Arbeitsblatt HT1: Poisson-Modell

1. Beispiele zur Poisson-Verteilung

Bearbeiten Sie die Datensätze (i) und (ii) gemäss den Teilaufgaben (a) – (c).

(i) **Radioaktivität.**
Rutherford und Geiger haben die Emission der α-Teilchen einer radioaktiven Substanz gemessen. In der Tabelle bezeichnet k die Anzahl der α-Teilchen, die in der Zeiteinheit ($= 7.5s$) beobachtet wurden; z_k gibt an, wie oft $k = 0, 1, \ldots, 14$ α-Teilchen beobachtet wurden (total 10'129 Zerfälle). (Datensatz ist im File `alpha.dat` gespeichert.)

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15+
z_k	57	203	383	525	532	408	273	139	49	27	10	4	0	1	1	0

(ii) **Verkehr.**
Die Poisson-Verteilung wird auch als Modell bei geringem Verkehr verwendet. Man geht dann davon aus, dass die Verteilung der Anzahl Autos in einem gegebenen Zeitintervall oder Gebiet genähert Poisson verteilt sind, falls der Verkehrsfluss ungefähr konstant und der Verkehr gering ist (damit sich die einzelnen Autos unabhängig voneinander bewegen können).

Die folgende Tabelle enthält die Anzahl Rechtsabbieger während 300 mal 3 Minuten Intervallen an einer Kreuzung. (Datensatz ist im File `verkehr.dat` gespeichert.)

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13+
z_k	14	30	36	68	43	43	30	14	10	6	4	1	1	0

a) Zeichnen Sie ein Balkendiagramm.

b) Bestimmen Sie den Mittelwert der Daten.

c) Zeichnen Sie die Werte der Poisson-Verteilung mit Parameter λ in das Stabdiagramm ein, indem Sie als Schätzer für den unbekannten Parameter λ den Mittelwert aus Unteraufgabe (b) verwenden.

R-Anleitung:

a)
```
> alpha <- read.table("alpha.dat", header=TRUE)          # Daten einlesen; oder mit Hilfe des Menüs
> plot(alpha$k, alpha$freq, type="h", lwd=6, col="blue", lend=2, xlab="k",
          ylab="Häufigkeit")        # Balkendiagramm; dafür wichtige Argumente: "type="h", lwd=6"
```

b)
```
> n <- sum(alpha$freq)                        # Anzahl Daten
> mu <- sum(alpha$freq*alpha$k)/n             # Schätzen von λ
> mu                                          # Mittelwert ansehen
```

c)
```
> yModell <- dpois(alpha$k, lambda=mu) * n                   # n · p̂_k bestimmen
> lines(alpha$k, yModell, type="b", lwd=2, col="red")        # n · p̂_k in Balkendiagramm einzeichnen
```

Bitte wenden!

2. Kennzahlen für beliebige Wahrscheinlichkeitsverteilungen

Das “Standard-Zufalls-Experiment” ist der Würfel. Wenn wir mit einem “fairen” Würfel würfeln, dann ist jede Augenzahl k, ($k = 1, 2, \ldots, 6$) gleich wahrscheinlich, also $P\langle X = k\rangle = \frac{1}{6}$.

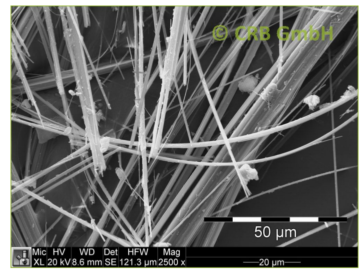
a) Zeichnen Sie im Fall des Würfels die Wahrscheinlichkeiten mit einem Balkendiagramm auf.

b) Berechnen Sie im Fall des Würfels den Erwartungswert, die Varianz und die Standardabweichung.

c) Berechnen Sie diese drei Kennzahlen für den Fall eines Würfels mit den Augenzahlen 1, 2, 3, . . . , 12.

3. Asbestfasern.

Um die Asbest-Konzentration zu messen, werden unter anderem Asbestfasern auf einem Filter ausgezählt. Dieser Teilaspekt wurde untersucht, indem Asbest in Wasser gelöst und gleichmässig über den Filter gesprüht wurde. Dann wurden 23 Proben von Filterstückchen mit einem Durchmesser von 3 *mm* genommen und die sich darauf befindenden Asbestfasern unter dem Elektronenmikroskop ausgezählt. Die Daten, die man erhielt, sind im Daten Frame `asbest1` gespeichert.



a) Bestimmen Sie das Maximum, das Minimum, die Spannweite und den Mittelwert der Daten.

Der vollständige Datensatz ist im Daten Frame `asbest2` gespeichert, da man noch weitere 52 Filterstückchen ausgezählt hatte.

b) Führen Sie mit dem vollständige Datensatz Teilaufgabe (a) nochmals durch. Was fällt Ihnen auf?

c) Zeichen Sie für beide Datensätze jeweils ein Stabdiagramm und legen Sie darüber die Kurve der Poisson-Verteilung, wobei für den unbekannten Parameter seine Schätzung eingesetzt werden soll. Was schliessen Sie aus diesen beiden Grafiken?

R-Anleitung:

a)
```
> # Daten als Vektor einlesen
> asbest1 <- read.table("CAS-DA_ModulA2-HT1_Daten/asbest1.dat", header=F)[,1]
> min(asbest1)   # ebenso mit max() und mean(); (Achtung! asbest1 ist ein Vektor)
```

c)
```
> par(mfrow=c(1,2))   # 2 Grafiken pro "Seite"
> a1 <- table(asbest1)                              # Häufigkeitstabelle erstellen
```

4. Segelflugunfälle.

Wenn die tödlichen Segelflugunfälle in der Schweiz seit 1986 betrachtet werden, erhält man einen Durchschnitt von 4.21 Unfällen pro Jahr. Im Jahr 2008 ereigneten sich 7 tödliche Unfälle, was zu Massnahmen Anlass gab. Betrachten Sie die obige Situation mit den Methoden der Statistik:

a) Welches Modell würden Sie verwenden, um die Anzahl Segelunfälle zu beschreiben? Warum?

b) Ist die Anzahl von 7 tödliche Unfällen gemäss Ihrem Modell aus der vorhergehenden Teilaufgabe plausibel? Beantworten Sie die Frage mit Faustregel P1 zum Niveau von 5%. (Freiwillig: Beantworten Sie die Frage exakt - siehe Skript.)