Diskrete und stetige Verteilungen

	diskrete Zufallsvariablen	stetige Zufallsvariablen
Dichtefunktion / PDF/PMF	$f(x) = P(X = x)$	$f(x) = F'(x) \neq P(X = x)$
Kumulative Verteilungsfunktion / CDF	$F(x) = P(X \leq x) = \sum_{u \leq x} f(u)$	$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du$
Wahrscheinlichkeiten	$P(a \leq X \leq b) = \sum_{a \leq x \leq b} f(x)$ $P(a < X \leq b) = \sum_{a < x \leq b} f(x)$ $P(a < X < b) = \sum_{a < x < b} f(x)$ $P(a < X) = 1 - F(a)$	$\left.\begin{matrix} P(a \leq X \leq b) \\ P(a < X \leq b) \\ P(a < X < b) \end{matrix}\right\} = \int_{a}^{b} f(x)dx$ $P(a < X) = 1 - F(a)$
Graphische Darstellung von f	Stabdiagramm	Graph
Erwartungwert	$E(X) = \sum_{x \in \mathbb{R}} f(x) \cdot x$	$E(X) = \int_{-\infty}^{\infty} f(x) \cdot x\,dx$
Varianz	$V(X) = \sum_{x \in \mathbb{R}} f(x) \cdot (x - E(X))^2$	$V(X) = \int_{-\infty}^{\infty} f(x) \cdot (x - E(X))^2\,dx$

Satz
Für diskrete und stetige Zufallsvariablen X und Y gelten die folgenden Regeln:

(1) Linearität des Erwartungswertes:

$$E(X+Y)=E(X)+E(Y) \text{ und } E(\alpha X)=\alpha E(X) \text{ mit } \alpha \in \mathbb{R}.$$

(2) Verschiebungssatz für die Varianz:

$$V(X)=E(X^2)-(E(X))^2$$

(3) $V(\alpha X+\beta)=\alpha^2 \cdot V(X)$ mit $\alpha, \beta \in \mathbb{R}$.

(4) Sind X und Y stochastisch unabhängig, so gilt:

$$V(X+Y)=V(X)+V(Y)$$

Bemerkung

Es ist effizienter, die Varianz mithilfe des Verschiebungssatzes zu berechnen als mithilfe der Definition.

Die Hypergeometrische Verteilung

Definition

Eine diskrete Zufallsvariable X heisst *hypergeometrisch verteilt* mit den Parametern n (Anzahl Ziehungen ohne Zurücklegen), N (Gesamtzahl aller Objekte) und M (Gesamtzahl aller Merkmalsträger), wenn ihre Dichtefunktion (PMF) gegeben ist durch

$$P(X=x)=\frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}}$$

Schreibweise: $X \sim H(N, M, n)$.

X zählt, wie oft bei der n-fachen Ziehung (nacheinander und ohne Zurücklegen) ein Merkmalsträger gezogen wird.

Satz
Für eine Zufallsvariable $X \sim H(N, M, n)$ gilt:

(1) $\mu=E(X)=n \cdot \frac{M}{N}$
(2) $\sigma^2=\mathrm{V}(X)=n \cdot \frac{M}{N} \cdot (1-\frac{M}{N}) \cdot \frac{N-n}{N-1}$
(3) $\sigma=S(X)=\sqrt{V(X)}$

Die Bernoulli Verteilung

Definition

Eine Zufallsvariable X heisst *Bernoulli-verteilt*, wenn sie nur zwei verschiedene Werte annehmen kann: den Wert 1 mit der Wahrscheinlichkeit $P(X=1)=p$ und den Wert 0 mit der Wahrscheinlichkeit $P(X=0)=1-p$.

Satz
Für Bernoulli-verteilte Zufallsvariablen X gilt:

(1) $E(X)=E(X^2)=p$.
(2) $V(X)=p\cdot(1-p)$

Die Binomialverteilung

Definition

Eine diskrete Zufallsvariable X heisst *binomialverteilt* mit den Parametern n (Anzahl Wiederholungen) und p (Wahrscheinlichkeit für ein Ergebnis 1), wenn ihre Dichtefunktion (PMF) gegeben ist durch

$$P(X=x)=\binom{n}{x}\cdot p^x\cdot(1-p)^{n-x}$$

Schreibweise: $X\sim B(n;p)$.
X zählt, wie oft bei der n-fachen Wiederholung eines Bernoulli-Experiments das Ergebnis 1 eintritt. Die Wahrscheinlichkeit für das Ergebnis 0 wird üblicherweise mit $q=1-p$ bezeichnet.

Die $B(n;p)$-verteilte Zufallsvariable X kann als Summe von n Bernoulli-verteilten Zufallsvariablen X_i aufgefasst werden: $X=\sum_{i=1}^n X_i$. Dabei hält X_i das Ergebnis des i-ten Experiments fest, und es gilt: $P(X_i=1)=p$.

Satz
Für eine Zufallsvariable $X\sim B(n;p)$ gilt:

(1) $\mu=E(X)=np$
(2) $\sigma^2=\mathrm{V}(X)=npq$
(3) $\sigma=S(X)=\sqrt{npq}$

Faustregel zur Approximation
Wenn die Bedingung $n\leq\frac{N}{20}$ erfüllt ist, kann die hypergeometrische Verteilung $H(N,M,n)$ gut durch die Binomialverteilung $B(n,\frac{M}{N})$ angenähert werden:
$H(N,M,n)\approx B(n,\frac{M}{N})$

Die Poisson Verteilung

Definition

Eine diskrete Zufallsvariable X heisst *poissonverteilt* mit dem Parameter $\lambda>0$ (durchschnittliche Anzahl Ereignisse pro betrachtetes Zeitintervall), wenn ihre Dichtefunktion (PMF) gegeben ist durch

$$P(X = x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}$$

Schreibweise: $X \sim Poi(\lambda)$.
X zählt die Anzahl der (stochastisch unabhängigen, gleichartigen) Ereignisse in einem betrachteten Zeitintervall.

Satz
Für eine Zufallsvariable $X \sim Poi(\lambda)$ gilt:

(1) $\mu = E(X) = \lambda$
(2) $\sigma^2 = \mathrm{V}(X) = \lambda$
(3) $\sigma = S(X) = \sqrt{\lambda}$

Faustregel zur Approximation
Wenn die Bedingung $n \geq 50$ und $p \leq 0.1$ erfüllt ist, kann die Binomialverteilung $B(n,p)$ gut durch die Poissonverteilung $Poi(n \cdot p)$ angenähert werden:
$B(n,p) \approx Poi(n \cdot p)$.

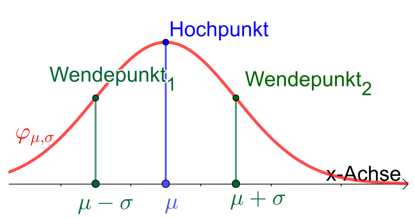
Gauss'sche Normalverteilung

Definition

Eine stetige Zufallsvariable X heisst *normalverteilt* mit den Parametern $\mu, \sigma \in \mathbb{R}$, $\sigma > 0$, wenn sie folgende Dichtefunktion (PDF) hat:

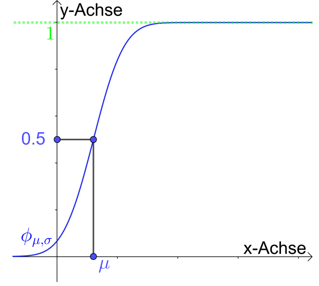
$$\varphi_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Schreibweise: $X \sim N(\mu; \sigma)$



Die kumulative Verteilungsfunktion (CDF) von $\varphi_{\mu,\sigma}(x)$ wird mit $\phi_{\mu,\sigma}(x)$ bezeichnet. Sie ist definiert durch:

$$\phi_{\mu,\sigma}(x) = P(X \leq x) = \int_{-\infty}^{x} \varphi_{\mu,\sigma}(t)\, dt = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$



Ist $\mu = 0$ und $\sigma = 1$, so spricht man von der *Standardnormalverteilung*. Ihre Dichtefunktion (PDF) wird mit φ bezeichnet; sie ist gegeben durch:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x^2}.$$

Ihre Verteilungsfunktion (CDF) $\phi_{0,1}(x)$ wird mit $\phi(x)$ bezeichnet. Schreibweise: $X \sim N(0;1)$.

Die Verteilungsfunktion der Normalverteilung kann nicht auf elementare Weise berechnet werden. Für ihre Werte gibt es Tabellen (Papula 12. Aufl. S. 514); die Tabellen beziehen sich allerdings immer auf die *Standard*normalverteilung.

Bemerkung

Die Dichtefunktion (PDF) $\varphi_{\mu,\sigma}(x)$ hat folgende Eigenschaften:

(a) Sie ist symmetrisch bezüglich der Geraden $x = \mu$.

(b) Sie hat Wendepunkte an den Stellen $\mu - \sigma$ und $\mu + \sigma$.

(c) Sie ist *normiert*, d.h. es gilt:

$$\int\limits_{-\infty}^{\infty} \varphi_{\mu,\sigma}(x)\,dx = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot \int\limits_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$$

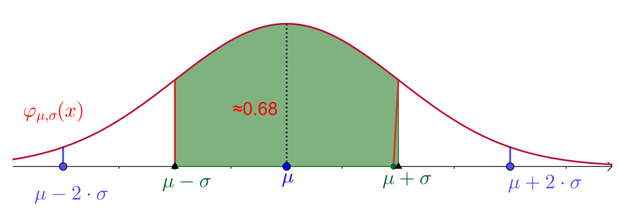
(d) Eine Änderung von μ bewirkt eine Verschiebung in *x*-Richtung; je grösser σ ist, desto breiter und niedriger wird die Glockenkurve.

(e) Für eine Zufallsvariable $X \sim N(\mu;\sigma)$ gilt: $E(X) = \mu$ und $V(X) = \sigma^2$.

Bemerkung

Bei einer Zufallsvariable *X*, die der Normalverteilung $N(\mu;\sigma)$ folgt, liegen

- ca. 68 % der beobachteten Werte zwischen $\mu - \sigma$ und $\mu + \sigma$,
- ca. 95 % der beobachteten Werte zwischen $\mu - 2\sigma$ und $\mu + 2\sigma$,
- ca. 99.7 % der beobachteten Werte zwischen $\mu - 3\sigma$ und $\mu + 3\sigma$.



Wichtige Eigenschaften einer $N(\mu;\sigma)$ -verteilten Zufallsvariable X

$\phi'_{\mu,\sigma}(x) = \varphi_{\mu,\sigma}(x)$

$\phi_{\mu,\sigma}(x) = \phi(\frac{x-\mu}{\sigma})$

$P(a \leq X \leq b) = \phi_{\mu,\sigma}(b) - \phi_{\mu,\sigma}(a)$

$$\begin{aligned} P(|X-\mu| \leq \varepsilon) &= P(\mu - \varepsilon \leq X \leq \mu + \varepsilon) \\ &= 2 \cdot \phi_{\mu,\sigma}(\mu + \varepsilon) - 1 \\ &= 1 - 2 \cdot \phi_{\mu,\sigma}(\mu - \varepsilon) \end{aligned}$$

In den Aussagen können $\leq$ Zeichen nach Belieben durch $<$ Zeichen ersetzt werden.

Zentraler Grenzwertsatz

Gegeben sind lauter identisch verteilte und stochastisch unabhängige Zufallsvariablen $X_1, X_2, \ldots$, alle mit demselben Erwartungswert μ und derselben Varianz σ^2 . Dann hat die Summe

$$S_n = \sum_{i=1}^{n} X_i$$

den Erwartungswert $n\mu$ und die Varianz $n\sigma^2$ und ist annähernd $N(n\mu; \sqrt{n}\sigma)$-verteilt.

Das arithmetische Mittel

$$\bar{X}_n = S_n / n$$

hat den Erwartungswert μ und die Varianz σ^2/n und ist annähernd $N(\mu; \frac{\sigma}{\sqrt{n}})$-verteilt.

Die Verteilungsfunktion (CDF) $F_n(u)$ der dazugehörigen standardisierten Zufallsvariablen

$$U_n = \frac{S_n - n\mu}{\sqrt{n} \cdot \sigma} = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$

konvergiert für $n \to \infty$ gegen die Verteilungsfunktion $\phi(u)$ der Standardnormalverteilung:

$$\lim_{n \to \infty} F_n(u) = \phi(u) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{u} e^{-\frac{1}{2}t^2} dt$$

Zuletzt geändert: Montag, 20. September 2021, 17:21