
Eine Verallgemeinerung des Cochran-Armitage-Tests auf mehrere Gruppen mit PROC LOGISTIC und PROC GENMOD

Oliver Kuß

Institut für Medizinische Epidemiologie, Biometrie und
Informatik

Universität Halle-Wittenberg

Oliver.Kuss@medizin.uni-halle.de

Heinrich Dickel

Abteilung Klinische Sozialmedizin

Universitätsklinikum Heidelberg

Heinrich_Dickel@med.uni-heidelberg.de

6. Konferenz für SAS[®]-Anwender
in Forschung und Entwicklung (KSFE)

28.2. - 1.3. 2002

Universität Dortmund

Programm

1. Das inhaltliche Problem

2. Trendtest für Raten: Der Cochran-Armitage-Test

3. Eine Verallgemeinerung des Cochran-Armitage-Tests auf mehrere Gruppen

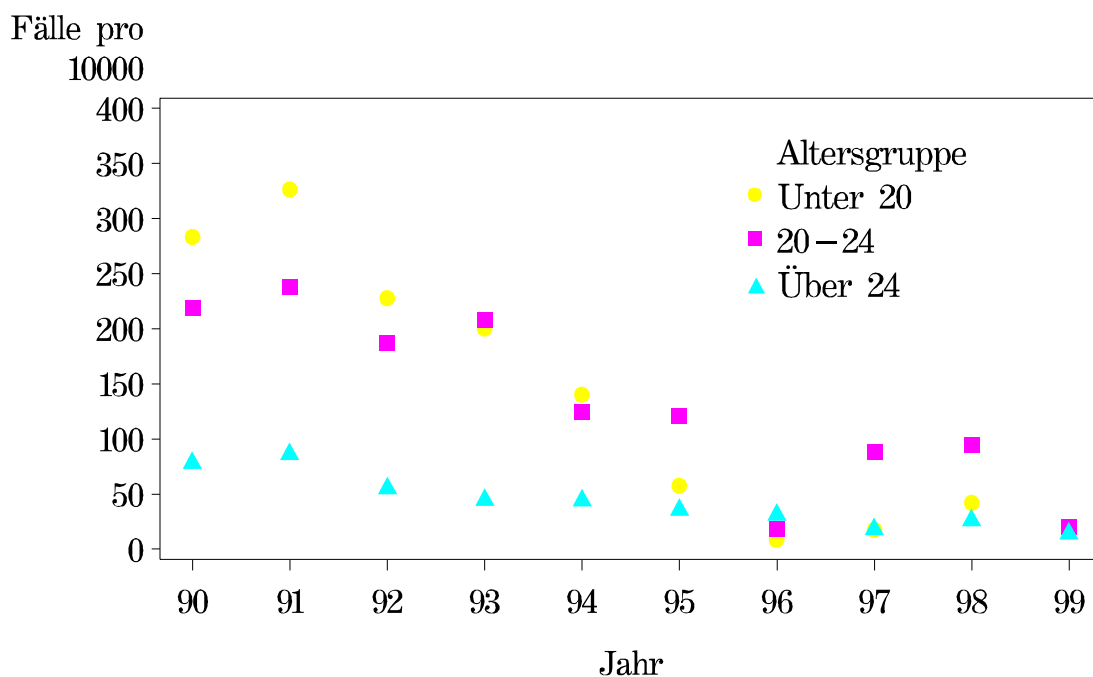
4. Diskussion/Ausblick

5. Literatur

1. Das inhaltliche Problem

BKH-N (Berufskrankheitenregister Haut Nordbayern):
Registrierung aller Verdachtsmeldungen auf das
Vorliegen einer berufsbedingten Hautkrankheit, 1990-
99

Jährliche beobachtete Inzidenzraten von Friseuren, nach
Altersgruppen getrennt



Klar abfallende Trends, aber auch unterschiedliche
Stärke???

Daten in SAS®:

```
data friseure;
  input year ageclass response count;
  cards;
  90 1 1 130
  90 1 0 4458
  90 2 1 71
  90 2 0 3166
  90 3 1 35
  90 3 0 4281
  . . .
;run;
```

2. Trendtest für Raten: Der Cochran-Armitage-Test

- Seit Version 6.12 in SAS[®] (TREND-Option im TABLES-Statement von PROC FREQ)

```
proc freq data=friseure;  
    tables year*response / trend;  
    weight count;  
    by ageclass;  
run;
```

- Viele gute statistische Eigenschaften, vor allem Alternative nicht notwendig linear
- Alternative Berechnung in SAS[®]: EXACT-Statement in PROC FREQ, PROC MULTTEST, PROC CORR
- Ergebnis für Gruppe der unter 20-Jährigen:

```
Cochran-Armitage Trend Test  
-----  
Statistic (Z)          11.3026  
One-sided Pr > Z      <.0001  
Two-sided Pr > |Z|    <.0001
```

3. Eine Verallgemeinerung des Cochran-Armitage-Test auf mehrere Gruppen

Tip auf SAS-Homepage:
Cochran-Armitage-Test ist äquivalent zu Score-Test im
logistischen Regressionmodell:

Überprüfung:

```
proc logistic data=friseure descending;  
  freq count;  
  model response=year;  
  by ageclass;  
run;
```

Output:

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
LR	159.7671	1	<.0001
Score	127.7480	1	<.0001
Wald	112.4861	1	<.0001

Idee: Nutze diese Äquivalenz und formuliere das Problem im Kontext der logistischen Regressionsmodelle

Konkret: Nimm Gruppenvariable (hier: Altersklasse) und deren Interaktionsterm mit der Klassenvariable (hier: Jahr) ins Modell auf.

PROC LOGISTIC und PROC GENMOD liefern zwei unterschiedliche Tests:

Wald-Test:

```
proc logistic data=friseure descending;
    freq count;
    class ageclass / param=glm;
    model response=ageclass year
            year*ageclass;
run;
```

Output:

```
Type III Analysis of Effects
                    Wald
Effect              DF      Chi-Square Pr > ChiSq

ageclass            2        13.0413    0.0015
year                 1        164.2423   <.0001
year*ageclass       2        11.6006    0.0030
```

Likelihood Ratio-Test:

```
proc genmod data=friseure descending;  
    freq count;  
    class ageclass;  
    model response=ageclass year  
            year*ageclass /  
            d=bin link=logit type3;  
run;
```

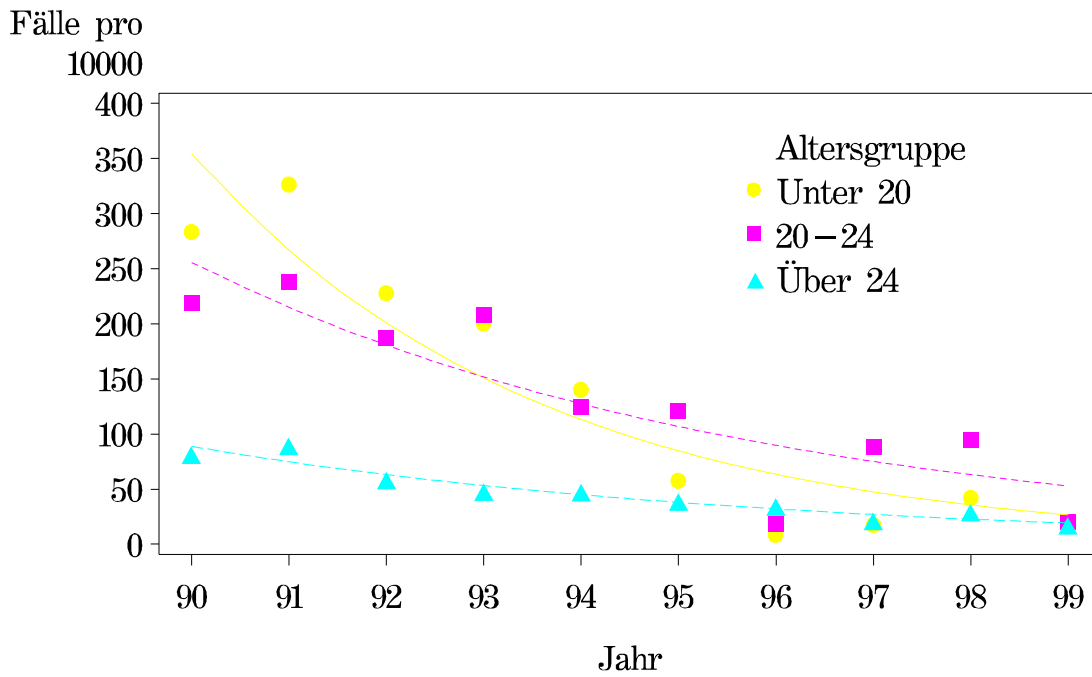
Output:

```
LR Statistics For Type 3 Analysis  
Chi-  
Source          DF   Square    Pr > ChiSq  
ageclass        2    13.48     0.0012  
year            1    200.14    <.0001  
year*ageclass   2    12.01     0.0025
```

Inhaltlich spannender: Paarweiser Vergleich

Vergleich	p-Wert
< 20 – 20-24	0.003
< 20 – >24	0.004
20-24 – >24	0.875

4. Diskussion/Ausblick



- Durch Interpretation des Cochran-Armitage-Tests im Kontext der logistischen Regressionsmodelle kann dieser leicht auf den Vergleich von Trends in mehreren Gruppen verallgemeinert werden
- Schätzung mit PROC LOGISTIC oder PROC GENMOD
- Statistische Eigenschaften unklar
→ Simulationsstudie

5. Literatur

- Dickel, H., Kuss, O., Blesius, C.R., Schmidt, A., Diepgen, T.L. (2001). Occupational skin diseases in Northern Bavaria between 1990 and 1999: a population-based study. *British Journal of Dermatology*, **145**, 453-462.
- Dickel, H., Kuss, O., Schmidt, A., Diepgen, T.L. (2001). Inzidenz berufsbedingter Hautkrankheiten in hautgefährdenden Berufsordnungsgruppen. *Der Hautarzt*, **52**, 615-623.
- Cochran, W.G. (1954). Some methods of strengthening the common χ^2 tests. *Biometrics*, **10**, 417-451.
- Armitage, P. (1955). Tests for linear trends in proportions and frequencies. *Biometrics*, **11**, 375-386.
- Williams, P.L. (1997). Trend tests for counts and proportions. In: Armitage, P., and Colton, T., [eds.] *Encyclopedia of Biostatistics*, Section 10: Categorical Data Analysis.
- Tarone, R.E., Gart, J.J. (1980). On the robustness of combined tests for trends in proportions. *Journal of the American Statistical Association*, **75**, 110-116.
- Cox, D.R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society, Series B*, **205**, 215-242.
- Margolin, B.H. (1988). Test for trend in proportions. In: Kotz, S., Johnson, N.L., Read, C.B., [eds.] *Encyclopedia of Statistical Sciences*, **9**, 334-336.
- SAS Institute Inc. (2001). Documentation Example 1 for PROC MULTTEST. http://ftp.sas.com/techsup/download/sample/samp_lib/statsampDocumentation_Example_1_00000034.html.
- SAS Institute Inc. (2001). How do I compute a trend test for proportions? http://www.sas.com/service/techsup/faq/stat_proc/multproc981.html.
- Agresti, A. (1990). *Categorical data analysis*. John Wiley & Sons, New York.
- Chen, J.J. (1993). Trend test for overdispersed proportions. *Biometrical Journal*, **35**, 949-958.
- Lloyd, C.J. (1999). *Statistical Analysis of Categorical Data*. John Wiley & Sons, New York.