

A Statistical Analysis of Speed Camera Data

Lee Fawcett

Newcastle University

HSS Seminar: Wednesday 24th March 2021

- **1996:** Government report: road safety cameras effective weapon in reducing casualty figures
- High implementation/running costs
- **1998:** Government allowed traffic authorities to recover these costs via speeding fines
- **2000** paper: Speed cameras an important part of the government's 2010 casualty reduction targets
- **April 2000:** two year pilot programme involving eight road safety camera partnerships (SCPs)
- Results at the end of 2000 prompted an earlier-than-expected national roll-out of SCPs

NSCP

- Joined the national programme in April 2003
- 56 mobile speed camera sites
- 'before' period (April 2001–March 2003) vs 'after' period (April 2004–March 2006)

Aims:

- To investigate changes in the number/severity of casualties
- To investigate changes in cost-of-treatment estimates

Speed cameras: Friend or foe?



Speed cameras: Friend or foe?



Speed cameras: Friend or foe?



Speed cameras: Friend or foe?



Regression To *[the]* Mean (RTM)

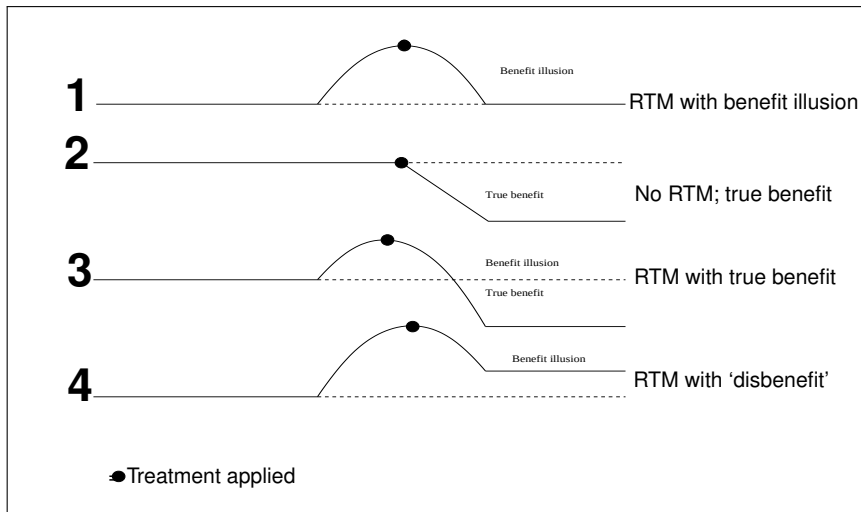
The 56 sites were chosen because of their unusually high casualty history (“hotspots”).

- The number of casualties is bound to decrease in any ‘after’ period, “...even if a garden gnome is used instead of a speed camera” (*Paul Smith, SafeSpeed*)

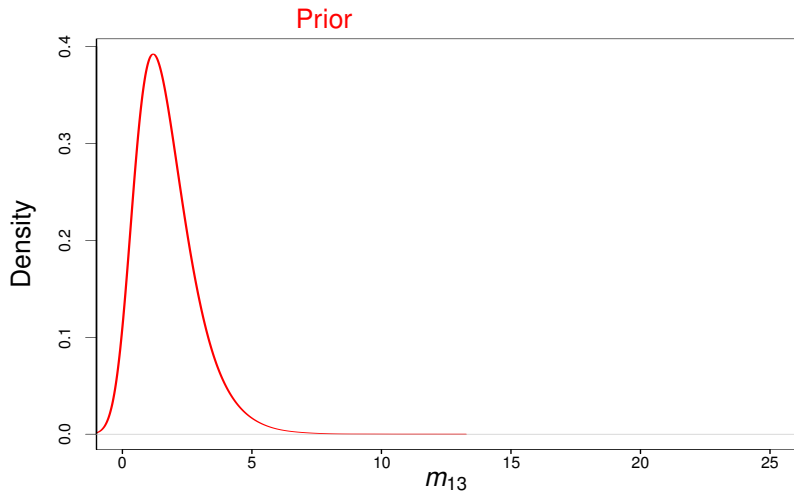


- Main consequence: ‘before’ versus ‘after’ will probably exaggerate the treatment effect
- Studies have shown that a reduction owing to RTM of between 20–30% is common

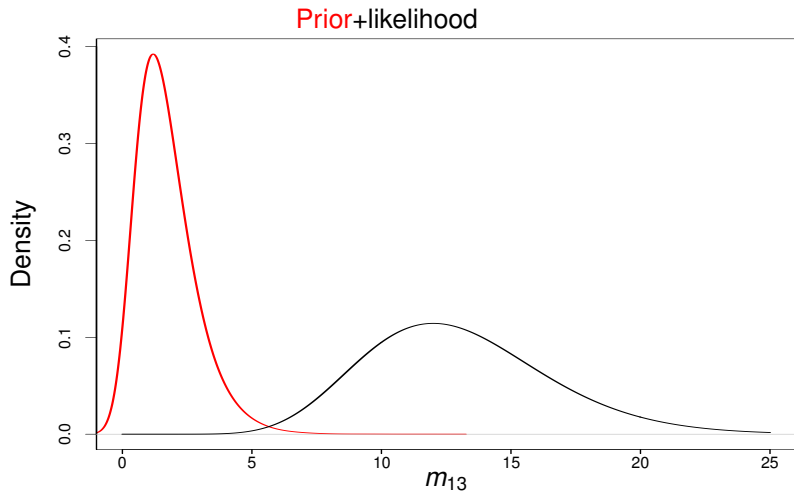
Regression To *[the]* Mean (RTM)



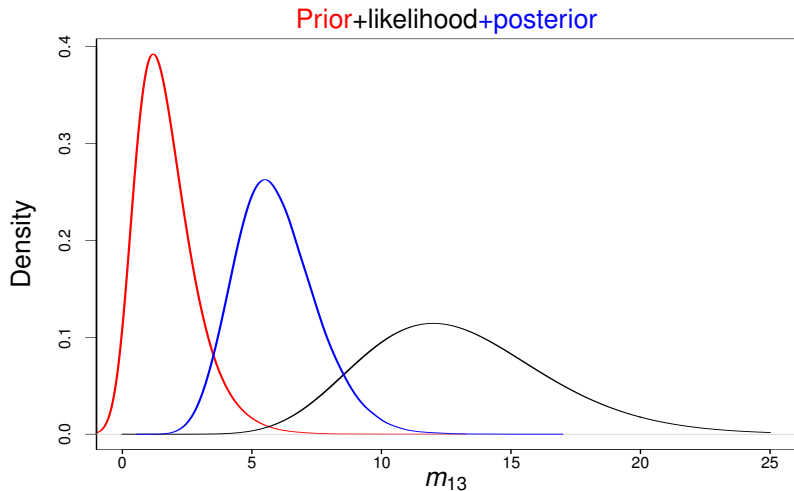
Solution: Empirical Bayes approach



Solution: Empirical Bayes approach



Solution: Empirical Bayes approach



A generalised linear modelling approach can be used to build a predictive accident model (PAM) for the prior mean μ_j , where data from non-treatment sites are used to estimate the regression coefficients:

$$\hat{\mu}_j = \exp \left\{ \hat{\beta}_0 + \sum_{p=1}^{n_p} \hat{\beta}_p x_{pj} \right\}.$$

Problem: are treatment/non-treatment sites exchangeable?

Application of Empirical Bayes

From 67 (non-speed camera) sites in Northumbria, we were given data relating to:

- x_1 : Speed limit (mph)
- x_2 : Average observed speed (mph)
- x_3 : 85th percentile speed (mph)
- x_4 : % of drivers over the speed limit
- x_5 : % of drivers at least 15mph over the speed limit
- x_6 : Daily traffic flow
- x_7 : Road classification (A, B, C, U)
- x_8 : Road type (single/dual/mixed)

Standard regression techniques were used to obtain the PAM.

Application of Empirical Bayes

- x_2 : Average observed speed (mph)
- x_4 : % of drivers over the speed limit
- x_6 : Daily traffic flow
- x_7 : Road classification (A, B, C, U)

This gives:

$$\hat{\mu}_j = \exp \{ 1.93 - 0.04x_{2j} - 0.01x_{4j} + 0.44x_{6j} + 0.67l_{1j} + 0.85l_{2j} + 1.06l_{3j} \}$$

EB estimates of casualty frequency

This PAM was then used to estimate μ_j , $j = 1, \dots, 56$, for each of our speed camera sites...

... and hence we obtain the **EB estimate of casualty frequency** for each of these sites, giving results like:

	$y_{j,\text{before}}$	μ_j	α_j	EB	$y_{j,\text{after}}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Site 13	12	1.71	0.59	5.95	2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Site 47	16	7.84	0.24	14.06	5
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Total	436			321	298

Site 13: Observed change: -10 ; after RTM: -4

Total: Observed change: -138 ; after RTM: -23

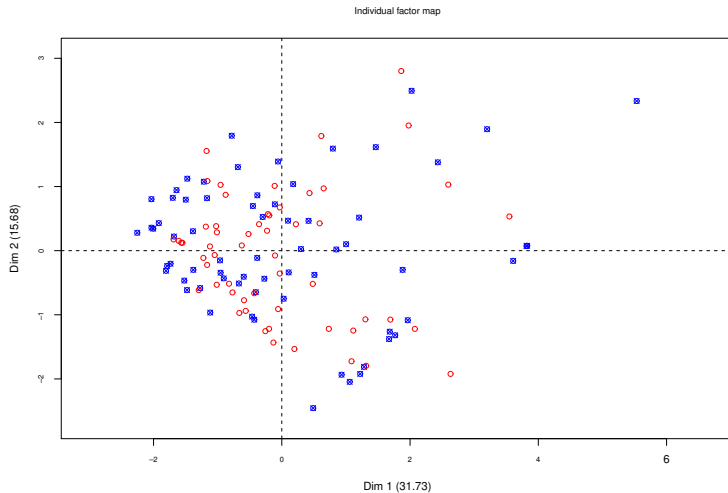
“Technique used only valid if reference sites and treatment sites can be considered exchangeable...”

- Principal Components Analysis/Multiple Factor Analysis on

		x_1	x_2	$\dots$	x_8
Reference	Site 1				
	$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$
	Site 67				
Treated	Site 1				
	$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$
	Site 56				

Plot scores for first PC against those for second PC, using different plotting character for “reference” and “treated”.

Reviewers' comments



- Permutation tests – e.g., we can compare individual variables between treatment and reference sites using:

$$\delta_p = |\bar{x}_p^{\text{TRT}} - \bar{x}_p^{\text{REF}}|, \quad p = 1, \dots, 8.$$

If the treatment and reference sites *are* exchangeable with respect to the explanatory variables, then the values of δ_p would not be significantly different to those obtained after a random allocation of sites to each group.

1. Randomly choose N permutations of “reference” and “treated” allocations
2. For each permutation Π_k , $k = 1, \dots, N$, find $\delta_p^{(\Pi_k)}$
3. Compare δ_p for the “real” allocation to the **permutation distribution** for δ_p
4. A p -value for the hypothesis H_0 : sites are exchangeable can be estimated as the proportion of permutations for which $\delta_p^{(\Pi_k)} \geq \delta_p$

Result: H_0 retained for all covariates.

Can also perform a permutation test on the **mean Mahalanobis distance** of each site in the treatment set to sites in the reference set:

$$\bar{D} = \frac{1}{56} \sum_{j=1}^{56} \sqrt{(\mathbf{X}_j^{\text{TRT}} - \bar{\mathbf{M}}^{\text{REF}})^T \Sigma^{-1} (\mathbf{X}_j^{\text{TRT}} - \bar{\mathbf{M}}^{\text{REF}})},$$

where $\bar{\mathbf{M}}^{\text{REF}} = (\bar{x}_1^{\text{REF}}, \dots, \bar{x}_1^{\text{REF}})$ and the covariance matrix Σ has (s, t) -th entry given by $\text{cov}(x_s^{\text{REF}}, x_t^{\text{REF}})$, $s, t = 1, \dots, 8$.

Result: H_0 retained!