

Fitting the Distribution of Inpatients' Hospital Days by ICD Categories

Hyejung Chang

Korea Health Industry Development Institute, Department of Health Industry Informatics

57-1 Nolyangjin-Dong Dongjac-Gu, Seoul, Korea

e-mail : hjchang@khidi.or.kr

Maengseok Noh

Korea Health Industry Development Institute, Department of Health Industry Informatics

57-1 Nolyangjin-Dong Dongjac-Gu, Seoul, Korea

e-mail : msnoh@khidi.or.kr

1. INTRODUCTION

This study focused on modeling the probability distribution of inpatients' hospital days. With the date of admission as an entry point and the discharge as an event, patients' lengths of stay (LOS) can be considered as the survival data from the entry to an event occurrence. Therefore, the survival function was fitted to the survey data of patients discharged from acute hospitals for one month. The LOS of all 279,089 patients were measured in days and analyzed.

The LOS is operationally defined as the period from the admission to the discharge permitted by medical doctors, for patients cured or treated with optimistic progress. Seven percent of censored patients were included in the sample of this study. They were patients discharged without permission by medical doctors such as discharged cases against medical advice, transferred or escaped cases, etc.

2. MODEL

The Proportional Hazard Model (PHM) is specified by the proportional relationship of hazard functions for each fixed effect (Cox, 1972). The fitting procedures are as follows: 1) examining the proportionality assumption for the survival hazard function by disease groups; and 2) fitting the survey data by strata using the PHM.

Although the proportionality is expected in many situations, specific levels of an important factor may produce hazard functions which differ markedly from the proportionality. The proportionality assumption is assessed by disease groups and covariates, using the figure of Kaplan-Meier estimators (Kaplan & Meier, 1958).

When the assumption is violated, we should consider a hazard function for an individual with covariates x in the j th stratum shown in the model (1) (Kalbfleisch & Prentice, 1980).

$$I_j(t | x, u) = I_{0j}(t) \exp(x^t \mathbf{b}), \quad j=1, \dots, M. \quad (1)$$

Age, sex, location of hospital, type of discharge, admission pathway, and emergency status were considered as covariates related to the distribution of LOS.

3. RESULTS

Figure 1 shows the relationship of a log minus log plot of estimated survival functions on log-transformed LOS for five grouped ICD categories. Since the curve of empirical survival functions are crossed each other, the assumption of proportionality turned out to be violated. However, the survival functions for all covariates followed the proportionality assumption.

Therefore, the model (1) with a stratum of disease group is fitted to the data of inpatients' hospital days. The significant impacts of covariates were obtained by Chi-squared tests. Using parameter estimates, the survival curves adjusted by other covariates were explored and explained by different disease groups.

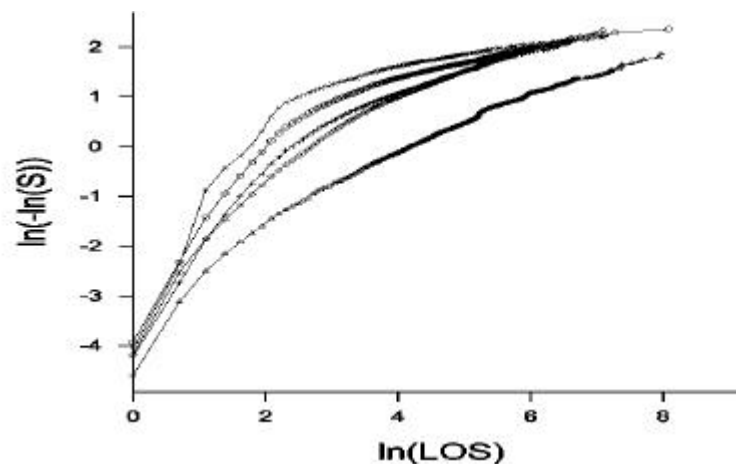


Fig. 1. Log minus log plot of estimated survival functions versus log LOS by ICD categories

4. CONCLUSION

This article introduced the distribution of inpatients' LOS in the context of the survival function. The distribution did not follow the proportionality assumption by disease groups, and then the hazard function with a stratum was considered and estimated. The estimated distribution function of inpatients' hospital stays can be utilized in various health services management circumstances such as the prediction of healthcare demands by service types or the policy planning on the supply of health resources.

REFERENCES

- Cox, D. R. (1972). Regression models and life tables. *Journal of Royal Statistical Society B*, **34**, 187-202.
- Kalbfleisch, J. D. & Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*. New York : Wiley.
- Kaplan, E.L. & Meier, P. (1958). Non-parametric estimation from incomplete observation. *Journal of the American Statistical Association*, **53**, 457-481.