

Lagemaße



Quiz Lagemaße

Dauer: 15 Minuten Abgabeformat: Textdatei

	Ja	Nein
1. Der Mittelwert ist ein Wert der auch in den Daten vorkommen muss.		X
2. Der Median ist ein Wert der nicht in den Daten vorkommen muss.	X	
3. Um den Mittelwert zu ermitteln müssen die Daten der Größe nach geordnet werden.		X
4. Extremwerte / Ausreißer beeinflussen den Mittelwert stärker als den Median.	X	
5. Der Mittelwert kann niemals gleich dem kleinsten oder größten Wert entsprechen.		X
6. Median und Mittelwert können den gleichen Wert annehmen.	X	



Diagramme

Wozu Diagramme und beschreibende Statistik?

Um einen erster Eindruck über Daten zu gelangen, Tendenzen erkennen.
Fehler in Daten erkennen. Hintergrundinformationen zur eigenen Statistik geben

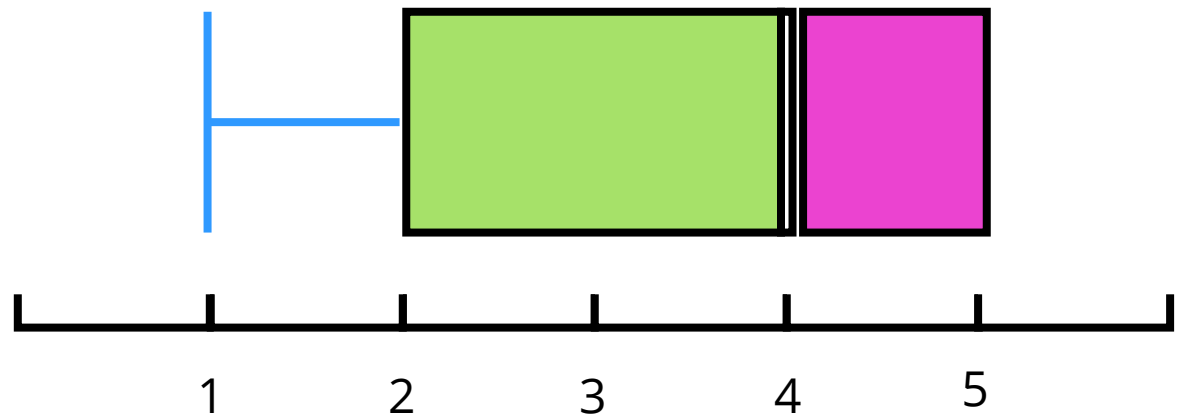


Boxplot zeichnen mit 15 Daten (einfachste Version)

4 2 2 5 2 4 4 4 4 5 3 5 2 5 1

Unsere Datenreihe: 1 2 2 **2** 2 3 4 **4** 4 4 4 **5** 5 5 5

Median = 4



Streuungsmaße



Varianz und Standardabweichung berechnen mit Zwischenschritten

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

	x_i	$x_i - \bar{x}_n$	$(x_i - \bar{x}_n)^2$
x_1	3	0,4	0,16
x_2	4	1,4	1,96
x_3	2	-0,6	0,36
x_4	1	-1,6	2,56
x_5	3	0,4	0,16

↓
 $\bar{x}_n = 2,6$

$$\sum_{i=1}^n = 5,2$$

$$5,2 * \frac{1}{n-1} = 1,3$$

$$s = \sqrt{1,3} = 1,14$$



Erwartungswert und Varianz bei bekannten Wahrscheinlichkeiten

Unendliches Zufallsexperiment

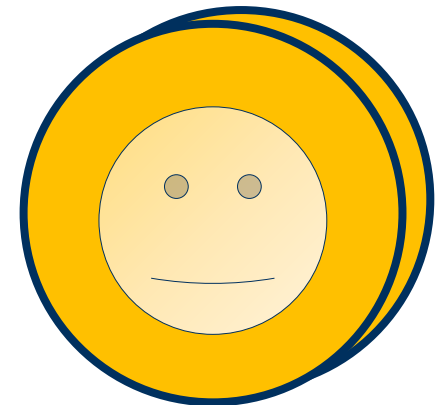
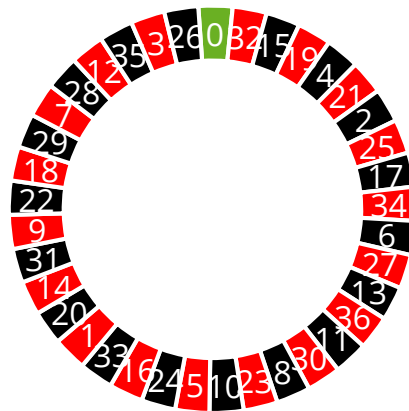
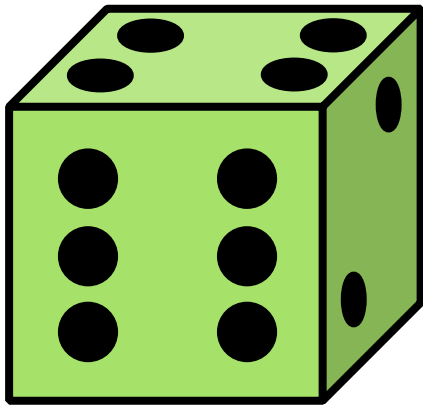
Wir wiederholen das Experiment unendlich oft (\rightarrow Gesetz der großen Zahlen).

Welchen durchschnittlichen Wert erhalten wir, wenn wir unendlich oft, eine Münze werfen, würfeln oder Roulette spielen?

Würfel: Mittelwert = 3,5

Roulette: Mittelwert = 18

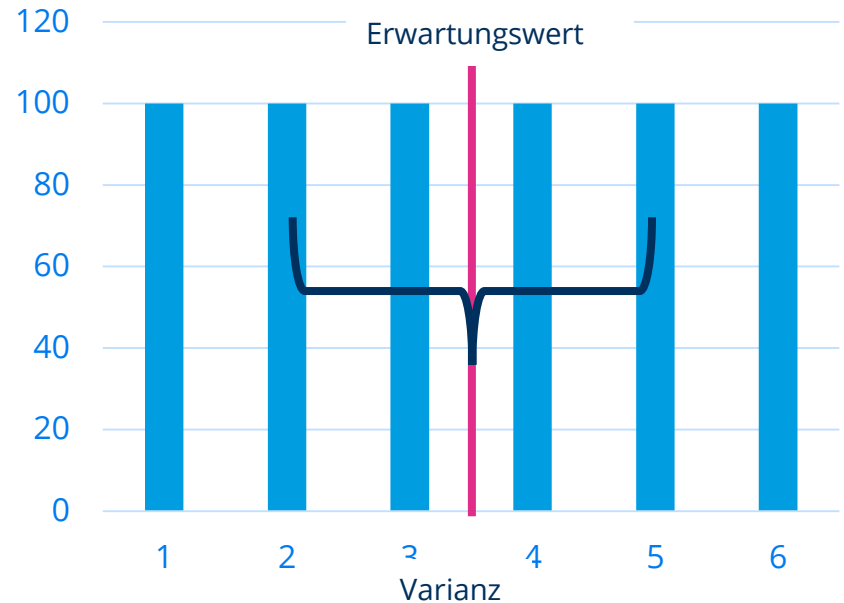
Definition für Münze notwendig: 0 = Kopf, 1 = Zahl. Mittelwert = 0,5



Varianz bei bekannten Wahrscheinlichkeiten

Erinnerung: Die Varianz zeigt an, wie sehr der Erwartungswert streut.

$$\text{Var}(X) = \sum_k (x_k - EX)^2 p_k$$



Aufgabe: Errechnen Sie die Varianz beim Würfel.

$$\text{Var}(X) = (1 - 3,5)^2 * \frac{1}{6} + (2 - 3,5)^2 * \frac{1}{6} + (3 - 3,5)^2 * \frac{1}{6} + (4 - 3,5)^2 * \frac{1}{6} + (5 - 3,5)^2 * \frac{1}{6} + (6 - 3,5)^2 * \frac{1}{6}$$

$$\text{Var}(X) = 2,91667$$

Einführung in die Stichprobentheorie

Warum werden Stichproben gezogen?

Man kann oft nicht alle

Menschen,

Einrichtungen,

Unternehmen

Länder

... analysieren / befragen.

Arten von Stichproben

Einfache Zufallsstichprobe:

Aus der Grundgesamtheit eine Zufallsziehung

Geschichtete Auswahl:

Grundgesamtheit teilen → aus jedem Teil eine Zufalls-Stichprobe

Klumpenstichprobe:

Grundgesamtheit teilen → aus jedem Klumpen alle erfassen

Mehrstufig Auswahl:

1. Realschulen in Dresden
2. Auswahl von 5 Stadtteilen
3. In den Stadtteilen zufällige Auswahl von Schulen oder Alle

Bewusste Auswahl:

Es werden nach Plan genau 50 % Frauen / Männer gewählt

Schneeballsystem:

bei schwieriger Erreichbarkeit. Ein Befragter empfiehlt weitere. noch mehr Möglichkeiten (wichtiger Fall, theoretisches Sampling, Quotens.,)

- All diesen Verfahren haben eine Regel
- Sie sind nicht willkürlich
- Die Grundgesamtheit ist immer angebar

Wieso entstehen zufällige Ergebnisse?

... weil keine Totalerhebung der Grundgesamtheit gemacht wird, sondern Stichproben gezogen werden.

Beispiel: In unserer Stichprobe erscheinen nur die roten Zahlen

Grundgesamtheit FH	Grundgesamtheit BA	Stichprobe FH	Stichprobe BA
500 €	500 €	500 €	500 €
500 €	500 €	500 €	500 €
500 €	500 €	500 €	
500 €	500 €	500 €	
1.000 €	1.000 €	1.000 €	1.000 €
1.000 €	1.000 €	1.000 €	1.000 €
1.000 €	1.000 €	1.000 €	1.000 €
1.000 €	1.000 €	1.000 €	
1.500 €	1.500 €	1.500 €	1.500 €
1.500 €	1.500 €	1.500 €	1.500 €
1.500 €	1.500 €		1.500 €
1.500 €	1.500 €		
2.000 €	2.000 €	2.000 €	2.000 €
2.000 €	2.000 €		2.000 €
2.000 €	2.000 €		2.000 €

	FH	BA
Unsere Stichprobe	1000 €	1318 €
In Wirklichkeit	1200 €	1200 €

Aus der Praxis: Stichprobengröße berechnen in der Praxis und warum wir das nicht lernen

Ich werde als Statistikberaterin öfters gefragt, wie groß eine Stichprobe sein muss. Es gäbe ja da einige Formeln.

Meine Antwort: Ja, diese Formeln können Sie in fast jedem Lehrbuch zur Statistik nachschauen. Jedoch wenn Sie in den Sozialwissenschaften und angrenzenden Wissenschaften unterwegs sind und Befragungen machen zu neuen Themen nützen Ihnen diese Formeln nichts. Denn Sie brauchen für diese Formeln immer bestimmtes Wissen über Ihre Grundgesamtheit. Das sie aber oft nicht haben. Wenn Sie das Wissen hätten, dann würden Sie die Befragung meist auch nicht machen. Und da bisse sich die Katze dann in den Schwanz.

Außerdem fragen Sie dann oft nicht nur nach einer Variable sondern nach mehreren, das hieße sie müssten genaue Infos über alle die zu erhebenden Variablen haben. Das haben aber selbst Profis nicht.

Falls Sie tatsächlich schon gute Vorinformationen über die Stichprobe haben, dann schlagen Sie die Formeln nach.

Allgemein kann man sich merken, dass Stichproben unter 30 jeden Kritiker zweifeln lassen. Pro zu analysierende Variable sollten Stichproben größer 30 sein. Wenn Sie also zwei Gruppen miteinander vergleichen wollen, brauchen Sie 60 Befragte. Da sie aber z.B. vorher nicht wissen können, ob z.B. wirklich 30 Befragte zur 1. und die anderen 30 zur 2. Gruppe gehören bräuchten Sie aus Sicherheitsgründen noch mehr Befragte ...

Ich könnte dazu noch mehr erzählen, teilweise habe ich das in meinen FAQs gemacht unter der Frage „... wie groß muss meine Stichprobe sein“: <https://tu-dresden.de/bu/wirtschaft/postgraduales/methodologicaladvisory/faq>

