Parameterschätzungen

Frage:
Wie kann man aus fehlerbehafteten Messungen sinnvolle Aussagen und Resultate extrahieren?
Unterscheidung der Fehlerquellen:

- Zufällige Fluktuationen
- Systematische Fehler (Methodenbedingt, Apparaturabhängig)

Die zufälligen Fluktuationen der Messungen folgen Wahrscheinlichkeistdichtenverteilung, die man a priori nicht kennt.
Eine Messung von n unabhängigen Zufallsvariablen $\{x_1, x_2, \ldots, x_n\}$ wird als Stichprobe bezeichnet.
Schätzung eines Parameters basierend auf dieser Stichprobe.
Schätzung ist selbst eine Zufallsgröß.

Eigenschaften einer Schätzung

Seit $\hat{\Theta}$ die Schätzung des wahren Wertes Θ_0.

1. Konstsitenz: $\lim\limits_{n\to} \hat{\Theta} = \Theta_0$
2. Erwartungstreue: $\langle \hat{\Theta} \rangle = \Theta_0$
3. Effizient: Varianz $\hat{\Theta}$ soll möglichst klein sein
4. Robustheit: Weitgehend stabile hinsichtlich falscher Annahmen oder fehlerhafter Daten.

Die letzten beiden Eigenschaften sind schwer zu vereinbaren.

Beispiel

Schätzung des Mittelwerts über das arithmetische Mittel

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (1)$$

1. Konsistenz: Folge des Gesetzes großer Zahlen
2. Erwartungstreue:

$$\langle\overline{x}\rangle = \langle\frac{1}{n}\sum_{i=1}^{n} x_i\rangle \qquad (2)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\langle x_i\rangle = \frac{1}{n} n\mu = \mu \qquad (3)$$

3. und 4. Robsustheit und Effizienz.
Das arihtm. Mittel ist für keine symmetrische Verteilung, mit der Ausnahme der Noramlverteilung, eine effizient oder robuste Schätzung.

Bessere Schätzung

Angepasster Mittelwert:
Hier werden die $(1-2r)n/2$ größten und kleinsten Werte aus der Mittelwertbildung entfernt.

$$r \to \frac{1}{2} \Rightarrow \text{arithmetisches Mittel} \quad (4)$$
$$r \to 0 \Rightarrow \text{Median} \quad (5)$$

Für eine unbekannte symmetr. Verteilung liefert der angepasste MW mit r=0.23 eine robuste und effiziente Schätzung des MW.

Mittelwert einer Gleichverteilung

Beste Schätzung des MW ist

$$\overline{x} = \frac{\check{x} + \hat{x}}{2} \tag{6}$$

mit

$$\check{x} = \text{kleinster Stichprobenwert} \tag{7}$$

$$\hat{x} = \text{größter Stichprobenwert} \tag{8}$$

Schätzung hat kleinste Varianz

$$\sigma^2 = \frac{1}{n^2} \tag{9}$$

Maximum-Likelihood-Methode

Prinzip:
Zu einer ZV x liegen die Messungen $x_1, \ldots, x_n$ vor mit zugrundeliegender Wahrscheinlichkeitsdichte $(f(x|a)$, wobei a für die Parameter der W'keitsdichte steht.
Aufgabe: Bestimme beste Schätzung $\hat{a}$ für die Parameter zu den vorliegenden Daten.
Annahme:
Mehrdimensionale W'keitsdichte der gemessenen Werte definiert die Funktion

$$L(a) = f(x_1|a)f(x_2|a)\cdots f(x_n|a) = \prod_i f(x_i|a) \tag{10}$$

$L(a)$ ist die Likelihood-Funktion an und gibt für die Stichprobe $\{x_i\}$ als Funktion des Parameters a die W'keit an, bei vorgegebenem a die Meßwerte zu erhalten.

Die beste Schätzung entspricht der wahrscheinlichsten Schätzung, und damit dem Maximum der Likelihoodfunktion.
Eine Wichtige Annahme ist dass die Wahrscheinlichkeitsdichten normiert sind

$$\int_{\Omega} f(x|a) = 1 \quad \forall a \tag{11}$$

Maximum ist definiert über

$$\frac{dL}{da} = 0 \tag{12}$$

Praktisch definiert man die Loglikelihood-Funktion

$$l(a) = \ln L(A) = \ln \prod_{i} f(x_i|a) = \sum_{i}^{n} \ln f(x_i|a) \tag{13}$$

Hiermit ist das Maximum

$$\frac{dl}{da} = 0 \tag{14}$$

Alternative Formulierung:
Negative-Loglikelihood-Funtkion

$$F(a) = -l(a) = -\sum_{i=1}^{n} \ln f(x_i|a) \tag{15}$$

Problem:
Man muss die W'keitsdichten kennen.

Beispiel

Teilchenquelle erzeugt 100 Teilchen, und der Teilchenstrahl passiert eine Kaskade von 3 Detektoren im Abstand von 2,3 und 4 m zur Quelle.

- 1. Detektor werden 8 Teilchen gemessen
- 2. Detektor werden 3 Teilchen gemessen
- 3. Detektor werden 1 Teilchen gemessen

Über Maximum-Likelihood-Methode kann der Reichweitenparameter λ abgeschätzt werden. Zurgundeliegende W'keitsdichte ist die Poissonverteilung

$$p(n|\nu) = \frac{\nu^n}{n!} e^{-\nu} \tag{16}$$

Das Zerfallsgesetz:

$$\nu(x) = \nu_0 e^{-\lambda x} \tag{17}$$

Definiere Loglikelihood-Funktion

$$F(a) = -L(a) = \sum_i^n \ln f(x_i|a) \tag{18}$$

Was ist in unserem Beispiel $f(x_i|a)$?

$$f(x_i|a) = p(n_i|\nu(x_i)) \tag{19}$$

wobei n_i die Zahl gemessener Teilchen und $\nu(x_i)$ die Zahl erwarteter Teilchen am Punkt x_i ist.

Fehler der MLM

Fall: 1 Parameter und asymptotischer Fall, d.h.
Stichprobenumfang $n \to \infty$
Entwickle die Loglikelihood-Funktion um das Minimum $\hat{a}$

$$F(a) = F(\hat{a}) + \frac{1}{2}\frac{d^2F}{da^2}(a-\hat{a})^2 + \ldots \tag{20}$$

Beachte $\frac{dF}{da}$ für $a = \hat{a}$ ist identisch Null.
Die Likelihood-Funktion ist dann

$$L(a) = \text{const} \cdot e^{-\frac{1}{2}\frac{d^2F}{da^2}(a-\hat{a})^2} \tag{21}$$

Annahme W'keistdichte entspricht einer Normalverteilung

$$L(a) = \text{const} \cdot e^{-a\frac{(a-\hat{a})^2}{2\sigma^2}} \tag{22}$$

$$\sigma = \left(\frac{d^2 F}{da^2} \right)^{-1/2} \tag{23}$$

Für den Fall, dass $F(a)$ nicht ohne weiteres entwickelt werden kann (asymmetrische Fehler) können die Fehler auch alternativ bestimmt werden

$$F(\hat{a} \pm r\sigma) = F(\hat{a}) + \frac{1}{2} r^2 \tag{24}$$

bzw. im asymetrischen Fall:

$$F(\hat{a} + \sigma_r) = F(\hat{a}) + \frac{1}{2} \tag{25}$$

$$F(\hat{a} - \sigma_l) = F(\hat{a}) + \frac{1}{2} \tag{26}$$