Objectives

- To provide a platform for the symposium presentations
- To provide an orientation to -- not mastery of -- Bayes’s Theorem
- To provide a resource for further study: www.richardwollert.com
Topics (slide numbers are in parentheses)

1. How recidivism rates are calculated from “frequency data” (4-18).

2. How frequency data may be converted to probability data (19-24).

3. How recidivism rates are calculated from probability data (25-27).

4. Bayes’s Theorem calculates recidivism rates from probability data (28-30). One version uses “likelihood ratios” (LRs; 31-33).

5. Discussion: The LR principle and corollaries, Bayesian symbols and mathematical formulas, implications, and advantages (34-45).
2 x 2 contingency tables are useful for understanding how levels of factors that identify offenders as being “at risk” for sexual recidivism are related to offending (Quinsey et al., 1998, p. 50).

- Risk factors are hypothetical constructs linked to recidivism
- Risk levels are contrasting facets subsumed by a risk factor
- At least two facets are always needed to have a risk factor
  - One that correlates with recidivism (the “positive pole”)
  - One that correlates with non-recidivism (the “negative pole”)
- Risk factors are sometimes called “independent variables”
1. Calculating recidivism rates from “frequency data”

Recidivism is a documented “outcome.” So is non-recidivism.

- Outcomes are “dependent variables.”

- Recidivism is commonly defined in the following ways.
  - Arrests
  - Charges
  - Convictions
1. Calculating recidivism rates from “frequency data”

- **2 x 2 Tables**
  - 2 rows of risk levels
  - 2 columns of outcomes

- **4 Contingency Boxes**
  - All offenders belong in one
  - No one belongs in more than one

<table>
<thead>
<tr>
<th>Risk Levels</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Outcome</td>
</tr>
<tr>
<td>Risk Level</td>
<td>Box A</td>
</tr>
<tr>
<td>Risk Level</td>
<td>Box C</td>
</tr>
</tbody>
</table>
1. Calculating recidivism rates from “frequency data”

- **Risk Levels**
  - One level of a risk factor represents an *elevation* of risk
    - Designated Row “E”
  - All others represent a *mitigation* of risk
    - Designated Row “M”

<table>
<thead>
<tr>
<th>Risk Levels</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td></td>
</tr>
</tbody>
</table>
1. Calculating recidivism rates from “frequency data”

- Outcomes
  - One outcome column represents **recidivism**
    - Designated column “R”
  - The other column represents **zero documentation of recidivism**
    - Designated column “Z”

<table>
<thead>
<tr>
<th>Risk Levels</th>
<th>R</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Four types, or “classes,” of offenders are defined in the contingency boxes when each row and column are combined

1. Elevated risks who have been documented as recidivists \((E \cdot R)\)*

2. Elevated risks with zero documentation as recidivists \((E \cdot Z)\)*

3. Mitigated risks who have been documented as recidivists \((M \cdot R)\)*

4. Mitigated risks with zero documentation as recidivists \((M \cdot Z)\)*

*The middle dot represents the number of offenders who have both \(E\) and \(R\), not the number of offenders with \(E\) multiplied by the number of offenders with \(R\).
1. Calculating recidivism rates from “frequency data”

4 boxes, 4 classes of offenders

- Elevated risks who have been documented as recidivists ($E\cdot R$)
- Elevated risks with zero documentation of recidivism ($E\cdot Z$)
- Mitigated risks who have been documented as recidivists ($M\cdot R$)
- Mitigated risks with zero documentation of recidivism ($M\cdot Z$)
1. Calculating recidivism rates from “frequency data”

The boxes report how many offenders in a group under study fall in each class

- **E·R, E·Z, M·R** and **M·Z**, plotted in the table as *contingency cells*, represent numerical values

- Numerical values are indicated with *bold, italicized* font
1. Calculating recidivism rates from “frequency data”

Useful terms

- Offenders in the E·R cell are called “true positives”
- Offenders in the E·Z cell are “false positives”
- Offenders in the M·R cell are called “false negatives”
- Offenders in the M·Z cell are “true negatives”

<table>
<thead>
<tr>
<th>Risk Levels</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>E·R (True positives)</td>
</tr>
<tr>
<td>M</td>
<td>M·R (False negatives)</td>
</tr>
</tbody>
</table>
1. Calculating recidivism rates from “frequency data”

- $E$, the sum of the frequencies for $E \cdot R$ and $E \cdot Z$, is recorded in the table’s margin.

<table>
<thead>
<tr>
<th>Risk Levels</th>
<th>Outcomes</th>
<th>Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>$E \cdot R$</td>
<td>$E \cdot Z$</td>
</tr>
<tr>
<td>M</td>
<td>$M \cdot R$</td>
<td>$M \cdot Z$</td>
</tr>
</tbody>
</table>

$E = E \cdot R + E \cdot Z$
1. Calculating recidivism rates from “frequency data”

- The “marginals” for \( M, R, \) and \( Z \) are calculated and recorded in the contingency table.

<table>
<thead>
<tr>
<th>Risk Levels</th>
<th>Outcomes</th>
<th>Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E )</td>
<td>( R )</td>
<td>( E \cdot R ) ( E \cdot Z ) ( E = E \cdot R + E \cdot Z )</td>
</tr>
<tr>
<td>( M )</td>
<td>( M \cdot R ) ( M \cdot Z ) ( M = M \cdot R + M \cdot Z )</td>
<td></td>
</tr>
<tr>
<td>Margin</td>
<td>( R = E \cdot R + M \cdot R ) ( Z = E \cdot Z + M \cdot Z )</td>
<td></td>
</tr>
</tbody>
</table>
### 1. Calculating recidivism rates from “frequency data”

- The sum of all offenders in the table ($N$) may be calculated by summing the column or row marginals.

- The sum of the row marginals equals the sum of the column marginals.

<table>
<thead>
<tr>
<th>Risk Levels</th>
<th>Margin</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>R</td>
</tr>
<tr>
<td>E</td>
<td>$E \cdot R$</td>
<td>$E \cdot Z$</td>
</tr>
<tr>
<td>M</td>
<td>$M \cdot R$</td>
<td>$M \cdot Z$</td>
</tr>
<tr>
<td></td>
<td>$R = E \cdot R + M \cdot R$</td>
<td>$Z = E \cdot Z + M \cdot Z$</td>
</tr>
</tbody>
</table>

**Note**: This represents a complete “Frequency Table”
Various simple formulas may be applied to contingency tables containing frequency data to answer important questions about recidivism. Each of the questions and formulas below refer to the frequency table.

<table>
<thead>
<tr>
<th>Question</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the overall recidivism rate $O$, often called the “base rate” or the “probability of recidivism” [i.e., $P(R)$]?</td>
<td>$O = R/N$</td>
</tr>
<tr>
<td>What is the recidivism rate $X$ for offenders with an elevated risk level?</td>
<td>$X = E \cdot R/E$</td>
</tr>
<tr>
<td>What’s the value of using risk level $E$ for identifying potential recidivists? (Positive Incremental Validity, or PIV)</td>
<td>$PIV = X - O$</td>
</tr>
</tbody>
</table>
## 1. Calculating recidivism rates from “frequency data”

Example 1, Slide 1. - Frequency data for Static-99R scores (Helmus, Thornton, Hanson, & Babchishin, 2011, Appendix 1) are compiled in the following contingency table.

<table>
<thead>
<tr>
<th>Risk Levels</th>
<th>Outcomes</th>
<th>Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>E (≥6)</td>
<td>$E \cdot R = 326$</td>
<td>$E \cdot Z = 898$</td>
</tr>
<tr>
<td>M (&lt;6)</td>
<td>$M \cdot R = 527$</td>
<td>$M \cdot Z = 6,355$</td>
</tr>
<tr>
<td>Margin</td>
<td>$R = 853$</td>
<td>$Z = 7,253$</td>
</tr>
</tbody>
</table>
1. Calculating recidivism rates from “frequency data”

Example 1, Slide 2. Static-99R recidivism rates that are obtained when the formulas for O and X are applied to the frequency data in the example contingency table.

<table>
<thead>
<tr>
<th>Question</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the overall rate O of recidivism for all offenders [i.e., the “base rate” or P(R)]?</td>
<td>[ O = \frac{853}{8,106} = .105 ]</td>
</tr>
<tr>
<td>What is the recidivism rate for offenders with an elevated risk level?</td>
<td>[ X = \frac{326}{1,224} = .266 ]</td>
</tr>
<tr>
<td>What is the positive incremental validity for E?</td>
<td>[ PIV = .266 - .105 = .161 ]</td>
</tr>
</tbody>
</table>
2. Converting frequency data into probability data

It’s not necessary to estimate recidivism rates from frequency data. An alternative – “Bayesian Analysis” – first converts the center boxes to “outcome ratios” (ORs)

- The OR for recidivists in the elevated risk category is the number of offenders in the $E \cdot R$ box divided by the total number of recidivists in the $R$ box.
  - I have abbreviated this result as “$E \cdot R\%$”

- The following formulas convert the center boxes to ORs:
  - $E \cdot R\% = E \cdot R/R$
  - $M \cdot R\% = M \cdot R/R$
  - $E \cdot Z\% = E \cdot Z/Z$
  - $M \cdot Z\% = M \cdot Z/Z$
2. Converting frequency data into probability data

This is a partial Bayesian data table that shows how ORs are computed from frequency data.

<table>
<thead>
<tr>
<th>Risk Levels</th>
<th>E</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>$E \cdot R % = E \cdot R / R$</td>
<td>$M \cdot R % = M \cdot R / R$</td>
</tr>
<tr>
<td>Z</td>
<td>$E \cdot Z % = E \cdot Z / Z$</td>
<td>$M \cdot Z % = M \cdot Z / Z$</td>
</tr>
<tr>
<td>Margin</td>
<td>$R = E \cdot R + M \cdot R$</td>
<td>$Z = E \cdot Z + M \cdot Z$</td>
</tr>
</tbody>
</table>
2. Converting frequency data into probability data

One more step is needed to complete our Bayesian table: Some marginals need to be converted to proportions.

- One key conversion is calculating the recidivism rate by dividing $R$ by $N$
  - The formula for this is $R\% = R/N$.

- The other is calculating the non-recidivism rate by dividing $Z$ by $N$
  - The formula for this is $Z\% = Z/N$.

- $R\%+Z\%$ equals 100%, so $1 - R\%$ always equals $Z\%$. 
## 2. Converting frequency data into probability data

This table summarizes how frequency data may be converted to ORs and proportions.

<table>
<thead>
<tr>
<th>Risk Levels</th>
<th>Outcomes</th>
<th>Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>$E \cdot R% = \frac{E \cdot R}{R}$</td>
<td>$E \cdot Z% = \frac{E \cdot Z}{Z}$</td>
</tr>
<tr>
<td>M</td>
<td>$M \cdot R% = \frac{M \cdot R}{R}$</td>
<td>$M \cdot Z% = \frac{M \cdot Z}{Z}$</td>
</tr>
<tr>
<td>Margin</td>
<td>$R% = \frac{R}{N}$</td>
<td>$Z% = \frac{Z}{N}$</td>
</tr>
</tbody>
</table>
2. Converting frequency data into probability data

Here’s Example 1, Slide 1 again: The frequency table for Static-99R scores.

<table>
<thead>
<tr>
<th>Risk Levels</th>
<th>Outcomes</th>
<th>Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>Z</td>
</tr>
<tr>
<td>E (≥6)</td>
<td>E·R = 326</td>
<td>E·Z = 898</td>
</tr>
<tr>
<td>M (&lt;6)</td>
<td>M·R = 527</td>
<td>M·Z = 6,355</td>
</tr>
<tr>
<td></td>
<td>R = 853</td>
<td>Z = 7,253</td>
</tr>
</tbody>
</table>
2. Converting frequency data into probability data

Here’s the Bayesian data table in which frequency data have been converted to ORs and proportions:

<table>
<thead>
<tr>
<th>Risk Levels</th>
<th>Outcomes</th>
<th>Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>Z</td>
</tr>
<tr>
<td>E (≥6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M (≤6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Margin</td>
<td>R% = 853/8106 = .105</td>
<td>Z% = 7253/8106 = .895</td>
</tr>
</tbody>
</table>

\[
R = \frac{326}{853} = 0.382 \\
Z = \frac{898}{7253} = 0.124 \\
E·R% = \frac{527}{853} = 0.618 \\
E·Z% = \frac{6355}{7253} = 0.876 \\
M·R% = \frac{853}{8106} = 0.105 \\
M·Z% = \frac{7253}{8106} = 0.895 \\
\]
3. How recidivism rates may be calculated from probability data

Example 1, Slide 2 used the following frequency formula to estimate the recidivism rate for offenders with high Static-99R scores

<table>
<thead>
<tr>
<th>Question</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the recidivism rate for offenders with an elevated risk level (high test score)?</td>
<td>( X = \frac{326}{1,224} = .266 )</td>
</tr>
</tbody>
</table>
### 3. How recidivism rates may be calculated from probability data

The recidivism rate for offenders with high Static-99R scores may also be calculated from ORs and proportions. Our example uses the four underlined numbers.

<table>
<thead>
<tr>
<th>Risk Levels</th>
<th>Outcomes</th>
<th>Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E (≥6)</strong></td>
<td>R</td>
<td>Z</td>
</tr>
<tr>
<td></td>
<td>$E\cdot R% = \frac{326}{853} = 0.382$</td>
<td>$E\cdot Z% = \frac{898}{7253} = 0.124$</td>
</tr>
<tr>
<td><strong>M (&lt;6)</strong></td>
<td>M•R% = $\frac{527}{853} = 0.618$</td>
<td>M•Z% = $\frac{6,355}{7253} = 0.876$</td>
</tr>
<tr>
<td><strong>Margin</strong></td>
<td>R% = $\frac{853}{8106} = 0.105$</td>
<td>Z% = $\frac{7253}{8106} = 0.895$</td>
</tr>
</tbody>
</table>
3. How recidivism rates may be calculated from probability data

The formula for this calculation is at the top. The bottom section shows how the ORs and proportions from our Bayesian table are combined.

\[
\text{Numerator} \quad / \quad \text{Denominator} \\
(ER\% \times R\%) \quad / \quad \sum [(ER\% \times R\%) + (EZ\% \times Z\%)] = \\
\]

\[
(.382 \times .105) \quad / \quad (.382 \times .105) + (.124 \times .895) = \\
.0401 \quad / \quad .0401 + .1110 = \\
.0401 \quad / \quad .1511 = .265
\]
4. Bayes’s Theorem calculates recidivism rates from probability data

A complete expression of our formula includes the left term:

\[ P(R|E) = \frac{(E \cdot R\% \times R\%)}{\sum [(E \cdot R\% \times R\%) + (E \cdot Z\% \times Z\%) ]} \]

- This is an informal expression of “Bayes’s Theorem”

- In probability language the first term in the formula, \( P(R|E) \), is read as: “The probability of recidivism on the condition of an elevated score”

- All the formula says is “I divide the probability of getting an offender with risk level E and outcome level R by the larger probability of getting an offender with risk level E”
4. Bayes’s Theorem calculates recidivism rates from probability data

There is a second important version of Bayes’s Theorem. It also uses the same underlined ORs and proportions to calculate $P(R|E)$, but combines them in a different way.

<table>
<thead>
<tr>
<th>Risk Levels</th>
<th>Outcomes</th>
<th>Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>E (≥6)</td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>M (&lt;6)</td>
<td></td>
</tr>
<tr>
<td>R%</td>
<td>E·R% = 326/853 = .382</td>
<td>E·Z% = 898/7253 = .124</td>
</tr>
<tr>
<td>Z%</td>
<td>M·R% = 527/853 = .618</td>
<td>M·Z% = 6,355/7253 = .876</td>
</tr>
<tr>
<td>R+Z</td>
<td>R% = 853/8106 = .105</td>
<td>Z% = 7253/8106 = .895</td>
</tr>
</tbody>
</table>

$R = \frac{853}{8106} = .105$  
$Z = \frac{7253}{8106} = .895$
4. Bayes’s Theorem calculates recidivism rates from probability data

This formula is called the “odds version of Bayes’s Theorem.” The top section expresses the formula. The bottom one shows the solution.

\[
\frac{\text{Numerator}}{\text{Denominator}} = \frac{(R\% / Z\%) \times (E \cdot R\% / E \cdot Z\%)}{(\text{numerator} + 1)} = \frac{(.105 / .895) \times (.382 / .124)}{(.1173 \times 3.0806)} = \frac{.3614}{1.3614} = .265
\]
A complete informal expression of the odds formula is:

\[ P(R|E) = \frac{(R\% / Z\%) \times (E \cdot R\% / E \cdot Z\%)}{\text{numerator} + 1} \]

- This version divides one OR by another – i.e., \( E \cdot R\% \) is divided by \( E \cdot Z\% \). The result of this type of operation is called a “likelihood ratio” (LR).

- “LRs” are “evidence probabilities” that quantify a test’s ability to correctly classify offenders. For example,
  - Recidivists vs. non-recidivists
  - Pedophiles vs. non-pedophiles

- A test must have an LR more than 1 to have positive incremental validity.
4. Bayes’s Theorem calculates recidivism rates from probability data. One version uses “LRs.”

Here are the ORs, LRs and recidivism rates for Elevated (≥ 6), Intermediate (2-5), and Mitigating (0-1) Static-99R scores. I added the “Intermediate” row to our table.

| Test Scores   | ORs: Recidivism | ORs: Zero Recidivism | LRs | P(R|E_s) |
|---------------|----------------|----------------------|-----|---------|
| Elevated      | $E\cdot R\%=.38$ | $E\cdot Z\%=.12$     | 3.17 | .27     |
| Intermediate  | $I\cdot R\%=.49$ | $I\cdot Z\%=.50$     | .98 | .10     |
| Mitigating    | $M\cdot R\%=.13$ | $M\cdot Z\%=.38$     | .34 | .04     |
4. Bayes’s Theorem calculates recidivism rates from probability data. One version uses “LRs.”

The ORs in column 2 add to 1 because R was the denominator when each was derived. Column 3 adds to 1 because Z was the denominator for calculating these proportions.

| Test Scores  | ORs: Recidivism | ORs: Zero Recidivism | LRs | P(R|Eₜ) |
|--------------|-----------------|----------------------|-----|---------|
| Elevated     | \( E \cdot R\% = .38 \) | \( E \cdot Z\% = .12 \) | 3.17 | .27     |
| Intermediate | \( I \cdot R\% = .49 \) | \( I \cdot Z\% = .50 \) | .98 | .10     |
| Mitigating   | \( M \cdot R\% = .13 \) | \( M \cdot Z\% = .38 \) | .34 | .04     |
Coherent scales have high LRs for high scores, moderate LRs for moderate scores, and low LRs for low scores (see columns 4 & 5). Low LRs are needed because they identify non-recidivists.

| Test Scores  | ORs: Recidivism | ORs: Zero Recidivism | LRs | P(R|E_s) |
|--------------|-----------------|----------------------|-----|---------|
| Elevated     | E·R% = .38      | E·Z% = .12           | 3.17| .27     |
| Intermediate | I·R% = .49      | I·Z% = .50           | .98 | .10     |
| Mitigating   | M·R% = .13      | M·Z% = .38           | .34 | .04     |

Discussion: The LR principle.
The “area under the Receiver Operating Characteristic” curve is a measure of predictive accuracy (Quinsey, 1998).

- Large areas are obtained when high test scores have large LRs and low test scores have small LRs
- Truncated test scores produce truncated LRs
  - The LR for those with high Static-99 scores in a “High Risk/High Needs” group is 1.78 (Wollert, 2010).
  - It’s 2.87 for high Static-99 scores in a full sample
- This difference shows that full sampling is the best choice: It will produce larger indicia of predictive accuracy than selective sampling
5. Discussion: Bayesian symbols.

Frequency data and Bayesian data give the same result if rounding error is absent. I have used informal symbols so far. Here are some formal symbols and how they are read:

<table>
<thead>
<tr>
<th>Informal Symbol</th>
<th>Formal Symbol</th>
<th>How the Symbol is Read</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, O</td>
<td>P(R)</td>
<td>Probability of recidivism</td>
</tr>
<tr>
<td>2. 1-O</td>
<td>P(Z)</td>
<td>Probability of non-recidivism</td>
</tr>
<tr>
<td>3. X</td>
<td>P(R</td>
<td>E)</td>
</tr>
<tr>
<td>4. E·R%</td>
<td>P(E</td>
<td>R)</td>
</tr>
<tr>
<td>5. E·Z%</td>
<td>P(E</td>
<td>Z)</td>
</tr>
<tr>
<td>6. E·R% / E·Z%</td>
<td>LR_e</td>
<td>Likelihood ratio for an elevated score.</td>
</tr>
</tbody>
</table>
5. Discussion: Bayesian symbols and mathematical formulas.

Here is a formal mathematical statement of the first version of Bayes’s Theorem that was described. Equivalent formulas are in Mossman (2006, p. 48) and Donaldson & Wollert (2008, p. 212)

\[
P(R|E) = \frac{P(R) \times P(E|R)}{P(R) \times P(E|R) + P(Z) \times P(E|Z)}
\]
Here is a formal mathematical statement of the odds version of Bayes’s Theorem. It is also cited in Mossman (2006, pp. 48-49), Wollert (2007, p. 194), and Wollert et al. (2010, p. 486-487)

\[
P(R|E) = \frac{\frac{P(R)}{P(Z)} \times LR_e}{1 + \left( \frac{P(R)}{P(Z)} \times LR_e \right)}
\]
Mathematically, Bayes’s Theorem is simply a matter of deduction.

Its scientific significance lies in its philosophical, operational, and applied implications.

It’s proof is 250 years old.
5. Discussion: Implications of Bayes’s Theorem

The term on the left side – also called the posterior probability – leads Bayesians to ask different questions about their theories and data than their counterparts who have a “frequentist” orientation.

\[
P(R|E) = \frac{P(R)}{P(Z)} x LR_e \\
1 + \left( \frac{P(R)}{P(Z)} x LR_e \right)
\]
5. Discussion: Implications of Bayes’s Theorem

The Bayesian vs. the Frequentist:

- The Frequentist asks, “What’s the probability I got my data set, which may or may not support my theory, on the basis of chance?”

- The Bayesian asks, “I’ve collected some new data. What’s the posterior probability that my theory is right?”

Isn’t this the ultimate question of interest? Perhaps its relevance attracts some scientists and practitioners to Bayesian methods.
5. Discussion: Implications of Bayes’s Theorem

The formula also implies a theory about the way science advances. It says that “new posterior probabilities \( P(R|E) \) are constructed as prior probabilities \( [P(R) \text{ & } P(Z)] \) are revised in the wake of new measurements \([LR_e]\).”

\[
P(R|E) = \frac{P(R) \times LR_e}{P(Z)} \left( \frac{P(R)}{P(Z)} x LR_e \right)
\]
5. Discussion: Advantages of Bayes’s Theorem

Operationally, Bayes’s Theorem is a practical tool for updating actuarials with minimal expense.

- Suppose the base rate has changed, but the measurement probabilities – the LRs – are stable.

- New data collection can be minimized because an updated actuarial table may be compiled by using Bayesian math to combine new base rate data with stable LRs.

- The Multisample Actuarial Table of Sex Offender Recidivism Rates ("MATS-1," Wollert et al, 2010) was constructed using this approach. It includes data from 9,305 sex offenders.
5. Discussion: Advantages of Bayes’s Theorem

Bayes’s Theorem is also a powerful logical tool

- It consists of three elements:
  - Prior probability
  - Evidence probability
  - Posterior probability

- If any two elements are quantified, the third may be deduced and used as a check on the whether the other elements are realistic (Wollert, 2007, p. 177 & p. 193)
Suppose an expert opines a sex offender with a high Static-99R score and a few risk factors outside Static-99R is likely (51%) to recidivate.

- The evaluator also opines the recidivism rate for sex offenders is 12%.
- These assumptions demand a LR of 7.63 to be true, but we know the LR for high Static-99R scores is only 3.17.
- It is illogical to argue that adding a few risk factors to Static-99R would raise the LR for the entire risk item battery to 7.63.
- The moral: To offer this opinion the expert needs to show that an LR of 7.63 characterizes the risk item battery that he or she has administered.
References

References


- Wollert, R. (March 9, 2010). The Use of Probability Mathematics in Sexually Violent Predator Evaluations. Training sponsored by the California Department of Mental Health Sex Offender Commitment Program, Seaside, CA.
Acknowledgment

The presenter wishes to acknowledge Ms. Kirsten Ziegler’s assistance with designing and formatting many of the slides in this presentation. Ms. Ziegler may be contacted at k.zig@hotmail.com.