HW 11 Solutions

DS: 2006

Question 1. Consider a cancer screening device. Supposing that $P(\textbf{Cancer}) = p$, $P(+|\textbf{Cancer}) = tp$, and $P(-|\textbf{Not Cancer}) = tn$, complete the following table with expressions of all joint, marginal, and conditional probabilities.

Work for each of the probabilities using the given information:

$P(\text{Not Cancer}) = 1 - P(\text{Cancer}) = 1 - p$ (*Complement*)

$P(-|\text{Cancer}) = 1 - P(+|\text{Cancer}) = 1 - tp$ (*Law of Total Probability*)

$P(+|\text{Not Cancer}) = 1 - P(-|\text{Not Cancer}) = 1 - tn$ (*Law of Total Probability*)

$P(+) = P(+|\text{Cancer})P(\text{Cancer}) + P(+|\text{Not Cancer})P(\text{Not Cancer}) = tp(p) + (1 - tn)(1 - p)$ (*Law of Total Probability*)

$P(-) = P(-|\text{Cancer})P(\text{Cancer}) + P(-|\text{Not Cancer})P(\text{Not Cancer}) = (1 - tp)p + tn(1 - p)$ (*Law of Total Probability*)

$P(+ \cap \text{Cancer}) = P(+|\text{Cancer})P(\text{Cancer}) = tp(p)$ (*Bayes Rule*)

$P(- \cap \text{Cancer}) = P(-|\text{Cancer})P(\text{Cancer}) = (1 - tp)p$ (*Bayes Rule*)

$P(+ \cap \text{Not Cancer}) = P(+|\text{Not Cancer})P(\text{Not Cancer}) = (1 - tn)(1 - p)$ (*Bayes Rule*)

$P(- \cap \text{Not Cancer}) = P(-|\text{Not Cancer})P(\text{Not Cancer}) = (tn)(1 - p)$ (*Bayes Rule*)

$P(\text{Cancer}|+) = \frac{P(+\cap\text{Cancer})}{P(+)} = \frac{tp(p)}{tp(p)+(1-tn)(1-p)}$

$P(\text{Cancer}|-) = \frac{P(-\cap\text{Cancer})}{P(-)} = \frac{(1-tp)p}{(1-tp)p+tn(1-p)}$

$P(\text{Not Cancer}|+) = \frac{P(+\cap\text{Not Cancer})}{P(+)} = \frac{(1-tn)(1-p)}{tp(p)+(1-tn)(1-p)}$

$P(\text{Not Cancer}|-) = \frac{P(-\cap\text{Not Cancer})}{P(-)} = \frac{(tn)(1-p)}{(1-tp)p+tn(1-p)}$

Table Showing Which Probabilities go Where

	Cancer	Not Cancer	Total
Test +	P(Cancer ∩ +)	P(Not Cancer ∩ -)	P(Test +)
Row	P(Cancer | +)	P(Not Cancer | +)	
Col	P(+ | Cancer)	P(+ | Not Cancer)	
Test -	P(Cancer ∩ -)	P(Not Cancer ∩ -)	P(Test -)
Row	P(Cancer | -)	P(Not Cancer | -)	
Col	P(- | Cancer)	P(- | Not Cancer)	
Total	P(Cancer)	P(Not Cancer)	1

Final Table

	Cancer	Not Cancer	Total
Test +	$tp(p)$	$(1-tn)(1-p)$	$tp(p)+(1-tn)(1-p)$
Row	$\frac{tp(p)}{tp(p)+(1-tn)(1-p)}$	$\frac{(1-tn)(1-p)}{tp(p)+(1-tn)(1-p)}$	
Col	tp	$1-tn$	
Test -	$(1-tp)p$	$(tn)(1-p)$	$(1-tp)p+tn(1-p)$
Row	$\frac{(1-tp)p}{(1-tp)p+tn(1-p)}$	$\frac{(tn)(1-p)}{(1-tp)p+tn(1-p)}$	
Col	$1-tp$	tn	
Total	p	$1-p$	1

Question 2: If p = 0.01, tp = 0.9, and tn = 0.8, what is the positive predictive value, $P($ Cancer$|+)$?

$$P(\text{Cancer}|+) = \frac{tp(p)}{tp(p)+(1-tn)(1-p)} = \frac{0.9(0.01)}{0.9(0.01)+(1-0.8)(1-0.01)} = \frac{1}{23} = 0.04347$$

Create a figure, with cancer incidence, $P($Cancer$)$, on the x-axis, going from 0 to 1 and the positive predictive value, $P($Cancer $|+)$, on the y-axis. Show how the positive predictive value changes as the incidence changes. Generate the figure under the assumption that the sensitivity, $P(+\|$Cancer$)$, is 0.9 and the specificity, $P(-\|$Not Cancer$)$, is 0.95.

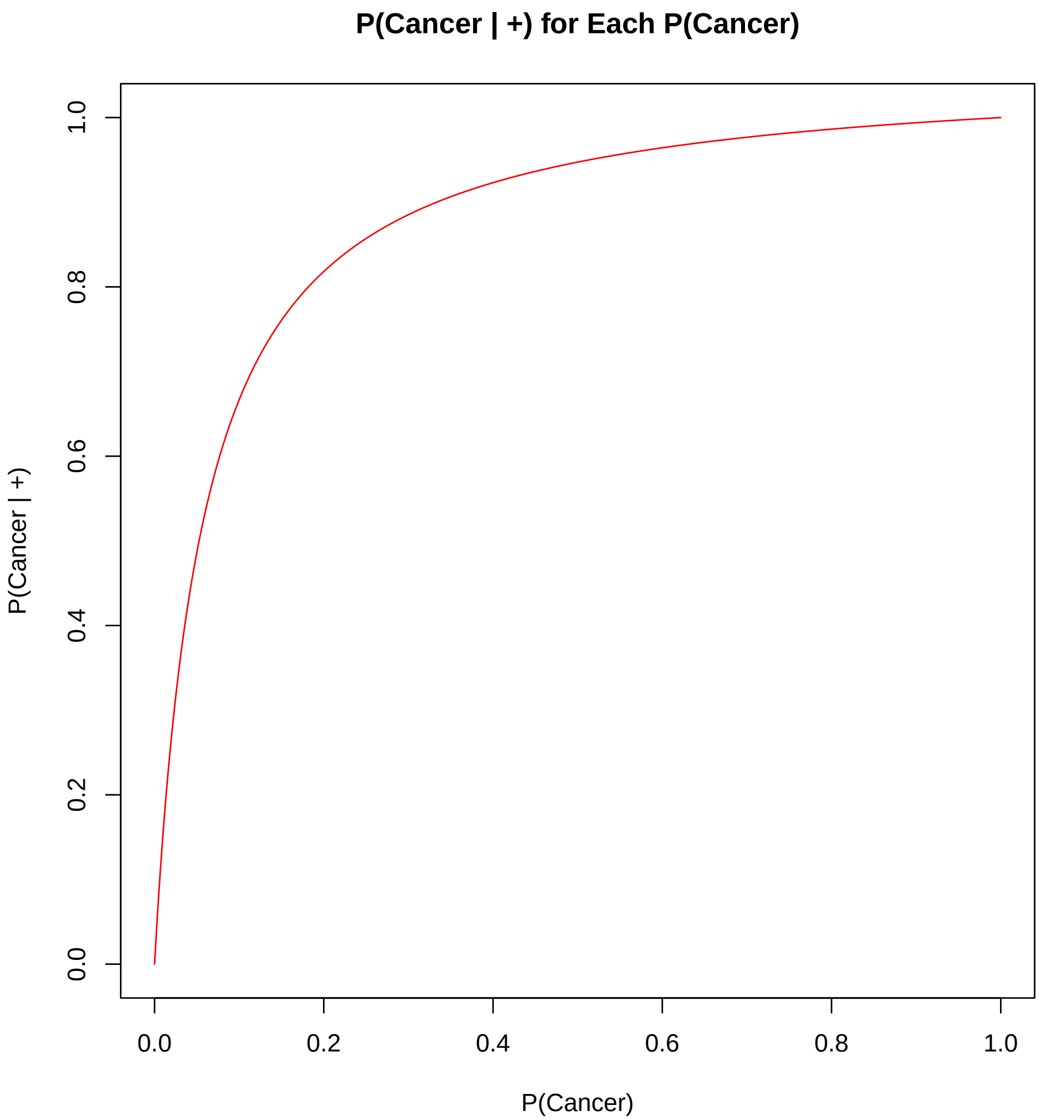
```
# assign the assumed values
tp <- 0.9
tn <- 0.95

# sequence of P(Cancer) values
p <- seq(0,1,length = 1000)
```

```r
# function to calculate the P(Cancer | +) from the table value
positive_predictive_value <- function(p){
  (tp * p) / ((tp * p) + ((1-tn) * (1- p)))
}

# find the P(Cancer | +) for each P(Cancer)
probabilities <- positive_predictive_value(p)

# generate the plot with P(Cancer) on the x-axis
# and P(Cancer | +) on the y-axis
{par(pty="s")
plot(x = p, y = probabilities,
     main = 'P(Cancer | +) for Each P(Cancer)',
     xlab = 'P(Cancer)',
     ylab = 'P(Cancer | +)',
     type = 'l',
     col = 'red')}
```



P(Cancer | +) for Each P(Cancer)
1.0
0.8
0.6
0.4
0.2
0.0
P(Cancer | +)
0.0
0.2
0.4
0.6
0.8
1.0
P(Cancer)