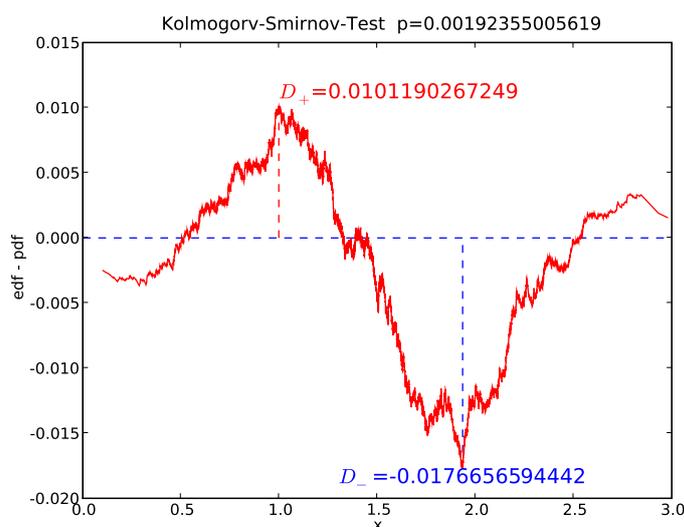


Übung zur Vorlesung “Statistische Methoden der Datenanalyse” H. Kolanoski, A. Schälicke – SS 2008

Übung 9

9.1 Kolmogorov-Smirnov-Test



- Erzeugen Sie näherungsweise Gauss-verteilte Zufallszahlen x_i aus der Summe dreier gleichverteilter Zufallszahlen z_i (z.B. $N=10000$ mal).
- Berechnen Sie Mittelwert und Varianz, und zeichnen Sie eine entsprechende Gauß-Verteilung.
- Bestimmen Sie die *empirische Verteilungsfunktion* (EDF, *empirical distribution function*) $S_{\text{EDF}}(x_i)$. Ordnen Sie dazu die x_i nach aufsteigenden Werten, und definieren Sie

$$S_{\text{EDF}}(x_i) = \frac{i}{N}. \quad (1)$$

Vergleichen Sie das Ergebnis mit der Verteilungsfunktion der Gauss-Verteilung (CDF, *cumulative distribution function*).

- Plotten Sie die Differenz zwischen der *empirischen Verteilungsfunktion* der Stichprobe und der Verteilungsfunktion der Gauss-Verteilung, und bestimmen Sie die maximale Abweichung

$$D = \sup |F(x) - S_{\text{EDF}}(x)|. \quad (2)$$

- Wie groß ist der p -Wert für den Kolmogorov-Smirnov-Test?

Hinweise zur Aufgabe 9.1

- Erzeugung von N Zufallszahlen aus der Summe von je m gleichverteilten Zufallszahlen geschieht am einfachsten durch

```
from numpy import *
N = 10000
m = 3
r = sum(rand(N,m),axis=1)
```

- Ein normiertes Histogramm mit vorgegeben Binning erzeugt man über

```
from pylab import *
bins=arange(0.,s,0.1)
hist(y,bins,normed=True)
```

- Wahrscheinlichkeitsdichte und Verteilungsfunktion der Gauß-Verteilung erhält man in Python aus:

```
from scipy import *

stats.norm.pdf(x,mu,sigma)    # probablity density function

stats.norm.cdf(x,mu,sigma)    # cumulative distribution function
```

- Ein *NumPy*-Array kann durch den Befehl `sort` geordnet werden:

```
x=sort(r)
```

- Maximum oder Minimum eines *NumPy*-Arrays kann man mit `max` bzw. `min` finden. Alternativ gibt es `argmax` und `argmin`, um den zugehörigen Index des Arrays zu erhalten.
- Den p -Wert für den Kolmogorov-Smirnov-Test erhält man aus

```
pval = stats.ksone.sf(d,N)
```

Sonstige Hinweise

- Vor der Verwendung von MINUIT, PYCUBA oder ROOT in Python müssen folgende Umgebungsvariablen gesetzt werden:

```
export PYTHONPATH=/users/eel/dreas/python/packages:/usr/global/root/lib
export LD_LIBRARY_PATH=/usr/global/root/lib:/users/eel/dreas/python/lib
```

Das geschieht am einfachsten in der Initialisierungsdatei der Shell, z.B. für `bash` in der Datei `.bashrc`.