Session 1: Probability

Exercise Solutions

Question 1

Suppose you roll a 6-sided die once.
Consider the following events:

E_1 = roll a 1
E_2 = roll an even number
E_3 = roll a 4, 5 or 6
E_4 = roll a 3 or 5

a) What is $Pr(E_4)$? **2/6**

b) Are E_2 and E_3 mutually exclusive? **No** E_2 and E_4? **Yes**

c) Find a mutually exclusive, exhaustive collection of events from among the defined events above. **E1, E2, E4**
Do the probabilities add to 1? **Yes**

d) What is $Pr(E_4^c)$? **4/6**

Question 2

Suppose we screen 10,000 people for a disease using a new screening test.
The data are summarized in the following table.

		Disease Status		
		Positive	*Negative*	
Test Result	*Positive*	9	80	89
	Negative	1	9,910	9,911
		10	9,990	**10,000**

a) What is P(test positive)? **89/10,000**

b) What is P(test positive and disease positive)? **9/10,000**

c) What is P(test positive or disease positive)? **90/10,000**

d) What is P(test positive | disease positive)? **9/10**

e) What is P(disease positive | test positive)? **9/89**

Question 3

In a group of 30 symptomatic women attending a clinic, some had cervical infections with *Chlamydia trachomatis* (C) or *Neisseria gonorrhea* (G), some were harboring both organisms, and some had neither. Seven women had C only, five women had G only and eight women had both (B).

a) What is the probability of chlamydia (C)? 7/30 + 8/30 = **15/30**

b) What is the probability of gonorrhea (G)? 5/30 + 8/30 = **13/30**

c) What is the probability of gonorrhea (G) or chlamydia (C)? 15/30 + 13/30 - 8/30 = **20/30**

d) Are gonorrhea and chlamydia mutually exclusive? **No.**

Question 4

A certain operation has a survival rate of 70%. If this operation is performed independently on three different patients, what is the probability all three operations will fail?

P(fail, fail, fail) = P(fail) x P(fail) x P(fail) = 0.3 x 0.3 x 0.3 = **0.027**

Question 5

If allele A has frequency 3/4 and allele a has frequency 1/4 , what are the prevalences of the 3 genotypes AA, Aa, and aa in the population (assuming random mating)?

P(AA) = (3/4) x (3/4) = **9/16**

P(Aa) = 2 x [3/4 x 1/4] = **6/16**

P(aa) = 1/4 x 1/4 = **1/16**

Exercises

Question 6

Suppose an influenza epidemic strikes a city. In two-child families, the older child has influenza 10% of the time (event A), the younger child has influenza 10% of the time (event B), and both children have influenza 2% of the time.

a) Are the events A and B independent?

No, since 0.1 x 0.1 ≠ 0.02

b) What is the probability that the older child has influenza if we know the younger child has influenza?

P(A | B) = P(A,B) / P(B) = 0.02/0.1 = **0.2**

Question 7

The following table gives the probability of a low birth weight (LBW) (< 2500 g) baby for different gestational ages. What is the overall probability of having a LBW infant in this population?

Length of Gestation	Proportion born at given gestational age	P(LBW) at given gestational age
< 20 weeks	0.0004	0.540
20 – 27 weeks	0.0059	0.813
28 – 36 weeks	0.0855	0.379
> 36 weeks	0.9082	0.035

prob = 0.0004 * 0.54 + 0.0059 * 0.813 + 0.0855 * 0.379 + 0.9082 * 0.035 = **0.069**

Question 8

An investigator wants to determine the characteristics of a screening test for bacterial vaginosis (BV). The investigator obtains screening results on 250 women and finds 50 who screen positive and 200 who screen negative. The "gold standard" (i.e., true measure of disease status) is a gram stain test but this test is more expensive to do. Therefore, the investigator decides to do gram stains on all 50 women who screened positive and a random sample of 50 of the women who screened negative. The following results are obtained:

		Gram stain		
		Positive	*Negative*	
Screening Test	*Positive*	44	6	50
	Negative	3	47	50
		47	53	**100**

a) Estimate the sensitivity and specificity of the screening test.
   sensitivity = 44/(44 + 3*4) = **0.786**
   specificity = 47/(6 + 47*4) = **0.242**

b) Estimate the positive predictive value (PPV) of the screening test.
   44/50 = **0.88**