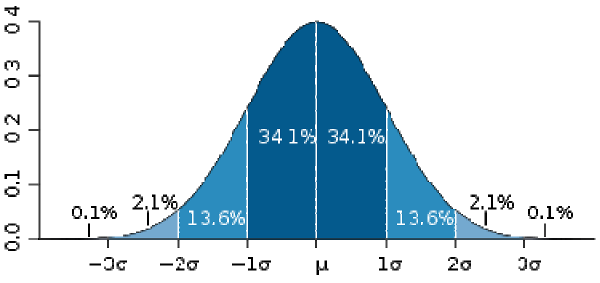
Vorlesung Gesamtbanksteuerung

Mathematische Grundlagen II

Dr. Klaus Lukas

Carsten Neundorf



Agenda

- Wiederholung stetige Renditen
- deskriptive Statistik
- Verteilungsparameter
- Erwartsungswert und Varianz
- Diskrete Zufallsvariable
- Stetige Zufallsvariable
- Die Normalverteilung
- Schätzfunktionen

Stetige Rendite

Sind S_s und S_t zwei Kurswerte eines Wertpapiers oder auch Indexes, so ergibt sich die stetige Rendite als natürlicher Logarithmus des Zuwachsverhältnisses:

$$r_{s,t} = \ln \frac{S_t}{S_s} = \ln S_t - \ln S_s$$

Die Datenreihen

GESAMTBANKSTEUERUNG

Im Verlaufe der Vorlesung sollen die eingeführten Konzepte jeweils auch auf Echtdaten angewandt werden.

Hierfür betrachten wir folgende Indizes:

Rentenindex REX
Aktienindex DAX

Agenda

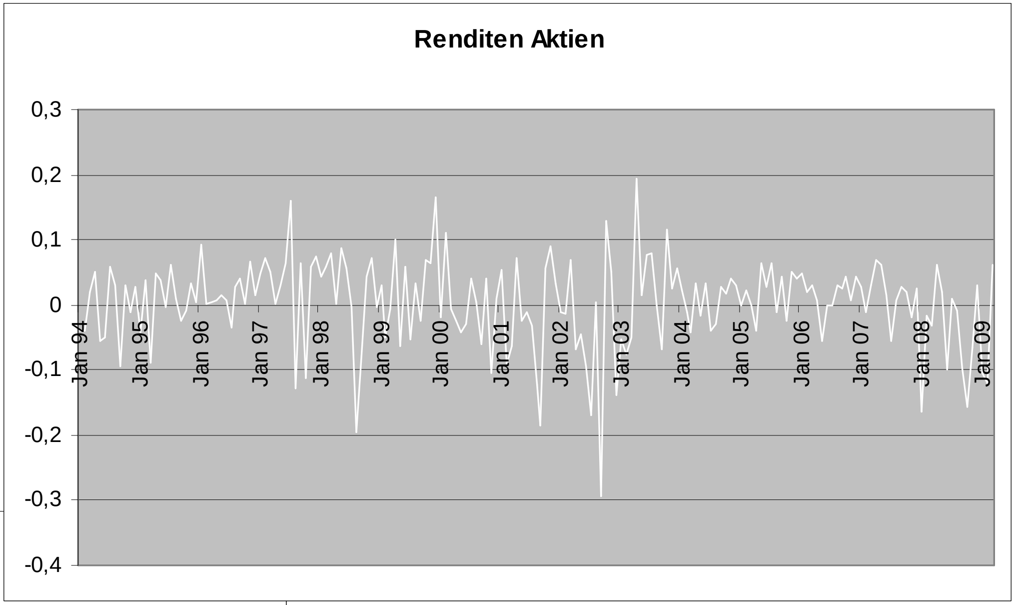
- Wiederholung stetige Renditen
- deskriptive Statistik
- Verteilungsparameter
- Erwartsungswert und Varianz
- Diskrete Zufallsvariable
- Stetige Zufallsvariable
- Die Normalverteilung
- Schätzfunktionen

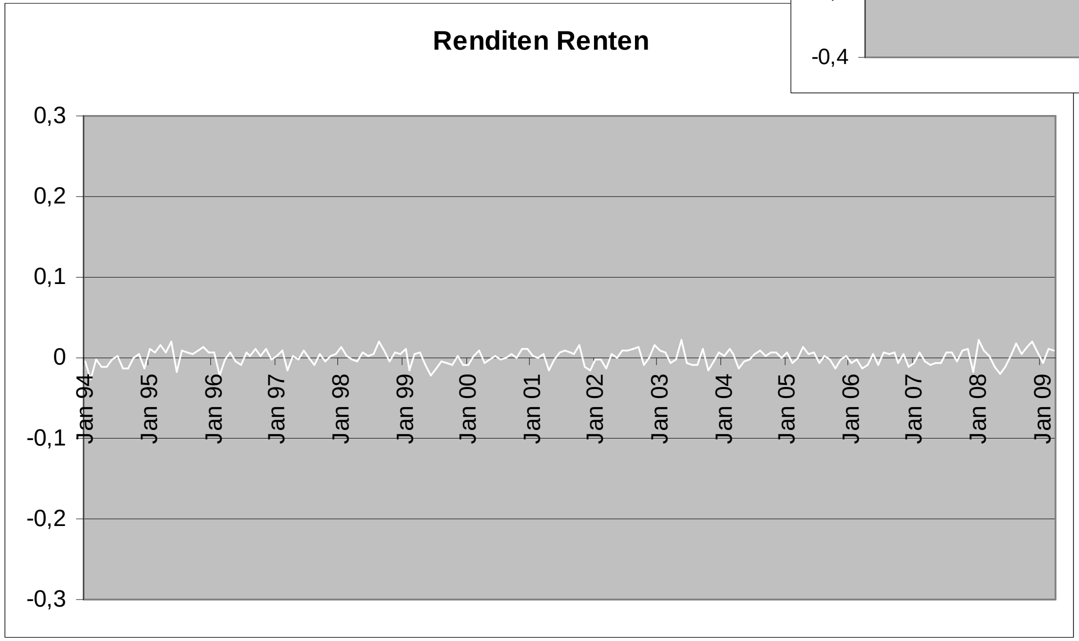
Ziele der deskriptiven Statistik

Die deskriptive Statistik zielt darauf ab, eine unüberschaubare Datenmenge durch möglichst wenige jedoch noch aussagekräftige Zahlen zu charakterisieren. Im Extremfall mittels lediglich einer Zahl.

Als erster Schritt, bevor irgendwelche Kennzahlen berechnet werden, bietet es sich an, die Daten erstmal graphisch darzustellen.

Graphische Zeitreihenanalyse





Grundbegriffe der Statistik

- Als *Grundgesamtheit* G wird die Menge aller statistischer Einheiten bezeichnet, über die man Aussagen gewinnen will. Diese muss klar umgrenzt sein.

- An den statistischen Einheiten werden interessante Größen beobachtet, die sogenannten *Merkmale* X.

- Typischerweise wird man nicht alle Einheiten der Grundgesamtheit in die Untersuchung einbeziehen, sondern ein nach bestimmten Kriterien ausgewählte Teilgesamtheit, die sogenannte *Stichprobe* vom Umfang n.

Grundbegriffe der Statistik

- Konkrete Werte eines Merkmals für eine bestimmte statistische Einheit wird *Ausprägungen des Merkmales* genannt. Abgekürzt mit x.

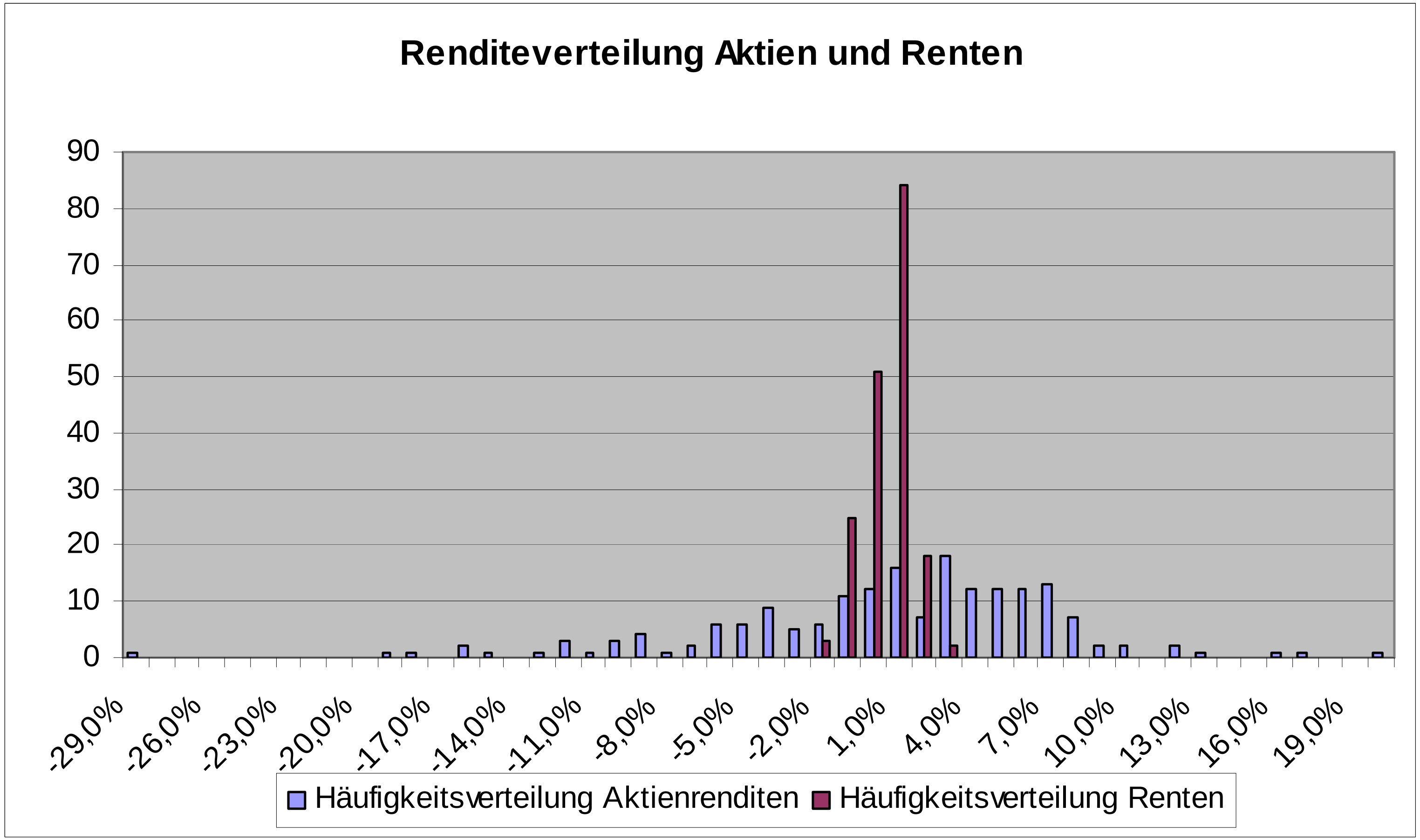
- Als absolute bzw. relative Häufigkeit einer Ausprägung a_j,j=1,...,k bezeichnet man die Anzahl bzw. den Anteil von Werten der Stichprobe, die mit a_j übereinstimmt. Die absoluten Häufigkeiten werden mit H_j bezeichnet, die relativen wir mit $h_j=H_j/n$.

- Häufigkeitsfunktion h(x)

$F(x) = \sum_{a_j < x} h_j$ wird als empirische Verteilungsfunktion bezeichnet

Häufigkeitsverteilung

GESAMTBANKSTEUERUNG



Agenda

- Wiederholung stetige Renditen
- deskriptive Statistik
- Verteilungsparameter
- Erwartsungswert und Varianz
- Diskrete Zufallsvariable
- Stetige Zufallsvariable
- Die Normalverteilung
- Schätzfunktionen

Lageparameter

GESAMTBANKSTEUERUNG

Der bekannteste Lageparameter ist der Durchschnittswert oder auch das arithmetische Mittel

$$\bar{x} = \frac{1}{n} * \sum_{i=1}^{n} x_i$$

Da jedem Merkmalswert das Gewicht 1/n zugeordnet wird, kann es vorkommen, dass der Mittelwert mit keinem beobachteten Wert übereinstimmt.

Lageparameter

Ein weiterer Lageparameter, der auch in den Beobachtungswerten enthalten ist, ist der Median.

Er ist durch die Eigenschaft definiert, dass mindestens 50% aller Werte kleiner oder gleich sind und mindestens 50% aller Werte auch größer oder gleich sind.

Um den Median zu ermitteln, müssen zuerst alle Ausprägungen der Grösse nach sortiert werden. Anschließend errechnet sich der Median wie folgt:

$$x_{Med} = x_{\left\lfloor \frac{n}{2} \right\rfloor + 1}$$

Lageparameter

GESAMTBANKSTEUERUNG

Will man nicht den „mittleren" Wert haben, sondern den p-größten, so kann man die Formel wie folgt abwandeln:

$$q_p = x_{\lfloor p*n \rfloor + 1}$$

Dies wird das p-Quantil genannt. p ist hierbei immer als %-Zahl anzusehen.

Streuungsparameter

GESAMTBANKSTEUERUNG

Unser Ziel ist es, mit möglichst wenigen Kennzahlen eine große Datenmenge zu beschreiben.

Am Beispiel der Renditeverteilung der Aktien und der Renten sehen wir aber, dass der Mittelwert die Daten nicht ausreichend beschreibt, da er für beide Datenreihen nahezu identisch bei 0,2% liegt, aber die größten und kleinsten Werte der Aktien deutlich mehr von der Mitte abweichen als die Werte der Renten.

Deshalb wird nun ein Maß für die „Streuung“ eingeführt.

Streuungsparameter

GESAMTBANKSTEUERUNG

Ein mögliches Maß für diese Streuung ist die mittlere quadratische Abweichung s^2, auch empirische Varianz bezeichnet. Als Formel ausgedrückt.

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

Hieraus wird dann die Standardabweichung s abgeleitet, indem aus s^2 die Wurzel gezogen wird.

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

Variationskoeffizient

Führt man die beiden Kennzahlen Standardabweichung und Mittelwert zusammen, so erhält man den sogenannten Variationskoeffizient V.

$$V = \frac{s}{\bar{x}}$$

Agenda

GESAMTBANKSTEUERUNG

GESAMTBANKSTEUERUNG

- Wiederholung stetige Renditen
- deskriptive Statistik
- Verteilungsparameter
- Erwartsungswert und Varianz
- Diskrete Zufallsvariable
- Stetige Zufallsvariable
- Die Normalverteilung
- Schätzfunktionen

Wahrscheinlichkeit und Zufallsereignis

GESAMTBANKSTEUERUNG

Bisher haben wir untersucht, wie man gegebene Datenmenge mit wenigen Kennzahlen möglichst einfach beschreiben kann.

Ein weiterer Aspekt der Stochastik ist die Beantwortung der Frage, welche Ergebnisse sind in der Zukunft möglich und wie wahrscheinlich sind sie.

Damit haben wir schon einen wesentlichen Begriff unbewusst verwendet, die Wahrscheinlichkeit.

Wahrscheinlichkeit und Zufallsereignis

GESAMTBANKSTEUERUNG

Die Wahrscheinlichkeit mathematisch exakt zu definieren, ist nicht ganz einfach. Dazu müssen erst noch zwei weitere Begriffe eingeführt werden.

Als Zufallsvorgang wird ein Geschehen bezeichnet, bei dem aus einer gegebenen Situation heraus mehrere sich gegenseitig ausschließende Folgesituationen möglich sind und es unsicher ist, welche eintritt.

Alle möglichen Folgesituationen werden als Ergebnismenge bezeichnet, eine konkrete Folgesituation als Elementarereignis.

Wahrscheinlichkeit und Zufallsereignis

GESAMTBANKSTEUERUNG

GESAMTBANKSTEUERUNG

Mit der Wahrscheinlichkeit eines Ereignisses A, kurz p(A), wollen wir die Chance für das Eintreten des Ereignisses beschreiben.

Mathematisch muss gelten:

1. $p(A) \geq 0 \;\; \textit{für jedes Ereignis } A$
2. $p(\textit{sicheres Ereignis}) = 1$
3. $p(A_1 \cup A_2 \cup A_3 \cup \ldots) = p(A_1) + p(A_2) + p(A_3) + \ldots$

Wahrscheinlichkeit und Zufallsereignis

Wir definieren

$$p(A) = \frac{Anzahl\ der\ \ für\ A\ günstigen\ Fälle}{Anzahl\ aller\ möglichen\ Fälle}$$

und als Gegenwahrscheinlichkeit q(A)

$$q(A) = 1 - p(A)$$

die Wahrscheinlichkeit, dass A nicht eintritt.

diskrete Zufallsvariable

Ordnet man den Ergebnissen eines Zufallsvorgangs Zahlen (Realisationen) zu, so erhält man eine Zufallsvariable.

Ein Beispiel: zweimal Würfel. Die möglichen Ergebnisse bestehen aus den 36 Zahlenpaaren $\omega=(i,j)$, $1\leq i,j\leq 6$. Für die Variable X=„Summe der Augenzahlen“ erhält man zu jedem Ergebnis (i,j) den Wert $x=i+j$. X wird dann zu einer Zufallsvariable mit der Ergebnismenge $\{2,3,\ldots,12\}$.

Es interessieren dann Ergebnisse wie

$\{X=4\}$=„Die Augenzahl ist 4“

diskrete Zufallsvariable

GESAMTBANKSTEUERUNG

Dies lässt sich durch die ursprünglichen Ergebnisse (i,j) ausdrücken. Es gilt

$$\{X=4\}=\{(1,3),(2,2),(3,1)\}$$

Um die Wahrscheinlichkeit dieses Ereignisses zu berechnen, führt man dies auch wieder auf die ursprünglichen Ergebnisse zurück.

$$P(X=4) = P(1,3) + P(2,2) + P(3,1) = \frac{3}{36} = \frac{1}{12}$$

diskrete Zufallsvariable

Eine Zufallsvariable X heißt diskret, falls sie nur endlich oder abzählbar unendlich viele Werte $x_1, x_2, \ldots, x_k, \ldots$ annehmen kann. Die Wahrscheinlichkeitsverteilung von X ist gegeben durch die Wahrscheinlichkeiten:

$$P(X = x_i) = p_i, \quad i = 1, 2, \ldots, k, \ldots$$

Wahrscheinlichkeitsfunktion

GESAMTBANKSTEUERUNG

Die Wahrscheinlichkeitsfunktion f(x) einer diskreten Zufallsvariablen X ist für $x \in \Re$ definiert als

$$f(x) = \begin{cases} P(X = x_i) = p_i & x = x_i \in \{x_1, x_2, \ldots, x_k, \ldots\} \\ 0, & sonst \end{cases}$$

Durch die Wahrscheinlichkeitsverteilung bzw. –funktion ist auch die Verteilungsfunktion $F(x)=P(X \leq x)$ gegeben.

$$F(x) = P(X \leq x) = \sum_{i: x_i \leq x} f(x_i)$$

Lageparameter diskreter Zufallsvariablen

GESAMTBANKSTEUERUNG

Ebenso, wie bei der Analyse von Kennzahlen, können auch für diskrete Zufallsvariable Erwartungswert, Varianz und Standardabweichung berechnet werden.

Der Erwartungswert errechnet sich mit Hilfe der Wahrscheinlichkeitsverteilung wie folgt:

$$E(X) = x_1 p_1 + \ldots + x_k p_k + \ldots = \sum_{i \geq 1} x_i p_i$$

und mit Hilfe der Wahrscheinlichkeitsfunktion f(x)

$$E(X) = x_1 f(x_1) + \ldots + x_k f(x_k) + \ldots = \sum_{i \geq 1} x_i f(x_i)$$

Statt E(X) findet auch das Symbol μ Verwendung.

Lageparameter diskreter Zufallsvariablen

GESAMTBANKSTEUERUNG

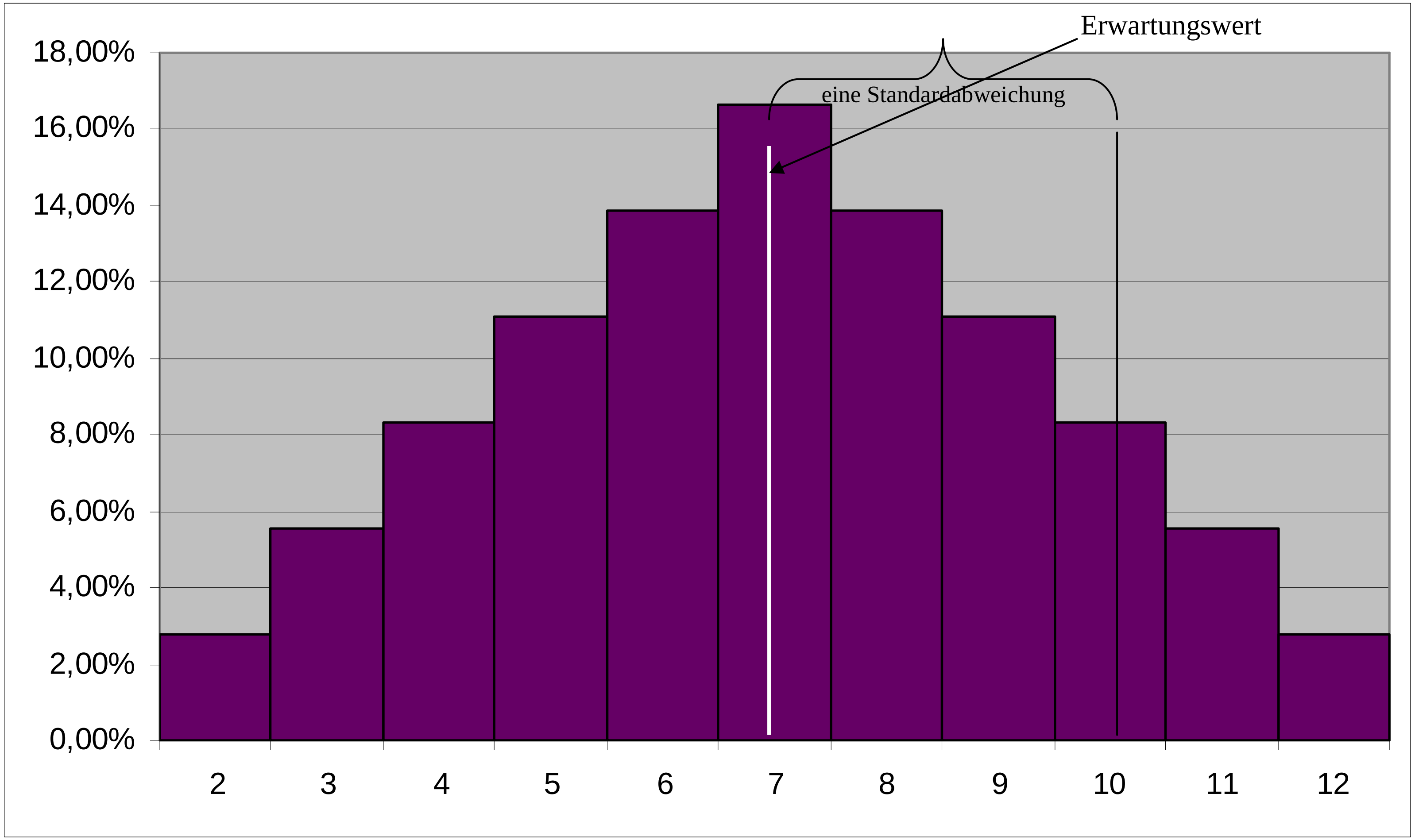
Die Varianz einer diskreten Zufallsvariable ist

$$\sigma^2 = Var(X) = (x_1 - \mu)^2 p_1 + \ldots + (x_k - \mu)^2 p_k + \ldots$$
$$= \sum_{i \geq 1} (x_i - \mu)^2 f(x_i)$$

Die Standardabweichung ist

$$\sigma = \sqrt{Var(X)}$$

Lageparameter diskreter Zufallsvariablen

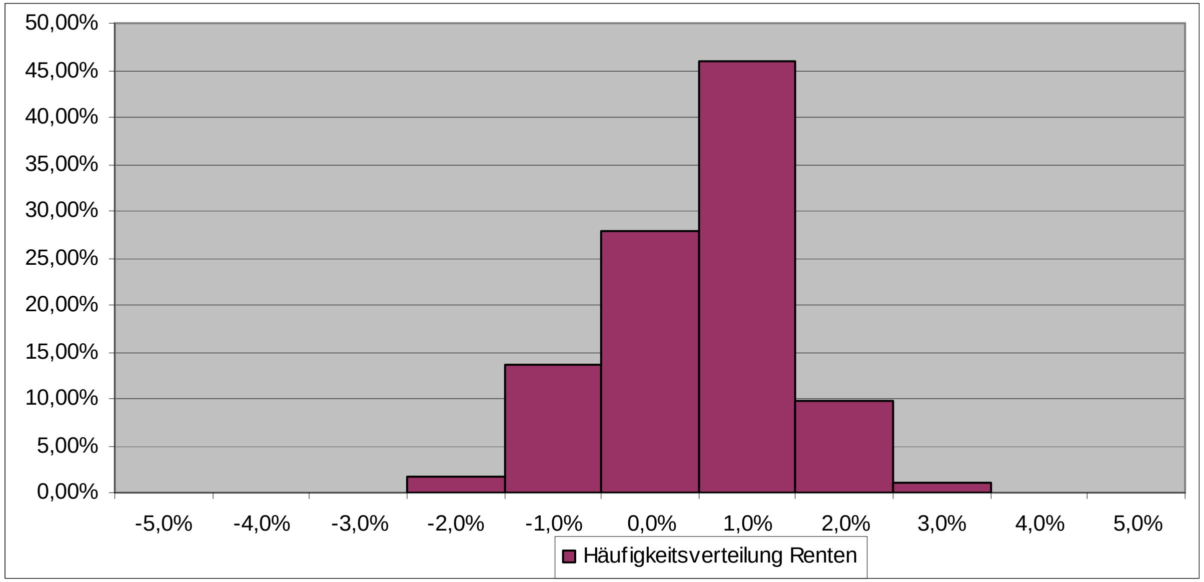


Agenda

GESAMTBANKSTEUERUNG

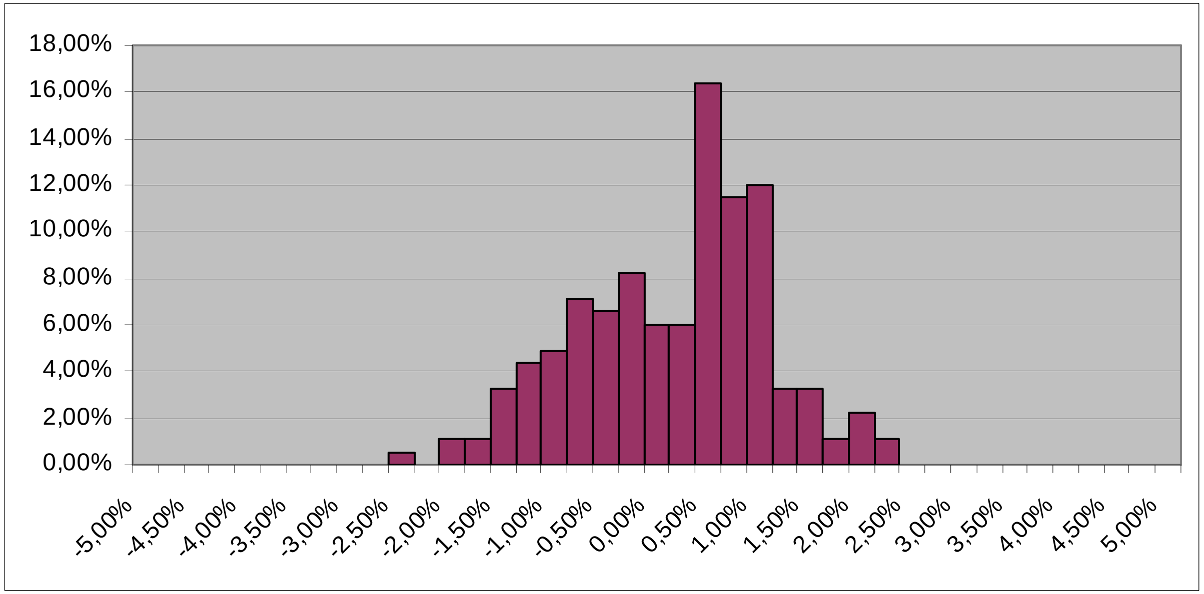
- Wiederholung stetige Renditen
- deskriptive Statistik
- Verteilungsparameter
- Erwartsungswert und Varianz
- Diskrete Zufallsvariable
- Stetige Zufallsvariable
- Die Normalverteilung
- Schätzfunktionen

stetige Zufallsvariablen



In den bisherigen Untersuchungen haben Zufallsvariablen untersucht, die eine feste Zahl von Wahrscheinlichkeiten haben. Die Wahrscheinlichkeit, dass ein Ereignis zwischen zwei Werten liegt, kann man dann an der Fläche der Balken ablesen. Teilt man die Balken feiner ein, kann man das immer noch an der Fläche der Balken ablesen.

stetige Zufallsvariablen



Führt man diesen Prozess fort und lässt die Breite immer mehr gegen null gehen, so nähert sich die Fläche dem Integral unter der Funktion an.

stetige Zufallsvariablen

Diese Überlegungen führen zur folgenden Definition von stetigen Zufallsvariablen, wobei gleichzeitig die Dichtefunktion f(x) definiert wird:

Eine Zufallsvariable X heißt stetig, wenn es eine Funktion $f(x) \geq 0$ gibt, so dass für jedes Intervall [a,b]

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx$$

gilt. Die Funktion f(x) heißt Wahrscheinlichkeitsdichte von X oder einfach nur Dichte.

stetige Zufallsvariablen

Noch ein paar Eigenschaften:

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$$

und

$$P(X = x) = 0 \quad \text{für alle } x \in \Re$$

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

stetige Zufallsvariablen

Aus der Definition von Dichten erhält man für die Verteilungsfunktion $F(x)=P(X\leq x)=P(-\infty<X \leq x)$

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

d.h. F(x) ist die Stammfunktion der Dichte.

Lageparameter stetiger Zufallsvariablen

Der Erwartungswert E(X) einer stetigen Zufallsvariable X mit Dichte f(x) ist

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx$$

Die Varianz Var(X) einer stetigen Zufallsvariable X mit Dichte f(x) ist

$$\sigma^2 = Var(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx.$$

$$\sigma = \sqrt{Var(X)}$$

wird Standardabweichung genannt.

Median und Quantile

Für $0<p<1$ ist das p-Quantil x_p die Zahl auf der x-Achse, für die

$$F(x_p)=p$$

gilt. Der Median x_{med} ist das 50%-Quantil, es gilt als

$$F(x_{med})=1/2.$$

Agenda

GESAMTBANKSTEUERUNG

- Wiederholung stetige Renditen
- deskriptive Statistik
- Verteilungsparameter
- Erwartsungswert und Varianz
- Diskrete Zufallsvariable
- Stetige Zufallsvariable
- Die Normalverteilung
- Schätzfunktionen

Die Normalverteilung

GESAMTBANKSTEUERUNG

Eine Zufallsvariable X heißt normalverteilt mit Parametern $\mu \in \Re$ und $\sigma^2 > 0$, kurz $X \sim N(\mu, \sigma^2)$, wenn sie die Dichte

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right), \quad x \in \Re$$

besitzt. Es gilt

$$E(X) = \mu, \qquad Var(X) = \sigma^2$$

Speziell für $\mu = 0$, $\sigma^2 = 1$ erhält man die Standardnormalverteilung N(0,1) mit der Dichte

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} x^2 \right)$$

Die Normalverteilung

GESAMTBANKSTEUERUNG

GESAMTBANKSTEUERUNG

Welche Eigenschaften hat diese Verteilung?

1. Sie ist symmetrisch zu μ, d.h. es gilt

$$f(\mu-x)=f(\mu+x)$$

2. Die Normalvertielung hat Glockenform mit dem Maximum an der Stelle μ und den Wendepunkten bei $\mu\pm\sigma$.

3. Je größer σ, desto schneller fällt die Kurve gegen 0.

Die Normalverteilung

Die Verteilungsfunktion lässt sich nicht analytisch berechnen und durch bekannte Funktionen in geschlossener Form schreiben. Deshalb muss F(x) durch numerische Verfahren berechnet werden. Entsprechende Tabellen sind in den Kalkulationsprogrammen enthalten.

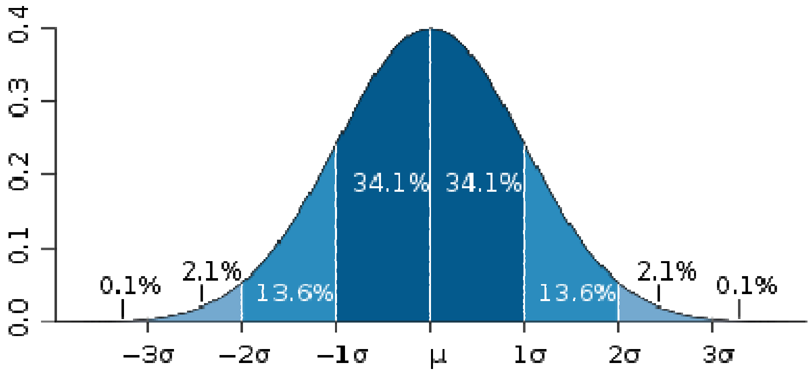
Wichtige Quantile für Standardnormalverteilung sind ebenfalls in den Programmen vorberechnet und in Tabellen hinterlegt.

Die Normalverteilung

GESAMTBANKSTEUERUNG

Zentrale Schwankungsintervalle

Für k=1 ist $P(\mu-\sigma \leq X \leq \mu+\sigma)=68{,}27\%$
Für k=2 ist $P(\mu-2\sigma \leq X \leq \mu+2\sigma)=95{,}45\%$
Für k=3 ist $P(\mu-3\sigma \leq X \leq \mu+3\sigma)=99{,}73\%$



Agenda

- Wiederholung stetige Renditen
- deskriptive Statistik
- Verteilungsparameter
- Erwartsungswert und Varianz
- Diskrete Zufallsvariable
- Stetige Zufallsvariable
- Die Normalverteilung
- Schätzfunktionen

Parameterschätzung

GESAMTBANKSTEUERUNG

Die Ziehung einer Stichprobe dient meistens dem Zweck, Informationen über das Verhalten eines Merkmales in der Grundgesamtheit zu gewinnen.

Dies ist mit einigen Unsicherheiten verbunden. Das zugrunde gelegte Modell ist eben nur ein Modell und die gemessenen Daten enthalten auch z.B. Messfehler.

Die in den Stichproben beobachteten Anteilswerte liefern einen Schätzer für den wahren Anteil in der Grundgesamtheit.

Punktschätzung

Ziel der Punktschätzung ist es, einen möglichst genauen Näherungswert eines Grundgesamtheitsparameters anzugeben.

Man unterscheidet dabei zwei Fälle:

1. Kennwerte einer beliebigen, unbekannten Verteilung
2. spezifische Parameter eines angenommenen Verteilungsmodells

Als Beispiel für den 2. Fall ist die Schätzung der Parameter μ,σ^2, wenn von einer Normalverteilung $N(\mu,\sigma^2)$ auszugehen ist.

Punktschätzung

GESAMTBANKSTEUERUNG

GESAMTBANKSTEUERUNG

Eine Schätzfunktion für den Grundgesamtheits-parameter θ ist eine Funktion

$$T=g(X_1,\ldots,X_n)$$

der Stichprobenvariablen $X_1,\ldots,X_n$. Der aus den Realisationen $x_1,\ldots,x_n$ resultierende numerische Wert

$$g(x_1,\ldots,x_n)$$

ist der zugehörige Schätzwert.

Punktschätzung

In der deskriptiven Statistik wurden Lage- und Streuungsparameter der Stichprobe bestimmt. Hinter diesen deskriptiven Parametern stehen Schätzfunktionen, deren Argumente Zufallsvariablen sind. Die resultierenden Realisationen dieser Schätzfunktionen entsprechen dann direkt den deskriptiven Lage- bzw. Streuungsparametern.

Punktschätzung

$$\overline{X} = g(X_1, \ldots X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i$$

ist dann eine Schätzfunktion für den Erwartungswert $\mu = E(X)$,

$$\tilde{S}^2 = g(X_1, \ldots X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i$$

eine Schätzfunktion für die Varianz $\sigma^2 = \mathrm{var}(X)$.

Anwendungsbeispiel