

Zufallsvariablen und Standardverteilungen

Dr. Thomas Zehrt

Universität Basel
Wirtschaftswissenschaftliches Zentrum
Abteilung Quantitative Methoden

Statistik II FS 2010

Outline

- 1 Erwartungswert, Varianz, Quantil und Modalwert
- 2 Unabhängige Zufallsvariablen
- 3 Standardisierung und arithmetisches Mittel
- 4 Die Ungleichung von Tschebyshev
- 5 Diskrete Standardverteilungen
 - Die (diskrete) Gleichverteilung
 - Die Zweipunktverteilung
 - Die Binomialverteilung
 - Die hypergeometrische Verteilung
 - Die Poisson-Verteilung
- 6 Stetige Standardverteilungen
 - Die stetige Gleichverteilung
 - Die Exponentialverteilung
 - Die Normalverteilung
 - Die Testverteilungen

Definition

X sei eine diskrete Zufallsvariable mit $\mathbb{R}_X = \{x_1, x_2, x_3, \dots, x_n\}$. Dann ist der Erwartungswert $E(X) = \mu$ von X definiert durch

$$E(X) = \sum_{i=1}^n x_i \underbrace{P(X = x_i)}_{p_i}$$

Die Varianz $Var(X) = \sigma^2$ von X (mit $\mu = E(X)$) ist:

$$Var(X) = \sum_{i=1}^n (x_i - \mu)^2 \underbrace{P(X = x_i)}_{p_i}$$

Definition

X sei eine stetige Zufallsvariable mit zugehöriger Dichte f . Dann ist der Erwartungswert $E(X) = \mu$ von X definiert durch

$$E(X) = \int_{-\infty}^{\infty} t f(t) dt$$

Die Varianz $Var(X) = \sigma^2$ von X (mit $\mu = E(X)$) ist:

$$Var(X) = \int_{-\infty}^{\infty} (t - \mu)^2 f(t) dt$$

Definition

Es seien X und Y zwei (beliebige) Zufallsvariablen mit $\mu_X = E(X)$ und $\mu_Y = E(Y)$. Dann ist die Kovarianz von X und Y definiert als:

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

Es gilt:

$$\text{Cov}(X, Y) = \text{Cov}(Y, X)$$

$$\text{Cov}(X, Y) = E(XY) - \mu_X\mu_Y$$

Satz (Rechenregeln für Erwartungswert und Varianz I)

Seien X, Y Zufallsvariablen und a, b, c reelle Zahlen. Dann gelten die folgenden Rechenregeln:

① $E(aX + bY + c) = a E(X) + b E(Y) + c$

② $Var(aX + b) = a^2 Var(X)$

③ *Verschiebungssatz der Varianz*

$$Var(X) = E(X^2) - [E(X)]^2$$

④ $Var(X \pm Y) = Var(X) + Var(Y) \pm 2Cov(X, Y)$

Aufgabe

- 1 Sei X die Augenzahl beim einmaligen Wurf eines fairen Würfels. Bestimmen Sie den Erwartungswert und die Varianz.
- 2 Sei X die Zufallsvariable aus dem Beispiel in der Einführung. Bestimmen Sie den Erwartungswert und die Varianz.

Lösung:

Definition

Ein p -Quantil x_p ($0 < p < 1$) ist der Wert der Verteilungsfunktion, für den $F(x_p) = p$ gilt.

Im Fall von diskreten Verteilungen wird als p -Quantil x_p der Wert genommen, für den gilt

$$F(x_p) \geq p \quad \text{und} \quad F(x) < p \quad \text{für} \quad x < x_p.$$

Man kann sich also x_p als den Wert einer Verteilung vorstellen, der die Wahrscheinlichkeitsmasse so teilt, dass links von x_p genau die Masse p und rechts von x_p genau die Masse $1 - p$ liegt.

Spezialfälle:

unteres Quartil: $x_{0.25}$, Median, Zentralwert: $x_{0.5}$, oberes Quartil: $x_{0.75}$

In der induktiven Statistik spielen p - und $(1 - p)$ -Quantile (zu $p = 0.025$, $p = 0.05$, ...) insbesondere in der Testtheorie eine wichtige Rolle.

Dort stellen diese Quantile Grenzen der Irrtumswahrscheinlichkeit dar, die die Grundlage für eine Testentscheidung bilden.

Definition

Der Modalwert D einer Verteilung ist derjenige Wert, bei dem die Dichte bzw. die Wahrscheinlichkeitsfunktion ihr Maximum hat. Liegt nur ein Maximum vor, so heisst die Verteilung unimodal.

Um einige wichtige Sätze formulieren zu können müssen wir den Begriff der Unabhängigkeit auf Zufallsvariablen übertragen. Diese Sätze stimmen i.A. **nicht** für beliebige Zufallsvariablen!

Definition

Zwei Zufallsvariablen X und Y heissen unabhängig, falls für **alle** $x \in \mathbb{R}_X$ und **alle** $y \in \mathbb{R}_Y$ die Ereignisse $(X \leq x)$ und $(Y \leq y)$ unabhängig sind, also falls

$$P((X \leq x) \text{ und } (Y \leq y)) = P(X \leq x) \cdot P(Y \leq y).$$

Bemerkung: Falls X und Y diskret sind, genügt es

$$P((X = x_i) \text{ und } (Y = y_j)) = P(X = x_i) \cdot P(Y = y_j).$$

für alle $x_i \in \mathbb{R}_X$ und für alle $y_i \in \mathbb{R}_Y$ zu zeigen.

Aufgabe

- ① *Ein Laplace-Würfel wird zweimal geworfen, X sei die Augenzahl beim ersten Wurf und Y die beim zweiten Wurf. Sind X und Y abhängig oder unabhängig?*
- ② *In einer Urne seien 6 Kugeln, durchnummeriert von 1 bis 6. Es werden zwei Kugeln **ohne** Zurücklegen gezogen. X sei die Nummer der ersten und Y die der zweiten Kugel. Sind X und Y abhängig oder unabhängig?*

Lösung:

Satz (Rechenregeln für Erwartungswert und Varianz II)

Seien X und Y **unabhängige** Zufallsvariablen. Dann gilt:

$$E(X \cdot Y) = E(X) \cdot E(Y)$$

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$$

$$\text{Cov}(X, Y) = 0$$

Definition

Eine Zufallsvariable heisst standardisiert, wenn sie den Erwartungswert 0 und die Varianz 1 hat. Sei X eine Zufallsvariable. Dann heisst die Zufallsvariable

$$Z = \frac{X - E(X)}{\sqrt{\text{Var}(X)}}$$

Standardisierung von X .

Aufgabe

Zeigen Sie, dass die Standardisierung einer Zufallsvariablen eine standardisierte Zufallsvariable ist.

Definition

Wir bezeichnen Zufallsvariablen X_1, X_2, \dots, X_n als unabhängig und identisch verteilt, falls alle X_i

- ein und dieselbe Verteilung besitzen und
- voneinander unabhängig sind.

Für die unabhängigen und identisch verteilten Zufallsvariablen X_1, X_2, \dots, X_n gelte nun $E(X_i) = \mu$ und $Var(X_i) = \sigma^2$ für $i = 1, 2, \dots, n$.

Dann heisst die neue Zufallsvariable

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

das arithmetische Mittel.

Es gilt

$$E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \cdot n \cdot \mu = \mu$$

und

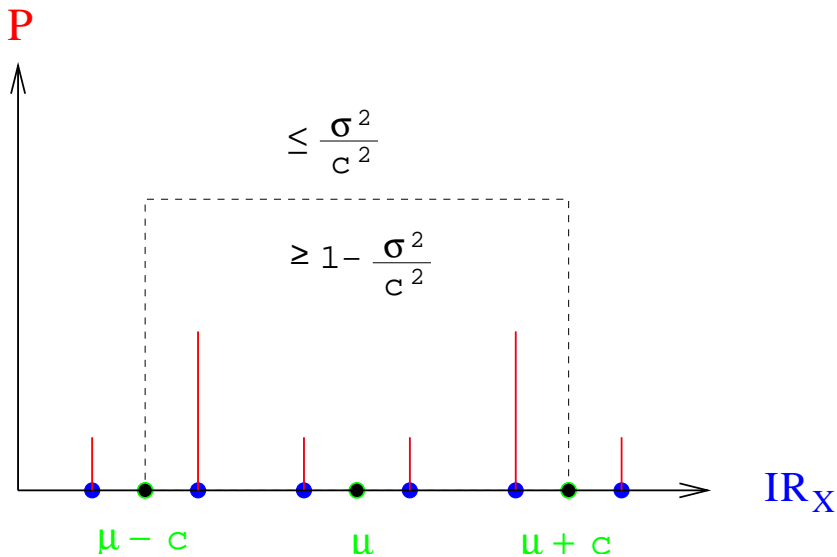
$$\begin{aligned} \text{Var}(\bar{X}) &= \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \text{Var}\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) \\ &= \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n} \end{aligned}$$

Satz (Die Ungleichung von Tschebyshev)

Sei X eine beliebige Zufallsvariable mit $\mu = E(X)$ und $\sigma^2 = \text{Var}(X)$.
Dann gelten für jede positive Zahl c die folgenden beiden
(äquivalenten) Ungleichungen:

$$P(|X - \mu| < c) \geq 1 - \frac{\sigma^2}{c^2} \quad \text{oder} \quad P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}$$

d.h. man kann relativ leicht die Wahrscheinlichkeit abschätzen, mit der X einen Wert innerherhalb (oder ausserhalb) des Intervalls $[\mu - c, \mu + c]$ annimmt.



Aufgabe

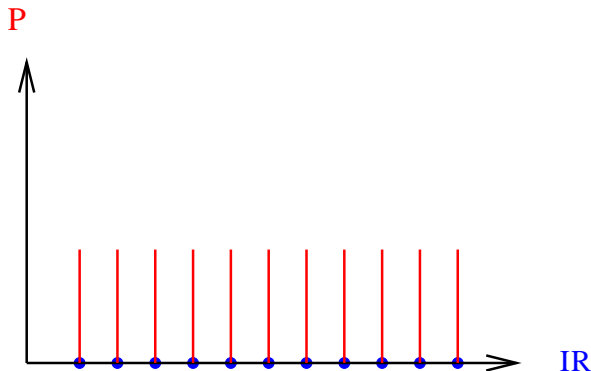
Von einer Zufallsvariablen X sei nur bekannt, dass sie den Erwartungswert 15 und die Varianz 4 besitzt. Wie gross ist $P(10 < X < 20)$ mindestens?

Lösung:

Eine diskrete Zufallsvariable X mit den Ausprägungen x_1, x_2, \dots, x_k heisst gleichverteilt, wenn für ihre Wahrscheinlichkeitsfunktion

$$P(X = x_i) = \frac{1}{k}$$

für alle $i = 1, 2, \dots, k$ gilt.



	x_i	x_1	x_2	\dots	x_k
Verteilung	$P(X = x_i)$	$\frac{1}{k}$	$\frac{1}{k}$	\dots	$\frac{1}{k}$
Erwartungswert	$E(X)$	$= \frac{1}{k} \sum_{i=1}^k x_i$			
Varianz	$Var(X)$	$= \frac{1}{k} \sum_{i=1}^k x_i^2 - E(X)^2$			

Aufgabe

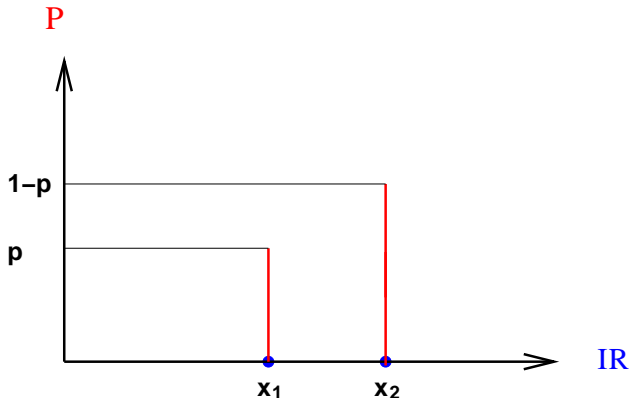
Beweisen Sie die Gleichung für $\text{Var}(X)$.

Lösung:

Eine diskrete Zufallsvariable X mit nur zwei Ausprägungen x_1 und x_2 heisst zweipunktverteilt.

Für ihre Wahrscheinlichkeitsfunktion gilt dann

$$P(X = x_i) = \begin{cases} p & \text{für } i = 1 \\ 1 - p & \text{für } i = 2 \end{cases}$$



Verteilung	x_i	x_1	x_2
	$P(X = x_i)$	p	$1 - p$
Erwartungswert	$E(X)$	=	
Varianz	$Var(X)$	=	

Eine diskrete Zufallsvariable X mit den Ausprägungen $0, 1, \dots, n$ heisst binomialverteilt, mit den Parametern n und p (und $q = 1 - p$), wenn für alle $x = 0, 1, \dots, n$ gilt:

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} = \binom{n}{x} p^x q^{n-x} = f_{Bi}(x; n, p)$$

	x_i	0	1	...	n	
Vert.	$P(X = x_i)$	$\binom{n}{0} p^0 q^n$	$\binom{n}{1} p^1 q^{n-1}$...	$\binom{n}{n} p^n q^0$	
Erwartw.	$E(X)$	$= n \cdot p$				
Varianz	$Var(X)$	$= n \cdot p \cdot (1 - p)$				

Wann tritt die Binomialverteilung auf?

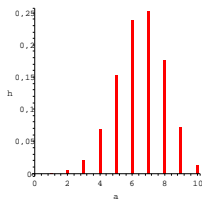
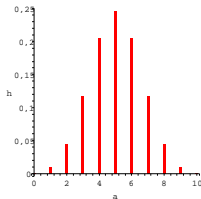
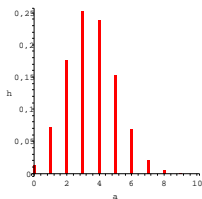
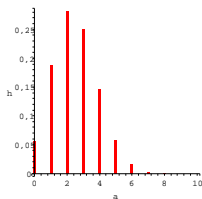
- Ein Experiment hat **zwei** mögliche Ergebnisse: Erfolg (E oder 1) und Misserfolg (E^c oder 0)
- Das Experiment wird eine feste Anzahl n mal durchgeführt.
- Die Durchführungen erfolgen unabhängig voneinander.
- Die Erfolgswahrscheinlichkeit p im Einzelexperiment ist konstant.

$X = \text{Anzahl Erfolge in den } n \text{ Versuchen}$

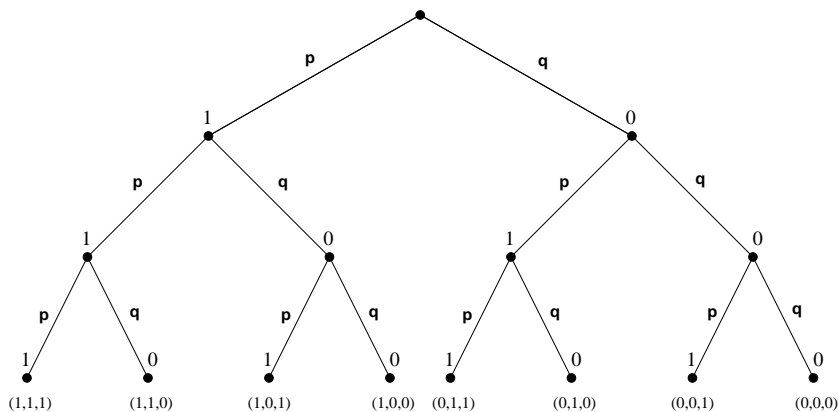
Dann ist X binomialverteilt.

Bezeichnung: $X \sim B(n; p)$

Verteilungen der Binomialverteilung für $n = 10$ und $p = 0.25, 0.35, 0.5, 0.65$.



Eine gute Möglichkeit, sich ein Experiment der vorne beschriebenen Art vorzustellen, ist der so genannte Binomialbaum.



Aufgabe

Bestimmen Sie die Wahrscheinlichkeit dafür, beim 10-maligen Werfen einer fairen Münze genau, mindestens oder höchstens 7-mal „Zahl“ zu erhalten.

Lösung:

Eine diskrete Zufallsvariable X mit den Ausprägungen $0, 1, 2, \dots, n$ heisst hypergeometrisch verteilt, mit den Parameter n, S, N (wobei $n \leq S$ und $n \leq N - S$), wenn für ihre Wahrscheinlichkeitsfunktion

$$P(X = x) = \frac{\binom{S}{x} \binom{N-S}{n-x}}{\binom{N}{n}} = f_{Hy}(x; N, S, n)$$

für alle $x = 0, 1, 2, \dots, n$ gilt.

x_i	0	1	...	n
Vert.				
$P(X = x_i)$	$\frac{\binom{S}{0} \binom{N-S}{n-0}}{\binom{N}{n}}$	$\frac{\binom{S}{1} \binom{N-S}{n-1}}{\binom{N}{n}}$...	$\frac{\binom{S}{n} \binom{N-S}{n-n}}{\binom{N}{n}}$
Erwartw.	$E(X) = n \cdot \frac{S}{N}$			
Varianz	$Var(X) = n \cdot \frac{S}{N} \left(1 - \frac{S}{N}\right) \left(\frac{N-n}{N-1}\right)$			

Wann tritt die hypergeometrische Verteilung auf?

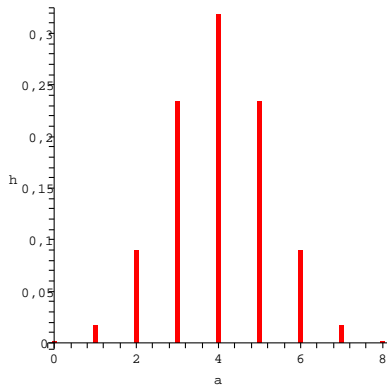
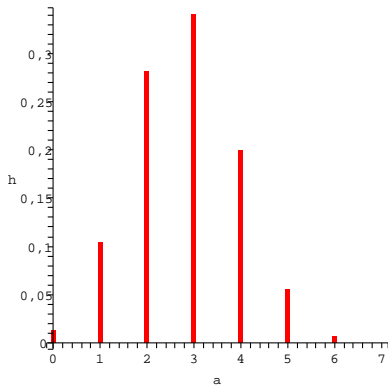
Von N Elementen seien S vom Typ E und $W = N - S$ vom Typ E^c . Eine Stichprobe vom Umfang n wird gezogen (ohne zurücklegen).

$X =$ Anzahl Elemente vom Typ E in dieser Stichprobe.

Dann ist X hypergeometrisch verteilt.

Verteilung der hypergeometrischen Verteilung für

- $S = 10, W = 15, n = 7$;
- $S = 15, W = 15, n = 8$



Aufgabe

In einer Urne befinden sich 10 schwarze und 15 weiße Kugeln. Es werden 7 Kugeln gezogen. Mit welcher Wahrscheinlichkeit befinden sich unter diesen 7 Kugeln genau 5 schwarze (und 2 weiße)?

Lösung:

Eine diskrete Zufallsvariable X mit den Ausprägungen $0, 1, 2, 3 \dots$ heisst poisson-verteilt, mit dem Parameter λ , wenn für ihre Wahrscheinlichkeitsfunktion

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

für alle $x = 0, 1, 2, 3 \dots$ gilt.

	x_i	0	1	2	3	...
Vert.	$P(X = x_i)$	$\frac{\lambda^0}{0!} e^{-\lambda}$	$\frac{\lambda^1}{1!} e^{-\lambda}$	$\frac{\lambda^2}{2!} e^{-\lambda}$	$\frac{\lambda^3}{3!} e^{-\lambda}$...
Erwartw.	$E(X)$	$= \lambda$				
Varianz	$Var(X)$	$= \lambda$				

Wann tritt die Poisson-Verteilung auf?

Ausgehend von zufälligen Ereignissen, die innerhalb eines Kontinuums mit der Intensitätsrate $\lambda > 0$ auftreten, sei

$X =$ Anzahl der Ereignisse innerhalb eines Kontinuums

Dann ist X poissonverteilt mit dem Parameter $\lambda > 0$.

Bezeichnung: $X \sim Po(\lambda)$

Aufgabe

An einer Tankstelle kommen durchschnittlich 4 (Intensitätsrate = λ) Fahrzeuge pro Minute an. X sei die Anzahl der in einer Minute ankommenden Fahrzeuge. Mit welcher Wahrscheinlichkeit kommen innerhalb einer Minute 0, 2 bzw. mindestens 3 Fahrzeuge an?

Lösung:

- Wahrscheinlichkeit, dass kein Fahrzeug ankommt:

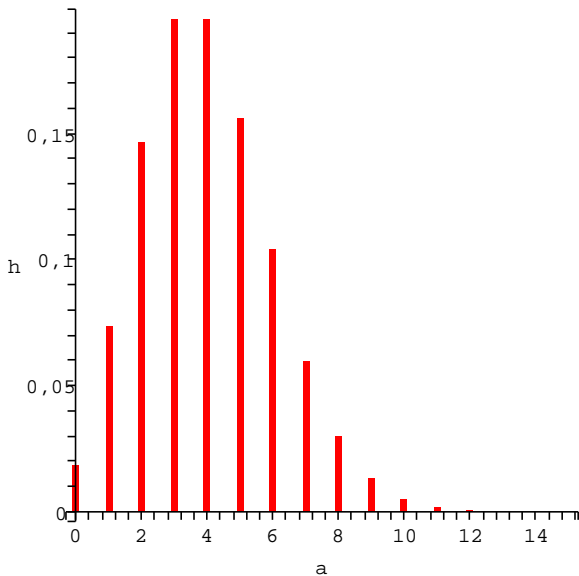
$$P(X = 0) = \frac{4^0}{0!} e^{-4} = 0.0183$$

- Wahrscheinlichkeit, dass genau zwei Fahrzeuge ankommen:

$$P(X = 2) = \frac{4^2}{2!} e^{-4} = 0.1465$$

- Wahrscheinlichkeit, dass mindestens drei Fahrzeuge ankommen:

$$P(X \geq 3) = 1 - P(X < 3) = 0.761$$



Eine stetige Zufallsvariable X mit der Dichte

$$f_{Re}(t) = \begin{cases} \frac{1}{b-a} & t \in [a, b] \\ 0 & \text{sonst} \end{cases}$$

heisst (stetig) gleichverteilt oder auch rechtecksverteilt auf dem Intervall $[a, b]$.

Aufgabe

Bestimmen Sie die Verteilungsfunktion F_{Re} der stetigen Gleichverteilung auf dem Intervall $[a, b]$ und skizzieren Sie f_{Re} und F_{Re} .

$$f_{Re}(t) = \begin{cases} \frac{1}{b-a} & t \in [a, b] \\ 0 & \text{sonst} \end{cases}$$

$$E(X) = \frac{a+b}{2}$$

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

Eine stetige Zufallsvariable X mit der Dichte

$$f_{Ex}(t; \lambda) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & \text{sonst} \end{cases}$$

heisst exponentialverteilt mit dem Parameter λ . Kurz: $X \sim Ex(\lambda)$

Aufgabe

Bestimmen Sie die Verteilungsfunktion F_{Ex} der Exponentialverteilung und skizzieren Sie f_{Ex} und F_{Ex} .

$$f_{Ex}(t; \lambda) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & \text{sonst} \end{cases}$$

$$E(X) = \frac{1}{\lambda}$$

$$\text{Var}(X) = \frac{1}{\lambda^2}$$

Aufgabe

Ermitteln Sie für eine mit $\lambda = 0.4$ exponentialverteilte Zufallsvariable X die folgenden Wahrscheinlichkeiten (bzw. bedingten Wahrscheinlichkeiten): $P(X < 3)$, $P(2 < X \leq 3)$, $P(X \geq 3|X > 2)$ und $P(X \leq 3|X > 2)$.

Lösungen:

0.6988, 0.1481, 0.6703, 0.3297

Aufgabe

Eine Glühlampe sei ununterbrochen in Betrieb, bis sie ausfällt. Die Zufallsgrösse X , die die (zufällige) Lebensdauer der Glühlampe angibt, sei exponentialverteilt.

Weiterhin sei bekannt, dass derartige Glühlampen eine mittlere Lebenszeit von 6000 Stunden haben.

- 1 Wie muss λ gewählt werden, damit $E(X)$ gleich der mittleren Lebenszeit ist?*
- 2 Wie gross ist die Wahrscheinlichkeit, dass die Glühlampe nach 6000 Stunden bereits ausgefallen ist?*
- 3 Eine Glühlampe sei bereits 3000 Stunden in Betrieb. Wie gross ist die Wahrscheinlichkeit, dass sie noch mindestens 3000 Stunden weiterbrennt?*

Weitere Beispiele für normalverteilte Zufallsvariablen:

- der Intelligenzquotient,
- die täglichen Renditen einer Firma,
- das Schlachtgewicht von Mastochsen ...

Gemeinsamkeit:
Es sind Grössen, die aus der Summe
vieler verschiedener (und kleiner)
zufälliger Einflüsse entstehen.

Einflüsse auf den Intelligenzquotienten:

Gene, Eltern, Geschwister, Schule, Lehrer, Ernährung, wie lange man als Säugling einen Schnuller hatte usw.

Eine stetige Zufallsvariable X heisst normalverteilt oder $N(\mu, \sigma^2)$ -verteilt, mit den Parametern

- μ ($-\infty < \mu < \infty$) und
- σ^2 ($0 < \sigma^2$),

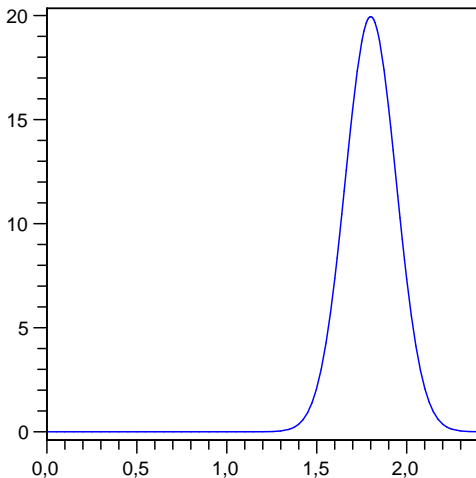
falls ihre Dichte durch die Funktion

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} =: \phi(t; \mu, \sigma^2)$$

gegeben ist.

Der Graph von f wird auch als Gaußsche Glockenkurve bezeichnet.

$$\phi(t; 1.8, 0.01^2) = \frac{1}{0.01\sqrt{2\pi}} e^{-\frac{(t-1.8)^2}{2 \cdot 0.01^2}}$$



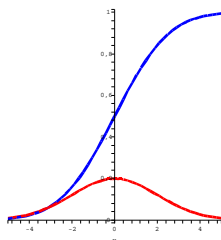
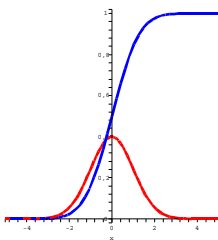
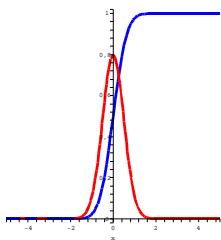
Die Verteilungsfunktion ist dann

$$F(x) = \int_{-\infty}^x f(t) dt = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt =: \Phi(x; \mu, \sigma^2).$$

Bemerkung:

Die Dichtefunktion f ist nicht elementar integrierbar, d.h. f hat keine Stammfunktion, die aus den elementaren Funktionen (Polynome, rationale Funktionen, e^x , $\log(x)$, $\sin(x)$, $\cos(x)$, ...) zusammengesetzt ist.

Verteilung (rot) und Verteilungsfunktion (blau) für $\mu = 0$ und $\sigma = 0.5, 1, 2$



Von spezieller Bedeutung ist die Standardnormalverteilung eine Normalverteilung mit den Parametern $\mu = 0$ und $\sigma^2 = 1$. Die Dichte ist

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} =: \phi(t)$$

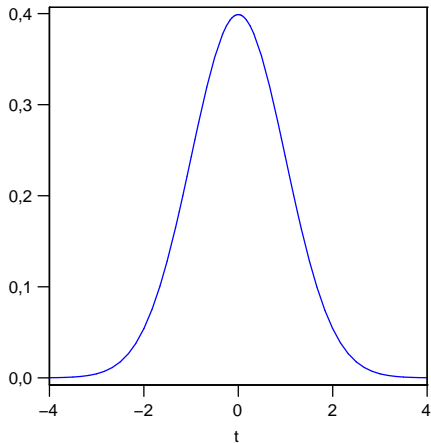
und die zugehörige Verteilungsfunktion schreibt sich als:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt =: \Phi(x).$$

Einfache Eigenschaften von Φ

- $\Phi(0) = 1/2$
- $\Phi(-x) = 1 - \Phi(x)$

$$\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$



Es gilt der Zusammenhang:

Ist die Zufallsvariable X $N(\mu, \sigma^2)$ -verteilt, so ist die Zufallsvariable $\frac{X - \mu}{\sigma}$ standardnormalverteilt:

$$\Phi(x; \mu, \sigma^2) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

Es genügt also, die Funktionswerte von $\Phi(x)$ bestimmen zu können, um die Funktionswerte jeder der Funktionen $\Phi(x; \mu, \sigma^2)$ zu berechnen.

Früher gab es umfangreiche Tafelwerke, in denen die Funktionswerte von Φ tabelliert waren.

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500	0.504	0.508	0.512	0.516	0.520	0.524	0.528	0.532	0.536
0.1	0.540	0.544	0.548	0.552	0.556	0.560	0.564	0.567	0.571	0.575
0.2	0.579	0.583	0.587	0.591	0.595	0.599	0.603	0.606	0.610	0.614
0.3	0.618	0.622	0.626	0.629	0.633	0.637	0.641	0.644	0.648	0.652
0.4	0.655	0.659	0.663	0.666	0.670	0.674	0.677	0.681	0.684	0.688
0.5	0.691	0.695	0.698	0.702	0.705	0.709	0.712	0.716	0.719	0.722
0.6	0.726	0.729	0.732	0.736	0.739	0.742	0.745	0.749	0.752	0.755
0.7	0.758	0.761	0.764	0.767	0.770	0.773	0.776	0.779	0.782	0.785
0.8	0.788	0.791	0.794	0.797	0.800	0.802	0.805	0.808	0.811	0.813
0.9	0.816	0.819	0.821	0.824	0.826	0.829	0.831	0.834	0.836	0.839
1.0	0.841	0.844	0.846	0.848	0.851	0.853	0.855	0.858	0.860	0.862
1.1	0.864	0.866	0.869	0.871	0.873	0.875	0.877	0.879	0.881	0.883
1.2	0.885	0.887	0.889	0.891	0.893	0.894	0.896	0.898	0.900	0.901
1.3	0.903	0.905	0.907	0.908	0.910	0.911	0.913	0.915	0.916	0.918
1.4	0.919	0.921	0.922	0.924	0.925	0.926	0.928	0.929	0.931	0.932
1.5	0.933	0.934	0.936	0.937	0.938	0.939	0.941	0.942	0.943	0.944
1.6	0.945	0.946	0.947	0.948	0.949	0.951	0.952	0.953	0.954	0.954
1.7	0.955	0.956	0.957	0.958	0.959	0.960	0.961	0.962	0.962	0.963
1.8	0.964	0.965	0.966	0.966	0.967	0.968	0.969	0.969	0.970	0.971
1.9	0.971	0.972	0.973	0.973	0.974	0.974	0.975	0.976	0.976	0.977
2.0	0.977	0.978	0.978	0.979	0.979	0.980	0.980	0.981	0.981	0.982

Aufgabe

Die Zufallsvariable sei normalverteilt mit $E(X) = 1$ und $Var(X) = 4$.
Drücken Sie die folgenden Wahrscheinlichkeiten durch Funktionswerte der Funktion Φ aus: $P(|X| > 4)$, $P(|X - 1| > 6)$ und $P(0 \leq X \leq 2)$.

Lösungen:

$$2 - \Phi(5/2) - \Phi(3/2), 2 - 2\Phi(3), 2\Phi(1/2) - 1$$

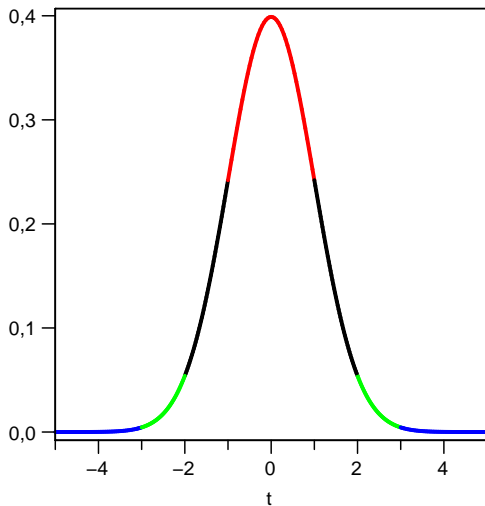
Es gelten die folgenden Aussagen:

- 1 $\phi(t)$ ist symmetrisch bezüglich der Achse $x = \mu$, d.h. es gilt $\phi(\mu + t) = \phi(\mu - t)$ für alle reellen Zahlen t .
- 2 $\phi(t)$ hat ein lokales Maximum an der Stelle $t = \mu$ und sonst keine Extremalstellen.
- 3 Die einzigen beiden Wendestellen von $\phi(t)$ sind $\mu + \sigma$ und $\mu - \sigma$.
- 4 $\lim_{t \rightarrow \pm\infty} \phi(t) = 0$
- 5 $E(X) = \mu$ und $Var(X) = \sigma^2$
- 6 Die folgenden Wahrscheinlichkeiten sind unabhängig von der konkreten Wahl der Parameter μ und σ :

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = 0.683$$

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0.955$$

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0.997$$



Beweis:

Die Zufallsvariable X sei $N(\mu, \sigma^2)$ -verteilt. Dann gilt:

$$\begin{aligned} & P(\mu - \sigma \leq X \leq \mu + \sigma) \\ &= \Phi(\mu + \sigma; \mu, \sigma) - \Phi(\mu - \sigma; \mu, \sigma) \\ &= \Phi\left(\frac{\mu + \sigma - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - \sigma - \mu}{\sigma}\right) \\ &= \Phi(1) - \Phi(-1) \\ &= 2\Phi(1) - 1 \\ &\approx 0.683. \end{aligned}$$

Zwischen vielen diskreten und stetigen Verteilungen bestehen Zusammenhänge der folgenden Art:

Satz (Lokaler Grenzwertsatz von de Moivre und Laplace)

Sei $X \sim B(n; p)$ eine binomialverteilte Zufallsvariable, $\mu = np$ und $\sigma = \sqrt{np(1-p)}$. Dann gilt

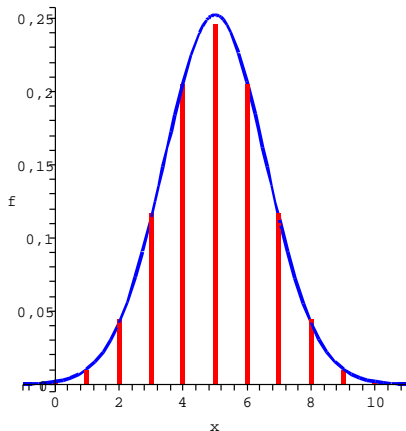
$$f_{Bi}(x; n, p) \approx \phi\left(x; np, \sqrt{np(1-p)}\right)$$

$$P(a \leq X \leq b) \approx \Phi\left(\frac{b - \mu + 0.5}{\sigma}\right) - \Phi\left(\frac{a - \mu - 0.5}{\sigma}\right)$$

mit hinreichender Genauigkeit, falls $n \cdot p \cdot (1-p) > 9$ gilt.

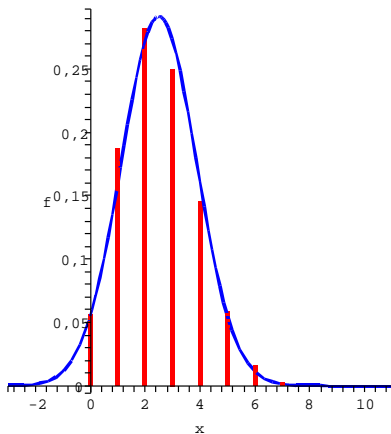
$n = 10$ und $p = 0.5 \rightarrow \mu = n \cdot p = 5$ und $\sigma = \sqrt{n \cdot p \cdot (1 - p)} \approx 1.58$

In der Skizze sehen Sie die Binomialverteilung $f_{Bi}(x; 10, 0.5)$ (als rotes Stabdiagramm) und die (blaue) Glockenkurve $\phi(x; 5, 1.58)$:



$n = 10$ und $p = 0.25 \longrightarrow \mu = n \cdot p = 2.5$ und $\sigma = \sqrt{n \cdot p \cdot (1 - p)} \approx 1.37$

In der Skizze sehen Sie die Binomialverteilung $f_{Bi}(x; 10, 0.25)$ (als rotes Stabdiagramm) und die (blaue) Glockenkurve $\phi(x; 2.5, 1.37)$:



Aufgabe

Bestimmen Sie approximativ die Zahl

$$P = \sum_{k=500}^{600} \binom{1000}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{1000-k} .$$

Lösung:

Z_1, Z_2, \dots, Z_n seien n unabhängige und identisch $N(0, 1)$ -verteilte Zufallsvariablen. Dann ist die Summe ihrer Quadrate $Z = \sum_{i=1}^n Z_i^2$ χ^2 -verteilt mit n Freiheitsgraden. Kurz: $Z \sim \chi_n^2$

Satz (Additionssatz)

Die Summe zweier unabhängiger χ_n^2 -verteilter bzw. χ_m^2 -verteilter Zufallsvariablen ist χ_{n+m}^2 -verteilt.

Wesentliches Beispiel: Für unabhängige Zufallsvariablen $X_i \sim N(\mu, \sigma^2)$ ($i = 1, 2, \dots, n$) gilt für die Stichprobenvarianz

$$S_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

tatsächlich

$$\frac{n-1}{\sigma^2} \cdot S_X^2 \sim \chi_{n-1}^2.$$

Sind X und Y unabhängige Zufallsvariablen mit $X \sim N(0, 1)$ und $Y \sim \chi_n^2$ so besitzt der Quotient

$$\frac{X}{\sqrt{Y/n}} \sim t_n$$

eine t_n -Verteilung (Student-Verteilung) mit n Freiheitsgraden.

Satz

Seien X_1, X_2, \dots, X_n unabhängige und identisch $N(\mu, \sigma^2)$ -verteilte Zufallsvariablen. Dann sind \bar{X} und S_X^2 unabhängig. Der folgende Quotient ist t_{n-1} -verteilt:

$$\frac{(\bar{X} - \mu)\sqrt{n}}{S_X} = \frac{(\bar{X} - \mu)\sqrt{n}}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}} \sim t_{n-1}$$

Sind X und Y unabhängige Zufallsvariablen mit $X \sim \chi_m^2$ und $Y \sim \chi_n^2$ so besitzt der Quotient

$$\frac{X/m}{Y/n} \sim F_{m,n}$$

eine Fisher'sche F-Verteilung mit (m, n) Freiheitsgraden.

Wichtigste Anwendung

Für $X \sim N(\mu_X, \sigma^2)$ und $Y \sim N(\mu_Y, \sigma^2)$ betrachten wir zunächst die beiden Stichprobenvarianzen

$$S_X^2 = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X})^2 \quad \text{Stichprobenumfang } m$$

$$S_Y^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2 \quad \text{Stichprobenumfang } n$$

Dann gilt für das Verhältnis der beiden Stichprobenvarianzen:

$$\frac{S_X^2}{S_Y^2} \sim F_{m-1, n-1}$$

Anmerkung: Ist eine Zufallsvariable W nach $F_{m,n}$ -verteilt, so ist $1/W$ natürlich $F_{n,m}$ -verteilt.