

**Julius Dufner, Uwe Jensen, Erich Schumacher**

# **Statistik mit SAS**

3., überarbeitete Auflage

*HLuHB Darmstadt*



**15914866**

B. G. Teubner Stuttgart · Leipzig · Wiesbaden

# Inhaltsverzeichnis

## Kapitel 1 Einführung in SAS

<b>1.1</b>	<b>Das SAS - Softwaresystem.....</b>	<b>13</b>
<b>1.2</b>	<b>Die SAS-Benutzeroberfläche .....</b>	<b>15</b>
<b>1.2.1</b>	<b>SAS-Fenstersystem .....</b>	<b>16</b>
<b>1.2.2</b>	<b>Arbeiten in Fenstern .....</b>	<b>18</b>
<b>1.3</b>	<b>Statistik-Komponenten.....</b>	<b>21</b>

## Kapitel 2 Das SAS-Programmsystem

<b>2.1</b>	<b>Ein einführendes Beispiel.....</b>	<b>23</b>
<b>2.1.1</b>	<b>DATA step und PROC step.....</b>	<b>24</b>
<b>2.1.2</b>	<b>SAS-Programm .....</b>	<b>25</b>
<b>2.1.3</b>	<b>Realisierung .....</b>	<b>27</b>
<b>2.2</b>	<b>Ergänzungen .....</b>	<b>32</b>
<b>2.2.1</b>	<b>SAS-Programm .....</b>	<b>32</b>
<b>2.2.2</b>	<b>Realisierung .....</b>	<b>33</b>
<b>2.2.3</b>	<b>Permanente SAS-Dateien .....</b>	<b>36</b>
<b>2.2.4</b>	<b>Regeln zur Programmgestaltung.....</b>	<b>37</b>
<b>2.3</b>	<b>Externe Daten.....</b>	<b>37</b>
<b>2.3.1</b>	<b>ASCII-Dateien .....</b>	<b>37</b>
<b>2.3.2</b>	<b>Dateien anderer Softwaresysteme.....</b>	<b>39</b>
<b>2.4</b>	<b>Die Programmiersprache SAS.....</b>	<b>40</b>
<b>2.4.1</b>	<b>SAS-Anweisungen .....</b>	<b>40</b>
<b>2.4.2</b>	<b>SAS-Programme .....</b>	<b>41</b>
<b>2.4.3</b>	<b>Beschreibung der benutzten Anweisungen .....</b>	<b>41</b>
<b>2.4.3.1</b>	<b>DATA step .....</b>	<b>42</b>
<b>2.4.3.2</b>	<b>PROC step.....</b>	<b>45</b>
<b>2.4.3.3</b>	<b>Anweisungen an beliebiger Stelle eines SAS-Programms.....</b>	<b>46</b>

## Kapitel 3 Beschreibende Statistik

<b>3.1</b>	<b>Eindimensionale Stichproben</b>	49
3.1.1	Graphische Darstellungen	50
3.1.1.1	Histogramme	50
3.1.1.2	Ausgabe und Export von SAS-Graphiken	55
3.1.1.3	Stabdiagramme	56
3.1.1.4	Kreisdiagramme	60
3.1.2	Statistische Maßzahlen	61
3.1.2.1	Lagemaße	62
3.1.2.2	Streuungsmaße	62
3.1.2.3	Formmaße	63
3.1.2.4	Statistische Maßzahlen mit SAS	64
<b>3.2</b>	<b>Zwei- und mehrdimensionale Stichproben</b>	69
3.2.1	Punktediagramme	69
3.2.2	Zusammenhangsmaße	70
3.2.3	Anpassung von Regressionsfunktionen	74
3.2.3.1	Prinzip der kleinsten Quadrate	74
3.2.3.2	Lineare Anpassung	77
3.2.3.3	Nichtlineare Anpassung	86
3.2.3.4	Ergänzungen zum DATA step	102

## Kapitel 4 Grundlagen der Wahrscheinlichkeitstheorie und Statistik

<b>4.1</b>	<b>Wahrscheinlichkeitstheorie</b>	105
4.1.1	Ereignisse, Stichprobenraum	106
4.1.2	Wahrscheinlichkeiten	106
4.1.3	Zufallsvariable	107
4.1.4	Einige spezielle Wahrscheinlichkeitsverteilungen	112
4.1.4.1	Diskrete Verteilungen	112
4.1.4.2	Absolut stetige Verteilungen	115
4.1.5	Grenzwertsätze	119
4.1.6	Testverteilungen	121
4.1.6.1	Die Chi-Quadrat ( $\chi^2$ ) -Verteilung	121
4.1.6.2	Die Studentsche t-Verteilung	122
4.1.6.3	Die F(isher)-Verteilung	123

<b>4.2</b>	<b>Grundlagen der beurteilenden Statistik</b>	124
4.2.1	Parameterschätzung	124
4.2.1.1	Punktschätzungen	124
4.2.1.2	Intervallschätzungen - Vertrauensintervalle	128
4.2.2	Tests	129

## Kapitel 5 Beurteilende Statistik - Grundlegende Verfahren

<b>5.1</b>	<b>Tests bei Normalverteilungsannahme</b>	132
5.1.1	Einstichproben-Tests	132
5.1.1.1	Test des Erwartungswertes – Einstichproben t-Test	132
5.1.1.2	Test der Varianz	138
5.1.2	Zweistichproben-Tests	141
5.1.2.1	Vergleich verbundener (gepaarter) Stichproben	141
5.1.2.2	Vergleich unabhängiger Stichproben – Der t-Test	142
<b>5.2</b>	<b>Anpassungstests</b>	148
5.2.1	Übersicht über einige Anpassungstests	148
5.2.2	Der Shapiro-Wilk Test	155
<b>5.3</b>	<b>Verteilungsfreie Verfahren - Nichtparametrische Methoden</b>	159
5.3.1	Einstichproben-Tests	159
5.3.1.1	Der Binomialtest	159
5.3.1.2	Test auf Zufälligkeit	162
5.3.2	Zwei- und k-Stichprobentests	165
5.3.2.1	Vergleich zweier verbundener Stichproben	165
5.3.2.2	Vergleich zweier unverbundener Stichproben	169
5.3.2.3	Vergleich mehrerer unabhängiger Stichproben - Der Kruskal-Wallis Test	173
5.3.2.4	Vergleich mehrerer verbundener Stichproben - Der Friedman Test	176
5.3.3	Kontingenztafeln – Unabhängigkeitstests und Homogenitätstests	179
5.3.3.1	Der Unabhängigkeitstest	180
5.3.3.2	Der exakte Test von Fisher	184
5.3.3.3	Der Homogenitätstest	188

## Kapitel 6 Varianzanalyse

<b>6.1</b>	<b>Einfaktorielle Varianzanalyse - fixe Effekte.....</b>	191
6.1.1	Varianzanalysemodell und F-Test.....	192
6.1.2	Gütfunktion und Wahl des Stichprobenumfangs .....	196
6.1.3	Durchführung in SAS – Beispiel 6_1 .....	198
6.1.4	Abweichungen von den Modellvoraussetzungen.....	201
6.1.5	Überprüfung von Modellvoraussetzungen .....	203
6.1.5.1	Test der Normalverteilungsannahme.....	203
6.1.5.2	Der modifizierte Levene-Test .....	205
6.1.6	Überparametrisierung des Modells .....	208
<b>6.2</b>	<b>Multiple Mittelwertsvergleiche.....</b>	209
6.2.1	Schätzung der Modellparameter.....	210
6.2.2	Vertrauensintervall und Test für eine Paardifferenz .....	211
6.2.3	Multiple Tests und simultane Vertrauensintervalle .....	212
6.2.3.1	Bonferroni- und Sidak-Test .....	212
6.2.3.2	Scheffe -Test .....	213
6.2.3.3	Tukey-Test und Tukey-Kramer-Test .....	214
6.2.3.4	Dunnett -Test für Vergleiche mit einer Kontrolle .....	215
6.2.4	Sidak- , Scheffe -Tests und lineare Kontraste in SAS .....	216
6.2.4.1	Sidak- und Scheffe -Tests in SAS .....	216
6.2.4.2	Lineare Kontraste in SAS .....	218
6.2.5	Wachstumsversuch, Tukey- und Dunnett -Tests in SAS .....	220
6.2.5.1	Vollständig zufällige Zuteilung mittels PROC PLAN.....	221
6.2.5.2	Auswertung in SAS .....	222
6.2.6	Vergleich simultaner Testprozeduren .....	227
6.2.6.1	Die Tests nach Bonferroni, Sidak, Scheffe, Tukey .....	227
6.2.6.2	Lineare Kontraste .....	228
6.2.6.3	Sequentielle Testprozeduren.....	229
6.2.6.4	Zusammenfassung .....	232
<b>6.3</b>	<b>Einfaktorielle Varianzanalyse - zufällige Effekte.....</b>	232
<b>6.4</b>	<b>Zweifaktorielle Varianzanalyse - Kreuzklassifikation .....</b>	235
6.4.1	Zweifaktorielle Varianzanalyse, fixe Effekte .....	236
6.4.1.1	Modell, F-Tests und paarweise Vergleiche.....	237
6.4.1.2	Durchführung in SAS – Beispiel 6_4 .....	240
6.4.2	Zweifaktorielle Varianzanalyse, zufällige Effekte .....	244
6.4.2.1	Modell und F-Tests .....	244
6.4.2.2	Durchführung in SAS .....	246

<b>6.4.3</b>	<b>Zweifaktorielles gemischtes Modell.....</b>	<b>248</b>
6.4.3.1	Gemischtes Modell und F-Tests.....	248
6.4.3.2	Durchführung in SAS .....	250
6.4.4	Eine Beobachtung pro Zelle.....	251
6.4.4.1	Modell und F-Tests .....	252
6.4.4.2	Durchführung in SAS .....	254
6.4.5	Höherfaktorielle kreuzklassifizierte Versuche .....	255
6.4.5.1	Dreifaktorielle kreuzklassifizierte Varianzanalyse .....	255
6.4.5.2	Durchführung in SAS .....	256
6.4.5.3	r-faktorielle kreuzklassifizierte Varianzanalyse .....	256
<b>6.5</b>	<b>Zweifaktorielle hierarchische Varianzanalyse.....</b>	<b>257</b>
6.5.1	Modell und F-Tests .....	258
6.5.2	Durchführung in SAS – Beispiel 6_5 .....	260
6.5.2.1	Tests .....	260
6.5.2.2	Schätzung der Varianzkomponenten .....	263
6.5.3	Höherfaktorielle Modelle.....	264
<b>6.6</b>	<b>Versuchsplanung - spezielle Randomisationsstrukturen .....</b>	<b>265</b>
6.6.1	Complete Randomized Designs.....	266
6.6.2	Randomisierte vollständige Blockanlagen .....	266
6.6.2.1	Modell , F-Tests und paarweise Vergleiche .....	268
6.6.2.2	Durchführung in SAS – Beispiel 6_6 .....	269
6.6.2.3	Modell mit zufälligen Blockeffekten .....	272
6.6.3	Zweifaktorielle Anlage in Blöcken .....	272
6.6.4	Split-Plot Anlage in Blöcken .....	274
6.6.4.1	Modell und F-Tests .....	274
6.6.4.2	Multiple Vergleiche .....	277
6.6.4.3	Durchführung in SAS – Beispiel 6_7 .....	280
<b>6.7</b>	<b>Unbalancierte Daten.....</b>	<b>288</b>
6.7.1	Zweifaktorielle Kreuzklassifikation, unbalancierte Daten, keine leeren Zellen.....	289
6.7.1.1	Modell.....	289
6.7.1.2	Beispiel 6_8 und R- Notation.....	291
6.7.1.3	Typ I- Quadratsummenzerlegung .....	295
6.7.1.4	Typ II- Quadratsummen .....	297
6.7.1.5	Typ III- Quadratsummenzerlegung .....	299
6.7.1.6	Durchführung in SAS – Beispiel 6_8 (fortgesetzt) .....	301
6.7.2	Paarweise Vergleiche adjustierter Erwartungswerte.....	303
6.7.2.1	Adjustierte Erwartungswerte – LSMeans .....	303

6.7.2.2	Durchführung in SAS – Beispiel 6_8 (fortgesetzt) .....	305
6.7.3	Modelle mit leeren Zellen – die Typ IV- Zerlegung .....	307
6.7.3.1	Schätzbare Funktionen und testbare Hypothesen .....	308
6.7.3.2	Typ IV- Quadratsummen.....	310
6.7.3.3	Typ IV-Zerlegung – Beispiel 6_9.....	310
6.7.3.4	Durchführung in SAS – Beispiel 6_9 .....	314
6.7.4	Auswertung mehrfaktorieller Modelle in SAS .....	319

## Kapitel 7 Lineare Regressionsanalyse

7.1	<b>Einfache lineare Regression</b> .....	322
7.1.1	Schätzung der Modellparameter.....	324
7.1.2	Univariate Vertrauensintervalle und Tests.....	327
7.1.3	Simultane Vertrauensbereiche und Tests.....	328
7.1.4	Durchführung in SAS – Beispiel 7_1 .....	329
7.1.5	Überprüfung der Modellannahmen.....	335
7.1.6	Ergänzungen .....	336
7.1.6.1	Prognose - Intervall für eine Beobachtung.....	336
7.1.6.2	Regression ohne Absolutglied .....	337
7.2	<b>Multiple lineare Regressionsanalyse</b> .....	340
7.2.1	Schätzung der Modellparameter.....	341
7.2.2	Univariate Vertrauensintervalle und Tests .....	344
7.2.3	Simultane Vertrauensbereiche und Tests.....	345
7.2.4	Überprüfung der Modellannahmen.....	347
7.2.5	Durchführung in SAS – Beispiel 7_2 .....	348
7.2.6	Techniken zur Modellauswahl.....	354
7.3	<b>Kovarianzanalyse</b> .....	357
7.3.1	Einfache Kovarianzanalyse .....	357
7.3.1.1	Schätzung der Modellparameter.....	359
7.3.1.2	Tests und paarweise Vergleiche .....	361
7.3.1.3	Durchführung in SAS – Beispiel 7_3 .....	364
7.3.1.4	Überprüfung von Modellannahmen .....	370
7.3.2	Erweiterungen des Kovarianzanalysemodells .....	373
	<b>Literaturverzeichnis</b> .....	374
	<b>Sachverzeichnis</b> .....	384