Zufallsvariable
Wahrscheinlichkeitsverteilung

Peter Büchel

HSLU I

ASTAT: Block 05

Zufallsvariable: Beispiel

- Zentraler Begriff in Statistik: *Zufallsvariable*
- Pack Spielkarten (Schweiz): 36 verschiedene Karten, mit 4 Farben und je den Werten $6, 7, 8, 9, 10$, Bube, Dame, König und Ass
- Ziehen nacheinander drei Karten, nach jedem Ziehen wieder zurücklegen
- Machen dies zweimal und erhalten folgendes Resultat:
  1. 6, Dame, König
  2. 8, Bube, Ass
- Frage: Welcher der Versuche ist „besser"?
- In dieser Form schwer vergleichbar

- Eine Lösung: Einzelnen Spielkarten werden Werte zugeordnet:
  - 6, 7, 8, 9 haben Wert 0
  - 10 hat den Wert 10
  - Bube hat den Wert 2
  - Dame hat den Wert 3
  - König hat den Wert 4
  - Ass hat den Wert 11
- Jetzt sind Ziehungen miteinander vergleichbar:
  1. 6, Dame, König $\rightarrow$ $0 + 3 + 4 = 7$
  2. 8, Bube, Ass $\rightarrow$ $0 + 2 + 11 = 13$
- Zweite Ziehung mit diesen Werten besser als die erste

Zufallsvariable

- Beispiel vorher: Situation kommt in der Stochastik häufig vor
- Zufallsexperiment mit dem Grundraum Ω
- Allen Elementarereignissen von Ω wird eine Zahl zugeordnet
- Zu jedem Elementarereignis ω gehört demnach eine Zahl

$$X(\omega) = x$$

- X: *Funktion*, die jedem Elementarereignis ω den Zahl x zugeordnet
- Diese Funktion wird *Zufallsvariable* genannt

Beispiel

- Ziehen Karten aus einem Stapel Spielkarten
- Jeder Karte wird eine Zahl zugeordnet:

$$\begin{aligned} \omega &= \text{As} & &\mapsto & X(\omega) &= 11 \\ \omega &= \text{König} & &\mapsto & X(\omega) &= 4 \\ &\vdots & & & &\vdots \\ \omega &= \text{Sechs} & &\mapsto & X(\omega) &= 0 \end{aligned}$$

- Mit Zahlen $X(\omega)$ kann z.B. „Durchschnitt" der gezogenen Karten berechnet werden
- Durchschnitt von „6, Dame, König" gleich $\frac{7}{3}$
- Für Elementarereignisse „6", „Dame" und „König" ohne Zahlen macht das Wort „Durchschnitt" keinen Sinn

Beispiel

- Werfen gemeinsam einen blauen und roten Würfel
- Grundraum Ω (siehe früher): Augenzahlen der Würfel:

$$\Omega = \{11, 12, \ldots 16, 21, 22, \ldots, 26, \ldots 66\}$$

- Auf Ω sind verschiedene Zufallsvariablen definierbar
- Sei X die Zufallsvariable für die Summe der Augenzahlen:
  - Dann gilt:

$$X(16) = 7 \qquad \text{oder} \qquad X(31) = 4$$

  - Die Werte, die die Zufallsvariable annehmen kann, wird als *Wertemenge* bezeichnet:

$$\mathbb{W}_X = \{2, 3, 4 \ldots, 11, 12\}$$

- Sei Y die Augenzahl des roten Würfels:
  - Dann gilt:

$$Y(16) = 6 \qquad \text{oder} \qquad Y(31) = 1 \qquad \text{oder} \qquad Y(13) = 3$$

  - Wertemenge:

$$\mathbb{W}_Y = \{1, 2, \ldots, 6\}$$

- Sei Z gleich 0 für alle Elementarereignisse:
  - Dann gilt:

$$Z(16) = 0 \qquad \text{oder} \qquad Z(31) = 0 \qquad \text{oder} \qquad Z(13) = 0$$

  - Wertemenge:

$$\mathbb{W}_Z = \{0\}$$

- Letztes Beispiel ist eine völlig legitime Zufallsvariable
- Wie sinnvoll diese ist, ist eine andere Frage

Beispiel

- Wählen zufällig eine Person aus
- Grundraum Ω: Alle Personen dieses Planeten
- Auch hier sind viele Zufallsvariablen denkbar:
  - X: Zufallsvariable, die jeder Person das Einkommen zuordnet
  - Y: Zufallsvariable, die jeder Person die Körpergrösse zuordnet
  - Z: Zufallsvariable, die jeder Person das Alter zuordnet
- Folgende Variablen sind *keine* Zufallsvariablen:
  - Variable V ordnet jeder Person das Geschlecht zu
  - Variable W ordnet jeder Person die zugehörige Nationalität zu
- „Resultat" einer Zufallsvariable *muss* eine *Zahl* sein

Definition Zufallsvariable

- Definition:

Zufallsvariable

Eine *Zufallsvariable* X ist eine *Funktion*:

$$\begin{array}{rccc} X: & \Omega & \to & \mathbb{R} \\ & \omega & \mapsto & X(\omega) \end{array}$$

Bemerkungen

- $\mathbb{R}$: Reelle Zahlen (Punkte der Zahlengeraden, alle Dezimalbrüche)
- Stochastik X (oder $Y, Z, \ldots$): Notation für *Zufallsvariable* (*Funktion*)
- Funktionen in der Mathematik oft in der Form:

$$y = f(x)$$

- Folgende Unterscheidung ist sehr wichtig:
  - *Zufallsvariable* werden mit *Grossbuchstaben* X (oder Y, Z) bezeichnet
  - Entsprechender *Kleinbuchstabe* x (oder y, z) stellt *konkreter Wert* dar, den die Zufallsvariable annehmen kann
  - Ereignis, bei dem die Zufallsvariable X den Wert x annimmt:

$$X = x$$

- Bei einer Zufallsvariable ist nicht die Funktion X zufällig, sondern nur das Argument ω:
  - Je nach Ausgang des Zufallsexperiments ω anderer Wert $x = X(\omega)$
- x heisst auch eine *Realisierung* der Zufallsvariablen X

Beispiele

- Spielkartenbeispiel: Realisierung $X = 11$ entspricht dem Ziehen eines Asses
- Würfelbeispiel: Realisierung $X = 8$ entspricht dem Würfeln der Augensumme 8

Wahrscheinlichkeitsverteilung einer Zufallsvariablen

- Schon gesehen: Berechnung W'keit $P(E)$ eines Ereignisses E
- Entsprechend: W'keit einer allgemeinen Realisierung x einer Zufallsvariable X

Beispiel: Spielkarten

- Zufallsvariable X: Wert einer gezogenen Spielkarte
- Wie gross ist W'keit ist, dass die gezogene Karte den Wert 4 hat
- Realisierung ist $X = 4$
- Zugehörige W'keit:

$$P(X = 4)$$

- Realisation $X = 4$ entspricht dem Ziehen eines Königs
- Gesucht W'keit, dass ein König gezogen wird:

$$P(X = 4) = P(\{\omega \mid \omega = \text{ ein König}\}) = \frac{4}{36} = \frac{1}{9}$$

- Vorgehen hier ist ausführlicher ausgeführt als unbedingt notwendig, aber es lässt sich auf nicht so einfache Beispiele verallgemeinern

Allgemein

- Definition:

> Die Werte einer Zufallsvariablen X (die möglichen Realisierungen von X) treten mit gewissen W'keiten auf. Die W'keit, dass X den Wert x annimmt, berechnet sich wie folgt:
>
> $$P(X = x) = P(\{\omega \mid \; X(\omega) = x\}) = \sum_{\omega; X(\omega) = x} P(\omega)$$

- Im Spielkartenbeispiel ist $x = 4$ und ω alle möglichen Könige, deren entsprechende W'keiten aufaddiert werden

Wahrscheinlichkeitsverteilung

- Vorher: W'keit einer Realisierung berechnet
- Jetzt: W'keiten *aller* Realisierungen berechnen
- Definition:

> **Wahrscheinlichkeitsverteilung**
>
> Für *jede* Realisierung einer Zufallsvariable wird die zugehörige W'keit berechnet $\quad \rightarrow \quad$ *W'keitsverteilung* dieser Zufallsvariablen

Beispiel

- Zufallsvariable X: Wert einer gezogenen Jasskarte
- Schon berechnet: W'keit

$$P(X=4)=\frac{1}{9}$$

- W'keit $P(X=0)$ mit Laplace-W'keit:

$$P(X=0)=\frac{16}{36}=\frac{4}{9}$$

- W'keit $P(X=2)$: Ziehen eines Buben:

$$P(X=2)=\frac{4}{36}=\frac{1}{9}$$

- W'keitsverteilung von X als Tabelle:

x	0	2	3	4	10	11
$P(X=x)$	4/9	1/9	1/9	1/9	1/9	1/9

- Werte für $P(X=1)$ oder $P(X=178)$ sind in Tabelle nicht aufgeführt
- Grund: Diese Werte können nicht gezogen werden
- Trotzdem eine W'keit zuordnen, nämlich die Zahl 0:

$$P(X=1)=0 \quad \text{oder} \quad P(X=178)=0$$

- Addieren alle Werte der W'keitsverteilung $\rightarrow$ 1
- Eine Realisierung *muss* gezogen werden
- Es gilt:

$$P(X=0)+P(X=2)+P(X=3)+P(X=4)+P(X=10)+P(X=11)=1$$

Beispiel: Augensumme zweier Würfel

- W'keitsverteilung für die Zufallsvariable X:

x	2	3	4	5	6	7	8	9	10	11	12
$P(X=x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

- Brüche sind hier ungekürzt: Man sieht besser, dass die Summe der W'keiten 1 geben muss

Allgemein

- Es gilt:

Wahrscheinlichkeitsverteilung

„Liste" von $P(X=x)$ für alle *möglichen* Werte $x_1, x_2, \ldots, x_n$ heisst *diskrete W'keitsverteilung* der diskreten Zufallsvariablen X

Es gilt immer:

$$P(X=x_1)+P(X=x_2)+\ldots+P(X=x_n)=1$$

Mit Summenzeichen:

$$\sum_{\text{alle möglichen } x} P(X=x)=1$$

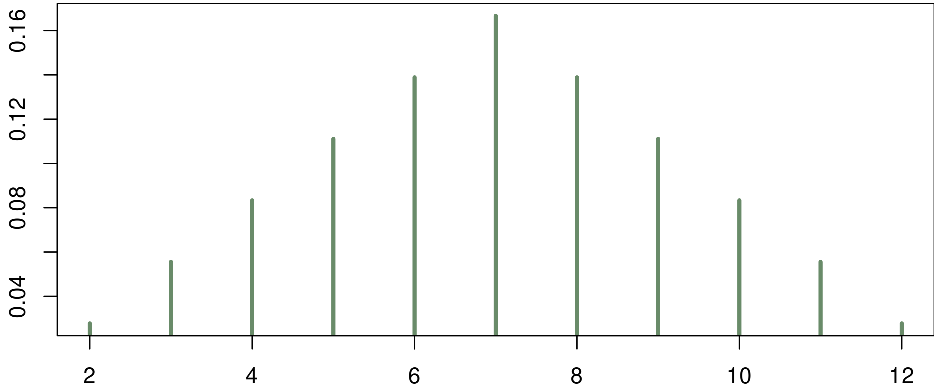
Alle W'keiten einer W'keitsverteilung ergeben 1

Beispiel: Augensumme zweier Würfel

- Schon gesehen:

x	2	3	4	5	6	7	8	9	10	11	12
$P(X=x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

- Skizze:



Beispiel: Augensumme zweier Würfel

- X: Zufallsvariable der geworfenen Augensumme
- Wie gross ist die W'keit, genau die Augensumme 6 zu würfeln?
  - Gesucht: $P(X=6)$:

$$P(X=6)=\frac{5}{36}$$

- Wie gross ist die W'keit, die Augensumme 6 oder 8 zu würfeln?
  - Gesucht: $P(X=6)+P(X=8)$:

$$P(X=6)+P(X=8)=\frac{5}{36}+\frac{5}{36}=\frac{10}{36}=\frac{5}{18}$$

- Wie gross ist die W'keit, höchstens die Augensumme 3 zu würfeln?
  - Gesucht:

$$P(X\leq 3)=P(X=2)+P(X=3)$$

  - Also:

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

- Wie gross ist die W'keit, mindestens die Augensumme 3 zu würfeln?
  - Gesucht:

$$P(X\geq 3)=P(X=3)+\ldots+P(X=12)$$

  - Einfacher:

$$P(X\geq 3)=1-P(X=2)$$

  - Also:

$$1-P(X=2)=1-\frac{1}{36}=\frac{35}{36}$$

- Wie gross ist die W'keit, eine Augensumme von 3 bis 5 zu würfeln?
  - Gesucht:

$$P(3 \leq X \leq 5) = P(X = 3) + P(X = 4) + P(X = 5)$$

  - Also:

$$P(3 \leq X \leq 5) = \frac{2}{36} + \frac{3}{36} + \frac{4}{36} = \frac{9}{36} = \frac{1}{4}$$

Kennzahlen einer Verteilung

- Beliebige (diskrete) Verteilung: Vereinfachend durch 2 Kennzahlen zusammengefasst:
  - *Erwartungswert* $\mathrm{E}(X)$: Mittlere Lage der Verteilung
  - *Standardabweichung* $\sigma(X)$: Streuung der Verteilung

- Diskrete Zufallsvariable X: Mögliche Werte $x_1, x_2, \ldots, x_n$
- Definition:

Erwartungswert und Standardabweichung

- *Erwartungswert*:

$$\begin{aligned}\mathrm{E}(X) &= x_1 \cdot P(X = x_1) + x_2 \cdot P(X = x_2) + \ldots + x_n \cdot P(X = x_n) \\ &= \sum_{\text{alle möglichen } x} xP(X = x)\end{aligned}$$

- *Varianz* und *Standardabweichung*:

$$\begin{aligned}\mathrm{Var}(X) &= (x_1 - \mathrm{E}(X))^2 \cdot P(X = x_1) + \ldots + (x_n - \mathrm{E}(X))^2 \cdot P(X = x_n) \\ &= \sum_{\text{alle möglichen } x} (x - \mathrm{E}(X))^2 P(X = x)\end{aligned}$$

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

Beispiel

- Wurf eines fairen Würfels: Alle 6 möglichen Zahlen gleiche W'keit geworfen zu werden
- Zufallsvariable X sei die geworfene Zahl
- Erwartungswert $\mathrm{E}(X)$:

$$\begin{aligned}\mathrm{E}(X) &= x_1 \cdot P(X = x_1) + x_2 \cdot P(X = x_2) + \ldots + x_6 \cdot P(X = x_6) \\ &= 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} \\ &= \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) \\ &= 3.5\end{aligned}$$

- Dieser Erwartungswert 3.5 nicht anderes als der Durchschnitt der Augenzahlen

- Wie lässt sich dieser Wert nun interpretieren?
- Würfeln fairen Würfel 100mal: Durchschnitt meistens *nicht* exakt 3.5
- Aber: Durchschnitt in der *Nähe* von 3.5
- Annäherung sollte immer besser werden, je mehr wir würfeln
- 100 Milliarden mal würfeln: Durchschnitt nicht *exakt* 3.5, aber sehr nahe dran
- Interpretation: Für sehr viele Würfe liegt Durchschnitt sehr nahe bei Erwartungswert

- Standardabweichung mit R berechnen:

```
x <- 1 : 6
p <- 1 / 6

E_X <- sum(x * p)

var_X <- sum((x - E_X)^2 * p)
sd_X <- sqrt(var_X)

sd_X
## [1] 1.707825
```

- D.h.: Abweichung „durchschnittlich" 1.7 von 3.5

Beispiel: Spielkarten

- Verteilung:

x	0	2	3	4	10	11
$P(X = x)$	4/9	1/9	1/9	1/9	1/9	1/9

- Ziehen aus dem Stapel eine Karte
- Welches ist der durchschnittliche Wert der Karte, die gezogen wird?
- Berechnen Erwartungswert $\mathrm{E}(X)$:

$$\mathrm{E}(X) = 0 \cdot \frac{4}{9} + 2 \cdot \frac{1}{9} + 3 \cdot \frac{1}{9} + 4 \cdot \frac{1}{9} + 10 \cdot \frac{1}{9} + 11 \cdot \frac{1}{9} = 3.33$$

- Zahlen untereinander in Tabelle werden multipliziert und dann addiert

- R:

```
x <- c(0, 2, 3, 4, 10, 11)
p <- 1 / 9 * c(4, 1, 1, 1, 1, 1)

E_X <- sum(x * p)
E_X
## [1] 3.333333
```

- Durchschnittlicher Wert, der zu erwarten ist, wenn Karte sehr oft gezogen und wieder in Stapel zurücklegt wird
- Viele Karten mit Wert 0 $\quad\rightarrow\quad$ Erwartungswert eher tief

- Varianz und die Standardabweichung:

$$\begin{aligned}\mathrm{Var}(X) = &(0-3.33)^2 \cdot \frac{4}{9} + (2-3.33)^2 \cdot \frac{1}{9} + (3-3.33)^2 \cdot \frac{1}{9} \\ &+ (4-3.33)^2 \cdot \frac{1}{9} + (10-3.33)^2 \cdot \frac{1}{9} + (11-3.33)^2 \cdot \frac{1}{9} \\ &= 16.67\end{aligned}$$

und

$$\sigma(X) = \sqrt{16.67} = 4.08$$

- R:

```
var_X <- sum((x - E_X)^2 * p)
sd_X <- sqrt(var_X)

sd_X
## [1] 4.082483
```

- „Mittlere" Abweichung 4.1: Eher gross wegen Werten 10, 11

Bemerkungen

- Erwartungswert einer diskreten Zufallsvariable: Gewichtetes arithmetisches Mittel von allen möglichen Werten, wobei die Werte mit ihrer W'keit gewichtet werden
- Erwartungswert: Oft auch mit μ_X bezeichnet:
  - Index X wird oft weglassen, falls Zufallsvariable klar
- W'keiten für alle Werte $x_1, x_2, \ldots, x_n$ gleich: Erwartungswert gerade arithmetisches Mittel der Werte
- Varianz das Quadrat der Abweichung eines Wertes der Zufallsvariable vom Erwartungswert mit der W'keit des Wertes gewichtet
- Standardabweichung hat dieselbe Einheit wie X: Einheit der Varianz deren Quadrat ist:
  - Z. B. X in Metern (m) gemessen $\rightarrow$ Var(X) in Quadratmeter (m^2)
  - $\sigma(X)$ wiederum die Dimension Meter (m)

Unterschied empirischer und theoretischer Kennzahlen

- Gesehen: Mittelwert:

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- *Empirische* Varianz:

$$\mathrm{Var}(x) = \frac{(x_1-\overline{x}_n)^2 + (x_2-\overline{x}_n)^2 + \cdots + (x_n-\overline{x}_n)^2}{n-1} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x}_n)^2$$

- Empirische Standardabweichung s_x:

$$s_x = \sqrt{\mathrm{Var}(x)} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x}_n)^2}$$

- Wie hängen diese Definition und die Definition für die Standardabweichung for eine Zufallsvariable X zusammen?
- Sehr genau unterscheiden:
  - Arithmetische Mittelwert $\overline{x}$: Aus *konkreten* Daten berechnet: Aus Messwerten $x_1, \ldots, x_n$ wird nach der Formel oben $\overline{x}_n$ berechnet
  - Erwartungswert E(X): *Theoretischer Wert*, der sich aus dem Modell der W'keitsverteilung ergibt
- Hoffnung: Arithmetisches Mittel $\overline{x}$ nähert für immer mehr Versuche den theoretischen Wert $\mu_X = \mathrm{E}(x)$ immer besser an, sofern die Daten der W'keitsverteilung von X folgen
- Ist dies nicht der Fall, so stimmt etwas am Modell nicht (z.B. alle Seiten beim Würfel haben gleiche W'keit geworfen zu werden)

Beispiel: Fairer Würfel

- Jede Seite die gleiche W'keit geworfen zu werden
- Gerechtfertigte Annahme aus *Symmetriegründen*: Alle Seiten gleich
- Erwartungswert:

$$E(X) = 3.5$$

- Werfen idealen, fairen Würfel $n = 10$ mal
- Obwohl der Würfel fair ist, wird der Durchschnitt nie genau 3.5 sein
- Idealer fairen Würfel: R:

```
x <- sample(1:6, size = 10, replace = T)
x
##  [1] 3 3 3 4 3 6 5 2 3 6
mean(x)
## [1] 3.8
```


- Durchschnitt 3.8: Einiges neben den zu *erwartenden* 3.5
- Würfel nochmals 10mal werfen: Normalerweise ein anderes Resultat
- Simulation mit R
- Simulieren 10mal 10 Würfe: Berechnen entsprechende Durchschnitte

```
for (i in 1:10)
{
x <- sample(1:6, size = 10, replace = T)
cat(mean(x)," ")
}
## 3.3  3.5  4.2  3.4  2.9  3  3.9  2.9  4.1  2.7
```

- Durchschnitte gehen von 2.7 bis 4.2
- Zwar in der „Nähe" von 3.5 aber nicht sehr genau

- Würfeln $n = 100$ Würfel:

```
x <- sample(1:6, size = 100, replace = T)
x
##   [1] 5 6 6 1 5 1 4 5 1 2 3 1 3 6 2 3 1 6 1 4 3 6 1 6 5
##  [26] 6 6 3 1 5 5 6 6 2 2 3 4 3 1 1 5 1 2 4 5 6 5 4 2 5
##  [51] 6 5 2 6 4 4 4 4 1 2 2 6 6 3 5 3 6 5 5 1 5 6 1 2 1
##  [76] 5 4 1 6 1 5 3 1 2 6 5 3 1 4 1 2 1 4 4 1 4 6 1 5 6
mean(x)
## [1] 3.57
```

- Mittelwert 3.57: Schon relativ nahe beim theoretischen Wert von 3.5

- Machen dies 10mal:

```
## 3.39  3.43  3.55  3.5  3.48  3.61  3.46  3.64  3.28  3.41
```

- Durchschnitte: Zwischen 3.28 und 3.64 ($\approx \pm 0.2$ vom Erwartungswert)
- Dasselbe für $n = 1000$ Würfe (auf drei Nachkommastellen):

```
## 3.475  3.49  3.435  3.437  3.407  3.479  3.567  3.474  3.498  3.565
```

- Durchschnitte: Zwischen 3.407 und 3.565 ($\approx \pm 0.1$ vom EW)
- Für $n = 1\,000\,000$: Durchschnitte auf 3 Nachkommastellen gerundet:

```
## 3.498  3.497  3.501  3.496  3.501  3.5  3.497  3.5  3.503  3.499
```

- Durchschnitte: Zwischen 3.498 und 3.503 ($\approx \pm 0.005$ vom EW)
- Durchschnitt: Für immer grössere n immer näher bei 3.5

- Annahme in Beispiel: Fairer Würfel gibt
- Dies ist *nicht* realistisch
- *Kein* realer Würfel ist wirklich symmetrisch, das heisst nicht alle Wurfzahlen sind gleich wahrscheinlich
- Können Würfel zu konstruieren, der *sehr* fair ist und alle Wurfzahlen mit einer W'keit von *fast* einem Sechstel vorkommen
- Aber *exakt* geht das nicht

- Analoges gilt für die Standardabweichung:
  - Empirische Standardabweichung: s_X aus *konkreten* Daten berechnet: Aus Messwerte $x_1, \ldots, x_n$ wird nach Formel oben s_X berechnet
  - Standardabweichung σ_X: Theoretischer Wert, der sich aus Modell der W'keitsverteilung ergibt
- Hoffnung: Empirische Standardabweichung s_X nähert für immer mehr Versuche den theoretischen Wert σ_X immer besser an, falls Daten der W'keitsverteilung von X folgen
- Ist dies nicht der Fall, so stimmt etwas am Modell nicht (z.B. alle Seiten beim Würfel haben gleiche W'keit geworfen zu werden)