Di, 20.05.08

Übung zur Vorlesung "Statistische Methoden der Datenanalyse"
H. Kolanoski, A. Schälicke – SS 2008

Übung 4 – Update

4.1 Neutrinoereignis

Mitglieder einer großen Kollaboration, die einen riesigen Proton-Zerfallsdetektor in einer Salzmine nahe Cleveland, Ohio, betreibt, wiesen 8 Neutrinos in ihrem Detektor in Koinzidenz mit der optischen Beobachtung der Supernova 1987A nach.

a) Wenn die mittlere Anzahl der nachgewiesenen Neutrinos zwei pro Tag beträgt, wie groß ist dann die Wahrscheinlichkeit, eine Fluktuation von acht oder mehr Neutrinos pro Tag zu beobachten?

b) Tatsächlich wurden die acht Neutrinos innerhalb eines zehnminütigen Intervalls beobachtet. Wie groß ist die Wahrscheinlichkeit, eine Fluktuation von acht oder mehr Neutrinos in einem zehnminütigen Intervall zu beobachten?

4.2 Bivariate Gauss-Verteilung

Die bivariate Gauss-Verteilung ist durch die Wahrscheinlichkeitsdichte

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{(x-\mu_x)^2}{2{\sigma_x}^2} - \frac{(y-\mu_y)^2}{2{\sigma_y}^2}\right) \tag{1}$$

beschrieben.

a) Überprüfen Sie die Normierung durch analytische Integration.
Hinweis: Eine Transformation in Polarkoordinaten und Variablensubstitution führt auf ein einfach zu lösendes Integral.

b) Bestimmen Sie analog zu a) die Wahrscheinlichkeit, Ereignisse innerhalb der 1σ-, 2σ- und 3σ-Ellipse zu finden.

c) Generieren Sie Zufallsereignisse (x, y) für eine bivariate Normal-Verteilung, d.h. für $\mu_{x,y} = 0$ und $\sigma_{x,y} = 1$.

d) Überprüfen Sie die Ergebnisse aus b) experimentell mit den in c) erzeugten Zufallszahlen.

b.w.

4.3 Korrelationsmatrix

Gegeben ist Stichprobe Ω_f einer zweidimensionalen Wahrscheinlichkeitsdichte $f(x, y)$ (Datei `test.dat` auf der Webseite der Übung).

a) Stellen Sie einige Ereignisse (z.B. die ersten 100) graphisch dar.

b) Bestimmen Sie die Mittelwerte μ_x und μ_y sowie alle Elemente der Kovarianzmatrix v_{ij}.

c) Ermitteln Sie aus den Ergebnissen von b) Erwartungswert μ_g und Varianz σ_g^2 der Funktion $g(x, y) = x \cdot y$.

d) Bestimmen Sie die orthogonale Transformation U, für die $U^{-1}VU$ diagonal wird.

e) Verwenden Sie U um selbst Zufallszahlen mit den in b) bestimmten Eigenschaften zu erzeugen. Vergleichen Sie das Ergebnis mit der Stichprobe.

Hinweise:

- Die Poisson-Verteilung ist in Python als
  `stats.poisson.pmf(k,lam)`
  bereit gestellt (alternativ kann man auch `factorial(k)` für eine eigene Definition verwenden).

- Normalverteilte Zufallszahlen können mit dem *SciPy* Modul durch
  `x=stats.norm.rvs(size=10)`
  erzeugt werden. Weiter Informationen zur Verwendung erhält man (wie üblich) mit
  `stats.norm?`

- **Achtung:** Einlesen einer ASCII-Datendatei geschieht am einfachsten durch:
  ```
  import pylab
  x,y = pylab.load('test.dat',unpack=True)
  ```
  oder auch direkt mit der komprimierten Datei
  ```
  x,y = pylab.load('test.dat.gz',unpack=True)
  ```
  In modernen Pythoninstallationen könnte auch `numpy.loadtxt` *verwendet werden.*

- Darstellung der ersten 100 Ereignisse als *Scatter-Plot* erfolgt mittels:
  `scatter(x[:100],y[:100])`

- Für die Berechnung von Varianz und Kovarianz wird die explizite Verwendung von `sum(x)` und `len(x)` nahegelegt.
  Alternativ ist eine Verwendung der Befehle aus dem `scipy.stats` Modul möglich (`mean(x), var(x), cov(x,y)`). (Achtung: `stats.var` und `var` unterscheiden sich!)

- Weiter Hinweise zur Matrixmanipulationen findet man z.B. im Abschnitt A.2.1 *Array-Arithmetik* in pythonForStat.pdf auf der Webseite der Übung.