Bayes linear kinematics in the analysis of failure rates and failure time distributions

Kevin J Wilson and Malcolm Farrow*

School of Mathematics and Statistics, Newcastle University, Newcastle upon Tyne, UK

October 11, 2010

Abstract

Collections of related Poisson or binomial counts arise, for example, from numbers of failures in similar machines or neighbouring time periods. A conventional Bayesian analysis requires a rather indirect prior specification and intensive numerical methods for posterior evaluations.

An alternative approach using Bayes linear kinematics [1] in which simple conjugate specifications for individual counts are linked through a Bayes linear belief structure is presented. Intensive numerical methods are not required. The use of transformations of the binomial and Poisson parameters is proposed. The approach is illustrated in two examples, one involving Poisson counts of failures, the other involving binomial counts in an analysis of failure times.

Keywords: Bayesian inference, Bayes linear kinematics, count data, failure rates, failure times.

1   Introduction

This paper is concerned with the analysis of collections of quantities with conditional binomial or Poisson distributions. Such situations occur frequently. For example, [2] investigated the numbers of defects on items in industrial experiments. The number of defects on each of thirty masks over each of eighteen runs was assigned a Poisson distribution with the rate parameter given by the expected number of defects on the item in that run. An example in [3] involved the numbers of ruptures in pipelines over six years. The pipelines were categorised into eight systems by depth, diameter and site and were further categorised by year. Each combination of system and year in these data corresponds to a Poisson random variable with a mean specific to that system and that year. In [4] the number of successful starts of each of a collection of emergency diesel generators (EDGs) for nuclear power stations followed a binomial distribution. The number of trials was the number of demands for that EDG and the unknown parameter of interest was the probability that the EDG started successfully.

In this paper a subjective analysis from the point of view of an interested party, termed "the expert", who has prior beliefs about the collection but who may, understandably, be unwilling or unable to specify a complete joint prior probability distribution over the whole set of unknowns is considered. Typically the individual binomial or Poisson parameters are not independent in the expert's prior beliefs. For example, in the case of the pipelines data, if a larger than expected number of ruptures were observed in one of the systems in the first year this may very well lead to a revision upwards of the expected numbers of ruptures in the same system in subsequent years.

Typically such data are analysed using a generalised linear model with, typically, the linear predictors related via a linear model to a set of coefficients which are given a multivariate normal prior distribution. See, for example, [5, 6]. Marginal predictive distributions are thus of rather complicated form, making prior elicitation more difficult. Computation of posterior distributions requires numerical methods, usually Markov chain Monte Carlo (MCMC). In this paper an alternative approach is proposed which

---

*Corresponding author: School of Mathematics and Statistics, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK. Email: malcolm.farrow@newcastle.ac.uk

gives conjugate marginal prior distributions to the unknown parameters of the binomial and Poisson distributions. The parameters are transformed to map them onto the whole real line and then Bayes linear methods are used to define a coherent covariance structure between these quantities. These are examples of the type of structure termed a Bayes linear Bayes graphical model by [1]. Updates as a result of observing data are achieved via Bayes linear kinematics, a form of Bayes linear analysis in which changes in belief about certain quantities upon observation of data are propagated through to the other quantities via Bayes linear updating rules.

The resulting analysis does not involve the use of a joint probability distribution for the transformed parameters and thus focusses more directly on the quantities of interest, with a relatively simple structure relating them, and, importantly, does not require the use of intensive numerical methods such as MCMC to evaluate revised beliefs when data are observed. Moreover the structure is preserved in the same form so that further updates can be made in the same way when more data are observed.

A brief review of generalised linear modelling approaches to collections of Poisson and binomial random quantities is given in Section 2.1. Section 2.2 provides an overview of Bayes linear methods and Section 2.3 gives an introduction to Bayes linear kinematics. Sections 3 and 4 respectively outline the Bayes linear kinematic approach as applied to failure rates and to failure time distributions in the form of life tables.

2 Bayes linear kinematics and Bayes linear Bayes analysis for count data

2.1 Generalised linear modelling

Counts $Y_1, \ldots, Y_n$ are observed. Suppose that the expert's beliefs are such that, conditional on the values of unknown parameters $\theta_1, \ldots, \theta_n$, either $Y_i \sim \text{Poisson}(\theta_i)$ or $Y_i \sim \text{bin}(N_i, \theta_i)$, where $\text{bin}(N_i, \theta_i)$ denotes a binomial distribution with known number of trials N_i, for $i = 1, \ldots, n$, and Y_i, Y_j are conditionally independent given θ_i, θ_j for $i \neq j$.

Typically a link function $g()$ is used to transform θ_i to a linear predictor $\eta_i = g(\theta_i)$ with $-\infty < \eta_i < \infty$. In a Poisson model the natural logarithm $\eta_i = \log(\theta_i)$ is commonly used leading to log-linear models ([7, 8]). In a binomial model common link functions used are the logit, $\eta_i = \log(\theta_i/[1 - \theta_i])$, probit, $\eta_i = \Phi^{-1}(\theta_i)$, where $\Phi^{-1}()$ is the inverse of the standard normal distribution function, and complementary log-log, $\eta_i = \log(-\log[1 - \theta_i])$, link functions ([7, 9]). The linear predictors $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)'$ are then related via a linear model $\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\gamma}$ to a vector of unknown coefficients $\boldsymbol{\gamma}$ which are then given a multivariate normal prior distribution. This induces a multivariate normal prior distribution over $\boldsymbol{\eta}$. This non-conjugate structure makes prior elicitation awkward and requires intensive numerical methods for posterior computation. The design matrix $\mathbf{X}$ contains the values of explanatory variables.

In this standard approach the likelihood and prior distributions are not conjugate. Therefore numerical methods, usually MCMC, are necessary to evaluate posterior distributions. This can seem rather heavy handed for apparently simple problems and can become a major obstacle in problems such as experimental design [10]. In such situations a method for analysing related Poisson and binomial distributions without the necessity for intensive numerical methods is desirable.

A problem which is closely related to the subject of this paper occurs in the analysis of non-Gaussian time series. The dynamic generalised linear model was proposed by [11]. Suppose, for example, that the observation Y_t at time t has conditionally a binomial or Poisson distribution with parameter θ_t, given the value of an unobserved state vector $\boldsymbol{\omega}_t$. As a means of seeing how to use information from observations in forecasting future observations, [11] suggest that a conjugate prior distribution for θ_t is interposed between θ_t and $\boldsymbol{\omega}_t$. The parameters of this prior distribution are determined by the first two moments of a transformed unknown $\eta_t = g(\theta_t)$ where $g()$ is a suitable link function. On observation of Y_t a conjugate update of the prior distribution for θ_t to the corresponding posterior distribution can be carried out and this, in turn, changes the moments of η_t. A full Bayesian analysis would generally involve intractable distributions for $\boldsymbol{\omega}_t$. Instead [11] do not specify a distribution for $\boldsymbol{\omega}_t$ but work simply in terms of its first two moments. On observation of Y_t the moments of $\boldsymbol{\omega}_t$ are updated by a Bayes linear adjustment. The state-evolution model then determines moments for $\boldsymbol{\omega}_{t+1}$, the state vector at the next time step. This in turn determines the moments for η_{t+1} and therefore the parameters of the conjugate prior distribution

for θ_{t+1} and so on. (Note that the notation here differs from [11]). The approach which is adopted in this paper is similar. However, whereas in the time-series context it is natural to think in terms of a time-ordered sequence of observations, adjustments and so on, in other contexts there is no such fixed ordering and, as described below, the question of commutativity with respect to the order of observations arises.

2.2 Bayes linear methods

In a traditional Bayesian analysis a full joint prior distribution is specified for all observables and unknown quantities such as parameters. Prior beliefs are then updated, by conditioning on the observations and using Bayes theorem, and posterior distributions are calculated. A Bayes linear analysis [12] differs from a full Bayesian analysis in that only first and second order moments are specified in the prior. Posterior (termed *adjusted*) moments are then calculated.

For example, for each random quantity Q in the analysis, specifications of $E_0(Q)$ and $Var_0(Q)$, the prior expectation and variance of Q, are made and for every two quantities Q_1 and Q_2 a prior covariance $Cov_0(Q_1, Q_2)$ is also specified. Consider two vectors $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)'$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_r)'$ where $\boldsymbol{\alpha}$ is a collection of quantities which shall be observed and $\boldsymbol{\beta}$ is a collection of quantities about which inferences are to be made. Suppose that a full second order prior specification has been made for the set $\boldsymbol{A} = \boldsymbol{\alpha} \cup \boldsymbol{\beta}$.

Bayes linear methods [12] offer a procedure by which beliefs about $\boldsymbol{\beta}$ are updated by a process of linear fitting on $\boldsymbol{\alpha}$ using the Bayes linear updating equations for the adjusted expectation and variance of $\boldsymbol{\beta}$ given $\boldsymbol{\alpha}$:

$$\begin{aligned}
E_1(\boldsymbol{\beta};\ \boldsymbol{\alpha}) &= E_0(\boldsymbol{\beta}) + Cov_0(\boldsymbol{\beta}, \boldsymbol{\alpha})Var_0^{-1}(\boldsymbol{\alpha})[\boldsymbol{\alpha} - E_0(\boldsymbol{\alpha})] \\
Var_1(\boldsymbol{\beta};\ \alpha) &= Var_0(\boldsymbol{\beta}) - Cov_0(\boldsymbol{\beta}, \boldsymbol{\alpha})Var_0^{-1}(\boldsymbol{\alpha})Cov_0(\boldsymbol{\alpha}, \boldsymbol{\beta}),
\end{aligned}$$

when $Var_0(\boldsymbol{\alpha})$ is invertible. When this matrix is not invertible a suitable generalised inverse such as the Moore-Penrose inverse can be used.

2.3 Bayes linear kinematics

Probability kinematics [13] is a method for updating probabilities of further events E when beliefs about the elements of a partition P change in some way. Bayes linear kinematics [1] is the corresponding kinematic form of a Bayes linear analysis in which the effect of changes in belief about some quantities, rather than actual observations on them, are propagated through to others within a Bayes linear structure. Define the full second order prior specification for some vector random quantity $\boldsymbol{Q}$ to be

$$S_0(\boldsymbol{Q}) = [E_0(\boldsymbol{Q}), Var_0(\boldsymbol{Q})],$$

where $E_0(\boldsymbol{Q})$ is a vector of prior expectations and $Var_0(\boldsymbol{Q})$ is a prior covariance matrix. Suppose that, rather than directly observing $\boldsymbol{\alpha}$ in the previous section, data are observed which simply change beliefs about $\boldsymbol{\alpha}$ in some way. Thus the second order specification is now $S_1(\boldsymbol{\alpha})$ rather than $S_0(\boldsymbol{\alpha})$. Then the specification $S_1(\boldsymbol{A})$ is a Bayes linear kinematic update [1] if it satisfies

$$E_0(\boldsymbol{\beta};\ \boldsymbol{\alpha}) = E_1(\boldsymbol{\beta};\ \boldsymbol{\alpha}), \quad Var_0(\boldsymbol{\beta};\ \boldsymbol{\alpha}) = Var_1(\boldsymbol{\beta};\ \boldsymbol{\alpha}),$$

where $E_i(\boldsymbol{\beta};\ \boldsymbol{\alpha})$ and $Var_i(\boldsymbol{\beta};\ \boldsymbol{\alpha})$ are the Bayes linear adjusted expectation and variance of $\boldsymbol{\beta}$ given $\boldsymbol{\alpha}$ using $S_i(\boldsymbol{A})$. These conditions lead to the Bayes linear kinematic updating equations:

$$\begin{aligned}
E_1(\boldsymbol{A}) &= E_0(\boldsymbol{A}) + Cov_0(\boldsymbol{A}, \boldsymbol{\alpha})Var_0^{-1}(\boldsymbol{\alpha})[E_1(\boldsymbol{\alpha}) - E_0(\boldsymbol{\alpha})], & (1) \\
Var_1(\boldsymbol{A}) &= Var_0(\boldsymbol{A};\ \boldsymbol{\alpha}) + Cov_0(\boldsymbol{A}, \boldsymbol{\alpha})Var_0^{-1}(\boldsymbol{\alpha})Var_1(\boldsymbol{\alpha})Var_0^{-1}(\boldsymbol{\alpha})Cov_0(\boldsymbol{\alpha}, \boldsymbol{A}). & (2)
\end{aligned}$$

This is also the case if $\boldsymbol{A}$ is replaced by $\boldsymbol{\beta}$ in the above equations. Now consider the situation in which there are p collections of random quantities $\boldsymbol{U}_1, \ldots, \boldsymbol{U}_p$ where $\boldsymbol{U}_k = (U_{k1}, \ldots, U_{kn_k})'$ for $k = 1, \ldots, p$. Suppose that a full second order prior specification has been made for $\boldsymbol{U} = \boldsymbol{U}_1 \cup \ldots \cup \boldsymbol{U}_p$ of the form $S_0(\boldsymbol{U}) = [E_0(\boldsymbol{U}), Var_0(\boldsymbol{U})]$ and that data information I_k is received which causes the beliefs about $\boldsymbol{U}_k$

3

to be updated to $S_1(\boldsymbol{U}_k;\ I_k) = [\mathrm{E}_1(\boldsymbol{U}_k), \mathrm{Var}_1(\boldsymbol{U}_k)]$. Then, as in (1) and (2), the Bayes linear kinematic update for $\boldsymbol{U}$ is

$$\mathrm{E}_1(\boldsymbol{U};\ I_k) = \mathrm{E}_0(\boldsymbol{U}) + \mathrm{Cov}_0(\boldsymbol{U}, \boldsymbol{U}_k)\mathrm{Var}_0^{-1}(\boldsymbol{U}_k)[\mathrm{E}_1(\boldsymbol{U}_k) - \mathrm{E}_0(\boldsymbol{U}_k)], \tag{3}$$

$$\mathrm{Var}_1(\boldsymbol{U};\ I_k) = \mathrm{Var}_0(\boldsymbol{U};\ I_k) + \mathrm{Cov}_0(\boldsymbol{U}, \boldsymbol{U}_k)\mathrm{Var}_0^{-1}(\boldsymbol{U}_k)\mathrm{Var}_1(\boldsymbol{U}_k)\mathrm{Var}_0^{-1}(\boldsymbol{U}_k)\mathrm{Cov}_0(\boldsymbol{U}_k, \boldsymbol{U}). \tag{4}$$

Now suppose that data are observed and beliefs updated once for each of $k = 1, \ldots, p$. A Bayes linear kinematic update can be made for $\boldsymbol{U}$ each time. As in probability kinematics, successive Bayes linear kinematic updates are not necessarily commutative. However, [1] give conditions under which the requirement of commutativity leads to a unique Bayes linear kinematic update. In the analyses in this paper each $\boldsymbol{U}_k$ is always a scalar U_k and a sufficient condition for a unique commutative update is

$$\mathrm{Var}_0^{-1}(U_k)\mathrm{Var}_1(U_k) < 1, \tag{5}$$

for all $k = 1, \ldots, p$. This solution, when it exists, is given by

$$\mathrm{E}_p(\boldsymbol{U};\ I_1, \ldots, I_p) = \mathrm{Var}_p(\boldsymbol{U})\left[\sum_{k=1}^{p}\mathrm{Var}_1^{-1}(\boldsymbol{U};\ I_k)\mathrm{E}_1(\boldsymbol{U};\ I_k) - (p-1)\mathrm{Var}_0^{-1}(\boldsymbol{U})\mathrm{E}_0(\boldsymbol{U})\right], \tag{6}$$

$$\mathrm{Var}_p(\boldsymbol{U};\ I_1, \ldots, I_p) = \left[\sum_{k=1}^{p}\mathrm{Var}_1^{-1}(\boldsymbol{U};\ I_k) - (p-1)\mathrm{Var}_0^{-1}(\boldsymbol{U})\right]^{-1}. \tag{7}$$

2.4 Bayes linear Bayes analysis for count data

2.4.1 Conjugate Bayesian updates

If $Y_i \mid \theta_i \sim \mathrm{Poisson}(\theta_i)$ then the natural conjugate prior distribution is a gamma(a_i, b_i) distribution and, given an observation y_i, the posterior distribution is gamma$(a_i + y_i,\ b_i + 1)$. More generally a known scale factor s_i, perhaps the time at risk, could be included. In this case $Y_i \mid \theta_i \sim \mathrm{Poisson}(s_i\theta_i)$ and the posterior distribution is gamma$(a_i + y_i,\ b_i + s_i)$, with $s_i = 1$ as a special case. The prior mean and variance are

$$\mathrm{E}_0(\theta_i) = \frac{a_i}{b_i}, \quad \mathrm{Var}_0(\theta_i) = \frac{a_i}{b_i^2} \tag{8}$$

and the posterior mean and variance are

$$\mathrm{E}_1(\theta_1) = \frac{a_i + y_i}{b_i + s_i}, \quad \mathrm{Var}_1(\theta_i) = \frac{a_i + y_i}{(b_i + s_i)^2}.$$

Notice that the posterior variance can be greater than the prior variance if y_i is sufficiently large.

If $Y_i \mid \theta_i \sim \mathrm{bin}(N_i, \theta_i)$ then the natural conjugate prior is a beta(a_i, b_i) distribution and, given an observation y_i, the posterior distribution is beta$(a_i + y_i,\ b_i + N_i - y_i)$. The prior mean and variance are

$$\mathrm{E}_0(\theta_i) = \frac{a_i}{a_i + b_i}, \quad \mathrm{Var}_0(\theta_i) = \frac{a_i b_i}{(a_i + b_i)^2(a_i + b_i + 1)} \tag{9}$$

and the posterior mean and variance are

$$\mathrm{E}_1(\theta_i) = \frac{a_i + y_i}{a_i + b_i + N_i}, \quad \mathrm{Var}_1(\theta_i) = \frac{(a_i + y_i)(b_i + N_i - y_i)}{(a_i + b_i + N_i)^2(a_i + b_i + N_i + 1)}.$$

Again it is possible for the posterior variance to be greater than the prior variance, for example with $a_i = 7$, $b_i = 1$, $N_i = 4$, $y_i = 2$.

2.4.2 Use of transformations

It would be possible to proceed by linking the parameters $\theta_1, \ldots, \theta_n$ in a Bayes linear structure. However there are advantages in transforming the parameters first. The transformed parameters $\eta_1, \ldots, \eta_n$ are then linked in a Bayes linear structure. Observing Y_i changes the mean and variance of η_i and the effect of this change is propagated by Bayes linear kinematics. The reasons for using the transformation are as follows. Firstly, in both cases, the range of θ_i is bounded. In the Poisson case $0 < \theta_i$ and, in the binomial case, $0 < \theta_i < 1$. The combination of linear updates with bounded parameter spaces seems undesirable both in terms of first and second moments. If information leads to adjustment of the expectation for a quantity towards a boundary, it seems clear that this adjustment should not continue to be linear as the boundary is approached. It is to be expected that variances will be affected by the proximity of a boundary and beliefs, when the mean is close to a boundary, will no longer be symmetric in the sense that deviations from the mean in either direction would be regarded in the same way. Similarly there are difficulties with covariances in bounded spaces where the tendency would be to imagine rather nonlinear relationships between unknowns close to boundaries. So it is desirable to transform the parameters onto unbounded spaces. Secondly, as seen above, it is possible for the variances of the untransformed parameters θ_i to increase when data are observed. While [1] (Theorem 5) give conditions for the existence of unique Bayes linear kinematic updates which allow some such variance increase, the transformations have the effect of making reductions in variance of the transformed parameters occur when observations are made, at least in most circumstances, and therefore allow use of the simpler sufficient condition given in Corollary 4 of [1].

Bayes linear kinematics, without transformation, gives a rule for adjusting beliefs about $\theta_1, \ldots, \theta_n$ by Bayes linear updates. Similarly Bayes linear kinematics, with the transformation, gives a Bayes linear rule for updating beliefs about $\eta_1, \ldots, \eta_n$, where there is a 1 - 1 relationship between η_i and θ_i. Any further use of conjugate Bayesian updating of beliefs about θ_j, given observation of Y_j, after already adjusting by observation of Y_i, relies on the idea that θ_j still has a distribution of the required conjugate form, whether or not a transformation is used. Similarly evaluating predictive distributions for new observations or credible intervals for $\theta_1, \ldots, \theta_n$ depends on such an idea. Additionally, when a transformation is used, this preserved conjugate form is required in order to convert back from the adjusted moments of η_j to the new distribution for θ_j.

Clearly, if adjustments were only ever made in one direction, eg. of beliefs about θ_j by observing Y_i, and this was never reversed to adjust beliefs about θ_i by observing Y_j, then it could simply be declared that the conditional distribution was the required conjugate distribution. (Such one-way belief adjustment might be appropriate, for example, in a time-series forecasting context, as in [11]). When commutativity, in the strong sense that conjugate updates of the marginal distributions of $\theta_1, \ldots, \theta_n$ are always appropriate, is required then this might be regarded as a pragmatic approximation which does not correspond exactly to a full Bayesian conditioning analysis. With no transformation, this assumption is made directly on the distributions of $\theta_1, \ldots, \theta_n$ under Bayes linear kinematic updates. With transformation, the assumption applies to the corresponding distributions of $\eta_1, \ldots, \eta_n$, eg, log-gamma distributions, in the same way.

2.4.3 Transformations in the Poisson and binomial cases

In the Poisson case the transformation $\eta_i = \log(\theta_i)$ is used. It is necessary to work with moments for both θ_i for the conjugate updates and η_i for the Bayes linear kinematic updates. The expectation and variance of θ_i are found using (8). The expectation and variance of η_i can be found using the fact that

$$\frac{1}{\Gamma(a_i)} \frac{d^r}{da_i^r} \Gamma(a_i) = \int_0^\infty [\log(z)]^r \frac{z^{a_i-1} e^{-z}}{\Gamma(a_i)} \, dz = \mathrm{E}\{[\log(b_i \theta_i)]^r\}.$$

Setting $r = 1$ gives

$$\mathrm{E}_0(\eta_i) + \log(b_i) = \mathrm{E}_0\{\log(b_i \theta_i)\} = \frac{d}{da_i}\{\log[\Gamma(a_i)]\} = \psi(a_i)$$

where $\psi()$ is the digamma function. Setting $r = 2$ gives

$$\frac{d^2}{da_i^2} \Gamma(a_i) = \frac{d}{da_i} \Gamma(a_i) \psi(a_i) = \Gamma(a_i)\psi_1(a_i) + \Gamma(a_i)\psi(a_i)^2,$$

where $\psi_1(x) = \frac{d}{dx}\psi(x)$ is the trigamma function. Thus

$$
\begin{aligned}
\text{Var}_0(\eta_i) &= \text{Var}_0(\log b_i\theta_i) = \text{E}_0[(\log b_i\theta_i)^2] - \text{E}_0[\log b_i\theta_i]^2 \\
&= \psi_1(a_i) + \psi(a_i)^2 - \psi(a_i)^2 = \psi_1(a_i).
\end{aligned}
$$

After the conjugate updates the expectations and variances of both θ_i and η_i remain of the same form but with a_i and b_i replaced with $a_i + y_i$ and $b_i + 1$ respectively. Thus for η_i they are

$$
\begin{aligned}
\text{E}(\eta_i \mid Y_i = y_i) &= \psi(a_i + y_i) - \log(b_i + 1) \\
\text{Var}(\eta_i \mid Y_i = y_i) &= \psi_1(a_i + y_i).
\end{aligned}
$$

Suppose that a full second order prior specification has been made for $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)'$ of the form $S_0(\boldsymbol{\eta}) = (\text{E}_0(\boldsymbol{\eta}), \text{Var}_0(\boldsymbol{\eta}))$. Observing $Y_i = y_i$ leads to the Bayes linear kinematic adjusted expectation and variance for $\boldsymbol{\eta}$ of

$$
\text{E}_1(\boldsymbol{\eta};\ Y_i = y_i) = \text{E}_0(\boldsymbol{\eta}) + \text{Cov}_0(\boldsymbol{\eta}, \eta_i)\text{Var}_0^{-1}(\eta_i)\left[\text{E}(\eta_i \mid Y_i = y_i) - \text{E}_0(\eta_i)\right],
$$

and

$$
\begin{aligned}
\text{Var}_1(\boldsymbol{\eta};\ Y_i = y_i) = \text{Var}_0(\boldsymbol{\eta}) &- \text{Cov}_0(\boldsymbol{\eta}, \eta_i)\text{Var}_0^{-1}(\eta_i)\text{Cov}_0(\eta_i, \boldsymbol{\eta}) \\
&+ \text{Cov}_0(\boldsymbol{\eta}, \eta_i)\text{Var}_0^{-1}(\eta_i)\text{Var}_0(\eta_i \mid Y_i = y_i)\text{Var}_0^{-1}(\eta_i)\text{Cov}_0(\eta_i, \boldsymbol{\eta}).
\end{aligned}
$$

which depend only on the prior specifications and fully Bayesian conjugate updates which have already been calculated. Here, for example, $\text{E}_1(\boldsymbol{\eta};\ Y_i = y_i)$ denotes the adjusted expectation after 1 observation has been made and the observation is given after the semicolon.

Now consider whether, having observed $\boldsymbol{y} = (y_1, \ldots, y_n)'$, there is a unique commutative Bayes linear kinematic update for $\boldsymbol{\eta}$. From [1] a unique commutative solution exists if

$$
\text{Var}_0^{-1}(\eta_i)\text{Var}(\eta_i \mid Y_i = y_i) < 1
$$

for all i. The variances are $\text{Var}_0(\eta_i) = \psi_1(a_i)$ and $\text{Var}(\eta_i \mid Y_i = y_i) = \psi_1(a_i + y_i)$. Each y_i must be a nonnegative integer. The trigamma function is monotonically decreasing on $\mathbb{R}_+$ and $\psi_1(x) \to 0$ as $x \to \infty$ so, as long as $y_i > 0$ for each i, $\text{Var}(\eta_i \mid Y_i = y_i) < \text{Var}_0(\eta_i)$ for all i and the uniqueness condition is met. If this is the case then the Bayes linear kinematic unique commutative solution is

$$
\text{E}_n(\boldsymbol{\eta};\ \boldsymbol{Y} = \boldsymbol{y}) = \text{Var}_n(\boldsymbol{\eta};\ \boldsymbol{Y} = \boldsymbol{y})\left[\sum_{i=1}^n \text{Var}_1^{-1}(\boldsymbol{\eta};\ Y_i = y_i)\text{E}_1(\boldsymbol{\eta};\ Y_i = y_i) - (n-1)\text{Var}_0^{-1}(\boldsymbol{\eta})\text{E}_0(\boldsymbol{\eta})\right]
$$

$$
\text{Var}_n(\boldsymbol{\eta};\ \boldsymbol{Y} = \boldsymbol{y}) = \left[\sum_{i=1}^n \text{Var}_1^{-1}(\boldsymbol{\eta};\ Y_i = y_i) - (n-1)\text{Var}_0^{-1}(\boldsymbol{\eta})\right]^{-1}.
$$

In fact, it is easily seen that these equations define a commutative update even if $y_i = 0$ for some i and this case satisfies the more general conditions in Theorem 5 of [1]. Hence the commutative update always exists in the gamma-Poisson case.

Having found the revised expectations and variances of every η_i, the means and variances of θ_i can be found by first solving the following equations for a_i^* and b_i^*.

$$
\begin{aligned}
\text{E}_n(\eta_i;\ \boldsymbol{Y} = \boldsymbol{y}) &= \psi(a_i^*) - \log(b_i^*) \\
\text{Var}_n(\eta_i;\ \boldsymbol{Y} = \boldsymbol{y}) &= \psi_1(a_i^*).
\end{aligned}
$$

The revised mean and variance of θ_i are then

$$
\text{E}_n(\theta_i;\ \boldsymbol{Y} = \boldsymbol{y}) = \frac{a_i^*}{b_i^*}, \quad \text{Var}_n(\theta_i;\ \boldsymbol{Y} = \boldsymbol{y}) = \frac{a_i^*}{b_i^{*2}}.
$$

In the binomial case any of the usual link functions $\mu_i = g(\theta_i)$ might be suitable. However the moments of μ_i are generally not straightforward so, for the Bayes linear kinematic updates the use of a related

quantity η_i, is proposed, which has mean and variance found from the mode of μ_i and the curvature at the mode of the log density of μ_i. Thus the link function is regarded as a "guide relationship" ([11]). For example, with the logistic link function, $\mu_i = \log[\theta_i/(1 - \theta_i)]$ and a beta(a_i, b_i) distribution for θ_i, it is easily shown that the mode of μ_i, that is the mean of η_i, is $m_i = \log(a_i/b_i)$ and the corresponding variance is $\text{Var}(\eta_i) = a_i^{-1} + b_i^{-1}$. This variance clearly decreases when any observation y_i is made.

For the failure-time application, as discussed in Section 4 below, the complementary log-log link function is chosen as it is more convenient for computation of the reliability $R(t) = \Pr(T \geq t)$, where T is a lifetime. Hence

$$\mu_i = g(\theta_i) = \log[-\log(1 - \theta_i)]. \tag{10}$$

If $\theta_i \sim \text{beta}(a_i, b_i)$, with density $f_{\theta,i}(\theta_i)$, and $\mu_i = \log[-\log(1 - \theta_i)]$ then the density of μ_i is

$$f_{\mu,i}(\mu_i) = f_{\theta,i}(\theta_i)\frac{d\theta_i}{d\mu_i} = \frac{\Gamma(a_i + b_i)}{\Gamma(a_i)\Gamma(b_i)}e^{\mu_i}\exp[-e^{\mu_i}]\theta_i^{a_i-1}(1 - \theta_i)^{b_i-1}.$$

Taking logs,

$$l_i(\mu_i) = \log\{f_{\mu,i}(\mu_i)\} = k_i + \mu_i - e^{\mu_i} + (a_i - 1)\log(\theta_i) + (b_i - 1)\log(1 - \theta_i)$$

where k_i is a constant. To find the mode, m_i,

$$\left(\frac{dl_i(\mu_i)}{d\mu_i}\right)_{m_i} = 1 - e^{m_i} + \left[\frac{(a_i - 1)}{\theta_{m,i}} - \frac{(b_i - 1)}{1 - \theta_{m,i}}\right]e^{m_i}\exp[-e^{m_i}] = 0, \tag{11}$$

where $\theta_{m,i} = 1 - \exp[-e^{m_i}]$, is solved numerically. The second derivative is

$$\begin{aligned}\frac{d^2l_i(\mu_i)}{d\mu_i^2} &= -e^{\mu_i} - \left[\frac{(a_i - 1)}{\theta_i^2} + \frac{(b_i - 1)}{(1 - \theta_i)^2}\right]e^{2\mu_i}\exp[-2e^{\mu_i}] \\ &\quad + \left[\frac{(a_i - 1)}{\theta_i} - \frac{(b_i - 1)}{1 - \theta_i}\right]e^{\mu_i}(1 - e^{\mu_i})\exp[-e^{\mu_i}]. \end{aligned} \tag{12}$$

The mean and variance of η_i can then be found as

$$\text{E}(\eta_i) = m_i, \quad \text{Var}(\eta_i) = -\left[\frac{d^2l_i(\mu_i)}{d\mu_i^2}\right]_{m_i}^{-1}.$$

Having made the conjugate updates, the same procedure can be applied but using $A_i = a_i + y_i$ and $B_i = b_i + N_i - y_i$ in place of a_i and b_i in the density and subsequent derivatives. Defining a Bayes linear structure for $\eta_1, \ldots, \eta_p$, i.e., specification of $\text{Cov}_0(\boldsymbol{\eta})$, allows the updates to be propagated to the other quantities in $\boldsymbol{\eta}$ via (3) and (4). Note that, once an adjusted mean and precision for η_j are found, (11) and (12) provide simultaneous linear equations in a_j^* and b_j^*, the new values of a_j and b_j, which are easily solved.

From (5) there is a unique commutative solution to the problem using Bayes linear kinematics if $\text{Var}_0^{-1}(\eta_i)\text{Var}(\eta_i \mid y_i) < 1$, $\forall i$. An analytic proof that this condition always holds is not yet available. However this has been investigated numerically. It is only necessary to consider the effect of a single observation $y = 1$ with $N = 1$. This is because this is equivalent to the observation $y = 0$ with $N = 1$ with a and b exchanged and any observation with larger N has the cumulative effect of a sequence of observations with $N = 1$. The increase in the precision of η_i given an observation $y_i = 1$ with $N_i = 1$ was investigated over a rectangular grid of values of (a_i, b_i) with $-1 \leq \log(a_i) \leq 12$ and $-1 \leq \log(b_i) \leq 12$ in steps of 0.1 and every value was positive.

3 Bayes linear kinematics for failure rates

3.1 Example: failure rates of piston-rings

Data are presented in [14], reproduced in [15], on the numbers of failures of piston-rings in four steam driven compressors over a number of years. Within each compressor there are three legs: north $(i = 1)$,

Table 1: ABOUT HERE

centre ($i = 2$) and south ($i = 3$). The south leg of each compressor is adjacent to the drive. The numbers of failures Y_{ij} in each leg i of each compressor j are given in Table 1.

Questions of interest for these data are

1. whether the rate of piston-ring failures varies between compressors

2. whether the rate of piston-ring failures varies between legs

3. whether the pattern of the location of failures is different for different compressors.

Given θ_{ij}, $Y_{ij} \mid \theta_{ij} \sim \text{Poisson}(\theta_{ij})$. Let $\eta_{ij} = \log(\theta_{ij})$. There are twelve conjugate updates to perform, one for each element of $\mathbf{Y} = (Y_{11}, \ldots, Y_{14}, Y_{21}, \ldots, Y_{24}, \ldots, Y_{34})'$.

3.2 Elicitation of prior beliefs

Now consider the process of specification of the expert's prior beliefs. Some general advice on elicitation is given in [16, 17, 18, 19]. Generally at least two statistics are elicited for each quantity of interest, one measure of centrality and at least one of spread. Quantiles are commonly chosen as non-statisticians tend to assess medians better than means for skewed distributions ([19, 20, 21]) and are generally poor at assessing variances. Whilst upper and lower quartiles are the most commonly used quantiles to estimate the spread of a distribution, [17, 22] recommend eliciting tertiles (33% and 67% points) as there is evidence this can reduce overconfidence.

In terms of the piston-rings example the elicitation process consists of finding the parameters a_{ij}, b_{ij} of the marginal gamma distributions and eliciting prior covariances between the η_{ij}'s.

3.2.1 Elicitation of prior expectations and variances

To find a_{ij} and b_{ij}, quantiles can be elicited for the gamma prior distribution of θ_{ij} and, following [17], the median $q_{ij}(1/2)$ and lower and upper tertiles $q_{ij}(1/3)$ and $q_{ij}(2/3)$ are chosen. To perform these elicitations questions are put to the expert in terms of the average number of failures per unit time over a very long period.

The compressors are identically designed and are all oriented the same way. Suppose that, *a priori*, there is no reason to believe any leg of any compressor would be more prone to failures than any other. Thus the marginal elicitation process reduces to eliciting a single median $q(1/2)$ and a single lower and upper tertile $q(1/3)$ and $q(2/3)$ for the failure rate θ in any leg of any compressor. Since three elicitations ($q(1/3)$, $q(1/2)$, $q(2/3)$) are made to determine two parameters a, b, in general there is no exact solution. However, in the case of the gamma distribution, the ratio of upper and lower tertiles, $q(2/3)/q(1/3)$ depends only on the shape parameter a and a numerical method can be used to find a value for a from this ratio. The elicited median $q(1/2)$ can then be used to find b.

3.2.2 Elicitation of prior covariances

To complete the second order prior specification, covariances for $\boldsymbol{\eta}$ must be specified. This is achieved by eliciting quantities involving the θ_{ij}'s. For each pair θ_{ij} and θ_{kl} with $(i, j) \neq (k, l)$, a prior covariance $\text{Cov}_0(\eta_{ij}, \eta_{kl})$ is required. This covariance is elicited by asking the expert to suppose that the value of θ_{kl}, the population average number of piston-ring failures per unit of time over a very long period, is now known and indicating that this has left the median for θ_{ij} unchanged at $q_{ij}(1/2)$. New tertiles, $q'_{ij}(1/3)$ and $q'_{ij}(2/3)$, are then elicited for θ_{ij} having learned θ_{kl}.

From these the parameters a'_{ij}, b'_{ij} of the gamma distribution can be found as when making the marginal prior specifications. Thus $\theta_{ij} \mid \theta_{kl} \sim \text{gamma}(a'_{ij}, b'_{ij})$. If the expert judges that θ_{ij} and θ_{kl} are unrelated then $q'_{ij}(1/3) = q_{ij}(1/3)$ and $q'_{ij}(2/3) = q'_{ij}(2/3)$ i.e., the elicited tertiles would remain unchanged as nothing has been learned about θ_{ij} by learning θ_{kl}. If the expert judges that there is a

relation between θ_{ij} and θ_{kl} then $q'_{ij}(1/3) > q_{ij}(1/3)$ and $q'_{ij}(2/3) < q_{ij}(2/3)$ i.e., the elicited tertiles will have moved closer together indicating a reduction in uncertainty about θ_{ij} having learned θ_{kl}. The closer together the tertiles become, the stronger the association between the two quantities.

Transforming back to η_{ij} gives $\text{Var}(\eta_{ij} \mid \eta_{kl}) = \psi_1(a'_{ij})$ and since

$$\frac{\text{Var}(\eta_{ij} \mid \eta_{kl})}{\text{Var}_0(\eta_{ij})} = 1 - \frac{\text{Cov}_0^2(\eta_{ij}, \eta_{kl})}{\text{Var}_0(\eta_{ij})\text{Var}_0(\eta_{kl})},$$

the modulus of the required prior covariance can be found. The sign of the covariance is determined by asking whether the expert's expectation for θ_{ij} would increase or decrease upon learning that θ_{kl} was greater than expected. This method is based on that which was used in the projects described in [23, 24, 18].

3.3   Results

Suppose that the expert settles on values of $q(1/3) = 11$ and $q(2/3) = 20$ for the lower and upper tertiles following the elicitation process. This leads to $a = 2.441$. If the expert also gives a median $q(1/2) = 15$, b is found to be 0.1411. If the four compressors are judged to be exchangable and the legs within each compressor are also regarded as exchangable (which, of course, might not be the case), the elicitation of a covariance structure can be reduced to the specification of three different covariances:

$$\text{Cov}_0(\eta_{ij}, \eta_{kl}) = \begin{cases} c_1, & \text{when } i = k, j \neq l, \\ c_2, & \text{when } i \neq k, j = l, \\ c_3, & \text{when } i \neq k, j \neq l. \end{cases}$$

Table 2 shows an example of elicited adjusted tertiles in the above three cases and the resulting adjusted gamma parameter values and covariances and correlation of η_{ij}.

Table 2: ABOUT HERE

A unique commutative Bayes linear kinematic solution can be found in this example. For this solution the adjusted values of the gamma parameters are calculated. Figure 1 shows the adjusted expectations

Figure 1: ABOUT HERE

of the θ_{ij}'s and adjusted 95% symmetric credible intervals for each of the 12 legs. The dashed line on the plot is the observed mean number of piston-ring failures in the time period, 166/12.

Table 3: ABOUT HERE

The first four locations correspond to the north leg, the next four to the centre leg and the final four to the south leg. A full list of locations along with posterior moments are given in Table 3. It appears that location 12, the south leg of compressor 4, has an unusually high rate of piston-ring failures.

In this example, a unique commutative Bayes linear kinematic adjustment also exists if the transformation is not used. The bracketed figures in Table 3 show the results. The prior specification was

derived from the same elicited tertiles as in the analysis with the transformation. The results are similar but generally a little lower. It seems that the effect of the observations which are less than the prior mean may be greater when no tansformation is used. While such an analysis without transformations is possible in this example, and will be possible in many cases, there will exist examples where it is not.

4 Bayes linear kinematics for failure time distributions

4.1 Applying Bayes linear kinematics to life table data

The second application is to the analysis of failure times. In this paper only the case where failure times are grouped into intervals as in a life table is considered. The authors intend to address the case of non-grouped data in a future paper. The analysis of time-to-event data is important in other fields as well as engineering, particularly in survival analysis in medicine. Information on Bayesian survival analysis can be found in [25] and [26].

Suppose that failure times are classified into p intervals so that interval i is $[t_i, t_{i+1})$ for $i = 1, \ldots, p$ with $t_1 = 0$ and $t_{p+1} = \infty$. Suppose initially that there is no censoring so an interval is recorded for the failure of every item. Let the number of failures of items in interval i be y_i and the number of items which have not failed at the start of interval i be N_i. Then $N_i = N_{i-1} - y_{i-1}$.

Given a parameter θ_i, the number of failures in interval i follows a binomial distribution

$$y_i \mid \theta_i \sim \text{bin}(N_i, \theta_i),$$

where θ_i is the unknown population probability that an item fails in interval i given that it has not failed by time t_i. A beta prior $\theta_i \sim \text{beta}(a_i, b_i)$ is adopted. Observation of y_i failures in interval i leads to a within interval update of $\theta_i \mid y_i \sim \text{beta}(A_i, B_i)$, where $A_i = a_i + y_i$ and $B_i = b_i + N_i - y_i$. The prior expectation and variance of θ_i are given by (9) with the posterior counterparts $\text{E}(\theta_i \mid y_i)$ and $\text{Var}(\theta_i \mid y_i)$ the same but using A_i and B_i instead of a_i and b_i.

In this application there is interest in the reliability function $R(t) = \Pr(T \geq t)$, where T is a lifetime. This can be expressed at each of the interval boundaries in terms of the conditional probabilities of failure up to that interval. That is

$$R(t_i) = \Pr(T \geq t_i) = \prod_{j=1}^{i-1} (1 - \theta_j).$$

Hence

$$\log[R(t_i)] = \sum_{j=1}^{i-1} \log(1 - \theta_j).$$

Therefore it would appear to be convenient to work in terms of the transformed quantity $\mu_j = -\log(1-\theta_j)$. However this quantity is still bounded below at zero and, in fact, will often be close to the boundary. Therefore the complementary log-log link function is chosen to remove the boundary effect. Hence $\mu_i = g(\theta_i) = \log[-\log(1 - \theta_i)]$.

As discussed in Section 2.4.3 a unique commutative Bayes linear kinematic solution exists at least over a wide range of prior specifications. It can be made sequentially in time by rearranging (6) and (7) to

$$
\begin{aligned}
\text{Var}_{(i)}(\boldsymbol{\eta}) &= \left(\text{Var}_{(i-1)}^{-1}(\boldsymbol{\eta}) + \text{Var}_i^{-1}(\boldsymbol{\eta}) - \text{Var}_0^{-1}(\boldsymbol{\eta}) \right)^{-1} \\
\text{E}_{(i)}(\boldsymbol{\eta}) &= \text{Var}_{(i)}(\boldsymbol{\eta}) \left(\text{Var}_{(i-1)}^{-1}(\boldsymbol{\eta}) \text{E}_{(i-1)}(\boldsymbol{\eta}) + \text{Var}_i^{-1}(\boldsymbol{\eta}) \text{E}_i(\boldsymbol{\eta}) - \text{Var}_0^{-1}(\boldsymbol{\eta}) \text{E}_0(\boldsymbol{\eta}) \right),
\end{aligned}
$$

where $\text{E}_{(i)}(\boldsymbol{\eta})$, $\text{Var}_{(i)}(\boldsymbol{\eta})$ denote the adjusted expectation and variance after observing $y_1, \ldots, y_i$ and $\text{E}_i(\boldsymbol{\eta})$, $\text{Var}_i(\boldsymbol{\eta})$ are the adjusted moments after observing just y_i.

Using the complementary log-log link function gives

$$\log[R(t_i)] = -\sum_{j=1}^{i-1} \exp(\mu_j).$$

To see what might reasonably be concluded about this quantity, the guide relationship (10) can again be considered. After observing data it might be reasonable to suppose that the result of such a nonlinear transformation is approximated by what happens if $\boldsymbol{\eta}$ has a multivariate normal distribution and $\log(1 - \theta_i) = -\exp(\eta_i)$. So $\exp(\boldsymbol{\mu}) \mid \boldsymbol{y}$ can be regarded as having approximately the moments of a multivariate lognormal distribution. If $\mathrm{E}(\eta_i) = M_i$, $\mathrm{Var}(\eta_i) = V_{ii}$ and $\mathrm{Cov}(\eta_i, \eta_j) = V_{ij}$ then, following this guideline and writing $w_i = -\log[1 - \theta_i] = \exp(\eta_i)$,

$$\mathrm{E}(w_i) \approx e^{M_i + V_{ii}/2}, \quad \mathrm{Var}(w_i) \approx e^{2M_i + V_{ii}}(e^{V_{ii}} - 1), \quad \mathrm{Cov}(w_i, w_j) \approx e^{M_i + M_j + (V_{ii} + V_{jj})/2}(e^{V_{ii} + V_{jj} + 2V_{ij}} - 1).$$

The approximate mean and variance of $-\sum_{j=1}^{i-1} w_j$ are then easily found.

4.2 Right Censoring

Non-informative right censoring can easily be introduced into this method using the actuarial assumption. If r_i is the number of right censored observations in interval i then the number of items which have not failed by the start of the interval is $N_i = N_{i-1} - y_{i-1} - r_{i-1}$. By the actuarial assumption, the number of items at risk on average in interval i, N_i' is $N_i' = N_i - r_i/2$. The binomial distribution given to y_i is then $y_i \sim \mathrm{bin}(N_i', \theta_i)$, and the analysis proceeds exactly as in the case of no censoring above.

4.3 Example: Centrifuge cloths

Data are given in [27] on the failure times of sugar centrifuge cloths. In all there are 229 cloths and all fail within 78 weeks. There is no censoring in the data. The data are presented in Table 4.

Table 4: ABOUT HERE

As for the piston-rings example, the elicitation process contains two stages: elicitation of the median and tertiles for the marginal beta distributions and elicitation of a coherent covariance structure for $\boldsymbol{\eta}$. The marginal elicitation process is very similar to that for the piston-rings example. The parameters a_i, b_i of the marginal beta distributions are found from the elicited median and tertiles. Details are omitted.

The prior values for the a_i's and the b_i's resulting from the elicitation process are given in Table 5.

Table 5: ABOUT HERE

Covariances between different elements of $\boldsymbol{\eta}$ can be elicited using a method similar to that used in the piston-rings example. The expert can be asked to imagine knowing the value of θ_i, from a very large experiment, and provide revised tertiles for θ_j given that the "true" value of θ_i was found to be equal to its prior median. As in the Poisson case, this leads to a calculation of the reduction in variance of η_j given knowledge of η_i and hence to the covariance of η_i and η_j.

However, with a large number of intervals it may be unappealing to consider all of the covariances individually. It may also be difficult to avoid accidental incoherence in the resulting covariance matrix. In any case it may well give more satisfactory results to adopt a more structured approach. Therefore ideas from [18] are used to give $\mathrm{Var}_0(\boldsymbol{\eta})$ a more structured form.

For example, bearing in mind the ordering of the time intervals, uncertainties about $\boldsymbol{\eta}$ might well be represented by a stationary process. Let $F_i = \eta_i - \mathrm{E}_0(\eta_i)$ so that F_i is a zero expectation quantity

which depends on time. Then $F_1, \ldots, F_p$ can be linked via a stationary process such as a first order autoregression, in which case

$$F_i = \phi F_{i-1} + \varepsilon_i, \ (i = 2, \ldots, p)$$

where $\mathrm{E}(\varepsilon_i) = 0$, $\mathrm{E}(\varepsilon_i^2) = v_\varepsilon$ and $\mathrm{E}(\varepsilon_i \varepsilon_j) = 0$ for $i \neq j$. For stationarity the initial variance of F_1 is set at the stationary value

$$\mathrm{Var}_0(F_1) = \frac{v_\varepsilon}{1 - \phi^2} = v_F.$$

The covariances between $\eta_1, \ldots, \eta_p$ are now given by $\mathrm{Cov}_0(\eta_i, \eta_j) = \phi^{|j-i|} v_F$. Thus covariances are weaker for intervals which are further apart. If a small number of covariances are elicited directly, the parameters can then be adjusted until the expert is happy with the result. Note that using a stationary process in this way implies that all of the variances of $\eta_1, \ldots, \eta_p$ are equal and this is likely to require a process of iterative adjustment of the assessed values of a_i and b_i. It is felt, however, that such a process is likely to lead to better prior assessments overall. Note also that this example illustrates an advantage of working in terms of the transformed quantities η_j since stationarity of beliefs would be implausible when working directly in terms of θ_j.

For the example the values $v_F = 0.453$, $\phi = 0.97$ were adopted and therefore $v_\varepsilon = 0.0268$.

The conjugate updates take place using $A_i = a_i + y_i$ and $B_i = b_i + N_i - y_i$ in place of a_i and b_i in (11), (12) to calculate $\mathrm{E}(\eta_i \mid y_i)$ and $\mathrm{Var}(\eta_i \mid y_i)$ respectively. These are then used in (3), (4) to calculate the Bayes linear kinematic update for $\boldsymbol{\eta}$ at each stage: $\mathrm{E}_1(\boldsymbol{\eta}; \ y_i)$ and $\mathrm{Var}_1(\boldsymbol{\eta}; \ y_i)$. The unique commutative Bayes linear kinematic solution is then given by

$$\mathrm{Var}_{(29)}(\boldsymbol{\eta}) \ = \ \left(\sum_{i=1}^{29} \mathrm{Var}_1^{-1}(\boldsymbol{\eta}; \ y_i) - 28 \mathrm{Var}_0^{-1}(\boldsymbol{\eta}) \right)^{-1}$$

$$\mathrm{E}_{(29)}(\boldsymbol{\eta}) \ = \ \mathrm{Var}_{(29)}(\boldsymbol{\eta}) \left(\sum_{i=1}^{29} \mathrm{Var}^{-1}(\boldsymbol{\eta}; \ y_i) \mathrm{E}_1(\boldsymbol{\eta}; \ y_i) - 28 \mathrm{Var}_0^{-1}(\boldsymbol{\eta}) \mathrm{E}_0(\boldsymbol{\eta}) \right).$$

Having performed the updates, posterior parameter values are found and are given in Table 6. It is clear that there has been a significant reduction in uncertainty upon observation of the data.

Table 6: ABOUT HERE

5   Conclusions

In this paper two applications of Bayes linear kinematics have been investigated, the first being the modelling of related Poisson distributions and the second in the analysis of life table data. In both cases taking transformations which mapped parameters onto an unrestricted scale allowed for more effective Bayes linear kinematic updates to be made by working on a scale in which linear fitting is more appropriate. Further, they allowed general comments to be made about when a unique commutative Bayes linear kinematic solution exists.

In the life table model a complementary log-log transformation was used as this allowed for a fairly straightforward calculation of the reliability function. Of course with the binomial distribution several transformations are possible.

The commutativity provided by the Bayes linear kinematic updates could be useful in situations where commutative solutions to problems do not currently exist. For example, in time series, they may provide an alternative to the non-commutative updates in the dynamic generalised linear models of [11] and, in survival analysis, to the dynamic Bayesian models of [28]. The authors intend to describe an application of the method to non-grouped lifetime data using an approach related to that of [28] in a future paper.

The use of Bayes linear kinematics avoids the necessity of using intensive numerical methods for posterior evaluations. As a result Bayes linear kinematics could have a role to play in the field of Bayesian experimental design and the authors intend to investigate this. In problems relating to the design of experiments the goal often is to maximise a function of utilities. In order to do this integrations have to be performed many times for different values of the design parameters. If each of these integrations requires Markov chain Monte Carlo methods then the analysis can quickly become computationally infeasible. A Bayes linear kinematic analysis provides an alternative approach in which posterior moments would be straightforwardly calculable even for complicated problems.

Apart from these computational advantages, of course, the Bayes linear kinematic approach makes careful assessment of genuine beliefs about relationships between quantities a practical proposition without the imposition of artificial distributional assumptions. Additional assumptions or approximations are required to interpret the results in terms of observable quantities or their untransformed moments but these are comparable to approximations which are traditionally used, for example, for confidence intervals for parameters of lifetime distributions.

Acknowledgement

References

[1] M. Goldstein and S. Shaw, "Bayes linear kinematics and Bayes linear Bayes graphical models," *Biometrika*, vol. 91, pp. 425–446, 2004.

[2] P. C. Wang, "Comparisons on the analysis of Poisson data," *Quality and Reliability Engineering International*, vol. 15, pp. 379–383, 1999.

[3] V. Fa-Si-Oen and A. Pievatolo, "The analysis of ruptures in metropolitan gas pipelines," *Quality and Reliability Engineering International*, vol. 16, pp. 17–22, 2000.

[4] H. F. Martz, P. H. Kvam, and L. R. Abramson, "Empirical Bayes estimation of the reliability of nuclear-power-plant emergency diesel generators," *Technometrics*, vol. 38, pp. 11–24, 1996.

[5] P. Dellaportas and A. F. M. Smith, "Bayesian inference for generalised linear and proportional hazards models via gibbs sampling," *Applied Statistics*, vol. 42, pp. 443–459, 1993.

[6] D. G. Clayton, "Generalised linear mixed models," in *Markov Chain Monte Carlo in Practice* (W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, eds.), pp. 275–301, Chapman and Hall, 1996.

[7] P. McCullagh and J. A. Nelder, *Generalized Linear Models*. Chapman and Hall, London, 1989.

[8] A. C. Cameron and P. K. Trivedi, *Regression Analysis of Count Data*. Cambridge University Press, 1998.

[9] P. Congdon, *Bayesian Statistical Modelling*. John Wiley and Sons, Chichester, UK, 2001.

[10] P. Müller, "Simulation-based optimal design," in *Bayesian Statistics 6* (J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, eds.), pp. 459–474, Oxford University Press, 1999.

[11] M. West, P. J. Harrison, and H. S. Migon, "Dynamic generalized linear models and Bayesian forecasting," *Journal of the American Statistical Association*, vol. 80, pp. 73–83, 1985.

[12] M. Goldstein and D. Wooff, *Bayes Linear Statistics*. John Wiley and Sons, Chichester, UK, 2007.

[13] R. C. Jeffrey, *The Logic of Decision*. McGraw-Hill, New York, 1965.

[14] O. L. Davies and P. L. Goldsmith, *Statistical Methods in Research and Production*. Longman, 1972.

[15] D. J. Hand, F. Daly, A. D. Lunn, K. J. McConway, and E. Ostrowski, *A Handbook of Small Data Sets*. Chapman and Hall, London, 1994.

[16] J. B. Kadane and L. J. Wolfson, "Experiences in elicitation," *The Statistician*, vol. 47, pp. 3–19, 1998.

[17] P. H. Garthwaite and A. O'Hagan, "Quantifying expert opinion in the UK water industry: an experimental study," *The Statistician*, vol. 49, pp. 455–477, 2000.

[18] M. Farrow, "Practical building of subjective covariance structures for large complicated systems," *The Statistician*, vol. 52, pp. 553–573, 2003.

[19] P. H. Garthwaite, J. B. Kadane, and A. O'Hagan, "Statistical methods for eliciting probability distributions," *Journal of the American Statistical Association*, vol. 100, pp. 680–700, 2005.

[20] C. R. Peterson, K. J. Snapper, and A. H. Murphy, "Credible interval temperature forecasts," *Bulletin of the American Meteorological Society*, vol. 53, pp. 966–970, 1972.

[21] A. H. Murphy and R. L. Winkler, "Reliability of subjective probability forecasts of precipitation and temperature," *Applied Statistics*, vol. 26, pp. 41–47, 1977.

[22] S. Barclay and C. R. Peterson, "Two methods for assessing probability distributions," Tech. Rep. 73-1, Decisions and Designs Inc., McLean Va., USA., 1973.

[23] T. Spiropoulos, *Decision Support for Management using Bayes Linear Influence Diagrams*. PhD thesis, University of Sunderland, Chester Road, Sunderland, SR1 3SD, UK, 1995.

[24] M. Goldstein, M. Farrow, and T. Spiropoulos, "Prediction under the influence: Bayes linear influence diagrams for prediction in a large brewery," *The Statistician*, vol. 42, pp. 445–459, 1993.

[25] J. P. Klein and M. L. Moeschberger, *Survival Techniques for Censored and Truncated Data*. Springer, 1997.

[26] J. G. Ibrahim, M.-H. Chen, and D. Sinha, *Bayesian Survival Analysis*. Springer, New York, 2002.

[27] J. F. Lawless, *Statistical Models and Methods for Lifetime Data*. John Wiley and Sons, Chichester, UK, 1982.

[28] D. Gamerman, "Dynamic Bayesian models for survival data," *Applied Statistics*, vol. 40, pp. 63–79, 1991.

Tables

Compressor	North	Centre	South	Total
1	17	17	12	46
2	11	9	13	33
3	11	8	19	38
4	14	7	28	49
Total	53	41	72	166

Table 1: Piston-ring failures.

Case (h)	t'_L	t'_U	a'	b'	c_h	ρ_h
1	12	18.5	4.412	0.2722	0.356	0.704
2	11.75	18.75	3.824	0.2331	0.322	0.639
3	11.25	19.25	2.960	0.1756	0.229	0.453

Table 2: Elicitation of covariances c_h and correlations ρ_h in cases $h = 1, 2, 3$.

Location	Compressor and leg		Posterior mean		95% interval			
1	1	North	16.033	(15.541)	10.347	(9.809)	22.944	(22.569)
2	2	North	11.544	(11.537)	6.914	(6.658)	17.342	(17.733)
3	3	North	12.133	(12.007)	7.278	(7.094)	18.208	(18.191)
4	4	North	14.632	(14.349)	9.145	(8.964)	21.389	(20.979)
5	1	Centre	14.474	(13.997)	9.309	(8.365)	20.761	(21.055)
6	2	Centre	9.472	(9.347)	5.464	(4.883)	14.566	(15.236)
7	3	Centre	9.655	(9.246)	5.477	(4.918)	14.997	(14.919)
8	4	Centre	10.494	(9.390)	5.854	(5.186)	16.466	(14.821)
9	1	South	16.062	(14.961)	9.831	(9.745)	23.798	(21.278)
10	2	South	14.252	(14.072)	8.793	(8.806)	21.007	(20.553)
11	3	South	17.664	(17.082)	11.603	(11.111)	24.979	(24.315)
12	4	South	23.898	(20.927)	16.701	(14.184)	32.365	(28.961)

Table 3: $E(\theta_{ij}; \mathbf{Y} = \mathbf{y})$ and 95% symmetric credible intervals for the 12 locations. The figures in brackets refer to an analysis without using transformations.

i	Weeks	N_i	y_i	i	Weeks	N_i	y_i	i	Weeks	N_i	y_i
1	[0,2)	229	24	11	[20,22)	55	12	21	[40,42)	12	2
2	[2,4)	205	36	12	[22,24)	43	5	22	[42,44)	10	2
3	[4,6)	169	27	13	[24,26)	38	4	23	[44,46)	8	2
4	[6,8)	142	23	14	[26,28)	34	4	24	[46,50)	6	0
5	[8,10)	119	15	15	[28,30)	30	1	25	[50,52)	6	4
6	[10,12)	104	9	16	[30,32)	29	4	26	[52,56)	2	0
7	[12,14)	95	12	17	[32,34)	25	4	27	[56,58)	2	1
8	[14,16)	83	11	18	[34,36)	21	5	28	[58,76)	1	0
9	[16,18)	72	13	19	[36,38)	16	2	29	[76,78)	1	1
10	[18,20)	59	4	20	[38,40)	14	2				

Table 4: The failure times of centrifuge cloths.

i	a_i	b_i	i	a_i	b_i	i	a_i	b_i
1	2.206	23.530	11	2.207	13.901	21	2.212	7.916
2	2.206	22.356	12	2.208	13.167	22	2.213	7.459
3	2.206	21.229	13	2.208	12.463	23	2.214	7.025
4	2.206	20.152	14	2.208	11.797	24	2.215	6.613
5	2.207	19.128	15	2.209	11.160	25	2.216	6.218
6	2.207	18.153	16	2.209	10.552	26	2.217	5.843
7	2.207	17.214	17	2.210	9.943	27	2.219	5.485
8	2.207	16.329	18	2.210	9.422	28	2.220	5.143
9	2.207	15.481	19	2.211	8.894	29	2.222	4.817
10	2.207	14.671	20	2.211	8.395			

Table 5: Prior marginal parameter values for centrifuge cloths.

i	a_i	b_i	i	a_i	b_i	i	a_i	b_i
1	46.811	307.394	11	35.312	229.984	21	17.137	132.024
2	70.365	398.553	12	29.495	208.556	22	16.091	123.654
3	66.580	379.546	13	26.095	195.514	23	14.917	114.216
4	59.691	355.585	14	23.954	185.460	24	13.490	103.069
5	49.315	327.076	15	21.860	172.720	25	13.093	98.747
6	42.225	302.348	16	22.688	175.539	26	10.926	85.579
7	42.424	296.904	17	22.803	172.466	27	9.642	77.917
8	41.360	280.024	18	22.235	165.332	28	8.397	69.988
9	40.276	261.804	19	19.745	150.169	29	7.556	64.539
10	33.793	229.087	20	18.251	140.299			

Table 6: Posterior marginal parameter values for centrifuge cloths.

Captions for Figures

Figure 1: Adjusted means and 95% symmetric credible intervals for θ_{ij} at the 12 locations.

APPENDIX

Notation

c_1, c_2, c_3	Covariances in the piston-rings example.
E_0, Var_0, Cov_0	Prior expectation, variance, covariance.
E_1, Var_1, Cov_1	Posterior expectation, variance, covariance.
E_k, Var_k, Cov_k	Expectation, variance, covariance after observing $Y_k = y_k$.
$f_0()$	A prior density function.
$F_1, \ldots, F_p$	Autoregressive uncertainty factors for transformed parameters in the lifetimes example.
$g()$	Link function.
$l(\mu)$	The log density of μ.
m_i	The mode of μ_i and mean of η_i.
M_i	Mean of transformed parameter in lifetimes example.
$N_1, \ldots, N_n$	Numbers of trials for binomial distributions.
$q_i(1/2), q_i(1/3), q_i(2/3)$	Median and lower and upper tertiles.
Q, Q_1, Q_2	Unknown quantities in a Bayes linear analysis.
$R(t)$	Reliability function.
$s_1, \ldots, s_n$	Known scale factors for Poisson means.
S_0, S_1	Prior and posterior Bayes linear (second order) specifications.
$t_1, \ldots, t_{p+1}$	Time-interval boundaries in the lifetimes example. $(t_{p+1} = \infty)$.
$\boldsymbol{U} = (\boldsymbol{U}_1, \ldots, \boldsymbol{U}_p)$	Unknowns in a Bayes linear kinematic analysis.
$\boldsymbol{U}_k = (U_{k1}, \ldots, U_{kn_k})$	Unknowns in a Bayes linear kinematic analysis.
v_ε, v_F	Innovation variance, stationary variance for uncertainty factors in the lifetimes example.
V_{ii}, V_{ij}	Variance, covariance of transformed parameters in lifetimes example.
$w_i = -\log[1 - \theta_i] = \exp(\eta_i)$	Quantity used in inference about reliability in lifetimes example.
$\mathbf{X}$	A design matrix.
$\boldsymbol{y} = (y_1, \ldots, y_n)'$	Observed values.
$\boldsymbol{Y} = (Y_1, \ldots, Y_n)'$	Observable counts.
$\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)'$	Unknown quantities in a Bayes linear analysis.
$\boldsymbol{\beta} = (\beta_1, \ldots, \beta_r)'$	Unknown quantities in a Bayes linear analysis. $\boldsymbol{A} = \boldsymbol{\alpha} \cup \boldsymbol{\beta}$.
$\boldsymbol{\gamma}$	Regression coefficients in a linear model.
$\Gamma(x)$, $\psi(x)$, $\psi_1(x)$	Gamma, digamma and trigamma functions.
$\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)'$	Transformed parameters.
$\varepsilon_2, \ldots, \varepsilon_p$	Autoregression innovations for transformed parameters in the lifetimes example.
$\theta_1, \ldots, \theta_n$	Binomial or Poisson parameters.
$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)'$	Transformed parameters (binomial case).
ϕ	Autoregressive parameter for uncertainty factors in the lifetimes example.