NOVEL STATISTICAL METHODS FOR COST-EFFECTIVENESS ANALYSIS

Thomas Mathew†
Joint work with: S.Zhai†, M. Flouri†, I.Bebu*, B.Egan**
†Department of Mathematics and Statistics, University of Maryland, Baltimore County
*Biostatistics Center, George Washington University **Department of Preventive Medicine and Biometrics, Uniformed Services University of the Health Sciences

UMBC
AN HONORS UNIVERSITY IN MARYLAND

INTRODUCTION AND MOTIVATION

MOTIVATION

Cost-effectiveness analysis is an integral part of health technology assessment, and addresses the question of whether a new treatment or other health care intervention offers good value for the money.
Cost is typically expressed in monetary terms, and the effectiveness is measured using variables such as survival time, Disability-Adjusted Life Years (DALYs), Quality-Adjusted Life Years (QALYs), et.
While evaluating treatments regarding their costs and health benefits, traditionally only population averages are compared. We wish to perform cost-effectiveness analysis, as it applies to the entire population. We propose to develop criteria that can bring out features not captured by the usual summary measures based on means; in particular, criteria involving medians, percentiles etc. This will enable us to answer questions such as: for a significant percentage of the population, is the increase in cost too large compared to the effectiveness gain? Statistical tolerance limits will be used to investigate such issues.

BACKGROUND

- Consider the scenario where two treatments are compared using two groups of patients. Let C_i^j (E_i^j) denote the cost (effect) of subject i in the j^{th} group, $i = 1, 2, \ldots, n_j$, $j = 1, 2$
- Traditional cost-effectiveness analysis is based on **Incremental Cost-Effectiveness Ratio** (ICER) and **Incremental Net Benefit** (INB)

$$\text{ICER} = \frac{\mu_C^1 - \mu_C^2}{\mu_E^1 - \mu_E^2}, \quad \text{INB} = \lambda(\mu_E^1 - \mu_E^2) - (\mu_C^1 - \mu_C^2)$$

μ_C^j (μ_E^j) : mean of cost (effect) in the j^{th} group
λ : Willingness-to-Pay amount

- ICER and INB are population parameters. The corresponding random variables are:

$$Y_{\text{ICER}} = \frac{C^1 - C^2}{E^1 - E^2}, \quad Y_{\text{INB}} = \lambda(E^1 - E^2) - (C^1 - C^2)$$

SETUP

Bivariate Normal

- Typically costs follow a log-normal distribution. However, the normal distribution is also used in practice.
- Under normality, $\begin{pmatrix} C_i^j \\ E_i^j \end{pmatrix} \sim N\left(\begin{pmatrix} C_i^j \\ E_i^j \end{pmatrix}, \Sigma^j\right)$
- When the costs are log-normal, we assume $\begin{pmatrix} \ln(C_i^j) \\ E_i^j \end{pmatrix} \sim N\left(\begin{pmatrix} C_i^j \\ E_i^j \end{pmatrix}, \Sigma^j\right)$

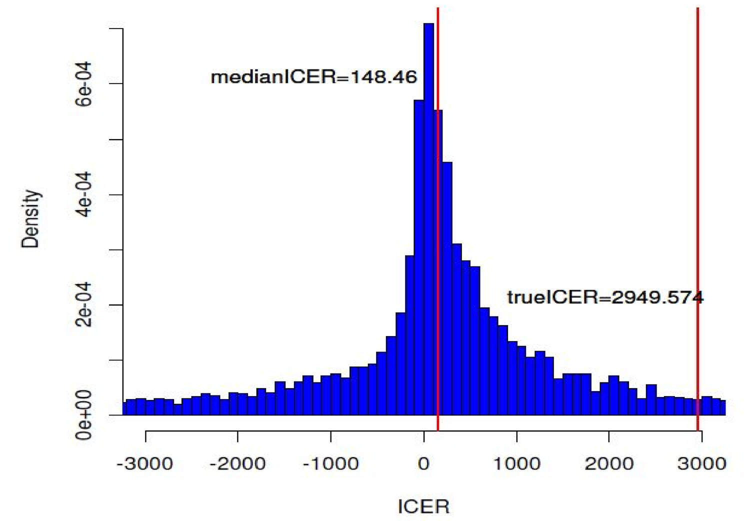
MEDIAN AND THE ICER

The median of Y_{ICER} could be drastically different from the value of ICER. In other words, the traditional ICER could be unrepresentative of the population.
In order to see this, consider the Lognormal-normal case with the following parameter values:

$$\mu^1 = \begin{pmatrix} \mu_C^1 \\ \mu_E^1 \end{pmatrix} = \begin{pmatrix} 8 \\ 10 \end{pmatrix}, \quad \Sigma^1 = \begin{pmatrix} 1.6 & -1.789 \\ -1.789 & 8 \end{pmatrix}$$

$$\mu^2 = \begin{pmatrix} \mu_C^1 \\ \mu_E^1 \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \end{pmatrix}, \quad \Sigma^2 = \begin{pmatrix} 1.2 & -1.342 \\ -1.342 & 6 \end{pmatrix}$$

The following histogram shows the distribution of Y_{ICER}, along with the median and the true value of ICER. Clearly, the ICER is unrepresentative of the distribution.



ONE-SIDED TOLERANCE LIMIT

- Tolerance limits are computed using random sample
- Lower tolerance limit, L, for Y_{INB}: $P_{\text{data}}[P(Y_{\text{INB}} \geq L|\text{data}) \geq p] = 1 - \alpha$
- Upper tolerance limit, U, for Y_{ICER}: $P_{\text{data}}[P(Y_{\text{ICER}} \leq U|\text{data}) \geq p] = 1 - \alpha$

We consider non-parametric tolerance limits based on order statistics [1]

- Let $X = (X_1, \ldots, X_N)$ be a sample from a distribution and $(X_{(1)}, \ldots, X_{(N)})$ the corresponding order statistics
- Want m such that $P_{X_{(m)}}[P_X(X \geq X_{(m)}) \geq p] = 1 - \alpha$
- $X_{(m)}$ is the intented lower tolerance
- m is defined as the largest integer such that
$$P(W \geq N - m + 1|N, 1 - p) \geq 1 - \alpha$$
where $W \sim Bin(N, 1 - p)$
- Similarly an upper tolerance limit can be defined

[1] Krishnamoorthy, K. and Mathew, T. (2009). Statistical Tolerance Regions: Theory, Applications, and Computation.Wiley, J. and Sons.

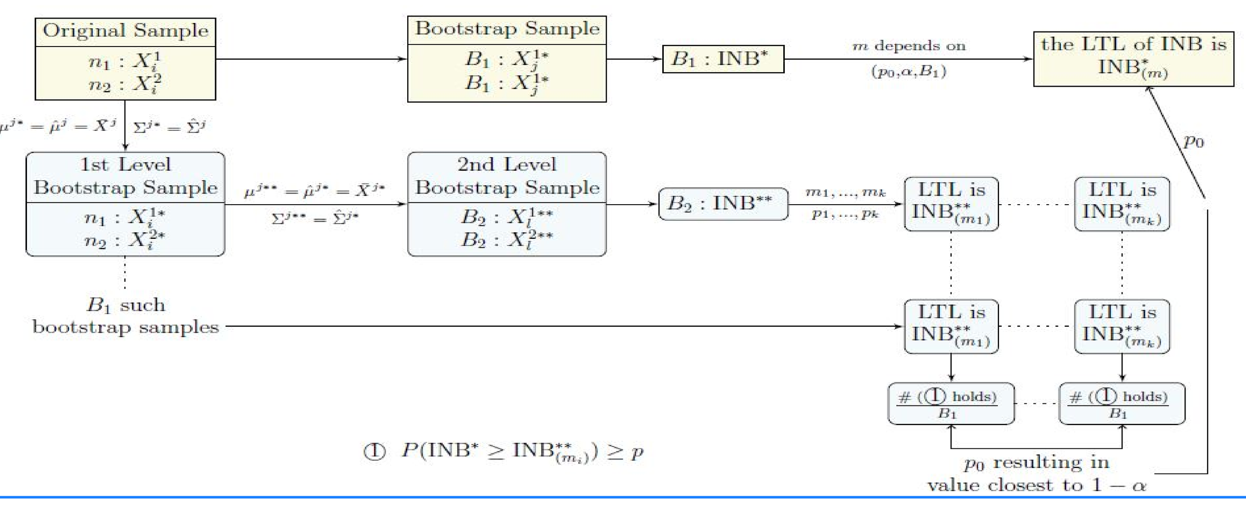
PARAMETRIC BOOTSTRAP WITH CALIBRATION

Samples are not directly available from the distributions of Y_{ICER} and Y_{INB}, so we proceed as follows:

- Generate pairs $\begin{pmatrix} C_i^{1*} \\ E_i^{1*} \end{pmatrix}$, $\begin{pmatrix} C_i^{2*} \\ E_i^{2*} \end{pmatrix}$ by parametric bootstrap, $i = 1, 2, \ldots, B_1$
- Compute $Y^*_{\text{ICER},i}$, $Y^*_{\text{INB},i}$, $i = 1, 2, \ldots, B_1$
- Use the order statistics based on $Y^*_{\text{ICER},i}$ and $Y^*_{\text{INB},i}$ to obtain tolerance limits

Coverage probabilities of the resulting tolerance limits are not satisfactory

- Use bootstrap calibration to improve the coverage
- Use B_2 second level bootstrap samples



SIMULATIONS

Bivariate Normal

$$\mu^1 = \begin{pmatrix} \mu_{11} \\ 10 \end{pmatrix}, \Sigma^1 = \begin{pmatrix} 500 & \sigma_{12}^{(1)} \\ \sigma_{21}^{(1)} & 8 \end{pmatrix}, \mu^2 = \begin{pmatrix} 485 \\ 10 \end{pmatrix}, \Sigma^2 = \begin{pmatrix} 450 & \sigma_{12}^{(2)} \\ \sigma_{21}^{(2)} & 6 \end{pmatrix}$$

where $\sigma_{12}^{(1)} = \rho\sqrt{500 \times 8}$, $\sigma_{12}^{(2)} = \rho\sqrt{450 \times 6}$, $\lambda = 25$

One-sided lower tolerance limit for INB:
$n_1 = n_2 = 50, 150$, $\mu_{11} = 500, 525$ and $\rho = -0.5, 0.1$

Lognormal-Normal

$$\mu^1 = \begin{pmatrix} \mu_{11} \\ 10 \end{pmatrix}, \Sigma^1 = \begin{pmatrix} 1.6 & \sigma_{12}^{(1)} \\ \sigma_{21}^{(1)} & 8 \end{pmatrix}, \mu^2 = \begin{pmatrix} 6 \\ 8 \end{pmatrix}, \Sigma^2 = \begin{pmatrix} 1.2 & \sigma_{12}^{(2)} \\ \sigma_{21}^{(2)} & 6 \end{pmatrix}$$

where $\sigma_{12}^{(1)} = \rho\sqrt{1.6 \times 8}$, $\sigma_{12}^{(2)} = \rho\sqrt{1.2 \times 6}$, $\lambda = 5000$

One-sided lower tolerance limit for INB:
$n_1 = n_2 = 50, 150$, $\mu_{11} = 8, 10$ and $\rho = -0.5, 0.1$

First level bootstrap sample size $B_1 = 5000$, second level bootstrap sample size $B_2 = 1000$, number of simulation runs $B_0 = 1000$.

RESULTS

Bivariate Normal with $p = 0.9$ and $1 - \alpha = 0.95$

n	μ_{11}	ρ	INB coverage	ICER coverage
50	500	-0.5	0.9585	0.9545
50	525	-0.5	0.9470	0.9630
150	500	-0.5	0.9535	0.9595
150	525	-0.5	0.9480	0.9575
50	500	0.1	0.9505	0.9505
50	525	0.1	0.9455	0.9560
150	500	0.1	0.9500	0.9595
150	525	0.1	0.9560	0.9555

Lognormal-Normal with $p = 0.9$ and $1 - \alpha = 0.95$

n	μ_{11}	ρ	INB coverage	ICER coverage
50	6	-0.5	0.9450	0.9440
50	8	-0.5	0.9430	0.9450
150	6	-0.5	0.9445	0.9505
150	8	-0.5	0.9515	0.9500
50	6	0.1	0.9525	0.9465
50	8	0.1	0.9475	0.9370
150	6	0.1	0.9485	0.9440
150	8	0.1	0.9475	0.9390

EXAMPLE

Canadian Implant Defibrilator Study (CIDS)[2]

Implant Cardioverter Defibrillator (ICD) vs. Amiodarone (Amd) in reducing the risk of death in survivors of Ventricular Tachycardia (VT).
$n_{\text{ICD}} = 212$ subjects, $n_{\text{Amd}} = 218$ subjects, $\lambda = 100000$

Bivariate Normal Data with:

$$\begin{pmatrix} \hat{\mu}_C^1 \\ \hat{\mu}_E^1 \end{pmatrix} = \begin{pmatrix} 87044 \\ 4.832 \end{pmatrix}, \Sigma^1 = \begin{pmatrix} 8462152 & 125.8 \\ 125.8 & 0.02418 \end{pmatrix}, \begin{pmatrix} \hat{\mu}_C^2 \\ \hat{\mu}_E^2 \end{pmatrix} = \begin{pmatrix} 38819 \\ 4.682 \end{pmatrix}, \Sigma^2 = \begin{pmatrix} 6497962 & 20.42 \\ 20.42 & 0.0244 \end{pmatrix}$$

Point Estimates:

$$\hat{\Delta}_c = 48225, \ \hat{\Delta}_e = 0.15, \ \widehat{\text{INB}} = \lambda\hat{\Delta}_e - \hat{\Delta}_c, \ \widehat{\text{ICER}} = \frac{\hat{\Delta}_c}{\hat{\Delta}_e}$$

100(1 − 2α)% Confidence Interval

Using Fieller's theorem

$$\frac{\widehat{\text{ICER}}\left(1 - z_{1-\alpha}^2 c \pm z_{1-\alpha}\sqrt{a + b - 2c - z_{1-\alpha}^2(ab - c^2)}\right)}{1 - z_{1-\alpha}^2 a}, \quad \widehat{\text{INB}} \pm z_{1-\alpha}\sqrt{u_\lambda}$$

where

$$a = \frac{\hat{V}(\hat{\Delta}_e)}{\hat{\Delta}_e^2}, b = \frac{\hat{V}(\hat{\Delta}_c)}{\hat{\Delta}_c^2}, c = \frac{\hat{C}(\hat{\Delta}_e, \hat{\Delta}_c)}{\hat{\Delta}_e^2, \hat{\Delta}_c)}, u_\lambda = \lambda^2\hat{V}(\hat{\Delta}_e) + \hat{V}(\hat{\Delta}_c) - 2\lambda\hat{C}(\hat{\Delta}_e, \hat{\Delta}_c)$$

- 95% Lower Confidence Limit for INB: -75785.45
- 95% Upper Confidence Limit for ICER: 83737.26

Tolerance Limits

- $p = 0.9, 1 - \alpha = 0.95$ Lower Tolerance limit for Y_{INB}: -63846.16
- $p = 0.9, 1 - \alpha = 0.95$ Upper Tolerance limit for Y_{ICER}: 898636.2
- $p = 0.03, 1 - \alpha = 0.95$ Lower Tolerance limit for Y_{INB}: 103.9
- $p = 0.2, 1 - \alpha = 0.95$ Upper Tolerance limit for Y_{ICER}: -127004.8

Conclusions

- 90% or more of individual INB values fall above: -63846.16
- 90% or more of individual ICER values fall below 898636.2
- Even though the average INB has a negative lower limit, at least 3% of the values are positive
- Even though the average ICER has a positive upper limit, at least 20% of the values are negative
- This shows that the traditional ICER analysis is inadequate for cost-effectiveness analysis

[2] Willan, A.R. and Briggs, A.H.(2006). Statistical analysis of cost-effectiveness data. Wiley:Chichester, UK.

ACKNOWLEDGMENT

The hardware used in the computational studies is part of the UMBC High Performance Computing Facility (HPCF). The facility is supported by the U.S. National Science Foundation through the MRI program (grant nos. CNS–0821258 and CNS–1228778) and the SCREMS program (grant no. DMS–0821311), with additional substantial support from the University of Maryland, Baltimore County (UMBC). See www.umbc.edu/hpcf for more information on HPCF and the projects using its resources.