5 FOLGEN STOCHASTISCHER GRÖSSEN

5.21 Gesetz der großen Zahlen

1. X_n sei eine binomialverteilte stochastische Größe, $X_n \sim B_{n,p}$. Gegen welche Konstanten konvergieren die folgenden stochastischen Größen fast sicher (für $n \to \infty$)?

   (a) X_n/n

   (b) $1 - X_n/n$

   (c) $(X_n/n)(1 - X_n/n)$

2. $X_1, X_2, \ldots$ sei ein UIV–Folge mit $X_i \sim f(x) = 6x(1-x)I_{(0,1)}(x)$.

   (a) Bestimmen Sie Mittelwert und Varianz von $\overline{X}_n = \sum_{i=1}^n X_i/n$. Wogegen konvergieren diese Größen für $n \to \infty$?

   (b) Wie lautet hier das Gesetz der großen Zahlen?

3. [R–Aufgabe] Eine Möglichkeit zur Bestimmung von π geht zurück auf Georges Buffon (1777): Zeichne ein Quadrat der Seitenlänge 2 (zentriert im Ursprung) und innerhalb des Quadrats zeichne einen Kreis mit Radius 1. Wähle zufällig einen Punkt im Quadrat; die Wahrscheinlichkeit, daß er innerhalb des Kreises liegt, ist $\pi/4$. Simulieren Sie diese Wahrscheinlichkeit, indem Sie z.B. $N = 1000$ derartige Punkte erzeugen und zählen, wieviele davon innerhalb des Kreises liegen; der Anteil dieser Punkte (multipliziert mit 4) ist ein Schätzwert für π. Veranschaulichen Sie die Schätzung durch Graphiken.

5.22 Zentraler Grenzverteilungssatz

*4. Zeigen Sie, daß für UIV–Folgen die Lindeberg–Bedingung erfüllt ist.

5. Für die Generierung von nach $N(0,1)$ verteilten Zufallszahlen gibt es auch die folgende approximative Methode: Man generiert 12 uniform (auf $(0,1)$) verteilte Zufallszahlen U_i und bildet:

$$\sum_{i=1}^{12} U_i - 6$$

   Dies liefert eine $N(0,1)$–verteilte Zufallszahl; für die Generierung einer weiteren nimmt man wieder 12 (neue) uniform verteilte Zufallszahlen, bildet den obigen Ausdruck, usw.

   (a) Geben Sie eine Erklärung für die Methode. (*Bem.:* Diese Methode kommt ohne die Invertierung der Verteilungsfunktion aus; allerdings braucht man für die Erzeugung *einer* normalverteilten Zahl 12 uniform verteilte Zahlen.)

   (b) [R–Aufgabe] Erzeugen Sie nach obiger Methode 500 normalverteilte Zufallszahlen mit Mittelwert 10 und Streuung 0.5 und stellen Sie das Ergebnis in Form eines Dichtehistogramms dar (mit zum Vergleich darüber gezeichneter theoretischer Dichte).

6. Ein Programm bestehe aus $n = 100$ Seiten Code und X_i sei die Anzahl der Fehler auf der i–ten Seite. Angenommen, die X_i's sind unabhängig und identisch poissonverteilt mit Mittel $\mu = 0.8$. Bestimmen Sie für die Gesamtzahl $Y = \sum_{i=1}^n X_i$ der Fehler mit Hilfe des zentralen Grenzverteilungssatzes einen approximativen Wert für $W\{75 < Y < 85\}$. Verwenden Sie bei der Berechnung die Stetigkeitskorrektur.

7. Wieviele (unabhängige) Versuche muß man durchführen, sodaß mit Wahrscheinlichkeit 0.8 das Ereignis A zumindest 5 Mal beobachtet wird, wenn die Wahrscheinlichkeit des Eintritts von A gleich 0.05 ist? Rechnen Sie mit Stetigkeitskorrektur.

8. [R–Aufgabe] Bestimmen Sie experimentell, ab welchem Wert von n der Stichprobenmittelwert $\overline{X}_n$ approximativ normalverteilt ist, wenn jedes X_i aus einer t–Verteilung mit 3 Freiheitsgraden stammt.
   [`clt-t.r`]

5.23 Fundamentalsatz der Statistik

9. Ermitteln und zeichnen Sie für die folgenden 20 Beobachtungen die empirische Verteilungsfunktion F_n^*.

0.10	−0.18	0.80	−0.80	0.58	0.14	−1.78	1.41	0.92	−0.28
−0.85	−0.63	2.17	−1.12	0.39	0.62	−0.69	0.24	−0.12	−0.16

10. Bestimmen Sie für die Daten des vorigen Beispiels (graphisch oder rechnerisch) die Stelle und den Wert des größten Abstands $\sup_x |F_n^*(x) - \Phi(x)|$ zwischen der empirischen Verteilungsfunktion und der Verteilungsfunktion der Standardnormalverteilung.

    [`dist.norm2.r`]

11. Während einer 6–monatigen Periode wurden die Zeiten (Stunden) zwischen aufeinander folgenden Crashes eines neu installierten Computersystems registriert. Die (der Größe nach geordneten) Zeiten waren wie folgt:

1	10	20	30	40	52	63	70	80	90	100
102	130	140	190	210	266	310	530	590	640	1340

   (a) Zeichnen Sie die empirische Verteilungsfunktion F_n^*.

   (b) Bestimmen Sie (graphisch oder rechnerisch) die Stelle und den Wert des größten Abstands zwischen der empirischen Verteilungsfunktion F_n^* und $\widehat{F}(x) = 1 - e^{-x/\overline{x}_n}$. Letztere ist die Verteilungsfunktion einer Exponentialverteilung, wobei τ durch das Stichprobenmittel $\overline{x}_n$ ersetzt wird. (Ist diese Ersetzung sinnvoll?)

   [`dist.exp2.r`]

5.24 Stichproben und Statistiken

12. Zeigen Sie, daß die Stichprobenvarianz S_n^2 auch wie folgt berechnet werden kann (*empirischer Verschiebungssatz*):

$$S_n^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right)$$

13. Zeigen Sie, daß die Stichprobenvarianz S_n^2 auch über die Differenzen der Beobachtungen berechnet werden kann:

$$S_n^2 = \frac{1}{2n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i - X_j)^2$$

14. $X_1, \dots, X_n$ sei eine Stichprobe aus einer Verteilung F mit Mittelwert μ und Varianz σ^2. Welche der folgenden Größen sind Statistiken?

$$(1)\ X_{(i)},\ i = 1, \dots, n \quad (2)\ \overline{X}_n \quad (3)\ S_n^2 \quad (4)\ \sqrt{n}\,\frac{\overline{X}_n - \mu}{S_n} \quad (5)\ F_n^*(x),\ x \in \mathbb{R} \quad (6)\ \sup_x |F_n^*(x) - F(x)|$$

15. Eine logarithmisch normalverteilte sG, $X \sim L(\mu, \sigma^2)$, wird n-mal unabhängig beobachtet, $X_1, \dots, X_n$.

   (a) Bestimmen Sie den Merkmalraum, den Parameterraum und den Stichprobenraum.

   (b) Ermitteln Sie die gemeinsame Dichtefunktion der Stichprobe.

   (c) Geben Sie geeignete Schätzfunktionen für μ und σ an.