STA 291
Fall 2009

LECTURE 17
THURSDAY, 22 OCTOBER

Le Menu

- **5 Probability (Review, mostly)**

- **21 Random Variables and Discrete Probability Distributions**

Suggested problems

- Suggested problems from the textbook:

  21.1 to 21.4,21.7 , 21.8, 21.10, and 21.11

Conditional Probability & the Multiplication Rule

$$P(A|B)=\frac{P(A\cap B)}{P(B)}, \text{provided } P(B)\neq 0$$

- Note: $P(A|B)$ is read as "the probability that A occurs given that B has occurred."
- Multiplied out, this gives *the multiplication rule:*

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

Multiplication Rule Example

- The multiplication rule:

$$P(A \cap B) = P(B) \times P(A \mid B)$$

- Ex.: A disease which occurs in .001 of the population is tested using a method with a false-positive rate of .05 and a false-negative rate of .05. If someone is selected and tested at random, what is the probability they are positive, and the method shows it?

Independence

- If events A and B are independent, then the events A and B have no influence on each other.

- So, the probability of A is unaffected by whether B has occurred.

- Mathematically, if A is independent of B, we write: $P(A|B) = \mathrm{P}(A)$

Multiplication Rule and Independent Events

Multiplication Rule for Independent Events: Let A and B be two independent events, then

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

Examples:

- Flip a coin twice. What is the probability of observing two heads?
- Flip a coin twice. What is the probability of getting a head and then a tail? A tail and then a head? One head?
- Three computers are ordered. If the probability of getting a "working" computer is 0.7, what is the probability that all three are "working" ?

Conditional Probabilities—Another Perspective

Example: Smoking and Lung Disease I

Joint Probabilities	Lung Disease	Not Lung Disease	*Row Totals*
Smoker	.12	.19	.31
Nonsmoker	.03	.66	.69
Column Totals	.15	.85	1.00

Conditional Probabilities—Another Perspective

Example: Smoking and Lung Disease I

Joint Probabilities	Lung Disease	Not Lung Disease	*Row Totals*
Smoker	.12	.19	.31
Nonsmoker	.03	.66	.69
Column Totals	.15	.85	1.00

Example: Smoking and Lung Disease II

Conditional Row Probabilities	Lung Disease	Not Lung Disease	*Row Totals*
Smoker	.12/.31 =.39	.19/.31 =.61	.31/.31 =1.00
Nonsmoker	.03/.69 =.04	.66/.69 =.96	.69/.69 =1.00
Smokers and Nonsmokers	.15	.85	1.00

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

Conditional Probabilities—Another Perspective

Example: Smoking and Lung Disease I

Joint Probabilities	Lung Disease	Not Lung Disease	*Row Totals*
Smoker	.12	.19	.31
Nonsmoker	.03	.66	.69
Column Totals	.15	.85	1.00

Example: Smoking and Lung Disease III

Conditional Column Probabilities	Lung Disease	Not Lung Disease	*Lung Disease and Not Lung Disease*
Smoker	.12/.15 =.80	.19/.85 =.22	.31
Nonsmoker	.03/.15 =.20	.66/.85 =.78	.69
Column Totals	.15/.15 =1.00	.85/.85 =1.00	1.00

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

Terminology

- $P(A \cap B) = P(A,B)$ joint probability of A and B (of the intersection of A and B)
- $P(A|B)$ conditional probability of A given B
- $P(A)$ marginal probability of A

Chapter 21: Random Variables

- A variable X is a **random variable** if the value that X assumes at the conclusion of an experiment cannot be predicted with certainty in advance.
- There are two types of random variables:

– **Discrete**: the random variable can only assume a finite or countably infinite number of different values (almost always a count)

– **Continuous**: the random variable can assume all the values in some interval (almost always a physical measure, like distance, time, area, volume, weight, ...)

Examples

Which of the following random variables are discrete and which are continuous?

a. X = Number of houses sold by real estate developer per week?

b. X = Number of heads in ten tosses of a coin?

c. X = Weight of a child at birth?

d. X = Time required to run a marathon?

Properties of Discrete Probability Distributions

Definition: A Discrete probability distribution is just a list of the possible values of a r.v. X, say (x_i) and the probability associated with each $P(X=x_i)$.

Properties:

1. All probabilities non-negative.
2. Probabilities sum to ________ .

$$P(x_i) \geq 0$$

$$\sum P(x_i) = 1$$

Example

The table below gives the # of days of sick leave for 200 employees in a year.

Days	0	1	2	3	4	5	6	7
Number of Employees	20	40	40	30	20	10	10	30

An employee is to be selected at random and let X = # days of sick leave.

a.) Construct and graph the probability distribution of X

b.) Find $P(X \leq 3)$

c.) Find $P(X > 3)$

d.) Find $P(3 \leq X \leq 6)$

Population Distribution vs. Probability Distribution

- If you select a subject randomly from the population, then the probability distribution for the value of the random variable X is the relative frequency (population, if you have it, but usually approximated by the sample version) of that value

Cumulative Distribution Function

Definition: The *cumulative distribution function*, or *CDF* is

$$F(x) = P(X \leq x).$$

Motivation: Some parts of the previous example would have been easier with this tool.

Properties:

1. For any value x, $0 \leq F(x) \leq 1$.
2. If $x_1 < x_2$, *then* $F(x_1) \leq F(x_2)$
3. $F(-\infty) = 0$ *and* $F(\infty) = 1$.

Example

Let X have the following probability distribution:

X	2	4	6	8	10
$P(x)$	.05	.20	.35	.30	.10

a.) Find $P (X \leq 6)$

b.) Graph the cumulative probability distribution of X

c.) Find $P (X > 6)$

Attendance Question #17

Write your name and section number on your index card.

Today's question: