

Early Detection of Reliability Problems Using Information From Warranty Databases

Huaiqing Wu

William Q. Meeker

Department of Statistics

Iowa State University

Ames, IA 50011

www.public.iastate.edu/~wqmeeker

August 16, 2001

1

Overview

- Background and motivation
- Example: automobile "VIN" and "Warranty Report" databases.
- Statistical formulation of the detection problem.
- Allocation of false alarm probabilities. Balance power over time.
- Critical values for the early detection monitoring procedure.
- Average run length.
- Summarization of monitoring procedure behavior.
- Concluding remarks.

2

Background and Motivation

- Serious, unanticipated reliability problems arise
- Early detection can save large amounts of money and customer good will
- Statistical detection rules in a warranty database can identify or warn of serious reliability problems long before they would otherwise be discovered
- Monitoring existing warranty data bases will add little additional costs

3

Previous Work

- General warranty topics: Blischke and Murthy (1994, 1996)
- General reviews of statistical methods: Robinson and McDonald (1991), Lawless and Kalbfleisch (1992), and Lawless (1998)
- Dealing with messy data problems: Suzuki (1985a, b), Kalbfleisch and Lawless (1988), Lawless, Hu, and Cao (1995), Hu, Lawless, and Suzuki (1998), Karim, Yamamoto, and Suzuki (2001a), and Wang and Suzuki (2001a, b)
- Detecting a change point from marginal count warranty data: Karim, Yamamoto, and Suzuki (2001b)

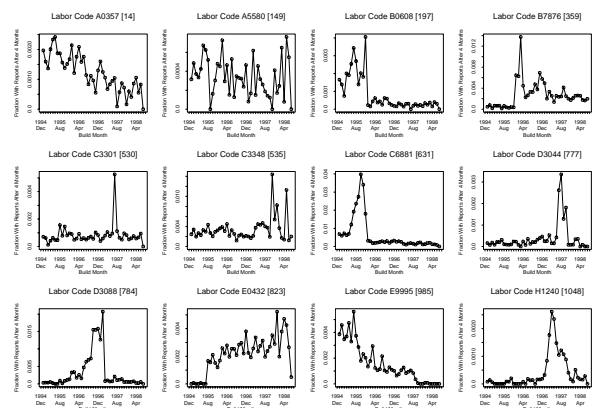
4

Automobile "VIN" and "Warranty Report" Databases

- VIN (production) database
 - ▶ One record for each automobile for a particular "platform" manufactured between January 1995 and August 1998
 - ▶ 566,406 unique VINs
 - ▶ Provides date of sale and other production information
- Warranty report database
 - ▶ Warranty reports received between January 1995 and November 1998
 - ▶ 1,350,675 records with 1,908 unique labor codes
 - ▶ Also contains cost, dealership code, and other information

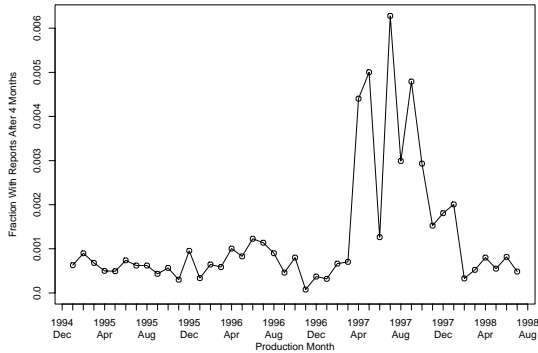
5

Fraction of Automobiles with Warranty Reports in the First Four Months of Service Versus Production Month for 12 Different Labor Codes



6

Retrospective View Giving the Fraction of Automobiles with Labor Code C0140 Warranty Reports in the First Four Months of Service versus Production Month



General Formulation of the Detection Problem

- Generalization of Shewhart process monitoring schemes
- Stratify by production period, time in service, and labor code
- One month used as the production/ time in service period in examples
- Data for a given production period accrue sequentially (due to variability in sale date and tracking by time in service)
- Nonparametric approach (based on report counts) most appropriate for detection
- Use a parametric model (e.g., Weibull or Lognormal) for prediction of future failures

Notation

- n_i units produced in period i
- n_{ij} units produced in period i and sold j periods after manufacture
- R_{ijk} number of warranty reports k th period in service for units manufactured in period i and sold in period $i+(j-1)$
- $S_{ijk} = \sum_{\ell=1}^j R_{i\ell k}$ is the cumulative number of reports for units manufactured in period i and sold in period $i+(j-1)$
- C_{ijk} critical limit for S_{ijk}
- λ_k intensity for service period k
- M number of future periods to monitor report intensities
- f_j fraction sold j months after production; $\sum_{j=1}^M f_j \leq 1$
- Superscript 0 (e.g., f_j^0) denotes historical or baseline value.

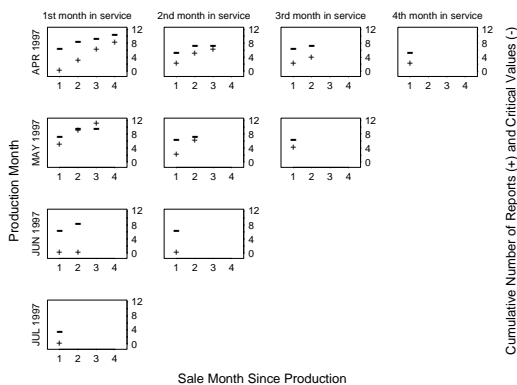
Formal Detection Rule Framework

- Multiple-parameter hypothesis

$$H_0: \lambda_1 \leq \lambda_1^0, \lambda_2 \leq \lambda_2^0, \dots, \lambda_M \leq \lambda_M^0$$
 versus $H_a: \lambda_1 > \lambda_1^0$ or $\lambda_2 > \lambda_2^0$ or ... or $\lambda_M > \lambda_M^0$,
- λ_k^0 values are from historical information or reliability targets
- α_k is the nominal false alarm probability for testing λ_k
- The overall false alarm probability for testing the hypothesis H_0 versus H_a is in is

$$\alpha^* = 1 - \prod_{k=1}^M (1 - \alpha_k^*) \leq 1 - \prod_{k=1}^M (1 - \alpha_k) = \alpha,$$
- Control and allocation of false alarm probabilities should be labor code specific.

**Sequential Test Monitoring Charts
Labor Code C0140 Warranty Reports**
+ Indicates Cumulative Number of Reports S_{ijk}
- Indicates Corresponding Critical Limit C_{ijk}

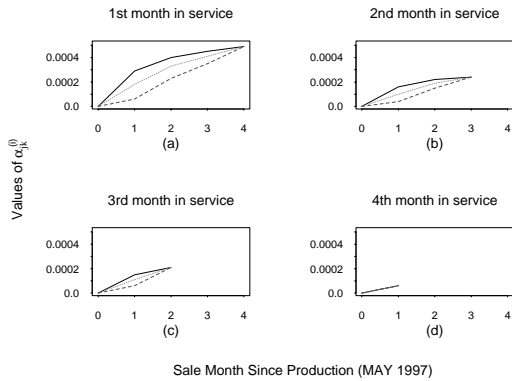


Control and Allocation of False Alarm Probabilities

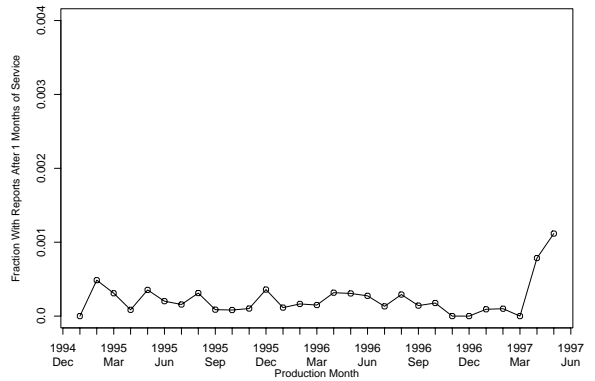
- Limit M , the number of periods monitored
- Simple rule: α_k to be proportional to the information available for testing H_0^k versus H_a^k :

$$\alpha_k = C(f_1^0 + \dots + f_{M-k+1}^0)\lambda_k^0,$$
- Must also allocate power for the sequential accumulation of data. We use the *error spending* approach developed Slud and Wei (1982) for sequential clinical trials.

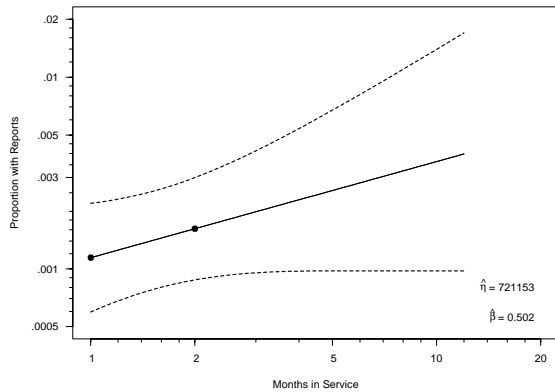
Typical Spending Functions for Production Month May 1997 for Labor Code C0140 With $\alpha = .1\%$ and $M = 4$, and $\rho = .5$, —; $\rho = 1$, \cdots ; and $\rho = 2$, - - -.



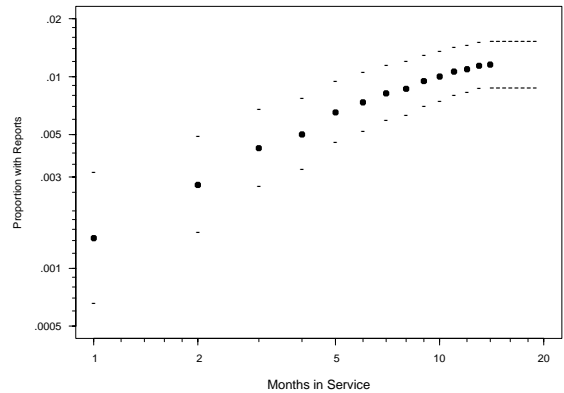
Nonparametric Estimate of Fraction With Reports After One Month of Service as a Function of Production Month for Data Available in July 1997.



Weibull Probability Plot and Weibull ML Estimate for Production Month May 1997 Based on Data Available in July 1997 (after two months in service)



Weibull Probability Plot for Production Month May 1997 Based on Data Available in November 1998



Average Run Length

- Probability of no alarm is triggered during the first $(\ell - 1)$ periods of monitoring

$$ARL = E(N) = \sum_{\ell=1}^{\infty} \Pr(N \geq \ell)$$

- Under H_0

$$\Pr(N \geq \ell) = \prod_{i=1}^{\ell-1} \gamma_{i, \ell-i} \geq (1 - \alpha)^{\ell-1}.$$

- Simple expression for ARL if n_i is constant and $f_{ij} = f_j^0, j = 1, \dots, M$
- Can also compute under specified H_a

Concluding Remarks and Extensions

- All 1,908 labor codes were investigated.
- Methods implemented on 48 interesting labor codes. Results summarized in the paper.
- Efficient monitoring should provide economical early detection, especially with modern computing/data storage capabilities
- Modeling past data is useful for setting the baseline report rates
- Should monitor different production lines/shifts separately
- Runs rules useful in some settings
- Detection also provides input to warranty cost forecasting algorithm