

Medizinische Statistik mit SAS-Analyst

Rainer Muche, Andreas Habel, Friederike Rohlmann

Universität Ulm

Abteilung Biometrie und Med. Dokumentation

89070 Ulm

rainer.muche@medizin.uni-ulm.de

Zusammenfassung

In der medizinischen Forschung, insbesondere im universitären Umfeld, werden sehr häufig statistische Auswertungen durchgeführt. In der Medizin gibt es drei typische Nutzergruppen mit unterschiedlichen Bedürfnissen und Anforderungen an die Statistiksoftware. Die Möglichkeiten der Nutzung von SAS-Analyst für diese Nutzer, speziell ungeübte, werden beschrieben. Als Resümee kann man zusammenfassen, dass SAS-Analyst ein ausgezeichnetes Tool mit kleinen Schwächen für Nutzer mit geringen Vorkenntnissen im medizinischen Umfeld ist. Eine Erweiterung der Nutzeroberfläche auf alle Möglichkeiten in SAS bzw. der Vertrieb einer Studentenversion SAS-Analyst würden die Einsatzmöglichkeiten wesentlich erweitern.

Keywords: Medizinstatistik, Biometrie, SAS-Analyst.

1 Einleitung

In der medizinischen Forschung, insbesondere im universitären Umfeld, werden sehr häufig statistische Auswertungen durchgeführt. Es gibt einige typische Spezifika in der Medizinstatistik, die diese von anderen Statistikbereichen unterscheidet:

- Daten: – kleine Datensätze
– geringe Anzahl Beobachtungen
– viele Zielgrößen
- Methoden: – viele Standardauswertungen (Deskription / Grafik)
– Konfidenzintervalle
– exakte Verfahren
– nichtparametrische Verfahren

Basierend auf der Grundvorlesung „Biomathematik für Humanmedizin“ sollte die Statistiksoftware mindestens folgende Auswertungen abdecken:

- Versuchsplanung
- Deskriptive Statistik (Lage- und Streumaße, Grafiken)
- Korrelation / lineare Regression
- Überlebenszeitanalysen (Kaplan-Meier-Kurven, Log-Rank-Test)
- Konfidenzintervalle (für Mittelwert, Median, Raten)
- Tests für Häufigkeiten (Chi-Quadrat-Test)
- Tests auf Lage (t-Tests, Wilcoxon-Tests)

Darüber hinaus sollten diese Verfahren exakt sowie einige wesentliche Regressionsmodelle (multiple lineare Regression, logistische Regression, Cox-Regression) und Varianzanalysen berechnet werden können.

2 Nutzergruppen für SAS in der Medizin

In der Medizin gibt es drei typische Nutzergruppen mit unterschiedlichen Bedürfnissen und Anforderungen an die Statistiksoftware.

Nutzerprofil „Student / Doktorand“:

während des Studiums ein bis zwei Auswertungen mit Standardmethoden in übersichtlichen Datensätzen.

Erwartungshaltung:

- kurze Einarbeitungszeit
- benutzerfreundliche Oberfläche
- Übertragbarkeit auf neue Fragestellungen
- wenig spezifische Programmkenntnisse als Voraussetzung
- Nutzbarkeit auf dem „eigenen“ Rechner
- Präsentationsgrafiken
- integriertes Dateneingabesystem

Nutzerprofil „klinischer Wissenschaftler“:

neben der ärztlichen Tätigkeit wissenschaftliche Aktivitäten über längeren Zeitraum, Einsatz von Standardmethoden und darüber hinaus gehende Verfahren.

Erwartungshaltung:

- wie „Student / Doktorand“

Nutzerprofil „Biometriker / Med. Dokumentar“:

professioneller Umgang mit der Software zur Berufsausübung, Einarbeitungszeit und benutzerfreundliche Oberfläche stehen nicht im Vordergrund.

Erwartungshaltung:

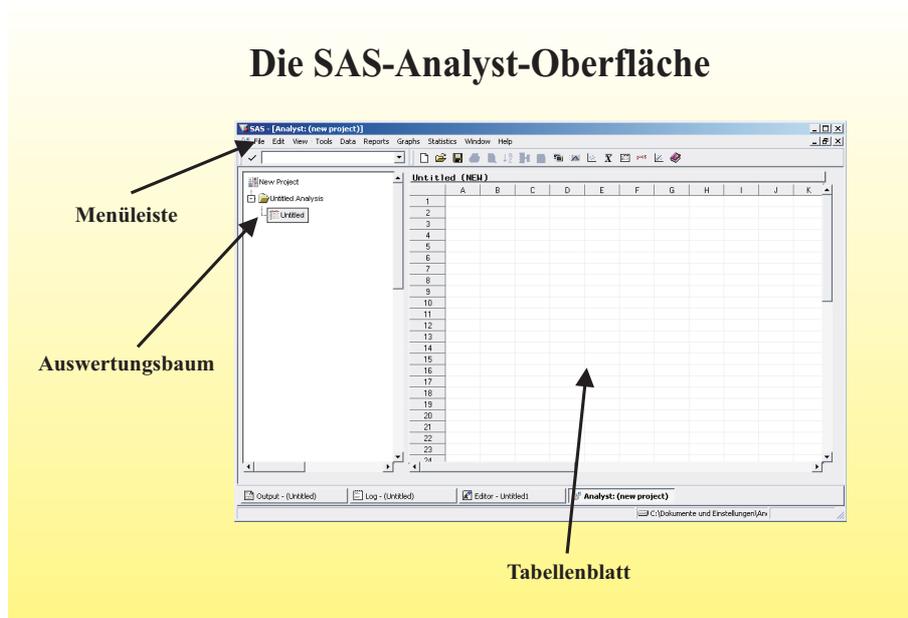
- langfristige Nutzbarkeit des Erlernten
- Programmiermöglichkeiten
- Übertragbarkeit auf neue Fragestellungen
- Dateneinlesemöglichkeiten
- Präsentationsgrafiken

3 SAS-Analyst (V8) – Kurzvorstellung

SAS-Analyst gehört zum Lieferumfang von SAS ab Version 8.0. Es wird automatisch mit installiert und kann über **Solutions** → **Analysis** → **Analyst** aufgerufen werden. Die folgende Abbildung zeigt den Aufbau der SAS-Analyst-Oberfläche.

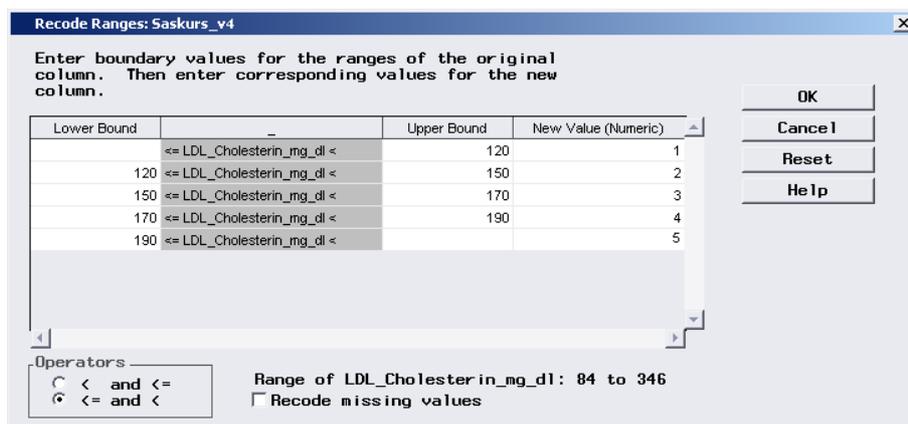
Folgende Eigenschaften besitzt die Oberfläche:

- Es ist eine maus- und menügesteuerte Oberfläche.
- Sie ist Windows angepasst.
- Es gibt Möglichkeiten zur Dateneingabe.
- Die entsprechenden SAS-Prozeduren werden per Mausklick implizit aufgerufen.
- Es gibt einige zusätzliche Statistikkösungen, die in der SAS-Oberfläche nicht enthalten sind.



4 Dateneingabe und Datenmanagement

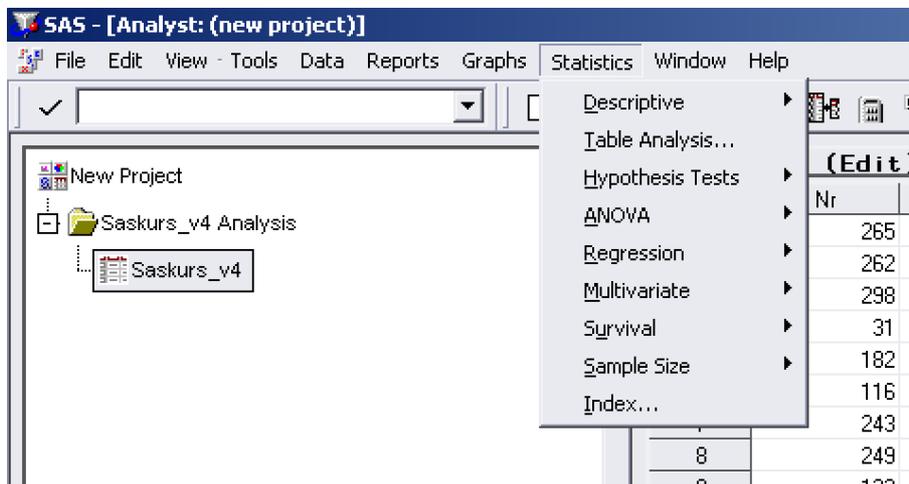
Daten können über den üblichen Import-Wizard in SAS-Analyst importiert werden. Darüber hinaus kann hier auch die Befehlsfolge **File** → **Open** zum Datenimport genutzt werden. Als einzulesende Dateiformate stehen SAS, Excel, Access, SPSS, ASCII und weitere zur Verfügung. Die direkte Dateneingabe in das Tabellenblatt in SAS-Analyst wird wegen der Unausgereiftheit des Programms nicht empfohlen (Details in [1]).



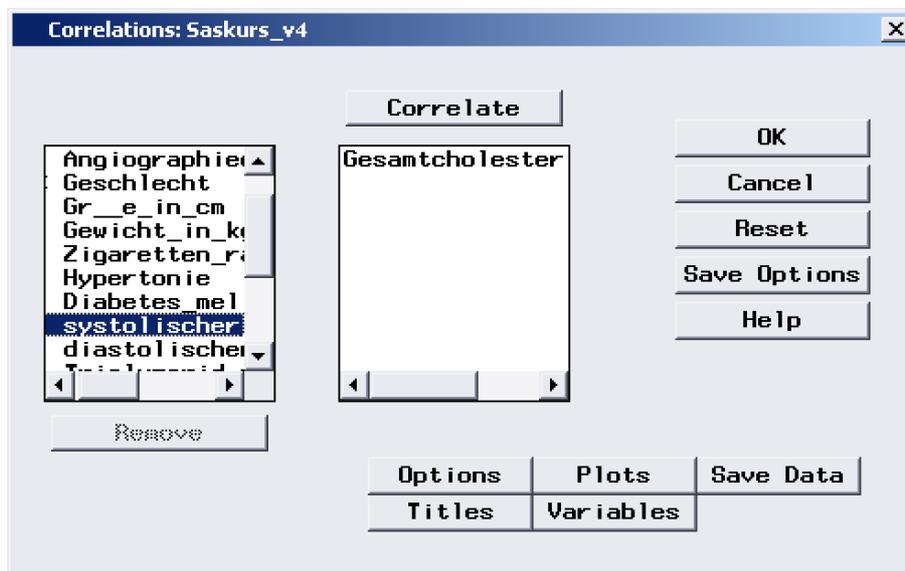
Über Menüs ist das Zusammenführen von Dateien möglich (SET/MERGE). Außerdem gibt es entsprechende Menüs zur Erzeugung neuer Variablen, Umcodierungen und Klassierungen (ein Beispiel in folgender Abbildung). Die Änderung von Variableneigenschaften ist ebenfalls menügesteuert möglich. Ein zur Verfügung stehender Datenfilter ermöglicht Untergruppenanalysen.

5 Auswertungsmöglichkeiten

Die folgende Abbildung zeigt die Auswertungsmöglichkeiten in SAS-Analyst übersichtsmäßig auf:



Die entsprechenden Auswahlmenüs sind folgendermaßen gestaltet und zeigen den üblichen Aufbau maus- und menügesteuerter Statistiksoftware (hier gezeigt am Beispiel der Korrelationsanalyseauswahl):



Folgende Auswertungsmöglichkeiten (mit Angabe der entsprechenden SAS-Prozeduren) gibt es:

- Deskriptive Statistik: PROC MEANS, PROC UNIVARIATE, PROC FREQ, PROC CORR
- Überlebenszeitanalyse: PROC LIFETEST, PROC PHREG
- Exakte Tests: Exakter Fisher Test (PROC FREQ)
unverb. Wilcoxon-Test (PROC NPAR1WAY)
verb. Wilcoxon-Test (PROC UNIVARIATE)
- Nichtparam. Tests: PROC NPAR1WAY
- Regressionsanalysen: PROC REG, PROC LOGISTIC, PROC PHREG
- Varianzanalysen: PROC ANOVA, PROC GLM
- Allgemeinere Modelle: PROC MIXED
- multivariate Analysen: PROC PRINCOMP, PROC CANCORR

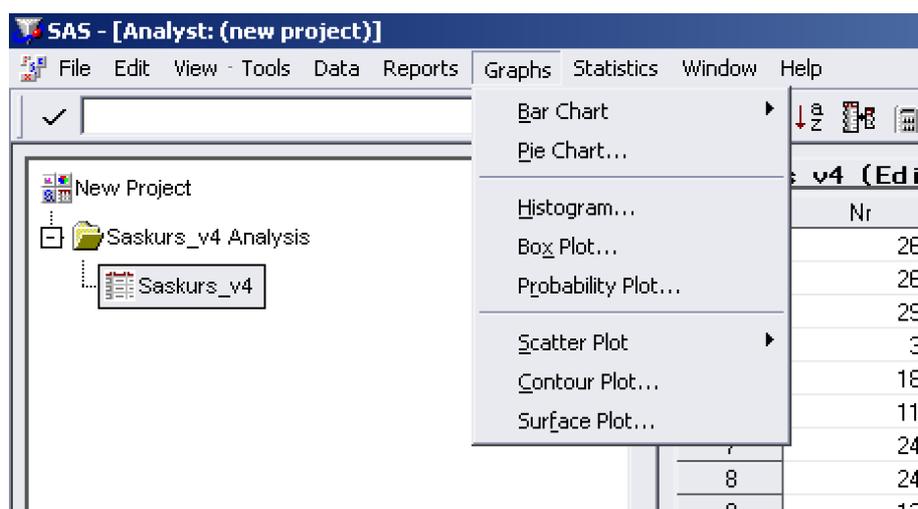
Insgesamt kann man aber über die SAS-Analyst – Oberfläche nicht alle vorhandenen SAS-Prozeduren aufrufen.

Folgende Erweiterungen gegenüber der Syntax-Ebene wurden implementiert:

- Fallzahlplanungsmöglichkeiten für die t-Test-Situation
- Power-Analysen
- Erzeugung von Source-Code

6 Grafikmöglichkeiten

In SAS-Analyst lassen sich verschiedene Grafikmöglichkeiten aufrufen (s. Abbildung und Aufzählung):



- Bar Chart und Pie Chart: PROC GCHART
- Box-Plot: PROC BOXPLOT
- Histogramm, Probability Plot: PROC UNIVARIATE
- Scatter Plot: PROC GPLOT
- Contour Plot: PROC GCONTOUR
- Surface Plot: PROC G3D

Weitere Grafikmöglichkeiten existieren in den Statistik-Menüs (z.B. für Power-Analysen).

Globale Grafikeinstellungen können im Menü **Tools** → **Graph Settings** festgelegt werden. Eine anschließende Grafikbearbeitung kann durchgeführt werden. Leider kann über SAS-Analyst auch nur der übliche SAS-Grafikeditor benutzt werden, der im Leistungsumfang und in der Bedienung doch nicht mehr dem heutigen Standard von Grafikbearbeitungsprogrammen entspricht. Ein Abspeichern der Grafiken in einem der üblichen Grafikformate (GIF, JPG, BMP, Postscript) ist möglich.

7 Daten- und Ergebnisausgabe

Unter dem Menüpunkt **Reports** gibt es die Möglichkeit, Daten auszudrucken. Dabei wird PROC PRINT zum Datenlisting benutzt. Ebenfalls stehen hier fünf verschiedene Tabellenstrukturen zur Verfügung, um Ergebnisse zu präsentieren. SAS-Analyst bedient sich hier der Prozedur PROC TABULATE. Man ist auf die fünf vorgegebenen Tabellenstrukturen festgelegt, aber für die meisten Anwendungen werden diese ausreichen.

Als ein Produkt der SAS-Version 8 lassen sich Ergebnisse, die mit SAS-Analyst erstellt werden, über das Output Delivery System (**ODS**) präsentabler darstellen. Es gibt die Möglichkeit der Ausgabe als übliches SAS-Listing oder im HTML-Format. Weitere Formate, die ab SAS-Version 8.1 dazugekommen sind, wie z.B. RTF, können allerdings nicht benutzt werden. Die Einstellung des Output-Stils lässt sich unter **Tools** → **Options** → **Preferences** → **Results** vornehmen.

8 Informationsmöglichkeiten

Für die SAS-Analyst – Oberfläche stehen ab der Version 8 einige Informationsmöglichkeiten zur Verfügung. Diese sind gerade für eher ungeübte Nutzer sehr wichtig, um bei Bedarf entsprechende Hilfemöglichkeiten zu haben.

Online-Hilfe in SAS-V8:

In der Online-Hilfe „SAS System Help“ gibt es einen entsprechenden Abschnitt mit Beschreibungen der SAS-Analyst – Oberfläche. Aufzurufen ist diese über: **HELP → SAS System Help → Using SAS Solutions and Tools → Using the ANALYST application for statistic tasks.**

Über die HELP-Buttons in den Statistik-Menüs werden z.B. auch direkt die entsprechenden Kapitel aus dieser Hilfe aufgerufen.

SAS OnlineDoc V.8 CD:

In der mit der SAS-Version ausgelieferten CD „SAS-OnlineDoc“ gibt es ein englisch-sprachiges Handbuch zur Nutzung von SAS-Analyst. Dieses Dokument kann über das Auswahlmenü oder direkt aufgerufen werden als **SASDOC/SASHTML/Analyst/Index.htm**

SAS-Analyst im Internet:

Auf den Internet-Seiten von SAS finden sich Informationen zu SAS-Analyst unter:

<http://www.sas.com/rnd/app/da/analyst.html> (aufgerufen 02/02).

Sekundärliteratur und Kurse:

Die Autoren haben ein deutsch-sprachiges Lehrbuch zur Nutzung von SAS-Analyst im Bereich der Medizinstatistik geschrieben [1]. In etwa monatlichen Abständen werden Kurse zur „Einführung in SAS-Analyst“ an der Universität Ulm von der Abteilung Biometrie angeboten. Bei Interesse und freien Kapazitäten können Auswertige an den Kursen teilnehmen (für Anfragen s. Adresse der Autoren).

9 Zusammenfassung

Zusammenfassend kann man folgende **Vorteile** der Nutzung von SAS-Analyst für die Zielgruppen Studenten und klinische Wissenschaftler festhalten:

- Die Oberfläche ist maus- und menügesteuert und Windows angepasst. Dadurch gibt es, da Vorkenntnisse aus dem Office-Paket meist vorhanden sind, kaum Probleme beim Erlernen der Nutzeroberfläche.
- Das Einlesen externer Daten ist problemlos möglich.
- Die Datenmanagement- / Statistik- und Grafikmöglichkeiten sind für die Mehrzahl der Anwendungen sicher ausreichend.
- Entsprechende Hilfen und Nachschlagewerke sind vorhanden, so dass die Nutzer auch im Eigenstudium die Software nutzen können.
- In der Oberfläche wird der SAS-Source-Code erzeugt, so dass mittelfristig die Nutzer auf die SAS-Syntax-Ebene wechseln können.

Folgende **Nachteile** der Nutzung von SAS-Analyst für die Zielgruppen Studenten und klinische Wissenschaftler meinen wir festhalten zu können:

- Das Tabellenblatt ist für Dateneingaben nicht einfach zu handhaben. Wir empfehlen deshalb eine externe Dateneingabe. Dieser Umstand ist deshalb ärgerlich, da ungeübte Nutzer nach Aufruf einer solchen Software zuerst das Tabellenblatt nutzen wollen und hier auf Schwierigkeiten stoßen. Dieser Eindruck wird dann leicht auf das gesamte Produkt übertragen und führt zur Ablehnung.
- Die Auswahl der multivariaten Methoden ist nicht nachvollziehbar.
- Der Grafikeditor und die Ergebnisausgabe (fehlende RTF-Ausgabe) haben nur beschränkte Möglichkeiten und führen dazu, dass weitere Bearbeitungssoftware notwendig wird, wenn die Ergebnisse präsentiert werden sollen.

Resumee:

Wir finden, dass SAS-Analyst als ein ausgezeichnetes Tool mit kleinen Schwächen für Gelegenheitsnutzer in der medizinischen Forschung eingesetzt werden kann. Dies gilt vor allem für Forschung im Umfeld der Universitätskliniken.

Erweiterungsmöglichkeiten sehen wir, neben den Verbesserungen der Grafik- und Ausgabemöglichkeiten, darin, dass die Oberfläche für alle SAS-Prozeduren zur Verfügung gestellt werden könnte. Zusätzlich könnte man SAS-Analyst parallel als abgespeckte SAS-Version für Studenten und Gelegenheitsnutzer anbieten mit entsprechend bezahlbarem Preis.

Literatur

- [1] Muche, R., Habel, A., Rohlmann, F. (2000). Medizinische Statistik mit SAS-Analyst. Springer-Verlag, Heidelberg.

