MATH 205: Statistical methods

Lecture 12: Bayes' rule

Tentative schedule

Date	Theme/Topic	Labs	Assignments
Aug 31	Syllabus		
Sep 2—9	Chapter 1: Describing dataset	Section 2: Handling data	
Sep 12—16	Chapter 2: Looking at Relationships	Section 3: Univariate data	
Sep 19—23	Chapter 3: Basic Ideas in Probability	Section 4: Bivariate Data	Homework 1 (due 09/23)
Sep 26—30	Chapters 3-4	Section 4: Correlation	
Oct 3—7	Chapter 4: Random variables and expectations	Section 6: Random data	Homework 2 (due 10/07)
Oct 10—14	Chapter 5: Useful distributions	Section 7: The central limit theorem	
Oct 17—21	Chapter 6: Samples and populations	Section 9: Confidence interval estimation	Homework 3 (due 10/21)
Oct 24—28	Review Midterm exam		Midterm: Oct 28 (lecture) Oct 24-26 (labs)
Oct 31—Nov 4	Chapter 7: The significance of evidence	Section 10: Hypothesis testing	
Nov 7—11	Goodness of Fit	Section 12: Goodness of Fit	Homework 4 (due 11/11)
Nov 14—18	Linear Regression	Section 13: Linear regression	
Nov 21–25	Thanksgiving break		
Nov 28 —Dec 2	One-Way Analysis of Variance	Section 15: Analysis of variance	Homework 5 (due 12/02)
Dec 5—7	Selected topics + Review		
Exam week			

Chapter 3: Basic ideas in probability

- Experiments, outcomes, events, and probability
- Independence
- Conditional probability

Conditional probability

Definition

Let $P(A) > 0$, the conditional probability of B given A, denoted by $P(B|A)$, is

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

Law of multiplication

$$P(B \cap A) = P(B|A)P(A)$$

Problem

We throw two fair six-sided dice. What is the conditional probability that the sum of spots on both dice is greater than six, conditioned on the event that the first die comes up five?

Properties

- Law of multiplication

$$P(B \cap A) = P(B|A)P(A)$$

- Law of total probability

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$$

- **Bayes' rule**

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

Law of total probability

Useful Facts 3.3 (Conditional Probability Formulas)
You should remember the following formulas:

- $P(\mathcal{B}|\mathcal{A}) = \dfrac{P(\mathcal{A}|\mathcal{B})P(\mathcal{B})}{P(\mathcal{A})}$
- $P(\mathcal{A}) = P(\mathcal{A}|\mathcal{B})P(\mathcal{B}) + P(\mathcal{A}|\mathcal{B}^c)P(\mathcal{B}^c)$
- Assume (a) $\mathcal{B}_i \cap \mathcal{B}_j = \varnothing$ for $i \neq j$ and (b) $\mathcal{A} \cap (\cup_i \mathcal{B}_i) = \mathcal{A}$; then $P(\mathcal{A}) = \sum_i P(\mathcal{A}|\mathcal{B}_i)P(\mathcal{B}_i)$

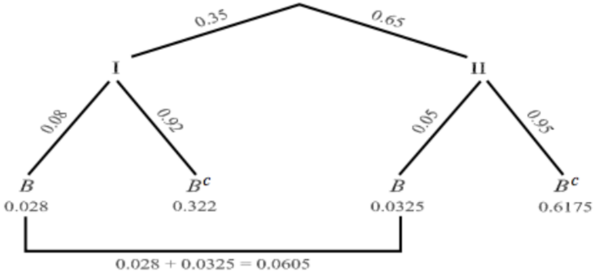
Law of total probability

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$$

Problem

An insurance company rents 35% of the cars for its customers from agency I and 65% from agency II. We also know that 8% of the cars of agency I and 5% of the cars of agency II break down during the rental periods.
What is the probability that a car rented by this insurance company breaks down?

Example



0.35
0.65
I
II
0.08
0.92
0.05
0.95
B
0.028
B^c
0.322
B
0.0325
B^c
0.6175
0.028 + 0.0325 = 0.0605

Bayes' rule

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

Problem

An insurance company rents 35% of the cars for its customers from agency I and 65% from agency II. We also know that 8% of the cars of agency I and 5% of the cars of agency II break down during the rental periods.
Assuming that a randomly selected car broke down, what is the probability that this car is from agency I?

A Pap smear is a procedure used to detect cervical cancer. For women with this cancer, there are about 16% false negatives:

$$P[\text{test negative}|\text{patient has cancer}] = 0.16$$

$$P[\text{test positive}|\text{patient has cancer}] = 0.84.$$

For women without cancer, there are about 10% false positives:

$$P[\text{test positive}|\text{patient does not have cancer}] = 0.10$$

$$P[\text{test negative}|\text{patient does not have cancer}] = 0.90$$

In the United States, there are about 8 women in 100,000 who have this cancer. Assume that a woman is taking the test. Given that the test is positive, what is the probability that she has cervical cancer?