# Rechnernutzung in der Physik

Institut für Experimentelle Teilchenphysik
Institut für Theoretische Teilchenphysik
Interfakultatives Institut für Anwendungen in der Informatik

Prof. Dr. U. Husemann, Dr. V. Shtabovenko
Dr. A. Mildenberger, Dr. Th. Chwalek
http://comp.physik.kit.edu

WS2019/20 – Blatt 09
Prog. bis Di., 28.01.2020

---

## Übungen zur Likelihood-Methode

Auf diesem Übungsblatt werden Grundlagen zur Likelihood-Methode am Beispiel der in der Praxis häufig auftretenden Poisson-Verteilung studiert sowie eine Parameteranpassung mit Störparametern durchgeführt.

### Aufgabe 17: Poisson-Likelihood *

Die Datei `Poisson.dat` enthält neun Zahlen, die dem Ergebnis der Messung eines Poisson-Prozesses entsprechen, also z. B. die Anzahlen der Klicks eines Geiger-Müller-Zählrohrs in aufeinanderfolgenden, gleich langen Zeitintervallen. In dieser Aufgabe sollen die Veränderungen der negativen Log-Likelihood-Funktion studiert werden, die sich ergeben, wenn nacheinander zunächst das erste und dann alle weiteren Ergebnisse berücksichtigt werden.
Die Vorlagen `basicDistributions.C` bzw. `nlLPoisson.py` enthalten Implementierungen der Poisson-Verteilung. Schreiben Sie Sie ausgehend davon eine Funktion zur Berechnung der negativen Log-Likelihood-Funktion der Poisson-Verteilung,

$$-\ln L(\lambda\,|\vec{x}) = -\sum_{i=1}^{n} \log\left(P(x_i, ; \lambda)\right) .$$

Stellen Sie $-\ln L(\lambda\,|\vec{x})$ für $n = 1, \ldots, 9$ grafisch als Funktion des Parameters $\lambda$ dar.

*Hinweis:* Es empfiehlt sich, eine allgemeine Funktion `negLogL(`$\lambda$`, P, `$\vec{x}$`)` zu schreiben, die aus der Verteilungsdichte $P(x; \lambda)$ die negative log-Likelihood-Funktion berechnet. Vergleichen Sie mit dem Beispiel aus der Vorlesung im Script `nl_poisson.py`.
Bestimmen Sie das Minimum der negativen Log-Likelihood-Funktion für den letzten Fall, bei dem alle Messwerte berücksichtigt werden. Sie können dazu einfach die Position des Minimums der für die grafische Ausgabe genutzten Log-Likelihood-Werte bestimmen. Den Index des Minimum eines Arrays erhalten Sie in `python` mit `np.argmin()`; in `ROOT` können Sie `TF1::GetMinimumX()` nutzen.
Vergleichen Sie Ihr Ergebnis mit der "klassischen" Lösung, dem Mittelwert aller Eingabewerte.

### Aufgabe 18: Profile Likelihood *

Wenn nur wenige Beobachtungen einer Größe vorhanden sind, ist eine Anpassung an eine Histogrammdarstellung nicht möglich, und man führt daher eine Parameteranpassung mit der Likelihood-Methode durch. Bei Experimenten treten häufig Störparameter auf, die für das physikalische Ergebnis irrelevant sind, aber dessen Unsicherheit beeinflussen. Die Berechnung und Analyse der Proflile-Likelihood für den jeweils interessierenden physikalischen Parameter ist eine einfache Möglichkeit, die durch Störparameter verursachten Unsicherheiten zu berücksichtigen.
In der Datei `tau_mu.dat` finden Sie 150 Werte gemessener Lebensdauern von Myonen, die im oder in der Nähe eines Myondetektors gestoppt wurden. Die Zeitdifferenz zwischen dem Nachweis des

einfallenden Myons und dem Eintreffen des Elektrons aus dem Zerfall ist die Messgröße. Bedingt durch das Messverfahren sind nur Zeiten zwischen 1.0 $\mu$s und 11.5 $\mu$s erfasst. Als Störgröße bei den Messungen tritt ein Untergrund aus Zufallskoinzidenzen von Signalen auf, die im Bereich von $\tau_{min}$=1.0 $\mu$s bis $\tau_{max}$ =11.5 $\mu$s als flach verteilt angenommen werden können. Die Verteilung der gemessenen Lebensdauern in diesem Bereich ist also gegeben durch die Summe einer Exponentialverteilung, parametrisiert durch die Myonlebensdauer $\tau_\mu$, und einer Gleichverteilung mit einem relativen Anteil $f_b$ an der Gesamtanzahl gemessener Zeitdifferenzen als Störparameter.

**Aufgabe:** Bestimmen Sie aus den Daten die Profile-Likelihood des Parameters $\tau_\mu$ und stellen Sie sie grafisch dar. Geben Sie (durch Ablesen der Grafik) das 68 %-Konfidenzintervall für die aus den Messdaten bestimmte Myonlebensdauer an.

*Hinweise zur Vorgehensweise:*
Geben Sie zunächst die Form der Verteilungsdichte der Zeitdifferenzen an:

$$pdf(t) \;=\; (1-f_b)\,C_n \exp\left(-t/\tau_\mu\right) \;+\; f_b/(t_{max}-t_{min})$$

Beachten Sie, dass die Verteilungsdichte auf Eins normiert sein muss - bestimmen Sie dazu die Normierungskonstante $C_n$ für die Exponentialfunktion im Wertebereich der Messwerte.
Wenn Sie die Daten als Histogramm auftragen, können Sie die ungefähren Werte der Parameter $\tau_\mu$ und $f_b$ abschätzen. Der Untergrundanteil $f_b$ ist klein (0 % bis 30 %), und die Myonlebensdauer beträgt 2.2 $\mu$s.
Die Bestimmung der Profile-Likelihood für einen Parameter und einen Störparameter gelingt mit einfachsten Mitteln, Sie brauchen dafür keinen komplizierten Optimierungscode: Wählen Sie z. B. 200 Werte $\tau_i$ für die Lebensdauer $\tau_\mu$ zwischen 1.5 $\mu$s und 2.9 $\mu$s. Nutzen Sie zur Berechnung der negativen Log-Likelihood-Funktion $-\ln L(\tau, f_b|\vec{t})$ eine Funktion analog zu Aufg. 16. Minimieren Sie dann für jeden der Werte $\tau_i$ die negative Log-Likelihood-Funktion bzgl. des Parameters $f_b$, indem Sie z. B. 200 Untergrundwerte $f_j^{(i)}$ zwischen 0. und 0.3 wählen und das Minimum $f_{\text{min}}^{(i)}$ mit `np.argmin()` (`python`) bzw. `TF1::GetMinimumX()` (`ROOT`) bestimmen. So erhalten Sie die Werte der Profile-Likelihood $-\ln L_p = -\ln L(\tau_i, f_{\text{min}}^{(i)}|\vec{t})$, die Sie gegen die Werte $\tau_i$ grafisch auftragen können.

---

Hinweis: Mit dem Rechnernamen `fphctssh.physik.uni-karlsruhe.de` können Sie von überall aus mittels ssh/scp Programm per Netzwerk auf einen Poolrechner zugreifen.

---