Good Value for the Cost

- A common goal of an economic analysis is to identify when we can be confident that one therapy is good value compared to another.
- A threat to such confidence arises because the economic result observed in an experiment may not truly reflect the result in the population.
  - Single sample drawn from a population.
- This form of uncertainty is referred to as sampling (or stochastic) uncertainty.

Outline

- Describe methods for identifying when we can and cannot be confident about a therapy’s value for the cost
  - Point estimates
  - Confidence intervals
  - Decision threshold
- Goal is to demonstrate the quantification and interpretation of sampling certainty by use of CI for CER, CI for NMB, and acceptability curves
  - I will not focus on the technical aspects of estimation.
  - Computer code is available on website
    http://www.uphs.upenn.edu/dgimhsr/stat%20cicer.htm
Steps for Expressing Sampling Uncertainty

1. Determine point estimate in the sample
2. Compare against decision threshold
3. Quantify statistical uncertainty relative to decision threshold
   - Confidence intervals
   - P-values

Example: Expressing Sampling Uncertainty for Clinical Outcomes

- Suppose you conduct a clinical trial to test the claim that drug A reduces blood pressure more than drug B.
- You find drug A reduces systolic blood pressure by 9.4 points.

1. What is the point estimate?
2. What is the decision threshold?
3. How would we express the uncertainty?

(the standard error of the 9.4 point change was 3.18)

ANSWERS

1. What is the point estimate? 9.4
2. What is the decision threshold? 0
3. How would we express the uncertainty?

Confidence interval: 9.4 ± (t*σ)

95% confidence interval = (3.16, 15.64)

We can be confident that a therapy is clinically effective when its confidence interval excludes our decision threshold; we can’t be confident when its interval includes our decision threshold

P-value = .0035
Decision Thresholds for Clinical Outcomes

• For risk differences or changes in blood pressure or cholesterol. Decision threshold = 0
  – Risk difference = 10%; 95% CI, 0.02 to 0.18
  – Risk difference = 10%, 95% CI, -0.05 to 0.25

• For odds ratios / relative risks. Decision threshold = ???
  – OR = 0.44; 95% CI, 0.23 to 0.86
  – OR = 0.44; 95% CI, 0.20 to 1.06

Sampling Uncertainty and Clinical Outcomes

1) Not confident A differs from B (A - B)

\[ \text{Less} \quad \rightarrow \quad \text{Decision threshold} \quad \rightarrow \quad \text{Greater} \]

2) Confident A greater than B (A - B)

\[ \text{Less} \quad \rightarrow \quad \text{Decision threshold} \quad \rightarrow \quad \text{Greater} \]

3) Confident A less than B (A - B)

\[ \text{Less} \quad \rightarrow \quad \text{Decision threshold} \quad \rightarrow \quad \text{Greater} \]

Sampling Uncertainty and Economic Outcomes

• Confidence statements about economic outcomes are also based on whether or not the confidence interval for the economic outcome includes the decision threshold
Example: Expressing Sampling Uncertainty in Cost Effectiveness Analysis

• Suppose you conducted an economic evaluation of two therapies and found that:
  – Therapy A on average cost $1000 more than therapy B, SE = $325, p=0.002
  – Therapy A on average yielded 0.01 QALYs more than therapy B, SE = 0.001925, p<0.0000
  – The correlation between the difference in cost and effect was -0.71, and there were 250 participants per group in the trial
1. What is the point estimate?
   $1000 / 0.01 = $100,000 / QALY saved
2. What is the decision threshold?
3. How would we express the uncertainty?

Example: Expressing Sampling Uncertainty in Cost Effectiveness Analysis:

2. What is the decision threshold?

• The decision threshold for a cost-effectiveness ratio is the maximum willingness to pay (W) for a unit of health outcome or maximum acceptable cost-effectiveness ratio (e.g., $30,000 GBP or 50-100,000 USD per QALY)
• What makes sampling uncertainty different from common clinical examples is the challenge in agreeing on a decision threshold
  – There is no single value of W that applies to all decision makers

Example: Expressing Sampling Uncertainty in Cost Effectiveness Analysis:

3. How do we express uncertainty?

• This is no different than how we would express uncertainty in common clinical examples
  – Confidence intervals
  – P-values
• The computation of these values are not standard because there are discontinuities in the distribution of the cost-effectiveness ratio
Cost-Effectiveness Plane

- Upper Left: Treatment Dominated
- Upper Right: Cost Effectiveness Ratio
- Lower Left: Cost Effectiveness Ratio
- Lower Right: Treatment Dominates

(-) Difference in Cost (+)
(-) Difference in Effect (+)

Cost-Effectiveness Plane

- Lines through the origin that each exclude \( \alpha/2 \) of the distribution of the difference in costs and effects

CI for CER

- Upper Limit
- Lower Limit

(-) Difference in Costs (+)
(-) Difference in Effects (+)

CI for CER for Example

\[ \Delta C = 1000; \ SE_C = 325; \ \Delta Q = 0.01; \ SE_Q = 0.001925; \]
\[ \rho = -.71; \ DOF = 498 \]

\[ \text{LL: 28,200; UL: 245,200} \]
Interpreting the CI for the CER

1) Not confident A differs from B

2) Confident A cost-effective compared to B

3) Confident B cost-effective compared to A

Sampling Uncertainty for Cost Effectiveness Ratios

- In this example of confidence intervals for cost-effectiveness ratios, the interpretation is the same as in the clinical example
  - If maximum willingness to pay falls within the confidence interval, we cannot be confident that the two therapies differ in their cost-effectiveness
  - If it falls outside the interval, we can be 95% confident that one of the therapies is cost-effective compared to the other

How is Sampling Uncertainty Different For the Case of Net Monetary Benefits (NMB)?

- NMB is an alternative method for expressing the results of a cost-effectiveness analysis
- NMB is a rearrangement of the cost-effectiveness decision rule W:
  \[ \text{CER} = \frac{(\Delta C)}{(\Delta E)} < W \]
  \[ \text{NMB} = W \times (\Delta E) - (\Delta C) > 0 \]
- The decision threshold for NMB: Always 0
- Decision rule: Interventions are cost-effective if NMB > 0
NMB

- NMB is continuous and is always defined (i.e., no odd statistical properties like the ratio)
- Unlike the cost-effectiveness ratio, the standard error of net benefits is always defined
- For a WTP of 50,000, NMB for experiment 1:
  \[(50,000 \times .01) - 1000 = -500\]
- Given that not all decision making bodies have an agreed upon maximum willingness to pay, we routinely estimate net benefit over the range of policy relevant values of willingness to pay

NMB on the Cost-Effectiveness Plane

- Net benefit is defined on the cost-effectiveness plane by a family of lines, all with a slope equal to the acceptability criterion
- Each line represents a single value of net benefit, which for NMB equals -intercept
- For the line passing through the origin, NMB = 0
  - Lines below and to the right of the net benefit=0 line have positive net benefits (i.e., acceptable cost-effectiveness ratios)
  - Lines above and to the left have negative net benefits
- Lines increase in value as one travels southeasterly
Constructing CI for NMB for Example

Difference in QALYs

Difference in Cost

WTP: 50,000; NMB: -500; 95% CI: -1284 to 284

CI for NMB for Multiple WTP

Interpreting CI for NMB

1) \( \lambda = 50,000 \): Not confident A differs from B (A - B)

2) \( \lambda = 250,000 \): Confident A net beneficial compared to B

3) \( \lambda = 10,000 \): Confident B net beneficial compared to A
CI for NMB vs. CI for CER

- Primary difference between interpretation of CI for CER and CI for NMB is that we compare CER to WTP whereas we build WTP into NMB.

Acceptability Curve

- We calculate the probability a therapy is acceptable by calculating the probability that it falls below a specified value of WTP (e.g., the maximum WTP).
- The acceptability criterion is defined on the cost-effectiveness plane as a line passing through the origin with a slope equal to WTP.
- The proportion of the distribution of the difference in cost and effect that falls below and to the right of this line is "acceptable" (i.e., has positive NMB); the proportion that is above and to the left of this line is "unacceptable".

Acceptability Criterion

- Diagram showing the cost-effectiveness plane with the acceptability criterion as a line passing through the origin with a slope equal to WTP.
Constructing the Acceptability Curve

Difference in QALYs

-0.005 0.000 0.005 0.010 0.015 0.020

Difference in Cost

-500 0 500 1000 1500 2000 2500

4000 Replicates; 100 = 2.5%

3900: 245,200
3600: 179,600
2800: 127,700
1200: 76,800
100: 28,200
400: 49,100

The Acceptability Curve

Willingness to Pay

0.00 0.25 0.50 0.75 1.00

Proportion

28,200 245,200

Experiment 1

Confidence interval for CER
CER CI: (28,200 to 245,200)

Confidence Frontier for NMB
CI intersect decision threshold (0) at 28,200 to 245,200

Acceptability Curve
Acceptability curve intersects 0.025 and 0.975 at 28,200 and 245,200
Conclusions

• If our goal is to identify when we can be confident that one therapy is good value compared to another, confidence intervals for cost-effectiveness ratios, confidence intervals for NMB, and acceptability curves all yield the same conclusions.

Confidence limits for cost-effectiveness ratios provide decision makers with concise information that allows them to determine – based on their own WTP – if they can be confident about a therapy’s value for the cost. Acceptability curves provide the added advantage of allowing decision makers to assess alternate levels of confidence if such alternative levels are of interest. The choice of method should be based on the most effective way to express the point estimate and the uncertainty to decision makers rather than basing the choice on the statistical properties of each estimator.

Does this hold in all situations?

• What about when the confidence interval for the cost-effectiveness ratio is undefined?
• What about when my confidence intervals don’t make any sense?
  – Sometimes what I thought was the lower confidence interval has a value that is higher than the point estimate and what I thought was the upper confidence interval …
YES!

• Confidence statements work in all cases
  • These statements are independent of the method.
    – NMB, CER, and Acceptability Curves all give the same message
    – Adoption recommendations are always independent of the method
• We’ll walk through the two odd examples that can be categorized as:
  – Pattern #3: Confidence interval is undefined
  – Pattern #2: Strange intervals

First:
Characteristics of Pattern 1:

• The difference in effect is statistically significant
• The above example is in Pattern 1:
• We know are observing a pattern 1 finding when:
  – The confidence interval for the cost-effectiveness ratio excludes the Y axis (i.e., LL < PE < UL)
  – Both NMB confidence limits curves intersect the decision threshold (0) once
  – The acceptability curve intersects horizontal lines drawn at both 0.025 and 0.975 on the Y axis.

Pattern 1 Findings (2)

One cannot be confident the two therapies differ from one another
One can be confident the more effective therapy is not good value
One can be confident the more effective therapy is good value
ΔC = 400, SE_C = 325 (p=0.22); ΔQ = 0.02, SE_Q = 0.02 (p = 0.32); ρ = .25; DOF = 498

Characteristics of Pattern 3

- The difference in effect is not significant
- We know we are observing a pattern 3 finding when:
  - The confidence interval for the CER is undefined
  - Neither NMB confidence limit curve intersects the decision threshold (0)
  - The acceptability curve never intersects horizontal lines drawn at either 0.025 and 0.975 on the Y axis
Pattern 3 Findings (2)

One cannot be confident the two therapies differ from one another

-∞ ← Willingness to Pay → ∞

In pattern 3 example, one can estimate the CI for Our WTP

<table>
<thead>
<tr>
<th>Difference in Cost</th>
<th>Difference in QALYs</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3000</td>
<td>-0.045</td>
</tr>
<tr>
<td>-1500</td>
<td>0.000</td>
</tr>
<tr>
<td>0</td>
<td>0.045</td>
</tr>
<tr>
<td>1500</td>
<td>0.090</td>
</tr>
<tr>
<td>3000</td>
<td>0.135</td>
</tr>
</tbody>
</table>

-51% Confidence Interval

Example for Pattern 2

• Suppose you conducted a second economic evaluation of two therapies and found that:
  – Therapy A on average cost 35 more than therapy B, SE = 777.5, p = 0.96
  – Therapy A on average yielded 0.04 QALYs more than therapy B, SE = 0.0224, p = 0.07
  – The correlation between the difference in cost and effect was 0.706; and there were 250 participants per group in the trial

• Point estimate CER:
  \[ \frac{35}{0.04} = \frac{875}{QALY\ saved} \]
\[ \Delta C = 35; \ SE_{C} = 777.5; \ \Delta Q = 0.04; \ SE_{Q} = 0.0224; \]
\[ \rho = .706; \ DOF = 498 \]

---

\[ \text{CI for CER for Pattern 2 example} \]

\[ \text{CI for CER for Experiment 2} \]

**Interpreting CI for CER for Experiment 2**

- For which WTP should we be confident about adoption and for which should we not be confident?
Characteristics of Pattern 2

- the difference in effect is not significant
- We know we are observing a pattern 2 finding when:
  - The confidence interval for the CER includes the Y axis (i.e., LL > UL > PE OR PE > LL > UL)
  - One NMB confidence limit curve intersects the decision threshold (0) twice; the other limit never intersects the decision threshold
  - The acceptability curve intersects a horizontal line drawn at either 0.025 and 0.975 on the Y axis twice and never intersects the other line

Pattern 2 Findings (2)
Review of 3 Patterns

**PATTERN 1**
- One cannot be confident the more effective therapy is not good value.
- One cannot be confident the two therapies differ from one another.
- One can be confident the more effective therapy is good value.

**Willingness to Pay**

**PATTERN 2**
- One cannot be confident that one of the therapies is good value.
- One can be confident the two therapies differ from one another.

**PATTERN 3**
- One cannot be confident the two therapies differ from one another.

Summary

- No matter which pattern of experiment is observed, confidence statements derived from CI for CER, CI for NMB and acceptability curves are identical.
- Even when CI for CER appear ill-behaved (e.g., LL > UL or limits undefined), confidence statements are clear and as unequivocal as those from CI for NMB and acceptability curves.
- Acceptability curves report CI for multiple levels of confidence and thus may provide more information to decision makers than CI for CER or a single set of NMB lines.

Extra slides
Adoption Recommendations (1)

• When we don’t consider sampling uncertainty:
  – Required to calculate CER/NMB for one set of outcomes:
    • CA significantly greater than CB and QA significantly greater than QB
  – Not required to compare cost and effect if:
    • CA significantly smaller than CB and QA significantly greater than QB
    • QA not significantly different from QB
    • CA not significantly different from CB
    • Therapies fail to differ significantly in both their cost and effect

Adoption Recommendations (2)

• When we consider sampling uncertainty, the last 3 of these adoption recommendations -- which recommend a therapy without the need for evaluating cost-effectiveness -- are reversed:
  – Calculate point estimates and confidence intervals when:
    • CA significantly greater than CB and QA significantly greater than QB
    • QA not significantly different from QB
    • CA not significantly different from CB
    • Therapies fail to differ significantly in both their cost and effect

How is it Possible for the Lower Limit to be Greater than the Upper Limit?

-0.045 0.000 0.045 0.090 0.135
Difference in QALYs

-3000 -2000 -1000 0 1000 2000 3000
Difference in Cost

UL 0 to 50,000
-0 to 0 oo
-50,000 0 to +oo
0 to 50,000
-0.045 0.000 0.045 0.090 0.135
Difference in QALYs

0 to 0 oo
0 to 50,000
-50,000 0 to +oo
How is it Possible that the Point Estimate is Not Bounded by the Limits?

CI for CER for Experiment 2a

\( \Delta C = 1985; \ SE_C = 1252.1; \ \Delta Q = 0.0001; \ SE_Q = 0.0089; \ p = .976; \ DOF = 498 \)