

# Tutorium - Foliensatz 10 aus Statistik 1

Thomas Haase

09.05.23

Kurze Wiederholung der Vorlesung

Stichprobe	Wahrscheinlichkeitsverteilung	Grundgesamtheit
Kennwertwerte	Parameter	Parameter
Mittelwert $\bar{x}$	Erwartungswert $\mu$	Mittelwert $\mu$
Standardabweichung $s$	Standardabweichung $\sigma$	Standardabweichung $\sigma$
Varianz $s^2$	Varianz $\sigma^2$	Varianz $\sigma^2$

# Wahrscheinlichkeit

**Wahrscheinlichkeit:** Maß für Chance, dass ein bestimmtes Ereignis eintritt

$$P(A) = \frac{\textit{positive outcomes}}{\textit{alle möglichen Ergebnisse}}$$

- Wahrscheinlichkeit = “*relative Häufigkeit eines Ereignis*”

# Wahrscheinlichkeit

**Wahrscheinlichkeit:** Maß für Chance, dass ein bestimmtes Ereignis eintritt

$$P(A) = \frac{\textit{positive outcomes}}{\textit{alle möglichen Ergebnisse}}$$

- ▶ Wahrscheinlichkeit = “*relative Häufigkeit eines Ereignis*”
- ▶ empirische Wahrscheinlichkeit = Schätzwert für Wahrscheinlichkeit

# Standardnormalverteilung

***Definition:***

(1)  $\mu = 0$

# Standardnormalverteilung

***Definition:***

(1)  $\mu = 0$

(2)  $\sigma = 1$

# Standardnormalverteilung

***Definition:***

(1)  $\mu = 0$

(2)  $\sigma = 1$

► ***Was passiert genau?***



# Z-Werte

- ▶ Was sagt der Z-Wert eines Rohwertes aus und wie wird er berechnet?

# Z-Werte

- ▶ Was sagt der Z-Wert eines Rohwertes aus und wie wird er berechnet?
- ▶  $z_i = \frac{x_i - \bar{x}}{s}$

# Z-Werte

- ▶ Was sagt der Z-Wert eines Rohwertes aus und wie wird er berechnet?
- ▶  $z_i = \frac{x_i - \bar{x}}{s}$
- ▶ Der Rohwert wird nichtmehr als absolutes Ergebnis (*“Lara hat Note 2 erreicht”*) sondern relativ zum Rest der Verteilung angegeben (*“Lara hat eine Note erreicht, die 2 Standardabweichungen über dem Durchschnitt der Klasse liegt”*)

# Z-Werte

- ▶ Was sagt der Z-Wert eines Rohwertes aus und wie wird er berechnet?
- ▶  $z_i = \frac{x_i - \bar{x}}{s}$
- ▶ Der Rohwert wird nichtmehr als absolutes Ergebnis (*“Lara hat Note 2 erreicht”*) sondern relativ zum Rest der Verteilung angegeben (*“Lara hat eine Note erreicht, die 2 Standardabweichungen über dem Durchschnitt der Klasse liegt”*)
- ▶ **Z-Transformation:** Z Werte für alle Rohwerte einer Verteilung berechnen

# Übung

berechnet die Z-Werte für folgende Roh-Werte (Es handelt sich um Punkte einer Klausur) und plottet

5,7,2,12,8

```
Werte <- c(5,4,2,12,9)
# Mittelwert
mean(Werte)
```

```
## [1] 6.4
```

```
# standard deviation
sd(Werte)
```

```
## [1] 4.037326
```

$$z_i = \frac{x_i - \bar{x}}{s}$$

# Lösung

```
Werte <- c(5,4,2,12,9)
scale(Werte)
```

```
##           [,1]
## [1,] -0.3467642
## [2,] -0.5944529
## [3,] -1.0898303
## [4,]  1.3870567
## [5,]  0.6439906
## attr(,"scaled:center")
## [1] 6.4
## attr(,"scaled:scale")
## [1] 4.037326
```

## Z-Transformation anschaulich gemacht

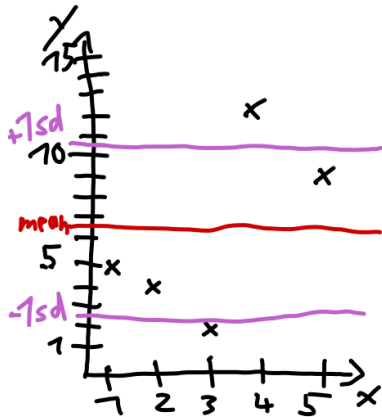
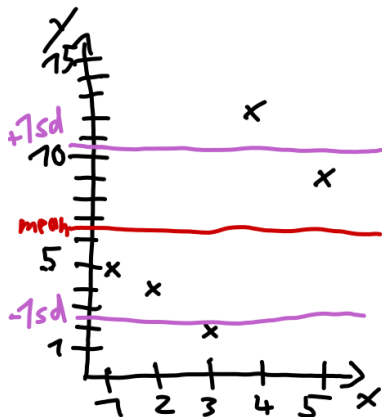


Figure 1: Rohwerte

5, 4, 2, 12, 9



Häufigkeit

$\leftarrow -1sd \quad +1sd \rightarrow$   
mean

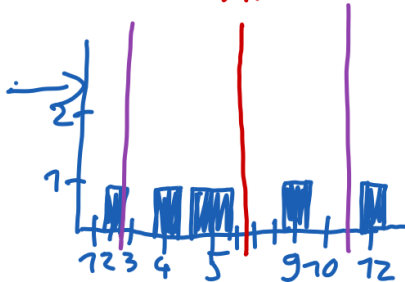
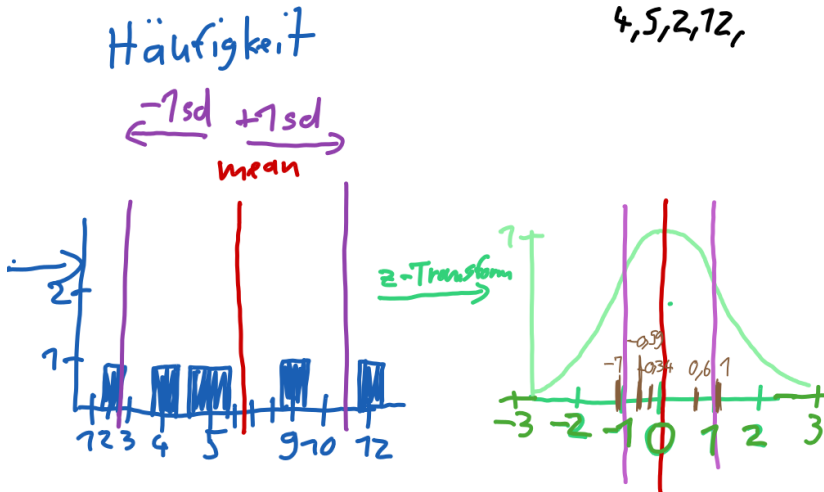


Figure 2: Häufigkeit



Das Beispiel besteht natürlich aus zu wenigen Werten, aber wenn man viele Werte einer Normalverteilten Variablen nimmt sie die typische Glockenförmige Verteilung an.

- Wenn wir mehr Klausurteilnehmer hätten würde die Kurve vermutlich Glockenförmig aussehen,



# Mit der Standardnormalverteilung arbeiten

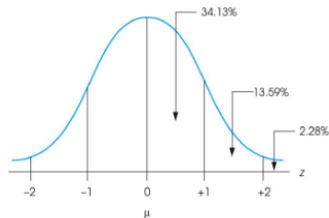
Was bringt uns diese Transformation? Selbstlernmodul - Zentrales Grenzwerttheorem - Wenn man irgendeine Variable aus der Bevölkerung unendlich oft misst, die unabhängig und identisch Verteilt ist und diese Z-Transformiert erhalten man **IMMER** eine Standardnormalverteilung.

- ▶ vorhin haben wir gesehen, dass die Standardnormalverteilung eine angepasste Häufigkeitsverteilung ist > - Die Fläche unter der Dichtefunktion beschreibt die Wahrscheinlichkeit, dass der jeweilige Wert auftritt, denn der Wert ist ja auch sooft in der Grundgesamtheit insgesamt verteilt. . .
- ▶ Normalverteilung:  $\mathcal{N}(\mu, \sigma^2)$
- ▶ Standardnormalverteilung:  $\mathcal{N}(0, 1)$

# Aufgabe

$N = 50.000$  Personen mit Körpergröße  $\mathcal{N}(175, 5)$  Wie groß sind 95% aller Personen?

# Lösung



Intervall	Flächenanteil
$[\mu - 1 \cdot \sigma; \mu + 1 \cdot \sigma]$	68.3%
$[\mu - 1.96 \cdot \sigma; \mu + 1.96 \cdot \sigma]$	95%
$[\mu - 2 \cdot \sigma; \mu + 2 \cdot \sigma]$	95.4%
$[\mu - 2.58 \cdot \sigma; \mu + 2.58 \cdot \sigma]$	99.0%
$[\mu - 3 \cdot \sigma; \mu + 3 \cdot \sigma]$	99.7%

-  $175 - 1.96 = 165,2$  -  $175 + 1.96 = 184,5$  - 95% aller Personen sind zwischen 165 und 184cm groß

## Tolle Videos zum üben

Das Video, welches wir im Tutorium geschaut hatten:

<https://youtu.be/uwhV0TAPmWc>

hier noch ein kurzes Video um das Verständnis zu schärfen:

<https://www.youtube.com/watch?v=2fzYE-Emar0>