STAT 509 Homework 1 Solution

Chong Ma

February 7, 2017

Problem 1

Consider the event A={tomorrow rains}, then $P(A) = 0.3$. Different sample space including the event A are possible as follows.

(i) Sample space S={24 hours tomorrow}

(ii) Sampel space S={all region covered by the local television station}

(iii) Sample space S={all local meteorologists}

(iv) Sample space S={all inhabitants of the region covered by the local televesion stations}

(v) Sample space S={all days forecasted by this local television weatherman}

Problem 2

Be familiar with set algebra, especially learing the *DeMorgan's Laws*. Note that $P(A) = 0.6, P(B) = 0.5, P(A \cap B) = 0.2$. Hence,

(a) $P(\bar{A}) = 1 - P(A) = 1 - 0.6 = 0.4$

(b) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.5 + 0.6 - 0.2 = 0.9$

(c) $P(\bar{A} \cap B) = P(B) - P(A \cap B) = 0.5 - 0.2 = 0.3$

(d) $P(\bar{A} \cup \bar{B}) = P(\overline{A \cap B}) = 1 - P(A \cap B) = 1 - 0.2 = 0.8$

(e) $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.2}{0.6} = \frac{1}{3}$

(f) Not independent, since $P(B) \neq P(B|A)$.

Problem 3

Note that A={a soil sample has higher copper content}, B={the mint is present} and $P(A) = 0.3, P(B) = 0.23, P(B|A) = 0.7$.

(a) $P(\bar{B}) = 1 - P(B) = 1 - 0.23 = 0.77$

(b) $P(A \cap B) = P(B|A)P(A) = (0.7)(0.3) = 0.21$

(c) $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.21}{0.23} = 0.913$

(d) No, since $P(B|A) \neq P(B)$.

Problem 4

Denote A={successfully finding the crashed plane}. Since $P(R_1) = 0.3$, $P(R_2) = 0.7$ and $P(A|R_1) = 0.8$, $P(A|R_2) = 0.4$, then

(a) $P(A \cap R_2) = P(A|R_2)P(R_2) = (0.4)(0.7) = 0.28$

(b)

$$\begin{aligned} P(\bar{A}) &= 1 - P(A) = 1 - \{P(A|R_1)P(R_1) + P(A|R_2)P(R_2)\} \\ &= 1 - \{0.8(0.3) + 0.4(0.7)\} = 0.48 \end{aligned}$$

(c) $P(R_1|A) = \frac{P(A \cap R_1)}{P(A)} = \frac{P(A|R_1)P(R_1)}{1-P(\bar{A})} = 0.4615$

Problem 5

(a) Denote A_1 = {component 1 is functioning}, A_2 = {component 2 is functioning} and A = {the system involving two components is functioning}. Note that $P(A_1) = P(A_2) = p$. Thus,

$$\begin{aligned} P(A) &= P(A_1 \cup A_2) = 1 - P(\overline{A_1 \cup A_2}) \\ &= 1 - P(\bar{A}_1 \cap \bar{A}_2) = 1 - P(\bar{A}_1)P(\bar{A}_2) \\ &= 1 - (1-p)^2 \end{aligned}$$

(b) Similarly, denote A_i = {component is functioning}, $i = 1, 2, \ldots, n$ and A = {the system involving n components is funct
Note that $P(A_i) = p, i = 1, 2, \ldots, n$. Thus,

$$\begin{aligned} P(A) &= P(\cup_{i=1}^n A_i) = 1 - P(\overline{\cup_{i=1}^n A_i}) \\ &= 1 - P(\cap_{i=1}^n \bar{A}_i) = 1 - \prod_{i=1}^n P(\bar{A}_i) \\ &= 1 - (1-p)^n \end{aligned}$$

(c) Solve $r_4 = 1 - (1-p)^4 = 0.999$, and we have

$$p = 1 - (1 - 0.999)^{1/4} = 0.8222$$

Problem 6

Note that sensitivity is $P(\text{Pos}|\text{Present}) = 0.95$ and specificity $= P(\text{Neg}|\text{Absent}) = 0.97$, hence $P(\text{Neg}|\text{Present}) = 1 - 0.95 = 0.05$ and $P(\text{Pos}|\text{Absent}) = 1 - 0.97 = 0.03$. Denote p = the probability that an individual is strep-
Where

$$\begin{aligned} \text{Pos} &= \text{test result is positive} \\ \text{Neg} &= \text{test result is negtive} \\ \text{Present} &= \text{the streptococcus is actually present} \\ \text{Absent} &= \text{test strptococcus is actually not present} \end{aligned}$$

(a)

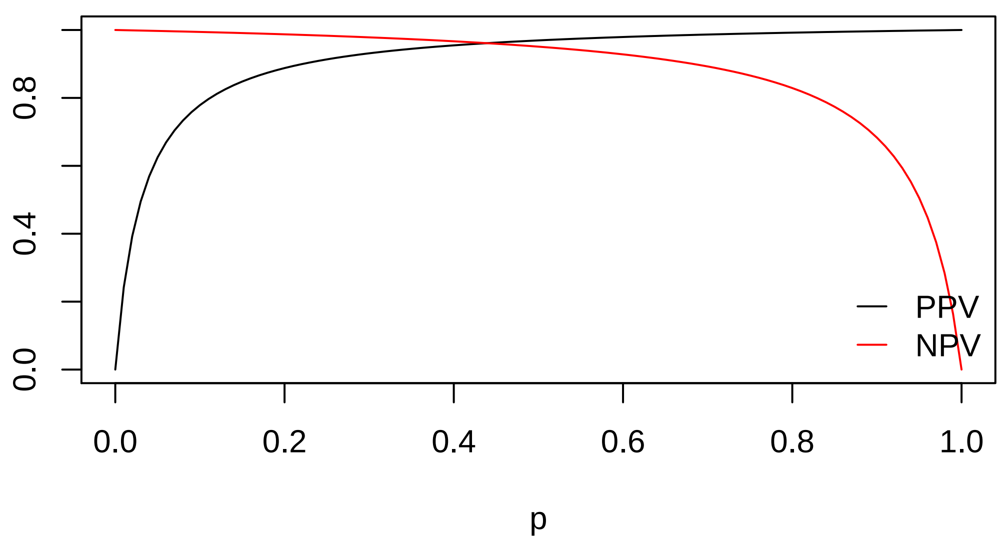
$$
\begin{aligned}
PPV(p) = P(\text{Present}|\text{Pos}) &= \frac{P(\text{Pos}|\text{Present})P(\text{Present})}{P(\text{Pos})} \\
&= \frac{P(\text{Pos}|\text{Present})P(\text{Present})}{P(\text{Pos}|\text{Present})P(\text{Present}) + P(\text{Pos}|\text{Absent})P(\text{Absent})} \\
&= \frac{0.95p}{0.95p + 0.03(1-p)}
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
NPV(p) = P(\text{Absent}|\text{Neg}) &= \frac{P(\text{Neg}|\text{Absent})P(\text{Absent})}{P(\text{Neg})} \\
&= \frac{P(\text{Neg}|\text{Absent})P(\text{Absent})}{P(\text{Neg}|\text{Absent})P(\text{Absent}) + P(\text{Neg}|\text{Present})P(\text{Present})} \\
&= \frac{0.97(1-p)}{0.97(1-p) + 0.05p}
\end{aligned}
$$

(b) Graph PPV and NPV as functions of p.

```
ppv<-function(p) 0.95*p/(0.95*p+0.03*(1-p))
npv<-function(p) 0.97*(1-p)/(0.97*(1-p)+0.05*p)
curve(ppv,0,1,col=1,xlab="p",ylab="")
curve(npv,0,1,col=2,add=TRUE)
legend("bottomright",c("PPV","NPV"),col=1:2,lty=rep(1,2),
       seg.len = 1,bty = "n")
```



(c) When $p = 0.10$, thus

$$PPV(p) = \frac{0.95p}{0.95p + 0.03(1-p)} = 0.7786885$$
$$NPV(p) = \frac{0.97(1-p)}{0.97(1-p) + 0.05p} = 0.9943052$$

If a person has a positive test, the predictive probability that the person is actually strep-positive is 0.78. If a person has a negative test, the predictive probability that the person is actually strep-negative is 0.99.

(d) Given the rapid test is positive, then the updated probability of being strep-positive is

$$PPV(0.1) = 0.78$$

Note the sensitivity and specificity for a throat-culture test are both 0.999, therefore the PPV for a throat-culture test is

$$PPV(p) = \frac{0.999p}{0.999p + 0.001(1-p)}$$

Given the throat-culture test is positive, using the updated $p = 0.78$, the new PPV is

$$PPV(0.78) = \frac{0.999(0.78)}{0.999(0.78) + 0.001(1-0.78)} = 0.9997$$

(e) Given the rapid test is negtive, then the updated probability of being strep-negtive is

$$NPV(0.1) = 0.994$$

Note the sensitivity and specificity for a throat-culture test are both 0.999, therefore the NPV for a throat-culture test is

$$NPV(p) = \frac{0.999(1-p)}{0.999(1-p) + 0.001p}$$

Given the throat-culture test is negtive, using the updated $p = 1 - 0.994 = 0.006$, the new NPV is

$$NPV(0.006) = \frac{0.999(0.994)}{0.999(0.994) + 0.001(1-0.994)} = 0.999994$$