Übung zu Empirische Ökonomie für Fortgeschrittene

Steffen Elstner, Klaus Wohlrabe, Steffen Henzel

SS 2009

Wichtige Verteilungen

Die Normalverteilung

Eine stetige Zufallsvariable mit der Wahrscheinlichkeitsdichte $f : \mathbb{R} \to \mathbb{R}$, $x \to \phi(x)$

$$\phi(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \tag{1}$$

mit dem Erwartungswert μ und Varianz σ^2 heißt normalverteilt. Man schreibt auch $x \sim N(\mu, \sigma^2)$. Die Verteilungsfunktion lautet dann:

$$\Phi(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) dt \tag{2}$$

- Die Form wird durch die Standardabweichung (σ) bestimmt. Die Lage durch den Erwartungswert.
- Die besondere Bedeutung der Normalverteilung beruht unter anderem auf dem zentralen Grenzwertsatz, der besagt, dass eine Summe von n unabhängigen, identisch verteilten Zufallsvariablen im Grenzwert $n \to \infty$ normalverteilt ist. Das bedeutet, dass man Zufallsvariablen dann als normalverteilt ansehen kann, wenn sie durch Überlagerung einer großen Zahl von Einflüssen entstehen, wobei jede einzelne Einflussgröße einen im Verhältnis zur Gesamtsumme unbedeutenden Beitrag liefert.

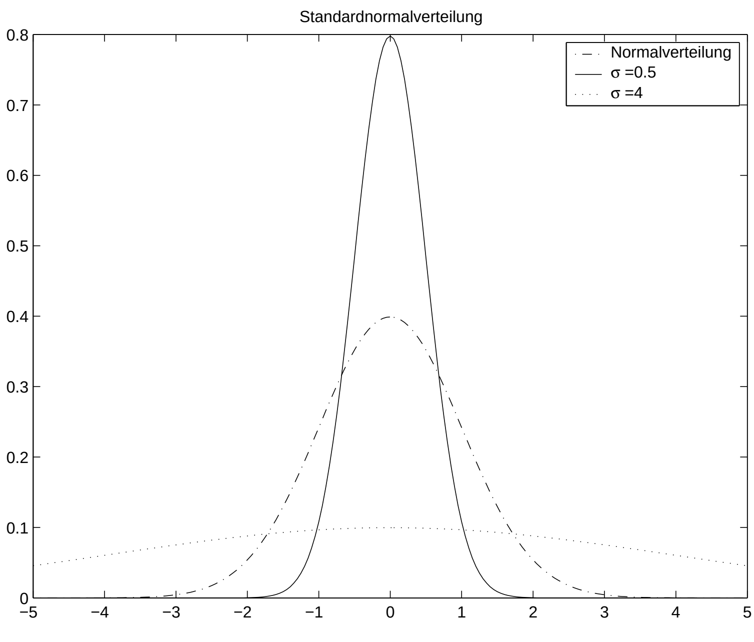

Abbildung 1: Normalverteilungen

Beispiel: Ablesen einer Wahrscheinlichkeit anhand der Tabelle der Standardnormalverteilung:

- gegeben: $Y \sim N(1,4)$; Wie groß ist die Wahrscheinlichkeit, das $Y > 3$?
- Zuerst standardisiert man Y, indem man den Erwartungswert abzieht und durch die Standardabweichung teilt: $z = \frac{Y-1}{\sqrt{4}}$; z ist dann standardnormalverteilt. $Y > 3$ entspricht $\frac{Y-1}{2} > \frac{3-1}{\sqrt{4}}$, d.h. $\frac{Y-1}{2} > 1$;
- Also gilt $\Pr(Y > 2) = \Pr\left[\frac{1}{2}(Y-1) > 1\right] = \Pr(z > 1) = 1 - \Phi(1) = 1 - 0.8413 = 0.1587$;
  (Die Werte der Verteilungsfunktion der Standardnormalverteilung finden sich im Anhang der meisten Statistiklehrbücher)
- Abbildung stellt das Beispiel graphisch dar, wobei $1 - \Phi(1)$ grau unterlegt ist. Man zieht den tabellierten Wert von eins ab, da man an den Werten rechts des kritischen Wertes interessiert ist.

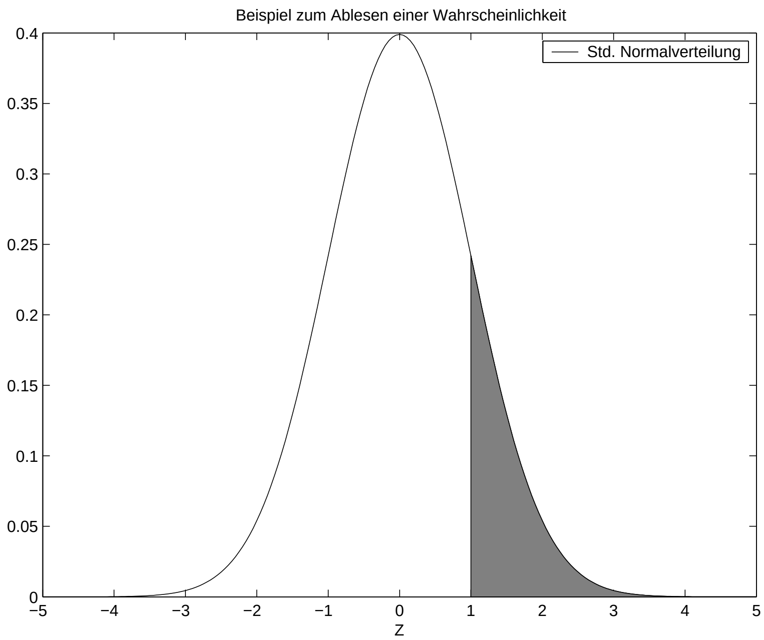

Abbildung 2: Beispiel zum Ablesen einer Wahrscheinlichkeit aus der Tabelle der Standardnormalverteilung

χ^2-Verteilung

Die Chi-Quadrat-Verteilung mit n Freiheitsgraden ist die Verteilung der Summe

$$\chi_n^2 = z_1^2 + ... + z_n^2 \tag{3}$$

n unabhängiger quadrierter standardnormalverteilter Zufallsvariablen, d. h. $z_k \sim \mathcal{N}(0,1)$ für $k = 1, \ldots, n$.

Für kleine n sind die Dichten deutlich linksteil. Für wachsendes n nähern sie sich der Gaußschen Glockenkurve an. Dies ist eine Folge des zentralen Grenzwertsatzes.

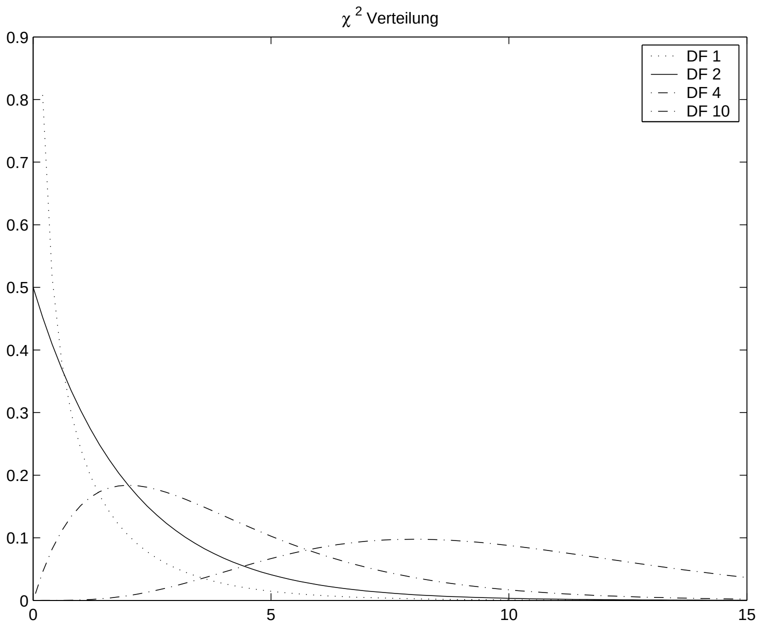

Abbildung 3: $\chi^2 - Verteilungen$

- Ist eine Zufallsvariable $Z \backsim N\left(0, \sigma^2\right)$, gilt

  $$\left(\frac{Z^2}{\sigma^2}\right) \backsim \chi_1^2$$

  Bemerkung: Chi-Quadrat-verteilt mit einem Freiheitsgrad

- Sind zwei *unabhängige* Zufallsvariablen Z_1 und Z_2 mit $Z_1 \backsim N\left(0, \sigma_1^2\right)$ und $Z_2 \backsim N\left(0, \sigma_2^2\right)$, gilt

  $$\left(\frac{Z_1^2}{\sigma_1^2} + \frac{Z_2^2}{\sigma_2^2}\right) \backsim \chi_2^2$$

  Dieser Ausdruck kann auch in Matrixschreibweise notiert werden:

  $$\begin{bmatrix} Z_1 & Z_2 \end{bmatrix} \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}^{-1} \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \backsim \chi_2^2$$

Bemerkung: Chi-Quadrat-verteilt mit zwei Freiheitsgraden

- Sind die beiden Zufallsvariablen korreliert, $\boldsymbol{z} = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \backsim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}\right)$ mit $\sigma_{21} = \sigma_{12} \neq 0$, gilt
$$\boldsymbol{z}'\sum\nolimits_z^{-1}\boldsymbol{z} \backsim \chi_2^2$$
Bemerkung: $\sum_z$ ist die Varianz-Kovarianz-Matrix von $\boldsymbol{z}$: $\sum_z = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$

- Das vorherige Resultat gilt auch analog, wenn $\boldsymbol{z}$ k Elemente enthält:
$$\boldsymbol{z}'\sum\nolimits_z^{-1}\boldsymbol{z} \backsim \chi_k^2$$

F-Verteilung

Aus den χ_m^2 und χ_n^2 Chi-Quadrat-verteilten Zufallsgrößen mit n bzw. m Freiheitsgraden lässt sich

$$F(m,n) = \frac{\frac{\chi_m^2}{m}}{\frac{\chi_n^2}{n}} \tag{4}$$

konstruieren. Dieser Ausdruck ist F-verteilt mit m und n Freiheitsgraden.

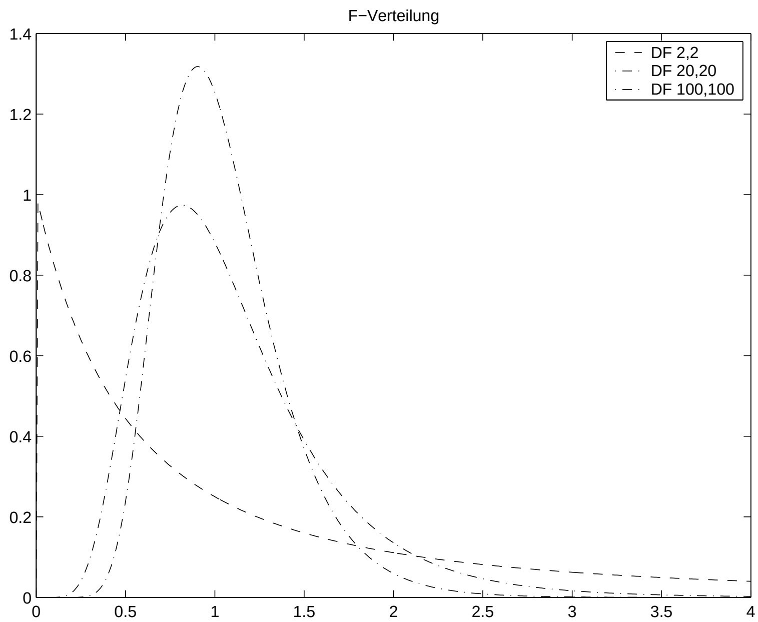

Abbildung 4: F-Verteilung

Student t-Verteilung

Die t-Verteilung mit n Freitheitsgraden ist gegeben als:

$$t_n = \frac{\mathcal{N}(0,1)}{\sqrt{\frac{\chi_n^2}{n}}} \tag{5}$$

Die Zählervariable muss unabhängig von der Nennervariable sein. Die Dichtefunktion der t-Verteilung ist dann symmetrisch bezüglich ihres Erwartungswertes 0. Eigenschaften:

- für $n > 30$ wird die t-Verteilung gut durch die Std. Normalverteilung approximiert
- für $n < 20$ hat die t-Verteilung mehr Masse an den Enden als die Normalverteilung

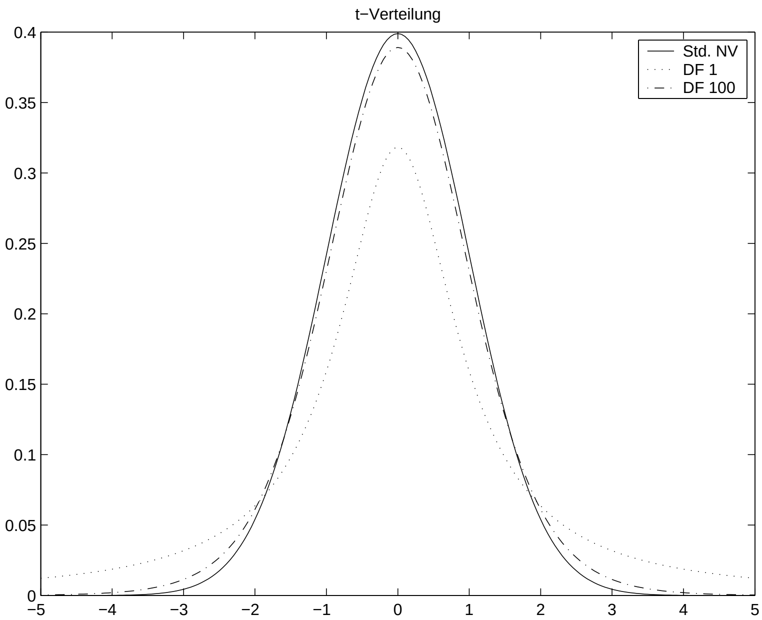

Abbildung 5: t-Verteilung

Bedingte und gemeinsame Wahrscheinlichkeiten bzw. Verteilungen

Gemeinsame Wahrscheinlichkeit

Die Wahrscheinlichkeit, dass die Zufallsvariable X den Wert x annimmt und gleichzeitig die Zufallsvariable Y den Wert y annimmt, nennt man die gemeinsame Wahrscheinlichkeit der beiden Ereignisse

$$Pr(Y = y, X = x)$$

Beispiel: Alte und neue Computer (aus Stock/Watson)

- Gleiche Anzahl an alten Computer (Ausprägung: $A = 0$) und neuen Computern (Ausprägung: $A = 1$), d.h. die Wahrscheinlichkeit einen neuen oder alten Computer zu erhalten beträgt: $Pr(A = 0) = Pr(A = 1) = 0,5$
- Die Anzahl der Computer, die abstürzt, wird mit der Ausprägung M definiert, so nimmt sie bspw. wenn ein Computer abstürzt den Wert 1 an

	$M = 0$	$M = 1$	$M = 2$	$M = 3$	$M = 4$	Total
$A = 0$	$Pr(A = 0, M = 0) = 0,35$	0,065	0,05	0,025	0,01	$Pr(A = 0) = 0,5$
$A = 1$	$Pr(A = 1, M = 0) = 0,45$	0,035	0,01	0,005	0,00	$Pr(A = 1) = 0,5$
Total	$Pr(M = 0) = 0,8$	0,1	0,06	0,03	0,01	1,0

Tabelle 1: Gemeinsame Wahrscheinlichkeiten, Bsp. aus Stock/Watson

Bedingte Wahrscheinlichkeit

Die Wahrscheinlichkeit, dass die Zufallsvariable Y den Wert y annimmt, gegeben, dass die Zufallsvariable X den Wert x annimmt wird beschrieben mit

$$Pr(Y = y|X = x)$$

Berechnet werden bedingten Wahrscheinlichkeiten mit dem Satz von Bays:

$$Pr(Y = y|X = x) = \frac{Pr(Y = y, X = x)}{Pr(X = x)}$$

Die Zufallsvariable X hat also bereits eine Ausprägung x angenommen. Mit dieser Information ändert sich unsere Betrachtung der Wahrscheinlichkeiten der Ausprägungen von Y.

Beispiel: Alte und neue Computer (aus Stock/Watson)

	$M=0$	$M=1$	$M=2$	$M=3$	$M=4$	Total
$Pr(M\|A=0)$	0,70	0,13	0,10	0,05	0,02	1,00
$Pr(M\|A=1)$	0,90	0,07	0,02	0,01	0,00	1,00

Tabelle 2: Bedingte Wahrscheinlichkeit von M gegeben A

Bedingter Erwartungswert

Der Erwartungswert, der der Zufallsvariable Y zuzuordnen ist, gegeben, dass die Zufallsvariable X den Wert x annimmt wird beschrieben mit

$$E\left(Y|X=x\right)$$

Der bedingte Erwartungswert wird berechnet durch

$$E\left(Y|X=x\right)=\sum_i y_i Pr\left(Y=y_i|X=x\right)$$

Hinweis: Ist der bedingte Erwartungswert unabhängig von der Ausprägung von X, dann sind Y und X unkorreliert.

Beispiel: Alte und neue Computer (aus Stock/Watson)

- der bedingte Erwartungswert der Ausfälle M, gegeben dass wir nur alte Computer betrachten, lässt sich wie folgt berechnen:

$$\begin{aligned} E\left(M|A=0\right) &= \sum_i M_i Pr(M=M_i|A=0) \\ &= 0\cdot 0,7+1\cdot 0,13+2\cdot 0,1+3\cdot 0,05+4\cdot 0,02 \\ &= 0,56 \end{aligned}$$

- der bedingte Erwartungswert der Ausfälle M, gegeben dass wir nur neue Computer betrachten, lässt sich wie folgt bestimmen:

$$\begin{aligned} E(M|A=1) &= \sum_i M_i P(M=M_i|A=1) \\ &= 0,14 \end{aligned}$$

Gesetz der iterierten Erwartungen

Der unbedingte Erwartungswert von Y ist gegeben durch

$$E\left(Y\right)=\sum_j E\left(Y|X=x_j\right)Pr\left(X=x_j\right)$$

oder

$$E(Y) = E[E(Y|X)]$$

Diesen Zusammenhang nennt man Gesetz der iterierten Erwartungen.

Beispiel: Alte und neue Computer (aus Stock/Watson)

$$\begin{aligned}
E(M) &= E[E(M|A)] \\
&= E(M|A=0)\,Pr(A=0) + E(M|A=1)\,Pr(A=1) \\
&= 0,56 \cdot 0,5 + 0,14 \cdot 0,5 \\
&= 0,35
\end{aligned}$$