



VS PLUS

Zusatzinformationen zu Medien des VS Verlags

Darstellung von Prüfverteilungen und Möglichkeiten der Berechnung ihrer Werte mit Hilfe von SPSS

Bernhard Dieckmann

in

Leila Akremi / Nina Baur / Sabine Fromm (Hrsg.)

**Datenanalyse mit SPSS für Fortgeschrittene 1:
Datenaufbereitung und uni- und bivariate Statistik**

Ein Arbeitsbuch

2011 / 3., überarbeitete und erweiterte Auflage

Darstellung von Prüfverteilungen und Möglichkeiten der Berechnung ihrer Werte mit Hilfe von SPSS

Bernhard Dieckmann

Zunächst sei erläutert, was eine Prüfverteilung ist. Zieht man aus einer gegebenen Grundgesamtheit wiederholt Zufalls-Stichproben mit dem Umfang n und notiert für jede der gezogenen Stichproben zum Beispiel den Kennwert „arithmetisches Mittel“, dann kann man diese Stichproben-Mittel in einer Häufigkeitsverteilung darstellen. Bei genügend häufiger Stichprobenziehung nähert sich diese Häufigkeitsverteilung einer mathematischen Funktion an, die als Prüf-Verteilung bezeichnet wird. Sie dient zur Abschätzung der Wahrscheinlichkeit dafür, dass ein Stichprobenkennwert in einem bestimmten Bereich der Prüfverteilung liegt.

In den Verfahren der Hypothesenprüfung werden folgende Prüfverteilungen angewendet:

Prüfverteilungen	Anwendungsbeispiele
Binomialverteilung	bei Hypothesen über Anteile
Standardform der Normalverteilung	bei Hypothesen über arithmetische Mittel
Chiquadrat-Verteilung	bei Hypothesen über die Differenz von erwarteten und realisierten Häufigkeiten in Kreuztabellen
t-Verteilung	bei Hypothesen über die Differenz von arithmetischen Mitteln
F-Verteilung	bei Hypothesen über den Quotienten von Treatment- und Fehlervarianz in der Varianzanalyse

Über diese Verteilungen informieren die folgenden kompakten Übersichten. Dabei unterscheiden wir folgende Bezeichnungen: Für „diskrete“ Funktionen lassen sich „Punktwahrscheinlichkeiten“ und „kumulierte Wahrscheinlichkeiten“ berechnen. Die Binomialverteilung ist eine diskrete Funktion. Für „stetige“ Funktionen lassen sich die „Wahrscheinlichkeitsdichte“ und „Wahrscheinlichkeitsfunktionen“ (Integrale) berechnen. Normalverteilung, Chiquadratverteilung, t-Verteilung und F-Verteilung sind stetige Funktionen.

Abszissen-Werte „diskreter“ Funktionen sind nur für ganze Zahlen definiert, wie die Augensumme beim Würfeln. Die Wahrscheinlichkeit, mit zwei Würfeln bei einmaligem Werfen eine Würfelaugensumme von z.B. 2 zu werfen, wird „Punktwahrscheinlichkeit“ genannt, weil sie sich nur auf den Punkt „2“ bezieht. Die Wahrscheinlichkeit, mit zwei Würfeln eine Würfelaugensumme von „weniger als 10“ zu werfen, wird „kumulierte Wahrscheinlichkeit“ genannt, weil sie durch die „Kumulierung“ (fortlaufende Addition) der Punktwahrscheinlichkeiten für die Punkte 2 bis 9 ermittelt wird. Eine „stetige Funktion“ wäre zum Beispiel die Zeit, in der Sportler 100 Meter laufen, sofern diese Zeit beliebig genau gemessen werden kann. Wenn diese Zeit in einer Grundgesamtheit von Berliner Fussballspielern „normalverteilt“ ist, ein arithmetisches Mittel von 13 Sekunden und eine Standardabweichung von 1 Sekunde hat, dann könnte interessieren, wie groß die Wahrscheinlichkeit ist, dass ein Sportler genau 12 Sekunden braucht. Wegen der hohen Messgenauigkeit ist diese Frage nicht beantwortbar. Beantwortbar ist aber die Frage, wie wahrscheinlich es ist, dass Sportler aus dieser Grundgesamtheit „zwischen 11 und 12“ Sekunden benötigen. Hierzu berechnet man die Fläche unterhalb des Funktionsgraphen einer Normalverteilung mit den oben angegebenen Parametern in dem gegebenen Intervall. Dieser Funktionsgraph wird auch „Wahrscheinlichkeitsdichte“ genannt. Die Fläche unter der Wahrscheinlichkeitsdichte in dem angegebenen Intervall ist ihr Integral in dem Intervall und dieses entspricht der Wahrscheinlichkeit, dass Zeit-Messwerte in diesem Intervall auftauchen.

Die Zuordnung dieser Bezeichnungen zueinander zeigt die folgende Übersicht:

	Einzelne Funktionswerte	Kumulierte Werte bzw. Fläche unter der Funktion
diskrete Funktionen	Punktwahrscheinlichkeit	kumulierte Wahrscheinlichkeit
stetige Funktionen	Wahrscheinlichkeitsdichte	Wahrscheinlichkeitsfunktion oder Integral oder Fläche unter der Dichtefunktion

1 Übersicht über Binomialverteilung

Verteilungstyp	Beispiel	Punktwahrscheinlichkeit Formel	Summenwahrscheinlichkeit Formel
Zwei diskrete komplementäre Merkmale A und B	Anteil von Inländern und Ausländern	$P_{n,k} = \binom{n}{k} p^k q^{n-k}$	$P_{0-u} = \sum_{k=0}^u \binom{n}{k} p^k q^{n-k}$
Terminologie	n = Stichprobengröße k = Laufvariable von 0 bis n p = Wahrscheinlichkeit des Auftretens von A q = 1-p = Wahrscheinlichkeit des Auftretens von B u: ein bestimmter Maximalwert von k		
Verlauf bestimmende Parameter: p und n	<i>Punktwahrscheinlichkeit:</i> für Messwerte von k = 0 bis k = n wenn p = 0,5: symmetrisch; stufenförmig; Maximum in der Mitte, sonst asymmetrisch. <i>Summenwahrscheinlichkeit:</i> von k = 0 bis k = n; in Stufen ansteigend Die Binomialverteilung ist durch n und p vollständig bestimmt.		
SPSS- Syntax	<i>Punktwahrscheinlichkeit für bestimmte Messwerte „k“:</i> COMPUTE Punktwahrsch = PDF.BINOM(k,n,p) . <i>Summenwahrscheinlichkeit für bestimmte Messwerte „k“:</i> COMPUTE Summenwahrsch = CDF.BINOM(k,n,p) . Graphik: GRAPH /BAR(GROUPED) = VALUE(Punktwahrsch Summenwahrsch) BY k .		

Wenn man sich mit SPSS eine numerisch exakte Vorstellung von einer Binomialverteilung mit zum Beispiel n = 15 und p = 0,6 machen will, kann man dies wie folgt erreichen: Man gebe im Dateneditor die Variable „k“ mit den Werten 0 bis 15 ein.

Abbildung 1: Daten in SPSS für die Binomialverteilung

5: Summenwahrsch ,00934766077542 Sichtbar: 3 von 3 Variablen

	k	Punktwahrsch	Summenwahrsch	var	var	var
1	0	,00	,00			
2	1	,00	,00			
3	2	,00	,00			
4	3	,00	,00			
5	4	,01	,01			
6	5	,02	,03			
7	6	,06	,10			
8	7	,12	,21			
9	8	,18	,39			
10	9	,21	,60			
11	10	,19	,78			
12	11	,13	,91			
13	12	,06	,97			
14	13	,02	,99			
15	14	,00	1,00			
16	15	,00	1,00			

Datenansicht Variablenansicht

IBM SPSS Statistics Prozessor ist bereit

Die Werte für die Punktwahrscheinlichkeiten einer Binomialverteilung mit $n = 15$ und $p = 0,6$ erhält man mit folgender Syntax (vergleiche oben):

```
COMPUTE Punktwahrsch = PDF.BINOM(k, 15, 0.6).
```

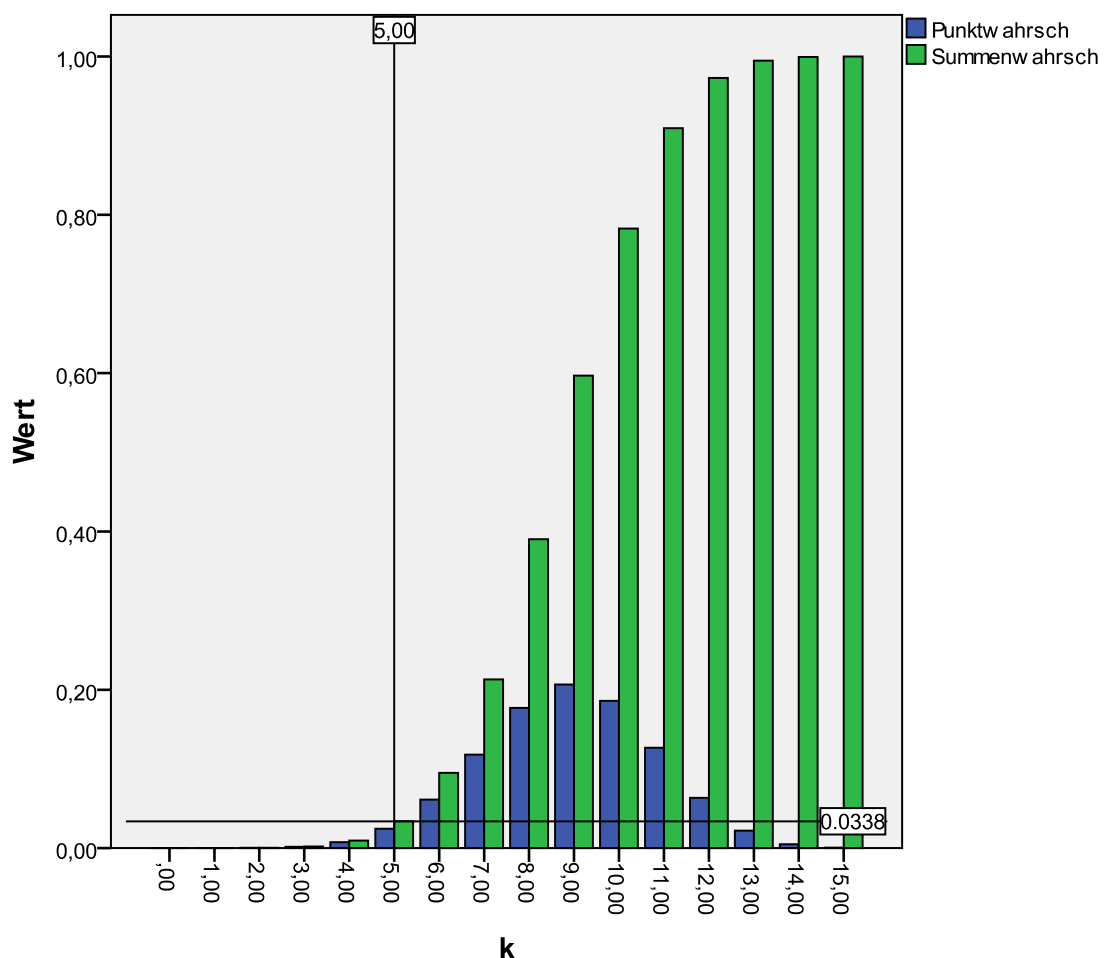
Die Werte für die Summenwahrscheinlichkeiten der gleichen Binomialverteilung erhält man mit folgender Syntax:

```
COMPUTE Summenwahrsch = CDF.BINOM(k, 15, 0.6).
```

Die Graphik der berechneten Wahrscheinlichkeiten erhält man mit folgender Syntax:

```
GRAPH  
/BAR(GROUPED) = VALUE(Punktwahrsch Summenwahrsch) BY k.
```

Abbildung 2: Punktwahrscheinlichkeiten und Summenwahrscheinlichkeiten einer Binomialverteilung mit $n = 15$ und $p = 0,6$



In die Graphik kann man mit SPSS Bezugslinien eintragen. Hier wurde der Punkt „5“ mit einer Bezugslinie versehen. Am Punkt 5 ist die Summenwahrscheinlichkeit gleich 0,0338. Was bedeutet dies? Würde man ein Experiment mit zwei möglichen Ergebnissen, bei dem das Ergebnis A die Wahrscheinlichkeit 0,6 hat, das Ergebnis B die Wahrscheinlichkeit 0,4 hat, einmal ausführen, dann kann man sehen, ob das Ergebnis A oder B ist. Wiederholt man das Experiment 15 mal (Stichprobengröße!), dann kann man zählen, wie oft A gekommen ist. Die Wahrscheinlichkeit, dass bei 15 Experimenten A genau 5 mal auftritt, ist gleich 0,0245 (oben nicht eingezeichnet). Die Summen-Wahrscheinlichkeit, dass A 0 mal oder 1 mal oder 2 mal ... oder 5 mal auftritt, ist gleich 0,0338. Dies ist die Summe der entsprechenden Punktwahrscheinlichkeiten.

Setzen wir für das Ereignis, dass aus einem Teich eine weißfleischige Forelle gefangen wird, das Symbol A, für das Ereignis, dass eine rotfleischige Forelle gefangen wird, das Symbol B, dann wird das Beispiel anschaulicher. Wichtig ist, dass bei der Binomialverteilung nur 2 einander ausschließende Ereignisse „erlaubt“ sind. Die Binomialverteilung wird in SPSS im sog. „Binomialtest“ benutzt (siehe Kapitel 11 im Lehrbuch-Text).

2 Übersicht über Normalverteilung

Verteilungstyp	Beispiel	Wahrscheinlichkeitsdichte	Integral (Fläche unter der Funktion der Wahrscheinlichkeitsdichte) = „Verteilungsfunktion“
Ein stetiges Merkmal	Körpergröße erwachsener Menschen (bei großen Stichproben; $n > 30$)	$f(x) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}$	$F(x) = P(x \leq X) = \int_{-\infty}^x f(x)dx$
Terminologie	x = Messwert μ = arithmetisches Mittel σ^2 = Varianz σ = Standardabweichung e = 2,718 π = 3,142		
Verlauf bestimmende Parameter μ und σ	Allgemeine Form der W.-dichte: von $x = -\infty$ ansteigend, Maximum bei $x = \mu$, abfallend bis $x = +\infty$; Wendepunkte der Funktion bei $\pm\sigma$. Standardform der W.-dichte (wenn $\mu=0$ und $\sigma=1$): von $z = -\infty$ ansteigend, Maximum bei $z = \mu=0$, abfallend bis $z = +\infty$. Wendepunkte der Funktion bei $z = \pm 1$. Immer symmetrisch. Integral: von $z = -\infty$ s-förmig kontinuierlich ansteigend bis zum Maximum 1. Wendepunkt bei μ . Das gesamte Integral (die Gesamtfläche unter der Dichtefunktion) ist immer = 1. Die Normalverteilung ist durch μ und σ vollständig bestimmt. Standardform: Bei der Standardform liegt ein Spezialfall vor, bei dem $\mu = 0$ und $\sigma = 1$ ist.		
SPSS- Syntax	<i>Dichtefunktion für Messwerte h:</i> COMPUTE Dichte = PDF.NORMAL(h,Mittel,Standardab). <i>Integral:</i> COMPUTE Integral = CDF.NORMAL(h,Mittel,Standardab). <i>Komplementäres Integral:</i> COMPUTE Integralkomple = 1 - CDF.NORMAL(h,Mittel,Standardab).		

Wenn man sich mit SPSS eine numerisch exakte Vorstellung von einer Standardnormalverteilung mit $\mu = 0$ und $\sigma = 1$ machen will, kann man dies wie folgt erreichen:

Man gebe im Dateneditor die Variable h mit den Werten von 1 bis 301 ein. Diese transformiere man mit der Syntax

```
COMPUTE h = ($casenum-150)/50.
EXECUTE.
```

Im Dateneditor befindet sich jetzt eine Variable h, deren Ausprägungen den Ausprägungen der Standardnormalverteilung entsprechen, so wie wir sie aus den Tabellen dieser Verteilung aus Statistikbüchern kennen. Wir können jetzt für diese Ausprägungen von SPSS errechnen lassen, welche Werte die Wahrscheinlichkeitsdichte, das Integral und das komplementäre Integral für die tabellierten Ausprägungen annehmen, wenn die Variable h eine Standardnormalverteilung wäre. Die hierfür erforderliche Syntax lautet:

```
COMPUTE wahrschDichte = PDF.NORMAL(h,0,1).
EXECUTE.
```

```
COMPUTE Integral = CDF.NORMAL(h,0,1).
EXECUTE.
```

```
COMPUTE Integralkomple = 1-(CDF.NORMAL(h,0,1)).
EXECUTE.
```

Abbildung 3: Daten in SPSS für die Normalverteilung

171 : Integral ,66275727315175 Sichtbar: 4 von 4 Variablen

	h	WahrschDichte	Integral	Integralkomple	var
141	-,18	,393	,429	,571	
142	-,16	,394	,436	,564	
143	-,14	,395	,444	,556	
144	-,12	,396	,452	,548	
145	-,10	,397	,460	,540	
146	-,08	,398	,468	,532	
147	-,06	,398	,476	,524	
148	-,04	,399	,484	,516	
149	-,02	,399	,492	,508	
150	,00	,399	,500	,500	
151	,02	,399	,508	,492	
152	,04	,399	,516	,484	
153	,06	,398	,524	,476	
154	,08	,398	,532	,468	
155	,10	,397	,540	,460	
156	,12	,396	,548	,452	
157	,14	,395	,556	,444	
158	,16	,394	,564	,436	
159	,18	,393	,571	,429	
160	,20	,391	,579	,421	
161	,22	,389	,587	,413	

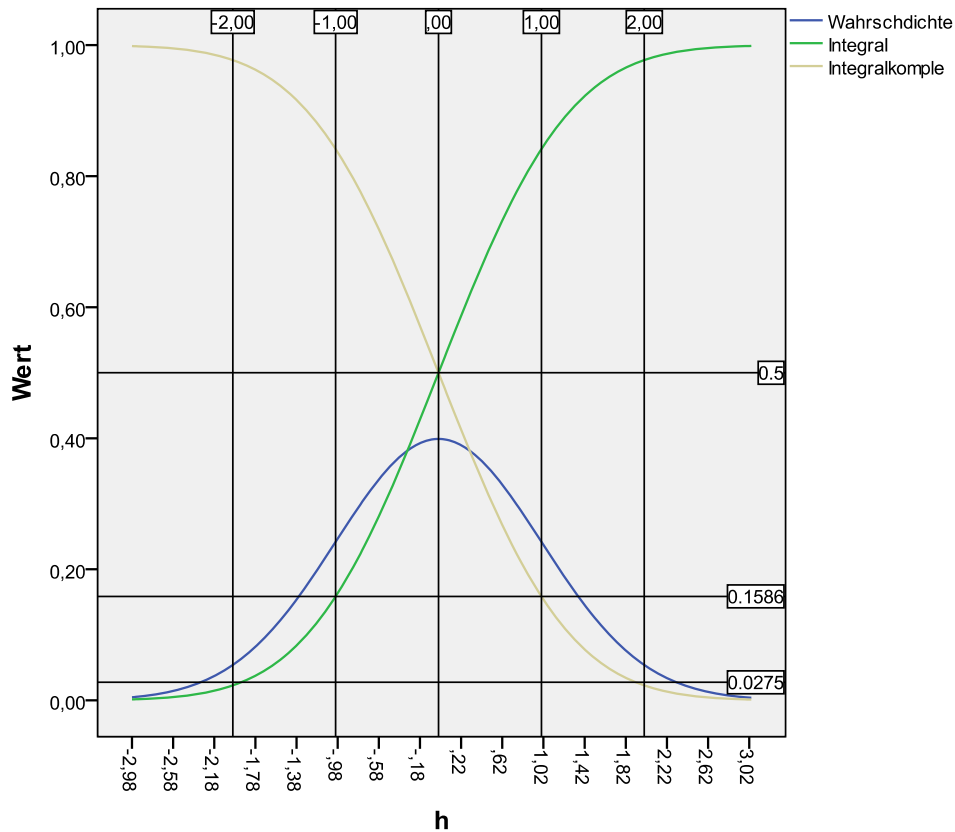
Datenansicht Variablenansicht

Der Dateneditor hierzu enthält obige Werte (Ausschnitt). Man kann an den Werten der Integrale zeigen, dass sich Integral und komplementäres Integral zu 1 addieren.

Wir können mit folgender Syntax die Wahrscheinlichkeitsdichte, deren Integral und deren komplementäres Integral in einer Graphik darstellen:

```
GRAPH
/LINE(MULTIPLE) = VALUE(WahrschDichte Integral Integralkomple) BY h.
```

Abbildung 4: Wahrscheinlichkeitsdichte Integral und komplementäres Integral für die Normalverteilung



Man kann dieser Graphik durch die eingefügten Bezugslinien zur Abszisse entnehmen, dass bei der Ausprägung -2 der Variable h (was minus 2 Standardabweichungen vom arithmetischen Mittel entspricht) das Integral der Funktion gleich 0,02275 beträgt; bei der Ausprägung -1 ist das Integral gleich 0,1586, bei der Ausprägung 0 ist es gleich 0,5. Das komplementäre Integral ist bei der Ausprägung von $h = -2$ gleich $1 - 0,0225$, bei der Ausprägung minus 1 gleich $1 - 0,1586$ etc.

Die Normalverteilung wird beim sogenannten z-Test benutzt, der auf große Stichproben ($n > 30$) anwendbar ist. Ist $n < 30$, wird der unten erwähnte t-Test benutzt.

3 Übersicht über die Chiquadratverteilung

Verteilungstyp	Beispiel	Wahrscheinlichkeitsdichte	Integral (Fläche unter der Funktion der Wahrscheinlichkeitsdichte)
Stetig; für n unabhängige Beobachtungen aus einer normalverteilten Population gilt: die Summe der quadrierten standardisierten Messwerte ist chiquadratverteilt mit n Freiheitsgraden	$\frac{ns^2}{\sigma^2} = \chi^2_{n-1}$ Der Quotient „n mal Stichproben-varianz, geteilt durch die Populationsvarianz“ ist chiquadratverteilt mit n-1 Freiheitsgraden, wenn die Population normalverteilt ist	für $\chi^2 = \sum_{i=1}^n z_i^2$ gilt $f(\chi^2) = \frac{e^{-\left(\frac{\chi^2}{2}\right)} * \chi^2\left(\frac{df}{2}-1\right)}{2^{\frac{df}{2}} * \Gamma\left(\frac{df}{2}\right)}$	$F(\chi^2) = P(\chi^2 \leq X) = \int_{-\infty}^x f(\chi^2) dx$
Terminologie	z_i^2 = standardisierter Messwert, quadriert χ^2 = Prüfverteilungs-Ausprägung e = 2,718 df = Freiheitsgrade Γ : Gammafunktion: Funktion ähnlich der Fakultät n!, auch für nichtganzzahlige Werte.		
Verlauf bestimmender Parameter: df	je nach Anzahl der df: wenn df = 1: von 0 kontinuierlich fallend, wenn df > 3: erst ansteigend, dann fallend. Mittelwert: df, Varianz: 2df Die Chiquadratverteilung ist durch df vollständig bestimmt.		
SPSS- Syntax	<i>Dichtefunktion für bestimmte Messwerte h:</i> COMPUTE (Dichte) = pdf.chisq (chiquadrat, df). <i>Integral:</i> COMPUTE (Integral_invers) = 1- cdf.chisq (chiquadrat, df).		

Wenn man sich mit SPSS eine numerisch exakte Vorstellung von einer Chiquadratverteilung z. B. mit 4 Freiheitsgraden machen will, kann man dies wie folgt erreichen:

Man gebe im Dateneditor die Variable h mit den Werten von 1 bis 64 ein. Diese transformiere man mit der Syntax

```
COMPUTE h = $casenum/4.
EXECUTE.
```

Hierdurch enthält die Variable h den unteren – relevanten – Wertebereich einer Chiquadratverteilung mit 4 Freiheitsgraden (von 0 bis 16), wobei pro Einheit vier Ausprägungen vorhanden sind.

Abbildung 5: Daten in SPSS für die Chiquadratverteilung

	h	Wahrschdichte	Integralkompl	var
1	,25	,06	,99	
2	,50	,10	,97	
3	,75	,13	,95	
4	1,00	,15	,91	
5	1,25	,17	,87	
6	1,50	,18	,83	
7	1,75	,18	,78	
8	2,00	,18	,74	
9	2,25	,18	,69	
10	2,50	,18	,64	
11	2,75	,17	,60	
12	3,00	,17	,56	
13	3,25	,16	,52	
14	3,50	,15	,48	
15	3,75	,14	,44	
16	4,00	,14	,41	
17	4,25	,13	,37	
18	4,50	,12	,34	
19	4,75	,11	,31	
20	5,00	,10	,29	
21	5,25	,10	,26	

Der Ausschnitt des Dateneditors in Abbildung 5 zeigt die Variable h und die mit der folgenden Syntax berechnete Wahrscheinlichkeitsdichte und deren komplementäres Integral:

Berechnung der Wahrscheinlichkeitsdichte:

```
COMPUTE Wahrschdichte = PDF.CHISQ(h,4).
EXECUTE.
```

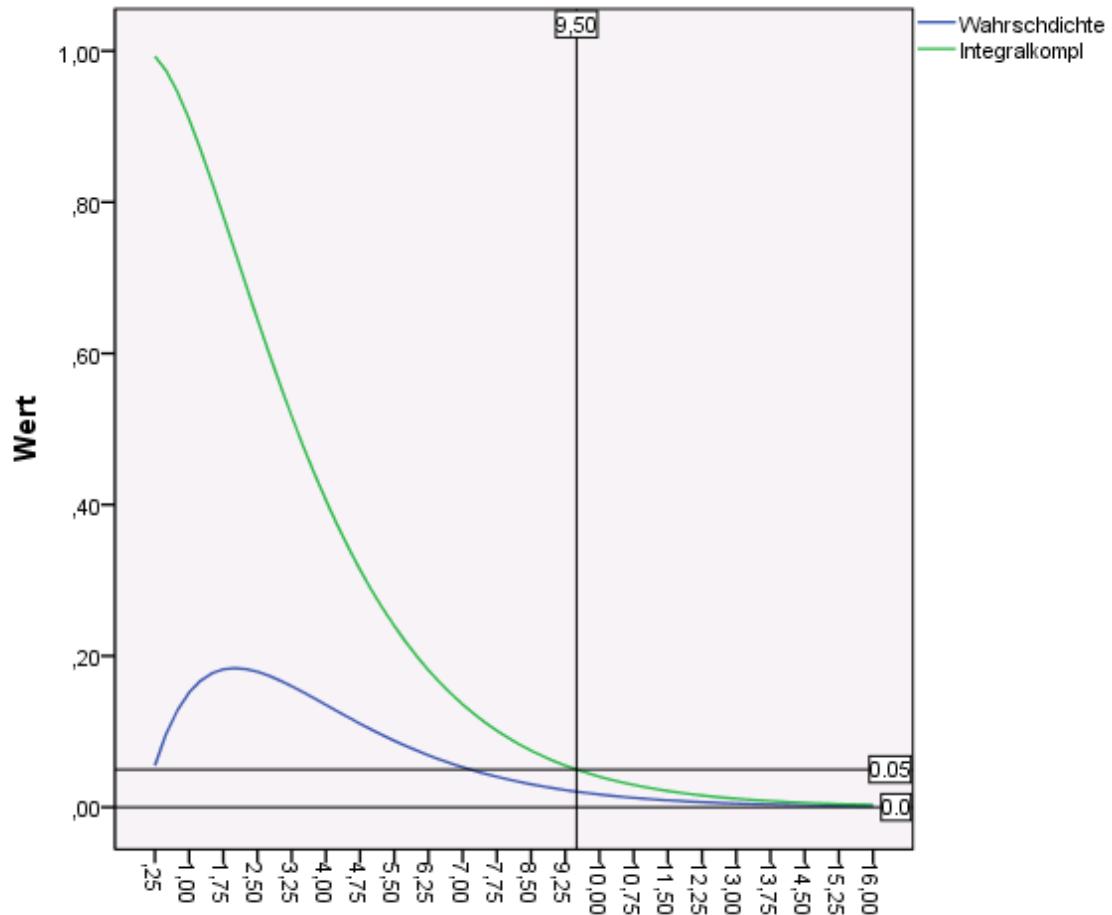
Berechnung des komplementären Integrals:

```
COMPUTE Integralkompl = 1-CDF.CHISQ(h,4).
EXECUTE.
```

Die Graphik dieser Funktionen erhalten wir folgendermaßen (siehe Abbildung 6):

GRAPH
/LINE(MULTIPLE) = VALUE(Wahrschdichte Integralkompl) BY h.

Abbildung 6: Wahrscheinlichkeitsdichte und komplementäres Integral für die Chiquadratverteilung



In diese Graphik wurden noch zwei Bezugslinien eingefügt: Eine senkrechte bei dem Wert: Chiquadrat (df_4) = 9,5 und für diesen Chiquadratwert das komplementäre Integral. Es beträgt 0,05, und markiert insofern die Grenze für die Beibehaltung einer entsprechenden Nullhypothese. Wäre das komplementäre Integral kleiner als 0,05, würde man – nach Maßgabe entsprechender Konventionen - die Nullhypothese ablehnen.

4 Übersicht über die t-Verteilung

Verteilungstyp	Beispiel	Wahrscheinlichkeitsdichte	Integral (Fläche unter der Funktion der Wahrscheinlichkeitsdichte)
Ein stetiges Merkmal	Benzinverbrauch eines Autos, kleine Stichprobe, $n < 30$	für $t_{df} = \frac{z}{\sqrt{\frac{\chi_{df}^2}{df}}}$ gilt: $h_{df}(t) = \frac{\Gamma \frac{df+1}{2}}{\sqrt{df\pi} \Gamma \frac{df}{2}} \left[1 + \frac{t^2}{df} \right]^{-\frac{df+1}{2}}$	$F(t) = P(t \leq X) = \int_{-\infty}^t f(t) dx$
Terminologie	df = Menge der Freiheitsgrade; Bestimmung der Freiheitsgrade je nach t-Test-Typ unterschiedlich t = Prüfverteilungs-Ausprägung $\pi = 3,141593$ df = Freiheitsgrade Γ: Gammafunktion		
Verlauf bestimmende Parameter	W.-dichte: von $-\infty$ ansteigend, Maximum bei 0, abfallend bis $+\infty$; immer symmetrisch Integral: von $-\infty$ s-förmig kontinuierlich ansteigend bis zum Maximum 1 Mittelwert: 0, Varianz: $df/(df-2)$ Die t-Verteilung ist durch df vollständig bestimmt.		
SPSS- Syntax	Dichtefunktion für bestimmte Messwerte w: COMPUTE Dichtedf = PDF.T(t, df). Integral: COMPUTE Integraldf = CDF.T(t, df) Komplementäres Integral: COMPUTE Integraldfkompl = 1-CDF.T(t, df)		

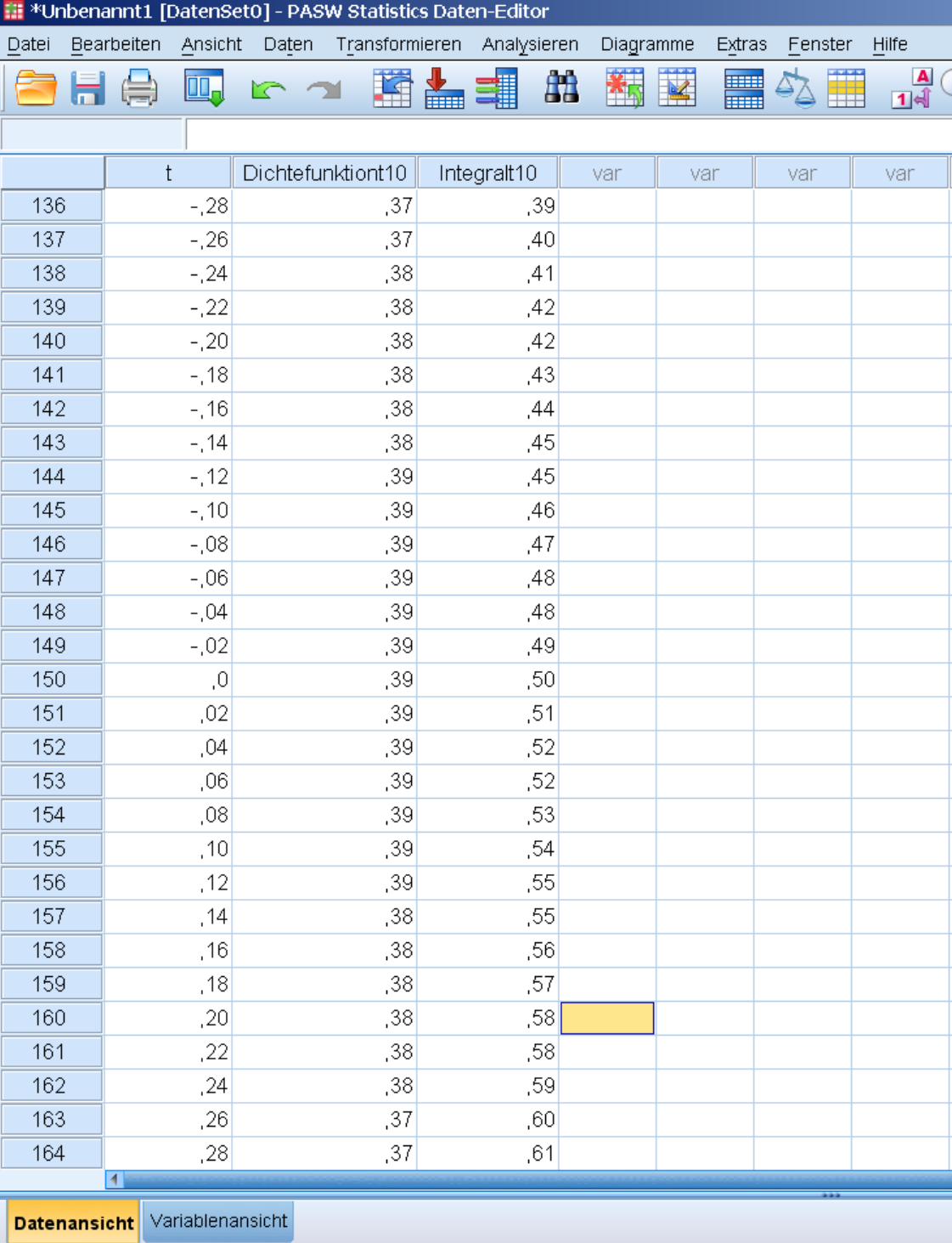
Wenn man sich mit SPSS eine numerisch exakte Vorstellung von einer t-Verteilung mit z.B. mit 10 Freiheitsgraden machen will, kann man dies wie folgt erreichen: Man gebe im Dateneditor – wie oben bei der Standardnormalverteilung schon praktiziert – die Werte von 1 bis 301 ein und transformiere diese wie folgt: Compute t = (\$casenum-150)/50. Die Dichtefunktion und das Integral erhält man durch die Syntax

COMPUTE Dichte10 = PDF.T(t,10).
 Integral: COMPUTE Integral10 = CDF.T(t, 10).

Abbildung 7: Daten in SPSS für die t-Verteilung

*Unbenannt1 [DatenSet0] - PASW Statistics Daten-Editor

Datei Bearbeiten Ansicht Daten Transformieren Analysieren Diagramme Extras Fenster Hilfe



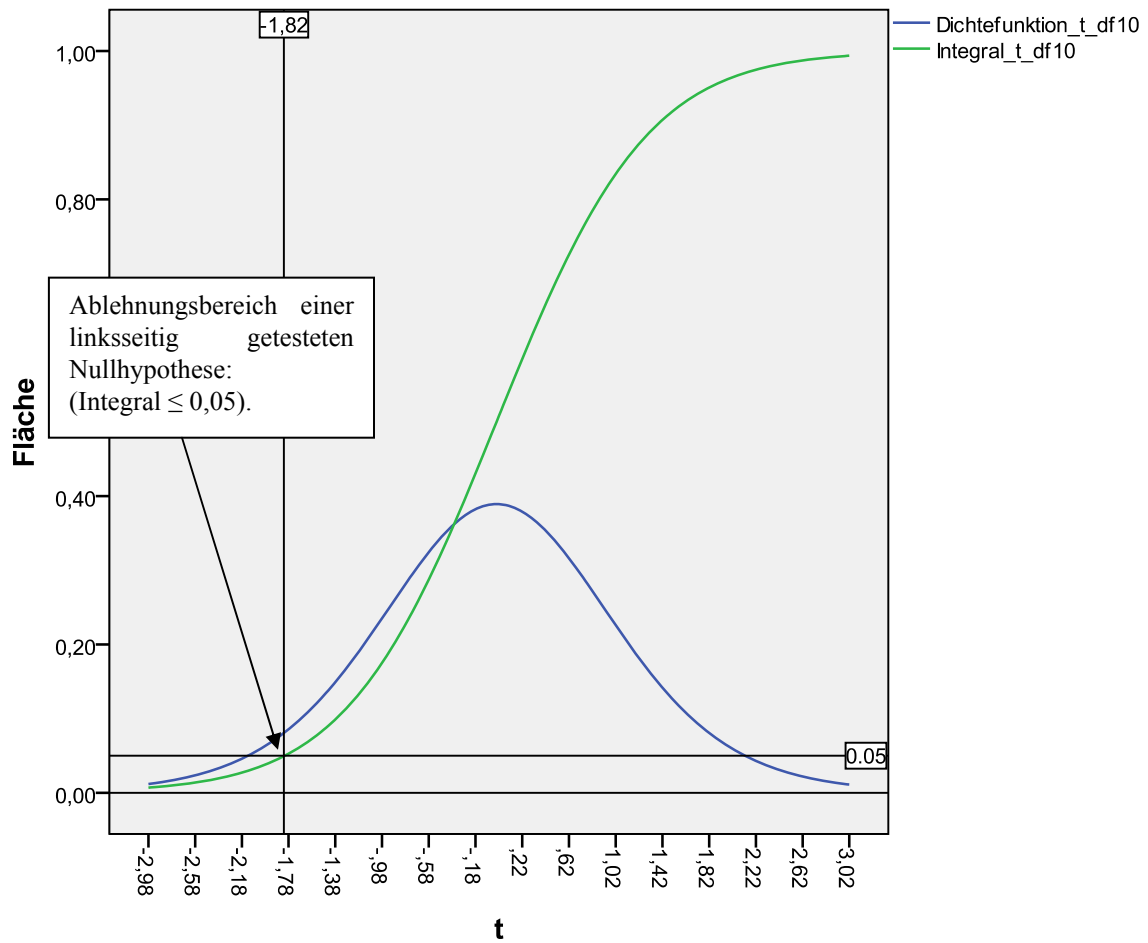
	t	Dichtefunktiont10	Integralt10	var	var	var	var
136	-,28	,37	,39				
137	-,26	,37	,40				
138	-,24	,38	,41				
139	-,22	,38	,42				
140	-,20	,38	,42				
141	-,18	,38	,43				
142	-,16	,38	,44				
143	-,14	,38	,45				
144	-,12	,39	,45				
145	-,10	,39	,46				
146	-,08	,39	,47				
147	-,06	,39	,48				
148	-,04	,39	,48				
149	-,02	,39	,49				
150	,0	,39	,50				
151	,02	,39	,51				
152	,04	,39	,52				
153	,06	,39	,52				
154	,08	,39	,53				
155	,10	,39	,54				
156	,12	,39	,55				
157	,14	,38	,55				
158	,16	,38	,56				
159	,18	,38	,57				
160	,20	,38	,58				
161	,22	,38	,58				
162	,24	,38	,59				
163	,26	,37	,60				
164	,28	,37	,61				

Datenansicht Variablenansicht

Die Graphik wird wie folgt erzeugt:

```
GRAPH
/LINE(MULTIPLE)= VALUE(Dichtefunktiont10 Integralt10) BY t.
```

Abbildung 8: Dichtefunktion und Integral einer t-Verteilung mit 10 Freiheitsgraden



Eingezeichnet ist der Ablehnungsbereich der Nullhypothese bei linksseitigem Hypothesentest bei Integralwert von bis zu 0,05 (bei einem t-Wert von bis zu -1,82).

5 Übersicht über die F-Verteilung

Verteilungs- typ	Beispiel	Wahrscheinlichkeitsdichte	Integral (Fläche unter der Funktion der Wahrscheinlichkeitsdichte)
stetig für ein Merkmal	Vergleich des durchschnitt- lichen Teergehalts verschiedener Zigarettenarten	für $F_{(m,n)} = \frac{\frac{\chi_m^2}{m}}{\frac{\chi_n^2}{n}}$ gilt: $g(m,n) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{m}{n}\right)^{\frac{m}{2}} \frac{(\chi)^{\frac{m}{2}-1}}{\left(1+\frac{m}{n}\chi\right)^{\frac{m+n}{2}}}$	$F(g) = P(g \leq X) = \int_{-\infty}^t f(g) dx$
Terminologie	χ_m^2 chiquadrat-verteilter Wert im Zähler der F-Verteilung; χ_n^2 chiquadrat-verteilter Wert im Nenner der F-Verteilung F = Prüfverteilungs-Ausprägung auf der Abszisse; m = df ₁ = Freiheitsgrade der Zählervarianz (Treatmentvarianz); n = df ₂ = Freiheitsgrade der Nennervarianz (Fehlervarianz) Γ: Gammafunktion g = Ordinatenwerte der Dichtefunktion der F-Verteilung F(g): Ordinatenwerte des Integrals der Dichtefunktion der F-Verteilung		
Verlauf be- stimmende Parameter: df ₁ und df ₂ , in der hier gebrauchten Terminologie: m und n	Eine Zufallsvariable, die F-verteilt ist, hat das arithmetische Mittel n/n-1 für für n > 2 und die Varianz $s^2 = \frac{2n^2(n+m-2)}{m(n-4)(n-2)^2}$ für n > 4 Für jede Kombination von Freiheitsgraden für n>4 gibt es eine F-Verteilung. Die Ordinate dieser Verteilungen steigt von 0 an bis zu einem Maximum und fällt dann kontinuierlich ab. Das Integral steigt s-förmig von 0 an und erreicht sein Maximum bei 1.		
SPSS-Syntax	<i>Dichtefunktion für bestimmte F-Werte:</i> COMPUTE (Dichte_def6_df16) = PDF.F(FWerte,df ₁ ,df ₂). <i>Inverses Integral:</i> COMPUTE (Integral_invers_df6_df16) = 1- CDF.F(FWerte, df ₁ ,df ₂)		

Wenn man sich mit SPSS eine numerisch exakte Vorstellung von einer F-Verteilung mit z.B. mit 6-Zähler- und 16 Nennerfreiheitsgraden machen möchte, gebe man die Zahlen 1-64 in den Dateneditor ein und erzeuge dann die Variable f wie folgt (einfachstes Verfahren für die Erzeugung des relevanten Wertebereichs im Dateneditor):

```
COMPUTE f = $casenum/16.
EXECUTE.
```

Die Dichtefunktion und das komplementäre Integral kann man dann wie folgt berechnen:

```
COMPUTE DichteF6_16 = PDF.F(f,6,16).
EXECUTE.

COMPUTE IntegralKomplF6_16 = 1-CDF.F(f,6,16).
EXECUTE.
```

Im Dateneditor zeigen sich dann folgende Daten (Auszug):

Abbildung 9: Daten in SPSS für die F-Verteilung

*Unbenannt1 [DatenSet0] - PASW Statistics Daten-Editor

Datei Bearbeiten Ansicht Daten Transformieren Analysieren Diagramme Extras Fenster Hilfe

44 : f 2,75

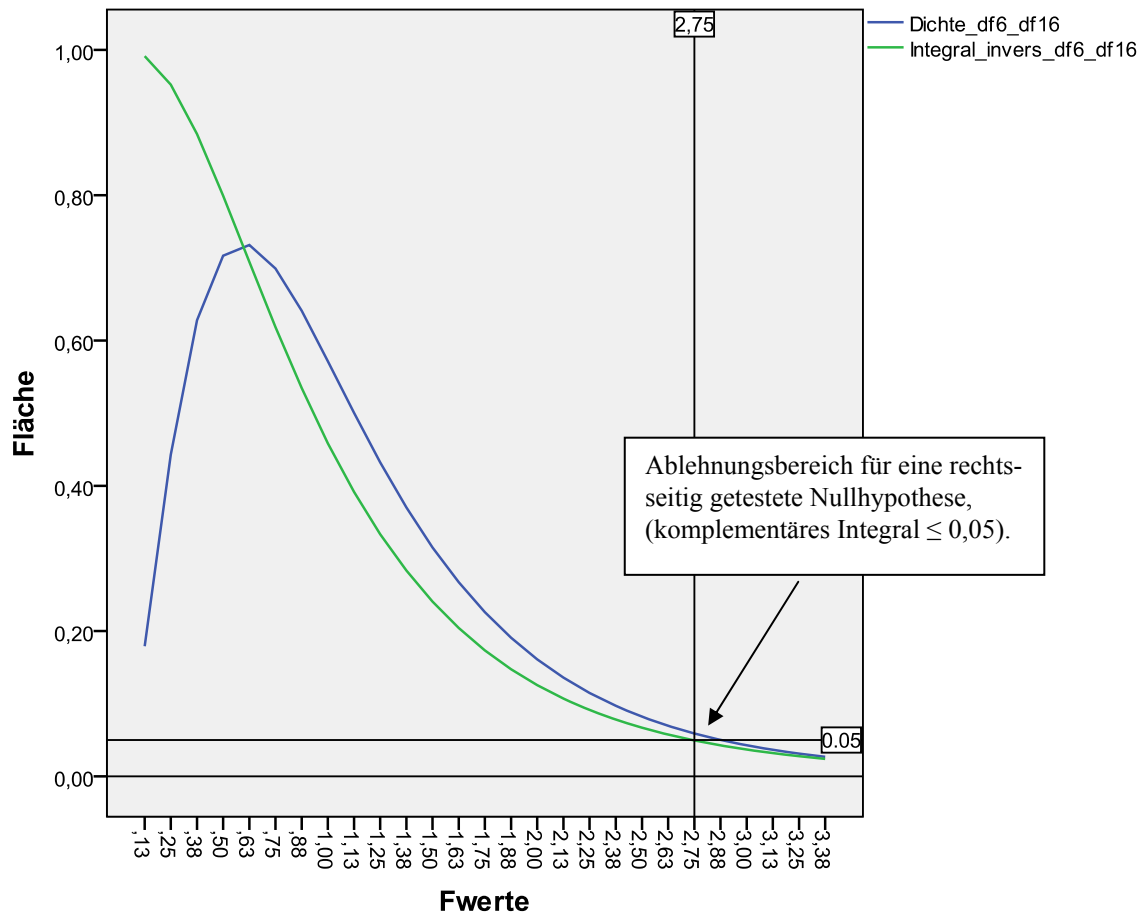
	f	DichteF6_16	IntegralKempl F6_16	var	var	var	var	var	var
28	1,75	,23	,17						
29	1,81	,21	,16						
30	1,88	,19	,15						
31	1,94	,18	,14						
32	2,00	,16	,13						
33	2,06	,15	,12						
34	2,13	,14	,11						
35	2,19	,12	,10						
36	2,25	,11	,09						
37	2,31	,11	,08						
38	2,38	,10	,08						
39	2,44	,09	,07						
40	2,50	,08	,07						
41	2,56	,08	,06						
42	2,63	,07	,06						
43	2,69	,06	,05						
44	2,75	,06	,05						
45	2,81	,05	,05						
46	2,88	,05	,04						
47	2,94	,05	,04						
48	3,00	,04	,04						
49	3,06	,04	,03						
50	3,13	,04	,03						
51	3,19	,03	,03						
52	3,25	,03	,03						
53	3,31	,03	,03						
54	3,38	,03	,02						
55	3,44	,02	,02						

Datenansicht Variablenansicht

Die Graphik der Dichtefunktion und des komplementären Integrals läßt sich wie folgt erzeugen:

GRAPH

/LINE(MULTIPLE) = VALUE(DichteF6_16 IntegralKemplF6_16) BY f.

Abbildung 10: Dichtefunktion und inverses Integral einer F-Verteilung mit $df_1 = 6$, $df_2 = 16$ 

Eingezeichnet ist der Ablehnungsbereich für eine rechtsseitig getestete Nullhypothese bei einem Wert des inversen Integrals $\leq 0,05$ (F-Wert $> 2,75$). Weil die Hypothese hier rechtsseitig getestet wurde, wurde auch das komplementäre Integral gewählt, um die 5%-Grenze zu zeigen.

Datenanalyse mit SPSS für Fortgeschrittene 1
Datenaufbereitung und uni- und bivariate Statistik
Akremi, L.; Baur, N.; Fromm, S. (Hrsg.)
2011, 339 S. 63 Abb., Softcover
ISBN: 978-3-531-17015-2