Einführung in die Stochastik
4. Übungsblatt

TECHNISCHE UNIVERSITÄT DARMSTADT

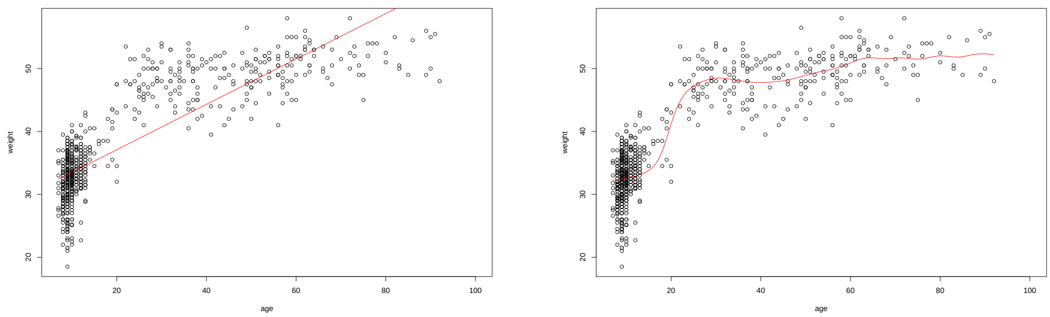
Fachbereich Mathematik
M. Kohler
A. Fromkorth
D. Furer

SS 2011
13.05.2011

Gruppen und Hausübung

Aufgabe 9 (4 Punkte)

Vor einigen Jahren haben amerikanische Biologen die amerikanische Walddrossel (eine Vogelart) beobachtet, um dabei unter anderem den Zusammenhang zwischen dem Alter und dem Gewicht der Vögel zu untersuchen. Die Durchführung der linearen und nichtparametrischen Regressionsschätzung hat folgende Bilder (siehe Abbildung 1) ergeben. Vergleichen Sie diese zwei Arten der Regressionsschätzung anhand der Bilder. Interpretieren Sie den Zusammenhang von Alter und Gewicht der amerikanischen Walddrossel, der aus der linearen bzw. nichtparametrischen Regression hervorgeht. Was können Sie über die Korrelation der Variablen Alter und Gewicht aussagen?



(a) lineare (b) nichtparametrische

Abbildung 1: Regressionsschätzungen

Lösung:

- Aus diesen Bildern ist zu erkennen, dass bei der linearen Regressionsschätzung eine Gerade an die Daten angepasst wird. Dahingegen passt man bei der nichtparametrische Regressionsschätzung allgemeine Funktionen an die Daten an, wobei man keine Annahmen über deren Bauart macht.
- Betrachtet man das Bild (a), so suggeriert die Gerade mit positiven Steigung, dass die Walddrosseln irgendwann mal “riesen groß“ werden, was natürlich nicht der Wahrheit entspricht. In der Abbildung (b) zu nichtparametrischen Regressionsschätzung ist auch eine permanente (bis auf das Intervall ca. [38,43]) Gewichtszunahme zu beabachten, diese “hält sich aber im Rahmen”
- Korrelation hat das gleiche Vorzeichen wie die Steigung der Regressionsgeraden. Insbesondere folgt hier: die Korrelation ist positiv.

Aufgabe 10 (4 Punkte)

Die Innenstadt von Mannheim ist in Häuserblöcken statt in Straßenzügen angelegt (siehe Abbildung 2). Deswegen hat sie auch den Beinamen Quadratestadt. In der Horizontalen sind die Nummern und in der Vertikalen die Buchstaben abgetragen, sodass jeder Häuserblock eindeutig durch z.B. A1 oder K3 identifiziert werden kann. Eine Touristengruppe, die sich im Block C6 aufhält, möchte von hier aus (Start) bis zum Block T6 (Ziel) den kürzesten Weg nehmen. (Einfachheitshalber wird angenommen, dass alle Blöcke gleich groß sind.) Wieviele verschiedene kürzeste Wege hat die Gruppe zur Auswahl.

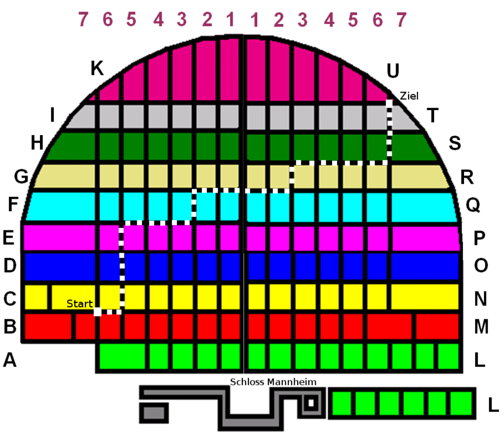


Abbildung 2: Innenstadt von Mannheim

Lösung: Als Erstes stellt man fest, dass die Gruppe um das Ziel zu erreichen genau $R = 12$ Mal nach “Rechts“ und $O = 7$ Mal nach “Oben” gehen muss. Demnach sollen die Turisten sich an $R + O$, also an genau 19 Kreuzungen für eine der Richtungen entscheiden. Bezeichnet man die Entscheidung nach “Rechts“ zu gehen mit einer 1 und nach ”Oben“ mit einer 0, so kann man den, in der Abbildung eingezeichneten, Weg wie folgt beschreiben:

$$1, 0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 0$$

Als Nächstes sieht man, um den Weg eindeutig zu beschreiben, genügt es z.B. die 7 ”Aufwärtsbewegungen” festzulegen. Dies ist aber nichts anderes, als die sieben 0 auf neunzehn Stellen zu verteilen. Das Problem entspricht dem Ziehen ohne Zurücklegen und ohne Berücksichtigung der Reihenfolge, also:

$$\binom{19}{7} = \frac{19!}{7!(19-7)!} = 50388$$

Somit gibt es 50388 Möglichkeiten am kürzesten von “Start” zu “Ziel” zu gelangen.

Aufgabe 11 (4 Punkte)

Ein Zufallsgenerator erzeugt mit Ziffern aus $\{0, 1, ..., 9\}$ Ziffernblöcke der Länge 4. Geben Sie mit Begründung die Wahrscheinlichkeiten für folgende fünf Ereignisse an:

(a) alle Ziffern verschieden

(b) genau ein Paar gleicher Ziffern

(c) genau zwei Paare gleicher Ziffern

(d) genau drei gleiche Ziffern

(e) vier gleiche Ziffern

Berechnen Sie zur Kontrolle die Summe aller Wahrscheinlichkeiten.

Lösung: $\Rightarrow$ Anzahl d. möglichen Fälle $= 10^4$

a) Wk., dass alle Ziffern verschieden sind = ?

Anschaulich: $a_1, a_2, a_3, a_4 \in \{0, 1, \ldots, 9\}$

Alle Ziffern verschieden: $\boxed{a_1 \mid a_2 \mid a_3 \mid a_4}$

Also: Für die erste Stelle gibt es 10 Möglichkeiten für a_1, für die zweite Stelle gibt es nur noch 9 Möglichkeiten für a_2 (da a_1 schon vergeben ist), für die dritte Stelle 8 und für die vierte Stelle 7 Möglichkeiten.

Insgesamt $10 \cdot 9 \cdot 8 \cdot 7 = 5040$ Möglichkeiten.

$\hookrightarrow$ Wk, dass alle Ziffern verschieden:

$$\frac{\text{Anz. günstiger Fälle}}{\text{Anz. möglicher Fälle}} = \frac{5040}{10^4} = 0.504 \mathrel{\widehat{=}} P(\mathbf{a}))$$

b) Wk., genau ein Paar gleicher Ziffern ist = ?

Anschaulich: $\boxed{a_1 \mid a_1 \mid a_3 \mid a_4}$

Also: Für die erste Stelle gibt es 10 Möglichkeiten für a_1 und somit für die zweite Stelle genau 1 Möglichkeiten (da a_1 schon vergeben ist), für die dritte Stelle gibt es 9 und für die vierte Stelle 8 Möglichkeiten.

Insgesamt $10 \cdot 1 \cdot 9 \cdot 8 = 720$ Möglichkeiten.

Anzahl d. Anordnungsmöglichkeiten: 6 siehe Skizze

a_1	a_1	a_3	a_4
a_1	a_3	a_1	a_4
a_1	a_3	a_4	a_1
a_3	a_1	a_1	a_4
a_3	a_1	a_4	a_1
a_3	a_4	a_1	a_1

$\hookrightarrow$ Wk, dass genau ein Paar gleicher Ziffern auftritt:

$$\frac{\text{Anz. günstiger Fälle}}{\text{Anz. möglicher Fälle}} = \frac{6 \cdot 720}{10^4} = 0.432 \mathrel{\widehat{=}} P(\mathbf{b}))$$

c) Wk., genau zwei Paare gleicher Ziffern = ?

Anschaulich: $\boxed{a_1 \mid a_1 \mid a_2 \mid a_2}$

Also: Für die erste Stelle gibt es 10 Möglichkeiten für a_1 und somit für die zweite Stelle genau 1 Möglichkeiten (da a_1 schon vergeben ist), für die dritte Stelle gibt es 9 und für die vierte Stelle ebenfalls genau 1 Möglichkeit.

Insgesamt $10 \cdot 1 \cdot 9 \cdot 1 = 90$ Möglichkeiten.

Anzahl d. Anordnungsmöglichkeiten: 3

$\hookrightarrow$ Wk, genau zwei Paare gleicher Ziffern:

$$\frac{\text{Anz. günstiger Fälle}}{\text{Anz. möglicher Fälle}} = \frac{90 \cdot 3}{10^4} = 0.027 \mathrel{\widehat{=}} P(\mathbf{c}))$$

d) Wk., genau drei gleiche Ziffern = ?

Anschaulich: $\boxed{a_1 \mid a_1 \mid a_1 \mid a_2}$

Insgesamt $10 \cdot 1 \cdot 1 \cdot 9 = 90$ Möglichkeiten.

Anzahl d. Anordnungsmöglichkeiten: 4

$\hookrightarrow$ Wk, genau drei gleiche Ziffern:

$$\frac{\text{Anz. günstiger Fälle}}{\text{Anz. möglicher Fälle}} = \frac{90 \cdot 4}{10^4} = 0.036 \mathrel{\widehat{=}} P(\mathbf{d}))$$

e) Wk., alle gleich = ?

Anschaulich: | a_1 | a_1 | a_1 | a_1 |

Insgesamt $10 \cdot 1 \cdot 1 \cdot 1 = 10$ Möglichkeiten.

$\hookrightarrow$ Wk, alle gleich:

$$\frac{\text{Anz. günstiger Fälle}}{\text{Anz. möglicher Fälle}} = \frac{10}{10^4} = 0.001 \mathrel{\hat{=}} P(\mathbf{e}))$$

Kontrolle: Summe aller Wk.: $P(\mathbf{a})) + P(\mathbf{b})) + P(\mathbf{c})) + P(\mathbf{d})) + P(\mathbf{e})) = 1$

Aufgabe 12 (4 Punkte)

a) (Erstes Lemma von Borel und Cantelli)
Sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum und sei $(A_n)_n$ eine Folge von Ereignissen mit

$$\sum_{n=1}^{\infty} P(A_n) < \infty.$$

Beweisen Sie, dass dann gilt

$$P(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k) = 0.$$

Hinweis: Begründen Sie zunächst

$$P(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k) \leq P(\cup_{k=N}^{\infty} A_k) \ (N \in \mathbb{N}).$$

Verwenden Sie diese Beziehung und schätzen Sie die Wahrscheinlichkeit rechts mit Hilfe der σ-Subadditivität ab.

b) Dozent K. fragt in jeder seiner mündlichen Prüfungen nach dem ersten Lemma von Borel und Cantelli. Da die Studenten sich untereinander absprechen, kann der n-te Prüfling die Frage nur mit Wahrscheinlichkeit $\frac{1}{n^2}$ nicht richtig beantworten. Wie groß ist die Wahrscheinlichkeit, dass bei sukzessiver Durchführung von unendlich vielen Prüfungen nur endlich viele der Prüflinge diese Frage nicht richtig beantworten?
Hinweis: Wenden Sie a) mit A_n ="Prüfling n beantwortet die Frage nicht richtig" an.

Lösung: a) Mit dem Hinweis gilt

$$P(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k) \overset{(1)}{\leq} P(\cup_{k=N}^{\infty} A_k) \overset{\sigma-subadd.}{\leq} \sum_{k=N}^{\infty} P(A_k) \overset{(2)}{\to} 0 \, (N \to \infty).$$

Zu (1): Sei $w \in \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k$. Dies bedeutet wiederum, dass w in unendlich vielen A_k liegt. Daraus folgt: es existiert mindestens ein $N_0 \geq N$ für alle $N \in \mathbb{N}$ mit $w \in A_{N_0}$. Demnach gilt $w \in \cup_{k=N}^{\infty} A_k$. Die Behauptung (1) folgt mit der Monotonie des Wahrscheinlichkeitsmaßes.

Zu (2): Diese Behauptung folgt aus der Konvergenz der Reihe

$$\sum_{n=1}^{\infty} P(A_n).$$

b) Der Hinweis und der Aufgabentext liefert

$$P(A_n) = \frac{1}{n^2}.$$

Wegen

$$\sum_{i=1}^{\infty} P(A_n) < \infty$$

kann man a) anwenden und es gilt

$$
\begin{aligned}
P(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k) &= \{\omega \in \Omega : \# \{n \in \mathbb{N} : \omega \in A_n\} = \infty\} \\
&= 0.
\end{aligned}
$$

D.h. das Ereignis, dass unendlich viele A_n eintreten, hat Wahrscheinlichkeit 0. Damit ist die Wahrscheinlichkeit, dass bei sukzessiver Durchführung von unendlich vielen Prüfungen nur endlich viele der Prüflinge diese Frage nicht richtig beantworten gleich dem Gegenereignis von $P(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k)$, also gleich 1.

Dieses Übungsblatt wird im Rahmen der Übungen am 16. bzw. 17.05.2011 besprochen. Ihre Ausarbeitungen geben Sie am 23. bzw. 24.05.2011 in Ihrere Übungsgruppe ab. Sie erhalten diese am 30. bzw. 31.05.2011 korrigiert zurück.