Solutions

Problem 1. a. For a positive test we have

\[ DLR_+ = \frac{P(+|\text{preg})}{P(+|\text{notpreg})} = \frac{sens}{1 - spec} \]

when \( spec = 0.52 \), we get \( DLR_+ = 1.56 \). The odds of pregnancy are 1.56 times the pre-test odds. When \( spec = 0.75 \), the odds of pregnancy are \( DLR_+ = 0.75/0.25 = 3 \), three times the pretest odds.

If the test is negative, we have

\[ DLR_- = \frac{P(-|\text{preg})}{P(-|\text{notpreg})} = \frac{1 - sens}{spec}. \]

When \( spec = 0.52 \), \( DLR_- = 0.48 \). The odds after the test are roughly half of the pre-test odds. When \( spec = 0.75 \), \( DLR_- = 0.33 \). The post-test odds are 1/3 of the pre-test odds.

b. Denote the prevalence by \( p \), then the positive predictive value

\[ PV^+ = \frac{P(+|\text{preg})P(\text{preg})}{P(+|\text{preg})P(\text{preg}) + P(+|\text{notpreg})P(\text{notpreg})} = \frac{0.75p}{0.75p + (1 - 0.635)(1 - p)}. \]

c. For the negative predictive value,

\[ PV^- = \frac{P(-|\text{notpreg})P(\text{notpreg})}{P(-|\text{notpreg})P(\text{notpreg}) + P(-|\text{preg})P(\text{preg})} = \frac{0.635(1 - p)}{0.635(1 - p) + (1 - 0.75)p}. \]

Problem 2. (a) Using the 20 cutoff for diagnosis of Alzheimer’s and the table as it appears above, we obtain the following summary of the data:

<table>
<thead>
<tr>
<th>Screen</th>
<th>Clinical Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D</td>
</tr>
<tr>
<td>+</td>
<td>12</td>
</tr>
<tr>
<td>-</td>
<td>34</td>
</tr>
<tr>
<td>Totals</td>
<td>46</td>
</tr>
</tbody>
</table>

Therefore \( sens = P(+|D) = \frac{12}{16} = 0.75 \). and \( spec = P(-|\overline{D}) = \frac{34}{46} = 0.74 \).

(b) We get

<table>
<thead>
<tr>
<th>Cut off</th>
<th>sens</th>
<th>1-spec</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.13</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0.19</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0.44</td>
<td>0.07</td>
</tr>
<tr>
<td>20</td>
<td>0.75</td>
<td>0.26</td>
</tr>
<tr>
<td>25</td>
<td>0.93</td>
<td>0.61</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
This gives the ROC curve

\[
\text{ROC curve}
\]

\[
\text{sensitivity}
\]

\[
\text{1-specificity}
\]

(c) The positive predictive value as a function of the prevalence of Alzheimer’s Disease can be expressed as:

\[
P_{V^+} = \frac{P(+ \mid D) \times P(D)}{P(+ \mid D) \times P(D) + P(+ \mid \bar{D}) \times P(\bar{D})}
\]

Using the definition of sensitivity and specificity,

\[
P_{V^+} = \frac{sens \times P(D)}{sens \times P(D) + (1 - spec) \times (1 - P(D))}
\]

With the values from the test,

\[
P_{V^+} = \frac{0.75 \times P(D)}{0.75 \times P(D) + 0.26 \times (1 - P(D))}
\]

The negative predictive value can be expressed as:

\[
P_{V^-} = \frac{P(- \mid D) \times P(D)}{P(- \mid D) \times P(D) + P(- \mid D) \times P(D)}
\]

Similarly

\[
P_{V^-} = \frac{0.74 \times (1 - P(D))}{0.74 \times (1 - P(D)) + 0.25 \times P(D)}
\]

To plot the \(P_{V^+}\) and \(P_{V^-}\) as a function of \(P(D)\), allowing the prevalence to vary from 0 to 1, we use the following code.
sens <- .75
spec <- .74
prev <- seq(0, 1, length = 1000)
ppv <- sens * prev / (sens * prev + (1 - spec) * (1 - prev))
npv <- spec * (1 - prev) / (spec * (1 - prev + (1 - sens) * prev))
plot(prev, ppv, type = "l")
plot(prev, npv, type = "l")

See the graphs below.

Problem 3. a. We have

\[ P(X = x) = \binom{10}{x} p^x (1 - p)^{10-x}. \]

For a given \( x \), the likelihood function vanishes at the boundaries, i.e. for \( p = 0 \) pr \( p = 1 \), unless \( x = 0 \) or \( x = n \). Now let us look for critical points: The log-likelihood is

\[ l(p, x) = \log(\binom{10}{x}) + x \log p + (n-x) \log(1-p) \]

Taking the derivative with respect to \( p \) and setting to zero gives

\[ \frac{dl}{dp} = \frac{x}{p} - \frac{10-x}{1-p} = 0. \]

We solve for \( p \) to obtain \( p = x/n \) and get a likelihood at \( p \) which is greater than zero.

In the cases \( x = n \) or \( x = 0 \) there is no critical point inside the interval from zero to one, so we have to explicitly compare the values. We confirm that again \( p = x/n \) gives the maximum.

b. For the plot, we use

```r
> plotPoints <- seq(0, 1, length = 100)
> n <- 10
> x <- 7
> plot(plotPoints, dbinom(x, n, plotPoints) / dbinom(x, n, x/n))
```
which gives

Since \( x = 7 \), the MLE estimate for \( p \) is \( 7/10 \). The likelihood ratio between \( p = 7/10 \) and \( p = 5/10 \) is

\[
\frac{(0.7)^7(1-0.7)^3}{(0.5)^7(1-0.5)^3} = 2.27
\]

so there is only weak relative evidence of a biased coin over a fair coin.

c. We compute the \( P \)-value

\[
P(X \geq 7) = P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10).
\]

For a fair coin, we get a probability of \( P(X \geq 7) = 0.17 \). So obtaining an “extreme” of this kind is not so infrequent.