2차원 품질보증데이터 모델링

백재욱*, 조진남**

한국방송통신대학교 정보통계학과*

동덕여자대학교 데이터정보학과**

Two-Dimensional Warranty Data Modelling

Jai Wook Baik*, Jin Nam Jo**

Department of Information Statistics, Korea National Open University*

Department of Data & Information Science, Dongduk Women's University**

Key words : Two-dimensional warranty data; Bivariate model; Marked point process; One-dimensional point process.

Abstract

Two-dimensional warranty data can be modelled using two different approaches: two-dimensional point process and one-dimensional point process with usage as a function of age. The first approach has three different models. First of all, bivariate model is appealing but is not appropriate for explaining warranty claims. Next, the rest of the two models (marked point process, and counting and matching on both directions independently) are more appropriate for explaining warranty claims. However, the second one (counting and matching on both directions independently) assumes that the two variables (variables representing the two-dimensions) are independent. Last of all, one-dimensional point process with usage as a function of age is also promising to explain the two-dimensional warranty claims. But the models or variations of them need more investigation to be applicable to real warranty claim data.

1. Introduction

The case of one-dimensional warranties has been a lot of attention. See Kalbfleisch, Lawless and Robinson (1991), Lawless and Kalbfleisch (1992), Lawless (1998) and Kalbfleisch and Lawless (1998). In contrast, the analysis and modelling of two-dimensional warranty data has received very little attention and in this paper we look at this issue.

See Iskandar (1993), Murthy, Iskandar and Wilson (1995), Singpurwalla and Wilson (1998), Kim and Rao (2000) and Yang and Nachlas (2001) for two-dimensional warranty data analysis.

Item failures are points on a plane with one axis representing age (t) and the other representing usage (x). Two different approaches have been employed to modelling item failures. The first is to model item failure by a two-dimensional point process formulation. See Iskandar (1993), Murthy, Iskandar and Wilson (1995) and Blischke and Murthy (1996, Chapter 7). The second approach involves modelling usage as a function of time so that failures are effectively modelled by a one-dimensional point process formulation. Iskandar (1993) suggests a linear model for usage given by $x(t) = \Lambda t$ where Λ is the usage rate and is modelled as a random variable to model the varying usage across the consumer population. We will elaborate on these two approaches. We will look at 3 different models in the first approach and at the second approach and consider how to get the number of claims within warranty limits on age and usage.

2. Bivariate modeling

We assume that (T, X) is a non-negative bivariate random variable with distribution function,

$$F(u,v) = \mathrm{P}\{T \le u, X \le v\}; u \ge 0, v \ge 0.$$

The survivor function is given by

$$\bar{F}(u,v) = \Pr\{T > u, X > v\} = \int_u^\infty \int_v^\infty f(s,t)dtds.$$

and, if $F(u,v)$ is differentiable, then the bivariate failure density function is given by

$$f(u,v) = \frac{\partial^2 F(u,v)}{\partial v \partial u}.$$

The hazard function is defined as

$$r(u,v) = f(u,v) / \bar{F}(u,v).$$

Then Figure 1 shows a risk set at age t and usage x.

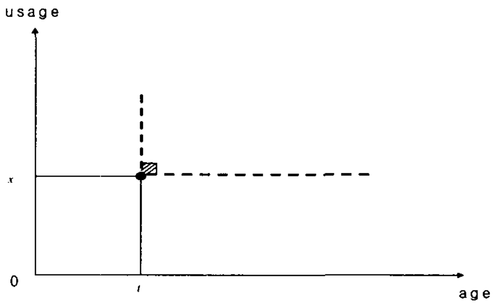

Figure 1. Risk set at (t,x)

If we let $N(W,U)$ be the number of claims within $[0,W) \times [0,U)$ then

$$P\{N(W,U) = n\}$$

$$= \frac{\exp[-\Lambda(W,U)]\Lambda(W,U)^n}{n!}, \ n \ge 0 \qquad (1)$$

where $\Lambda(t,x) = \int_0^t \int_0^x \lambda(u,v)dvdu$.

Note that $E[N(W,U)] = \Lambda(W,U)$.

Example 1: Bivariate Weibull model

Consider Lu and Bhattacharyya's bivariate Weibull model (see Lu and Bhattacharyya (1990)) where the survivor function is

$$\bar{F}(t,x) = \exp\{-[(t/\theta_1)^{\beta_1/\delta} + (x/\theta_2)^{\beta_2/\delta}]^\delta \qquad (2)$$

with $\theta_1, \theta_2, \beta_1, \beta_2 > 0$ and $0 < \delta \le 1$.

Therefore, hazard function is

$$\lambda(t,x)=\alpha_1^{-1}\alpha_2^{-1}\beta_1\beta_2(t/\alpha_1)^{\beta_1/\delta-1}(x/\alpha_2)^{\beta_2/\delta-1}[(t/\alpha_1)^{\beta_1/\delta}+(x/\alpha_2)^{\beta_2/\delta}]^{\delta-2}$$
$$\times\{[(t/\alpha_1)^{\beta_1/\delta}+(x/\alpha_2)^{\beta_2/\delta}]^{\delta}+1/\delta-1\}. \quad (3)$$

Suppose that $\theta_1=2, \beta_1=1.5, \theta_2=3, \beta_2=2.0, \delta=0.5.$

Then $E(T)=\theta_1\Gamma(1/\beta_1+1)=1.81$ (years), $E(X)=\theta_2\Gamma(1/\beta_2+1)=2.66$ (10^3 Km) and $E(N(5,10])=\Lambda(5,10)=15.588$.

3. Marked point process

3.1 In case X_i~NHPP with $\lambda_U(x)$

First we assume that T_i follows a non-homogeneous Poisson process (NHPP) with $\lambda_W(t)$. Therefore, if we let $N(W)$ be the number of failures until warranty period W then

$$P\{N(W)=n\}$$
$$=\frac{\exp[-\Lambda_W(W)]\Lambda_W(W)^n}{n!},\ n\geq 0 \quad (4)$$

where $\Lambda_W(t)=\int_0^t\lambda_W(u)du$.

Now

$$P\{N(W,U)=k\}$$
$$=\sum_{j=k}^{\infty}P\{N(W,U)=k\mid N(W)=j\}P\{N(W)=j\}$$
$$=P\{N(W,U)=k\mid N(W)=k\}P\{N(W)=k\}+\sum_{j=k+1}^{\infty}P\{N(W,U)=k\mid N(W)=j\}P\{N(W)=j\}$$
$$=P\{X_k<U\}P\{N(W)=k\}+\sum_{j=k+1}^{\infty}P\{N(W,U)=k\mid N(W)=j\}P\{N(W)=j\},$$

see Figure 2

$$=P\{X_k<U\}\ P\{N(W)=k\}+P\{X_k<U,X_{k+1}>U\}\sum_{j=k+1}^{\infty}P\{N(W)=j\}$$
$$=(1-\sum_{i=0}^{k-1}P\{N(U)=i\})\frac{e^{-\Lambda_W(W)}\Lambda_W(W)^k}{k!}+P\{N(U)=k\}\ (1-\sum_{i=0}^{k}\frac{e^{-\Lambda_W(W)}\Lambda_W(W)^i}{i!})$$
$$=(1-\sum_{i=0}^{k-1}\frac{e^{-\Lambda_U(U)}\Lambda_U(U)^i}{i!})\frac{e^{-\Lambda_W(W)}\Lambda_W(W)^k}{k!}+\frac{e^{-\Lambda_U(U)}\Lambda_U(U)^k}{k!}(1-\sum_{i=0}^{k}\frac{e^{-\Lambda_W(W)}\Lambda_W(W)^i}{i!}) \quad (5)$$

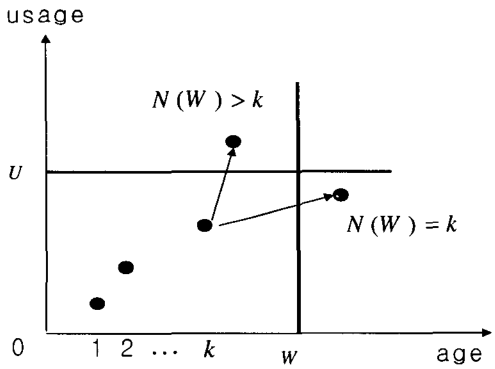

Figure 2. In case $N(W,U)=k$

Example 2: Marked point process (X_i~NHPP with $\lambda_U(x)$)

For W=5, U=10, $\theta_1=2, \beta_1=1.5, \theta_2=3, \beta_2=2.0$

$P\{N(W,U)=n\}$ is as shown in Figure 3.

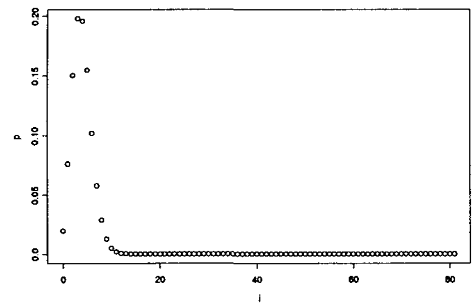

Figure 3. Plot of $P\{N(W,U)=n\}$

In this case $E(N(5,10))$ =3.952847.

3.2 In case $(\tilde{T}_i, \tilde{X}_i)$~ $f(t,x)$

First let $(\tilde{T}_i, \tilde{X}_i)$ be the inter-arrival age and mileage between claims:

$$(\tilde{T}_i, \tilde{X}_i)=(T_i - T_{i-1}, X_i - X_{i-1}).$$

We also assume that T_i~NHPP with $\lambda_W(t)$. Now we assume that

$$(\tilde{T}_i, \tilde{X}_i) \sim f(t,x) \; (= f(t)\; f(x\,|\,t))$$

follows some bivariate distributions. For instance, consider Gumbel's bivariate exponential distribution :

$$F(t,x) = (1-e^{-t})(1-e^{-x})(1+\alpha e^{-t-x}),\quad t,x \geq 0. \quad (6)$$

See Figure 4 for the density function f(t,x).

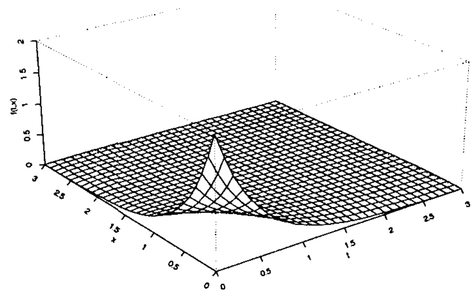

Figure 4. Plot of $f(t,x)$

Properties of $F(t,x)$ are as follows (see Johnson and Kotz (1972)):

$$f(t,x) = e^{-(t+x)}\{1+\alpha(2e^{-t}-1)(2e^{-x}-1)\}$$
$$f(x\,|\,t) = e^{-x}\{1+\alpha(2e^{-t}-1)(2e^{-x}-1)\}$$
$$E(X\,|\,T=t) = 1+\alpha/2-\alpha e^{-t} \quad (7)$$
$$Var(X\,|\,T=t) = 1+\alpha/2-\alpha^2/4-\alpha(1-\alpha)e^{-t}-\alpha^2 e^{-2t}$$
$$E(TX) = 1+\alpha/4$$
$$Corr(T,X) = \alpha/4$$

Now $E\{N(W,U)\}$ can be obtained using simulation as in the following procedure :

Set $i=1$

a) Generate t_i (note that t_i is from NHPP with $\lambda_W(t_i)$)

b) For t_i calculate $\tilde{t}_i = t_i - t_{i-1}$ (set t_0 =0) and set $\tilde{t}_i = t$.

c) Get $\tilde{x}_i$ from $E(\tilde{x}_i\,|\,\tilde{t}_i)$

d) If t_i>W or x_i>U stop and $N(W,U)=i$

Otherwise $i \leftarrow i+1$ and go back to a).

Example 3: Marked point process (Gumbel's bivariate distribution for $f(t,x)$)

Simulation results for $E\{N(W,U)\}$ with W=5 and U=2, 5, 10 for α =0 and 1 after 1,000 simulations are as follows :

		α	
		0	1
	2	1	1.129
U	5	2.234	2.694
	10	3.928	4.078

4. Counting and matching on both directions independently

Assume that N_t=number of events over $[0,t)$, N_x =number of events over $[0,x)$, N_C = Min(N_t, N_x).

Assume also that minimal repair is undertaken on both time and usage axis: both N_t~NHPP with $\lambda_W(t)$ and N_x~NHPP with $\lambda_U(x)$.

Then,

$$P\{N_C = n\}$$
$$= P\{N_t = n\}P\{N_x \geq n\} + P\{N_t \geq n+1\}P\{N_x = n\}$$
$$= \frac{e^{-\Lambda_W(t)}\Lambda_W(t)^n}{n!}\sum_{k=n}^{\infty}\frac{e^{-\Lambda_U(U)}\Lambda_U(U)^k}{k!} + \frac{e^{-\Lambda_U(x)}\Lambda_U(x)^n}{n!}\sum_{k=n+1}^{\infty}\frac{e^{-\Lambda_W(t)}\Lambda_W(t)^k}{k!} \quad (8)$$

Example 4: Independent case ($\delta = 1$) with $W = 5, U = 10, \theta_1 = 2, \beta_1 = 1.5, \theta_2 = 3, \beta_2 = 2.0$

Note that,

$$\Lambda_W(5) = \int_0^5 \lambda_W(t)dt = (5/\theta_1)^{\beta_1} = 3.953.$$

$$\Lambda_U(10) = \int_0^{10} \lambda_U(x)dx = (10/\theta_2)^{\beta_2} = 11.111$$

Therefore, the probability density function is given by $P\{N_C = n\}$

$$= \frac{e^{-3.953}3.953^n}{n!}\sum_{k=n}^{\infty}\frac{e^{-11.111}11.111^k}{k!} + \frac{e^{-11.111}11.111^n}{n!}\sum_{k=n+1}^{\infty}\frac{e^{-3.953}3.953^k}{k!} \quad (9)$$

The plot of $P\{N_C = n\}$ is as shown in Figure 5. Therefore $E(N(5,10))$=3.914.

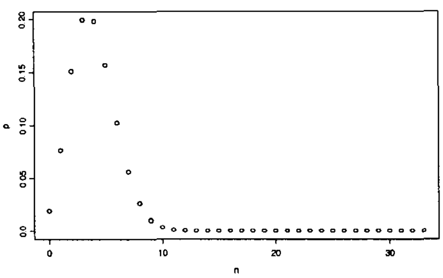

Figure 5. Plot of $P\{N(W,U) = n\}$

5. One-dimensional approach

This approach involves modelling usage as a function of time so that failures are effectively modelled by a one-dimensional point process formulation. Iskandar (1993) suggests a linear model for usage given by x(t)=Rt where R is a random variable for usage rate with the distribution function G(r) and is modelled to represent the varying usage across the consumer population. According to chapter 8 of Blischke and Murthy (1994) if we let $N(t\,|\,r)$ be the number of failures over the interval $[0,t)$ conditional on $R = r$ then

$$P\{N(t) = n\} = \int_0^{\infty} P\{N(t\,|\,r) = n\}dG(r)$$

where,

$$P\{N(t|r)=n\}=\frac{\{\int_0^t \lambda_W(s|r)ds\}^n \exp\{-\int_0^t \lambda_W(s|r)ds\}}{n!}$$

and if we let $N(W,U|r)$ be the number of failures under warranty conditional on $R=r$ then $E(N(W,U))$ =

$$\int_0^{\gamma_1} \Lambda_W(W|r)dG(r)+\int_{\gamma_1}^{u} \Lambda_W(U/r|r)dG(r)$$

where $\gamma_1 = U/W$.

Suppose that $\lambda_W(t|r)=\frac{\beta}{\alpha(r)}(\frac{t}{\alpha(r)})^{\beta-1}$ where we require that $\alpha(r)$ decreases as r increases.

Then we can consider the following two types of models for $\alpha(r)$:

Model 1. $\alpha(r)=\alpha_0+\alpha_1/r$.

Model 2. $\alpha(r)=\alpha_0+\alpha_1 e^{-r}$.

Suppose that $r \sim U(0,u)$. Then ,

$$E(N(W,U)) = \frac{1}{u}\{\int_0^{\gamma_1}(\frac{W}{\alpha_0+\alpha_1/r})^{\beta}dr + \int_{\gamma_1}^{u}(\frac{U/r}{\alpha_0+\alpha_1/r})^{\beta}dr\} \quad \text{for Model 1}$$

$$E(N(W,U)) = \frac{1}{u}\{\int_0^{\gamma_1}(\frac{W}{\alpha_0+\alpha_1 e^{-r}})^{\beta}dr + \int_{\gamma_1}^{u}(\frac{U/r}{\alpha_0+\alpha_1 e^{-r}})^{\beta}dr\} \quad \text{for Model 2}$$

Example 5. $E\{N(W,U)\}$ for Model 1

Suppose that the mean usage rate m is 2 (it could be that a car is driven 20,000km per year on the average) and u=2m. Then for W=5 and U=10, γ_1=U/W(=2).

Now the numerical computations for Model 1 are as follows:

α_0	α_1	α	β	$E\{N(W,U)\}$
0.5	3	2	1.5	2.559045
1	2	2	1.5	2.48144976603834
1.5	1	2	1.5	2.54971772060991
1.9	0.2	2	1.5	2.85361380389008
2	0	2	1.5	3.134186

6. Discussions

Two-dimensional warranty data modeling has been tried. Three different models have been attempted in the two-dimensional point process approach. Bivariate modeling technique is appealing but is not appropriate for explaining warranty claims. This model may be appropriate for spatial model in two-dimensional problem. Marked point process is appealing to warranty claim data analysis. But it is not yet sure whether, for instance, the assumption of $(\tilde{T}_i, \tilde{X}_i) \sim f(t,x)$ is valid irrespective of i in real warranty claim data. If not, a more complicated model is to be entertained. Still another model of counting and matching on both directions independently is also appealing. But it lacks in applicability in that it assumes that the two variables are independent. Last of all, one-dimensional point process with usage as a function of age is also promising to explain the two-dimensional warranty claim. But it assumes only linear relationship between age and mileage. In reality mileage to age ratio may be stiff at the beginning and less stiff later as time goes by or vice versa. At the moment both marked point process and one-dimensional point process with usage as a function of age (not just linear relationship) are tried to real warranty claim data.

References

[1] Blischke, W.R., and Murthy, D.N.P.(1994), Warranty Cost Analysis, Marcel Dekker, Inc., New York.

[2] Blischke, W.R. and Murthy, D.N.P. [Eds.](1996), Product Warranty Handbook, Marcel Dekker, Inc., New York.

[3] Iskandar, B.P.(1993), "Modelling and Analysis of Two Dimensional Warranties Policies," Doctoral Dissertation, The University of Queensland, Australia.

[4] Johnson, N.L. and Kotz, S.(1972), Distributions in Statistics: Continuous Multivariate Distributions, John Wiley & Sons, Inc.

[5] Kalbfleisch, J.D., Lawless, J.F. and Robinson, J.A.(1991), "Methods for the analysis and prediction of warranty claims," Technometrics, 33, 273-28

[6] Kalbfleisch, J.D. and Lawless, J.F.(1998), "Estimation of Reliability in Field-Performance Studies(with Discussion) ," Technometrics, 30, 365-388.

[7] Kim, H.G. and Rao, B.M.(2000), "Expected warranty cost of a two-attribute free-replacement warranties based on a bivariate exponential distribution," Computers and Industrial Engineering, 38, 425-434.

[8] Lawless, J.F.(1998), "Statistical analysis of product warranty data," International Statistical Review, 66, 41-60.

[9] Lawless, J.F. and Kalbfleisch, J.D.(1992), "Some issues in the collection and analysis of field reliability data," in Survival Analysis: State of the Art (J.P. Klein and P.K. Goel eds.), Kluwer, Amsterdam, 141-152.

[10] Lu, J.C. and Bhattacharyya, G.K.(1990), "Some new constructions of bivariate Weibull models," Ann. Inst. Statist. Math, 42, 543-549

[11] Murthy, D.N.P., Iskandar, B.P and Wilson, R.J.(1995), "Two dimensional failure free warranties: Two-dimensional point process models," Operations Research, 43, 356-366.

[12] Singpurwalla, N.D. and Wilson, S.P.(1998), "Failure models indexed by two scales," Advances in Applied Probability, 30, 1058-1072.

[13] Yang, S-C and Nachlas, J.A.(2001), "Bivariate reliability and availability modelling," IEEE Transactions on Reliability, 50, 26-35.