SS 2014

Zentralübung zur Vorlesung Diskrete Wahrscheinlichkeitstheorie

Dr. Werner Meixner

Fakultät für Informatik
TU München

http://www14.in.tum.de/lehre/2014SS/dwt/uebung/

5. Juni 2014

ZÜ VI

Übersicht:

1. Übungsbetrieb	Termine, Fragen, Probleme
2. Thema	Gamma- und Normalverteilung Exponentialverteilung Erlangverteilung
3. Vorbereitung	Blatt 8

1. Übungsbetrieb

Übungsgruppen in den beiden Nachpfingstwochen:
Montags-, Dienstags- und Mittwochsgruppen
in der zweiten Woche,
Donnerstags- und Freitagsgruppen
in der ersten Woche

Fristverlängerung Hausaufgabenabgabe Blatt 8:
Abgabe bis 18.6., 10 Uhr

Aktuelle Fragen, Anregungen?

2. Thema Gamma- und Normalverteilung

Ähnlich wie bei diskreten Verteilungen, gibt es bei kontinuierlichen Verteilungen

Familien zusammengehöriger Verteilungen,

wie z.B.

die Gammaverteilungen und die Normalverteilungen.

Zu den Gammaverteilungen zählen insbesondere die Erlangverteilungen, mit der Exponentialverteilung als Spezialfall.

2.1 Gammaverteilung

Definition

Eine kontinuierliche Zufallsvariable X ist Gammaverteilt von der Ordnung r mit Parameter λ mit reellen Werten $r > 0$ und $\lambda > 0$, i.Z. $X \sim \text{Gamma}(r, \lambda)$, falls für die Dichte f_X von X gilt

$$f_X(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} \cdot I_{(0,\infty)}(x)\,,$$

mit der Gammafunktion

$$\Gamma(r) = \int_0^\infty t^{r-1} e^{-t}\, \mathrm{d}t\,.$$

Erinnerung

$\Gamma(r+1) = r \cdot \Gamma(r)$,
$\Gamma(n) = (n-1)!$ für $n \in \mathbb{N}$ und
$\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Eigenschaften

1. $\mathbb{E}[X] = \frac{r}{\lambda}$,
2. $\mathrm{Var}[X] = \frac{r}{\lambda^2}$,
3. Faltungseigenschaft:

$$X \sim \text{Gamma}(r,\lambda),\;\; Y \sim \text{Gamma}(s,\lambda) \;\text{ und } X, Y \text{ unabhängig}$$
$$\implies X + Y \sim \text{Gamma}(r+s,\lambda).$$

Allgemeine Aufgabenstellung:

Bestimmen Sie für $n \in \mathbb{N} \setminus \{0\}$ mit einer geeigneten Induktion die Verteilungsfunktion einer Gamma(n, λ)-verteilten ZV.

Bemerkung: Gesucht ist ein integralfreier Ausdruck für die Verteilungsfunktion.

In VA 1 lösen wir diese Aufgabe für $n = 3$.

Die Verteilung der Summe Y von n unabhängigen, mit Parameter λ exponentialverteilten ZV $X_1, X_2, \ldots, X_n$ heißt

Erlang-Verteilung der Ordnung n.

Sie besitzt für $x > 0$ die angegebene Dichte

$$f_Y(x) = \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} \,.$$

Man berechnet die Dichte durch wiederholte Faltung.

Die Exponentialverteilung ist Spezialfall der Erlangverteilung mit $n = 1$.

Die Verteilungsfunktion

$$F_Y(x) = \int_{-\infty}^{x} f_Y(t)\, dt$$

berechnet man durch partielle Integration.

Dabei tritt eine Rekursionsformel

$$F_{S_n}(x) = A + F_{S_{n-1}}(x)$$

auf, wobei S_n jeweils für Y steht.

Den gesuchten Ausdruck erhält man dann durch Aufsummieren.

2.2 Normalverteilung

Definition

Eine kontinuierliche Zufallsvariable X heißt normalverteilt mit den Parametern $\mu \in \mathbb{R}$ und $\sigma \in \mathbb{R}^+$, i.Z. $X \sim \mathcal{N}(\mu, \sigma^2)$, falls für die Dichte f_X von X gilt

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)} .$$

$\mathcal{N}(0,1)$ heißt Standardnormalverteilung.

Verteilungsfunktion

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \, \mathrm{d}t$$

Bezeichnung:

$$\Phi(x; \mu, \sigma) := F(x) \, .$$

$$\Phi(x) := \Phi(x; 0, 1) \, .$$

Eigenschaften

1. $\mathbb{E}[X] = \mu$,
2. $\mathrm{Var}[X] = \sigma^2$,
3. Faltungseigenschaft:

   $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2), \;\; X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$
   und X_1, X_2 unabhängig mit $a_1, a_2 \in \mathbb{R}, a_1^2 + a_2^2 \neq 0$

   $$\implies a_1 X_1 + a_2 X_2 \sim \mathcal{N}(a_1\mu_1 + a_2\mu_2, a_1^2\sigma_1^2 + a_2^2\sigma_2^2).$$

4. Lineare Transformation:

   $X \sim \mathcal{N}(\mu, \sigma^2)$ und $a, b \in \mathbb{R}$ mit $a \neq 0$

   $$\implies aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)\,.$$

3. Vorbereitung Blatt 8

3.1 VA 1

Seien X_1 und X_2 unabhängige exponentialverteilte Zufallsvariablen mit Parametern λ_1 bzw. λ_2.

1. Berechnen Sie die Dichtefunktion von $Y = X_1 + X_2$ durch Anwendung der Faltungsformel

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(x) \cdot f_{X_2}(y - x)\,\mathrm{d}x$$

   und vereinfachen Sie das Ergebnis im Fall $\lambda_1 = \lambda_2$ so weit wie möglich.

2. Seien X_1, X_2, X_3 unabhängig exponentialverteilt mit gleichem Parameter λ und $Y = X_1 + X_2 + X_3$. Berechnen Sie die Verteilungsfunktion F_Y in geschlossener Form.

(1)

Berechnen Sie die Dichtefunktion von $Y = X_1 + X_2$ durch Anwendung der Faltungsformel

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(x) \cdot f_{X_2}(y - x)\,\mathrm{d}x$$

und vereinfachen Sie das Ergebnis im Fall $\lambda_1 = \lambda_2$ so weit wie möglich.

Lösung

Es gilt $f_{X_1}(x) = \lambda_1 e^{-\lambda_1 x}$ und $f_{X_2}(x) = \lambda_2 e^{-\lambda_2 x}$. Damit folgt, wenn $\lambda_1 \neq \lambda_2$ gilt, für $y \geq 0$:

$$
\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} \lambda_1 e^{-\lambda_1 x} \cdot \lambda_2 e^{-\lambda_2 (y-x)} \, \mathsf{d}x \\
&= \lambda_1 \lambda_2 e^{-\lambda_2 y} \int_0^y e^{(\lambda_2 - \lambda_1) x} \, \mathsf{d}x \\
&= \lambda_1 \lambda_2 e^{-\lambda_2 y} \left[\frac{e^{(\lambda_2 - \lambda_1) x}}{\lambda_2 - \lambda_1} \right]_{x=0}^{x=y} \\
&= \lambda_1 \lambda_2 e^{-\lambda_2 y} \cdot \frac{e^{(\lambda_2 - \lambda_1) y} - 1}{\lambda_2 - \lambda_1} \\
&= \lambda_1 \lambda_2 \cdot \frac{e^{-\lambda_1 y} - e^{-\lambda_2 y}}{\lambda_2 - \lambda_1} \, .
\end{aligned}
$$

Im Fall $\lambda_1 = \lambda_2 =: \lambda$ gilt

$$
\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} \lambda e^{-\lambda x} \cdot \lambda e^{-\lambda(y-x)} \, \mathsf{d}x \\
&= \lambda^2 e^{-\lambda y} \int_0^y \mathsf{d}x \\
&= \lambda^2 y e^{-\lambda y} \, .
\end{aligned}
$$

Für $y \leq 0$ folgt in allen Fällen direkt $f_Y(y) = 0$.

Seien X_1, X_2, X_3 unabhängig exponentialverteilt mit gleichem Parameter λ und $Y = X_1 + X_2 + X_3$. Berechnen Sie die Verteilungsfunktion F_Y in geschlossener Form.

Lösung

Wir wenden die Faltungsformel noch einmal an, und zwar auf $f_{X_1+X_2}$ und f_{X_3} wie folgt.

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X_1+X_2}(x) \cdot f_{X_3}(y-x)\,\mathsf{d}x \\ &= \int_0^y \lambda^2 x e^{-\lambda x} \cdot \lambda e^{-\lambda(y-x)}\,\mathsf{d}x \\ &= \lambda^3 e^{-\lambda y} \cdot \int_0^y x\,\mathsf{d}x \\ &= \frac{\lambda^3 y^2}{2} e^{-\lambda y}\,. \end{aligned}$$

Die Verteilungsfunktion F_Y kann nun durch Integration der Dichtefunktion f_Y berechnet werden, wie es im Folgenden ausführlich dokumentiert wird.

Wir wenden insbesondere partielle Integration an.

$$
\begin{aligned}
F_Y(y) &= \int_{-\infty}^{y} f_Y(x)\,\mathrm{d}x \\
&= \int_0^y \frac{\lambda^3 x^2}{2} e^{-\lambda x}\,\mathrm{d}x \\
&= (-\lambda^2)\int_0^y \frac{x^2}{2}\cdot(-\lambda)e^{-\lambda x}\,\mathrm{d}x \\
&= (-\lambda^2)\left[\frac{x^2}{2}\cdot e^{-\lambda x}\right]_{x=0}^{x=y} - (-\lambda^2)\int_0^y x\cdot e^{-\lambda x}\,\mathrm{d}x \\
&= (-\lambda^2)\frac{y^2}{2}e^{-\lambda y} - \lambda\int_0^y x\cdot(-\lambda)e^{-\lambda x}\,\mathrm{d}x \\
&= (-\lambda^2)\frac{y^2}{2}e^{-\lambda y} - \lambda\left(\left[x\cdot e^{-\lambda x}\right]_{x=0}^{x=y} - \int_0^y e^{-\lambda x}\,\mathrm{d}x\right) \\
&= \text{(Fortsetzung nächste Folie)}
\end{aligned}
$$

$$
\begin{aligned}
&= (-\lambda^2)\frac{y^2}{2}e^{-\lambda y} \;-\; \lambda\left(\left[x\cdot e^{-\lambda x}\right]_{x=0}^{x=y} - \int_0^y e^{-\lambda x}\,\mathsf{d}x\right)\\
&= (-\lambda^2)\frac{y^2}{2}e^{-\lambda y} \;-\; \lambda y e^{-\lambda y} - \int_0^y (-\lambda)e^{-\lambda x}\,\mathsf{d}x\\
&= (-\lambda^2)\frac{y^2}{2}e^{-\lambda y} \;-\; \lambda y e^{-\lambda y} - \left[e^{-\lambda x}\right]_{x=0}^{x=y}\\
&= 1 \;-\; \frac{\lambda^2 y^2}{2}e^{-\lambda y} \;-\; \lambda y e^{-\lambda y} \;-\; e^{-\lambda y}\,.
\end{aligned}
$$

Bemerkung

Seien $X_1, X_2, \ldots, X_n$ unabhängige mit Parameter λ exponentialverteilte Zufallsvariable. Die Zufallsvariable $Y = X_1 + \ldots + X_n$ besitzt die sogenannte Erlang-Verteilung

$$F_Y(y) = 1 - \sum_{i=0}^{n-1} \frac{(\lambda y)^i}{i!} \cdot e^{-\lambda y}\,.$$

3.2 VA 2

Sei X eine kontinuierliche Zufallsvariable.

1. Zeigen Sie: Falls $X \sim \mathcal{N}(2, \frac{1}{2})$, dann gilt $2X + 1 \sim \mathcal{N}(5, 2)$.
2. Seien $d_1, d_2, c \in \mathbb{R}$ mit $d_1 < d_2$ und $c > 0$. Berechnen Sie $a, b \in \mathbb{R}$, so dass für $Y = aX + b$ gilt

$$\Pr[d_1 \leq X \leq d_2] = \Pr[-c \leq Y \leq c]\,.$$

(1)

Zeigen Sie: Falls $X \sim \mathcal{N}(2, \frac{1}{2})$, dann gilt $2X + 1 \sim \mathcal{N}(5, 2)$.

Lösung

Seien μ und σ der Erwartungswert bzw. die Varianz von X, d. h. $\mu = 2$ bzw. $\sigma^2 = \frac{1}{2}$.

$Y = 2X + 1$ ist eine lineare Transformation von X und nach Satz der Vorlesung deshalb normalverteilt mit Erwartungswert $2\mu + 1 = 5$ bzw. Varianz $2^2\sigma^2 = 2$. W. z. b. w.

(2)

Seien $d_1, d_2, c \in \mathbb{R}$ mit $d_1 < d_2$ und $c > 0$.
Berechnen Sie $a, b \in \mathbb{R}$, so dass für $Y = aX + b$ gilt

$$\Pr[d_1 \leq X \leq d_2] = \Pr[-c \leq Y \leq c]\,.$$

Lösung

Sei $a > 0$. Dann gilt

$$d_1 \leq X \leq d_2 \quad \Longleftrightarrow \quad ad_1 + b \leq Y \leq ad_2 + b\,.$$

Wir lösen für a, b die Gleichungen

$$ad_1 + b = -c \quad \text{und} \quad ad_2 + b = c\,.$$

Lösung:

$$a = \frac{2c}{d_2 - d_1}\,, \qquad b = \frac{d_1 + d_2}{d_1 - d_2} \cdot c\,.$$

3.3 VA 3

Wir betrachten unabhängige stetige Zufallsvariablen X und Y, die beide auf dem Intervall $[0, 1] \subseteq \mathbb{R}$ gleichverteilt sind.
Sei $Z = \max\{X, Y\}$.

1. Berechnen Sie die Verteilungsfunktion F_Z.
2. Bestimmen Sie eine Funktion $u : [0, 1] \to [0, 1]$, so dass $u(X)$ die gleiche Verteilung wie Z besitzt.

(1)

Berechnen Sie die Verteilungsfunktion F_Z.

Lösung

Die gemeinsame Dichte der Zufallsvariablen X, Y sei $f_{X,Y}(x, y)$.

Aufgrund der Unabhängigkeit von X, Y gilt für $(x, y) \notin [0, 1] \times [0, 1]$ die gemeinsame Dichte 0 und für $(x, y) \in [0, 1] \times [0, 1]$

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y) = 1 \cdot 1 = 1\,.$$

Wir berechnen die Verteilungsfunktion $F_Z(z)$.

Offenbar gilt zunächst
$F_Z(z) = 0$ bzw. $F_Z(z) = 1$ für $z \leq 0$ bzw. $1 \leq z$.

Für $z \in [0,1] \subseteq \mathbb{R}$ gilt

$$\begin{aligned}
F_Z(z) &= \Pr[\max\{X,Y\} \leq z] \\
&= \Pr[X \leq z, Y \leq z] \\
&= \int_{[0,z]\times[0,z]} f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y \\
&= \int_{[0,z]\times[0,z]} 1\,\mathrm{d}x\mathrm{d}y \\
&= z^2\,.
\end{aligned}$$

(2)

Bestimmen Sie eine Funktion $u : [0, 1] \to [0, 1]$, so dass $u(X)$ die gleiche Verteilung wie Z besitzt.

Lösung

Aus der Vorlesung ist bekannt, dass wir eine Simulation von F_Z aus der Inversen von F_Z erhalten können.
Wir rechnen direkt und setzen die Invertierbarkeit von u voraus.
Sei $Y = u(X)$.

Dann gilt

$$\begin{aligned} F_Y(y) &= \Pr[Y \le y] = \Pr[u(X) \le y] \\ &= \Pr[X \le u^{-1}(y)] \\ &= F_X(u^{-1}(y)) \\ &= u^{-1}(y)\,. \end{aligned}$$

Aus der Gleichung $F_Z(y) = F_X(u^{-1}(y))$ folgt nun $y^2 = u^{-1}(y)$, mithin

$$u(x) = \sqrt{x}\,.$$