8.8 Funktionen von Zufallsvariablen

Sei X eine Zufallsvariable mit Dichte $f(x)$ und $\phi(x)$ eine reellwertige Funktion, suchen wir die Verteilung von

$$Y = \phi(X)$$

(i) den Wertebereich von Y bestimmen
(ii) die Verteilungsfunktion von $Y, G(y)$ finden

$$G(y) = P(Y \leq y)$$

$$= P(\phi(X) \leq y)$$

(iii) $g(y) = \frac{\partial G}{\partial y}$

Beispiel

$$f(x) = 3x^2 \quad 0 \leq x \leq 1$$

$$\phi(x) = \log_e(x)$$

$$(i) \quad Y = \log_e X \Rightarrow -\infty < y \leq 0$$

$$(ii) \quad G_Y(y) = P(\log_e X \leq y) = P(X \leq e^y) = e^{3y}$$

$$(iii) \quad g(y) = 3e^{3y} \quad -\infty < y \leq 0$$

$$= 0 \quad \text{sonst}$$

8.8.1 Die Verteilung einer Verteilungsfunktion

$$\phi(x) = F(x)$$

$$(i) \qquad Y = F(X) \Rightarrow 0 \leq y \leq 1$$

$$(ii) \quad G_Y(y) = P(Y \leq y) = P(F_X(X) \leq y)$$

Nehmen wir an, dass $F_X(X)$ umkehrbar ist

$$= P(X \leq F_X^{-1}(y))$$

$$= y$$

$$\Rightarrow Y \sim U(0,1)$$

Wenn die Verteilungsfunktion von einer Zufallsvariable X umkehrbar ist, ist $F(X)$ eine gleichverteilte Zufallsvariable.

8.8.2 Formel

Im allgemeinen, für $\phi(x)$ umkehrbar und strikt monoton und differenzierbar gilt

$$g(y) = f(\phi^{-1}(y)) \left|\frac{\partial x}{\partial y}\right| \qquad (1)$$

8.8.3 Die Ungleichung von Jensen

Sei X eine Zufallsvariable, $E[X] < \infty$ und $g(x)$ eine konvexe Funktion

$$[\exists \quad \lambda(a) \quad s.t. \quad g(x) \geq g(a) + \lambda(a)(x-a) \quad \forall a]$$

Dann gilt $E[g(X)] \geq g(E[X])$

Beweis:
Setzen wir $a = E[X]$

$$g(x) \geq g(E[X]) + \lambda(E[X])(x - E[X])$$

$$E[g(X)] \geq g(E[X]) + \lambda(E[X])(E[X] - E[X])$$

8.8.3.1 Jensen Beispiel

Sei

$$g(x) = -\log x$$

$$\Rightarrow E[-\log X] \geq -\log E[X]$$

$$E[\log X] \leq \log E[X]$$

Sei $P(X = x_i) = p_i \quad i = 1, \ldots n \quad x_i > 0 \quad \forall i$
$\log E[X] = \log \sum_{i=1}^{n} x_i p_i \geq E[\log X] = \sum_{i=1}^{n} p_i \log x_i$

$$\Rightarrow \sum_{i=1}^{n} x_i p_i \geq \prod_{i=1}^{n} x_i^{p_i}$$

Für $p_i = \frac{1}{n} \quad \forall i$ haben wir bewiesen, dass
der arithmetische Mittel $\geq$ der geometrische Mittel

8.8.4 Nichtmonotone Funktionen

$$\phi(x) = x^2$$

$$\text{für} \quad X \sim U(-1, 1)$$

$$\begin{aligned} f(x) &= 0.5 \quad -1 \leq x \leq 1 \\ &= 0 \qquad \text{sonst} \end{aligned}$$

Wertebereich für $Y = X^2$ ist $0 \leq y \leq 1$

$$\begin{aligned} G(y) &= P(Y \leq y) \\ &= P(X^2 \leq y) \\ &= P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= F(\sqrt{y}) - F(-\sqrt{y}) \end{aligned}$$

$$\Rightarrow g(y) = \frac{1}{2\sqrt{y}}$$

8.9 Kombinationen von Zufallsvariablen

8.9.1 Faltungen

Seien X_1, X_2 unabhängige reellwertige Variablen deren Verteilungen Dichten f_1, f_2 haben. Die Verteilung von $S = X_1 + X_2$ ist

$$P(S \leq s) = \int\int_{x_1+x_2 \leq s} f_1(x_1) f_2(x_2) \mathrm{d}x_1 \mathrm{d}x_2$$

Mit der Substitution

$$u = x_1 + x_2 \quad v = x_2$$

$$P(S \leq s) = \int_{-\infty}^{s} \left(\int_{-\infty}^{\infty} f_1(u-v) f_2(v) \mathrm{d}v \right) \mathrm{d}u$$

$$(f_1 * f_2)(u) = \int_{-\infty}^{\infty} f_1(u-v) f_2(v) \mathrm{d}v \qquad (1)$$

heißt die Faltung von f_1 und f_2 und ist die Dichte von $X_1 + X_2$.

8.9.2 Exponentialverteilungen

$$X \sim E(\lambda) \quad Y \sim E(\lambda) \quad \text{u.i.v.}$$

$$W = X + Y$$

$$h(w) = \int_0^w \lambda e^{-\lambda x} \lambda e^{-\lambda(w-x)} \mathrm{d}x$$

$$= \lambda^2 w e^{-\lambda w}$$

Für

$$X_i \sim E(\lambda) \quad i = 1, \ldots n \quad \text{u.i.v.}$$

gilt im allgemeinen

$$W = \sum_{i=1}^{n} X_i \sim \text{Erlang}(n, \lambda)$$

$$h(w) = \frac{(\lambda w)^{n-1}}{(n-1)!} \lambda e^{-\lambda w}$$

eine Sonderform der Gamma Verteilung

$$h(w) = \frac{(\lambda)^r}{\Gamma(r)} w^{r-1} e^{-\lambda w}$$

8.9.3 Gleichverteilungen

$$X \sim U(0,1) \quad Y \sim U(0,1) \quad \text{u.i.v.}$$

$$W = X + Y$$

mit Wertebereich

$$0 \le w \le 2$$

$$\begin{aligned}
h(w) &= \int_0^w f(x)g(w-x)\mathrm{d}x \quad 0 \le w \le 1 \\
&= \int_0^w \mathrm{d}x \\
&= w \\
&= \int_{w-1}^1 f(x)g(w-x)\mathrm{d}x \quad 1 \le w \le 2 \\
&= \int_{w-1}^1 \mathrm{d}x \\
&= 2 - w
\end{aligned}$$

8.9.4 Normalverteilungen cf. §7.4.4

$$X \sim N(0,1) \quad Y \sim N(0,1) \quad \text{u.i.v.}$$

$$W = X + Y$$

$$h(w) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(w-x)^2} \mathrm{d}x$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(2x^2-2xw+w^2)} \mathrm{d}x$$

$$2x^2 - 2xw + w^2 = \left(\sqrt{2}x - \frac{w}{\sqrt{2}}\right)^2 + \frac{1}{2}w^2$$

$$h(w) = \frac{e^{-\frac{1}{4}w^2}}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\sqrt{2}x - \frac{w}{\sqrt{2}})^2} \mathrm{d}x$$

Setzen wir $s = \sqrt{2}x - \frac{w}{\sqrt{2}}$ dann ist $h(w) = \frac{e^{-\frac{1}{4}w^2}}{2\sqrt{\pi}}$ und

$$W \sim N(0,2)$$

8.9.5 Normalverteilungen im allgemeinen

$$X \sim N(\mu_1, \sigma_1^2) \quad Y \sim N(\mu_2, \sigma_2^2) \quad \text{unabhängig}$$

$$\Rightarrow W = X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

Weiterhin gilt

$$V = X - Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$$

und für $\{X_i, i = 1, \ldots n\}$ u.i.v. mit $X_i \sim N(\mu, \sigma^2)$

$$W = \sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2)$$

Aber für $W = nX$ gilt

$$W \sim N(n\mu, n^2\sigma^2)$$

Für $W = \frac{1}{n} \sum\limits_{i=1}^{n} X_i$ gilt

$$W \sim N(\mu, \frac{\sigma^2}{n})$$

8.9.6 Beispiele cf. §8.3

Leute im Aufzug

Sei W_i das Gewicht einer Person i und sei

$$W_i \sim N(\mu, \sigma^2)$$

Das Gesamtgewicht von n (unabhängig ausgewählten) Leuten $W = \sum_{i=1}^{n} W_i$ hat dann die Verteilung

$$N(n\mu, n\sigma^2)$$

Wir suchen n, so dass für eine vorgegebene Grenze L

$$P(W \leq L) \geq 0.999$$

Erlaubtes Gepäck

Gegeben N Sitze und eine Gesamtgewichtsgrenze für Passagiere und Gepäck von T, suchen wir b so dass

$$P(W = \sum_{i=1}^{N} W_i \leq T - bN) \geq 0.999$$

Schokoriegel

Sei V das Gewicht eines Schokoriegels und $V \sim N(\mu, \sigma^2)$. Auf der Hülle steht das Gewicht s. Wir sollten μ und σ so setzen, dass $P(V \geq s) \geq 0.95$.

8.9.7 Beispiel — Gleichverteilung

Ein Stück Holz der Länge 1 wird an einer zufälligen Stelle gebrochen. Es bleiben zwei Teile, X und $Y = 1 - X$. O.B.d.A. $X \geq Y$.

Wenn wir eine Gleichverteilung für die Bruchstelle annehmen, ist die Dichte von X

$$f(x) = 2 \qquad 0.5 \leq x \leq 1$$

Wir können Fragen beantworten, wie

$$P(X > 2Y) =$$

$$P(X - Y > 0.5) =$$

$$P(X \quad \text{und} \quad Y > 0.3333) =$$

Wenn wir den größeren Teil nochmals zufällig brechen und den größeren Teil davon Z nennen, was ist

$$P(Z > Y)?$$

Die Dichte von Z

$$g(z) = \frac{2}{x} \qquad \frac{x}{2} \leq z \leq x$$

$$P(Z > 1 - X) = \int_{\frac{1}{2}}^{1} 2 \int_{\max(\frac{x}{2}, 1-x)}^{x} \frac{2}{x} \mathrm{d}z \mathrm{d}x$$

$$\frac{x}{2} > 1 - x \quad \text{wenn} \quad x > \frac{2}{3}$$

$$P(Z > 1 - X) = \int_{\frac{1}{2}}^{\frac{2}{3}} \frac{4}{x} \int_{(1-x)}^{x} \mathrm{d}z \mathrm{d}x + \int_{\frac{2}{3}}^{1} \frac{4}{x} \int_{\frac{x}{2}}^{x} \mathrm{d}z \mathrm{d}x$$

$$= \int_{\frac{1}{2}}^{\frac{2}{3}} \frac{4}{x}(2x - 1) \mathrm{d}x + \int_{\frac{2}{3}}^{1} \frac{4}{x} \frac{x}{2} \mathrm{d}x$$

$$= [8 - 4 \log x]_{\frac{1}{2}}^{\frac{2}{3}} + \frac{2}{3}$$

$$= 2 + 4 \log \frac{3}{4} = 0.8493$$