lösungsblatt: Wert der Teststatistik)

(2)

5. Eine Stichprobe unter den bestbezahlten Fußballspielern wurde im Jahr 2010 erhoben und unter anderem das Jahreseinkommen (in Mio. Pfund) erfragt.

	Name	Einkommen	Herkunft
1	Ronaldo	11.3	Europa
2	Ibrahimovic	10.4	Europa
3	Eto-o	9.1	Afrika
4	Messi	9.1	Amerika
5	Kaka	8.7	Amerika
6	Adebayor	7.4	Afrika
7	Bezema	7.4	Europa
8	Tevez	7.0	Amerika
9	Ronaldinho	6.5	Amerika
10	Gerrard	6.5	Europa
11	Terry	6.5	Europa
12	Lampard	6.5	Europa
13	Henry	6.5	Europa
14	Xavi	6.5	Europa
15	Alves	6.1	Europa
16	Raul	5.6	Europa
17	Ferdinand	5.6	Europa
18	Toure	5.6	Europa
19	Ballack	5.6	Europa
20	Kanoute	5.2	Afrika
21	Robinho	5.2	Amerika
22	Casillas	5.2	Europa
23	Deco	5.2	Europa
24	Rooney	5.2	Europa
25	Valdez	5.2	Europa
26	Drogba	4.8	Afrika

$\frac{1}{n}$



- a) Mit den log-normalverteilten Daten: Überprüfen Sie grafisch mit einem Wahrscheinlichkeitsnetz, ob die logarithmierten Jahreseinkommen der Spieler normalverteilt sind. Schätzen Sie das Mittel und die Standardabweichung der log-Jahreseinkommen unter Verwendung des Wahrscheinlichkeitsnetzes. (3)
- b) Überprüfen Sie die logarithmierten Jahreseinkommen (unter Normalverteilungsannahme), ob für die amerikanischen und afrikanischen Fußballer die Varianzen übereinstimmen (Signifikanzniveau $\alpha = 0.05$). (2)
- c) Testen Sie (unter Normalverteilungsannahme der logarithmierten Werte) 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)

6. Betrachten Sie die Daten aus dem vorigen Beispiel.

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

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)*