MAS2317/3317: Introduction to Bayesian Statistics

Dr. Lee Fawcett

Case Study 1: Speed Cameras and Regression to the Mean

Semester 2, 2014—2015

Background

In 2007, official statistics revealed that 222,146 people were reported as injured as a result of road traffic accidents in Great Britain.

- 2,222 killed
- 24,690 seriously injured
- Huge economic and human cost on society

Road casualty reduction is therefore a key aim of government road safety policy with new road safety measures continually being tried and tested.

Background

- **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

Controversy

One of the first road safety camera partnerships was that linked to the Cleveland Police Force.

***Hartlepool Mail*, April 2003:**

"*Speed Cameras reduce road deaths by 70%*"

Similar messages were reported in the local/national media praising the use of speed cameras:

"*Speed cameras save lives*"

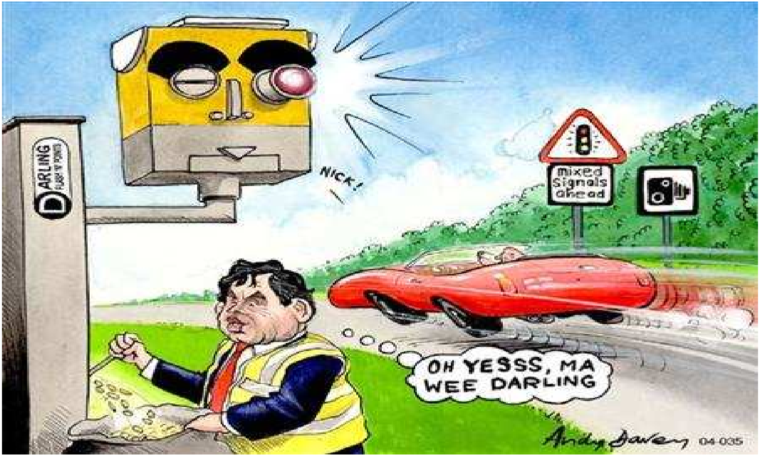
However, the rapid growth in speed camera activity, and subsequent increase in members of the public being punished for speeding offences, prompted a vigorous debate over the value of speed cameras.

Controversy

Paul Smith, founder of *SafeSpeed*: Speed cameras:

- “Are just another government tax”
- “Do not account for Regression to the Mean”
- “Are no more effective than lots of garden gnomes”

Speed cameras: Friend or foe?



Speed cameras: Friend or foe?

Speed cameras: Friend or foe?

Speed cameras: Friend or foe?

Assessment of effectiveness

The road safety camera partnerships selected sites for speed cameras based on their previous casualty record.

For example, the Cleveland Safety Camera Partnership had identified 72 "**blackspot**" locations which had observed at least 5 serious casualties over a two year observation period (2000-2002).

At each of these sites a speed camera was installed, and then the number of casualties was observed over a two year treatment period (2002-2004).

The results seemed conclusive:

- In the "before" period, there were 361 casualties...
- ...in the "after" period, there were 108 casualties...
- ...giving the 70% reduction quoted in the Hartlepool Mail

Organisations like SafeSpeed claimed such before–after studies were flawed, as they did not account for the phenomenon of **Regression To the Mean**, or RTM.

In fact, they claimed that such studies **exaggerated the effectiveness of speed cameras**.

Indeed, they had the support of leading statisticians and RTM made the headlines of leading national newspapers.

The **Northumbria Safety Camera Partnership** was formed in 2001.

In 2004 I was asked to provide statistical support in their assessment of the effectiveness of speed cameras across the northeast.

"We'd like you to produce some statistics that prove that speed cameras save lives and reduce the financial burden of road traffic accidents to the NHS"

"Teesside SCP only quoted percentage changes... we'd like some hard statistics, maybe paired t–tests and the like"

I felt uncomfortable about this.

- I was clocked by a speed camera three days after passing my driving test in 2002
  - I was doing **32mph** in a 30 zone
  - I got 3 points
  - I got a £60 fine
- Hard statistics: "*t*–tests and the like"??

Then I realised how much statisticians can charge for their services!

Regression To the Mean (RTM)

Regression To the Mean (RTM)

I came across the debate for and against speed cameras during my research.

In particular, I thought about RTM and how we could account for this.

One challenge was to explain to my colleagues in the Partnership what RTM actually was!

Hypothesis: Placing a piece of paper under a die causes it to decrease the number of sixes it rolls.

The experiment

- Take ten dice
- Roll each twenty times noting the number of sixes for each
- Place a piece of paper under the two “highest scoring” dice
- Roll each of these dice another twenty times on top of the paper and count the number of sixes observed again
- The second total is almost always lower than the first, *proving* that the piece of paper decreased the number of sixes rolled by the dice

Regression To the Mean (RTM)

I actually did this with a group of A level students at a “Maths & Stats engagement talk” at South Tyneside College recently, and obtained the following results:

Student	1	2	3	4	5	6	7	8	9	10
Sixes	3	3	1	1	0	7*	4	5*	2	2
Sixes						2		1		

- **Before:** 12
- **After:** 3
- 75% reduction!

Accounting for RTM

So, the argument is that

- because speed cameras are installed at sites with an abnormally high casualty record
- these sites would probably see a reduction in casualties in the after period anyway
- even without the cameras!

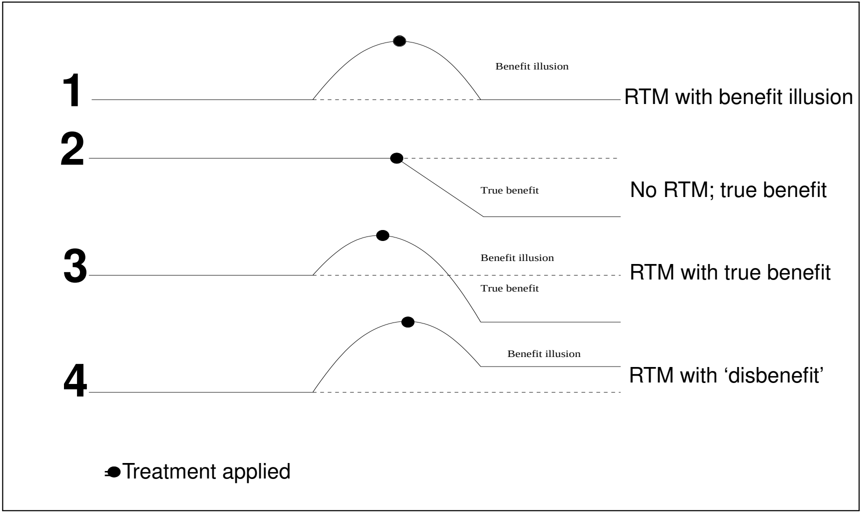
Perhaps the cameras do have a role to play, but we shouldn't really attribute the *entire* reduction to the cameras.

Some of this might have happened anyway!

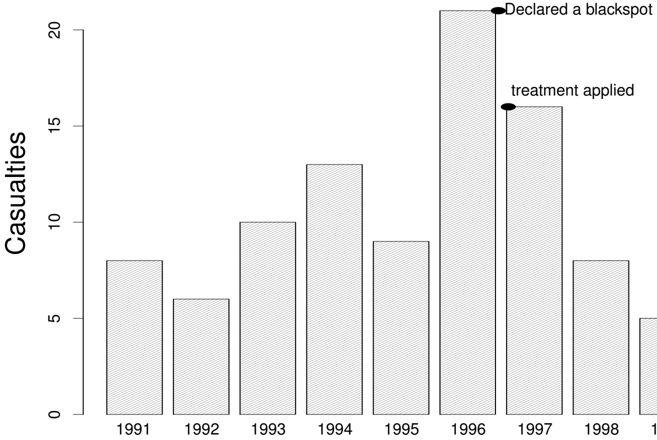
Ideal world: *randomly* choose some sites for speed cameras. Ethical concerns...

Even more ideal world: *randomly* choose some sites for speed cameras then observe what happens in a "parallel universe"... not practical.

Accounting for RTM



Accounting for RTM



The Northumbria Safety Camera Partnership

- 154 casualties at 17 blackspots in South Tyneside in a “before” period
- 41 casualties in the “after” period
- Rather than just compare before and after figures, I thought I would try to account for RTM – using a **Bayesian analysis**!

A statistical model for casualty frequency

Suppose we assume that Y_i, $i = 1, \ldots, 17$ represent casualty frequencies at each blackspot site in South Tyneside.

Assume that

$$Y_i \sim Po(\theta_i)$$

Let y_i be the (abnormally high) observed number of casualties from these Poisson distributions in the observation period.

How can we estimate the number of casualties that we would *normally* expect to see at the speed camera sites?

Formulating a prior for θ_i

A gamma prior might be reasonable; in fact, similar studies often use:

$$\theta_i \sim Ga(g, h_i).$$

- The Gamma is the **conjugate prior** for the Poisson (see Chapter 3)
- The gamma is defined over the positive real line
- Constant gamma "shape" g for all speed camera sites
- Site–specific gamma "rate" h_i
- How do we choose (g, h_i)?

Prior specification

Northumbria police recorded several variables at a total of 84 sites in South Tyneside

- Let's call the other 67 (non–camera) site our **reference sites**
- As these were *not* chosen for speed cameras, their casualty frequency was more "normal"
- A **simple linear regression** analysis gives the following for the reference sites:

$$Y_j = -4.288 + 0.157X_j + \epsilon_j,$$

where X_j and Y_j are the average observed vehicle speed, and number of casualties, at each reference site j, $j = 1, \ldots, 67$.

Prior specification

Idea: Use the regression equation obtained at the reference sites on average observed vehicle speeds at our blackspot sites to predict more "normal" casualty frequencies here!

Thus, for each speed camera site i, $i = 1, \ldots, 17$, we obtain

$$\mu_i = -4.288 + 0.157X_i.$$

Recall the gamma prior for θ_i: $\theta_i \sim Ga(g, h_i)$.

Using $h_i = g/\mu_i$ gives

- $E(\theta_i) = \mu_i$ and
- $Var(\theta_i) = \mu_i^2/g$

Discussions with a road safety expert suggested that $g = 2.5$ would be good.

Obtaining the posterior for θ_i

Recall from lectures that

$$\text{posterior} \propto \text{prior} \times \text{likelihood}$$

We have a $Ga(g, g/\mu_i)$ prior for θ_i and a Poisson likelihood.

This gives (see Assignment 2)

$$\theta_i | y_i \sim Ga\left(g + y_i, \frac{g}{\mu_i} + 1\right),$$

with **posterior mean**

$$E(\theta_i | y_i) = \alpha_i \mu_i + (1 - \alpha_i) y_i,$$

where

$$\alpha_i = \frac{g}{g + \mu_i}.$$

Assessing the effectiveness of speed cameras

Instead of comparing **before** with **after**, compare the **posterior mean** with the after figure.

- This “tones down” the abnormally high figure from the before period
- It does so using an estimate of casualty frequency we might “usually” expect to see at the blackspot sites
- But we don’t “chuck away” the abnormally high value – it *did* happen!
- A Bayesian approach provide a compromise between the two – recall the posterior mean is a weighted sum of both μ_i and y_i

Assessing the effectiveness of speed cameras

In other words, we combine

- our **prior beliefs** about what we might normally expect to see at these blackspot sites (using the reference sites)
- with our **observed** (abnormally high) value
- to form our **posterior beliefs**, a synthesis of the two!

Example: site 1

Recall that we have a Poisson distribution for the casualty frequencies at each site, and a gamma prior for the Poisson rate for the precise form of this prior), i.e.

$$\begin{aligned} Y_i|\theta_i &\sim Po(\theta_i) \\ \theta_i &\sim Ga(g, h_i). \end{aligned}$$

Recall also that this gives the gamma posterior

$$\theta_i|y_i \sim Ga\left(g + y_i, \frac{g}{\mu_i} + 1\right),$$

with mean

$$E(\theta_i|y_i) = \alpha_i\mu_i + (1 - \alpha_i)y_i,$$

where

$$\alpha_i = \frac{g}{g + \mu_i}.$$

We were told from the road safety expert that $g = 2.5$.

Example: site 1

The average speed in the before period at site 1 was $x_1 = 55$ mph, and we observed $y_1 = 20$ casualties.

Substituting this into our predictive accident model obtained from the reference set gives

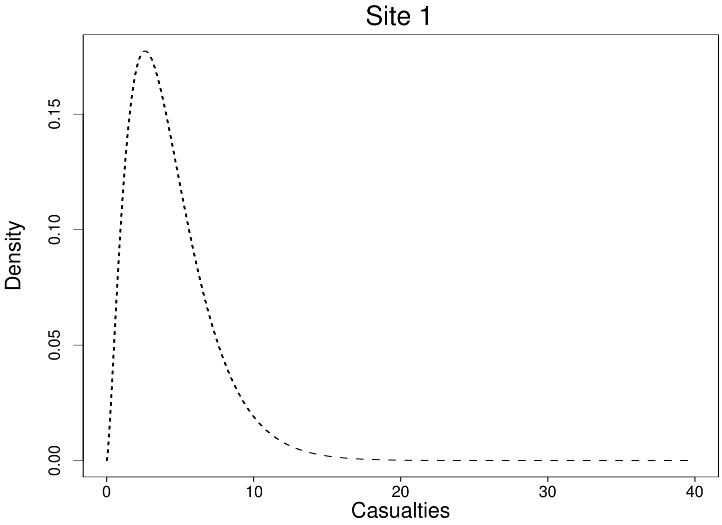
$$\mu_1 = -4.288 + 0.157 \times 55 = 4.347 \text{ casualties.}$$

Thus, we have the following prior and posteriors:

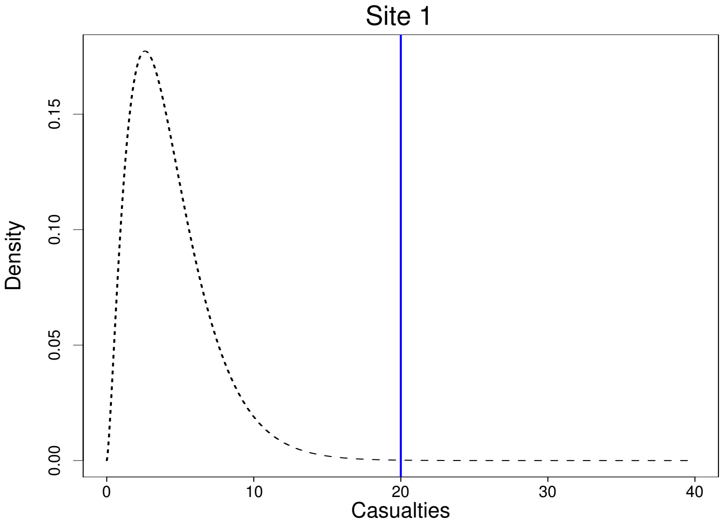
$$\begin{aligned} \theta_1 &\sim Ga\,(g = 2.5, h_1 = 2.5/4.347) \qquad \text{i.e.} \\ &\sim Ga(2.5, 0.575), \qquad \text{and} \end{aligned}$$

$$\begin{aligned} \theta_1 | y_1 = 20 &\sim Ga\left(g + y_1, \frac{g}{\mu_1} + 1\right) \qquad \text{i.e.} \\ &\sim Ga(22.5, 1.575). \end{aligned}$$

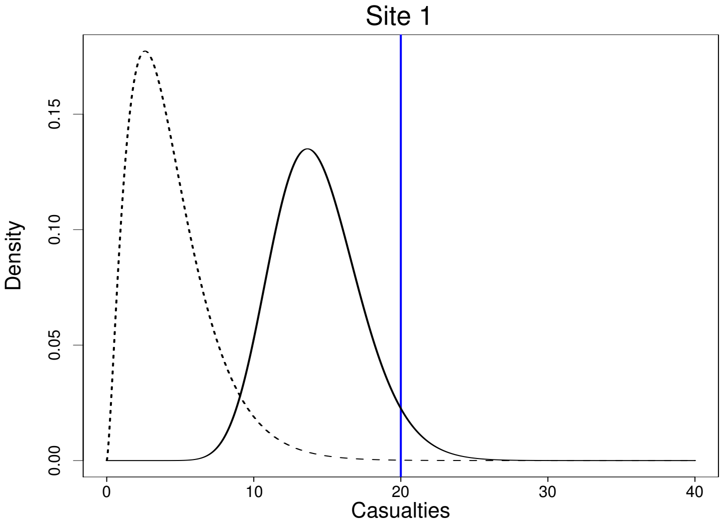
Example: site 1



Example: site 1



Example: site 1



Some results

Site	y_i	μ_i	α_i	$E(\theta_i \| y_i)$	$y_{i,\text{after}}$	Difference	
						Observed	After RTM
1	20	4.35	0.36	14.29	0	–20	–14.29
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
10	8	2.31	0.52	5.04	3	–5	–2.04
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Total	**154**			**90**	**41**	**–113**	**–49**

- Reduction from 154 $\rightarrow$ 90:"would have happened anyway"
- Remaining reduction from 90 $\rightarrow$ 41: speed cameras
- Without RTM: 73% reduction: exaggerates the effect of the cameras!
- After RTM: 54% reduction: more realistic!

Summary

- The Bayesian framework provides a means of accounting for the phenomenon of Regression To the Mean
- Without doing so will exaggerate the effectiveness on the speed cameras
- RTM has now become a recognised effect in studies such as this, and the Bayesian approach is now routinely used in official procedures when it is required to assess the effectiveness of such road safety measures (e.g. speed cameras, mini roundabouts, speed humps,...)