

Einführung

Der Ausgang physikalischer Experimente ist in der Regel mit Unsicherheiten behaftet, das heißt, das Resultat ist unvorhersagbar, zufällig. Diese Unsicherheit kann zwei unterschiedliche Ursachen haben:

- eine **Unsicherheit im Messprozess**, die zu **Messfehlern** führt;
- der grundsätzlich **statistische Charakter von physikalischen Prozessen** (statistisches Verhalten in Vielteilchensystemen, zum Beispiel Molekülbewegung in Gasen, oder quantenmechanische Prozesse, die nur Wahrscheinlichkeitsaussagen zulassen).

Um physikalische Experimente interpretieren zu können, benötigt man deshalb statistische Methoden, die in dieser Vorlesung selektiv und auf einem einführenden Niveau behandelt werden. Beispiele für die Anwendung statistischer Methoden sind:

- Bestimmung von **Wahrscheinlichkeiten** für das Auftreten von Ereignissen, häufig als Funktion einer oder mehrere Variablen, für die man dann Wahrscheinlichkeitsverteilungen erhält.
- Bestimmung der **Unsicherheit einer Messgröße**. Die Angabe eines Messergebnisses ohne einen Messfehler ist sinnlos!

Beispiel: Die Messung der Lichtgeschwindigkeit zu $2.8 \cdot 10^8$ m/s ist konsistent mit dem festgelegten Wert $2.99792458 \cdot 10^8$ m/s, wenn der Fehler der Messung zum Beispiel zu etwa ± 0.2 abgeschätzt wird:

$$c = (2.8 \pm 0.2) \cdot 10^8 \text{ m/s}$$

Bei der Angabe

$$c = (2.8 \pm 0.01) \cdot 10^8 \text{ m/s}$$

wird man sich andererseits wundern müssen, ob das eine große Entdeckung ist oder ob eher Quellen von Unsicherheit unberücksichtigt geblieben sind.

Es gibt zwei unterschiedliche Quellen von Unsicherheiten in einem Messprozess:

- **statistische Fehler**, die in der Regel experimentell bestimmt werden können;

- **systematische Fehler**, zu deren Abschätzung häufig die Erfahrung eines guten Experimentators notwendig ist.
- Beurteilung der **Signifikanz von Messsignalen** basiert auf der Bestimmung der Messfehler (Beispiel: das Signal einer kosmischen Radioquelle über einem Hintergrundrauschen).
Die zu erwartende Signifikanz eines experimentellen Ergebnisses sollte bereits bei der **Vorbereitung des Experimentes** berücksichtigt werden. So könnte man zum Beispiel mit statistischen Methoden festlegen, welcher Anteil der Messzeit bei dem obigen Beispiel für die Messung des Hintergrundes verwendet werden soll. Solche Planungen sind natürlich besonders wichtig, wenn die Experimente sehr zeitaufwendig und/oder kostspielig sind.
- **Entscheidung über Modellhypothesen**, die die Daten beschreiben: wann kann eine Hypothese akzeptiert werden, wann sollte sie verworfen werden, in welchem Bereich liegen die Parameter eines Modells.
- **Ausgleichsrechnung**: statistisch korrekter Ausgleich von Messwerten, die ein System überbestimmen (mehr Messungen als freie Parameter). Beispiele sind die Anpassung von Modellen an Daten und Bestimmung von Modellparametern oder die Berücksichtigung von Zwangsbedingungen in der Rekonstruktion von Teilchenreaktionen aus gemessenen Viererimpulsen.
- Berechnung komplizierter Prozesse durch **Simulationen**: die sogenannte Monte-Carlo-Methode bedient sich dabei statistischer Methoden. Sie findet zum Beispiel Verwendung bei der Bestimmung der Nachweiswahrscheinlichkeit eines Detektors oder bei der Analyse von Produktionsabläufen, Vorratshaltung oder Finanzierungsmodellen in der Wirtschaft.
- **Entfaltung**: Rückrechnung einer “wahren” Verteilung aus einer gemessenen mit Berücksichtigung von Auflösungs- und Effizienz-Effekten.
- **Klassifizierung**: Einteilung von Ereignissen in Klassen auf der Basis der, im allgemeinen multivariaten, Messwerte. Es gibt Klassifikationsalgorithmen, die auf die Erkennung der richtigen Klasse eines Ereignisses trainiert werden können, wie zum Beispiel Neuronale Netze (**‘statistisches Lernen’**)

Bei der Analyse von Daten kann man in der Regel auf Statistik- und Datenanalyseprogramme auf Computern zurückgreifen. Die Anwendung solcher Programme setzt aber ein gutes Verständnis der statistischen Methodik und sorgfältige Analysen der jeweils vorliegenden Problematik voraus.