Australian Government
Civil Aviation Safety Authority
www.casa.gov.au

PETA – Underlying Statistical Principles

Publishing information

Published by:	Civil Aviation Safety Authority
Postal address:	GPO Box 2005, Canberra ACT 2601
Office:	16 Furzer Street, Phillip ACT 2606
Telephone:	131 757 +61 2 6217 1111 (from outside Australia)
Facsimile:	+61 7 3144 7575
Email:	safetysystems@casa.gov.au
Internet:	https://www.casa.gov.au/

Trend Analysis Tool for CASA – Underlying Statistical Principles

Generalised Linear Modelling Framework

Safety incident data is count data, which is frequently modelled assuming a Poisson distribution. The Poisson distribution gives the probability of k independent events occurring in a specified time interval (and assumes that the times between events are independent). If clusters are present in the data, the assumption of independence will not be satisfied and, while relationships can still be modelled, formal statistical tests might not give an accurate assessment of statistical significance.

The consequence of assuming that aviation safety incidents have a Poisson distribution is that generalised linear models (GLMs) must be used for model fitting, rather than ordinary linear regression which assumes normality. A feature of GLMs is the use of a link function – a transformation applied to the expected value of the response variable to ensure fitting a linear model is reasonable. For the Poisson distribution we use the logarithmic link. A linear model is fitted to the logarithm of the expected (or mean) count at a given time, with the observed counts assumed to follow a Poisson distribution whose mean is equal to that expected mean.

Thus the tool uses a GLM framework with a Poisson distribution and log link.

One key consideration in the analysis of safety incidents is not just the incident itself, but the number of incidents relative to opportunity. Opportunity for incidents, or exposure, might be difficult to measure. Where it is possible to develop an exposure measure the generalised linear model framework outlined above can be extended to allow for exposure by including it as an offset. Then we require a value for exposure corresponding to each quarter at which there is a count of incidents.

Dispersion

Under the Poisson model, the variance of the observed count Y at any given time t is equal to its variance at that time. Sometimes the variance in the data is larger than this. This is called overdispersion. Occasionally the variance is smaller than expected and this is called underdispersion.

By preference, these models for count data are fitted using the Poisson distribution and tests comparing nested models (one model is a special case of the other, connected by a line in Figure 1) are based on the chi-square distribution for the change in deviance. This is a likelihood ratio test. However, if overdispersion is identified in model M1 in Figure 1, this is flagged, and the quasi-likelihood method and F-tests are used instead.

Dispersion can be estimated by the Pearson chi-squared goodness of fit statistic, the sum of $(\text{observed count} - \text{expected count})^2/\text{expected count}$, divided by the residual degrees of freedom. It is expected to be approximately one in the case of truly Poisson data. The Pearson chi-squared goodness of fit test can be used to test for overdispersion.

The Base Models

The primary model fitted by Data Analysis Australia has two terms, a linear trend term called "Trend" and a `final observation' effect which we label "LastSpecial". The final observation effect is an indicator variable which is set to one for the final observation and zero elsewhere, and will be significant if the final observation is `special' in some way and doesn't fit in with the pattern established over the proceeding time periods. Once this model is fitted, a model-selection technique (backwards elimination) is used to determine whether the model can be simplified. The final (best-fitting) model is selected from among special cases of the primary models as the simplest model that adequately describes the data. The final model is one of four possibilities:

M1. A model with both a linear trend and final observation effect;

M2. A model with a final observation effect;

M3. A model with a linear trend; or

M4. A constant model.

The relationships between these models are displayed in Figure 1. An arrow connecting a model A to a model B indicates that model B is a special case of model A – we say model B is nested within model A, and a likelihood ratio test can be performed to compare the models, or, equivalently, to test whether the term that has been dropped is statistically significant.

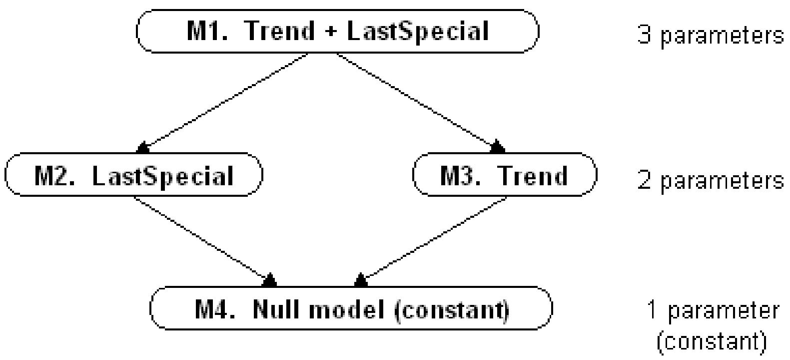

Figure 1. Lattice diagram of nested models for model selection. An arrow from a model leads to a simpler model that is a special case of the first model. Likelihood ratio tests can be used to compare two such models.

Note that Model M2 is not nested in model M3, nor vice versa, but model M4 is nested within model M1, via either M2 or M3. (A vertical arrow could be added to connect models M1 and M4, and a formal likelihood ratio test could be performed comparing models M1 and M4. However models M2 and M3 cannot be compared in this way.)

Model Extensions

In the procedure implemented by Data Analysis Australia for the CASA safety incident screening, we focus initially on selecting the best model from among models M1 to M4. We also consider some generalisations of models M1 to M4 to flag other possible changes in the aveage count, including a change in the level (M5), non-linearity (M6), and finally we consider the possible effects of annual seasonal cycles (models M7 to M10). These are displayed in Figure 2.

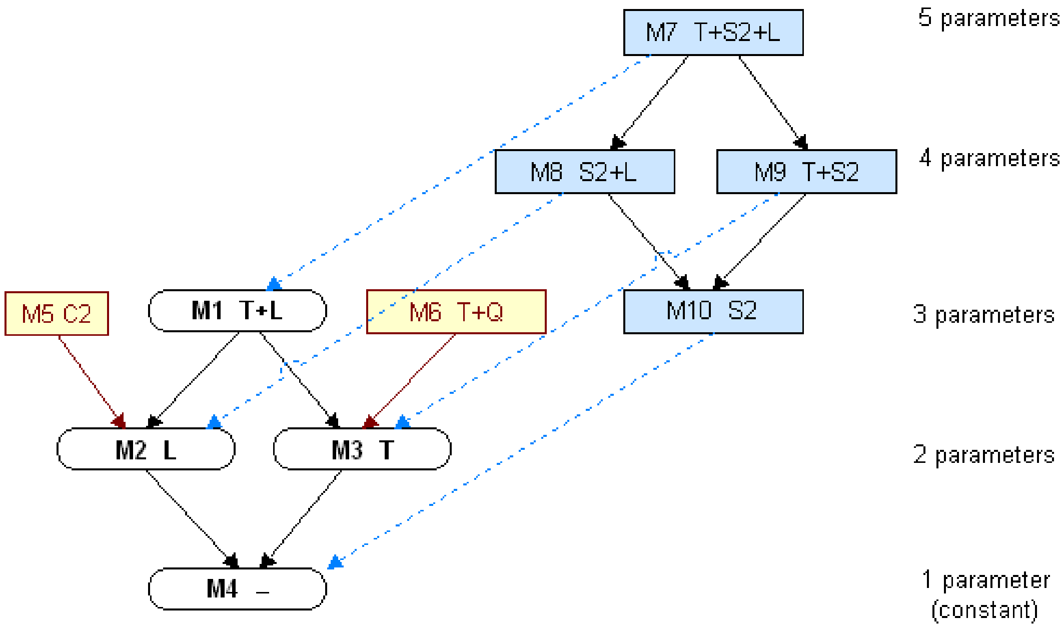

Figure 2. Figure 1 with the addition of all the other models examined. C2 indicates a change in level at an estimated Changepoint (two parameters), T indicates (linear) Trend, Q indicates Quadratic trend, S2 indicates Seasonality (two parameters), and L indicates Last observation special.

The additional models are:

- Change in level (M5) – begins at one level, and at an estimated changepoint jumps (up or down) to another level. Three parameters must be estimated. Model M2 is a special case of this, with the changepoint set between the last two observations, and thus, in the lattice diagram in Figure 2, M5 sits above M2 with an arrow leading to M2. If model M5 is found to be statistically significantly better than M2, a flag for Change in level is raised.
- Quadratic trend model (M6) – enables a crude test for evidence of nonlinearity, by adding a squared term to the linear trend model M3. Model M6 is a generalisation of M3 and sits above it in Figure 2. If model M6 is statistically significantly better than M3 (the quadratic term is significant), then a flag for Nonlinearity is raised.
- Seasonality (M7-M10) - we examine possible annual seasonal cycles in a limited set of models by adding a simple sinusoidal cyclic seasonal term (which requires two parameters, which can be regarded as amplitude and phase) to M1 to M4.

The Seasonality flag is triggered if an acceptable model with seasonality (one of M7 to M10) is significantly different from the corresponding model without seasonality (among M1 to M4) – this is a two degree of freedom test because it involves dropping two parameters simultaneously – it is meaningless to consider the amplitude or the phase in isolation – unlike the other comparisons which are all tests with one degree of freedom.

Summary of Flags

The following flags are displayed in the workbook:

- **Trend**. A linear trend is evident in the data. A sloping arrow (↗ or ↘) indicates the direction.
- **Last Special**. The last observation deviates from the pattern (linear trend or constant) for the prior observations. The direction and strength of evidence is indicated by bold (🡅, 🡇) or finer (↑, ↓) arrows, and colour-coded to highlight this most important indicator of something different happening in the last quarter.
- **Second-last Special**. When the last observation is omitted, the second-last observation deviates from the prior pattern. Indicated similarly to the Last Special flag (🡅, ↑, ↓ or 🡇). Outcome might differ from the Last Special flag in the previous run of the tool due to using slightly different data.
- **Last Quarter Changes Previous Model**. This flag will be triggered if the best-fitting model out of M1 to M4 based on the data excluding the last quarter is inconsistent with the best-fitting model including the latest quarter, for example if there was no linear trend prior to the last quarter, but when the last quarter is included a linear trend becomes evident, or vice versa. A Constant model previously and a Last Special model now is not flagged here, nor is a Trend model previously and a Trend plus Last Special model now. A brief text indication of the previous model is given.
- **Underdispersion**. Displayed as ^^^, this flag indicates that, in the model including linear Trend and Last Special, there was less variability in the data than expected from a Poisson distribution – the counts were unusually consistent. A *possible* cause of this is inadequate reporting, and therefore this might warrant checks on the adequacy of reporting. Poisson models are still fitted for all the screening tests, despite the underdispersion.
- **Overdispersion**. Displayed as /\/\/\, this flag indicates that, in the model including linear Trend and Last Special, there was greater variability in the data than expected from a Poisson distribution – there is unexplained extra variation. In the models fitted and statistical tests used, a quasi-Poisson model is used to adjust for this.

- **Excess zeros**. This flag triggers if the number of zeros in the data is unusually large for a Poisson distribution, and is displayed as 000. This could be evidence of underreporting in some of the quarters. It is tested by fitting a zero-inflated Poisson version of the model including linear Trend and Last Special and examining the statistical significance of the term for zero inflation. Zero-inflation may be one explanation for overdispersion, but we cannot say if it is the sole cause.
- **Outliers**. If the data contains one or more extreme outliers, this can affect other tests. For example a single extreme outlier can increase the variation so much that trends appear not significant and therefore would not be flagged. Therefore if this flag triggers, the reader should be aware that other characteristics of the data might be obscured and the graphical plot should be examined. The number displayed is the count of such outliers.
- **Change in Level**. Generalises the Last Special model to allow a change in level at any time point. The flag triggers if the Change in Level model is statistically significantly better than the Last Special model. Displayed as a step symbol, indicating either a step up (_—) or a step down (—_).
- **Nonlinearity**. Generalises the linear Trend model by adding a quadratic term to provide a crude test for nonlinearity. The flag triggers if the quadratic term is statistically significant. The symbol displayed indicates whether the departure from linearity tends to be upwards (convex) or downwards (concave).
- **Seasonality**. Models M1 to M4 are fitted with the addition of a simple annual cycle to investigate seasonality. This flag triggers if the seasonality is statistically significant, and the symbol ~~ is displayed to indicate the presence of an annual cycle.
- **Overdispersion Removed**. Overdispersion indicates extra variation from unexplained sources. Sometimes that variation can be explained by including other terms such as a change in level, quadratic trend or seasonality, and the apparent overdispersion disappears in these models. If this occurs it is flagged by a code to indicate which model or models reduced the overdispersion: C for Change in Level, N for Nonlinearity, and S for Seasonality.