| **Ue Statistik u. Wahrscheinlichkeitsth. f. Inf.** | Di 12:00 – 18:00 |
|---|---|
| 107.251 W 2004\|5<br>http://www.statistik.tuwien.ac.at/RISueb/ | **7.Blatt** |
| Werner GURKER Tel.: 58801–10724<br>Spr.: Di u. Do 11–12 e–mail: W.Gurker@tuwien.ac.at | 7. Dez. 2004 |



1. $X$ sei eine geometrisch verteilte sG mit $p = 1/2$. Ermitteln Sie die Wahrscheinlichkleitsfunktion $p_Z(z)$ von $Z = (X-2)^2$. Weisen Sie nach, daß:

$$\sum_{z \in M_Z} p_Z(z) = 1$$

2. Ermitteln Sie die Dichte $f_Z(z)$ von $Z = |X|$ für $X \sim N(0,1)$. Beachten Sie dabei, daß $|x|$ auf $\mathbb{R}$ keine umkehrbar eindeutige Funktion ist und daher *Satz 14.2* aus der Vorlesung nicht (unmittelbar) anwendbar ist. Stellen Sie $f_Z(z)$ graphisch dar.

3. (a) Bestimmen Sie die Konstante $c$ so, daß die folgende Funktion eine Dichte ist:

$$f(x) = c\,x^2\,e^{-x^2}, \;\; 0 < x < \infty \quad (f(x) = 0 \text{ sonst})$$

   (b) Es gelte $X \sim f(x)$; ermitteln Sie die Dichte von $Z = X^2$. Verwenden Sie dabei den Transformationssatz für Dichten (*Satz 14.2* aus der Vorlesung). Welche Verteilung hat $Z$ ?

4. $X$ sei eine stetige sG mit Dichte $f(x)$, $x \in \mathbb{R}$. Ermitteln Sie einen allgemeinen Ausdruck für die Dichte von $Z = X^2$. Überprüfen Sie die Gültigkeit des Ausdrucks an Hand des Beispiels aus der Vorlesung mit $X \sim N(0,1)$.

5. Erläutern Sie, wie man Beobachtungen mit den Verteilungen von **Bsp 5-3** erzeugen kann. Generieren Sie jeweils 5 Beobachtungen für (a) und (b) auf Basis der folgenden uniform verteilten Zufallszahlen (mit R erzeugt: `runif(5)`):

0.1184 0.9029 0.6623 0.5566 0.4964

6. (a) Die Methode der Inversion der Verteilungsfunktion zur Simulation von Beobachtungen von sGn läßt sich auch bei diskreten (oder gemischten) Größen anwenden. Dabei stützt man sich auf eine verallgemeinerte Definition der Inversen einer Verteilungsfunktion:

$$F^{-1}(u) := \inf\{x \,:\, F(x) \geq u\}, \quad 0 < u < 1$$

   Erläutern Sie graphisch, was diese Definition praktisch bedeutet.

   (b) Ein Algorithmus zur Erzeugung von binomialverteilten Beobachtungen, $X \sim B_{n,p}$, lautet wie folgt:

   (1) Erzeuge eine (auf $(0,1)$) uniform verteilte Zahl $U$;

   (2) $c = \dfrac{p}{1-p}$, $i = 0$, $\alpha = (1-p)^n$, $F = \alpha$;

   (3) Wenn $U \leq F$, setze $X = i$, stop;

   (4) $\alpha = c\,\dfrac{n-i}{i+1}\,\alpha$, $F = F + \alpha$, $i = i + 1$;

   (5) Gehe zu (3).

Zeigen Sie, daß dieser Algorithmus tatsächlich Beobachtungen einer Binomialverteilung mit den Parametern $n$ und $p$ generiert. *Wie oft wird (3) im Mittel zur Erzeugung einer Binomial–Beobachtung aufgerufen?

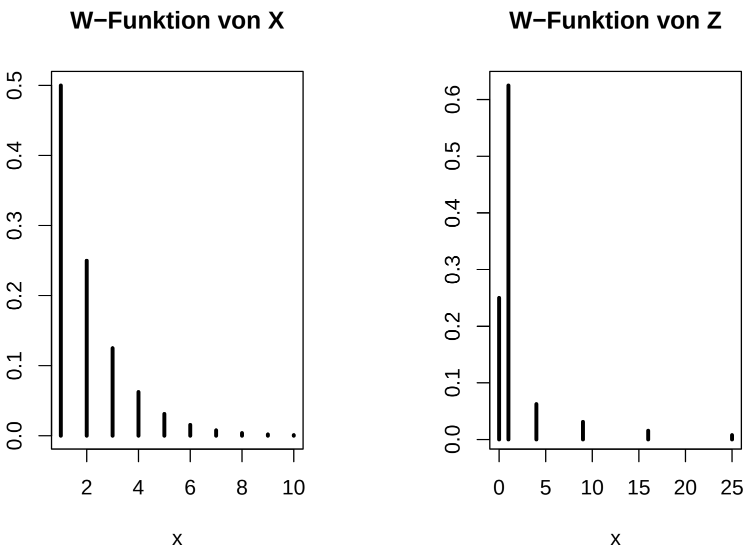
<u>Lösungen zum 7. Blatt</u>

1. Merkmalraum von $Z$: $\quad M_Z = \{0, 1, 4, 9, 16, \ldots\}$

$$Z = 0 \iff X = 2$$
$$Z = 1 \iff X = 1 \text{ oder } X = 3$$
$$Z = z \iff X = \sqrt{z} + 2 \text{ für } z = 4, 9, 16, \ldots$$

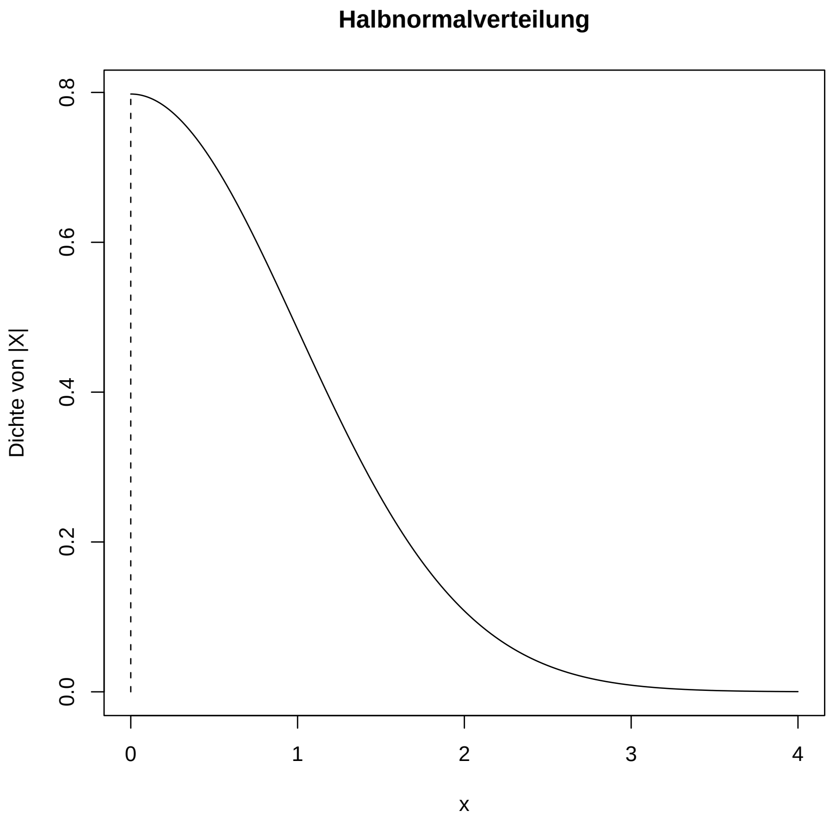
$$p_Z(z) = \begin{cases} p_X(2) = \dfrac{1}{4} & z = 0 \\[2ex] p_X(1) + p_X(3) = \dfrac{5}{8} & z = 1 \\[2ex] p_X(\sqrt{z}+2) = \dfrac{1}{4}\left(\dfrac{1}{2}\right)^{\sqrt{z}} & z = 4, 9, 16, \ldots \end{cases}$$

$$\sum_{z \in M_Z} p_Z(z) = \frac{1}{4} + \frac{5}{8} + \frac{1}{4}\sum_{k=2}^{\infty}\left(\frac{1}{2}\right)^k = \frac{7}{8} + \frac{1}{16}\underbrace{\sum_{k=0}^{\infty}\left(\frac{1}{2}\right)^k}_{=2} = 1$$




2. $F_{|X|}(x) = W\{|X| \le x\} = W\{-x \le X \le x\} = 2\Phi(x) - 1, \; x \ge 0$

$$f_{|X|}(x) = \frac{dF_{|X|}(x)}{dx} = 2\,\varphi(x) = \sqrt{\frac{2}{\pi}}\,e^{-x^2/2}, \; x \ge 0$$




3. (a) Substitution: $t = x^2, \quad x = \sqrt{t}, \quad \dfrac{dx}{dt} = \dfrac{1}{2\sqrt{t}}$

$$c \int_0^\infty x^2\, e^{-x^2}\, dx = \frac{c}{2} \int_0^\infty t^{1/2}\, e^{-t}\, dt = \frac{c}{2} \int_0^\infty t^{3/2-1}\, e^{-t}\, dt = \frac{c}{2}\, \Gamma\left(\frac{3}{2}\right) = \frac{c\sqrt{\pi}}{4}$$

$$\Longrightarrow \quad c = \frac{4}{\sqrt{\pi}}$$

(b) $z = x^2, \quad x = \sqrt{z}, \quad \dfrac{dx}{dz} = \dfrac{1}{2\sqrt{z}}$

$$f_Z(z) = f_X(\sqrt{z}) \left| \frac{dx}{dz} \right| = \frac{4\,z}{\sqrt{\pi}}\, e^{-z}\, \frac{1}{2\sqrt{z}} = \frac{2\sqrt{z}}{\sqrt{\pi}}\, e^{-z}, \; z > 0$$

Entspricht einer $Gam\left(\dfrac{3}{2}, 1\right)$–Verteilung.

4. $F_{X^2}(x) = W\{X^2 \leq x\} = W\{-\sqrt{x} \leq X \leq \sqrt{x}\} = F_X(\sqrt{x}) - F_X(-\sqrt{x}), \; x \geq 0$

$$f_{X^2}(x) = \frac{dF_{X^2}(x)}{dx} = \frac{1}{2\sqrt{x}} \left[ f_X(\sqrt{x}) + f_X(-\sqrt{x}) \right], \; x > 0$$

Beispiel: $X \sim N(0,1)$

$$f_{X^2}(x) = \frac{1}{2\sqrt{x}} \left[ \frac{1}{\sqrt{2\pi}}\, e^{-x/2} + \frac{1}{\sqrt{2\pi}}\, e^{-x/2} \right] = \frac{1}{\sqrt{2\pi x}}\, e^{-x/2} = \frac{x^{1/2-1}\, e^{-x/2}}{\sqrt{\pi}\, 2^{1/2}}, \; x > 0$$

$$\left( \widehat{=} \; Gam\left(\frac{1}{2}, 2\right) \equiv \chi_1^2 \right)$$

5. Dazu ist die Verteilungsfunktion zu invertieren:

   (a) $x = F^{-1}(u) = -\ln(-\ln u)$

| $u$ | 0.1184 | 0.9029 | 0.6623 | 0.5566 | 0.4964 |
|---|---|---|---|---|---|
| $x$ | −0.7579 | 2.2814 | 0.8866 | 0.5346 | 0.3561 |

   (b) $x = F^{-1}(u) = \tan\left[\pi\left(u - \frac{1}{2}\right)\right]$

| $u$ | 0.1184 | 0.9029 | 0.6623 | 0.5566 | 0.4964 |
|---|---|---|---|---|---|
| $x$ | −2.5633 | 3.1758 | 0.5592 | 0.1797 | −0.0113 |

   *Anm.:* Cauchy–Verteilung $\equiv$ $t_1$–Verteilung; R: `rt(5,1)`.

6. (b) Mit $X \sim p_X(x)$, $x \in \mathbb{Z}$, und $U \sim U_{(0,1)}$ gilt:

$$\begin{aligned} W\{F_X^{-1}(U) = k\} &= W\{F_X(k-1) < U \leq F_X(k)\} \\ &= F_X(k) - F_X(k-1) \\ &= W\{X = k\}, \quad k \in \mathbb{Z} \end{aligned}$$

Um eine Beobachtung von $X$ zu generieren nimmt man also die kleinste ganze Zahl $k$, sodaß $F_X(k) \geq u$, für eine vorher generierte uniform verteilte Zufallszahl $u$. Der angegebene Algorithmus realisiert dies für eine binomialverteilte sG. Zur schnelleren Berechnung von $F_X$ wird dabei eine Rekursion verwendet (Rekursionsanfang: $W\{X = 0\} = (1-p)^n$):

$$\begin{aligned} W\{X = i+1\} &= \binom{n}{i+1} p^{i+1}(1-p)^{n-i-1} \\ &= \frac{p}{1-p}\,\frac{n-i}{i+1}\binom{n}{i} p^i (1-p)^{n-i} \\ &= c\,\frac{n-i}{i+1}\,W\{X = i\}, \; i = 0, 1, 2, \ldots, n-1 \end{aligned}$$

*Schritt (3) wird genau dann $k$ Mal aufgerufen, wenn $U$ zwischen $F_X(k-2)$ und $F_X(k-1)$ liegt; die Wahrscheinlichkeit dafür ist $p_X(k-1)$. Der Erwartungswert für die Anzahl $Y$ der Aufrufe von (3) ist also:

$$\mathbb{E}(Y) = \sum_{k=1}^{\infty} k\,p_X(k-1) = \sum_{k=0}^{\infty}(k+1)\,p_X(k) = 1 + \sum_{k=0}^{\infty} k\,p_X(k) = 1 + np$$