STOCHASTIK

Tschebyscheff-Ungleichung

Einführung
mit Anwendungsbeispielen

Datei Nr. 36111

Friedrich W. Buckel

Stand 12. April 2010

INTERNETBIBLIOTHEK FÜR SCHULMATHEMATIK

www.mathe-cd.de

Inhalt

Im Text 36112 findet man Aufgaben zur Anwendung der Tschebyscheff-Ungleichung.

1 Wiederholung: Standardabweichung

Die Standardabweisung wurde eingeführt, um ein Maß dafür zu bekommen, um wie viel die einzelnen Werte einer Messung „durchschnittlich“ vom Mittelwert abweichen. Es handelt sich also um eine **statistische Größe**, die beschreibt, wie sich eine Stichprobe verhält.

Verwendet man diesen Begriff vorausschauend für die Zukunft, dann begibt man sich in die **Wahrscheinlichkeitsrechnung**, wo die Standardabweichung eben diese Abweichung *voraussagt*.

Bei einer Stichprobe wird etwas gezählt oder gemessen. Dazu verwendet man eine Variable, die sogenannte Zufallsvariable, die man meist mit X bezeichnet. Zu X gehört eine Definitionsmenge $S = \{x_1 ; x_2 ; ... ; x_k\}$, man nennt sie auch den Ereignisraum. Diese k Werte können also bei der Erfassung unserer Stichprobe für X auftreten. Jeder mit einer gewissen Häufigkeit (statistisch gesehen) bzw. Wahrscheinlichkeit. Damit gelingt es leicht, den statistischen Mittelwert bzw. in der Vorausschau den zu erwartenden Mittelwert, also den Erwartungswert, zu berechnen.

Beispiel 1: Für die Produktion von Schrauben ist eine regelmäßige Maschinenkontrolle notwendig, damit der Durchmesser dieser Schrauben nicht allzu sehr von der Sollstärke abweicht. Die betreffende Maschine soll Schauben mit dem Durchmesser d = 12,5 mm produzieren.
Messungen ergaben folgende Tabelle: (Die 3. und 6. Spalte wurde erst nach der Berechnung des Erwartungswertes eingetragen!)

X	h bzw. p	$\lvert X-E(X)\rvert$	X	h bzw. p	$\lvert X-E(X)\rvert$
12,0	0,02	0,448	12,5	0,40	0,052
12,1	0,05	0,348	12,6	0,10	0,152
12,2	0,08	0,248	12,7	0,05	0,252
12,3	0,10	0,148	12,8	0,03	0,352
12,4	0,15	0,048	12,9	0,02	0,452

Hier ist X die Zufallsvariable „Dicke der Schraube“. h ist die bei Prüfungen ermittelte relative Häufigkeit, die man dann als (empirische) Wahrscheinlichkeit übernimmt.
Der Ereignisraum $S = \{12{,}0 ; 12{,}1 ; ... ; 12{,}9\}$ umfasst k = 10 Werte.
Ermitteln wir zuerst den Mittelwert bzw. Erwartungswert für X:

$$\boxed{E(X) = x_1 \cdot p_1 + x_2 \cdot p_2 + ... + x_{10} \cdot p_{10} = 12{,}448}$$

Kennt man ihr, kann man die Absolutbeträge der Abweichungen der Messwerte von diesem Erwartungswert in die 3. und 6. Spalte eintragen (was bereits geschehen ist).
Der Mittelwert dieser Abweichungen berechnet man mit der Formel

$$\boxed{\overline{X} = |x_1 - E| \cdot p_1 + |x_2 - E| \cdot p_2 + ... + |x_k - E| \cdot p_k}$$

$$\overline{X} = 0{,}448 \cdot 0{,}02 + 0{,}348 \cdot 0{,}05 + ... + 0{,}452 \cdot 0{,}002 = 0{,}1364$$

Um diesen Wert weichen durchschnittlich die Durchmesser vom zu erwartenden Mittelwert ab. Doch mit diesem Wert wird kaum gearbeitet, weil es sich gezeigt hat, dass man mit der **Standardabweichung** eine bessere Auswertungsmöglichkeit hat.

Unter der Varianz V(X) einer Zufallsgröße versteht man:

$$\boxed{V(X) = (x_1 - E)^2 \cdot p_1 + (x_2 - E)^2 \cdot p_2 + ... + (x_k - E)^2 \cdot p_k}$$

Die Wurzel daraus nennt man die **Standardabweichung**: $\boxed{\sigma = \sqrt{V(X)}}$

Für diese Varianz werden die also Abweichungen (Differenzen) quadriert und dann mit der zugehörigen relativen Häufigkeit bzw. Wahrscheinlichkeit multipliziert.

Oft schreibt man anstatt p_1 den Ausdruck $P(X = x_1)$ hin, damit klarer wird, was man meint, nämlich die Wahrscheinlichkeit dafür, dass X den Wert x_1 annimmt:

$$\boxed{V(X) = (x_1 - E)^2 \cdot P(X = x_1) + (x_2 - E)^2 \cdot P(X = x_2) + ... + (x_k - E)^2 \cdot P(X = x_k)}$$

Verwendet man noch das Summenzeichen, dann bleibt diese Kurzformel übrig:

$$\boxed{V(X) = \sum_{i=1}^{k} (x_i - E)^2 \cdot P(X = x_i)}$$

womit man sagen will, dass man für jeden Index i von 1 bis k diese Produkte berechnet und diese dann addiert, also genau wie es in der Formel darüber steht.

Für unser Beispiel ergibt dies

$$V(X) = 0{,}448^2 \cdot 0{,}02 + 0{,}348^2 \cdot 0{,}05 + ... + 0{,}452^2 \cdot 0{,}02 = 0{,}031896$$

Schaut man sich die zugehörige Maßeinheit an, dann erhält man wegen des Quadrierens natürlich mm^2. Also muss man am Ende wieder die Wurzel ziehen, dies ergibt dann endlich das, was man Standardabweichung nennt:

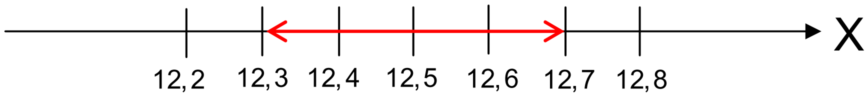
$$\sigma(X) = \sqrt{V(X)} = 0{,}1785945 \approx 0{,}18 \ (mm)$$

Ich will hier nicht weiter darauf eingehen, warum diese Standardabweichung nützlicher ist, als die mittlere Abweichung $\overline{X}$ von der Seite zuvor. Wir wollen Sie hier vorerst nur als *ein* Maß für die Abweichung der Werte einer Zufallsvariablen X von ihrem Erwartungswert E verstehen.

Eines erkennt man sofort: Diese Standardabweichung fällt groß aus, wenn stark von E abweichende Werte mit großen Wahrscheinlichkeiten auftreten. Ist umgekehrt diese Standardabweichung klein, dann treten auch keine großen Abweichungen auf.

2 Eine neue Fragestellung

Bei der automatischen Herstellung in unserer Schraubenfabrik betrachten wir nun wieder die Schrauben vom Solldurchmesser 12,5 mm. Der Hersteller gibt eine Toleranzgrenze von 0,2 mm für die Fertigung vor, d.h. die zulässigen Durchmesser dürfen nur im Bereich $\left[12{,}3\,;\,12{,}7\right]$ liegen (alle Werte ab jetzt in mm).



Dieses Intervall kann man durch Ungleichungen auf diese Arten beschreiben:

$$12{,}3 \le X \le 12{,}7 \qquad \left|-12{,}5\right.$$

Oder: $-0{,}2 \le X - 12{,}5 \le 0{,}2$ (2)

Und das heißt: $|X - 12{,}5| \le 0{,}2$ (3)

Alle drei Ungleichungen haben als Lösungsmenge das Intervall $\left[12{,}3\,;\,12{,}7\right]$. Mit etwas Übung erkennt man, dass (2) und (3) genau von den Zahlen X gelöst wird, die vom Solldurchmesser 12,5 höchstens den Abstand 0,2 haben. Gerade die Betragsungleichung wird dafür meistens verwendet.
(Man lerne diese Schreibweise! Siehe auch Text 41001 – Seite 9.)

Die Kontrolle ergibt nun bei dieser Maschine tatsächlich den Erwartungswert (d. h. vorhergesagter Mittelwert) von 12,5 und eine Standardabweichung von $\sigma(X) = 0{,}08$.

Mit welcher Wahrscheinlichkeit ist daher eine beliebige Schraube brauchbar (d.h. innerhalb dieses Toleranzintervalls), oder mit anderen Worten: Wie viel Prozent der Schrauben werden brauchbar sein?
Dies ist für einen produzierenden Betrieb eine sehr wichtige Fragestellung.

Die Mathematik hilft hier auf relativ einfache Weise weiter. Ich will hier gleich die Lösung zeigen, im nächsten Abschnitt werden wir begründen, warum diese Methode zum richtigen Ergebnis führt.
Das geeignete Werkzeug heißt **TSCHEBYSCHEFFsche Ungleichung**. Sie sieht so aus.

$$\boxed{P(|X - E| \ge c) \le \frac{\sigma^2}{c^2}}$$

Sie gibt eine Wahrscheinlichkeit an, und zwar nicht die vom Toleranzbereich $|X - 12{,}5| \le 0{,}2$ sondern vom Gegenteil, denn in der Formel steht ja $|X - 12{,}5| \ge 0{,}2$.
Wir erhalten also eine Aussage über die Wahrscheinlichkeit dafür, dass ein Schraubendurchmesser außerhalb des Toleranzbereiches liegt, eine Schraube also eigentlich unverkäuflich ist.
Die Zahl c ist das Toleranzmaß, also bei uns die 0,2 (mm), und σ die Standardabweichung.

$$P(|X - 12{,}5| \ge 0{,}2) \le \frac{0{,}08^2}{0{,}2^2} = 0{,}16$$

Also sind höchstens 16 % dieser Schrauben unbrauchbar.
Ist das nicht eine Supermethode?