Educational Material

Datenanalyse
WS 2001/00

Author(s):
Pruys, H.

Publication Date:
2001

Permanent Link:
https://doi.org/10.3929/ethz-a-004444215

7 Zeitreihen, Exponential- und Poissonverteilung

Unter Zeitreihen versteht man im allgemeinen eine Reihe von Zeitpunkten. Auf dem beiliegenden Zeitungsauschnitt sind einige interessante Zeitreihen dargestellt — es handelt sich um die Daten von Überschwemmungskatastrophen. Genauso liessen sich die Zeitpunkte zu anderen außergewöhnlichen metereologischen Ereignissen, Erdbeben, Börsencrashs, Flugzeugabstürzen, Geburten, radioaktiven Zerfällen, usw. darstellen. In allen Fällen interessiert, wie die Zeiten verteilt sind — und insbesondere, ob sie zufällig verteilt sind. Um diese Frage beantworten zu können, sollten wir uns erst einmal klarmachen, wie Zufallsreihen aussehen. Unter Zuhilfenahme des Skripts (Kapitel 2.3) finden wir, daß die Anzahl an Ereignissen in einem endlichen Zeitabschnitt der Reihe Poissonverteilt ist und daß die Zeitabstände innerhalb dieser Reihen exponentialverteilt sind. In dieser Übung wollen wir letztere Eigenschaft ausnutzen, um zufällige Zeitreihen zu generieren. Die zugrundeliegende Poissonstatistik werden wir dann überprüfen und die Reihe visuell mit den in der Zeitung abgedruckten vergleichen.
Die Exponentialverteilung hat folgende Form:

$$f(t;\lambda) = \lambda e^{-\lambda t}$$

($1/\lambda$ ist die mittlere Zeit)
Hieraus ergibt sich für unsere Reihe $t_{1\dots n}$ von zufällig verteilten Zeitpunkten:

$$f(t_{i+1}) = t_i + \lambda e^{-\lambda(t_{i+1}-t_i)}$$
$$f(t_1) = \lambda e^{-\lambda(t_1)}$$

mit frei wählbaren λ. Die verbleibende Frage ist, wie man in der Praxis(MATLAB) exponentialverteilte Zufallszahlen erzeugt. Im Skript Kapitel 3.1 finden wir alle nötigen Informationen. Insbesondere ist Formel (3.6) ein einfaches 'Kochrezept', um aus gleichverteilten Zahlen x_i, die wir in Matlab problemlos erzeugen können, exponentialverteilte Zahlen y_i zu machen:

$$y_i = \frac{-lnx_i}{\lambda}$$

Nachdem wir das theoretische Rüstzeug zusammen haben, kommen wir jetzt zum praktischen Teil, also wie auch in den anderen Übungen zur Programmbeschreibung:

- Wahl von λ und der Länge der Zeitreihe t_{max}
- Erzeugung einer ausreichend großer Reihe von gleichverteilten Zufallszahlen (`rand`)
- Umwandlung in eine exponentialverteilte Reihe (s.o., `log`)
- Umwandlung in eine zufällige Zeitreihe (s.o.,`cumsum`)
- Darstellung der Zeitreihe (NICHT das Ende der Reihe mit darstellen oder in späteren Schritten verwenden! Warum nicht!?). Man verwende den Befehl `plot`
- Überprüfung, ob die Anzahl der Zeiten in einem endlichen Intervall wirklich poissonverteilt ist, wie man es von der Theorie her erwarten würde. Im einzelnen heißt das:
  - Wahl einer Intervallbreite — am besten in Einheiten von $1/\lambda$. Es sollten im Mittel einige wenige Punkte in jedem Intervall liegen.
  - Mittels des Befehls `hist` die Zeitpunkte in jedem Intervall zählen und anschließend die 'vergröberte' Reihe darstellen (`bar`). Man vergleiche die Beispielfigur weiter unten, falls unklar sein sollte, was man unter der 'vergröberten' Reihe zu verstehen hat.
  - Die Anzahl der Einträge in den einzelnen Intervallen histogrammieren (`hist`).
  - Dieses Histogramm darstellen und mit einer Poissonverteilung vergleichen — die Verteilung also mit in die Figur hineinzeichnen, am besten mit Fehlerbalken (man verwende die Befehle `poisson` und `plot_data`).

Das Programm sollte man einige Male starten, um einen Eindruck von der Dynamik solcher Zeitreihen zu bekommen. Am Schlußs schaue man sich nochmals den Zeitungsausschnitt an und ziehe seine eigenen Schlüsse.

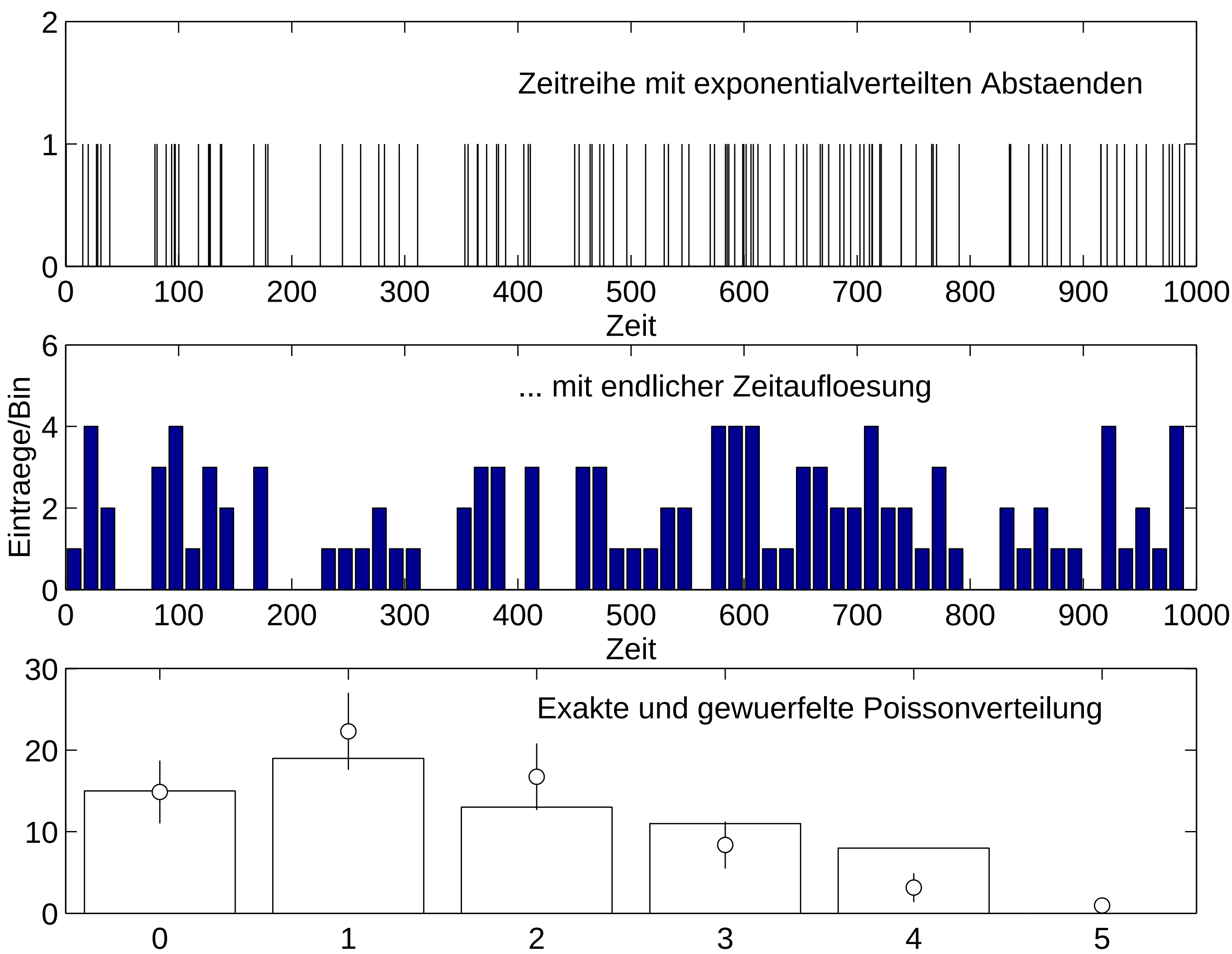

Abbildung 4: Eine zufällige Zeitreihe, ihre Vergröberung und die abgeleitete Poissonverteilung