Überblick über Tests für gleichverteilte (Pseudo-)Zufallszahlengeneratoren

1 Statistische Tests (Gleichverteilung)

- **χ^2-Test** (siehe Statistik-LV, Test auf Anzahl der Punkte in gewählten Partitionen)
- **Kolmogorov-Smirnov-Test** (siehe Statistik-LV, Test der empirischen Verteilungsfunktion gegen die wahre)

2 Empirische Tests (Gleichverteilung, Zufälligkeit)

Gleichverteilung und Zufälligkeit kann mit diesen Tests nur ausgeschlossen, nicht bewiesen werden. Daher müssen mehrere dieser Tests in Kombination durchgeführt werden, um ein gewisses Maß an Sicherheit zu erlangen.
Für die $U(0,1)$-verteilte Folge $\langle U_n \rangle$ definiere die ganzzahlige Hilfsfolge $\langle Y_n \rangle = \langle \lfloor dU_n \rfloor \rangle$, mit $d \in \mathbb{N}$ passend gewählt.

- **Frequenztest** (frequency test): $\langle U_n \rangle$ ist gleichverteilt (χ^2- oder KS-Test), ebenso jede Teilfolge (z.B. $\langle U_{2n} \rangle$, $\langle U_{n^2} \rangle$, etc.)
- **Serientest** (serial test): Wenn $\langle U_n \rangle$ gleichverteilt auf $[0,1]$, sind auch die Paare $\langle (Y_{2j}, Y_{2j+1}) \rangle$ gleichverteilt auf $\{0,\ldots,d-1\} \times \{0,\ldots,d-1\} \Rightarrow \chi^2$- oder KS-Test; verallgemeinerbar auf Tripel, etc.
- **Lückentest** (gap test): Betrachte Intervall $J = [\alpha, \beta] \subseteq [0,1]$ der Länge $p = \beta - \alpha$.

  Die ZV G mit $U_{k-1} \in J$, $U_k \notin J$, $U_{k+1} \notin J$, $\ldots$, $U_{k+G-1} \notin J$, $U_{k+g} \in J$ (Länge der Lücke, die nicht in J liegt) ist geometrisch verteilt mit Parameter $1-p \Rightarrow \chi^2$-Test auf geom. Verteilung

  Implementierung: Wähle $0 < h \in \mathbb{N}$ und zähle Anzahl der Lücken der Länge $0, 1, \ldots, h-1$ und $\geq h$, bis genügend Daten gesammelt. Die Wahrscheinlichkeiten für die einzelnen Längen betragen $p_r = p(1-p)^r$ für $0 \leq r < h$ und $p_{\geq h} = (1-p)^h$ für eine Länge $\geq h$.
- **Run test**: Betrachte Länge k von auf- oder absteigenden Folgen (Runs)

  $$x_{n-1} \geq x_n < x_n + 1 < \cdots < x_{n+k-1} \geq x_{n+k}\,.$$

  Die Längen von aufeinanderfolgenden Runs sind nicht unabhängig! Daher kann ein χ^2-Test nicht direkt durchgeführt werden.
- **Permutationstest** (permutation test): Die s-Tupel $(x_n, x_{n+1}, \ldots, x_{n+s-1})$ mit $s \geq 2$ und $0 \leq n < N$ haben $s!$ mögliche Ordnungen, die alle gleiche Wahrscheinlichkeit besitzen $\Rightarrow \chi^2$-Test auf diskrete Gleichverteilung

- **Test auf Serielle Korellation** (serial corellation): Bestimme die Korellation von $(U_0, U_1, \ldots, U_{n-1})$ mit den direkten Nachfolgern $(U_1, U_2, \ldots, U_n)$. Ebenso mit zyklischen Verschiebungen $(U_q, U_{q+1}, \ldots, U_{n-1}, U_0, U_1, \ldots, U_{q-1})$ mit $0 < q < n$.
- **Test auf beliebige Teilfolgen**: Jede beliebige Teilfolge einer gleichverteilten Zufallsfolge muss wieder gleichverteilt sein $\Rightarrow$ obige Tests auch auf beliebige Teilfolgen anwendbar.
- und viele weitere... (siehe etwa [Knu81, Fis96, Nie92])

3 Theoretische Tests

Mathematische Sätze, die aus der gegebenen Konstruktion des Pseudo-RNG bewiesen werden. Nur möglich, wenn Generator genau bekannt.

4 Spektraltest

Alle als gut bekannten Generatoren erfüllen Spektraltest, alle als schlecht bekannten versagen.
Idee:
Sei $x_0, x_1, \ldots$ eine periodische rationale Folge mit gemeinsamem Nenner m und Periode $per(\langle x - n \rangle) = T$. Betrachte die Exponentialsumme

$$\mathcal{E}(\mathbf{h}) = \frac{1}{T} \sum_{n=0}^{T-1} \exp\left(2\pi i \sum_{i=1}^{s} h_i x_{n+i-1}\right) .$$

Nach der Theorie der Exponentialsummen hat für exakt gleichverteilte (Punkte mit exakt gleichem Abstand) diese Summe den Wert

$$\mathcal{E}(\mathbf{h}) = \begin{cases} 1, & \text{für } \mathbf{h} \equiv 0 \mod m \\ 0, & \text{sonst} \end{cases}$$

Abweichung davon ist Maß für Nicht-Gleichverteilung der Punkte.
Theorie dazu: [Knu81]

Literatur

[Fis96] George S. Fishman. *Monte Carlo.* Springer Series in Operations Research. Springer-Verlag, New York, 1996. Concepts, algorithms, and applications.

[Knu81] Donald E. Knuth. *The art of computer programming. Vol. 2.* Addison-Wesley Publishing Co., Reading, Mass., second edition, 1981. Seminumerical algorithms, Addison-Wesley Series in Computer Science and Information Processing.

[Nie92] Harald Niederreiter. *Random number generation and quasi-Monte Carlo methods*, volume 63 of *CBMS-NSF Regional Conference Series in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.