Übungsrunde 9, Gruppe 2

LVA 107.369, Übungsrunde 8, Gruppe 2, 12.12.

Markus Nemetz, markus.nemetz@tuwien.ac.at, TU Wien, 12/2006

1 3.75

1.1 Angabe

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 $X \rightsquigarrow \mathcal{N}(0, 1)$.

1.2 Wichtige Begriffe:

1.2.1 Verteilungsfunktion

In der Wahrscheinlichkeitstheorie kann die (kumulative) Verteilungsfunktion $F(x)$ verwendet werden, um die Wahrscheinlichkeitsverteilung einer reellen Zufallsvariable X zu beschreiben.

In der Regel spricht man einfach von einer Verteilungsfunktion. Die explizite Kennzeichnung als kumulativ kann helfen, Verwechslungen mit der Wahrscheinlichkeitsfunktion oder Dichtefunktion zu vermeiden.

Wenn die Verteilungsfunktion F stetig differenzierbar ist, dann definiert ihre Ableitung eine Dichtefunktion

$$f = F' = \frac{\mathrm{d}\, F(x)}{\mathrm{d}\, x}$$

1.2.2 Dichtefunktion

Eine Dichtefunktion, Wahrscheinlichkeitsdichte oder Wahrscheinlichkeitsdichtefunktion (WDF) (engl.: probability density function (pdf)) dient in der Mathematik der Beschreibung von Wahrscheinlichkeitsverteilungen.

Die Wahrscheinlichkeiten der einzelnen Ausprägungen einer stetigen Zufallsvariablen können (im Gegensatz zum diskreten Fall der Wahrscheinlichkeitsfunktion) nicht angegeben werden, denn die Wahrscheinlichkeiten für jede einzelne Ausprägung müssen streng genommen 0 gesetzt werden. Es lassen sich nur Wahrscheinlichkeiten $f(x)\,\mathrm{d}\,x$ dafür angeben, dass die Werte innerhalb eines Intervalls $\mathrm{d}\,x$ um x liegen. Die Funktion $f(x)$ heißt dann Dichtefunktion. Die Wahrscheinlichkeit, dass die Zufallsvariable Werte zwischen a und b annimmt, wird dann allgemein definiert als das Integral über diese Funktion mit den Integrationsgrenzen a und b.

Beispielsweise fragt man nicht, wie viele Personen exakt 1,75 Meter groß sind, sondern z. B., wie viele Personen zwischen 1,75 und 1,76 m groß sind. Denn die Wahrscheinlichkeit, dass eine Person auf beliebig viele Nachkommastellen genau 1,75 Meter groß ist, ist theoretisch und praktisch gleich Null (Nullmenge).

Durch Integration über ein nach unten unendliches Intervall erhält man die Verteilungsfunktion. Eine Zufallsvariable X heißt genau dann stetig verteilt oder kontinuierlich verteilt, wenn die dazugehörige Verteilungsfunktion stetig ist.

Auch für stetige Verteilungen lassen sich Momente angeben, die wichtigsten sind der Erwartungswert und die Varianz.
Definition: Eine integrierbare Funktion f heißt (Wahrscheinlichkeits-)Dichte oder Dichtefunktion der Zufallsvariable X, wenn

$$\int_a^b f(x)\,\mathrm{d}x = \mathrm{P}(a \leq X \leq b)$$

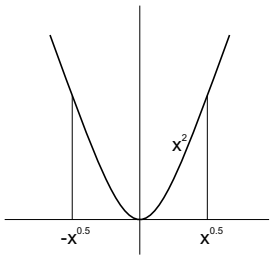
1.2.3 Transformationssatz

(Aus http://wss.uni-hd.de/resources/stat1/2003w/gsm/gsm_formel.ps)
(Transformationssatz für eindimensionale Zufallsvariablen). Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, $X : \Omega \to \mathbb{R}$ eine Zufallsvariable $h : X(\Omega) \to \mathbb{R}$ eine streng monotone Funktion und $Y := h(X)$. Dann gilt für $y \in h(X(\Omega))$

1. für die Verteilungsfunktion:
   - $F_Y(y) = F_X(h^{-1}(y))$ sofern h streng monoton wachsend ist
   - $F_Y(y) = 1 - F_X(h^{-1}(y))$ sofern h streng monoton fallend ist
2. a) für die Verteilung einer diskreten Zufallsvariablen $X : P_Y(y) = P_x(h^{-1}(y))$
   b) für die Dichte einer stetigen Zufallsvariablen X, wenn h differenzierbar ist:

$$f_y(y) = \begin{cases} f_x(h^{-1}(y))|(h^{-1})'(y)| = f_X(h^{-1}(y))|\frac{1}{h'}(h^{-1}(y))| \\ 0\,\text{sonst} \end{cases}$$

Geometrisch verdeutlicht zu unserem Beispiel:



1.3 Lösung des Beispiels

Zunächst ermittelt man die Verteilungsfunktion von $Z = X^2$:

$$F_Z = W\{Z \leq x\} = W\{X^2 \leq x\} = W\{-\sqrt{x} \leq X \leq \sqrt{x}\} = \\ F_X\{\sqrt{x}\} - F_X\{-\sqrt{x}\}, \qquad x \geq 0$$

Ableitung ergibt die Dichte:

$$f_Z(x) = F_Z' = \frac{1}{2\sqrt{x}}(f_X(\sqrt{X}) + f_X(-\sqrt{X})), \qquad x > 0$$

Beispiel für $X \rightsquigarrow \mathcal{N}(0,1)$:

$$f_Z(x) = \frac{1}{2\sqrt{x}}(\frac{1}{\sqrt{2\pi}}e^{-\frac{x}{2}} + \frac{1}{\sqrt{2\pi}}e^{-\frac{x}{2}}) =$$
$$\frac{1}{2\sqrt{x}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x}{2}} = \frac{x^{-\frac{1}{2}}e^{-\frac{x}{2}}}{\sqrt{2\pi}}, \qquad x > 0$$

(Dichte einer χ_1^2-Verteilung mit einem Freiheitsgrad)

2 3.77

2.1 Angabe

Die sG X sei uniform verteilt auf $(-\pi/2, -\pi/2)$. Bestimmen Sie die Verteilung (Dichte) von $Y = \tan X$.

2.2 Lösung des Beispiels

Die Verteilungsfunktion von X lautet (Nach Transformationssatz):

$$F_X(x) = \frac{2x+\pi}{2\pi}, \qquad -\frac{\pi}{2} \leq x \leq \frac{\pi}{2}$$

Für $Y = \tan X$ lautet nun die Verteilungsfunktion:

$$F_Y(x) = W\{\tan(X) \leq x\} = 1 - W\{\tan(X) > x\} =$$
$$1 - W\{-\arctan(x) \leq X \leq \arctan(x)\} =$$
$$1 - [\frac{2\arctan(x)+\pi}{2\pi} - \frac{-2\arctan(x)+\pi}{2\pi}] =$$
$$1 - \frac{2\arctan(x)}{\pi}, \ 0 \leq x \leq 1$$

Die Dichtefunktion erhält man durch die Ableitung von $F_Y(x)$:

$$f_Y(x) = F_Y'(x) = \frac{2}{\pi(\sqrt{1+x^2})}, \qquad 0 \leq x \leq 1$$

3 3.80

3.1 Angabe

Ein Algorithmus zur Erzeugung von poissonverteilten Beobachtungen $X \rightsquigarrow P_\mu$ lautet wie folgt:

1. Erzeuge eine (auf $(0,1)$) uniform verteilte Zahl U;
2. Setze i = 0, $\alpha = e^{-\mu}$, $F = \alpha$;
3. Wenn $U \leq F$, setze $X = i$, stop;
4. Setze $\alpha = \frac{\mu\alpha}{i+1}, F = F + \alpha, i = i + 1$;
5. Gehe zu (3).

Zeigen Sie, daß dieser Algorithmus Beobachtungen einer Poissonverteilung mit dem Parameter (= Mittelwert) μ generiert. *Wie oft wird (3) im Mittel zur Erzeugung einer Poisson-Beobachtung aufgerufen?

3.2 Lösung des Beispiels

Der angegebene Algorithmus realisiert die Prozedur für eine poissonverteilte sG. Zur schnelleren Berechnung von F_X wird dabei eine Rekursion verwendet (Rekursionsanfang $W\{X = 0\} = e^{-\mu}$):

$$
\begin{aligned}
a_0 &= e^{-\mu} \cdot i_0 = 0 \\
a_1 &= \frac{\mu \cdot e^{-\mu}}{i_0 + 1} \\
a_2 &= \frac{\mu \cdot \mu \cdot e^{-\mu}}{1 \cdot 2} \\
a_2 &= \frac{\mu \cdot \mu \cdot \mu \cdot e^{-\mu}}{1 \cdot 2 \cdot 3} \\
&\vdots \\
W\{X = i+1\} = \frac{\mu^{i+1} \cdot e^{-\mu}}{(i+1)!} &= \frac{\mu}{i+1} \cdot \frac{\mu^i \cdot e^{-\mu}}{i!} = \frac{\mu}{\mathbf{i+1}} \cdot \mathbf{W\{X = i\}}
\end{aligned}
$$

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

$$\mathbb{E} = \sum_{k=1}^{\infty} k \cdot p_X(k-1) = \sum_{k=0}^{\infty} (k+1) \cdot p_X(k) =$$
$$\underbrace{\sum_{k=0}^{\infty} k \cdot p_X(k)}_{\mathbb{E}(X)} + \underbrace{\sum_{k=0}^{\infty} p_X(k)}_{=1} = \mathbb{E}(X) + 1$$

Für eine P_μ-Verteilung gilt $\mathbb{E}(Y) = \mu + 1$.

4 4.2

4.1 Angabe

Zeigen Sie, daß die folgende Funktion nicht die Verteilungsfunktion eines 2-dimensionalen stochastischen Vektors (X, Y) sein kann:

$$F(x,y) = \begin{cases} 1 & x + 2y \geq 1 \\ 0 & x + 2y < 1 \end{cases}$$

4.2 Theoretische Grundlagen: Mehrdimensionale Verteilung

Die gemeinsame Wahrscheinlichkeitsverteilung mehrerer Zufallsvariablen nennt man multivariate Verteilung oder auch mehrdimensionale Verteilung.
Die n-dimensionale Verteilungsfunktion an der Stelle $(x_1, \ldots, x_n)$ ist die Verteilungsfunktion

$$F(x_1, \ldots, x_n) = P[X_1 \leq x_1, \ldots, X_n \leq x_n].$$

Um Verwechslungen zu vermeiden, werden skalare Zufallsvariablen groß geschrieben, Zufallsvektoren jedoch klein. Matrizen und Vektoren werden unterstrichen.

Man betrachtet p Zufallsvariablen X_j $(j = 1, \ldots, p)$, jeweils mit einem Erwartungswert E(Xj) und der Varianz $V(X_j)$.
Man interessiert sich für die gemeinsame Wahrscheinlichkeit, dass alle X_j höchstens gleich einer jeweiligen Konstanten x_j sind, also

$$P(X_1 \leq x_1; X_2 \leq x_2; \ldots; X_p \leq x_p) = F_X(x_1; x_2; \ldots, x_p).$$

Multivariate Zufallsvariablen werden i.A. in Matrixform dargestellt. Man fasst die Zufallsvariablen in einem $(p \times 1)$-Zufallsvektor $\underline{x}$ zusammen:

$$\underline{x} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix}.$$

Für die obige gemeinsame Wahrscheinlichkeit erhält man

$$F_x(\underline{x}) = F_X \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}.$$

Die Erwartungswerte befinden sich im Erwartungswertvektor

$$E(\underline{x}) = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{pmatrix}$$

.

4.3 Lösung des Beispiels

Besser ist, man widerlegt direkt über die Eigenschaften der Verteilungsfunktion, sondern geht den Weg übers Ereigniskalkül:

$$F(x, y) = W\{X \leq x, Y \leq y\} = W(\omega \in M | X(\omega) \leq x \text{ und } Y(\omega) \leq y\})$$

Jetzt definieren wir zwei Ereignisse:

$$A = \{\omega \in M | X(\omega) \leq 10\} \cap \{\omega \in M | Y(\omega) \leq 0\}$$
$$B = \{\omega \in M | X(\omega) \leq 0\} \cap \{\omega \in M | Y(\omega) \leq 10\}$$

Wir betrachten: $F(10, 0) = 1$: Das entspricht im Ereigniskalkül $W(A) = 1$. Dann betrachten wir $F(0, 10) = 1$: Das entspricht im Ereigniskalkül $W(B) = 1$.
Laut Additionstheorem gilt:

$$W(A \cup B) = W(A) + W(B) - W(A \cap B) = 1 + 1 - W(A \cap B) = 2 - W(A \cap B)$$

Was ist nun $W(A \cap B)$?

$$W(A \cap B) = W(\omega \in M | X(\omega) \leq 0 \cap \omega \in M | Y(\omega) \leq 0)$$

Das entspricht aber $F(0, 0) = 0$.
Also ergibt $W(A \cup B) = 2 - 0 = 2$ einen ungültigen Wert für die Wahrscheinlichkeit. Die 'Wahrscheinlichkeitsverteilung' liefert also ungültige Werte.

5 4.6

5.1 Angabe

Angenommen, 12 Karten werden aus einem gewöhnlichen (gut durchmischten) Kartenpaket gezogen. X_1 sei die Zahl der gezogenen Einser (Asse), X_2 die Zahl der Zweier, X_3 die Zahl der Dreier und X_4 die Zahl der Vierer. Bestimmen Sie einen Ausdruck für die gemeinsame Verteilung von (X_1, X_2, X_3, X_4), d.h. bestimmen Sie $p(x_1, x_2, x_3, x_4) = W\{X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4\}$, wenn die Ziehungen (a) ohne Zurücklegen, (b) mit Zurücklegen erfolgen.

5.2 Theoretische Grundlagen: Diskrete mehrdimensionale Verteilung

Eine diskrete Verteilung auf $\mathbb{R}^m$ ist durch eine höchstens abzählbare Menge M von Elementen $(x_1, \ldots, x_n) \in \mathbb{R}^m$, die keinen Häufungspunkt besitzt und zugehörige Punktwahrscheinlichkeiten $p(x_1, \ldots, x_n) > 0$ geben, so dass gilt:

$$\sum_{(x_1,\ldots,x_n)\in M} p(x_1, \ldots, x_n) = 1.$$

Die Wahrscheinlichkeit einer m-dimensionalen Borel-Menge $B \subseteq \mathbb{R}^m$ ist dann ggegeben durch:

$$\sum_{(x_1,\ldots,x_n)\in B,\, (x_1,\ldots,x_n)>0} p(x_1, \ldots, x_,)$$

Definition einer Borel-Menge:
Für einen gegebenen topologischen Raum Ω ist die borelsche Ω-Algebra definiert als die kleinste μ-Algebra, die die offenen Mengen von Ω enthält.

- Eine μ-Algebra auf einer Grundmenge Ω ist eine Menge von Teilmengen, die die Grundmenge enthält und die bezüglich Komplementbildung und abzählbarer Vereinigung abgeschlossen ist. Eine Grundmenge zusammen mit einer auf ihr erklärten μ-Algebra heißt auch Messraum.
- Da für jeden topologischen Raum Ω die Potenzmenge $\mathcal{P}(\Omega)$ eine μ-Algebra ist, existiert auch die borelsche μ-Algebra zu Ω.
- Eine borelsche μ-Algebra ermöglicht es somit, einen topologischen Raum in kanonischer Weise mit der zusätzlichen Struktur eines Messraums auszustatten. Im Hinblick auf diese Struktur heißt der Raum dann auch Borel-Raum.

5.2.1 Multinomialverteilung

Auch Polynomialverteilung genannt. n unabhängige Widerholungen eines Versuches mit r Versuchsausgängen ($r \geq 2$), $\Omega = \{(a_1, \ldots, a_m) : a_i \in \{b_1, \ldots, b_r\}\}$, $A_{n_1,\ldots,n_r} = \{(a_1, \ldots, a_n) : n_1 \text{ mal } b_1, \ldots, n_r \text{ mal } b_r\}$, $n_1 + \cdots + n_r = n$, p_j Wahrscheinlichkeit, dass b_j im i-ten Versuch $\forall$i, $(p_1 + \cdots + p_r = 1)$,

$$P(A_{n_1,\ldots,n_r}) = \frac{n!}{n_1! \cdot \cdots \cdot n_r!} \cdot p_1^{n_1} \cdot \cdots \cdot p_r^{n_r}$$

5.3 Theoretische Grundlagen: mehrdimensionale hypergeometrische Verteilung

Auswahl von n Elementen aus M möglichen ohne Zurücklegen; die M Elemente sind in r Gruppen mit $M_1, \ldots, M_r$ Elementen jeweils geteilt, $M_1 + \cdots + M_r = M$,$\Omega = \{[a_1, \ldots, a_n] : a_i \in \{1, \ldots, M\}, a_k \neq a_j \, (k \neq j)\}$,$A_{n_1,\ldots,n_r} = \{[a_1, \ldots, a_n] : a_k \neq a_j \, (k \neq j)\}, n_1$ El. aus 1. Gr. $, \ldots, n_r$ El. aus r-ter Gr. $\}$,

$$P(A_{n_1,\ldots,n_r}) = \frac{\binom{M_1}{n_1} \cdot \cdots \cdot \binom{M_r}{n_r}}{\binom{M}{n}}$$

$r = 2$ - **hypergeometrische Verteilung**
2 Gruppen, gut/schlechtm unter n ausgewählten k gute. A_k

$$P(A_k) = \frac{\binom{M_1}{k} \cdot \cdots \cdot \binom{M - M_1}{n - k}}{\binom{M}{n}}$$

für $M \to \infty$ $, \frac{M_1}{M} \to p$ $, 0 < p < 1$:

$$P(A_k) \to \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$$

(Approximation der hypergeometrischen Verteilung durch die Binomialverteilung)

5.4 Lösung des Beispiels

a) (Hyperhgeometrisch)

$$P(A_{n_1}, \ldots, A_{n_r}) = \frac{\binom{M_1}{n_1} \cdot \cdots \cdot \binom{M_r}{n_r}}{\binom{M}{N}} = \frac{1}{\binom{52}{12}} \underbrace{\binom{4}{x_1}\binom{4}{x_2}\binom{4}{x_3}\binom{4}{x_4}}_{4-\text{dimensional}} =$$

$$\binom{52 - 16}{12 - y}, \qquad y = x_1 + x_2 + x_3 + x_4$$

b) (Multinomial)

$$\frac{n!}{n_1! \cdot \cdots \cdot n_r!} p_1^{n_1} \cdot \cdots \cdot p_r^{n_r} = \frac{12!}{x_1! x_2! x_3! x_4!} (\frac{4}{52})^{x_1} (\frac{4}{52})^{x_2} (\frac{4}{52})^{x_3} (\frac{4}{52})^{x_4} (\frac{36}{12})^{12-y}$$

6 4.7

6.1 Angabe

Zwei regelmäßige Tetraeder mit den Augenzahlen 1, 2, 3, 4 werden geworfen. Bestimmen Sie - in Tabellenform - die gemeinsame Verteilung von X = untenliegende Augenzahl des 1. Tetraeders und Y = größere der beiden untenliegenden Augenzahlen.

6.2 Theoretische Grundlagen: Diskrete mehrdimensionale Verteilung

Siehe Beispiel 4.6.!

6.3 Lösung des Beispiels

Tabekke der Punktwahrscheinlichkeiten $p(X, Y)$: (senkrecht X, waagrecht Y):

$\downarrow X \mid \rightarrow Y$	1	2	3	4 | RV_X
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$ | $\frac{4}{16}$
2	0	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{1}{16}$ | $\frac{4}{16}$
3	0	0	$\frac{3}{16}$	$\frac{1}{16}$ | $\frac{4}{16}$
4	0	0	0	$\frac{4}{16}$ | $\frac{4}{16}$
RV_Y	$\frac{1}{16}$	$\frac{3}{16}$ | $\frac{5}{16}$	$\frac{7}{16}$	

Beispiel Berechnung $(2, 2)$: der erste ist zwei, das Maximum ist auch 2, also kann der zweite Würfel 1 oder 2 sein - daraus folgt: 2/16 (1/16 für $(2, 1)$ + 1/16 für $(2, 2)$)
Ergibt Randverteilung für $X = \frac{1}{4}$ und gesamt 1.
Allgemein:

$$P_X(y) = \frac{2y - 1}{16}$$