SS 2004

Diskrete Strukturen II

Ernst W. Mayr

Fakultät für Informatik

TU München

http://www14.in.tum.de/lehre/2004SS/ds/index.html.de

Exponentialverteilung als Grenzwert der geometrischen Verteilung

Erinnerung: Die Poisson-Verteilung lässt sich als Grenzwert der Binomialverteilung darstellen.

Wir betrachten eine Folge geometrisch verteilter Zufallsvariablen X_n mit Parameter $p_n = \lambda/n$. Für ein beliebiges $k \in \mathbb{N}$ ist die Wahrscheinlichkeit, dass $X_n \leq k \cdot n$, gleich

$$\begin{aligned}
\Pr[X_n \leq kn] &= \sum_{i=1}^{kn} (1 - p_n)^{i-1} \cdot p_n = p_n \cdot \sum_{i=0}^{kn-1} (1 - p_n)^i \\
&= p_n \cdot \frac{1 - (1 - p_n)^{kn}}{p_n} = 1 - \left(1 - \frac{\lambda}{n}\right)^{kn} .
\end{aligned}$$

Wegen $\lim_{n\to\infty}(1-\frac{\lambda}{n})^n = e^{-\lambda}$ gilt daher für die Zufallsvariablen $Y_n := \frac{1}{n}X_n$, dass

$$\begin{aligned}
\lim_{n\to\infty} \Pr[Y_n \le t] &= \lim_{n\to\infty} \Pr[X_n \le t\cdot n] \\
&= \lim_{n\to\infty} 1 - \left(1-\frac{\lambda}{n}\right)^{tn} \\
&= 1 - e^{-\lambda t}.
\end{aligned}$$

Die Folge Y_n der (skalierten) geometrisch verteilten Zufallsvariablen geht also für $n \to \infty$ in eine exponentialverteilte Zufallsvariable mit Parameter λ über.

2.3 Mehrere kontinuierliche Zufallsvariablen

Mehrdimensionale Dichten

Definition: Zu zwei kontinuierlichen Zufallsvariablen X, Y wird der zugrunde liegende gemeinsame Wahrscheinlichkeitsraum über $\mathbb{R}^2$ durch eine integrierbare (gemeinsame) Dichtefunktion $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}_0^+$ mit

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) \,\mathrm{d}\, x \,\mathrm{d}\, y = 1$$

beschrieben.

Für ein Ereignis $A \subseteq \mathbb{R}^2$ (das aus abzählbar vielen geschlossenen oder offenen Bereichen gebildet sein muss) gilt

$$\Pr[A] = \int_A f_{X,Y}(x, y) \,\mathrm{d}\, x \,\mathrm{d}\, y.$$

Unter einem Bereich B verstehen wir dabei Mengen der Art

$$B = \{(x,y) \in \mathbb{R}^2 \mid a \le x \le b, c \le y \le d\} \quad \text{mit } a,b,c,d \in \mathbb{R}.$$

Dabei können die einzelnen Intervallgrenzen auch „offen“ sein.

Analog zum eindimensionalen Fall ordnen wir der Dichte $f_{X,Y}$ eine (gemeinsame) Verteilung $F_{X,Y} : \mathbb{R}^2 \to [0,1]$ zu:

$$F_{X,Y}(x,y) = \Pr[X \le x, Y \le y] = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(u,v)\,\mathsf{d}\,u\,\mathsf{d}\,v.$$

Randverteilungen und Unabhängigkeit

Definition: Sei $f_{X,Y}$ die gemeinsame Dichte der Zufallsvariablen X und Y. Die Randverteilung der Variablen X ist gegeben durch

$$F_X(x) = \Pr[X \leq x] = \int_{-\infty}^{x} \left[\int_{-\infty}^{\infty} f_{X,Y}(u,v) \,\mathrm{d}\, v \right] \mathrm{d}\, u.$$

Analog nennen wir

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,v) \,\mathrm{d}\, v$$

die Randdichte von X. Entsprechende Definitionen gelten symmetrisch für Y.

Definition:

Zwei kontinuierliche Zufallsvariablen X und Y heißen unabhängig, wenn

$$\Pr[X \leq x, Y \leq y] = \Pr[X \leq x] \cdot \Pr[Y \leq y]$$

für alle $x, y \in \mathbb{R}$ gilt.

Dies ist gleichbedeutend mit

$$F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y)\,.$$

Differentiation ergibt

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)\,.$$

Für mehrere Zufallsvariablen $X_1, \ldots, X_n$ gilt analog: $X_1, \ldots, X_n$ sind genau dann unabhängig, wenn

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdot \ldots \cdot F_{X_n}(x_n)$$

bzw.

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n)$$

für alle $x_1, \ldots, x_n \in \mathbb{R}$.

Warten auf mehrere Ereignisse.

Satz 41: Die Zufallsvariablen $X_1, \ldots, X_n$ seien unabhängig und exponentialverteilt mit den Parametern $\lambda_1, \ldots, \lambda_n$. Dann ist auch $X := \min\{X_1, \ldots, X_n\}$ exponentialverteilt mit dem Parameter $\lambda_1 + \ldots + \lambda_n$.

Beweis: Der allgemeine Fall folgt mittels Induktion aus dem für $n = 2$. Für die Verteilungsfunktion F_X gilt:

$$\begin{aligned} 1 - F_X(t) &= \Pr[X > t] = \Pr[\min\{X_1, X_2\} > t] \\ &= \Pr[X_1 > t, X_2 > t] \\ &= \Pr[X_1 > t] \cdot \Pr[X_2 > t] \\ &= e^{-\lambda_1 t} \cdot e^{-\lambda_2 t} = e^{-(\lambda_1 + \lambda_2)t}. \end{aligned}$$

q. e. d.

Anschaulich besagt Satz 41, dass sich die Raten addieren, wenn man auf das erste Eintreten eines Ereignisses aus mehreren unabhängigen Ereignissen wartet. Wenn beispielsweise ein Atom die Zerfallsrate λ besitzt, so erhalten wir bei n Atomen die Zerfallsrate $n\lambda$ (wie uns auch die Intuition sagt).

Poisson-Prozess.

Wir hatten bei der Diskussion der geometrischen und der Poisson-Verteilung festgestellt:

Wenn der zeitliche Abstand der Treffer geometrisch verteilt ist, so ist ihre Anzahl in einer festen Zeitspanne binomialverteilt.

Im Grenzwert $n \longrightarrow \infty$, wobei wir die Trefferwahrscheinlichkeit mit $p_n = \lambda/n$ ansetzen, konvergiert die Binomialverteilung gegen die Poisson-Verteilung und die geometrische Verteilung gegen die Exponentialverteilung konvergiert. Im Grenzwert $n \longrightarrow \infty$ erwarten wir deshalb die folgende Aussage:

Wenn man Ereignisse zählt, deren zeitlicher Abstand exponentialverteilt ist, so ist die Anzahl dieser Ereignisse in einer festen Zeitspanne Poisson-verteilt.

Seien $T_1, T_2 \ldots$ unabhängige exponentialverteilte Zufallsvariablen mit Parameter λ. Die Zufallsvariable T_i modelliert die Zeit, die zwischen Treffer $i-1$ und i vergeht.

Für den Zeitpunkt $t > 0$ definieren wir

$$X(t) := \max\{n \in \mathbb{N} \mid T_1 + \ldots + T_n \leq t\}.$$

$X(t)$ gibt also an, wie viele Treffer sich bis zur Zeit t (von Zeit Null ab) ereignet haben. Es gilt:

Fakt 42: Seien $T_1, T_2, \ldots$ unabhängige Zufallsvariablen und sei $X(t)$ für $t > 0$ wie oben definiert. Dann gilt: $X(t)$ ist genau dann Poisson-verteilt mit Parameter $t\lambda$, wenn es sich bei $T_1, T_2, \ldots$ um exponentialverteilte Zufallsvariablen mit Parameter λ handelt.

Zum Zufallsexperiment, das durch $T_1, T_2, \ldots$ definiert ist, erhalten wir für jeden Wert $t > 0$ eine Zufallsvariable $X(t)$. Hierbei können wir t als Zeit interpretieren und $X(t)$ als Verhalten des Experiments zur Zeit t. Eine solche Familie $(X(t))_{t>0}$ von Zufallsvariablen nennt man allgemein einen stochastischen Prozess. Der hier betrachtete Prozess, bei dem $T_1, T_2, \ldots$ unabhängige, exponentialverteilte Zufallsvariablen sind, heißt Poisson-Prozess und stellt ein fundamentales und zugleich praktisch sehr bedeutsames Beispiel für einen stochastischen Prozess dar.

Beispiel: Wir betrachten eine Menge von Jobs, die auf einem Prozessor sequentiell abgearbeitet werden. Die Laufzeiten der Jobs seien unabhängig und exponentialverteilt mit Parameter $\lambda = 1/30[1/s]$. Jeder Job benötigt also im Mittel $30s$.

Gemäß Fakt 42 ist die Anzahl von Jobs, die in einer Minute vollständig ausgeführt werden, Poisson-verteilt mit Parameter $t\lambda = 60 \cdot (1/30) = 2$.

Die Wahrscheinlichkeit, dass in einer Minute höchstens ein Job abgearbeitet wird, beträgt $e^{-t\lambda} + t\lambda e^{-t\lambda} \approx 0{,}406$.

Summen von Zufallsvariablen

Satz 43:

Seien X und Y unabhängige kontinuierliche Zufallsvariablen. Für die Dichte von $Z := X + Y$ gilt

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) \cdot f_Y(z - x) \,\mathsf{d}\, x\,.$$

Beweis: Nach Definition der Verteilungsfunktion gilt

$$F_Z(t) = \Pr[Z \leq t] = \Pr[X + Y \leq t] = \int_{A(t)} f_{X,Y}(x, y) \,\mathsf{d}\, x \mathsf{d}\, y$$

wobei $A(t) = \{(x, y) \in \mathbb{R}^2 \mid x + y \leq t\}$.

Aus der Unabhängigkeit von X und Y folgt

$$\begin{aligned} F_Z(t) &= \int_{A(t)} f_X(x) \cdot f_Y(y)\,\mathrm{d}\,x\mathrm{d}\,y \\ &= \int_{-\infty}^{\infty} f_X(x) \cdot \left(\int_{-\infty}^{t-x} f_Y(y)\,\mathrm{d}\,y \right) \mathrm{d}\,x. \end{aligned}$$

Mittels der Substitution $z := x + y$, $\mathrm{d}z = \mathrm{d}y$ ergibt sich

$$\int_{-\infty}^{t-x} f_Y(y)\,\mathrm{d}\,y = \int_{-\infty}^{t} f_Y(z-x)\,\mathrm{d}\,z$$

und somit

$$F_Z(t) = \int_{-\infty}^{t} \left(\int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,\mathrm{d}\,x \right) \mathrm{d}\,z\,.$$

q. e. d.

Satz 44: (Additivität der Normalverteilung)

Die Zufallsvariablen $X_1, \ldots, X_n$ seien unabhängig und normalverteilt mit den Parametern μ_i, σ_i $(1 \leq i \leq n)$. Es gilt: Die Zufallsvariable

$$Z := a_1 X_1 + \ldots + a_n X_n$$

ist normalverteilt mit Erwartungswert $\mu = a_1\mu_1 + \ldots + a_n\mu_n$ und Varianz $\sigma^2 = a_1^2\sigma_1^2 + \ldots + a_n^2\sigma_n^2$.

Beweis: Wir beweisen zunächst den Fall $n = 2$ und $a_1 = a_2 = 1$. Nach Satz 43 gilt für $Z := X_1 + X_2$, dass

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_{X_1}(z-y) \cdot f_{X_2}(y) \,\mathrm{d}\, y \\ &= \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \underbrace{\left(\frac{(z-y-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right)}_{=:v} \right) \mathrm{d}\, y. \end{aligned}$$

Wir setzen

$$
\begin{aligned}
\mu &:= \mu_1 + \mu_2 \\
\sigma^2 &:= \sigma_1^2 + \sigma_2^2 \\
v_1 &:= (z - \mu)/\sigma \\
v_2^2 &:= v - v_1^2
\end{aligned}
$$

Damit ergibt sich

$$v_2^2 = \frac{(z - y - \mu_1)^2}{\sigma_1^2} + \frac{(y - \mu_2)^2}{\sigma_2^2} - \frac{(z - \mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}.$$

woraus wir

$$v_2 = \frac{y\sigma_1^2 - \mu_2\sigma_1^2 + y\sigma_2^2 - z\sigma_2^2 + \mu_1\sigma_2^2}{\sigma_1\sigma_2\sigma}$$

ermitteln.

Damit folgt für die gesuchte Dichte

$$f_Z(z) = \frac{1}{2\pi \cdot \sigma_1 \cdot \sigma_2} \cdot \exp\left(-\frac{v_1^2}{2}\right) \cdot \int_{-\infty}^{\infty} \exp\left(-\frac{v_2^2}{2}\right) \mathrm{d}\, y.$$

Wir substituieren noch

$$t := v_2 \text{ und } \mathrm{d}\, t = \frac{\sigma}{\sigma_1 \sigma_2} \mathrm{d}\, y$$

und erhalten

$$f_Z(z) = \frac{1}{2\pi \cdot \sigma} \cdot \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right) \cdot \int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}\, t.$$

Mit Lemma 35 folgt, dass $f_Z(z) = \varphi(z; \mu, \sigma)$ ist.

Daraus erhalten wir die Behauptung für $n = 2$, denn den Fall $Z := a_1 X_1 + a_2 X_2$ für beliebige Werte $a_1, a_2 \in \mathbb{R}$ können wir leicht mit Hilfe von Satz 36 auf den soeben bewiesenen Fall reduzieren. Durch Induktion kann die Aussage auf beliebige Werte $n \in \mathbb{N}$ verallgemeinert werden. *q. e. d.*

Momenterzeugende Funktionen für kontinuierliche Zufallsvariablen

Für diskrete Zufallsvariablen X haben wir die momenterzeugende Funktion

$$M_X(s) = \mathbb{E}[e^{Xs}]$$

eingeführt. Diese Definition kann man unmittelbar auf kontinuierliche Zufallsvariablen übertragen. Die für $M_X(s)$ gezeigten Eigenschaften bleiben dabei erhalten.

Beispiel: Für eine auf $[a,b]$ gleichverteilte Zufallsvariable U gilt

$$\begin{aligned} M_U(t) &= \mathbb{E}[e^{tX}] = \int_a^b e^{tx} \cdot \frac{1}{b-a}\,\mathsf{d}\,x \\ &= \left[\frac{e^{tx}}{t(b-a)}\right]_a^b \\ &= \frac{e^{tb}-e^{ta}}{t(b-a)}\,. \end{aligned}$$

Für eine standardnormalverteilte Zufallsvariable $N \sim \mathcal{N}(0,1)$ gilt

$$\begin{aligned} M_N(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{t\xi} e^{-\xi^2/2} \,\mathsf{d}\,\xi \\ &= e^{t^2/2} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(t-\xi)^2/2} \,\mathsf{d}\,\xi \\ &= e^{t^2/2}\,. \end{aligned}$$

Daraus ergibt sich für $Y \sim \mathcal{N}(\mu, \sigma^2)$ wegen $\frac{Y-\mu}{\sigma} \sim \mathcal{N}(0,1)$

$$\begin{aligned} M_Y(t) &= \mathbb{E}[e^{tY}] \\ &= e^{t\mu} \cdot \mathbb{E}[e^{(t\sigma)\cdot\frac{Y-\mu}{\sigma}}] \\ &= e^{t\mu} \cdot M_N(t\sigma) \\ &= e^{t\mu+(t\sigma)^2/2}\,. \end{aligned}$$