Tutorium

Wahrscheinlichketstheorie und Frequentistische Inferenz

BSc Psychologie WiSe 2022/23

Belinda Fleischmann

(8) Transformationen der Normalverteilung

Selbstkontrollfragen

1. Erläutern Sie den Begriff der Transformation einer Zufallsvariable.
2. Erläutern Sie die zentrale Idee der Transformationstheoreme.
3. Erläutern Sie die Bedeutung der Standardtransformationen für die Statistik.
4. Geben Sie das Summentransformationstheorem wieder.
5. Geben Sie das Mittelwerttransformationstheorem wieder.
6. Geben Sie das Z-Transformationstheorem wieder.
7. Geben Sie das χ^2-Transformationstheorem wieder.
8. Beschreiben Sie die WDF der t-Verteilung in Abhängigkeit ihrer Freiheitsgrade.
9. Geben Sie das T-Transformationstheorem wieder.
10. Geben Sie das F-Transformationstheorem wieder.

Wdhl.: Realisierung von Zufallsvariablen

Der einzelne Wert, den eine Zufallsvariable bei jedem Durchgang eines Zufallsvorgangs annimmt, heißt eine **Realisierung der Zufallsvariable**. Mithilfe eines Computers lassen sich Zufallsexperimente simulieren und Realisierungen von Zufallsvariablen erhalten.

Realisierungen von normalverteilten Zufallsvariablen erhält man in R mit `rnorm()`, wobei die Syntax für Realisierungen von n unabhängig und identisch verteilten Zufallsvariablen $\xi_i \sim N(\mu, \sigma^2), i = 1, \ldots, n$ durch `rnorm(n,mu,sigma)` gegeben ist.

```
rnorm(1,0,1)                    # \xi_i \sim N(0,1)
```

```
> [1] -1.4
```

```
rnorm(1,10,1)                   # \xi_i \sim N(10,1)
```

```
> [1] 10.3
```

```
rnorm(3,5,sqrt(2))              # \xi_i \sim N(5,2), i = 1,2,3 (u.i.v.)
```

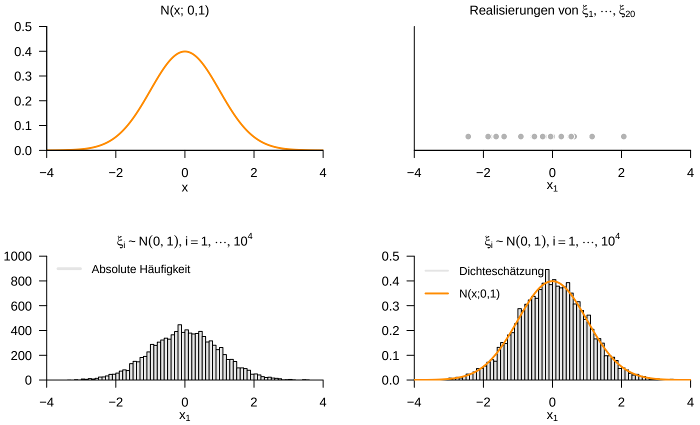
```
> [1] 1.55 4.99 5.88
```

```
rnorm(1e1,5,sqrt(2))            # \xi_i \sim N(5,2), i = 1,...,10 (u.i.v.)
```

```
>  [1] 6.62 2.42 4.65 4.65 4.60 4.22 5.89 7.92 2.69 5.72
```

Realisierungen von Zufallsvariablen

Die empirische Verteilung unabhängig und identisch simulierter Zufallsvariablenrealisationen entspricht der Verteilung der Zufallsvariable. Die empirische Verteilung stellt man mit Histogrammen (Häufigkeitsverteilungen) oder histogrammbasierten Dichteschätzern dar.



1. Erläutern Sie den Begriff der Transformation einer Zufallsvariable.

- *Transformation* einer ZV meint die Anwendung einer Funktion auf Zufallsvariablen sowie die arithmetische Verknüpfung mehrerer Zufallsvariablen.
- Die zentrale Fragestellung dabei ist: "Wenn die Zufallsvariable ξ normalverteilt ist, wie ist dann eine Zufallsvariable υ, die sich durch Transformation von ξ ergibt, verteilt?"
- Die Anwendung einer (deterministischen) Funktionauf eine zufällige Größe ergibt im Allgemeinen wieder eine zufällige Größe.

Formal:

Theorem (Transformation eines Zufallsvektors)

$\xi : \Omega \to \mathcal{X}$ sei ein Zufallsvektor und $f : \mathcal{X} \to \mathbb{R}^m$ sei eine multivariate vektorwertige Funktion. Dann ist

$$\upsilon : \Omega \to \mathbb{R}, \omega \mapsto \upsilon(\omega) := (f \circ \xi)(\omega) := f(\xi(\omega)) \tag{1}$$

ein Zufallsvektor.

Bemerkungen

- Ist ξ eine Zufallsvariable, dann gilt $\mathcal{X} := \mathbb{R}$ und $f : \mathbb{R} \to \mathbb{R}$. Wir schreiben in diesem Fall in der Regel einfach $\upsilon := f(\xi)$ und nennen υ die *transformierte Zufallsvariable*.

Beispiel für Transformation einer normalverteilen ZV

Es sei ξ eine normalverteilte ZV $\xi \sim N(0,1)$ und $\upsilon := \xi^2$ ihre Transformation.

Aufgrund der Quadrierung können wir bereits analytisch feststellen, dass υ nicht die gleiche Verteilung haben kann wie xi.

Wir können das auch mithilfe einer Simulation zeigen.

```
# Simulationsspezifikation
n       = 1e4                          # Anzahl von u.i.v Realisierungen (ZVen)
mu      = 0                            # Erwartungswertparameter von \xi
sigsqr  = 1                            # Varianzparameter von \xi

# Quadrieren einer Zufallsvariable
x       = rnorm(n, mu, sqrt(sigsqr))   # Realisierungen x_i, i = 1,....,n von \xi
y       = x - 4.2                      # Realisierungen y_i = x_i^2 von \ups

# Ausgabe der ersten acht Werte
print(x[1:8], digits = 2)
```

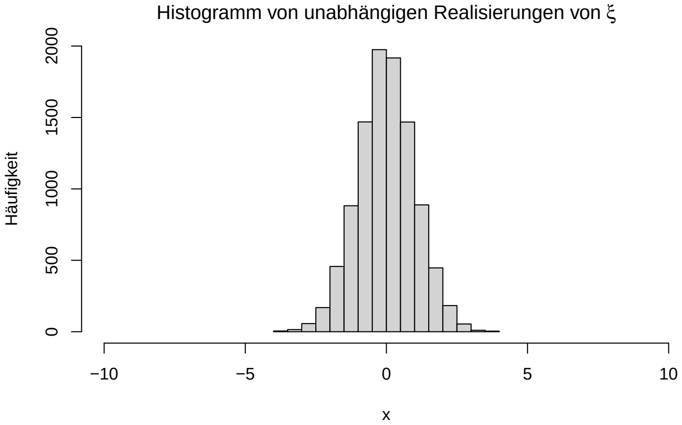
```
> [1] -1.863 -0.522 -0.053  0.543 -0.914  0.468  0.363 -1.305
```

```
print(y[1:8], digits = 2)
```

```
> [1] -6.1 -4.7 -4.3 -3.7 -5.1 -3.7 -3.8 -5.5
```

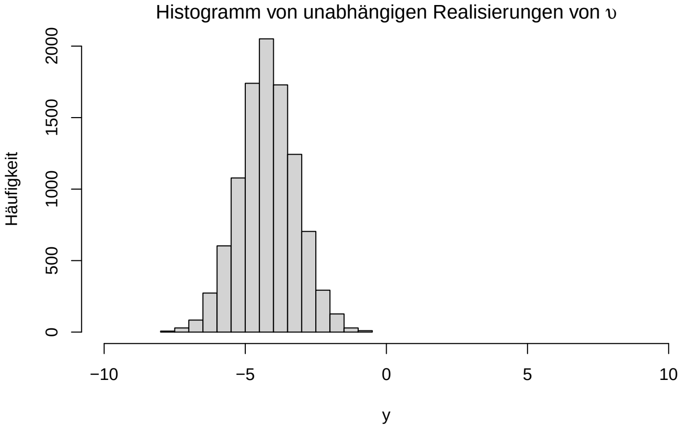
Beispiel für Transformation einer normalverteilen ZV (fortgeführt)

Simulation der Transformation normalverteilter Zufallsvariablen in R



Beispiel für Transformation einer normalverteilen ZV (fortgeführt)

Simulation der Transformation normalverteilter Zufallsvariablen in R



2. Erläutern Sie die zentrale Idee der Transformationstheoreme.

- Mit den **Transformationstheoremen** haben wir einige generelle Werkzeuge zum Berechnen der WDFen von transformierten Zufallsvariablen.
- Diese Werkzeuge sind von allgemeiner Form:
  - "Wenn ξ eine Zufallsvariable mit WDF p_ξ ist,
  - und $\upsilon := f(\xi)$ die durch f transformierte Zufallsvariable ist,
  - dann gilt für die WDF von υ die folgende Formel: $p_\upsilon := \{\text{Formel}\}$".

3. Erläutern Sie die Bedeutung der Standardtransformationen für die Statistik.

Standardtransformationen spielen inbesondere in der frequentistischen Inferenz zentrale Rollen, weil

(1) die Zentralen Grenzwertsätze die Annahme additiv unabhängig normalverteilter Störvariablen, und damit normalverteilter Daten, rechtfertigt,
(2) wie wir in der nächsten Vorlesungseinheit sehen werden, es sich bei Schätzern und Statistiken um Transformationen von Zufallsvariablen handelt, und
(3) Konfidenzintervalle und Hypothesentests durch die Verteilungen ihrer jeweiligen Statistiken charakterisiert und gerechtfertigt sind.

Diese Aussagen sind von der allgemeinen Form:

- "Wenn $\xi_i, i = 1, ..., n$ unabhängig und identisch normalverteilte Zufallsvariablen sind und
- $\upsilon := f(\xi_1, ..., \xi_n)$ eine Transformation dieser Zufallsvariablen ist,
- dann ist die WDF von υ durch die Formel $p_\upsilon :=$ {Formel} gegeben,
- und man nennt die Verteilung von υ *Verteilungsname*".

4. Geben Sie das Summentransformationstheorem wieder.

Theorem (Summe unabhängig normalverteilter Zufallsvariablen)

Für $i = 1, \ldots, n$ *seien* $\xi_i \sim N(\mu_i, \sigma_i^2)$ *unabhängige normalverteilte Zufallsvariablen. Dann gilt für die Summe* $\upsilon := \sum_{i=1}^{n} \xi_i$ *, dass*

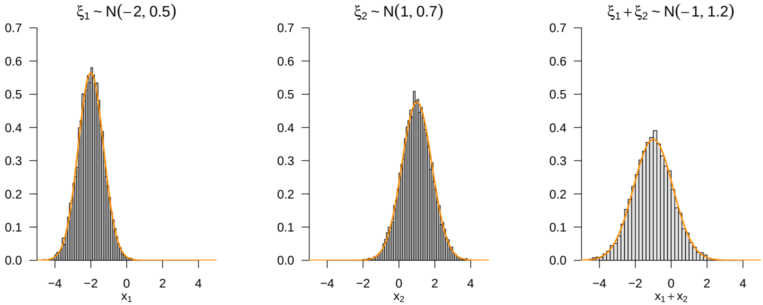
$$\upsilon \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right) \tag{2}$$

Für unabhängige und identisch normalverteilte Zufallsvariablen $\xi_i \sim N(\mu, \sigma^2)$ *gilt folglich*

$$\upsilon \sim N(n\mu, n\sigma^2). \tag{3}$$

Anwendungsfälle: Mittelwerttransformation, Generalisierung der zentralen Grenzwertsätze

Beispiel



5. Geben Sie das Mittelwerttransformationstheorem wieder.

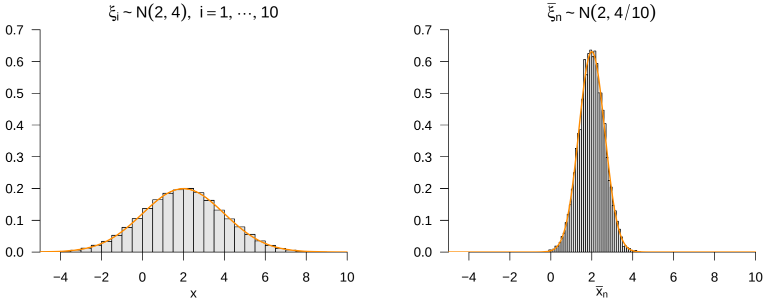
Theorem (Stichprobenmittel von u.i.v. normalverteilten Zufallsvariablen)

Für $i = 1, \ldots, n$ seien $\xi_i \sim N(\mu, \sigma^2)$ unabhängig und identisch normalverteilte Zufallsvariablen. Dann gilt für das Stichprobenmittel $\bar{\xi}_n := \frac{1}{n}\sum_{i=1}^{n} \xi_i$, dass

$$\bar{\xi}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right). \tag{4}$$

Anwendungsfälle: Analyse von Erwartungswertschätzern, Generalisierung der zentralen Grenzwertsätze

Beispiel



Anmerkung:

- die linke Abbildung zeigt *eine* der 10 ZVen $\xi_1, \ldots, \xi_1 0$

6. Geben Sie das Z-Transformationstheorem wieder.

Theorem (Z-Transformation)

Es sei $y \sim N(\mu, \sigma^2)$ eine normalverteilte Zufallsvariable. Dann ist die Zufallsvariable

$$Z := \frac{y - \mu}{\sigma} \tag{5}$$

eine z-verteilte Zufallsvariable, es gilt also $Z \sim N(0, 1)$.

- y bezeichnet hier eine normalverteilte Zufallsvariable. Werte, die diese Zufallsvariable annehmen kann, bezeichnen wir in der Folge mit $\tilde{y}$.
- Anwendungsfälle: Die Z-Konfidenzintervallstatistik und die Z-Teststatistik

Z-Zufallsvariable und ihre Verteilung

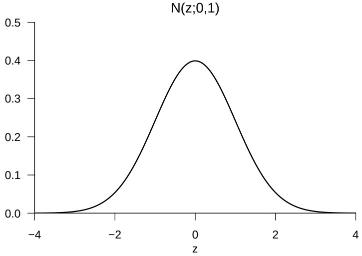
Definition (z-Zufallsvariable)

Z sei eine Zufallsvariable mit Ergebnisraum $\mathbb{R}$ und WDF

$$p : \mathbb{R} \to \mathbb{R}_{>0}, z \mapsto p(z) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z^2\right). \tag{6}$$

Dann sagen wir, dass Z einer *z-Verteilung (oder Standardnormalverteilung)* unterliegt und nennen Z eine *z-Zufallsvariable*. Wir kürzen dies mit $Z \sim N(0,1)$ ab. Die WDF einer z-Zufallsvariable bezeichnen wir mit

$$N(z; 0, 1) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z^2\right). \tag{7}$$



7. Geben Sie das χ^2-Transformationstheorem wieder.

Theorem (χ^2-Transformation)

$Z_1, ..., Z_n \sim N(0, 1)$ seien unabhängig und identisch verteilte z-Zufallsvariablen. Dann ist die Zufallsvariable

$$U := \sum_{i=1}^{n} Z_i^2 \tag{8}$$

eine χ^2-verteilte Zufallsvariable mit n Freiheitsgraden, es gilt also $U \sim \chi^2(n)$. Insbesondere gilt für $Z \sim N(0, 1)$ und $U := Z^2$, dass $U \sim \chi^2(1)$.

Anwendungsfälle: U-Konfidenzintervallstatistik, t- und f-Zufallsvariablen

χ^2-Zufallsvariable und ihre Verteilung

Definition (χ^2-Zufallsvariable)

U sei eine Zufallsvariable mit Ergebnisraum $\mathbb{R}_{>0}$ und WDF

$$p : \mathbb{R}_{>0} \to \mathbb{R}_{>0}, u \mapsto p(u) := \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} u^{\frac{n}{2}-1} \exp\left(-\frac{1}{2}u\right), \tag{9}$$

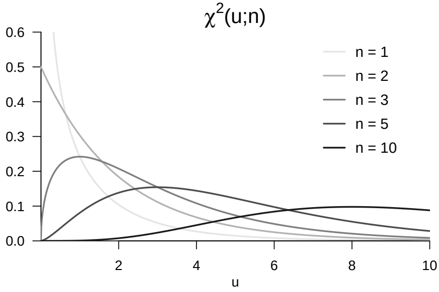
wobei Γ die Gammafunktion bezeichne. Dann sagen wir, dass U einer χ^2-Verteilung mit n Freiheitsgraden unterliegt und nennen U eine χ^2-Zufallsvariable mit n Freiheitsgraden. Wir kürzen dies mit $U \sim \chi^2(n)$ ab. Die WDF einer χ^2-Zufallsvariable bezeichnen wir mit

$$\chi^2(u; n) := \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} u^{\frac{n}{2}-1} \exp\left(-\frac{1}{2}u\right). \tag{10}$$

Bemerkung

- Die WDF der χ^2-Verteilung entspricht der WDF $G\left(u; \frac{n}{2}, 2\right)$ einer Gammaverteilung.

χ^2-Zufallsvariable und ihre Verteilung (fortgeführt)



Steigendes n verbreitert $\chi^2(u; n)$ und verschiebt Masse zur größeren Werten.

8. Beschreiben Sie die WDF der t-Verteilung in Abhängigkeit ihrer Freiheitsgrade.

- Steigendes n verschiebt Wahrscheinlichkeitsmasse aus den Ausläufen zum Zentrum.
- Ab $n = 30$ gilt $T(t; n) \approx N(0, 1)$.

t-Zufallsvariable und ihre Verteilung

Definition (t-Zufallsvariable)

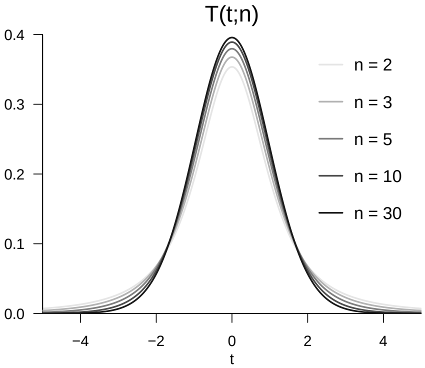
T sei eine Zufallsvariable mit Ergebnisraum $\mathbb{R}$ und WDF

$$p : \mathbb{R} \to \mathbb{R}_{>0}, t \mapsto p(t) := \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \tag{11}$$

wobei Γ die Gammafunktion bezeichne. Dann sagen wir, dass T einer t-Verteilung mit n Freiheitsgraden unterliegt und nennen T eine t-Zufallsvariable mit n Freiheitsgraden. Wir kürzen dies mit $T \sim t(n)$ ab. Die WDF einer t-Zufallsvariable bezeichnen wir mit

$$T(t; n) := \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}. \tag{12}$$

t-Zufallsvariable und ihre Verteilung (fortgeführt)



9. Geben Sie das T-Transformationstheorem wieder.

Theorem (T-Transformation)

$Z \sim N(0,1)$ sei eine z-Zufallsvariable, $U \sim \chi^2(n)$ sei eine χ^2-Zufallsvariable mit n Freiheitsgraden, und Z und U seien unabhängige Zufallsvariablen. Dann ist die Zufallsvariable

$$T := \frac{Z}{\sqrt{U/n}} \tag{13}$$

eine t-verteilte Zufallsvariable mit n Freiheitsgraden, es gilt also $T \sim t(n)$.

Anwendungsfälle: T-Konfidenzintervallstatistik und die T-Teststatistik

10. Geben Sie das F-Transformationstheorem wieder.

Theorem (F-Transformation)

$V \sim \chi^2(n)$ und $W \sim \chi^2(m)$ seien zwei unabhängige χ^2-Zufallfsvariablen mit n und m Freiheitsgraden, respektive. Dann ist die Zufallsvariable

$$F := \frac{V/n}{W/m} \tag{14}$$

eine f-verteilte Zufallsvariable mit n, m Freiheitsgraden, es gilt also $F \sim f(n, m)$.

Definition (f-Zufallsvariable)

F sei eine Zufallsvariable mit Ergebnisraum $\mathbb{R}_{>0}$ und WDF

$$p_F : \mathbb{R} \to \mathbb{R}_{>0}, f \mapsto p_F(f) := m^{\frac{m}{2}} n^{\frac{n}{2}} \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \frac{f^{\frac{m}{2}-1}}{\left(1+\frac{m}{n}f\right)^{\frac{m+n}{2}}}, \tag{15}$$

wobei Γ die Gammafunktion bezeichne. Dann sagen wir, dass F einer f-Verteilung mit n, m Freiheitsgraden unterliegt und nennen F eine f-Zufallsvariable mit n, m Freiheitsgraden. Wir kürzen dies mit $F \sim f(n, m)$ ab. Die WDF einer f-Zufallsvariable bezeichnen wir mit

$$F(f; n, m) := m^{\frac{m}{2}} n^{\frac{n}{2}} \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \frac{f^{\frac{m}{2}-1}}{\left(1+\frac{m}{n}f\right)^{\frac{m+n}{2}}}. \tag{16}$$

f-Zufallsvariable und ihre Verteilung (fortgeführt)