9.1 Bedingte Wahrscheinlichkeiten im stetigen Fall

(cf. §4.1 - §4.3 für den diskreten Fall)

Sei $f_{XY}(x,y)$ die gemeinsame Dichte der Zufallsvariablen X und Y. Aus Satz 7.7.1 gilt, dass

$$f_{XY}(x,y) = f_X(x) f_Y(y)$$

nur, wenn X und Y unabhängig sind. Sonst verwenden wir

$$\begin{aligned}
P(x_1 < X \le x_2) &= \int_{x_1}^{x_2} \int_y f_{XY}(x,y) \mathrm{d}y \mathrm{d}x \\
&= \int_{x_1}^{x_2} \int_y f_{Y|X}(y|x) f_X(x) \mathrm{d}y \mathrm{d}x \\
&= \int_{x_1}^{x_2} \left(\int_y f_{Y|X}(y|x) \mathrm{d}y \right) f_X(x) \mathrm{d}x \\
&= \int_{x_1}^{x_2} f_X(x) \mathrm{d}x
\end{aligned}$$

$f_X(x)$ ist die Randverteilung (Marginalverteilung) von X
$f_{Y|X}(y|x)$ ist die bedingte Verteilung von $Y|X$

9.2 Bayes im stetigen Fall

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

oder

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{\int_x f_{Y|X}(y|x)f_X(x)\mathrm{d}x}$$

z.B. Die Leistungen zwei Sportler, A und B, werden im Voraus mit den Verteilungen $f_A(t_A)$ und $f_B(t_B)$ geschätzt. Wenn Sie nachher nur die Information haben, dass A tatsächlich einen kleineren Wert als B erreichte, wie sieht Ihre Verteilung für A dann aus?

$$f(t_A|t_A < t_B) = \frac{f_A(t_A)\int_{t_A}^{\infty} f_B(t_B)\mathrm{d}t_B}{\int_A f_A(t_A)\int_{t_A}^{\infty} f_B(t_B)\mathrm{d}t_B\mathrm{d}t_A}$$

Mit $T_A \sim E(\lambda_A)$ und $T_B \sim E(\lambda_B)$ findet man

$$f(t_A|t_A < t_B) = (\lambda_A + \lambda_B)e^{-(\lambda_A+\lambda_B)}$$

9.3 Längen von Intervallen und Wartezeiten

9.3.1 Durchschittliche Intervalllänge Sei die Dichte der Länge eines Zeitintervalls, X, $f(x)$ sein ($x \geq 0$). Wir wählen per Zufall einen Zeitpunkt aus, wie groß ist das Intervall um diesen Punkt?

$$g(y) \propto yf(y) = \frac{yf(y)}{E[X]}$$

weil

$$\int_y g(y)\mathrm{d}y = 1 \quad \text{und} \quad \int_x xf(x)\mathrm{d}x = E[X]$$

Dann

$$E[Y] = \int_o^\infty y^2 f(y)\mathrm{d}y = \frac{E[X^2]}{E[X]} = \frac{V[X] + E[X]^2}{E[X]}$$

$$\text{z.B.} \quad X \sim E(\lambda) \Rightarrow E[Y] = \frac{2}{\lambda}$$

Der Durchschnitt der getroffenen Intervalle ist zweimal so lang wie das durchschnittliche Intervall.

9.3.2 Verteilung von Wartezeiten

$$
\begin{aligned}
P(\text{ Wartezeit } \le t | \text{ Länge } = y) &= 1 \qquad y \le t \\
&= \frac{t}{y} \qquad y > t
\end{aligned}
$$

Hier wird angenommen, dass die Ankunfstzeit innerhalb eines Intervalls gleichverteilt ist — was sollte man sonst annehmen?

$$
\begin{aligned}
H(t) &= P(\text{ Wartezeit } \le t) \\
&= \int_0^t g(y) \mathrm{d}y + \int_t^\infty \frac{t}{y} g(y) \mathrm{d}y \\
&= \int_0^t \frac{x}{\mu_X} f(x) \mathrm{d}x + \int_t^\infty \frac{t}{\mu_X} f(x) \mathrm{d}x \\
&= \frac{1}{\mu_X} \left(\int_0^t x f(x) \mathrm{d}x + t(1 - F(t)) \right)
\end{aligned}
$$

und

$$
\begin{aligned}
\frac{\mathrm{d}H}{\mathrm{d}t} &= h(t) \\
&= \frac{1}{\mu_X} (t f(t) + 1 - F(t) - t f(t)) \\
&= \frac{1 - F(t)}{\mu_X}
\end{aligned}
$$

$$\text{wo} \quad \mu_X = E[X]$$

Wartezeiten Beispiel

Sei $X \sim E(\lambda)$

$$\begin{aligned} f(x) &= \lambda e^{-\lambda t} \\ h(t) &= \frac{1}{\mu_X} e^{-\lambda t} \\ &= \lambda e^{-\lambda t} \end{aligned}$$

also auch $E(\lambda)$. Dank der Symmetrie hat die Zeit seit dem letzten Bus auch die Exponentialverteilung.

Ist es möglich, dass E[Wartezeit] $>$ E[Intervalllänge]?

$$\begin{aligned} E[W] &= \int_0^\infty \frac{t(1-F(t))}{\mu_X} \mathrm{d}t \\ &= \frac{1}{\mu_X}\left[\frac{1}{2}t^2(1-F(t)\right]_0^\infty + \frac{1}{2\mu_X}\int_0^\infty t^2 f(t)\mathrm{d}t \\ &= \frac{1}{2}\frac{E[X^2]}{E[X]} \end{aligned}$$

$$\Rightarrow E[W] > E[X] \text{ wenn } E[X^2] > 2E[X]^2$$

z.B. in der Familie

$$f(x) = p\lambda e^{-\lambda x} + (1-p)\mu e^{-\mu x} \qquad 0 < p < 1$$

Mischungen von Exponentialverteilungen

10 Momente und Charakteristische Funktionen

$E[X^k]$ ist das k.Moment einer Verteilung.
$E[(X - E[X])^k]$ ist das k. zentrierte Moment.

z.B. (1) $\quad X \sim U(0,1)$

$$\begin{aligned} E[X^k] &= \int_0^1 x^k \mathrm{d}x \\ &= \frac{1}{k+1} \\ E[(X-E[X])^k] &= \int_0^1 \left(x - \frac{1}{2}\right)^k \mathrm{d}x \\ &= 0 \qquad\qquad k = 2m+1 \\ &= \frac{(1/2)^k}{k+1} \qquad k = 2m \end{aligned}$$

(2) $\quad X \sim N(0,1)$

$$\begin{aligned} E[X^k] &= 0 \qquad\qquad k = 2m+1 \\ &= (k-1)(k-3).....1 \qquad k = 2m \end{aligned}$$

10.1 Momenterzeugende Funktionen

Für diskrete Zufallsvariablen kann man eine Verteilung durch die WEF zusammenfassen. Für stetige ZV hat man an die MEF dafür gedacht:

$$M_X(t) = E[e^{tX}] \qquad (1)$$

Für X stetig mit Dichte f ist die MEF mit einer Laplace Transformation verwandt da

$$M(t) = \int e^{tx} f(x) \mathrm{d}x$$

Falls $M(t) < \infty$ in einem offenen Intervall um 0, gilt nach Taylor

$$M(t) = \sum_{k=0}^{\infty} \frac{E[X^k]}{k!} t^k \qquad (2)$$

Beispiele: (1) $X \sim U(0,1) \Rightarrow E[e^{tX}] = \frac{1}{t}(e^t - 1)$

und (2) $X \sim E[\lambda] \Rightarrow E[e^{tX}] = \left(1 - \frac{t}{\lambda}\right)^{-1} \quad (t < \lambda)$

Es gibt zwei Probleme mit diesen Funktionen:

(a) Das bestimmende Integral ist nicht immer endlich.

(b) Verteilungen werden nicht eindeutig durch Momente definiert.

10.1.1 Lognormal Gegenbeispiel

Sei X lognormalverteilt. Dann ist $Y = \log_e X \sim N(\mu, \sigma^2)$. Für $Y \sim N(0,1)$ hat X die Dichte

$$f(x) = \frac{1}{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(\log x)^2} \qquad x > 0$$

Betrachten wir die Funktion

$$f_a(x) = f(x)(1 + a\sin(2\pi\log x)) \quad -1 \leq a \leq 1$$

Wir zeigen, dass f_a eine Dichte ist, die dieselben Momente als f hat. Da $f_a(x) \geq 0$ genügt es zu zeigen, dass

$$D = \int_0^\infty x^k f(x)\sin(2\pi\log x)\mathrm{d}x = 0 \quad k = 0, 1, 2, \ldots$$

Nach der Transformation $t = \log_e x$ haben wir

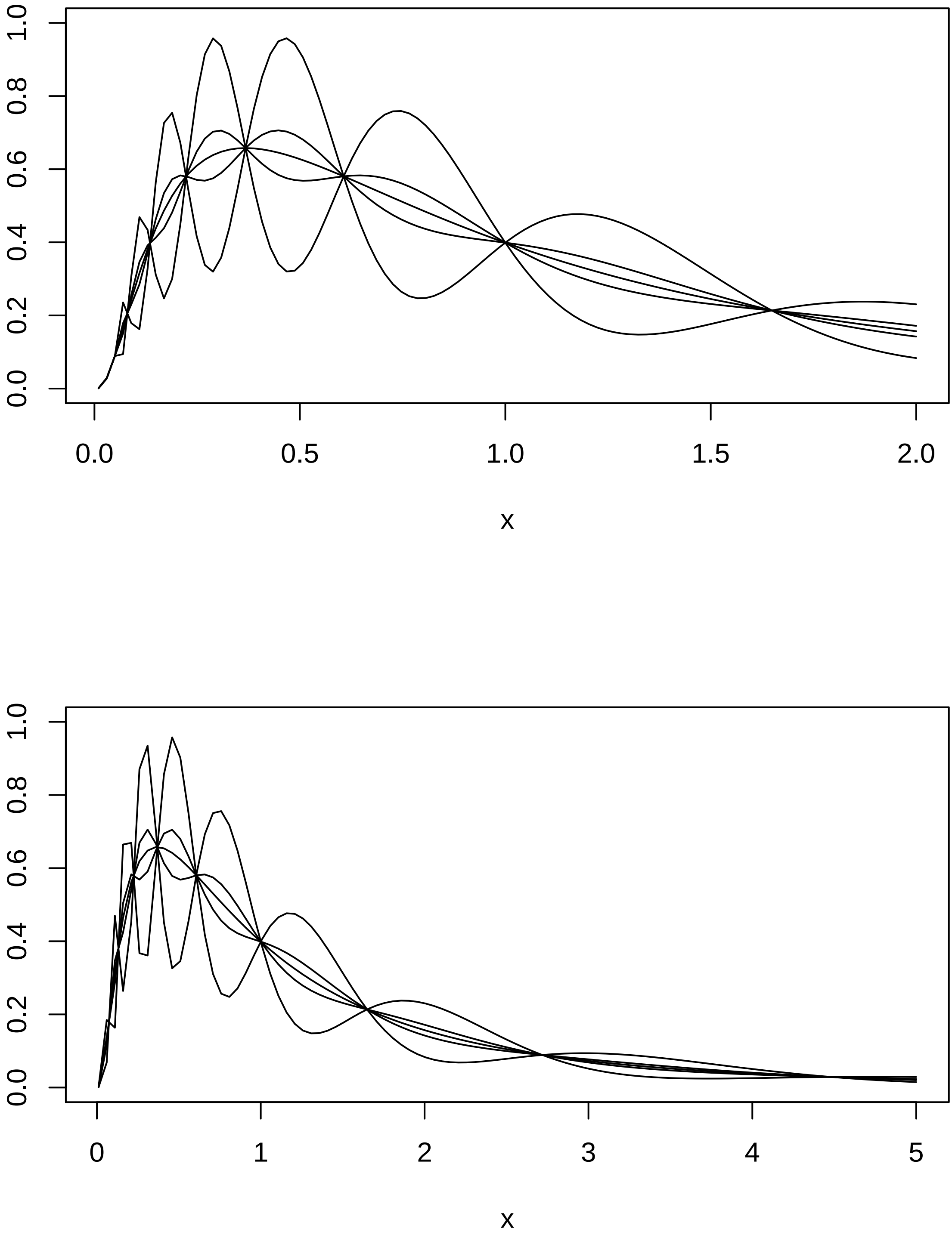
$$D = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2 + kt}\sin 2\pi t\mathrm{d}t$$

Nach der weiteren Transformation $y = t - k$ erhalten wir

$$D = e^{-\frac{1}{2}k^2}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2}\sin 2\pi y\mathrm{d}y = 0$$

weil die Funktion ungerade ist.

Lognormalvariante: Dichten mit gleicher MGF (a=−0.5,−0.1,0,0.1,0.5)



10.2 Charakteristische Funktionen

$$\zeta_X(t) = E[e^{itX}] = \sum_{k=0}^{\infty} \frac{(it)^k}{k!} E[X^k]$$

ist die Charakteristische Funktion der ZV X.

1. $\zeta_X(t)$ existiert $\forall t \in \mathbb{R}$

2. $\zeta_X(t)$ ist die Fouriertransformation der Dichte

3. Die Beziehung zwischen F_X und $\zeta_X(t)$ ist umkehrbar eindeutig.

z.B. $X_i \sim U(0,1) \quad i = 1, \dots n$ u.i.v.
Die C.F. von X_i ist $\frac{1}{it}(e^{it} - 1)$
und die C.F. von der Summe $S = \sum_{i=1}^{n} X_i$ ist

$$\left(\frac{1}{it}\right)^n (e^{it} - 1)^n$$

10.2.1 Resultate für Charakteristische Funktionen

- Für unabhängige $X_1, X_2, \ldots, X_n$ ist die charakteristische Funktion der Summe

$$S = \sum_{j=1}^{n} X_j$$

das Produkt der einzelnen c.f.'s

$$\zeta_S(t) = E\left[e^{it\sum X_j}\right] = \prod_{j=1}^{n} E\left[e^{itX_j}\right] = \prod_{j=1}^{n} \zeta_{X_j}(t)$$

- Sei X eine reellwertige ZV mit c.f. $\zeta_X(t)$ und sei $E[X^k] < \infty$ für $k = 1, 2, \ldots$ dann folgt

$$i^k E[X^k] = \zeta_X^(k)(0)$$

- Seien X, Y unabhängige ZV mit c.F. $\zeta_X(t), \zeta_Y(t)$

$$a, b \in \mathbb{R} \quad Z = aX + b \Rightarrow \zeta_Z(t) = e^{itb}\zeta_X(t)$$

$$\zeta_{X+Y}(t) = \zeta_X(t)\zeta_Y(t)$$

z.B.

$$X \sim N(\mu_X, \sigma_X^2), \quad Y \sim N(\mu_Y, \sigma_Y^2)$$

und

$$S = X + Y$$

$$\zeta_X(t) = e^{i\mu_X t - \frac{1}{2}\sigma_X^2 t^2}$$

$$\zeta_{X+Y}(t) = e^{i(\mu_X+\mu_Y)t - \frac{1}{2}(\sigma_X^2+\sigma_X^2)t^2}$$

$$\Rightarrow S \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_X^2)$$

wegen der Eindeutigkeit der c.F.

- Sei X eine reellwertige ZV mit Dichte f und sei g eine stetige Funktion auf $\mathbb{R}$ mit $Y = g(X)$ dann

  $$\zeta_Y(t) = E_Y[e^{itY}] = E_X[e^{itg(X)}] = \int_{-\infty}^{\infty} e^{itg(x)} f(x) \mathrm{d}x$$

- Sei f die Dichte von der ZV X, dann ist

  $$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \zeta_X(t) \mathrm{d}t$$

  $\forall x$ wo f differenzierbar ist.