KIT-Fakultät für Physik
Institut für Experimentelle Teilchenphysik
Institut für Theorie der Kondensierten Materie
Prof. Dr. Ulrich Husemann (ETP)
Dr. Nils Faltermann (ETP)
Dr. Andreas Poenicke (TKM)

Praktikum: Computergestützte Datenauswertung

Sommersemester 2021

Übungsblatt Nr. 4 Bearbeitung bis: 02.06., 18 Uhr

Aufgabe 4.1: Kovarianzmatrix

Wir betrachten in dieser Aufgabe einen kleinen Datensatz bestehend aus verschiedenen unabhängigen Messungen von einer Gruppe verschiedener Messgrößen. Den Datensatz entnehmen Sie bitte aus der Vorlage `B4A1_start.ipynb`.

a) Die Daten liegen als `numpy` Array vor, wobei jede Reihe einer unabhängigen Messung und jede Spalte einer Messgröße entspricht. Berechnen Sie zunächst den Vektor der Erwartungswerte der verschiedenen Messgrößen über alle Messungen.

b) Berechnen Sie nun explizit die Kovarianzmatrix für den gesamten Datensatz ohne die Verwendung von `numpy.cov`. Überlegen Sie sich zunächst, wie die Kovarianzmatrix allgemein in höheren Dimensionen als Matrixgleichung definiert ist. Sie können die Kovarianzmatrix dann unter Zuhilfenahme von z.B. `numpy.matmul` oder dem Operator `@` relativ einfach berechnen.

Aufgabe 4.2: Korrelation von Bin-Inhalten

In dieser Aufgabe soll die Häufigkeitsverteilung eines Histogramms, sowie die Korrelation zwischen einzelnen Bins näher untersucht werden. Dazu gibt es eine Vorlage, welche Sie ergänzen können (`B4A2_start.ipynb`).

a) Füllen Sie als Experiment $N = 100$ in $[0, 1]$ gleichverteilten Zufallszahlen in ein Histogramm mit 5 Bins zwischen $[0, 1]$. Die Bins enthalten nun im Mittel je $N/5$ Einträge. Diese Häufigkeit N_i in den einzelnen Bins soll nun weiter untersucht werden: Wiederholen Sie dieses Experiment 10'000 mal und bilden Sie Arrays n_i der Einzelhäufigkeiten N_i. Erzeugen Sie damit die Histogramme der in Bin 1 und Bin 2 gefundenen Häufigkeiten n_1 bzw. n_2. Welche Verteilung der Bin-Inhalte erwarten Sie?

b) Zweidimensionale Histogramme sind eine anschauliche Methode, Korrelationen zwischen den Bin-Inhalten zu untersuchen. Stellen Sie entsprechend die Einträge, n_i, in den Bins 1 und 2 als zweidimensionales Histogramm dar (`matplotlib.pyplot.hist2d`). Lassen Sie sich die aus dem zweidimensionalen Histogramm bestimmten Korrelationskoeffizienten ausgeben. Hierzu können Sie z.B. die Funktion `hist2dstat` aus dem Paket `PhyPraKit` verwenden.

Aufgabe 4.3: Fehlerfortpflanzung

Wir wollen in dieser Aufgabe das Gesetz der Fehlerfortpflanzung für zwei verschiedene Fälle testen.

a) Erzeugen Sie zwei Datensätze von je 1000 normalverteilten Zufallszahlen x_i mit $(\mu_x, \sigma_x) = (1.5, 0.5)$ und y_i mit $(\mu_y, \sigma_y) = (0.6, 0.15)$. Bilden Sie die Summe von Paaren von Zufallszahlen, $v_i = x_i + y_i$, und stellen Sie die Summen v_i als Histogramm grafisch dar.

b) Berechnen Sie mit Hilfe des Fehlerfortpflanzungsgesetzes die Standardabweichung σ_v der v_i und zeichnen Sie zum Vergleich eine Normalverteilung mit $(\mu_v = \mu_x + \mu_y, \sigma_v)$ ein.

c) Bilden Sie nun das Verhältnis von je zwei der beiden Zahlen, $w_i = x_i/y_i$, und stellen Sie das Histogramm der Quotienten w_i grafisch dar. Berechnen Sie wieder mit Hilfe des Fehlerfortpflanzungsgesetzes die Standardabweichung σ_w der w_i und zeichnen Sie zum Vergleich eine Normalverteilung mit $(\mu_w = \mu_x/\mu_y, \sigma_w)$ ein. Vergleichen Sie den Mittelwert und die Standardabweichung der Normalverteilung mir den Werten, die Sie direkt aus dem Histogramm erhalten. Wie gut stimmen das Histogramm und die Normalverteilung überein?