The Utility of Bayesian Inference in Instrumental Variables Models

Donald B. Rubin
Harvard University
Example: Randomized Experiment with Noncompliance, Sommer and Zeger Vitamin A Data

<table>
<thead>
<tr>
<th>Row</th>
<th>True Compliance Type</th>
<th>Treatment Assignment</th>
<th>Treatment Received</th>
<th>$Y_{obs}$</th>
<th>Number of Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>?</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11514</td>
</tr>
<tr>
<td>2</td>
<td>?</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>74</td>
</tr>
<tr>
<td>3</td>
<td>N</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2385</td>
</tr>
<tr>
<td>4</td>
<td>N</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>34</td>
</tr>
<tr>
<td>5</td>
<td>C</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>9663</td>
</tr>
<tr>
<td>6</td>
<td>C</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>12</td>
</tr>
</tbody>
</table>

## Results of Three Standard MoM Analyses & IVE

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimate</th>
<th>Calculation</th>
<th>Row Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITT</td>
<td>-0.0026</td>
<td>$\frac{12 + 34}{9663 + 2385 + 12 + 34} - \frac{74}{11514 + 74}$</td>
<td>3, 4, 5, &amp; 6 vs. 1 &amp; 2</td>
</tr>
<tr>
<td>As-treated</td>
<td>-0.0065</td>
<td>$\frac{12}{9663 + 12} - \frac{34 + 74}{11514 + 2385 + 34 + 74}$</td>
<td>5 &amp; 6 vs. 1, 2, 3, &amp; 4</td>
</tr>
<tr>
<td>Per protocol</td>
<td>-0.0052</td>
<td>$\frac{12}{9663 + 12} - \frac{74}{11514 + 74}$</td>
<td>5 &amp; 6 vs. 1 &amp; 2</td>
</tr>
<tr>
<td>Instrumental Variable</td>
<td>-0.0031</td>
<td>$\frac{\text{ITT}}{\text{Proportion(Compliers)}}$</td>
<td>Requires Exclusion Restriction</td>
</tr>
</tbody>
</table>

IVE: MoM CACE Analysis

\[ ACE = p_N \cdot NACE + p_C \cdot CACE \]

\[-0.0025 = 0.2 \cdot NACE + 0.8 \cdot CACE \]

\[-0.0025 = 0.8 \cdot CACE \implies CACE = -0.0025/0.8 = -0.0031 \]
Bayesian Analysis of Sommer & Zeger Data
Posterior Distribution of CACE

This Bayesian analysis considers true compliance type to be missing data for people assigned control, and multiply imputes them using observed outcomes.
Bayesian Analysis of Sommer & Zeger Data, Marginal Posterior Distributions with and without Exclusion Restriction

<table>
<thead>
<tr>
<th>Estimand</th>
<th>Exclusion restriction</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Median</th>
<th>5&lt;sup&gt;th&lt;/sup&gt; percentile</th>
<th>95&lt;sup&gt;th&lt;/sup&gt; percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>CACE</td>
<td>No</td>
<td>3.1</td>
<td>2.5</td>
<td>3.2</td>
<td>-0.9</td>
<td>7.0</td>
</tr>
<tr>
<td>ITT&lt;sub&gt;y(n)&lt;/sub&gt;</td>
<td>No</td>
<td>0.5</td>
<td>10.1</td>
<td>0.2</td>
<td>-14.1</td>
<td>17.5</td>
</tr>
<tr>
<td>CACE</td>
<td>Yes</td>
<td>3.1</td>
<td>1.2</td>
<td>3.1</td>
<td>1.2</td>
<td>5.1</td>
</tr>
</tbody>
</table>

Bayesian Analysis of Sommer & Zeger Data
Marginal Posterior Distribution of CACE

Bayesian Analysis of Sommer & Zeger Data
Marginal Posterior Distribution of $ITT_Y^{(n)}$ Without Exclusion

Bayesian Analysis of Sommer & Zeger Data
Joint Posterior Distribution of CACE and $\text{ITT}_{Y}^{(n)}$

Hypothetical Example Illustrating Frequentist Superiority of Bayes over IVE (and MLE), for CACE under both Exclusion Restrictions and Monotonicity; $Z_i$ is treatment assignment

| $C_i$           | $P(C_i|\pi)$ | $Y_i|C_i, Z_i = 0, \pi$ | $Y_i|C_i, Z_i = 1, \pi$ |
|----------------|--------------|-------------------------|-------------------------|
| True complier  | 0.25         | $N(0.1, 0.16)$          | $N(0.9, 0.49)$          |
| Never taker    | 0.45         | $N(1.0, 0.25)$          | $N(1.0, 0.25)$          |
| Always taker   | 0.30         | $N(0.0, 0.36)$          | $N(0.0, 0.36)$          |

From *True complier* row, CACE = 0.9 – 0.1 = 0.8
Monotonicity means there are no defiers
Exclusion restrictions: no effect of $Z_i$ on $Y_i$

One Sample Illustrating Bayes, Standard IVE and Standard MLE-based Methods

Hypothetical Example Illustrating Frequentist Superiority of Bayes over IVE (MoM) and MLE, Frequentist Evaluation under Monotonicity and Exclusion Restrictions

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Mean bias</th>
<th>Median bias</th>
<th>Root mean squared error</th>
<th>Median absolute error</th>
<th>Coverage rate</th>
<th>Median width</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posterior mean</td>
<td>-0.10</td>
<td>-0.07</td>
<td>0.48</td>
<td>0.30</td>
<td>0.91</td>
<td>1.61</td>
</tr>
<tr>
<td>Posterior median</td>
<td>-0.08</td>
<td>-0.06</td>
<td>0.51</td>
<td>0.32</td>
<td>0.74</td>
<td>1.11</td>
</tr>
<tr>
<td>MLE</td>
<td>-0.14</td>
<td>-0.12</td>
<td>0.51</td>
<td>0.31</td>
<td>0.91</td>
<td>2.78</td>
</tr>
<tr>
<td>IVE</td>
<td>0.55</td>
<td>0.13</td>
<td>2.31</td>
<td>0.54</td>
<td>0.91</td>
<td>2.78</td>
</tr>
</tbody>
</table>


There exists a variety of more advanced examples of Bayesian superiority.