4 Erzeugung spezieller Verteilungen

4.1 Erzeugung diskreter Zufallsvariablen

Ziel ist es, eine diskrete zufällige Variable X zu erzeugen:

$$X: \quad \begin{pmatrix} x_1 & x_2 & \ldots & x_m \\ p_1 & p_2 & \ldots & p_m \end{pmatrix}.$$

Dabei gelte $p_i \geq 0$ ($i = 1, \ldots, m$) und $\sum_{i=1}^{m} p_i = 1$.

Zerlegen das Intervall $[0, 1]$ in Teilintervalle I_j,

$$I_j = \left[\sum_{k=0}^{j-1} p_k, \sum_{k=0}^{j} p_k \right[, \qquad (p_0 = 0)$$

Sei u eine Pseudozufallszahl.

$$X = x_j \quad \text{falls} \quad u \in I_j$$

4.2 Erzeugung stetiger Zufallsvariablen

Ziel: Erzeugen $X \sim F)$.

Es sei $U \sim R(0,1)$ eine gleichverteilte Pseudozufallszahl mit Dichtefunktion:

$$h(u) = \begin{cases} 1 & \text{, falls } 0 \leq u < 1; \\ 0 & \text{, sonst.} \end{cases}$$

Wir betrachten nun eine Transformation $X := \varphi(U)$, wobei φ monoton wachsend sei. Die dabei entstehende Zufallsgröße X ist ebenfalls stetig verteilt und für ihre Dichte gilt (nach der Transformationsformel für Dichten (vgl. Abschnitt 7, Satz 2.22)):

$$f_X(x) = h\left(\varphi^{-1}(x)\right) \cdot \left|\frac{d\varphi^{-1}(x)}{dx}\right|.$$

Wir wählen nun $\varphi := F^{-1}$. Dann erhalten wir:

$$f_X(x) = h(F(x)) \cdot \frac{dF(x)}{dx} = f(x).$$

Damit besitzt die Zufallsgröße $X = F^{-1}(U)$ die gewünschte Verteilungsfunktion.

Erzeugung einer normalverteilten Zufallsvariablen

Ziel: $X \sim \mathcal{N}(0,1)$ erzeugen,

$$F(x) := \Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt.$$

Erzeugung einer solchen Zufallsgröße:

- Quantilmethode (siehe oben)
- Zentraler Grenzwertsatz
- Box-Müller Transformation

Quantilmethode

$U \sim R(0,1)$. $\quad X := \Phi^{-1}(u)$ ist standardnormalverteilt, denn es gilt:

$$f_X(x) = h(\Phi(x)) \cdot \frac{d\Phi(x)}{dx} = \frac{d\Phi(x)}{dx} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Problem: $\Phi^{-1}(u)$ ist etwas aufwendig zu berechnen.

Ziel: $X \sim \mathcal{N}(\mu, \sigma^2)$ erzeugen,

$$Y := \mu + \sigma \cdot \Phi^{-1}(U) \sim \mathrm{N}(\mu, \sigma^2).$$

Zentraler Grenzwertsatz (vgl. Satz 3.18, Seite 316).

$U_1, \ldots, U_n \sim R(0,1)$ unabhängig. Erwartungswert und Varianz sind

$$\mu \;:=\; EU_i = \int_0^1 x\,dx = \tfrac{1}{2}$$

$$\sigma^2 \;:=\; E\left(U_i - \tfrac{1}{2}\right)^2 = \tfrac{1}{12}$$

Nach dem zentralen Grenzwertsatz gilt:

$$\lim_{n\to\infty} P\left(\frac{\sum_{i=1}^n U_i - n\cdot\mu}{\sqrt{n}\cdot\sigma} < x\right) = \Phi(x).$$

Einsetzen:

$$\lim_{n\to\infty} P\left(\frac{\sum_{i=1}^n U_i - \frac{n}{2}}{\sqrt{\frac{n}{12}}} < x\right) = \Phi(x).$$

Definieren wir also eine Zufallsgröße

$$X := \frac{\sum_{i=1}^n U_i - \frac{n}{2}}{\sqrt{\frac{n}{12}}},$$

so ist diese für hinreichend großes n angenähert standardnormalverteilt.

Bsp. 4.6 *Es sei $n = 12$. Wir erhalten dann folgende Zufallsgröße X:*

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

Diese Approximation ist in der Regel ausreichend. Man braucht jedoch 12 Pseudozufallszahlen, um eine standardnormalverteilte Zufallsgröße zu erhalten.

Der Aufwand bei dieser Methode ist also ziemlich hoch.

Satz 4.5 (Box–Müller**–Transformation)** *Seien $U, V \sim R(0,1)$ unabhängig. Dann sind die Zufallsgrößen*

$$\begin{aligned} X &= \sqrt{-2 \cdot \ln U} \cdot \cos(2\pi V) \\ Y &= \sqrt{-2 \cdot \ln U} \cdot \sin(2\pi V) \end{aligned}$$

unabhängig und standardnormalverteilt, $X, Y \sim \mathrm{N}(0,1)$.

Beweis: vgl. Beispiel 2.35, Seite 264. $\square$

Erzeugung exponentialverteilter Zufallsvariablen

Es sei $U \sim R(0,1)$ eine Pseudozufallszahl. Erzeugt werden soll eine Zufallsgröße $X \sim \mathrm{EX}(\lambda)$ mit der Verteilungsfunktion:

$$F(x) = \begin{cases} 1 - e^{-\lambda \cdot x} & \text{, falls } x \geq 0; \\ 0 & \text{, sonst.} \end{cases}$$

Dazu wird folgende Transformation verwendet (vgl. Beispiel 2.25, Seite 230):

$$X := F^{-1}(U) = -\frac{1}{\lambda} \cdot \ln(1-u) \geq 0.$$

Erzeugung einer binomialverteilten Zufallsvariable

Variante 1: Seien $X_i \sim Bi(1,p)$. Dann ist $X = \sum_{i=1}^n X_i$ binomialverteilt mit Parametern (n,p).

Variante 2: (Intervallmethode)

Zerlegen das Intervall $(0,1)$ in disjunkte Teilintervalle der Länge p_k, $p_k = \binom{n}{k} p^k (1-p)^{n-k}$, der Einzelwahrscheinlichkeiten, etwa

$$
\begin{aligned}
(0,1) &= \bigcup_{i=0}^{n} I_i \\
&= (0,p_0] \cup (p_0, p_0+p_1] \cup (p_0+p_1, p_0+p_1+p_2] \cup \cdots \\
&\qquad \cup (1 - \sum_{i=0}^{n-1} p_i, 1)
\end{aligned}
$$

Sei $U \sim R(0,1)$.

$$X = i \quad \text{falls} \quad U \in I_i.$$

Erzeugung einer POISSON–Verteilten Zufallsvariable

Es ist jetzt eine POISSON–verteilte Zufallsgröße X zu erzeugen, d.h.

$$P(X=i)=\frac{\lambda^i}{i!}\cdot e^{-\lambda} \qquad (i=0,1,2,\ldots).$$

Variante 1: Intervallmethode

Variante 2: (Über die Exponentialverteilung)

Satz 4.6 *Es seien* $Y_1,\ldots,Y_k$ *unabhängige exponential-verteilte Zufallsgrößen und* $Y^{(k)} := \sum_{i=1}^{k} Y_i$*, Dann gilt für die Dichte der Zufallsvariable* $Y^{(k)}$*:*

$$f_{Y^{(k)}}(y)=\begin{cases} \frac{\lambda^k}{(k-1)!}\cdot y^{k-1}\cdot e^{-\lambda\cdot y} & \text{, falls } y\geq 0;\\ 0 & \text{, sonst.}\end{cases}$$

Diese Funktion ist die Dichte der sogen. ERLANG–Verteilung mit Parametern (k,λ).

Beweis: Wir beweisen die Aussage mittels vollständiger Induktion. Es sei $y \geq 0$.

IA: Da $Y^{(1)} = Y_1$ exponentialverteilt,

$$f_{Y^{(1)}}(y) = \lambda \cdot e^{-\lambda \cdot y}.$$

IV: Es sei die Aussage für k gültig.

IS: Wir zeigen sie für $k+1$. Es gilt:

$$Y^{(k+1)} = Y^{(k)} + Y_{k+1}.$$

Nun besitzt Y_{k+1} als exponentialverteilte Zufallsgröße dieselbe Dichtefunktion wie die zufällige Variable $Y^{(1)}$. Folglich können wir die Funktion $f_{Y^{(k+1)}}$ mittels Faltung der Dichtefunktionen $f_{Y^{(k)}}$ und $f_{Y^{(1)}}$ darstellen. Daher erhalten wir:

$$\begin{aligned}
f_{Y^{(k+1)}}(y) &= \int\limits_0^\infty f_{Y^{(k)}}(x) \cdot f_{Y^{(1)}}(y-x)\,dx \\
&= \int\limits_0^y \frac{\lambda^k}{(k-1)!} \cdot x^{k-1} \cdot e^{-\lambda\cdot x} \cdot \lambda \cdot e^{-\lambda\cdot(y-x)}\,dx \\
&= \int\limits_0^y \frac{\lambda^{k+1}}{(k-1)!} \cdot x^{k-1} \cdot e^{-\lambda\cdot y}\,dx \\
&= \frac{\lambda^{k+1}}{(k-1)!} \cdot e^{-\lambda\cdot y} \cdot \int\limits_0^y x^{k-1}\,dx \\
&= \frac{\lambda^{k+1}}{k!} \cdot y^k \cdot e^{-\lambda\cdot y}
\end{aligned}$$

$\square$

Satz 4.7 *Sind Y_i ($i \in \mathbb{N}$) unabhängige, exponentialverteilte Zufallsgrößen ($Y_i \sim \mathrm{EX}(\lambda)$, $i \in \mathbb{N}$), so ist die wie folgt definierte Zufallsvariable Y* Poisson*–verteilt mit Parameter λ:*

$$Y := \inf\left\{k\colon\ \sum_{i=1}^{k+1} Y_i > 1\right\} \sim \mathrm{PO}(\lambda).$$

Es gilt also:

$$P(Y = i) = \frac{\lambda^i}{i!} \cdot e^{-\lambda} \qquad (i = 1, 2, \ldots).$$

Beweis: Es gilt:

$$\begin{aligned}
P(Y = k) &= P(\sum_{i=1}^{k} Y_i \leq 1, \sum_{i=1}^{k+1} Y_i > 1) \\
&= P(\sum_{i=1}^{k} Y_i \leq 1, Y_{k+1} > 1 - \sum_{i=1}^{k} Y_i) \\
&= \int_0^1 P(Y_{k+1} > 1 - T | T = t) f_T(t)\, dt \\
&= \int_0^1 P(Y_{k+1} > 1 - t) f_T(t)\, dt \\
&= \int_0^1 e^{-\lambda(1-t)} \cdot \frac{\lambda^k}{(k-1)!} t^{k-1} e^{-\lambda t}\, dt \\
&= e^{-\lambda} \lambda^k \int_0^1 \frac{t^{k-1}}{(k-1)!}\, dt \\
&= e^{-\lambda} \frac{\lambda^k}{k!},
\end{aligned}$$

wobei $T = Y^{(k)} = \sum_{i=1}^{k} Y_i$ Erlang-verteilt ist. $\square$

Erzeugung einer geometrisch verteilten Zufallsvariable

Variante 1: Zur Erzeugung einer geometrisch verteilten Zufallsvariablen $X \sim Geo(p)$ seien $Y_i \sim Bi(1,p)$ Bernoulli verteilte Zufallsvariablen und

$$X = \min\{n : Y_n = 1\}$$

Variante 2: Sei $Y \sim Exp(\lambda)$, d.h. $F(y) = 1 - e^{-\lambda y}$. Die Zufallsvariable $\lfloor Y \rfloor$ ist geometrisch verteilt mit $p = 1 - e^{-\lambda}$.

Beweis: Es gilt:

$$\begin{aligned}
P(\lfloor Y \rfloor = k) &= P(k \leq Y < k+1) \\
&= F(k+1) - F(k) \\
&= (1 - e^{-\lambda(k+1)}) - (1 - e^{-\lambda k}) \\
&= e^{-\lambda k}(1 - e^{-\lambda}) = (1-p)^k p
\end{aligned}$$

$\square$