Confidence intervals for laboratory sonic boom annoyance tests

Jonathan Rathsam
Andrew Christian

Statistical Engineering Knowledge Exchange Workshop
Crystal City, Virginia
April 12, 2016
Acknowledgments

• Commercial Supersonic Technology Project
  – Jacob Klos, Alexandra Loubeau, Jerry Rouse, Kevin Shepherd

• Design Environment for Novel Vertical Lift Vehicles Subproject
  – Ran Cabell and Colin Theodore
Outline

1. Sonic boom research
2. Annoyance caused by sonic boom vibrations
3. Confidence interval estimation methods
   a. Delta Method
   b. Bootstrap (Parametric and Non-Parametric)
   c. Bayesian Posterior Estimation
4. Results
Supersonic Flight

-Flying above speed of sound continuously produces shock wave
-Sound of shock wave is a sonic boom

-Business travelers, cargo shippers, and traveling public
-Market potential validated in numerous studies [Henne 2005]
-New US-led aircraft manufacturing sector
Historic sonic boom highlights

- 1947  Chuck Yeager first flies supersonically
- 1964  Sonic boom tests end early due to public complaints
- 1973  Supersonic flight forbidden over land
- 2003  Shaped Sonic Boom Demonstration
- 2016  NASA announces preliminary design of supersonic X-plane

Supersonic X-plane
Motivation

• Aircraft noise regulators (FAA, ICAO) considering allowing commercial supersonic flight

• Community annoyance prediction model
  - Link predicted booms to community annoyance
  - Support new regulations
  - Support aircraft designers

[Sound Level [dB]]

[Fidell, et al. 2012]
Laboratory study

• Is there a vibration penalty?
  – increment in sound level that yields same annoyance increment as realistic vibration

• If so, how great? (high and low vibration)
Test Method

Reference contains sound *and* vibration

Which event is more annoying?
Test Method

Reference contains sound *and* vibration

Which event is more annoying?
Test Method

Reference contains sound and vibration

Which event is more annoying?

% Choosing Test Sound

Point of Subjective Equality (PSE)

Ref. Level

Test Sound Level

First
Second

0.2
0.4
0.6
0.8
1.0
Test Method

Reference contains sound *and* vibration

Which event is more annoying?
Test Method

Reference contains sound *and* vibration

Which event is more annoying?

![Graph showing the point of subjective equality (PSE) and interval estimate between sound and vibration levels.](image)
Research Question

• What is most appropriate interval estimation technique?
  a. Delta Method
  b. Bootstrap: parametric
  c. Bootstrap: non-parametric
  d. Bayesian Posterior Estimation

– Two research groups had same question
Delta Method: Theory

Logistic Regression Equation

\[
\text{Pr}(y_i = 1) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}
\]

Point of Subjective Equality (PSE)

\[
PSE = \frac{-\beta_0}{\beta_1}
\]

Taylor Series Approximation to Variance of PSE [Morgan 1992]

\[
\text{Var} (\text{PSE}) = \frac{1}{\beta_1^2} \left[ \text{Var}(\beta_0) + \text{PSE}^2 \times \text{Var}(\beta_1) + 2 \times \text{PSE} \times \text{Cov}(\beta_0, \beta_1) \right]
\]

Delta Method Confidence Interval

\[
PSE \pm z_{\left(1-\frac{\alpha}{2}\right)} \sqrt{\text{Var}(\text{PSE})}
\]
Delta Method: Application

- PSE = 82.6 dB
- 95% Conf. Interval = 81.3—83.9 dB
- Speed: 1 GLM fit
- Notes:
  - Closed form
  - Unknown failure modes
Bootstrap Analysis: Background

• Suppose we ran this test many times...

• Each subject of our test represents many similar subjects in the population

• Resample to simulate many experiments
Bootstrap Analysis: Parametric

- Use GLM to fit data from single experiment
  - $\langle \beta_0, \beta_1 \rangle$
  - $\text{Cov}(\beta_0, \beta_1)$

- Resample from multivariate distribution
Bootstrap Analysis: Non-parametric

- Create new datasets by sampling with replacement from raw data
- For each new dataset, generate a PSE
## Results: Guidance Table

<table>
<thead>
<tr>
<th>Method</th>
<th>PSE</th>
<th>PSE Interval min—max</th>
<th>Longest Operation</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delta</td>
<td>82.6</td>
<td>81.3—83.9</td>
<td>1 GLM fit (fastest)</td>
<td>• Closed form&lt;br&gt;• Unknown failure modes</td>
</tr>
<tr>
<td>Bootstrap: Parametric</td>
<td>82.6</td>
<td>81.2—83.9</td>
<td>Sorting N resampled PSEs (2nd fastest)</td>
<td>• Resamples are normally distributed&lt;br&gt;• Observable failure models (e.g. negative slope)</td>
</tr>
<tr>
<td>Bootstrap: Nonparametric</td>
<td>82.6</td>
<td>81.3—83.9</td>
<td>N GLM fits (slowest)</td>
<td>• Fewest assumptions&lt;br&gt;• Not suitable for low-n binomial data</td>
</tr>
</tbody>
</table>
Bayesian Posterior Estimation

\[ p(\beta_0, \beta_1 | \text{Data}) \propto L(\text{Data} | \beta_0, \beta_1) \times p(\beta_0, \beta_1) \]

**Posterior**  **Likelihood**  **Prior**

- Uses all data in each calculation
- Previously analytical only
- Markov Chain Monte Carlo sampling methods evaluate posterior for arbitrary likelihoods and priors
- Evaluated in R [Kruschke 2014]
## Results: Guidance Table

<table>
<thead>
<tr>
<th>Method</th>
<th>PSE</th>
<th>PSE Interval min—max</th>
<th>Longest Operation</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delta</td>
<td>82.6</td>
<td>81.3—83.9</td>
<td>1 GLM fit (fastest)</td>
<td>• Closed form</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Unknown failure modes</td>
</tr>
<tr>
<td>Bootstrap: Parametric</td>
<td>82.6</td>
<td>81.2—83.9</td>
<td>Sorting N resampled PSEs (2nd fastest)</td>
<td>• Resamples are normally distributed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Observable failure models (e.g. negative slope)</td>
</tr>
<tr>
<td>Bootstrap: Nonparametric</td>
<td>82.6</td>
<td>81.3—83.9</td>
<td>N GLM fits (slowest)</td>
<td>• Fewest assumptions</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Not suitable for low-n binomial data</td>
</tr>
<tr>
<td>Bayesian Posterior Estimation</td>
<td>82.6</td>
<td>81.4—83.9</td>
<td>N likelihood evaluations (2nd slowest)</td>
<td>• Most flexible (can include prior information)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Diagnostics needed to ensure proper MCMC performance</td>
</tr>
</tbody>
</table>
Research questions revisited (1)

• What is most appropriate interval estimation technique among four standard solutions?
  - Results from all methods are functionally equivalent
  - Delta Method used because fastest to calculate
  - BPE is recommended because it has fewest assumptions

• Return to sonic boom annoyance
Research Questions Revisited (2)

Which event is more annoying?

![Graph showing the percentage of choosing test sound vs. reference level with vibration penalty labeled.]

![Graph showing vibration penalty vs. reference sound level in dB.]

Reference Sound Level [dB]

Vibration Penalty [dB]
Research questions revisited (2)

• Is there a vibration penalty? Yes
  0 – 5 dB for low vibration and 5 – 10 dB for high vibration
Thank You

References:


Backup Slides
Bootstrap: Paramteric

The GLM-Logit model returns two parameters:

• \(<\beta_0, \beta_1>\) -- ML estimators of the logit parameters
• Cov(\(\beta\)) -- Covariance of these parameters:

Resample from resulting multivariate normal distribution
Bootstrap: Non-parametric

• Create resampled data sets by drawing from the initial raw data (with replacement).
• Run the GLM on each resampled set to produce the ML Logit fit for that set (discard the covariance).
• Use these fits to generate the resampled PSEs.
Point Clouds
Point Clouds

BPE Method

Non-Parametric Bootstrap

\[ \beta_1 \]

\[ \beta_0 \]
Are vibrations from a sonic boom annoying?

- “...sonic booms experienced inside were less acceptable than those experienced outside presumably because of ...the rattling and shaking of items within the structure, and the actual vibration of the structure itself.” [Nixon and Borsky 1966]

Kryter, et al. 1968

Rathsam, et al. 2014
Research Motivation

• Aircraft noise regulators (FAA, ICAO) considering allowing commercial supersonic flight

• Community annoyance prediction model
  - Link predicted booms to community annoyance
  - Support new regulations
  - Support aircraft designers