Übung 3

1. (**Zählt doppelt**: Ein Punkt für den Computer-Teil) Die Anzahl X der Erdbeben während einer Zeitspanne wird oft mit einer Poisson-Verteilung modelliert. Vom 21. bis 25. Oktober 1996 wurden in einem bestimmten Gebiet 30 Erdbeben der Stärke grösser als 4 beobachtet.

a) Schätze $\widehat{\lambda}$, den Parameter der Poissonverteilung für die Anzahl Erdbeben (mit einer Stärke grösser 4) in einer Periode von 5 Tagen in diesem Gebiet.

b) Bestimme mit Hilfe einer Faustregel das 95%-Vertrauensintervall für λ.

c) Bestimme das exakte 95%-Vertrauensintervall $]\lambda_u, \lambda_o[$, dessen untere Grenze λ_u und dessen obere Grenze λ_o gegeben sind durch

$$P_{\lambda_u}[X \geq 30] = 0.025 \qquad P_{\lambda_o}[X \leq 30] = 0.025$$

R-Anleitung: `> ppois(30,λ)` # $P_\lambda[X \leq 30]$

Hinweis: Verwende die mit der Faustregel bestimmten Intervall-Grenzen als Startwerte für λ_u resp. λ_o und bestimme dann die richtigen Werte (auf eine Stelle nach dem Komma genau) durch ausprobieren.

d) Ergänze folgende Aussagen durch Vermutungen, die aufgrund Deiner obigen Resultate naheliegend sind.

i) $\lambda \leq \lambda_u \Longrightarrow P_\lambda[X \geq 30] \leq \ldots$

ii) $\lambda \in]\lambda_u, \lambda_o[\Longrightarrow P_\lambda[X \geq 30] \geq \ldots, \quad \lambda \in]\lambda_u, \lambda_o[\Longrightarrow P_\lambda[X \leq 30] \geq \ldots$

iii) $\lambda \geq \lambda_o \Longrightarrow P_\lambda[X \leq 30] \leq \ldots$

e) Herr Thöny von einer Rückversicherungsgesellschaft behauptet, die „wahre mittlere Erdbebenanzahl" sei 14.5 Beben der Stärke grösser als 4 in einer Zeitspanne von fünf Tagen. Ist diese Behauptung (unsere zu prüfende Nullhypothese) mit unserer Beobachtung von 30 Erdbeben „verträglich"? Beantworte die Frage auf jede der folgenden Arten:

1. Lösungsweg: Liegt $\lambda_0 = 14.5$ im approximativen Vertrauensintervall?
2. Lösungsweg: Liegt $\lambda_0 = 14.5$ im exakten Vertrauensintervall?
3. Lösungsweg: Bestimme mittels einer Faustregel den Annahmebereich $]x_u, x_o[$ des Tests, ob $\lambda_0 = 14.5$ sein kann und prüfe, ob die Beobachtung $x = 30$ im Annahmebereich liegt.
4. Lösungsweg: Bestimme den exakten Annahmebereich $]x_u, x_o[$ für einen Test auf dem 5% Niveau, ob $\lambda_0 = 14.5$ sein kann.
   Die *kritischen Werte* x_u und x_o (beide ganzzahlig!) sind durch

$$P_{14.5}[X \leq x_u] \leq^* 0.025 \qquad P_{14.5}[X \geq x_o] \leq^* 0.025$$

gegeben, und $\leq^*$ soll heissen: "gleich oder kleiner, aber möglichst gross".
Prüfe, ob die Beobachtung $x = 30$ im Annahmebereich liegt.
R-Anleitung: `> ppois(x,14.5)` # $P_{14.5}[X \leq x]$

f) Beschreibe mit eigenen Worten in wenigen Sätzen die Beziehung zwischen Vertrauensintervallen und Tests.

2. (Simulationsaufgabe) In dieser Aufgabe lassen wir uns vom Computer Zufallszahlen zu einem vorgegebenen λ simulieren, berechnen für jede Zufallszahl das zugehörige Vertrauensintervall und schauen, wie oft λ im berechneten Intervall liegt.

a) Wie oft sollte das wahre λ innerhalb des 95%-Vertrauensintervalls liegen?

b) Simuliere 1'000 Zufallszahlen, die poissonverteilt sind mit $\lambda = 41.3$ und berechne die zugehörigen Vertrauensintervalle mit Hilfe der Faustregel. Wie oft liegt das wahre λ im Vertrauensintervall? Plotte die Vertrauensintervalle und drucke sie aus. Was siehst Du?

R-Anleitung:

```
set.seed(912)                    ## initalisiert Zufallsgenerator ==> reproduzierbares Resultat!
zz <- rpois(1000, lambda = ?)                           ## 1000 Zufallszahlen generieren
ki.unten <- zz - 2 * sqrt(zz)                           ## untere Grenze der Konfidenzintervalle
ki.oben <- ?                                            ## obere Grenze der KI
ki.enthaelt.41.3 <- 41.3 >= ki.unten & 41.3 <= ki.oben  ## logischer Vektor
prozent.enthalten <- sum(ki.enthaelt.41.3) / ?          ## sum ==> Anzahl TRUE's

postscript("ki.ps")  ## alle plots, die folgen, ins file ki.ps plotten, bis zum Befehl dev.off()
plot(c(1,1000), c(41.3, 41.3), type = "l", ylim = c(min(ki.unten), max(ki.oben)))
for (i in 1:100*10)                                     ## trage jedes zehnte KI ein
  lines(c(i, i), c(ki.unten[i], ki.oben[i]))
dev.off()                                               ## plotten in ein file beenden
```

Dann in neuem Terminal folgenden Befehl schreiben: **gv ki.ps** und im Menü *File* mit *Printer: nearest* drucken.

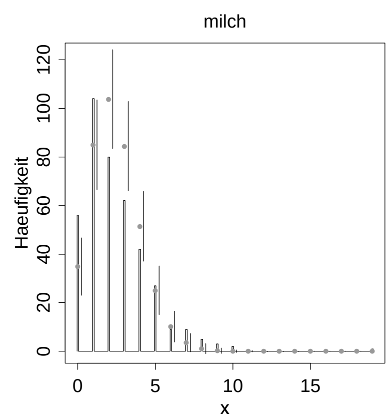
c) (Freiwillig) Führe dasselbe nun für $\lambda = 2$ durch (ohne den grafischen Teil). Stimmt die Prozentzahl noch? Weshalb nicht?

3. Beim Testen von Milch auf bakteriologische Verunreinigungen wird $0.01ml$ Milch über eine Glasfläche von $1cm^2$ verteilt. Diese Fläche wird in 400 gleich grosse Quadrätchen unterteilt und die Bakterienkolonien darin unter dem Mikroskop ausgezählt worden. In einem Versuch wurde folgendes Ergebnis erzielt:

k	0	1	2	3	4	5	6	7	8	9	10	19
z_k	56	104	80	62	42	27	9	9	5	3	2	1

a) Könnte sich aufgrund des Versuchs die Poissonverteilung als Modell für die Anzahl Kolonien pro Quadrätchen eignen?

b) Unten sind die gezählten Daten als Stabdiagramm dargestellt. Zusätzlich sind bei jedem Stäbchen die Bereiche eingezeichnet, in denen sich die Anzahl Quadrätchen mit soviel Bakterien befinden sollte, sofern es sich um eine Poissonverteilung handelt.



Beurteilen Sie nun aufgrund der Grafik, ob die Poissonverteilung ein geeignetes Modell darstellt.

c) Welche Gründe könnten für obiges Resultat sorgen?

Abgabe: Montag, 17. April, in der Pause der Vorlesungsstunde (UMNW: Di, 24. April)