

Vertiefung NWI: 11. Vorlesung zur Wahrscheinlichkeitstheorie

Mittwoch, 18.6.2014

11. Zufallsgrößen mit Dichten

Erinnerung

Der Satz von de Moivre-Laplace zeigt für eine mit Parametern n und p binomialverteilte Zufallsgröße S_n , daß

$$\mathbb{P}\left\{a < \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right\} \rightarrow \int_a^b \varphi(x) dx \quad \text{für } n \rightarrow \infty$$

mit $\varphi \geq 0$ und $\int_{-\infty}^{\infty} \varphi(x) dx = 1$.

Motivation

Für große n können wir also die Wahrscheinlichkeit, daß die ZV $(S_n - np)/\sqrt{np(1-p)}$ in einem Intervall $(a, b]$ liegt, durch ein Integral approximieren. Wir möchten nun allgemeiner Zufallsgrößen mit Werten im Kontinuum (z.B. $(-\infty, \infty)$, $[a, \infty)$, (a, ∞) , $[a, b]$, $[a, b)$, usw.) mit Hilfe solcher Integrale einführen.

Definition: Dichte

Eine Funktion $f : \mathbb{R} \rightarrow [0, \infty)$, die stückweise stetig ist, heißt *Dichte*, sofern

$$\int_{-\infty}^{+\infty} f(x) dx = 1$$

gilt.

Erläuterung

Für eine Funktion $g \geq 0$, die auf jedem kompakten Intervall $[-N, +N]$ Riemann-integrierbar ist, ist die Folge $\{\int_{-N}^{+N} g(x) dx\}_{N \in \mathbb{N}}$ monoton wachsend in N . Ist die Folge unbeschränkt, so setzen wir $\int_{-\infty}^{+\infty} g(x) dx = \infty$, andernfalls

$$\int_{-\infty}^{+\infty} g(x) dx = \lim_{N \rightarrow \infty} \int_{-N}^{+N} g(x) dx .$$

Beispiele von Dichten (vgl. Bilder im Anhang)

1. Dichte der Standardnormalverteilung, vgl. Satz von de Moivre-Laplace,

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} .$$

2. Dichte der Normalverteilung mit den Parametern $\mu \in \mathbb{R}$ und $\sigma > 0$
(Wir werden später sehen, daß $\mu \in \mathbb{R}$ die Rolle des Erwartungswerts und $\sigma^2 > 0$ jene der Varianz spielen.)

$$\varphi(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} .$$

Es gilt $\varphi(x; \mu, \sigma^2) \geq 0$ für alle x . Somit zeigt die Substitution $y = (x-\mu)/\sigma$, daß $x \mapsto \varphi(x; \mu, \sigma^2)$ eine Dichte ist: Für $N \rightarrow \infty$ gilt

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-N}^N e^{-(x-\mu)^2/2\sigma^2} dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{(-N-\mu)/\sigma}^{(N-\mu)/\sigma} e^{-y^2/2} \sigma dy \rightarrow \int_{-\infty}^{\infty} \varphi(y) dy = 1$$

3. Dichte der uniformen Verteilung auf $[a, b]$ für $-\infty < a < b < \infty$

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{für } x \in [a, b] \\ 0 & \text{sonst} \end{cases}$$

f ist Dichte, weil $f \geq 0$ und

$$\int_{-\infty}^{\infty} f(x) dx = \int_a^b \frac{1}{b-a} dx = 1$$

Beachte: f ist das Analogon zur Gleichverteilung im Fall diskreter Zufallsgrößen. Wie die Gleichverteilung kann auch die uniforme Verteilung nur auf beschränkten Mengen definiert werden.

4. Dichte der Exponentialverteilung zum Parameter $\alpha > 0$

$$f(x) = \begin{cases} 0 & \text{für } x < 0 \\ \alpha e^{-\alpha x} & \text{für } x \geq 0 \end{cases}$$

f ist Dichte, weil $f \geq 0$ und

$$\int_{-\infty}^{\infty} f(x) dx = \lim_{N \rightarrow \infty} \int_0^N \alpha e^{-\alpha x} dx = - \lim_{N \rightarrow \infty} \int_0^N \frac{d}{dx} (e^{-\alpha x}) dx = - \lim_{N \rightarrow \infty} e^{-\alpha x} \Big|_{x=0}^N = 1$$

Definition: Dichte und Verteilungsfunktion einer Zufallsgröße

f sei eine Dichte. Dann heißt eine Zufallsgröße X *absolutstetig mit Dichte f* , falls

$$F_X(x) := \mathbb{P}\{X \leq x\} = \int_{-\infty}^x f(y) dy \quad \forall x \in \mathbb{R}$$

Die Funktion F_X heißt *Verteilungsfunktion* von X .

Warnung

Angenommen, wir haben eine ZV X derart, daß

$$(*) \quad \mathbb{P}\{a < X \leq b\} = \int_a^b f(x) \, dx$$

für alle a, b mit $a < b$ (mit einer über ganz \mathbb{R} integrierbaren Funktion f). Dann gilt für jedes n

$$0 \leq \mathbb{P}\{X = a\} \leq \mathbb{P}\{a - 1/n < X \leq a\} = \int_{a-1/n}^a f(x) \, dx \rightarrow 0 \quad \text{für } n \rightarrow \infty$$

Das heißt, für jedes a gilt

$$\mathbb{P}\{X = a\} = 0.$$

Für diskrete Zufallsgröße kann somit keine Dichte f existieren, weil $(*)$ nicht für alle $a < b$ gelten kann!

Bemerkungen

- Im Rahmen dieser Vorlesung nehmen wir einige Einschränkungen in Kauf: Wir zeigen z.B. nicht, unter welchen Voraussetzungen eine Dichte existiert. Auch spezifizieren wir den Wahrscheinlichkeitsraum nicht. Hier gilt es zu beachten, daß $\mathbb{P}\{X \in A\}$ in der Regel nicht für *alle* Teilmengen von \mathbb{R} definiert werden kann. Uns soll genügen, daß alles gutgeht, wenn wir uns auf Mengen A beschränken, die durch höchstens abzählbare Vereinigungen und Schnitte von Intervallen dargestellt werden können.
- Vergleichen Sie die bekannten Eigenschaften der Verteilungsfunktion für diskrete Zufallsgrößen mit jenen der Verteilungsfunktion einer Zufallsgröße mit Dichte:
 - Verteilungsfunktionen sind monoton wachsend.
 - $\lim_{x \rightarrow -\infty} F_X(x) = 0$ und $\lim_{x \rightarrow \infty} F_X(x) = 1$
 - Falls die Zufallsgröße X eine Dichte hat, so ist die Verteilungsfunktion F_X stetig. Für diskrete Zufallsgrößen dagegen ist die Verteilungsfunktion stückweise konstant.

(Machen Sie zwei Skizzen!)

- Dichten sind nicht eindeutig bestimmt.
(Verschiedene Dichten ergeben sich, wenn wir ein Dichte in einzelnen Punkten ändern.)
- Ist die Dichte f stetig in einem Punkt a , so gilt

$$f(a) = \left. \frac{dF_X(y)}{dy} \right|_{y=a}$$

Dies erlaubt, aus der Verteilungsfunktion einer Zufallsgröße eine Dichte zu bestimmen.

Beispiele (vgl. Bilder im Anhang)

3. Verteilungsfunktion der uniformen Verteilung

$$F_X(x) = \begin{cases} 0 & \text{für } x < a \\ \int_a^x \frac{1}{b-a} dy = \frac{x-a}{b-a} & \text{für } x \in [a, b] \\ \int_{-\infty}^{\infty} f(y) dy = 1 & \text{für } x > b \end{cases}$$

4. Verteilungsfunktion der Exponentialverteilung

$$F_X(x) = \begin{cases} 0 & \text{für } x < 0 \\ \int_0^x \alpha e^{-\alpha y} dy = -e^{-\alpha y} \Big|_{y=0}^x = 1 - e^{-\alpha x} & \text{für } x \geq 0 \end{cases}$$

1. Verteilungsfunktion Φ der Standardnormalverteilung (siehe Tabelle)

2. Verteilungsfunktion Φ_{μ, σ^2} der Normalverteilung

$$\Phi_{\mu, \sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x e^{-(z-\mu)^2/2} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-\mu)/\sigma} e^{-y^2/2} dy = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

(mit der gleichen Substitution wie oben)

Sprechweise

Eine ZG heißt normalverteilt, uniform verteilt, exponentialverteilt, etc, falls sie eine entsprechende Dichte hat.

Rechenregeln

Hat X eine Dichte f , so gelten für alle $-\infty \leq a < b \leq \infty$

- $\mathbb{P}\{a < X < b\} = \mathbb{P}\{a \leq X \leq b\} = \mathbb{P}\{a \leq X < b\} = \mathbb{P}\{a < X \leq b\}$
- $\mathbb{P}\{a < X \leq b\} = \int_a^b f(x) dx = \mathbb{P}\{X \leq b\} - \mathbb{P}\{X \leq a\} = F_X(b) - F_X(a)$
- $\mathbb{P}\{X \leq b\} = F_X(b) = \int_{-\infty}^b f(x) dx$
- $\mathbb{P}\{a < X\} = \int_a^{\infty} f(x) dx = 1 - \int_{-\infty}^a f(x) dx = 1 - \mathbb{P}\{X \leq a\} = 1 - F_X(a)$

Beispiel

Berechnen der Wahrscheinlichkeit, daß eine exponentialverteilte Zufallsgröße mit Parameter α einen Wert größer als a annimmt bzw. einen Wert zwischen a und b .

- Direktes Ausrechnen des entsprechenden Integrals
- Erst Berechnen der Verteilungsfunktion, dann der Wahrscheinlichkeit

Definition: Erwartungswert und Varianz

Sei X eine Zufallsgröße mit einer Dichte f .

- Falls $\int_{-\infty}^{\infty} |x|f(x) dx < \infty$, so sagen wir, daß *der Erwartungswert von X existiert*. Er ist in diesem Fall gegeben durch

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x) dx = \lim_{N \rightarrow \infty} \int_{-N}^N xf(x) dx .$$

(*Warnung*: So dürfen wir nur rechnen, wenn wir bereits gezeigt haben, daß der Erwartungswert existiert, denn für Dichten f , die gerade Funktionen sind, gilt $\int_{-N}^N xf(x) dx = 0$ für alle N , auch wenn der Erwartungswert ggf. gar nicht existiert.)

- Falls der Erwartungswert von X existiert, definieren wir die *Varianz von X* als

$$\text{Var}(X) = \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2 f(x) dx .$$

Ist $\text{Var}(X) < \infty$, so sagen wir, daß die Varianz existiert.

Bemerkung

Vergleichen Sie mit dem Fall diskreter Zufallsgrößen.

Beispiele

3. Uniforme Verteilung (direktes Ausrechnen – warum ist der Erwartungswert $(a + b)/2$?)
Der Erwartungswert existiert, weil $f = 0$ außerhalb des kompakten Intervalls $[a, b]$.

$$\begin{aligned}\mathbb{E}[X] &= \int_a^b x \frac{1}{b-a} dx = \frac{1}{2(b-a)} x^2 \Big|_{x=a}^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2} \\ \mathbb{E}[X^2] &= \int_a^b x^2 \frac{1}{b-a} dx = \frac{1}{3(b-a)} x^3 \Big|_{x=a}^b = \frac{b^3 - a^3}{3(b-a)} = \frac{1}{3}(b^2 - ab + a^2) \\ \text{Var}(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{b^2 - ab + a^2}{3} - \frac{a^2 + 2ab + b^2}{4} = \frac{(a-b)^2}{12}\end{aligned}$$

4. Exponentialverteilung (verwende partielle Integration)

Der Erwartungswert existiert genau dann, wenn die erste Zeile der folgenden Rechnung einen endlichen Wert ergibt (was der Fall ist):

$$\begin{aligned}\mathbb{E}[X] &= \int_0^\infty \underbrace{x}_{\downarrow} \underbrace{\alpha e^{-\alpha x}}_{\uparrow} dx = - \lim_{N \rightarrow \infty} x e^{-\alpha x} \Big|_{x=0}^N + \int_0^\infty e^{-\alpha x} dx = 0 + \frac{1}{\alpha} \int_{-\infty}^\infty f(x) dx = \frac{1}{\alpha} \\ \mathbb{E}[X^2] &= \int_0^\infty \underbrace{x^2}_{\downarrow} \underbrace{\alpha e^{-\alpha x}}_{\uparrow} dx = - \lim_{N \rightarrow \infty} x^2 e^{-\alpha x} \Big|_{x=0}^N + 2 \int_0^\infty x e^{-\alpha x} dx = 0 + \frac{2}{\alpha} \mathbb{E}[X] = \frac{2}{\alpha^2} \\ \text{Var}(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{2}{\alpha^2} - \left(\frac{1}{\alpha}\right)^2 = \frac{1}{\alpha^2}\end{aligned}$$

1. Standardnormalverteilung

Die Existenz der Erwartungswerts folgt aus der folgenden Überlegung: Es existiert ein $x_0 > 0$ mit $x e^{-x^2/2} \leq e^{-x^2/4}$ für alle $x \geq x_0$. Daher gilt für alle $N \geq x_0$:

$$\int_{-N}^N |x| \varphi(x) dx \leq 2 \int_0^{x_0} x \varphi(x) dx + 2 \int_{x_0}^N \frac{1}{\sqrt{2\pi}} e^{-x^2/4} dx \leq 2x_0^2 + 2 \int_0^\infty e^{-x^2/4} dx < \infty$$

Wir berechnen nun Erwartungswert und Varianz:

$$\begin{aligned}\mathbb{E}[X] &= \int_{-\infty}^\infty x \varphi(x) dx \\ &= \lim_{N \rightarrow \infty} \left[\int_{-N}^0 x \varphi(x) dx + \int_0^N x \varphi(x) dx \right] = \lim_{N \rightarrow \infty} \left[- \int_0^N y \varphi(y) dy + \int_0^N x \varphi(x) dx \right] = 0 \\ \text{Var}(X) &= \mathbb{E}[X^2] \\ &= \int_{-\infty}^\infty \underbrace{x}_{\downarrow} \underbrace{x \varphi(x)}_{\uparrow} dx = \lim_{N \rightarrow \infty} \left[x(-\varphi(x)) \Big|_{x=-N}^N - \int_{-N}^N (-\varphi(x)) dx \right] = 0 + \int_{-\infty}^\infty \varphi(x) dx = 1\end{aligned}$$

2. Normalverteilung (verwende Substitution und Wissen über Dichte, Erwartungswert und Varianz der Standardnormalverteilung)

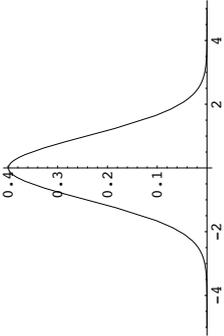
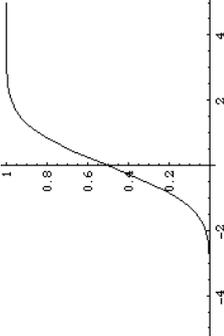
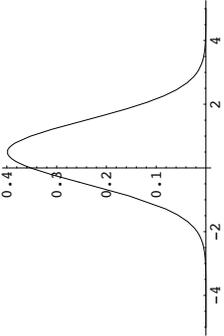
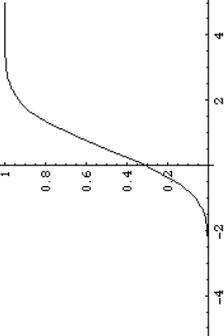
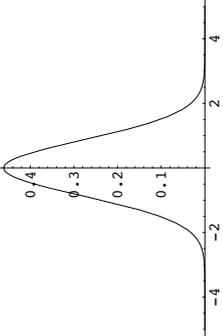
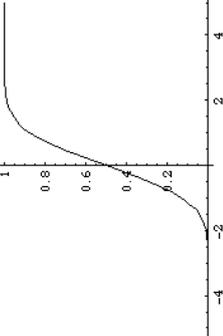
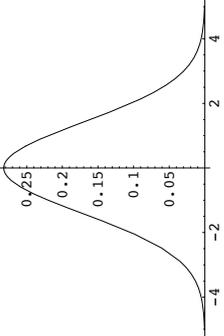
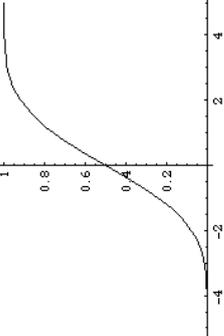
Existenz des Erwartungswerts: Analog zur obigen Rechnung für die Standardnormalverteilung

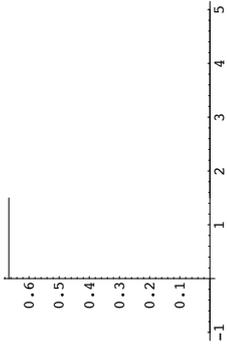
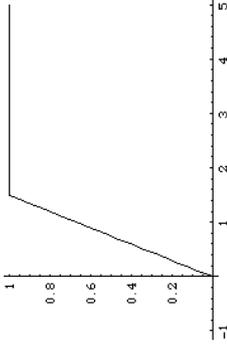
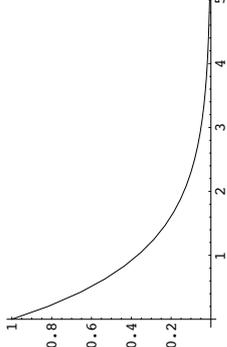
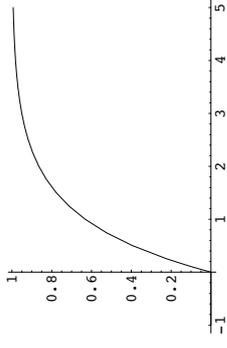
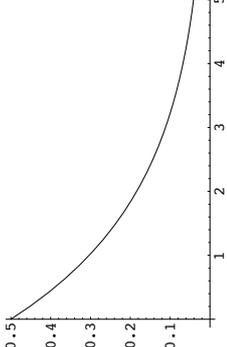
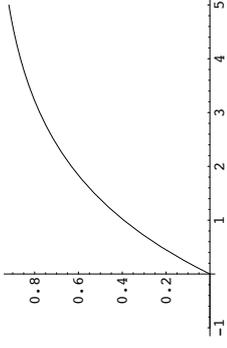
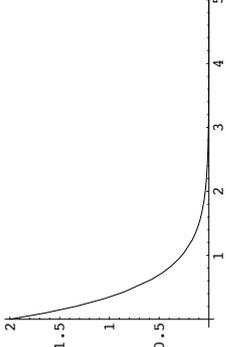
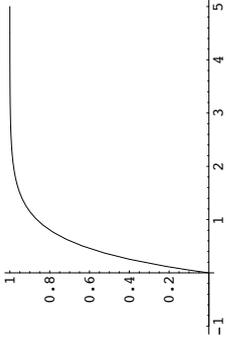
$$\begin{aligned}\mathbb{E}[X] &= \int_{-\infty}^\infty x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} dx = \int_{-\infty}^\infty (\sigma y + \mu) \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \\ &= \sigma \int_{-\infty}^\infty y \varphi(y) dy + \mu \int_{-\infty}^\infty \varphi(y) dy = \sigma \cdot 0 + \mu \cdot 1 = \mu \\ \mathbb{E}[X^2] &= \int_{-\infty}^\infty x^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} dx = \int_{-\infty}^\infty (\sigma y + \mu)^2 \varphi(y) dy \\ &= \sigma^2 \int_{-\infty}^\infty y^2 \varphi(y) dy + 2\mu\sigma \int_{-\infty}^\infty y \varphi(y) dy + \mu^2 \int_{-\infty}^\infty \varphi(y) dy \\ &= \sigma^2 \cdot 1 + 2\mu\sigma \cdot 0 + \mu^2 \cdot 1 = \sigma^2 + \mu^2 \\ \text{Var}(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \sigma^2\end{aligned}$$

Bemerkung

Die bekannten Rechenregeln gelten weiter:

- Linearität des Erwartungswerts
- $\text{Var}(aX + b) = a^2 \text{Var}(X)$ für $a, b \in \mathbb{R}$
- $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$
- Tschebyscheffsche Ungleichung
- Definition der bedingten Wahrscheinlichkeit

Verteilung	Parameter	Dichte	Verteilungsfunktion
Standardnormal- verteilung	$(\mu = 0, \sigma = 1)$		
Normalverteilung	$\mu = 1/2, \sigma = 1$		
	$\mu = 0, \sigma = \sqrt{3}/2$		
	$\mu = 0, \sigma = \sqrt{2}$		

Verteilung	Parameter	Dichte	Verteilungsfunktion
<p>uniforme Verteilung</p>	<p>$a = 0, b = 3/2$</p>		
<p>Exponential- verteilung</p>	<p>$\alpha = 1$</p>		
	<p>$\alpha = 1/2$</p>		
	<p>$\alpha = 2$</p>		

Zufallsgrößen mit Dichten

Verteilung	Parameter	Dichte	Verteilungsfunktion	Erwartungswert	Varianz
Standardnormalverteilung	$(\mu = 0, \sigma = 1)$	$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$	$\Phi(x) = \int_{-\infty}^x \varphi(y) dy$	0	1
Normalverteilung	$\mu \in \mathbb{R}, \sigma > 0$	$\varphi(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$	$F(x) = \int_{-\infty}^x \varphi(y; \mu, \sigma) dy$	μ	σ^2
uniforme Verteilung	$a, b \in \mathbb{R}, a < b$	$f(x) = \begin{cases} \frac{1}{b-a} & \text{für } x \in [a, b] \\ 0 & \text{sonst} \end{cases}$	$F(x) = \begin{cases} 0 & \text{für } x < a \\ \frac{x-a}{b-a} & \text{für } a \leq x \leq b \\ 1 & \text{für } x > b \end{cases}$	$(a+b)/2$	$(a-b)^2/12$
Exponentialverteilung	$\alpha > 0$	$f(x) = \begin{cases} \alpha e^{-\alpha x} & \text{für } x \geq 0 \\ 0 & \text{sonst} \end{cases}$	$F(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 - e^{-\alpha x} & \text{für } x \geq 0 \end{cases}$	$1/\alpha$	$1/\alpha^2$