Ausarbeitung zum Seminarthema „Verteilungsfunktionen von Zufallsvariablen, Erwartung und bedingte Erwartung“

Burim Ramosaj

27. Februar 2013

1 Verteilungsfunktionen

1.1 Definition

Eine Funktion $F : \mathbb{R} \to \mathbb{R}$ heißt *(eindimensionale) Verteilungsfunktion*, falls Sie folgende Eigenschaften besitzt:

i.) F ist nicht fallend, d.h. es gilt: $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$

ii.) F ist rechtsseitig stetig und die linksseitigen Grenzwerte existieren

iii.) $\lim_{x \to -\infty} F(x) = 0, \quad \lim_{x \to \infty} F(x) = 1$

Nun sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum. Jede messbare Funktion $X : (\Omega, \mathscr{A}) \mapsto (\mathbb{R}, \mathscr{B})$ heißt Zufallsvariable. Dabei gilt folgendes: $P_X(B) = P(X^{-1}(B)), B \in \mathscr{B}$. Mit den Eigenschaften des Wahrscheinlichkeitsmaßes kann man zeigen, dass folgende Funktion eine (eindimensionale) Verteilungsfunktion ist.

$$F_X(x) \quad = \quad P_X((-\infty, x]), \quad x \in \mathbb{R}$$

1.2 Definition

Sei nun $(X_1, \ldots, X_n)$ ein n-dimensionaler Zufallsvektor. Eine Funktion $G : \mathbb{R}^n \to \mathbb{R}$ heißt *(mehrdimensionale) Verteilungsfunktion*, falls sie folgende Eigenschaften erfüllt:

i.) $G(x_1, \ldots, x_n)$ ist in jedem Argument nicht fallend

ii.) $G(x_1, \ldots, x_n)$ ist in jedem Argument rechtsseitig stetig

iii.) $G(x_1, \ldots, x_n) \to 0$ für mindestens ein $x_i \to -\infty$
$G(x_1, \ldots, x_n) \to 1$, für alle $x_j \to \infty, j = 1, \ldots, n$

iv.) Falls $a_i \leq b_i, i = 1, \ldots, n$, und $\Delta_{a_i b_i} G(x_1, \ldots, x_n)$
$= G(x_1, \ldots, x_{i-1}, b_i, \ldots, x_n) - G(x_1, \ldots, x_{i-1}, a_i, \ldots, x_n) \geq 0$

$\Rightarrow \Delta_{a_1 b_1} \Delta_{a_2 b_2} \cdots \Delta_{a_n b_n} G(x_1, \ldots, x_n) \geq 0$

1.3 Eigenschaften von Verteilungen von Zufallsvariablen

1.3.1 Motivation

Das folgende Beispiel soll zeigen, dass zwei in Verteilung gleiche Zufallsvariablen, wenn man sie mit einer weiteren Zufallsvariablen multipliziert, nicht unbedingt dieselbe Verteilung besitzen müssen. Hierfür sei $X \overset{f.s}{>} 0$ eine symmetrische Zufallsvariable und $Y = -X$. Dann ist es offensichtlich, dass

$$Y \overset{d}{=} X$$

Nun sei $Z = Y$. Dann gilt, dass $ZX = -X^2$, und $ZY = (-X)(-X) = X^2$. Folglich kann ZX, ZY nicht identisch verteilt sein, weil ZX f.s negativ und ZY f.s. positiv ist.

Wir können jedoch mittels messbarer Funktionen zwei in Verteilung unterschiedliche Zufallsvariablen so transformieren, dass sie wieder dieselbe Verteilung besitzen.

1.3.2 Über die Transformation verschiedener Verteilungen

Zunächst führen wir die Gamma- und die Beta-Funktion ein, die für das folgende Beispiel relevant sind.

i.) Die *Gamma-Funktion*:
$\Gamma : \mathbb{R}_+ \to \mathbb{R}, \quad \Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx$

ii.) Die *Gamma-Verteilung*:

Sei $\lambda > 0, p > 0$. X heißt *Gamma-Verteilt*, falls X eine absolut stetige Zufallsvariable ist mit Dichte $f_X(x) = \frac{\lambda^p x^{p-1} e^{-\lambda x}}{\Gamma(p)} \mathbb{1}\{x \geq 0\}$

iii.) Die *Beta-Funktion*:

Sei $p > 0, q > 0$. Die *Beta-Funktion* hat folgende Gestalt:

$$B(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1} dt, \quad B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

Für dieses Beispiel betrachten wir 3 paarweise unabhängige Zufallsvariablen mit unterschiedlichen Verteilungen:

- $X \sim N(0,1) \quad \Rightarrow \quad X^2 \sim \Gamma(\frac{1}{2}, \frac{1}{2})$
- $Y \sim Exp(1) \quad \Rightarrow \quad Y \sim \Gamma(1,1)$
- Z besitze folgende Dichtefunktion: $f_Z(x) = \frac{1}{\pi\sqrt{x(1-x)}} \mathbb{1}\{x \in [0,1]\}$

Wir transformieren diese und betrachten nun folgende Zufallsvariablen:

$$\tilde{X} = \log(\tfrac{1}{2}X^2), \quad \tilde{Y} = \log(Y), \quad \tilde{Z} = \log(Z)$$

Für eine Gamma-Verteilte Zufallsvariable X mit Parameter $\lambda > 0, p > 0$ gilt folgendes:

$$\mathbb{E}X^{it} = \int_0^\infty x^{it} \frac{\lambda^p x^{p-1} e^{-\lambda x}}{\Gamma(p)} dx = \frac{\lambda^p}{\Gamma(p)} \int_0^\infty x^{(it+p)-1} e^{-\lambda x} dx \overset{\lambda x = u}{=} \frac{\lambda^p}{\Gamma(p)} \int_0^\infty \frac{1}{\lambda^{it+p-1}} u^{it+p-1} e^{-u} \frac{1}{\lambda} du = \frac{\Gamma(it+p)}{\lambda^{it}\Gamma(p)}$$

Berechnen wir nun die charakteristischen Funktionen der drei Zufallsvariablen.

$$\varphi_{\tilde{X}}(t) = \mathbb{E}(e^{it\tilde{X}}) = \mathbb{E}(e^{it\log(\frac{1}{2}X^2)}) = \tfrac{1}{2^{it}}\mathbb{E}(X^2)^{it} \overset{\lambda=\frac{1}{2},p=\frac{1}{2}}{=} \frac{\Gamma(it+\frac{1}{2})}{\Gamma(\frac{1}{2})} = \frac{\Gamma(it+\frac{1}{2})}{\sqrt{\pi}}$$

$$\varphi_{\tilde{Y}}(t) = \mathbb{E}(e^{it\tilde{Y}}) = \mathbb{E}(Y)^{it} \overset{\lambda=1,p=1}{=} \Gamma(it+1)$$

$$\varphi_{\tilde{Z}}(t) = \mathbb{E}(Z)^{it} = \int_0^1 x^{it} \frac{1}{\pi} x^{-\frac{1}{2}} (1-x)^{-\frac{1}{2}} dx = \frac{1}{\pi} \int_0^1 x^{(it+\frac{1}{2})-1} (1-x)^{\frac{1}{2}-1} dx$$
$$\overset{p=it+\frac{1}{2},q=\frac{1}{2}}{=} \frac{1}{\pi} B(it+\frac{1}{2}, \frac{1}{2}) = \frac{\Gamma(it+\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(it+1)\pi} = \frac{1}{\sqrt{\pi}} \frac{\Gamma(it+\frac{1}{2})}{\Gamma(it+1)}$$

Weil die Zufallsvariablen paarweise unabhängig sind und für die charakteristische Funktionen gilt:

$$\varphi_{\tilde{X}}(t) = \varphi_{\tilde{Y}}(t)\varphi_{\tilde{z}}(t), \quad \forall t > 0, \quad \Rightarrow \quad \tilde{X} \overset{d}{=} \tilde{Y} + \tilde{Z} \quad \Rightarrow \quad X^2 \overset{d}{=} 2YZ$$

1.3.3 Über den Zusammenhang stark unimodaler und unimodaler Verteilungen

1.3.3.1 Definition

Eine Verteilung heißt *unimodal*, wenn dessen Dichtefunktion nur einen Gipfel besitzt. Eine Verteilungsfunktion G heißt *stark unimodal*, wenn die Faltung $G * F$, für jede unimodale Verteilungsfunktion F, wieder unimodal ist.

Es gibt einige Verteilungen, die stark unimodal sind, wie z.B. die Normalverteilung, die Gleichverteilung auf einem beliebigen Intervall $[a, b]$, die Gamma-Verteilung mit Parameter $\lambda \geq 1$ etc. Jedoch haben wir, wie in den Foliensätzen gezeigt, ein Beispiel gefunden, in dem die Faltung einer unimodalen Verteilung nicht wieder unimodal sein muss.

Wir zeigen nun, dass die starke Unimodularität eine stärkere Eigenschaft ist als die herkömmliche Unimodularität.

Dazu sei F_k die Verteilungsfunktion der Gleichverteilung auf der Menge $\{0, 1, \dots, k\}, k \in \mathbb{N}$, wobei $F_i, F_j, \quad i \neq j$ Verteilungsfunktionen von unabhängigen Zufallsvariablen sind. Ferner sei $F = \frac{1}{2}(F_0 + F_{m+1})$ für ein festes $m \geq 3, m \in \mathbb{N}$. Dann ist F unimodal, aber die Faltung $G = F * F$ nicht. Die Verteilung G ist auf der Menge $\{0, 1, 2, \dots, 2m - 2\}$ konzentriert und falls $g_i, \quad i = 0, 1, 2, \dots, 2m-2$ die Punkte der Atome sind, dann können wir leicht erkennen, dass für $X_0 \sim F, X_{m+1} \sim F$ Folgendes gilt:

$$g_0 = G(0) = (\tfrac{1}{2}(F_0(0) + F_{m+1}(0)))^2 = \tfrac{1}{4}(1 + \tfrac{1}{m})^2 = \tfrac{1}{4}(1 + 2m^{-1} + m^{-2})$$

$$\begin{aligned} g_1 &= P(X_0 + X_{m+1} = 1) = P(X_0 = 0, X_{m+1} = 1) + P(X_0 = 1, X_{m+1} = 0) \\ &= 2P(X_0 = 0)P(X_{m+1} = 1) = 2F(0)(F(1) - F(0)) = 2(F(1)F(0) - F(0)^2) \\ &= 2(\frac{1}{4}(1 + 3m^{-1} + 2m^{-2}) - \frac{1}{4}(1 + 2m^{-1} + m^{-2})) = \frac{1}{4}(2m^{-1} + 2m^{-2}) \end{aligned}$$

Analoges Vorgehen zur Berechnung von $g_2 = G(2) - G(1)$ führt zu folgendem Ergebnis:

$$4g_0 = 1 + 2m^{-1} + m^{-2}, \quad 4g_1 = 2m^{-1} + 2m^{-2}, \quad 4g_2 = 2m^{-1} + 3m^{-2}$$

Weil $m \geq 3$, gilt : $g_1 < \min(g_0, g_2)$. G ist somit nicht unimodal, weil mindestens g_0 bzw. g_2 Gipfelpunkte sind. Wir erkennen damit, dass F zwar unimodal ist, aber nicht stark unimodal.

1.4 Mehrdimensionale Verteilungsfunktionen

1.4.1 Definitionsvergleich

In diesem Beispiel wollen wir die Unterschiede zwischen Definition 1.1 und Definition 1.2 genauer erläutern und dabei auf die zentrale Eigenschaft iv.)[1] eingehen. Die Punkte i.) - iii.) scheinen sowohl bei Definition 1.1 als auch bei Definition 1.2, im mehrdimensionalen Sinne übertragen, gleich. Sie ist, wie jetzt gezeigt wird, essentiell für die n-dimensionale Verteilungsfunktion.

[1] $a_i \leq b_i, i = 1, \dots, n, \Delta_{a_i b_i} G(x_1, \dots, x_n) = G(x_1, \dots, x_{i-1}, b_i, \dots, x_n) - G(x_1, \dots, x_{i-1}, a_i, \dots, x_n) \geq 0 \Rightarrow \Delta_{a_1 b_1} \Delta_{a_2 b_2} \cdots \Delta_{a_n b_n} G(x_1, \dots, x_n) \geq 0$

Dazu konstruieren wir ein Beispiel, das die Bedingungen i.) - iii.) aus Definition 1.2 erfüllt, aber nicht den Punkt iv.). Sei folgende Funktion definiert:

$$G(x,y) = \begin{cases} 0, & \text{wenn } x < 0 \text{ oder } y < 0 \\ \min(1, \max(x,y)), & \text{sonst.} \end{cases}$$

Die Funktion G lebt im 1. Quadranten und nimmt nur Werte zwischen 0 und 1 an. D.h. für zunehmende x und/oder y ist die Funktion durch 1 beschränkt und durch 0 vice versa. Es ist offensichtlich, dass $G(x,y)$ sowohl in x als auch in y nicht-fallend und rechtsseitig stetig ist. Somit sind die Bedingungen i.) - iii.) erfüllt. Lediglich iv.) wird verletzt. Dazu betrachten wir das Rechteck im 1. Quadranten und zwei unabhängig Zufallsvariablen X, Y:

$$R = [\tfrac{1}{3}, 1] \times [\tfrac{1}{3}, 1]$$

$P((X,Y) \in R) = P(\frac{1}{3} \leq X \leq 1, \frac{1}{3} \leq Y \leq 1) \overset{\text{unabhängig}}{=} P(\frac{1}{3} \leq X \leq 1)P(\frac{1}{3} \leq Y \leq 1) = (P(X \leq 1) - P(X \leq \frac{1}{3}))(P(Y \leq 1) - P(Y \leq \frac{1}{3})) = G(1,1) - G(1, \frac{1}{3}) - G(\frac{1}{3}, 1) + G(\frac{1}{3}, \frac{1}{3}) = 1 - 1 - 1 + \frac{1}{3} = -\frac{2}{3}$

Was nicht sein kann. Somit ist G keine mehrdimensionale Verteilungsfunktion.

1.4.2 Über die Stetigkeit n-dimensionaler Verteilungsfunktionen

Eine Zufallsvariable X sei auf dem Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$ definiert. Für das Wahrscheinlichkeitsmaß P soll dabei gelten, dass es keine Atome besitzt, also $P(\{\omega : X(\omega) = x\}) = 0, \forall x \in \mathbb{R}$. Dann ist die Verteilungsfunktion stetig in jedem $x \in \mathbb{R}$. Fraglich ist, ob diese Eigenschaft sich auch auf $\Omega \subseteq \mathbb{R}^n, n \geq 2$ übertragen lässt. Im Allgemeinen muss dies jedoch nicht der Fall sein. Dazu betrachten wir folgende 2-dimensionale Verteilungsfunktion

$$F(x,y) = \begin{cases} xy, & \text{wenn } 0 \leq x < 1, 0 \leq y < \frac{1}{2} \\ \frac{x}{2}, & \text{wenn } 0 \leq x < 1, \frac{1}{2} \leq y < \infty \\ y, & \text{wenn } 1 \leq x < \infty, 0 \leq y < 1 \\ 1, & \text{wenn } x \geq 1, y \geq 1 \\ 0, & \text{sonst} \end{cases}$$

Man kann leicht zeigen, dass die obige Funktion die Voraussetzungen der Definition 1.2 erfüllt und somit eine Verteilungsfunktion des Zufallsvektors (X,Y) ist. Dazu betrachtet man folgende Abbildung der Urbilder dieser Funktion:

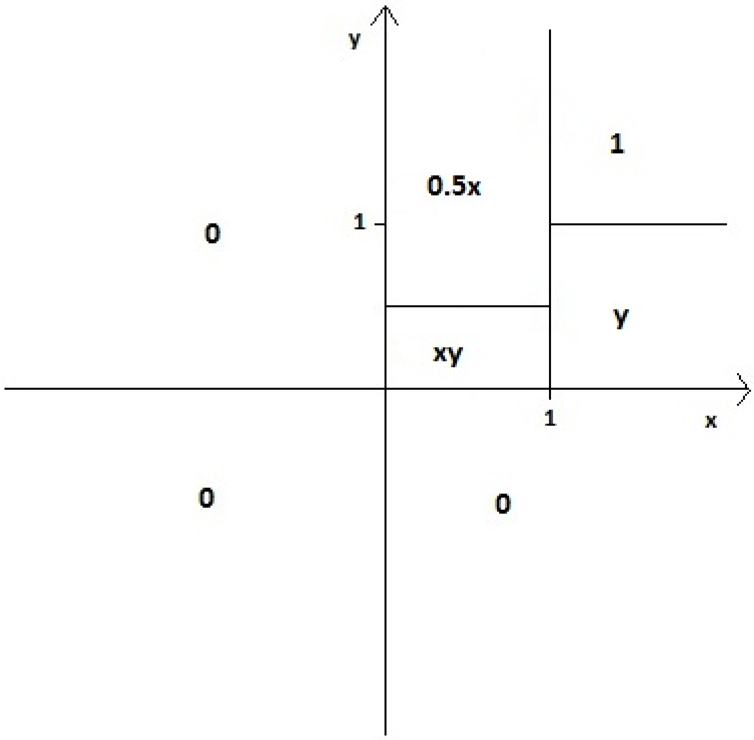

Abbildung 1: Urbild von F

Man erkennt, dass die Funktion auf dem Quadranten $Q = \{(x, y) : 0 \leq x < \infty, 0 \leq y < \infty\}$ lebt. Zudem hat jeder Punkt $(x, y) \in Q$ ein Wahrscheinlichkeitsmaß von Null. Also könnte man wie im eindimensionalen Fall vermuten, dass diese Funktion stetig ist, was jedoch falsch ist. Betrachtet man alle diejenigen Punkte aus der Menge $M = \{(1^-, y) : y \in [\frac{1}{2}, \infty)\}$, so erkennt man, dass diese Sprungstellen der Funktion sind. Denn wertet man die Funktion an diesen Stellen aus, so gilt für $(1^-, y) \in M$: $F(1^-, y) = \frac{1^-}{2} \approx \frac{1}{2}$. Jedoch entspricht dieser Wert nicht dem Wert in $(1, y_0), y_0 \geq 1 : F(1, y_0) = 1 \neq \frac{1}{2}$. Der Grund dafür liegt in der Existenz einer Hyperebene mit positivem Wahrscheinlichkeitsmaß.

1.4.3 Über den Zusammenhang multivariater und Rand-Verteilungsfunktionen

Falls der Zufallsvektor $(X_1, \ldots, X_n)$ die Verteilungsfunktion $F(x_1, \ldots, x_n)$ besitzt, dann gilt für die Randverteilungen $F_k(x_k) = P(X_k \leq x_k), k = 1, \ldots, n$, dass diese eindeutig bestimmt sind. Das folgende Beispiel soll dabei die Rückrichtung widerlegen. Dazu betrachten wir Randdichten, die f.s. übereinstimmen, jedoch verschiedene gemeinsame Dichtefunktionen besitzen.

$$f(x_1, x_2) = \begin{cases} \frac{1}{4}(1 + x_1 x_2), & \text{wenn } -1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1 \\ 0, & \text{sonst} \end{cases}$$

$$g(x_1,x_2)=\begin{cases}\frac{1}{4}, & \text{wenn } -1\leq x_1\leq 1, -1\leq x_2\leq 1\\ 0, & \text{sonst}\end{cases}$$

Dann gilt für die jeweilige Randdichtefunktion von f bzw. g :

$$f_1(x_1)=\int_{-1}^{1} f(x_1,x_2)dx_2=\int_{-1}^{1}\frac{1}{4}(1+x_1x_2)dx_2=\frac{1}{2}\mathbb{1}\{x_1\in[-1,1]\}$$

Analog berechnet man die übrigen Randdichten und kommt zu folgenden Ergebnissen:

$$f_2(x_2)=\tfrac{1}{2}\mathbb{1}\{x_2\in[-1,1]\},\quad g_1(x_1)=\tfrac{1}{2}\mathbb{1}\{x_1\in[-1,1],\quad g_2(x_2)=\tfrac{1}{2}\mathbb{1}\{x_2\in[-1,1]\}$$

und folglich: $f_1=g_1,\quad f_2=g_2$, aber $f\neq g$

1.4.4 Stetige Verteilungsfunktionen

Im Foliensatz hatten wir eine stetige, multivariate Dichtefunktion konstruiert, dessen Randdichten in einem Punkt unstetig sind. Wir erweitern dieses Beispiel um mehrere Unstetigkeitspunkte, d.h. wir haben eine stetige multivariate Dichtefunktion, dessen Randdichtefunktionen nicht nur einen Unstetigkeitspunkt besitzen, sondern mehrere. Dazu sei nochmals f und f_1 gegeben.

$$f(x,y)=\tfrac{1}{2\sqrt{2\pi}}|x|\exp(-|x|-\tfrac{1}{2}x^2y^2),\quad f_1(x)=\begin{cases}0, & \text{wenn } x=0\\ \frac{1}{2}\exp(-|x|), & x\neq 0\end{cases}$$

Man konstruiere eine neue Dichtefunktion $g=\sum_{n=1}^{\infty}\underbrace{2^{-n}f(x-r_n,y)}_{=f_n},\quad r_i\in\mathbb{Q},\forall i.$

Da f in $\mathbb{R}^2$ beschränkt ist, f_n monoton ist, gilt nach dem Satz von Beppo-Levi, dass sich Integral und unendliche Reihe vertauschen lassen und die Funktionen f,g in $\mathbb{R}^2$ gleichmäßig konvergieren. Für die Dichte g_1 gilt also:

$$g_1(x)=\int_{-\infty}^{\infty}\sum_{n=1}^{\infty}2^{-n}f(x-r_n,y)\overset{Beppo-Levi}{=}\sum_{n=1}^{\infty}\int_{-\infty}^{\infty}2^{-n}f(x-r_n,y)dy\overset{s.o.}{=}\sum_{n=1}^{\infty}2^{-n}f_1(x-r_n)$$

Obwohl g überall stetig ist, gilt für dessen Randdichte g_1, dass diese in jedem rationalen Punkt $r_1,r_2,\ldots$ unstetig ist, denn $f_1(r_k)=0,\forall k\in\mathbb{N},\quad f_1(x)\neq 0,\forall x\notin\mathbb{Q}.$

1.5 Erwartungswerte und bedingte Erwartung

1.5.1 Eigenschaften der Bedingten Erwartung

Bevor wir uns mit einigen Beispielen aus dem Bereich der Erwartungswerte und der bedingten Erwartung beschäftigen, liste ich hier nochmals die Eigenschaften der bedingten Erwartung ohne Beweis auf. Hierbei sei $\mathscr{D}$ eine Teil-σ-Algebra von $\mathscr{A}$.

i.) $X \overset{f.s.}{=} c, \quad c = \text{ const } \Rightarrow \quad \mathbb{E}(X|\mathscr{D}) \overset{f.s.}{=} c$

ii.) $X_1 \overset{f.s}{\leq} X_2 \Rightarrow \quad \mathbb{E}(X_1|\mathscr{D}) \overset{f.s.}{\leq} \mathbb{E}(X_2|\mathscr{D})$

iii.) X_1, X_2 seien integrierbar, $c_1, c_2 \in \mathbb{R},$ dann ist
$\mathbb{E}(c_1X_1 + c_2X_2|\mathscr{D}) \overset{f.s.}{=} c_1\mathbb{E}(X_1|\mathscr{D}) + c_2\mathbb{E}(X_2|\mathscr{D})$

iv.) $\mathbb{E}(\mathbb{E}(X|\mathscr{D})) \overset{f.s.}{=} \mathbb{E}X$

v.) $\mathscr{D}_1 \subset \mathscr{D}_2 \subset \mathscr{A} \Rightarrow \mathbb{E}(\mathbb{E}(X|\mathscr{D}_2)|\mathscr{D}_1) \overset{f.s.}{=} \mathbb{E}(X|\mathscr{D}_1)$

vi.) Falls X unabhängig bzgl. $\mathscr{D}$ ist, (d.h. X ist unabhängig von $\mathbb{1}_A, A \in \mathscr{D}$), dann gilt: $\mathbb{E}(X|\mathscr{D}) \overset{f.s.}{=} \mathbb{E}X$.

vii.) falls $X \quad \mathscr{D}$- messbar ist, und $E(|XY|) < \infty \Rightarrow \quad \mathbb{E}(XY|\mathscr{D}) \overset{f.s.}{=} X\mathbb{E}(Y|\mathscr{D})$

1.5.2 Linearitätseigenschaften des Erwartungswertes

Es gilt zwar im Allgemeinen, dass $\mathbb{E}(\sum\limits_{k=0}^{n} X_k) = \sum\limits_{k=0}^{n} \mathbb{E}(X_k) \quad (*)$, falls $\mathbb{E}(X_i) < \infty, \forall i \in \{0, 1, \ldots, n\}$. Das folgende Beispiel zeigt, dass die Endlichkeit des Erwartungswertes eine wichtige Voraussetzung für die Gültigkeit von $(*)$ ist. Dazu sei $W \sim U[0, 1]$. Wir definieren nun weitere Zufallsvariablen X, Y, Z wie folgt:

- $X = Y = \tan(\frac{\pi W}{2})$
- $Z = -2\tan(\frac{\pi W}{2})$

Dann gilt für den Erwartungswert von X, Y : $\int\limits_0^1 \tan(\frac{\pi x}{2})dx = \frac{2}{\pi}(-1)\log(\cos(\frac{\pi x}{2}))|_0^1 = \infty$.
Also ist
$\mathbb{E}(X+Y+Z) = \mathbb{E}(2\tan(\frac{\pi W}{2}) - 2\tan(\frac{\pi W}{2})) = \mathbb{E}(0) = 0 \neq \mathbb{E}(X)+\mathbb{E}(Y)+\mathbb{E}(Z) = \infty - \infty$

1.5.3 Über die Existenz des Momentes der Ordnung (-1)

Unter welchen Voraussetzungen exisitiert $\mathbb{E}(X)^{-1}$? Das Beispiel zeigt, dass es eine hinreichende, aber nicht notwendige Bedingung hierfür gibt. Die folgende Bedingung wurde von Piegorsch und Casella (1985) bewiesen.

Satz 1.5.3.1
Sei X eine ZV mit Dichtefunktion $f(x), x \in (0, \infty)$, welche stetig ist und folgende Grenzwertbeziehung erfüllt:

$$\lim_{x \to 0} x^{-\alpha} f(x) = 0 \text{ für ein } \alpha > 0$$

Dann gilt: $\mathbb{E}(X)^{-1} < \infty$.

Wir zeigen, dass Satz 1.5.3.1 für die Existenz von $\mathbb{E}(X)^{-1}$ nicht notwendig sein muss. Sei $\{f_n, n \geq 1\}$ mit:

$f_n(x) = \frac{1}{|\log^n(x)| \int_0^c |\log^n u|^{-1} du}, \quad c \in (0,1), \quad 0 < x < c.$
Somit ist f_n eine Dichtefunktion einer ZV X_n, da diese positiv ist und durch die Normierung mittels dem Integral im Nenner diese zu 1 im Bereich $(0, c)$ integriert wird. Zudem besitzt die Funktionenfolge $\{f_n\}_{n \in \mathbb{N}}$ folgende Eigenschaften:

- $\int_0^c |\log^n u|^{-1} du = -\int_0^c (\log(u))^{-n} du = -\int_0^c (\frac{1}{\log(u)})^n du \leq -\int_0^c (\frac{1}{x})^n dx \overset{n \geq 1}{=} \frac{1}{n-1} x^{-n+1}|_0^c = \frac{1}{n-1} c^{1-n} < \infty$
- $\lim_{x \to 0} f_n(x) = 0, \quad \lim_{x \downarrow 0} x^{-\alpha} f_n(x) = \infty \quad \forall \alpha > 0$

Es ist offensichtlich, dass das Integral $\int_0^c |\log^n u|^{-1} du$ für $n = 1$ divergiert $\forall c \in (0, 1)$. Dennoch existiert $\mathbb{E}(X_n)^{-1}$ für $n \geq 2, n \in \mathbb{N}$, obwohl die Voraussetzung des Satzes nicht erfüllt ist.

$$\mathbb{E}(X_n)^{-1} = \frac{1}{\int_0^c |\log(u)^n|^{-1} du} \int_0^c (x|\log^n x|)^{-1} dx = \underbrace{\frac{1}{\int_0^c |\log(u)^n|^{-1} du}}_{<\infty} \left(\frac{1}{1-n} (\log(x))^{1-n}\right)|_0^c < \infty$$

1.5.4 Über die Gültigkeit des Fubini-Theorems

Seien $(\Omega_1, \mathscr{A}_1, P_1)$ und $(\Omega_2, \mathscr{A}_2, P_2)$ zwei Wahrscheinlichkeitsräume. Dann gibt es genau ein Wahrscheinlichkeitsmaß P, definiert auf dem messbaren Raum $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \times \mathscr{A}_2)$, sodass gilt:

$$P(A_1 \times A_2) = P_1(A_1)P_2(A_2), \quad A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2$$

Zudem gilt für jede nicht-negative Zufallsvariable X, definiert auf $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \times \mathscr{A}_2, P)$

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} X(\omega_1, \omega_2) P(d(\omega_1, \omega_2)) &= \int_{\Omega_1} \int_{\Omega_2} X_{\omega_1}(\omega_2) P_2(d\omega_2) P_1(d\omega_1) \\ &= \int_{\Omega_2} \int_{\Omega_1} X_{\omega_2}(\omega_1) P_1(d\omega_1) P_2(d\omega_2), \end{aligned}$$

falls $\int X dP$ existiert. Die Existenz ist eine wesentliche Voraussetzung des Satzes von Fubini und somit der Gültigkeit von (1). Dies kann anhand eines Gegenbeispieles gezeigt werden. Sei Z eine nicht-integrierbare, nicht-negative Zufallsvariable auf $(\Omega_1, \mathscr{A}_1, P_1)$. Definiere X auf dem Raum $\Omega_1 \times \{0,1\}$ mit zugehöriger $\sigma-$Algebra $\mathscr{A} = \mathscr{B}(\Omega_1 \times \{0,1\})$. Für die Argumente $(\omega, 0), (\omega, 1) \in \Omega \times \{0,1\}$ gelte folgende Identität :

$$X(\omega, 0) \overset{f.s.}{=} Z(\omega), \quad X(\omega, 1) \overset{f.s.}{=} -Z(\omega)$$

Man erkennt, dass die so definierte Zufallsvariable (1) nicht erfüllt, da die erste Gleichung 0 ergibt und die letzte ∞.

Zu guter Letzt betrachten wir noch ein Beispiel aus dem Bereich der bedingten Erwartung.

1.5.5 Über die Identität $\mathbb{E}(\mathbb{E}(X \mid Y)) = \mathbb{E}(X)$

Damit $\mathbb{E}(\mathbb{E}(X \mid Y)) = \mathbb{E}(X)$ gilt, wird die Existenz von $\mathbb{E}(|X|)$ vorausgesetzt. Wir wollen diese Voraussetzung anhand eines Beispieles überprüfen. Sei nun Y eine positive Zufallsvariable mit Parameter $\nu > 0$ und folgender Dichte:

$$(1) \qquad g_\nu(y) = (\tfrac{1}{2}\nu)^{\frac{1}{2}\nu} (\Gamma(\tfrac{1}{2}\nu))^{-1} y^{\frac{1}{2}\nu - 1} e^{-\frac{1}{2}\nu y}, \quad y > 0$$

Man erkennt, dass $Y \sim \Gamma(\frac{1}{2}\nu, \frac{1}{2}\nu)$. Angenommen die bedingte Dichte von X gegeben $Y = y$ hat für $y > 0$ folgende Gestalt:

$$(2) \qquad f(x|y) = (2\pi)^{-\frac{1}{2}} y^{\frac{1}{2}} e^{-\frac{1}{2}yx^2}, \quad x \in \mathbb{R}$$

Folglich gilt:

$$\mathbb{E}(X|Y=y) = \int_{-\infty}^{\infty} x f(x|y) dx = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} y^{\frac{1}{2}} e^{-\frac{1}{2}yx^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{yx}{y^{\frac{1}{2}}} e^{-\frac{1}{2}yx^2} dx =$$
$$\frac{-1}{\sqrt{2\pi y}} \int_{-\infty}^{\infty} (-yx) e^{-\frac{1}{2}yx^2} dx = -\frac{1}{\sqrt{2\pi y}} e^{-\frac{1}{2}yx^2} \Big|_{x=-\infty}^{\infty} = 0$$

Des Weiteren folgt aus (1) und der Annahme (2), dass X folgende Dichte besitzen muss:

$$(3) \qquad h_\nu(x) = \Gamma(\tfrac{1}{2}(\nu+1))(\Gamma(\tfrac{\nu}{2})\sqrt{\pi\nu})^{-1}(1+\tfrac{x^2}{\nu})^{-\frac{1}{2}(\nu+1)}, \quad x \in \mathbb{R}$$

X ist hierbei Student-verteilt mit $\nu > 0$ Freiheitsgraden. Für $\nu = 1$ entspricht dies der Dichte der Cauchy-Verteilung. Bekanntlich existiert deren Erwartungswert nicht. Also gilt nun insgesamt:

$$\mathbb{E}(\mathbb{E}(X|Y)) \overset{\mathbb{E}(X|Y)=0}{=} \mathbb{E}(0) = 0 \overset{\nu=1}{\neq} \mathbb{E}(X) = \infty$$

Der Grund für diesen angeblichen Widerspruch liegt darin, dass wir mit (2) begonnen haben und daraus $\mathbb{E}(X|Y) = 0$ folgern konnten. Aus (1) und (2) konnten wir (3) ableiten, welche eine Dichte einer nicht-integrierbaren Zufallsvariablen ist.

Literatur

[1] Spodarev, Evgeny. Elementare Wahrscheinlichkeitstheorie und Statistik, Vorlesungsskript. Universität Ulm (2011)

[2] Spodarev, Evgeny. Stochastik I, Vorlesungsskript. Universität Ulm (2009)

[3] Stoyanov, Jordan. Counterexamples in Probability, 2nd Edition. Wiley, Chichester [u.a.] 1997.