Statistical modelling of exposure-response relationships in the analysis of multiple transportation noises

Sonja Kuhnt[1], Christoph Schürmann[2]

[1] *Universität Dortmund, Fachbereich Statistik, D-44227 Dortmund, Germany, Email: kuhnt@statistik.uni-dortmund.de*

[2] *Universität Dortmund, Fachbereich Statistik, D-44227 Dortmund, Germany, Email: schuermann@statistik.uni-dortmund.de*

Introduction

Being part of the research network 'Quiet Traffic', section 'Noise Effects', the project group 'Assessment of annoyance for combined sources' is interested in how combined noise sources of both road and railway traffic affect the annoyance and subjectively sensed loudness of individuals. Preceding the main study, an experiment has been carried out to develop suitable statistical models for the relationship of sound levels and annoyance and loudness, respectively. Additional covariates such as gender, time of day of exposure and laboratory are also considered.

Experimental design

The experiment was designed such that three laboratories in different cities of Germany each carried out the same design. In each laboratory, 24 subjects were exposed to both railway and road traffic noise. For each kind of noise, sound levels of 40, 52, 70, 82 dB were presented twice, and levels of 46, 58, 64 and 76 dB were presented once, all of them in random order. After each sound presentation the subjects recorded their annoyance and perceived loudness on a category subdivided rating scale (see [1]). One half of the trials was carried out in the morning, the other one in the afternoon. Each group consisted of the same number of males and females.

Methods

Regression model

A subject's response was measured as a natural number ranging between 1 and 50 and can be viewed as a random variable with binomial distribution $I\!B(n, \pi)$, where $n = 49$ and π is unknown. This allows the application of a generalized linear model (GLM) that extends the regular linear regression to the case of some non-normal response variables [2]. Here, a GLM models the exposure-response relationship between covariates X_j, $j = 1, \ldots, p$ and annoyance (or loudness). Multiplying a given value $X_j = x_j$ with its parameter β_j and summing up for all j gives the so-called predictor variable

$$\eta = \beta_0 + \sum_{j=1}^{p} \beta_j x_j. \tag{1}$$

The predictor is linked to the expectation value $E(Y) = \pi$ of Y by a certain link function g. In this situation, the so-called complementary log-log link is a suitable choice (see [3]). It allows us to compute the expected annoyance (or loudness) $E(Y) = g^{-1}(\eta)$ for a given predictor η by

$$E(Y) = (1 - exp(-exp(\eta))) \cdot 49 + 1. \tag{2}$$

The estimation of the parameters β_j by the maximum likelihood method then shows, how they affect the response and if their effect is significant.

If our model is valid for the data, the observations should have variance $Var(Y) = 49 \cdot \pi \cdot (1 - \pi)$. However, it is possible that the actual variance is ϕ-times greater than the theoretical one, which is called overdispersion.

Model selection

We aim at a model equation that explains the observed data well, but is also as sparse as possible. This can be achieved by performing a backward directed stepwise procedure: Starting with a regression equation that includes all factors (main effects and all interactions), we stepwise remove those factors having the largest p-values as long as they are larger than the significance level of $\alpha = 0.05$. Main effects are only removed if all their interactions have been removed before. An intercept $x_0 = 1$ is always included.

Coding of variables

The factors are coded as presented in Table 1. Interactions are not listed as they are simply the product of the codings of the respective covariates. For example, the levels of the interaction of gender and kind of noise $x_{G,K}$ equal $x_G \cdot x_K$. In fact, the experiment has been performed in the laboratories at three different sites, but for illustrative purposes we need only involve two of them in our analysis: So, $x_L = -1$ and $x_L = 1$ refer to the Eichstätt and the Dortmund laboratory, respectively.

Covariate	Factor	Coded levels
Noise level	x_N	40, 46, ..., 82
Kind of Noise	x_K	-1 (rail), +1 (road)
Laboratory	x_L	-1 (lab 1), +1 (lab 2)
Gender	x_G	-1 (male), +1 (female)
Time of day	x_T	-1 (morning), +1 (afternoon)

Table 1: Coding of factor levels.

Results

Annoyance

Table 2 presents the factors x_j of the final regression equation for annoyance together with the parameter estimates $\hat{\beta}_j$ and p-values. Besides, the overdispersion parameter is estimated by $\hat{\phi} = 4.25$.

x_j	$\hat{\beta}_j$	p-value
$x_0 \equiv 1$	-2.518	< 0.0001
x_N	0.038	< 0.0001
x_K	-0.054	< 0.0001
x_L	0.189	0.0012
x_G	-0.228	< 0.0001
x_T	0.061	< 0.0001
$x_{G,N}$	0.003	0.0033
$x_{G,L}$	0.028	0.0226
$x_{G,T}$	-0.030	0.0138
$x_{L,N}$	-0.002	0.0471

Table 2: Regression results for annoyance: Resulting factors and their estimated parameters.

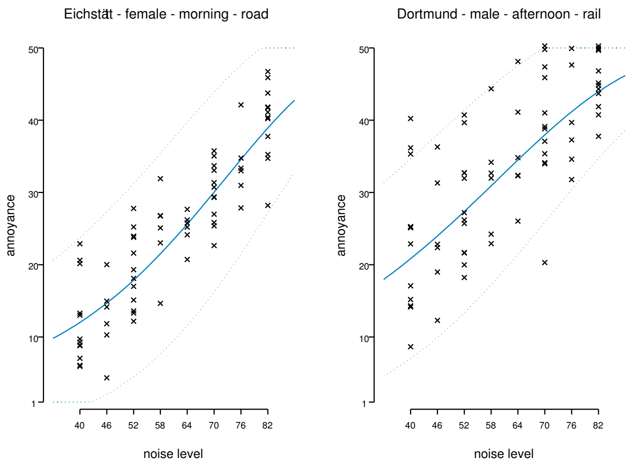

Figure 1: Observed (asterisks) and estimated annoyance (solid lines) with 95% confidence intervals (dotted lines) for two strata.

As an illustration, Figure 1 shows estimated and observed annoyance for varying noise levels for those two strata that make for the most extreme values. A covariate's relevance may best be interpreted if compared to the effect of noise level: Here, for example, the parameter estimate for kind of noise is $\hat{\beta}_K = -0.054$. So, if the kind of noise changes from -1 to $+1$, the predictor augments by 0.108 which is the same as if the noise level was increased by 2.84 dB. Thus, to yield equal annoyance, road noise has to be about 2.84 dB louder than rail noise.

Loudness

Loudness and annoyance are highly correlated ($\rho = 0.943$). It is thus not surprising that the loudness' final model equation contains nearly the same factors and yields similar parameter estimates (Table 3). Here, the estimated overdispersion parameter is $\hat{\phi} = 3.08$. Parameter estimates can be interpreted as for annoyance (see above).

Figure 2 shows observed and estimated values for the perceived loudness. The strata with the highest and smallest estimated ratings are the same as for annoyance. A closer look at both graphs reveals that for any noise level loudness is rated slightly higher than annoyance.

x_j	$\hat{\beta}_j$	p-value
$x_0 \equiv 1$	-2.618	< 0.0001
x_N	0.041	< 0.0001
x_K	-0.052	< 0.0001
x_L	0.050	< 0.0001
x_G	-0.167	0.0009
x_T	0.059	< 0.0001
$x_{G,N}$	0.002	0.0158
$x_{G,L}$	0.022	0.0356
$x_{G,T}$	-0.057	< 0.0001
$x_{L,T}$	0.054	< 0.0001

Table 3: Regression results for loudness: Resulting factors and their estimated parameters.

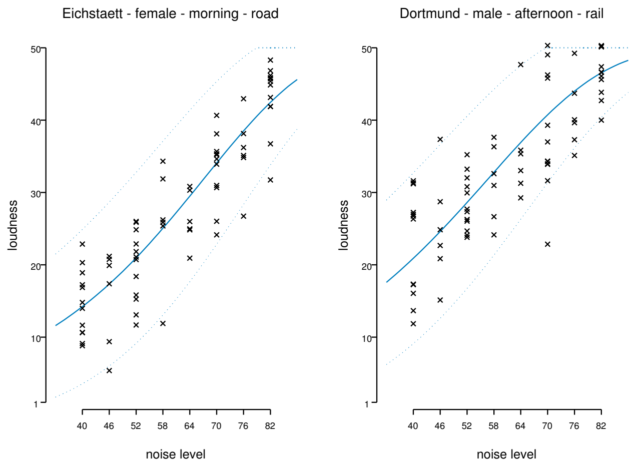

Figure 2: Observed (asterisks) and estimated loudness (solid lines) with 95% confidence intervals (dotted lines) for two strata.

Acknowledgements

The support of the German Ministry of Education and Research (BMBF) is gratefully acknowledged.

References

[1] Hellbrück, J. Category-Subdivision Scaling - A Powerful Tool in Audiometry and Noise Assessment. In: Fastl H, Kuwano S, Schick A, eds. Recent trends in research. Festschrift for Seiichiro Namba. Oldenburg: BIS, 317-336, 1996.

[2] McCullagh P, Nelder JA. Generalized Linear Models. Chapman & Hall, London, 1989.

[3] Kuhnt S, Schürmann Ch, Griefahn B. Annoyance from Multiple Transportation Noise: Statistical Models and Outlier Detection. Submitted for publication.