

Statistische Methoden der Datenanalyse

Markus Schumacher, Stan Lai, Florian Kiss

Übung III

12.11.2013, 13.11.2013

Anwesenheitsaufgaben

Aufgabe 15 Zentraler Grenzwertsatz

Der Zentrale Grenzwertsatz besagt, dass die Summe von ausreichend vielen unabhängigen Zufallsvariablen durch eine Gaussverteilung beschrieben werden kann. Um den zentralen Grenzwertsatz zu demonstrieren, sollen in dieser Übung mehrere im Intervall $[0,1]$ gleichverteilte Zufallsvariablen aufaddiert werden, und mit einer Gausschen Verteilungsfunktion verglichen werden.

Erstellen Sie dazu ein ROOT-Skript, das folgende Komponenten enthält:

- (i) Einen Zufallszahlengenerator vom Typ `TRandom3`.
- (ii) Erzeugen Sie mittels dieses Zufallszahlengenerators gleichverteilte Zufallszahlen zwischen 0 und 1. Benutzen Sie eine `for`-Schleife, um dies genau `NVar`-mal zu machen (am besten erstellen Sie `NVar` als Ganzzahlvariable am Anfang des Makros, diese Zahl wird später noch benötigt).
- (iii) Füllen Sie die Summe dieser `NVar` Zufallszahlen in ein Histogramm mit geeigneten Binngrenzen. Benutzen Sie eine Ganzzahlvariable `int nbins`, um sich die Anzahl der Bins im Histogramm zu merken, da dies später noch benötigt wird.
- (iv) Schachteln Sie dieses Vorgehen in eine zweite `for`-Schleife, in der Sie das Zufallsexperiment `nexp`-mal wiederholen. Auch diese Zahl merken Sie sich, indem Sie eine Ganzzahlvariable am Anfang des Skriptes erstellen. Achten Sie darauf, den Zufallszahlengenerator und das Histogramm ausserhalb der `for`-Schleifen zu erstellen.
- (v) Zeichnen Sie das Histogramm auf dem Bildschirm.
- (vi) Zeichnen Sie zum Vergleich eine Gaussfunktion ein. Dies kann mit Hilfe der folgenden Befehle, die nach dem Zeichnen des Histogramms erledigt werden müssen, gemacht werden:

```
TF1 funktion = TF1("funktion",
  "[0]/([1]*sqrt(2*acos(-1)))*exp(-(x-[2])**2/(2*[1]**2))",0,NVar);
funktion.SetParameter(0,NVar*nexp/nbins);
funktion.SetParameter(1,sqrt(NVar/12.));
funktion.SetParameter(2,NVar/2.);
funktion.Draw("same");
```

Dabei ist `NVar` die Anzahl der in jedem Zufallsexperiment aufaddierten Variablen, `nexp` die Anzahl der durchgeführten Zufallsexperimente und `nbins` die Anzahl der Histogrammbins.

Starten Sie zunächst mit nur einer Zufallsvariablen (`NVar=1`). Erhöhen Sie nun die Anzahl schrittweise. Wieviele gleichverteilte Zufallsvariablen sollte man mindestens addieren, um eine gute Annäherung an die Gaussfunktion zu erhalten?

Zusatzaufgabe 16 *Produkt zweier gleichverteilter Zufallszahlen*

In Hausaufgabe 14 haben Sie gezeigt, dass die Wahrscheinlichkeitsdichtefunktion des Produktes zweier gleichverteilter, unabhängiger Größen x_1, x_2 zwischen 0 und 1 gegeben ist durch $-\ln z$, $z = x_1 x_2$.

- (i) Schreiben Sie ein ROOT-Makro, um zwei Zufallszahlen zu generieren, ihr Produkt zu bilden, in ein Histogramm zu füllen und es schliesslich auszugeben. Erzeugen Sie sich dazu einen Zufallszahlengenerator vom Typ `TRandom3`, erstellen innerhalb einer `for`-Schleife zwei Zufallszahlen (`Uniform()`) und füllen das Produkt in ein Histogramm.
- (ii) Normieren Sie das Histogramm auf die Gesamtanzahl der Einträge.
- (iii) Zeichnen Sie die erwartete Wahrscheinlichkeitsdichtefunktion ein, indem Sie die ROOT-Klasse `TF1` benutzen. Die dazu notwendige Syntax lautet

```
TF1 funktion=TF1("funktion", "-log(x)/[0]", 0, 1);  
funktion.SetParameter(0, nbins);  
funktion.Draw("same");
```

Beachten Sie, dass hierbei durch die Anzahl der Histogrammbins `nbins` zu dividieren ist.

Stimmt das Ergebnis mit der Erwartung überein?

Hausaufgaben

Aufgabe 17 Verkorkter Wein

5 Punkte

- (i) Sie fahren an die Mosel, um 50 Flaschen leckeren Riesling zu kaufen. Sie wissen aber, dass 4% der Flaschen leider verkorkt und der enthaltene Wein verdorben sein werden. Berechnen Sie die Wahrscheinlichkeiten $P(2)$ bzw. $P(3)$, dass 2 bzw. 3 der 50 Flaschen verkorkt sind.
- (ii) Ein Freund von Ihnen verzichtet törrichterweise auf Riesling, und zieht Liebfrauenmilch vor. Er bittet Sie 10 Flaschen Liebfrauenmilch mitzubringen. Die Wahrscheinlichkeit einer verkorkten Flasche für Liebfrauenmilch sei 20%. Was sind die Wahrscheinlichkeiten $P(2)$ bzw. $P(3)$, dass 2 bzw. 3 der 10 Flaschen verkorkt sind?
- (iii) Ihr Freund stellt Professor Schumacher die selbe Frage wie oben. Weil er keine Lust hat, alles genau zu berechnen, möchte er lieber eine Abschätzung durch die Poissonverteilung nutzen. Vergleichen Sie Professor Schumachers Abschätzung, mit den Antworten aus den Teilen (i) und (ii). Was beobachten Sie?

Aufgabe 18 Variablentransformation

5 Punkte

Betrachtet seien zwei Zufallsvariablen x_1 und x_2 mit zugehörigen Wahrscheinlichkeitsdichtefunktionen $f_1(x_1)$ und $f_2(x_2)$. Die beiden Zufallsvariablen seien unabhängig, das heißt für die kombinierte Wahrscheinlichkeitsdichteverteilung gilt:

$$f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$$

Zeigen Sie: Falls x_1 und x_2 in zwei neue Zufallsvariablen y_1 und y_2 transformiert werden, so dass $y_1 = y_1(x_1)$ und $y_2 = y_2(x_2)$ gilt, so sind auch die beiden neuen Zufallsvariablen unabhängig voneinander.

Aufgabe 19 Fourierfaltung

5 Punkte

Gegeben seien zwei unabhängige Zufallsvariablen x und y , die beide gleichverteilt sind zwischen 0 und 1, d.h. ihre Wahrscheinlichkeitsdichtefunktion ist gegeben durch

$$g(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{sonst} \end{cases}$$

und analog durch $h(y)$. Benutzen Sie die Fourier-Faltung, um zu zeigen, dass die Wahrscheinlichkeitsdichtefunktion $f(z)$ von $z = x + y$ gegeben ist durch:

$$f(z) = \begin{cases} z & 0 < z < 1 \\ 2 - z & 1 < z < 2 \\ 0 & \text{sonst.} \end{cases}$$

Aufgabe 20 Korrelierte Messgeräte

5 Punkte

Zwischen der an einem Leiter angelegten Spannung U , dem Widerstand R und dem fließenden Strom I besteht der Zusammenhang (Ohmsches Gesetz)

$$U = R \cdot I.$$

Sie haben ein Gerät, das den Strom und eines das den Widerstand R messen kann. Sie wissen allerdings, dass die Messungen des Strom und des Widerstands korreliert sind, da beide Geräte von der selben Firma hergestellt wurden. Der Korrelationskoeffizient betrage $\rho_{I,R} = 0.5$.

Drücken Sie die Abhängigkeit der Varianz der Spannung $V(U)$ durch die Varianzen $V(R)$, $V(I)$ aus.