AN OFFPRINT FROM

BAYESIAN STATISTICS 4

Proceedings of the Fourth Valencia International Meeting

Dedicated to the memory of Morris H. DeGroot, 1931–1989

April 15–20, 1991

Edited by
J. M. Bernardo
J. O. Berger
A. P. Dawid
and
A. F. M. Smith

CLARENDON PRESS · OXFORD
1992

BAYESIAN STATISTICS 4, pp. 675–688
J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, (Eds.)

Bayesian Models for Quality Assurance

TELBA Z. IRONY, CARLOS A. de B. PEREIRA and RICHARD E. BARLOW
The George Washington University, USA, Universidade de São Paulo, Brazil
and *University of California at Berkeley, USA*

SUMMARY

Two models for analyzing discrete time series of quality data are suggested. They enable analysts to extract information from data and from experts in a very effective way. The Additive model should be used for processes that degrade as time goes by (processes that age, for instance). The Multiplicative model fits processes that improve with time (processes that depend on learning). In the case that the process varies, sometimes getting better and sometimes getting worse, both models should be alternated according to the process behavior. Simulation is used to validate the proposed models and to evaluate their performances.

Keywords: QUALITY ASSURANCE; QUALITY AUDITS; PRODUCTION PROCESS; QMP – QUALITY MEASUREMENT PLAN; INFLUENCE DIAGRAM; STATISTICAL CONTROL; EXCHANGEABILITY; GAMMA DISTRIBUTION; BETA DISTRIBUTION; POISSON DISTRIBUTION.

1. INTRODUCTION

Quality audits are performed by inspectors in production processes in order to report product quality to management. A quality audit is a structured system of inspections done continuously on a sampling basis. Sampled product is inspected and defects are assessed whenever the product fails to meet engineering requirements. The results are combined into a rating period and compared to a quality standard which is a target value of defects per unit reflecting a trade-off between manufacturing cost, operating costs and customer need.

The models presented in this paper aim to replace the Quality Measurement Plan (QMP), which was implemented throughout AT&T Technologies in 1980 (see Hoadley (1981)). The QMP is a statistical method for analyzing discrete time series of quality audit data consisting of the expected number of defects given standard quality. The method is heuristic because the model's exact solution is mathematically intractable.

Two alternative models, namely the Additive model and the Multiplicative model, are presented in this paper. They are exact, tractable and much easier to use than the QMP.

The Additive model was formulated in order to deal with production processes that degrade as time goes by (processes that age, for instance). The Multiplicative model is appropriate for processes that improve with time (e.g. processes that depend on learning). In cases where the process varies, sometimes getting better and sometimes getting worse (behavior of an out-of-control process) both models should be alternated. This procedure is straightforward due to the fact that the posterior distributions in both models are of the same nature.

The models enable the analyst to extract information from the data and from the experts that are involved in the problem in a very effective way. The expert knowledge is intrinsically incorporated into the models. The information extracted from the available data updates the expert opinion via Bayes' theorem. Influence diagrams were very helpful to summarize the dynamics of the modeling process.

Simulation has been used to validate the proposed models and to evaluate their performance.

2. NOTATION AND ASSUMPTIONS

Suppose there are T rating periods: $t = 1, \ldots, T$ (current period). For period t, we have the following data from the audit:

n_t = audit sample size; x_t = number of defects in the audit sample;

s = standard number of defects per unit.

$e_t = sn_t$ = expected number of defects in the sample when the quality standard is met.

The assumptions are the following: x_t has a Poisson distribution with mean $n_t\lambda_t$, where λ_t is the defect rate per unit. If λ_t is reparametrized on a quality index scale, the result is:

$\theta_t = \lambda_t/s$ = quality index at rating period t. In other words, $\theta_t = 1$ is the standard value. Therefore, we can write: $(x_t|\theta_t) \sim \text{Poi}(e_t\theta_t)$.

The parameter of interest is θ_T, the current quality index. The objective is to derive the posterior distribution of θ_T given the past data, d_{T-1}, and current data, x_T.

Here $d_{T-1} = (x_1, \ldots, x_{T-1})$ and d_0 is a constant.

The standard quality on the quality index scale is "one". "Two" means twice as many defects as expected under the standard. Hence, the larger the quality index, the worse the process.

3. THE ADDITIVE MODEL

3.1. *Assumptions*

This model is adequate for processes that degrade with time. It starts with a quality index θ, which may be thought of as the quality index for previous ratings. At each rating period t, an increment δ_t, also unknown, is added so that the quality index at rating period $t = 1$ will be $\theta_1 = \theta + \delta_1$, the quality index at rating period $t = 2$ will be $\theta_2 = \theta + \delta_1 + \delta_2$ and at $t = T$, the quality index will be $\theta_T = \theta + \delta_1 + \ldots + \delta_T$.

Here we are assuming that e_t is constant for all periods t (and we will call it e) but the model may be easily extended for the case in which e_t varies from period to period.

This model does not require the assumption of exchangeability between lots, allowing changes in the quality index from period to period. The following influence diagram represents the Additive model:

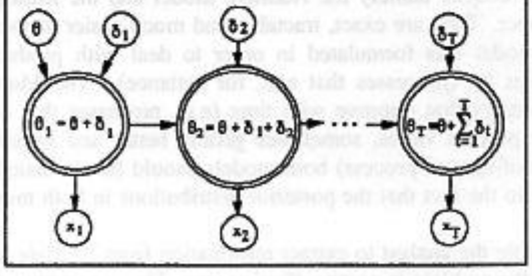

Figure 1.

Since usually many factors affect quality, there is a central limit theorem effect. Therefore unimodal prior distributions for θ and for the δ's are reasonable choices. A convenient assessment is a gamma distribution. Hence:

$\theta|\alpha_0, \beta_0 \sim \text{Gamma}(\alpha_0, \beta_0)$

$\delta_t|\alpha_t, \beta_t \sim \text{Gamma}(\alpha_t, \beta_t)$ where $\beta_t = \beta_0 + (t-1)e$ for $t = 1, \ldots, T$.

The choice of α_0 and β_0 is completely free, and will reflect the analyst's experience about the initial quality index θ. The assessment of prior distributions for the δ's is more delicate. The choice of α_t is also free but β_t is determined by β_0 and by the period in which the rating is being made. This feature makes the model work nicely. In order to fit a gamma distribution with any mean she pleases, the analyst has to pick the right α_t. Consequently, there will be a trade-off between the mean and the variance of the δ's. Usually the choice of the mean is more meaningful since the prior mean of δ_t expresses the average amount by which the analyst judges that the quality index has increased from period $t-1$ to period t. On the other hand, it is reasonable that the variance of δ_t becomes smaller and smaller as t increases and the analyst becomes more acquainted with the production process.

If it is judged that the means of the δ's are about the same for all rating periods, the variances will be decreasing. Whenever the variance of δ_t becomes too small, the assessment procedure must start all over again. In other words, t must be reset to 1 and new gamma distributions should be assessed for θ and δ_1. These distributions should incorporate all the knowledge the analyst has gathered up to the current rating period. Then, α_0, β_0 and α_1 will be chosen freely and suitable means and variances for θ_1 will be assessed.

At any rating period T, the quality index is given by $\theta_T = \theta + \sum_{t=1}^{T} \delta_t$. It will be shown in the sequel that the posterior distribution for θ_T will be gamma with shape parameter $\sum_{t=0}^{T} \alpha_t + \sum_{t=1}^{T} x_t$ and scale parameter $(\beta_0 + Te)^{-1}$.

The influence diagram that represents this model has to be constructed as time goes by due to the dynamic nature of the model. The system may change at each rating period and three new nodes will be added to the influence diagram.

3.2. *The Solution*

Let us start with rating period $t = 1$. Based on previous knowledge about the production process, prior distributions for both, θ and δ_1 should be assessed. The corresponding influence diagram is:

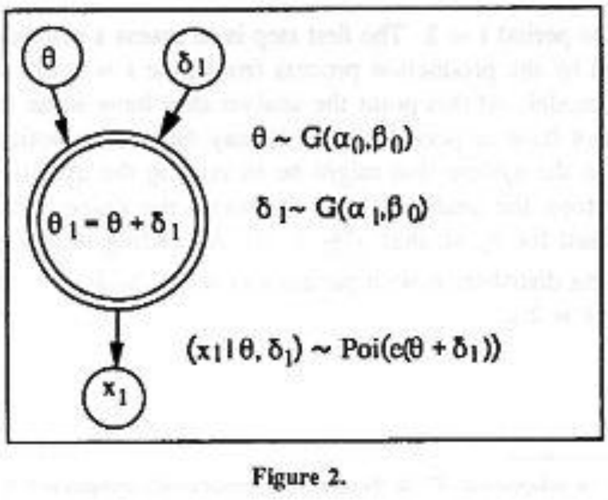

Figure 2.

Let $\theta \sim G(\alpha_0, \beta_0), \delta_1 \sim G(\alpha_1, \beta_0)$ and $\theta \amalg \delta_1 | \alpha_0, \alpha_1, \beta_0$[(1)].

Then $\theta_1 = (\theta + \delta_1) \sim G(\alpha_0 + \alpha_1, \beta_0)$.

Consequently:

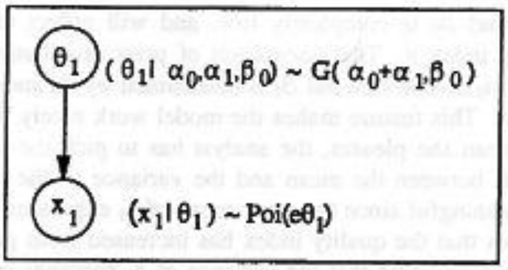

Figure 3.

Our objective is to compute the posterior distribution of θ_1, the current quality index at rating period $t = 1$, given the number of defects found in that period, x_1. This corresponds to an arc reversal operation in the influence diagram framework.

The updated influence diagram is:

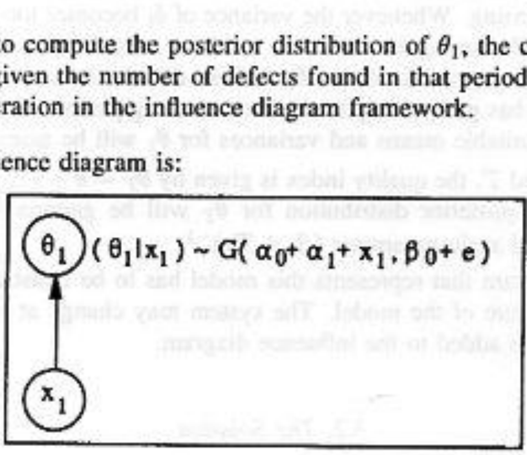

Figure 4.

Now we proceed to period $t = 2$. The first step is to assess a prior distribution for δ_2, the increment experienced by the production process from time $t = 1$ to $t = 2$, and here comes the advantage of this model. At this point the analyst may have some information about the process that she did not have at period $t = 1$. It may have been noticed, for instance, that there is a small flaw in the system that might be increasing the quality index by an amount Δ, on average. Therefore, the analyst is able to choose the shape parameter of the gamma distribution, α_2, assessed for δ_2, so that $\frac{\alpha_2}{\beta_0+e} = \Delta$. According to this reasoning, δ_2 will be assessed using a gamma distribution with parameters $\alpha_2 = \Delta(\beta_0 + e)$ and $\beta_2 = \beta_0 + e$. The influence diagram for $t = 2$ is:

(1) The symbol $\amalg$ means "is independent of". In this case, θ is conditionally independent of δ_1 given $\alpha_0, \alpha_1, \beta_0$.

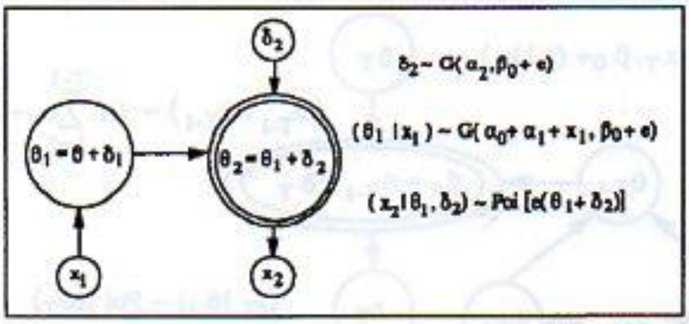

Figure 5.

By the same argument used previously, $\theta_2 = \theta_1 + \delta_2 \sim G\left(\sum_{t=0}^{2} \alpha_t + x_1, \beta_0 + e\right)$. Then:

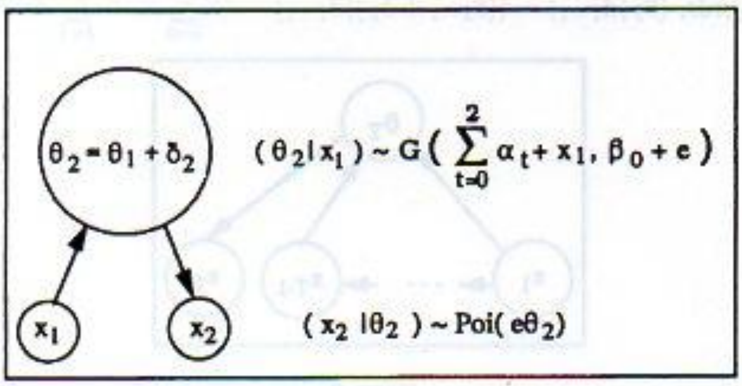

Figure 6.

The posterior distribution of θ_2 given x_1 and x_2 will be a gamma distribution with parameters $\left(\sum_{t=0}^{2} \alpha_t + \sum_{t=1}^{2} x_t\right)$ and $(\beta_0 + 2e)$. The resulting influence diagram is:

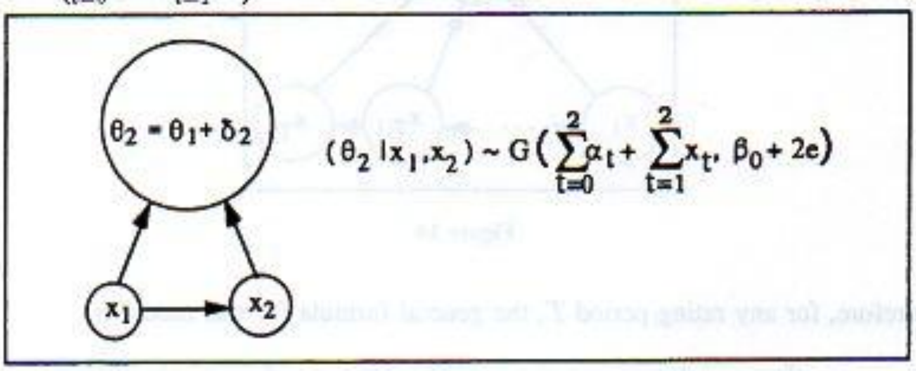

Figure 7.

By induction we are able to get the general influence diagram for any rating period $t = T$.

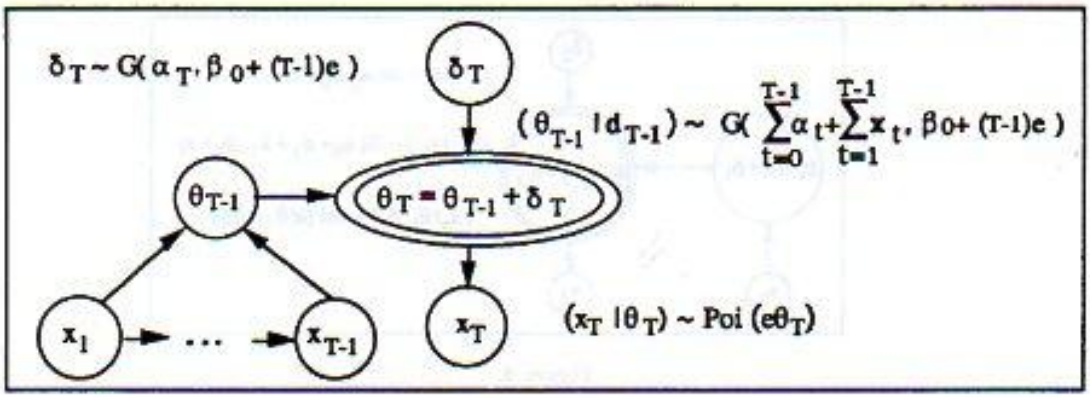

Figure 8.

In other words, $(\theta_T|d_{T-1}) = ((\theta_{T-1}+\delta_T)|d_{T-1}) \sim G(\sum_{t=0}^{T}\alpha_t + \sum_{t=1}^{T-1} x_t, \beta_0 + (T-1)e)$.

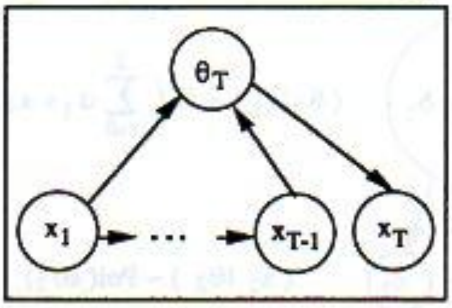

Figure 9.

The final influence diagram is obtained by performing the arc reversal operation.

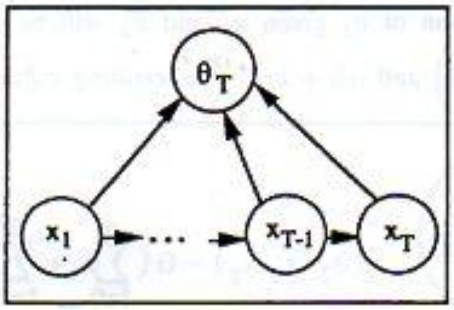

Figure 10.

Therefore, for any rating period T, the general formula for this model is:

$$(\theta_T|d_T) \sim G(\sum_{t=0}^{T}\alpha_t + \sum_{t=1}^{T} x_t, \beta_0 + Te), \text{ where } \theta_T = \theta + \sum_{t=1}^{T}\delta_t \text{ and } d_T = (x_1, \ldots, x_T).$$

4. THE MULTIPLICATIVE MODEL

4.1. *Assumptions*

This model should be used for processes that are judged to get better as time goes by. In this case, we also start with the initial quality index θ. At each rating period t, the quality index will be given by $\theta_t = \theta_{t-1}(1-\delta_t)$ where $0 \le \delta_t \le 1, t = 1, \ldots, T$ and $\theta_0 = \theta$. δ_t is the proportion by which the analyst judges that the process got better at time t.

As before, $(x_t|\theta_t) \sim \text{Poi}(e\theta_t)$ where e is the expected number of defects in the sample when the quality standard is met. We remain with the assumption that e is constant for all periods $t = 1, \ldots, T$. Again, the assumption of exchangeability is not needed and the following influence diagram will represent the Multiplicative model:

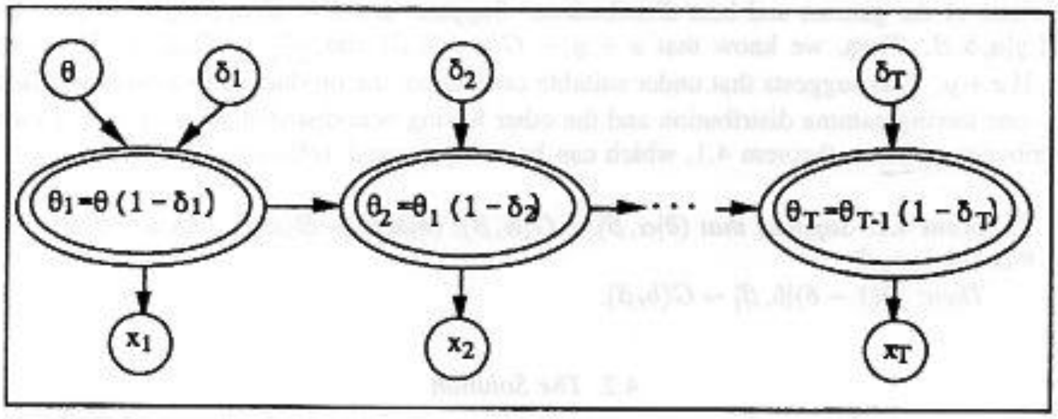

Figure 11.

The influence diagram representation requires an assessment of a joint distribution for the random quantities appearing in the model. The easiest thing to do is to assess prior distributions for θ and for the δ's, since it is been already assumed that $(x_t|\theta_t) \sim \text{Poi}(e\theta_t)$.

The gamma distribution is a sensible assessment for θ.

Since $0 \le \delta_t \le 1$, beta distributions are reasonable assessments for the δ's.

After T rating periods, the analyst is interested in the posterior distribution for the quality index θ_T. The posterior distribution obtained for θ_T when the Multiplicative model is used is gamma. In symbols, $(\theta_T|\boldsymbol{d}_T) \sim G(\alpha_T, \beta_T)$. If the assessment for δ_T is a beta distribution with parameters a_T and $b_T(\delta_T \sim B(a_T, b_T))$, then $\alpha_T = b_T + x_T$. If the assessment for the initial quality index θ is a $G(\alpha_0, \beta_0)$, then $\beta_T = \beta_0 + Te$.

The Multiplicative model requires some constraints on the parameters of the distributions assessed to the δ's. These constraints will be understood as the model is explained and the solution is worked out. The first step is to assess a gamma distribution for $\theta : \theta \sim G(\alpha_0, \beta_0)$. The assessment for δ_1 should be of the form:

$\delta_1 \sim B(a_1, b_1)$ where $a_1 + b_1 = \alpha_0$ and $B(a_1, b_1)$ means a beta distribution with parameters a_1 and b_1.

At rating period t, we will have:

$(\theta_{t-1}|\boldsymbol{d}_{t-1}) \sim G(\alpha_{t-1}, \beta_{t-1})$ and $\delta_t \sim B(a_t, b_t)$ where $a_t + b_t = \alpha_{t-1}$ for $t = 1, \ldots, T$.

In other words, the assessments for the δ's must be such that the parameters of the beta distribution assessed for δ_t will depend upon the parameters of the posterior distribution of θ_{t-1} given $\boldsymbol{d}_{t-1}$.

The choice of α_0 and β_0, which will reflect the analyst's opinion about the process average, is free. The choice of a_t and b_t is constrained by the relation: $a_t + b_t = \alpha_{t-1}$. This means that there is a trade-off between the mean and the variance of δ_t and the analyst must keep it in mind. It is more intuitive to assess the mean of δ_t because it expresses the mean proportion by which the analyst judges the process is changing. Nevertheless, this choice is not completely free because it could lead to an unreasonable variance.

As in the Additive model, the influence diagram that represents this model has to be constructed as time goes by. At each rating period, new nodes referring to the current period are added to the influence diagram and the assessments of these nodes are made based upon the posterior distribution of the nodes referring to previous periods.

Before we proceed with the solution for this model, let us explore an interesting characteristic of the gamma and beta distributions. Suppose that $x \sim G(a, \beta), y \sim G(b, \beta)$ and $x \amalg y | a, b, \beta$. Then, we know that $x + y \sim G(a + b, \beta)$ and $\frac{x}{x+y} \sim B(a, b)$. Moreover, $\frac{x}{x+y} \amalg x + y$. This suggests that under suitable conditions, the product of two random quantities, one having gamma distribution and the other having beta distribution, may have gamma distribution. Hence, theorem 4.1, which can be easily proved, follows:

Theorem 4.1. *Suppose that* $(\theta|\alpha, \beta) \sim G(\alpha, \beta)$, $(\delta|a, b) \sim B(a, b)$ *with* $a + b = \alpha$, *and* $\theta \amalg \delta | a, b, \alpha, \beta$.

Then: $[\theta(1 - \delta)|b, \beta] \sim G(b, \beta)$.

4.2. *The Solution*

Starting with rating period $t = 1$, we have the following influence diagram.

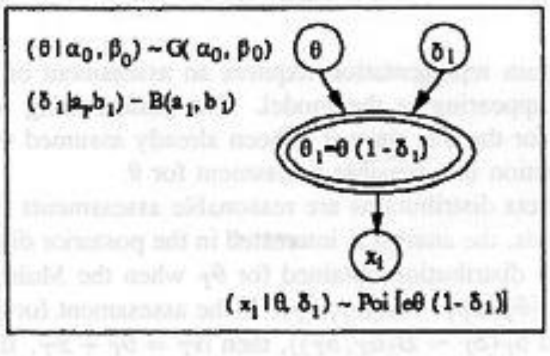

Figure 12.

By theorem 4.1, this diagram may be redrawn in the following way:

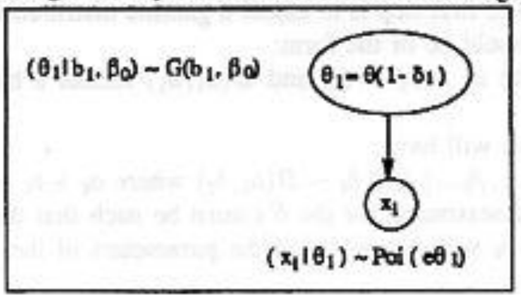

Figure 13.

The objective is to reverse the arc and to compute the posterior distribution of θ_1 given x_1.

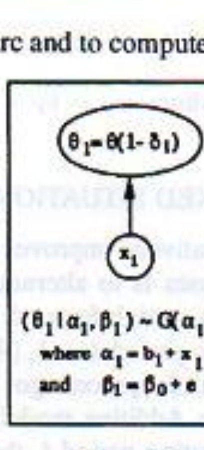

Figure 14.

The general solution, for any rating period $t = T$, may be obtained by induction. It is illustrated in the following sequence of influence diagrams:

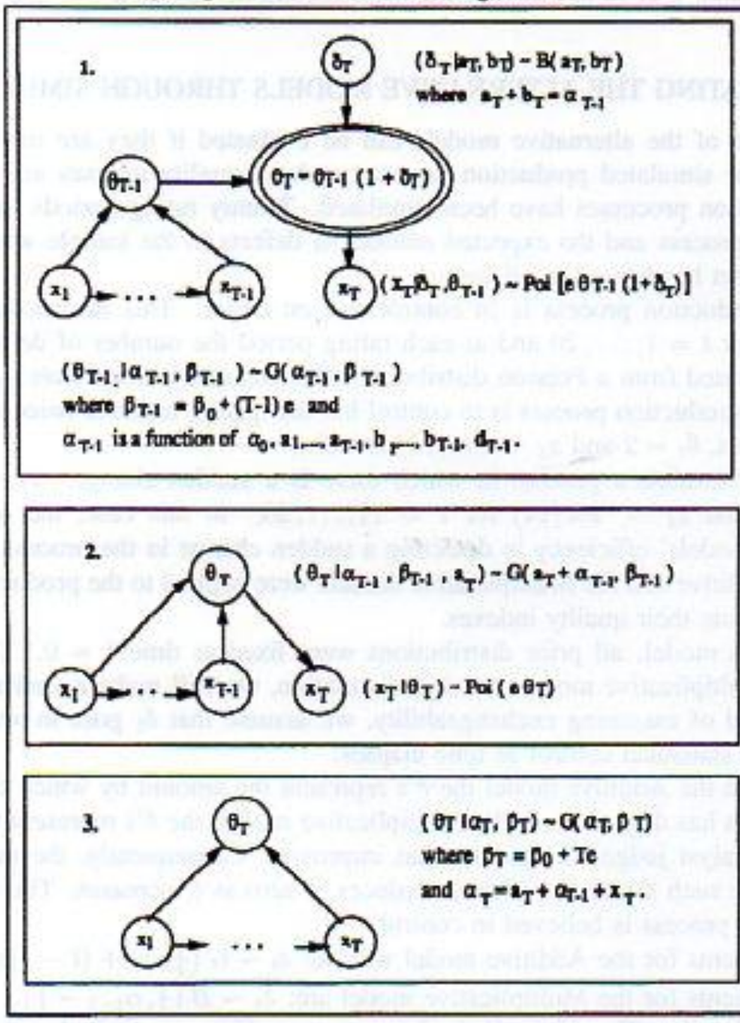

Figure 15.

In other words, the posterior distribution of the current quality index is given by:

$$(\theta_T|\alpha_T, \beta_T) \sim G(\alpha_T, \beta_T) \text{ where } \alpha_T = b_T + x_T \text{ and } \beta_T = \beta_0 + Te.$$

5. MIXED SITUATIONS

A situation in which the quality alternatively improves and deteriorates is surely of some interest. The best solution for those cases is to alternate the use of the Multiplicative and the Additive models. This procedure is straightforward because in both cases the posterior distribution of the quality index at any period $t-1, (\theta_{t-1}|d_{t-1})$ is a gamma distribution. Consequently, if the analyst believes that the process got better at period t, she should switch to the Multiplicative model even if the Additive model was used up to that time. If it is believed that the process got worse at rating period t, the Additive model should be used.

Due to the mathematical tractability of both models, this is a much better idea than to use always the Multiplicative model, with the multipliers being allowed to span a range containing the value 1 or to use always the Additive model allowing the "increments" δ_t to be negative.

6. EVALUATING THE ALTERNATIVE MODELS THROUGH SIMULATION

The performance of the alternative models can be evaluated if they are used to construct control charts for simulated production processes whose quality indexes are known. Four different production processes have been simulated. Twenty rating periods have been considered in each process and the expected number of defects in the sample when the quality standard (e) is met has been 7 in all cases.

The first production process is in control and on target. This means that the quality index, θ_t, is 1 for $t = 1, \ldots, 20$ and at each rating period the number of defects in the lot, x_t, will be generated from a Poisson distribution with mean 1e. In symbols $x_t \sim \text{Poi}(7)$.

The second production process is in control but the quality index is twice as large as the target. In symbols, $\theta_t = 2$ and $x_t \sim \text{Poi}(14)$ for all t.

Finally, we simulate a process in which there is a sudden change. $x_t \sim \text{Poi}(7)$ for $t = 1, \ldots, 10$ and $x_t \sim \text{Poi}(14)$ for $t = 11, \ldots, 20$. In this case, the objective is to investigate the models' efficiency in detecting a sudden change in the process.

Both the Additive and the Multiplicative models were applied to the production processes in order to estimate their quality indexes.

In Hoadley's model, all prior distributions were fixed at time $t = 0$. To analyze the Additive and Multiplicative models through simulation, we will make a similar assumption. However, instead of assuming exchangeability, we assume that δ_t goes to zero that is, the process tends to statistical control as time elapses.

Recall that in the Additive model the δ's represent the amount by which the analyst believes the process has degraded. In the Multiplicative model, the δ's represent the proportion by which the analyst judges the process has improved. Consequently, the assessments for the δ's should be such that their influence reduces to zero as t increases. This is appropriate in cases that the process is believed in control.

The assessments for the Additive model will be: $\delta_t \sim G\left(\frac{1}{t}, \alpha_0 + (t-1)e\right)$.

The assessments for the Multiplicative model are: $\delta_t \sim B\left(\frac{1}{t}, \alpha_{t-1} - \frac{1}{t}\right)$.

By minimizing the effect of the δ's in this way, we achieve a certain degree of objectivity in the evaluation of the alternative models. It should be pointed out that the main difference between the Additive and the Multiplicative models is the role of the δ's. Since the δ's have no influence in our case, we do not expect to find a remarkable difference between the performances of the Additive and the Multiplicative models in the simulated processes.

Both models require a prior probability for the initial quality index θ. We begin with a gamma distribution having parameters $\alpha_0 = \beta_0 = 5$, which is equivalent to assessing mean 1 and variance 0.20 for θ. This is done for all simulated production processes.

The charts plotted in the sequel represent the analysis of the production processes via the alternative models. They are based upon the posterior probabilities for the quality index given the simulated data. The dots represent the posterior mean for the quality index at each rating period. A whisker represents one posterior standard deviation. Consequently, the intervals depicted in the charts are two standard deviations long and are centered on the posterior mean.

Processes in Statistical Control

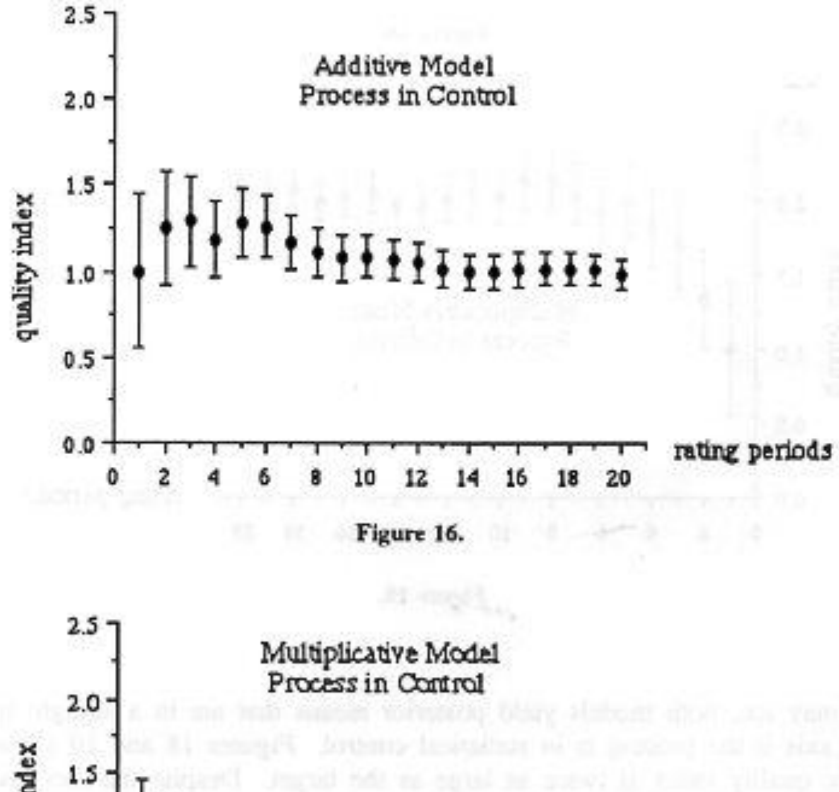

Figure 16.

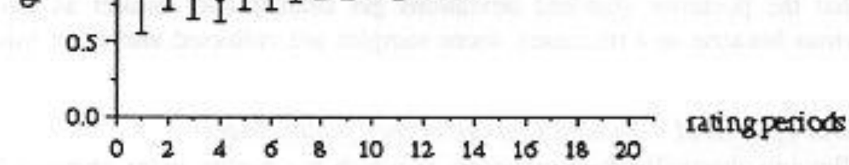

Figure 17.

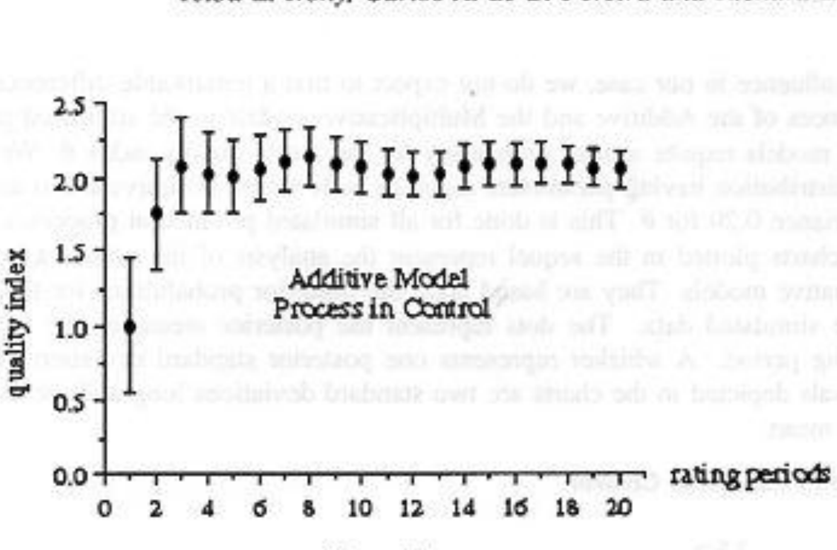

Figure 18.

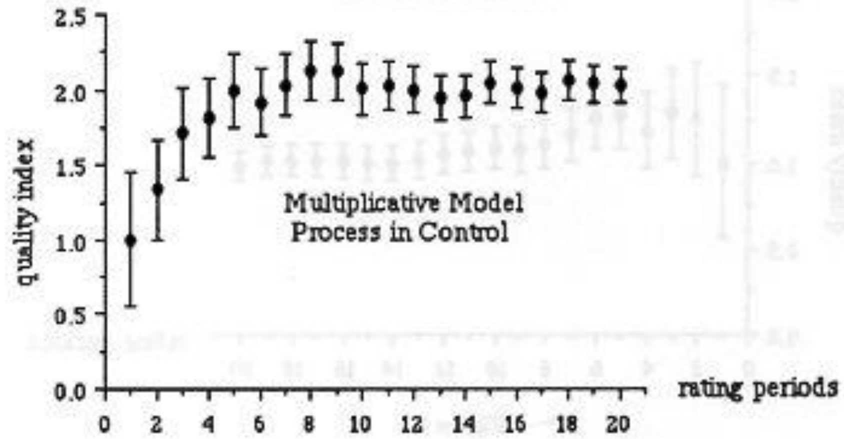

Figure 19.

As we may see, both models yield posterior means that are in a straight line parallel to the time axis if the process is in statistical control. Figures 18 and 19 show processes in which the quality index is twice as large as the target. Despite the fact that the prior mean assessed for θ was 1, the models were able to capture the information provided by the simulated data and the posterior mean for θ_t approximates 2 for $t \geq 2$.

Note that the posterior standard deviations get smaller and smaller as time goes by. This is obvious because as t increases, more samples are collected and more information is extracted.

Processes Out of Control – Sudden Change in the Quality Index

The following charts illustrate processes in which the quality index changes from 1 to 2 in the 11^{th} rating period.

It is easy to observe in the charts that from rating period 11 on, the posterior means of the quality indexes start to increase systematically. Despite the fact that there is a sudden change in the quality index from 1 to 2 at rating period 11, the charts do not show a jump

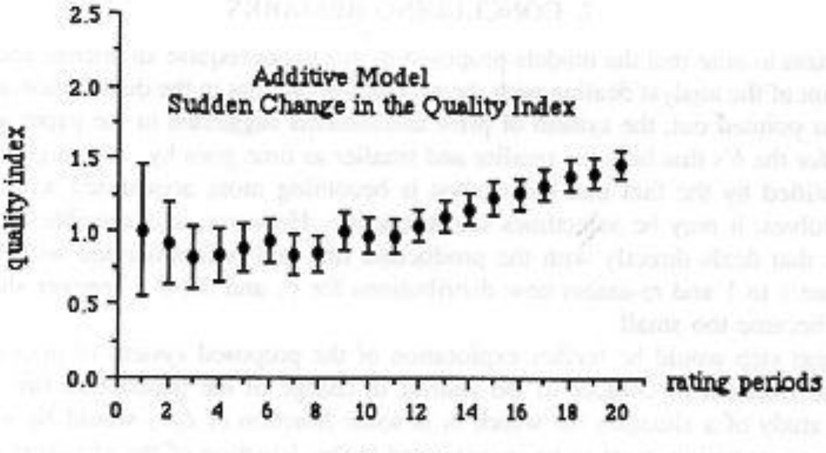

Figure 20.

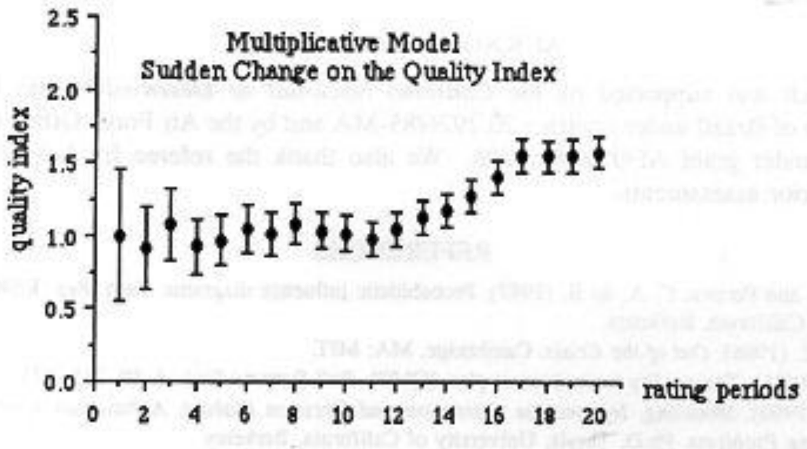

Figure 21.

from 1 to 2. Instead, they show a gradual change starting at period 11. This happens because by period 11, a great deal of information supporting the fact that the quality index should be around 1 have been assimilated. Good past history tempers an observed change. The more rating periods we have in the past, the greater is the inertia preventing the indication of a sudden change. In other words, the models are robust against statistical "jitter". They do not overreact to a few more defects.

Whenever a change of trend in the quality index is detected, one must suspect that there might be a sudden change in the process. In that case, the recommended procedure is to restrict the use of past data and to start recalculating the quality index from the point the change is detected. For example, in our case, the trend of change in the quality index started at period 11. Therefore, at period 11, t should be reset to 1, a wide prior distribution based on past data and expert judgement should be assessed to the quality index and the process must start all over.

7. CONCLUDING REMARKS

It is important to note that the models proposed in this paper require an intense and interactive participation of the analyst dealing with the production process in the distribution assessments.

As was pointed out, the system of prior assessments suggested in the paper will produce variances for the δ's that become smaller and smaller as time goes by. Although this behavior can be justified by the fact that the analyst is becoming more acquainted with the process as time evolves, it may be sometimes too restrictive. However, it is sensible to believe that an analyst that deals directly with the production line and is familiarized with it should be able to reset t to 1 and re-assess new distributions for δ_1 and for θ whenever she thinks the variances became too small.

The next step would be further exploration of the proposed system of priors that would allow more freedom of choices to the analyst in charge of the production line. Perhaps, a simulated study of a situation for which δ_t is some function of δ_{t-1} would be worthwhile.

Another interesting point to be investigated is the detection of the changing time for the cases in which the production process gets out of control. There is some literature about this topic but the specifics about quality assurance in production processes are worth exploring.

ACKNOWLEDGEMENTS

This research was supported by the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* of Brazil under contract 20.2926/85-MA and by the Air Force Office of Scientific Research, under grant AFOSR-86-0208. We also thank the referee for helpful comments about the prior assessments.

REFERENCES

Barlow, R. E. and Pereira, C. A. de B. (1987). Probabilistic influence diagrams. *Tech. Rep.* **ESRC 87-7**, University of California, Berkeley.

Deming, W. E. (1986). *Out of the Crisis*. Cambridge, MA: MIT.

Hoadley, B. (1981). The quality measurement plan (QMP). *Bell Systems Tech. J.* **60**, 215–273.

Irony, T. Z. (1990). *Modeling, Information Extraction and Decision Making. A Bayesian Approach to Some Engineering Problems*. Ph.D. Thesis, University of California, Berkeley.