ETH Zürich HS 2014
RW, D-MATL, D-MAVT
Prof. Marloes Maathuis

Koordinator
Lukas Gonon

Stochastik
Serie 8

1. Wir betrachten eine stetige Verteilung mit folgender Dichte:

$$f(x) = \begin{cases} \frac{\alpha}{x^{\alpha+1}} & x \geq 1 \\ 0 & x < 1 \end{cases}$$

wobei $\alpha > 0$ ein unbekannter Parameter ist. Wir wollen den Parameter α aus einer Stichprobe schätzen.

a) Bestimme die Likelihood- und die log-Likelihood-Funktion basierend auf n unabhängigen identisch verteilten Beobachtungen $x_1, \ldots, x_n$ einer Zufallsvariablen mit obiger Dichte.

b) Bestimme den zugehörigen Maximum-Likelihood-Schätzer für α. Schreibe zuerst die allgemeine Formel für n Beobachtungen hin und berechne den Schätzer dann für die folgende konkrete Stichprobe:

x_1	x_2	x_3	x_4	x_5
12.0	4.0	6.9	27.9	15.4

c) Bestimme den Momentenschätzer für α, wieder zuerst allgemein basierend auf n unabhängigen Beobachtungen $x_1, \ldots, x_n$ und dann für obige Stichprobe.

Hinweis: Der Momentenschätzer setzt voraus, dass $\alpha > 1$ ist; warum?

d) Vergleiche den Maximum-Likelihood- und den Momentenschätzer für obige Stichprobe. Ist der Momentenschätzer hier sinnvoll?

2. Eine Tankstelle veranschlagt für einen Ölwechsel mindestens α Minuten. Die tatsächlich benötigte Zeit X variiert natürlich im Bereich $X \geq \alpha$ und ist von Kunde zu Kunde verschieden. Man kann jedoch annehmen, dass diese Zeit durch eine exponentialverteilte Zufallsvariable gut wiedergegeben wird. Die Zufallsvariable X besitze somit die Dichtefunktion

$$f_X(x) := \begin{cases} e^{\alpha - x} & x \geq \alpha, \\ 0 & \text{sonst,} \end{cases}$$

d.h. $X = \alpha + Z$, wobei $Z \sim \mathrm{Exp}(1)$. Um α schätzen zu können, wurde von 10 zufällig ausgewählten Kunden die für den Ölwechsel benötigte Arbeitszeit in Minuten notiert:

4.2 3.1 3.6 4.5 5.1 7.6 4.4 3.5 3.8 4.3,

woraus sich der empirische Mittelwert von $\overline{x}_{10} = 4.41$ ergibt. Bestimme (zunächst allgemein[1]) den

a) Maximum-Likelihood Schätzer,

[1] Basierend auf n unabhängigen Beobachtungen von X.

Bitte wenden!

b) Momentenschätzer

für α und werte diesen aus für die angegebene Stichprobe.

3. Wir betrachten die geometrische Verteilung. Zur Erinnerung: dies ist eine diskrete Verteilung mit Wahrscheinlichkeitsfunktion

$$P(X = x) = (1 - p)^{x-1} p, \quad x = 1, 2, 3, \ldots$$

wobei $0 < p < 1$ die Erfolgswahrscheinlichkeit ist. Eine geometrisch verteilte Zufallsvariable beschreibt die Anzahl Versuche bis zum ersten Erfolg bei unabhängigen Bernoulliversuchen mit Erfolgswahrscheinlichkeit p. Wir wollen den Parameter p aus einer Stichprobe schätzen.

a) Bestimme die Likelihood-Funktion basierend auf n unabhängigen identisch verteilten Beobachtungen $x_1, \ldots, x_n$ einer geometrisch verteilten Zufallsvariablen.

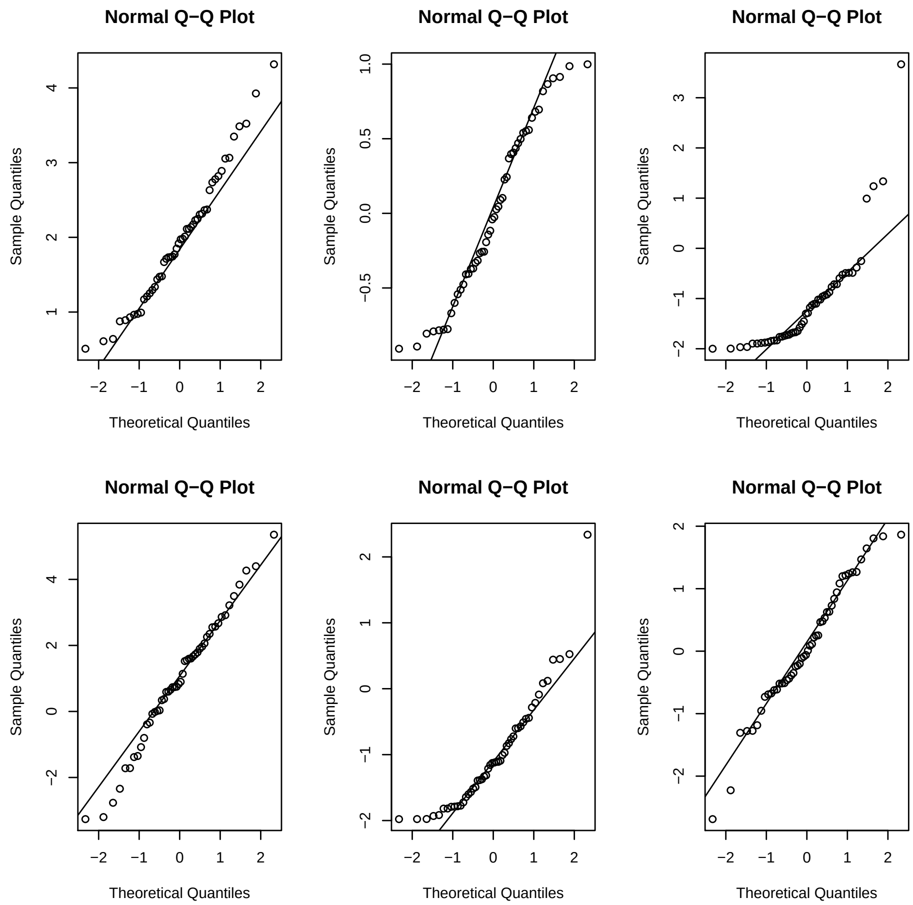
b) Bestimme den zugehörigen Maximum-Likelihood-Schätzer für p.

c) Bestimme den Momentenschätzer für p (wieder basierend auf n unabhängigen Beobachtungen $x_1, \ldots, x_n$). Vergleiche mit dem Maximum-Likelihood-Schätzer.

d) Angenommen du hast nur eine einzige Beobachtung x einer geometrisch verteilten Zufallsvariablen. Wir können das Experiment dann auch folgendermassen interpretieren: es wurden x unabhängige Bernoulliversuche durchgeführt und dabei wurde genau ein Erfolg beobachtet. Schreibe die Likelihood-Funktion für dieses Experiment auf: was ist der Unterschied zu der in a) gefundenen Likelihood-Funktion (für $n = 1$)? Warum erhälst du den gleichen Maximum-Likelihood-Schätzer?

4. Die folgenden sechs Figuren stellen Quantil-Quantil-Plots mit Stichprobenumfang 50 verschiedener Verteilungen gegen die Normalverteilung dar. Es wurden also 50 Datenpunkte entsprechend einer gegebenen Verteilung simuliert (dies für sechs Verteilungen) und jeder dieser sechs Datensätze wird nun in einem Normalplot eingezeichnet.

Siehe nächstes Blatt!



Folgende Verteilungen wurden simuliert: Je einmal $\mathcal{N}(0,1)$, $\mathcal{N}(2,1)$, $\mathcal{N}\left(1,2^2\right)$, $\text{Uni}[-1,1]$ sowie zweimal die Verteilung von $Y := Z - 2$ wobei $Z \sim \text{Exp}(1)$.

Bestimme bei jeder Figur welche Verteilung simuliert wurde und begründe deine Entscheidung.

Abgabe: Donnerstag 13. November