

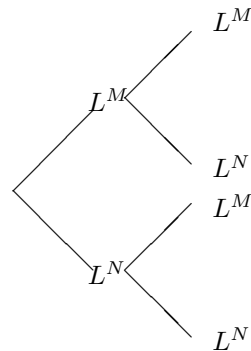
**MAT 2379, Introduction to biostatistics**

**Solution to Assignment 1**

*Due date: Monday September 26, 2011*

Total = 20 marks

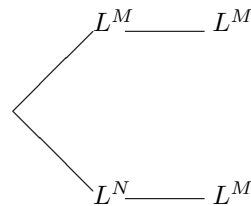
**Problem 2.1.** (2 marks) (a)



Female Gamete	Male Gamete	
	$\frac{1}{2}L^M$	$\frac{1}{2}L^N$
$\frac{1}{2}L^M$	$\frac{1}{4}L^M L^M$ (type M)	$\frac{1}{4}L^M L^N$ (type MN)
$\frac{1}{2}L^N$	$\frac{1}{4}L^M L^N$ (type MN)	$\frac{1}{4}L^N L^N$ (type N)

The offspring can have type M blood with probability 1/4, type N blood with probability 1/4, and type MN blood with probability 1/2.

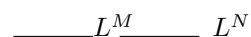
(b)



Female Gamete	Male Gamete	
	$\frac{1}{2}L^M$	$\frac{1}{2}L^N$
$L^M$	$\frac{1}{2}L^M L^M$ (type M)	$\frac{1}{2}L^M L^N$ (type MN)

The offspring can type M or type MN blood, each with probability 1/2.

(c)



Female Gamete	Male Gamete $L^N$
$L^M$	$L^M L^N$ (type MN)

The offspring has type MN blood with probability 1.

**Problem 2.5.** (2 marks) (a) A flower is purple only if its genotype is  $CCPP$ ,  $CCPp$ ,  $CcPP$  or  $CcPp$ . The flower is white in all the other cases, i.e. when its genotype is  $ccPP$ ,  $ccPp$ ,  $CCpp$ ,  $Ccpp$ ,  $ccpp$ .

(b) We cross  $CcPp \times CcPp$ .

Female Gamete	Male Gamete			
	$\frac{1}{4}CP$	$\frac{1}{4}Cp$	$\frac{1}{4}cP$	$\frac{1}{4}cp$
$\frac{1}{4}CP$	$\frac{1}{16}CC PP$ (purple)	$\frac{1}{16}CC Pp$ (purple)	$\frac{1}{16}Cc PP$ (purple)	$\frac{1}{16}Cc Pp$ (purple)
$\frac{1}{4}Cp$	$\frac{1}{16}CC Pp$ (purple)	$\frac{1}{16}CC pp$ (white)	$\frac{1}{16}Cc Pp$ (purple)	$\frac{1}{16}Cc pp$ (white)
$\frac{1}{4}cP$	$\frac{1}{16}Cc PP$ (purple)	$\frac{1}{16}Cc Pp$ (purple)	$\frac{1}{16}cc PP$ (white)	$\frac{1}{16}cc Pp$ (white)
$\frac{1}{4}cp$	$\frac{1}{16}Cc Pp$ (purple)	$\frac{1}{16}Cc pp$ (white)	$\frac{1}{16}cc Pp$ (white)	$\frac{1}{16}cc pp$ (white)

The probability that the offspring has purple flowers is  $9/16$ .

**Problem 3.2.** (2 marks) (a) The percentage is  $100\% - 1.3\% = 98.7\%$ .

(b) Let  $E$  be the event that the person speaks English and  $F$  be the event that the person speaks French. We know that  $P(E \cap F') = 0.599$ ,  $P(F \cap E') = 0.016$  and  $P(E \cup F) = 0.987$ . From the Venn diagram, we have:

$$P(E \cup F) = P(E \cap F') + P(F \cap E') + P(E \cap F).$$

Hence

$$\begin{aligned} P(E \cap F) &= P(E \cup F) - P(E \cap F') - P(F \cap E') \\ &= 0.987 - 0.599 - 0.016 = 0.372 \end{aligned}$$

**Problem 3.5.** (2 marks) Let  $A$  be the event that the tomato has an increased resistance to pests and  $B$  the event that the tomato has a longer shelf life. We know that  $P(A) = 0.75$ ,  $P(B) = 0.5$  and  $P(A \cap B) = 0.3$ .

- (a)  $P(A \cap B') = P(A) - P(A \cap B) = 0.75 - 0.3$ .
- (b)  $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.75 + 0.5 - 0.3 = 0.95$ .
- (c)  $P(A' \cap B) = P(B) - P(A \cap B) = 0.5 - 0.3 = 0.2$ .
- (d)  $P(A' \cap B') = 1 - P(A \cup B) = 1 - 0.95 = 0.05$ .

**Problem 4.3.** (2 marks) (a) Let  $F$  be the event that a randomly chosen man has prostate cancer. Let  $A$  be the event that his PSA level is greater than 10,  $B$  be the event that his PSA level is between 4 and 10, and  $C$  be the event that his PSA level is smaller than 4. We know that  $P(F|A) = 0.67$ ,  $P(F|B) = 0.25$  and  $P(F|C) = 0.05$ . By the total probability rule,

$$\begin{aligned} P(F) &= P(F|A)P(A) + P(F|B)P(B) + P(F|C)P(C) \\ &= (0.67)(0.15) + (0.25)(0.10) + (0.05)(0.75) \\ &= 0.1005 + 0.025 + 0.0375 = 0.163. \end{aligned}$$

(b) By the Bayes' rule,

$$P(A|F) = \frac{P(F \cap A)}{P(F)} = \frac{P(F|A)P(A)}{P(F)} = \frac{0.1005}{0.163} = 0.62.$$

**Problem 4.4.** (2 marks) We observed

	Diseased	Non-diseased
Test +	197	8
Test -	3	92
Total	200	100

(a) The false positive rate is  $P(T + |U-) = 8/100 = 0.08$ . The false negative rate is  $P(T - |U+) = 3/200 = 0.015$ .

(b) The sensitivity is  $P(T + |U+) = 197/200 = 0.985$ .

The specificity is  $P(T - |U-) = 92/100 = 0.92$ .

(c) The positive predictive value is

$$\begin{aligned} \text{PPV} = P(U + |T+) &= \frac{P(T + |U+)P(U+)}{P(T + |U+)P(U+) + P(T + |U-)P(U-)} \\ &= \frac{(197/200)(0.15)}{(197/200)(0.15) + (8/100)(0.85)} = 0.685. \end{aligned}$$

The negative predictive value is

$$\begin{aligned} \text{NPV} = P(U - |T-) &= \frac{P(T - |U-)P(U-)}{P(T - |U-)P(U-) + P(T - |U+)P(U+)} \\ &= \frac{(92/100)(0.85)}{(92/100)(0.85) + (3/200)(0.15)} = 0.9971. \end{aligned}$$

(d) The positive predictive value is

$$\begin{aligned} \text{PPV} = P(U + |T+) &= \frac{P(T + |U+)P(U+)}{P(T + |U+)P(U+) + P(T + |U-)P(U-)} \\ &= \frac{(197/200)(0.01)}{(197/200)(0.01) + (8/100)(0.99)} = 0.1106. \end{aligned}$$

The negative predictive value is

$$\begin{aligned} \text{NPV} = P(U - |T-) &= \frac{P(T - |U-)P(U-)}{P(T - |U-)P(U-) + P(T - |U+)P(U+)} \\ &= \frac{(92/100)(0.99)}{(92/100)(0.99) + (3/200)(0.01)} = 0.9998. \end{aligned}$$

**Problem 5.1.** (2 marks) (Solution 1) Let  $A_1, A_2, A_3, A_4$  be the event that the donor's blood type is O, A, B, respectively AB. Let  $B_1, B_2, B_3, B_4$  be the event that the blood type of the receiving individual is O, A, B, respectively AB. The event  $A_i$  is independent of  $B_j$ , for any  $i = 1, 2, 3, 4$  and  $j = 1, 2, 3, 4$ . The event that the transfusion is successful can be written as the following union of disjoint events:

$$C = (A_1 \cap B_1) \cup (A_1 \cap B_2) \cup (A_1 \cap B_3) \cup (A_1 \cap B_4) \cup (A_2 \cap B_2) \cup (A_2 \cap B_4) \cup (A_3 \cap B_3) \cup (A_3 \cap B_4) \cup (A_4 \cap B_4).$$

Hence,

$$\begin{aligned} P(C) &= P(A_1)P(B_1) + P(A_1)P(B_2) + P(A_1)P(B_3) + P(A_1)P(B_4) + P(A_2)P(B_2) + \\ &\quad P(A_2)P(B_4) + P(A_3)P(B_3) + P(A_3)P(B_4) + P(A_4)P(B_4) \\ &= (0.46)(0.46) + (0.46)(0.42) + (0.46)(0.09) + (0.46)(0.03) + (0.42)(0.42) + \\ &\quad (0.42)(0.03) + (0.09)(0.09) + (0.09)(0.03) + (0.03)(0.03) \\ &= 0.6607. \end{aligned}$$

The probability that the transfusion is not successful is:

$$P(C') = 1 - P(C) = 1 - 0.6607 = 0.3393.$$

**Problem 5.3.** (2 marks) (a) Let  $p_1$  be the percentage of the male population and  $p_2$  be the percentage of the female population. We know  $p_1/p_2 = 0.83$ . Hence  $p_1 = (0.83)p_2$ . On the other hand,  $p_1 + p_2 = 1$ . We obtain:  $(0.83)p_2 + p_2 = 1$ , i.e.  $(1.83)p_2 = 1$ . Hence  $p_2 = 1/1.83 = 0.546$  and  $p_1 = 1 - 0.546 = 0.454$ .

(b) Let  $F$  and  $M$  be the event that the person is a female, respectively a male. Let  $D$  be the event that the person suffers from depression. We know that  $P(F) = 0.546$ ,  $P(M) = 0.454$ ,  $P(D|F) = 0.16$  and  $P(D|M) = 0.08$ . By the total probability rule,

$$\begin{aligned} P(D) &= P(D|F)P(F) + P(D|M)P(M) = (0.16)(0.546) + (0.08)(0.454) \\ &= 0.08736 + 0.03632 = 0.12368. \end{aligned}$$

(c) Since  $P(D) \neq P(D|F)$ ,  $D$  and  $F$  are not independent.

(d) By the Bayes' rule,

$$P(F|D) = \frac{P(F \cap D)}{P(D)} = \frac{0.08736}{0.12368} = 0.706.$$

**Problem 8.5.** (2 marks) A random person is chosen in Africa. Let  $A$  be the event that the person had access to safe drinking water, and  $B$  the event that the person suffers from waterborne disease. We know that  $P(A) = 0.45$ ,  $P(B|A) = 0.32$  and  $P(B|A') = 0.88$ .

(a) By the total probability rule,

$$P(B) = P(B|A)P(A) + P(B|A')P(A') = (0.32)(0.45) + (0.88)(0.55) = 0.144 + 0.484 = 0.628$$

The incidence rate of waterborne diseases in Africa is 62.8%.

(b) By the Bayes' rule,

$$P(A'|B) = \frac{P(B|A')P(A')}{P(B)} = \frac{0.484}{0.628} = 0.77$$

**Problem 8.7.** (2 marks) (a) Let  $A_i$  be the event that the  $i$ -th selected fly is black, for  $i = 1, 2$ . Using the total probability rule:

$$\begin{aligned} P(A_2) &= P(A_2|A_1)P(A_1) + P(A_2|A_1')P(A_1') \\ &= \frac{4}{24} \cdot \frac{5}{25} + \frac{5}{24} \cdot \frac{20}{25} = \frac{1}{5} \end{aligned}$$

Since  $P(A_2|A_1) = 4/24 = 1/6 \neq 1/5 = P(A_2)$ ,  $A_1$  and  $A_2$  are not independent.

(b) As in part (a), we have:

$$\begin{aligned} P(A_2) &= P(A_2|A_1)P(A_1) + P(A_2|A_1')P(A_1') \\ &= \frac{1,999}{9,999} \cdot \frac{2,000}{10,000} + \frac{2,000}{9,999} \cdot \frac{8,000}{10,000} = \frac{1}{5} \end{aligned}$$

Since  $P(A_2|A_1) = 0.2 \approx 1,999/9,999 = P(A_2)$ ,  $A_1$  and  $A_2$  are (almost) independent.