Laura Whiting
300019191

25



## MAT 2379, Introduction to biostatistics

### Assignment 1

*Deadline: Before 3 pm on Thursday, My 23*

- Solve the following problems using a calculator permitted by the Faculty of Science (TI30, TI34, Casio fx-260 and Casio fx-300).
- Late assignments are not accepted.
- There are 6 questions.
- You should complete ALL the questions in the assignments. It is possible, however, that not all the questions are going to be marked. You will not be informed beforehand which questions are going to be marked.

**Question 1:** Problem 1.6 from the textbook.

**Question 2:** Problem 2.8 from the textbook.

**Question 3:** Problem 3.2 from the textbook.

**Question 4:** Due to an inappropriate nutrition program, the chicken in a large poultry farm have developed some nutritional diseases. It is estimated that 4% of the poultry have the fatty liver syndrome, a condition characterized by obesity and an enlarged fatty liver, and 3% suffer from cage layer fatigue, a condition which results in soft bones, similar to osteoporosis. 4% of the chickens have at least one of the conditions, but not both.

(a) A chicken is randomly selected from this farm. What is the probability that it will have both conditions?

(b) A chicken is randomly selected from this farm. What is the probability that it has none of the conditions?

**Question 5:** There is a new diagnostic test for a disease that occurs in about 1.5% of the population. The test is not perfect. Its sensitivity is 99% and its specificity is 97%. A person is chosen at random from this population and the tests indicates that this person has the disease. Compute the positive predictive value of the test.

3.3 in book

**Question 6:** The nuchal translucency test is a special ultrasound scan which is widely used as a screening test for Down's syndrome in early pregnancy. The test measures the fluid under the skin at the back of the baby's neck and can be used to determine the risk of having a baby with Down's syndrome. The following table gives the test results for a sample of 1,000 pregnant women, with the age between 35 and 40:

| | Down Syndrome Baby | Normal Baby | Total |
|---|---|---|---|
| Test + | 5 | 55 | 60 |
| Test - | 3 | 937 | 940 |
| Total | 8 | 992 | 1000 |

(a) Compute the false positive rate and the false negative rate.

(b) Compute the sensitivity and the specificity of the test.

+9 - Assignment 1

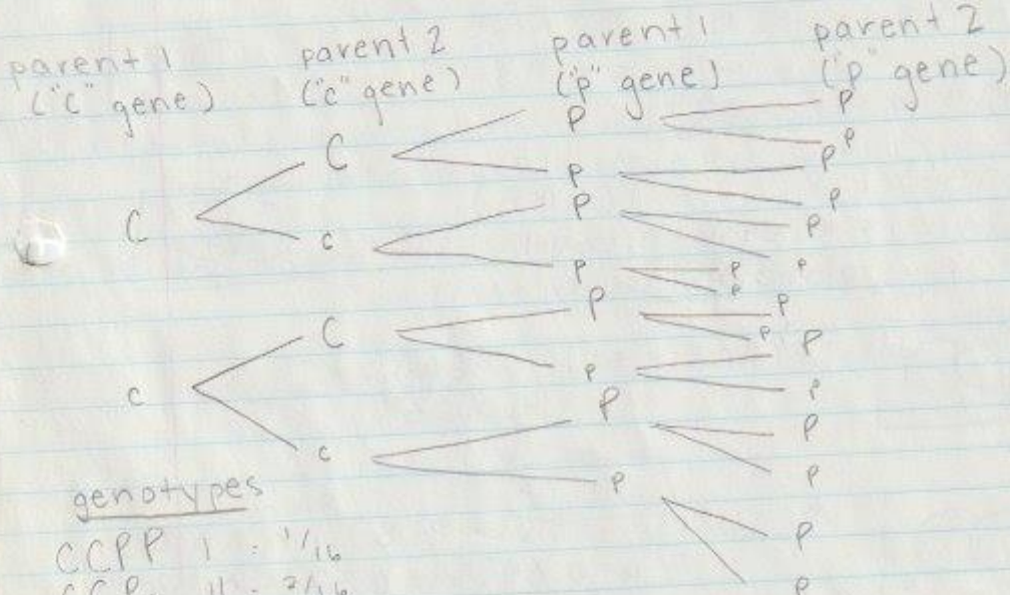
Laura Whiting
300019191

1.6. a) All the possible genotypes for a pea plant with white flowers:

+1

| |
|---|
| ccpp |
| Ccpp |
| CCpp |
| ccPp |
| ccPP |

b) CcPp × CcPp




genotypes

CCPP | : 1/16
CCPp || : 2/16
CCpp | : 1/16
CcPP || : 2/16
CcPp |||| : 4/16
Ccpp || : 2/16
ccPP | : 1/16
ccPp || : 2/16
ccpp | : 1/16

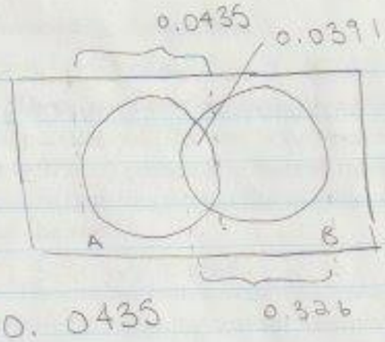
phenotypes

white : 7/16
purple : 9/16

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

t6

2.8




A = A+ in chemistry
B = A+ in biology

$P(A) = 50/1150 = 0.0435$

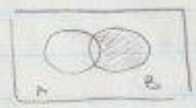
$P(B) = 375/1150 = 0.326$

$P(A \cap B) = 45/1150 = 0.0391$

a) $P(A' \cap B) = P(B) - P(A \cap B)$

$= 0.326 - 0.0391$

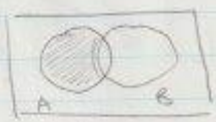
$= \boxed{0.287}$



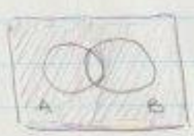

b) $P(A \cap B') = P(A) - P(A \cap B)$

$= 0.0435 - 0.0391$

$= \boxed{0.0044}$




c)

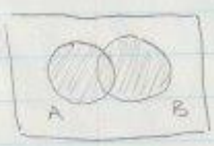



$P(A \cap B)' = 1 - P(A \cap B)$

$= 1 - 0.0391$

$= \boxed{0.961}$

d)

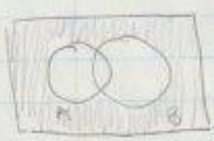



$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$= 0.0435 + 0.326 - 0.0391$

$= \boxed{0.330}$

e)




$P(A \cup B)' = 1 - P(A \cup B)$

$= 1 - 0.330$

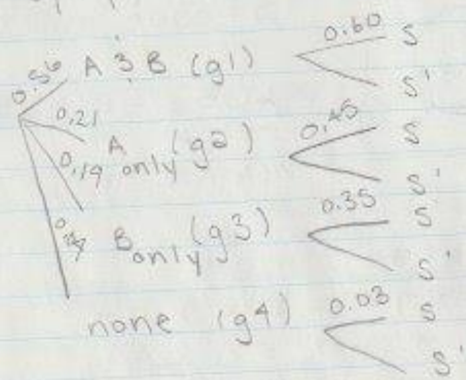
$= \boxed{0.670}$

3.2.

A = restless sleep
B = excessive sleepiness during day

15

symptom




S = sleep apnea

S' = no sleep apnea

a) P(total patients with sleep apnea) = $P(S \cap (A \cap B)) + P(S \cap (A \cap B')) + P(S \cap (A' \cap B)) + P(S \cap (A \cup B)')$

total P(sleep apnea) = $P(S \mid A \cap B)P(A \cap B) + P(S \mid A \cap B')P(A \cap B') + P(S \mid A' \cap B)P(A' \cap B) + P(S \mid (A \cup B)')P((A \cup B)')$

$= (0.60)(0.56) + (0.45)(0.21) + (0.35)(0.19) + (0.03)(0.04)$

$= 0.4982 = \boxed{49.8\%}$

b) $P(A \mid S) = \frac{P(S \mid A)P(A)}{P(S)}$

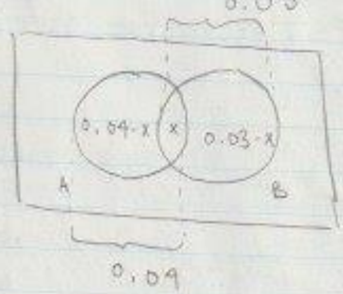
$= \frac{P(S \mid A \cap B)P(A \cap B)}{P(S)} + \frac{P(S \mid A \cap B')P(A \cap B')}{P(S)}$

$= \frac{(0.60)(0.56)}{0.4982} + \frac{(0.45)(0.21)}{0.4982}$

$= 0.6744 + 0.1897$

$= \boxed{0.864}$

question 4:




A = fatty liver syndrome
B = cage layer fatigue

$P(A \cup B) - P(A \cap B)$

or

$P(A \cap B') + P(A' \cap B) = 0.04$

a) $P(A \cup B) - P(A \cap B) = 0.04$

$(P(A) + P(B) - x) - (x) = 0.04$

$0.04 + 0.03 - x - x = 0.04$

$0.07 - 2x = 0.04$

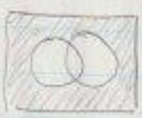
$-0.07 \qquad -0.07$

$\frac{-2x}{-2} = \frac{-0.03}{-2}$

$x = 0.015$

$\therefore P(A \cap B) = \boxed{0.015}$

b) $P(A \cup B)' = 1 - P(A \cup B)$



$= 1 - (P(A) + P(B) - P(A \cap B))$

$= 1 - (0.04 + 0.03 - 0.015)$

$= 1 - 0.055$

$= \boxed{0.945}$

question 5:

D = has disease
D′ = not have disease

+4

sensitivity = P(test + | D) = 0.99

specificity = P(test − | D′) = 0.97

P(D) = 0.015

$$PPV = (D \mid test +) = \frac{TP}{TP + FP}$$

$$= \frac{P(test + \cap D)}{P(test + \cap D) + P(test + \cap D')}$$

$$= \frac{P(test + \mid D)\,P(D)}{P(test + \mid D)\,P(D) + P(test + \mid D')\,P(D)}$$

$$= \frac{(0.99)(0.015)}{(0.99)(0.015) + (1 - 0.97)(1 - 0.015)}$$

$$= \frac{(0.99)(0.015)}{(0.99)(0.015) + (0.03)(0.985)}$$

$$= \frac{0.01485}{0.01485 + 0.02955}$$

PPV = **0.334**

+4

question 6: D: has Down's

D' = does not have Down's.

a) false positive rate:

$P(\text{test} + | D') = \frac{P(\text{test} + \cap D')}{P(D')} = \frac{65}{992} = \boxed{0.0655}$

not written as probabilites (divided by total number of patients) because will cancel anyways to give same $P(\text{test}+|D')$.

false negative rate:

$P(\text{test} - | D) = \frac{P(\text{test} - \cap D)}{P(D)} = \frac{3}{8} = \boxed{0.375}$

b) sensitivity = $P(\text{test} + | D) = \frac{P(\text{test} + \cap D)}{P(D)}$

$= \frac{5}{8} = \boxed{0.625}$

c) specificity = $P(\text{test} - | D') = \frac{P(\text{test} - \cap D')}{P(D')}$

$= \frac{937}{992} = \boxed{0.946}$