STA 291
Summer 2008

Lecture 5

Casinos

- "There is no such thing as luck. It is all mathematics." Casinos make money on their games because of the mathematics behind the games.
- With a few notable exceptions, the house always wins - *in the long run* - because of the mathematical advantage the casino enjoys over the player.
- Because of a famous mathematical result called the law of large numbers, a casino is guaranteed to win *in the long run*.
- In the gambling industry, nothing plays a more important role than mathematics.

6. Probability

- Abstract but necessary because this is the mathematical theory underlying all statistical inference
- Fundamental concepts that are very important to understanding *Sampling Distribution, Confidence Interval*, and *P-Value*
- Our goal for Chapter 6 is to learn the rules involved with assigning probabilities to events

Probability: Basic Terminology

- **Experiment:** Any activity from which an outcome, measurement, or other such result is obtained.
- **Random (or Chance) Experiment:** An experiment with the property that the outcome cannot be predicted with certainty.
- **Outcome:** Any possible result of an experiment.
- **Sample Space:** The collection of all possible outcomes of an experiment.
- **Event:** A specific collection of outcomes.
- **Simple Event:** An event consisting of exactly one outcome.

Experiments, Outcomes, Sample Spaces, and Events

Examples:

1. Flip a coin
2. Flip a coin 3 times
3. Roll a die
4. Draw a SRS of size 50 from a population
5. Time your commuting time in a weekday morning.
6. A football game between two chosen teams
7. Give AIDS patient a treatment and record how long he/she lives.

- There are often more than one way to assign probability to a sample space. (could be infinitely many ways)
- Statistical techniques we will learn later help identify which one "reflecting the truth".
- In Chap. 6/7 we do not identify which probability is more "true" but just learn the consequences of a (legitimate) probability assignment.

legitimate probability

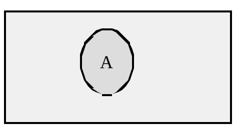
- P(A) has to be a number between 0 and 1, inclusive
- The probability of the sample space S must be 1: P(S) =1
- When we able to list all simple events, O_i summation over i for $P(O_i)$ must be 1

$$\Sigma\, P(O_i) = 1$$

P(A) could be computed by sum of $P(O_i)$ over those O inside A.

Complement

- Let *A* denote an event.
- **The complement of an event** *A***:** All the outcomes in the sample space *S* that do not belong to the event *A*. The complement of *A* is denoted by A^C



S

Law of Complements:

$$P(A^C) = 1 - P(A)$$

Example: If the probability of getting a "working" computer is 0.7, What is the probability of getting a defective computer?

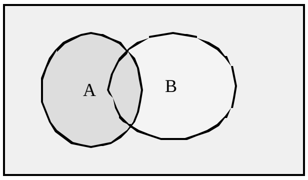
Union and Intersection

- Let *A* and *B* denote two events.
- **The union of two events:** All the outcomes in *S* that belong to at least one of *A* or *B*. The union of *A* and *B* is denoted by $A \cup B$
- **The intersection of two events:** All the outcomes in *S* that belong to both *A* and *B*. The intersection of *A* and *B* is denoted by $A \cap B$

Additive Law of Probability

Let A and B be two events in a sample space S. The probability of the union of A and B is

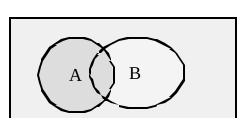
$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$



S

Using Additive Law of Probability

Example: At a large University, all first-year students must take chemistry and math. Suppose 15% fail chemistry, 12% fail math, and 5% fail both. Suppose a first-year student is selected at random. What is the probability that this student failed at least one of the courses?



S

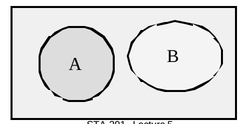
Disjoint Events

- Let *A* and *B* denote two events.
- **Disjoint (mutually exclusive) events:** *A* and *B* are said to be disjoint if there are no outcomes common to both *A* and *B*.
- Using notation, this is written as $A \cap B = \varnothing$
- Note: The last symbol denotes the null set or the empty set.

Disjoint Events

Let A and B be two events in a sample space S. The probability of the union of two disjoint (mutually exclusive) events A and B is

$$P(A \cup B) = P(A) + P(B).$$



S

Assigning Probabilities to Events

- The probability of an event is a value between 0 and 1.
- In particular:
  - 0 implies that the event will not occur
  - 1 implies that the event will occur
- How do we assign probabilities to events?

Assigning Probabilities to Events

- There are different approaches to assigning probabilities to events
- Objective
  - equally likely outcomes (classical approach)
  - relative frequency
- Subjective

Equally Likely Approach (Laplace)

- The equally likely outcomes approach usually relies on symmetry/geometry to assign probabilities to events.
- As such, we do not need to conduct experiments to determine the probabilities.
- Suppose that an experiment has only *n* outcomes. The equally likely approach to probability assigns a probability of 1/*n* to each of the outcomes.
- Further, if an event *A* is made up of *m* outcomes, then *Prob*(*A*) = *m*/*n*.

Equally Likely Approach

- Examples:

1. Select a SRS of size 2 from a population
2. Roll a fair die
   - The probability of getting “5” is 1/6
   - This does not mean that whenever you roll the die 6 times, you definitely get exactly one “5”
   - The probability of the event “4 or above” is

Counting method

- At every step, you always have **k** choices
- There are **m** steps

- Total number of choices = k k k k …k = k to m power

- Example: pick 3 lotto [10 to 3 power]

Counting method II

- Ranking of **n** distinct objects

- You first decide who is #1, then who is #2, etc.

- Total choices = n(n-1)(n-2)(n-3)…(3) (2) 1
  = n !

Counting method III

- Number of ways to select a group of **m** out of **n** candidates [no order]

n(n-1)(n-2)…(n-m+1)
----
m(m-1)(m-2)…3 2 1

Relative Frequency Approach (von Mises)

- The relative frequency approach borrows from calculus' concept of limit.
- Here's the process:

1. Repeat an experiment *n* times.
2. Record the number of times an event *A* occurs. Denote that value by *a*.
3. Calculate the value *a/n*

Relative Frequency Approach

- We could then define the probability of an event *A* in the following manner:

$$\text{Prob}(A) = \lim_{n\to\infty} \frac{a}{n}$$

- Typically, we can't can't do the "*n to infinity*" in real-life situations, so instead we use a "large" *n* and say that

$$\text{Prob}(A) \approx \frac{a}{n}$$

Example

Relative Frequency Approach: Relative frequency of an event occurring in an infinitely large number of trials

Time Period	Number of Male Live Births	Total Number of Live Births	Relative Frequency of Live Male Birth
1965	1927.054	3760.358	0.51247
1965-1969	9219.202	17989.360	0.51248
1965-1974	17857.860	34832.050	0.51268

Online Applet: Coin Tossing

Relative Frequency Approach

- What is the formal name of the device that allows us to use "large" *n*?
- Law of Large Numbers:
  - As the number of repetitions of a random experiment increases,
  - the chance that the relative frequency of occurrence for an event will differ from the true probability of the event by more than any small number
  - approaches 0.

Subjective Probability Approach

- A subjective probability relies on a person to make a judgment as to how likely an event will occur.
- The events of interest are usually events that cannot be replicated easily or cannot be modeled with the equally likely outcomes approach.
- As such, these values will most likely vary from person to person.
- The only rule for a subjective probability is that the probability of the event must be a value in the interval [0,1].

Probabilities of Events

Let A be the event A = {o_1, o_2, …, o_k}, where o_1, o_2, …, o_k are k different outcomes. Then

$$P(A) = P(o_1) + P(o_2) + \cdots + P(o_k)$$

Problem: The number on a license plate is any digit between 0 and 9. What is the probability that the first digit is a 3? What is the probability that the first digit is less than 4?

Probability tables

- One row of outcomes, one row of corresponding probabilities.
- R x C probability tables: when the outcomes are classified by two features

Example: Smoking and Lung Disease

	Lung Disease	Not Lung Disease	*Marginal (smoke status)*
Smoker	0.12	0.19	
Nonsmoker	0.03	0.66	
Marginal (disease status)			

- Equivalent to a table with 4 entries:

(smoker & lung disease) 0.12
(smoker & not lung disease) 0.19
(nonsmoker & lung disease) 0.03
(nonsmoker & not lung disease) 0.55

But the R x C table reads much better

- From the R x C table we can get a table for smoker status alone, or disease status alone.
- Those are called marginal probabilities

One way street

- Given the joint probability table, we can figure out the marginal probability
- Given the marginal, we may not determine the joint: there can be several different joint tables that lead to identical marginal.

Example: Smoking and Lung Disease

	Lung Disease	Not Lung Disease	*Marginal (smoke status)*
Smoker	0.02	0.29	
Nonsmoker	0.13	0.56	
Marginal (disease status)			

Same marginal, different joint.

Using the table

- P(smoker and lung disease) = 0.02
- P(smoker or lung disease) =

Conditional Probability

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

- Note: *P*(*A*|*B*) is read as "the probability that *A* occurs given that *B* has occurred."
- Note: *P(A* and *B)* = *P(A|B)·P(B)*

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*) = *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" ?

Terminology

- *P(An 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*

Example: Medical Screening

- Screenings are routinely performed on patients
- Examples
  - PSA for men [determining evidence of prostate cancer]
  - Pap for women [cervical cancer]
- These screening tests are not 100% accurate
- False-positive test result: The patient does not have the disease, but the test shows positive
- False-negative result: The patient does have the disease, but the test produces a negative result

Example: Medical Screening

- The Pap smear is the standard test for cervical cancer.
- False-positive rate = 0.636
- False-negative rate = 0.180
- Family history and age are factors that must be considered when assigning a probability of cervical cancer.
- Suppose that the proportion of women a patient's age and with her family history that have cervical cancer is 2%.
- Determine the effects a positive and negative Pap smear test have on the probability that the patient has cervical cancer.

Example

	A_1	A_2
B_1	.20	.15
B_2	.60	.05

Determine whether the events are independent from the following joint probabilities?

Example

	Cash	Credit Card	Debit Card
Under $20	.09	.03	.04
$20~$100	.05	.21	.18
Over $100	.03	.23	.14

a. What proportion of purchase was paid by debit card?

b. Find the probability that a credit card purchase was over $100

c. Determine the proportion of purchases made by credit card or by debit card.

7. 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
  - **Continuous:** the random variable can assume all the values in some interval

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 _______ .

$$0 \le P(x_i) \le 1$$

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

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 $P(X \le 3)$.
c.) Find $P(X \ge 3)$.
d.) Find $P(3 \le X \le 6)$.

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 population distribution of that variable
- Example:

X=number of sick days/height/grade of a randomly chosen person

Cumulative Distribution Function

Definition: The ***cumulative distribution function***, or **CDF is** $F(x) = P(X \le x)$.

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

Properties:
1. For any value x, $0 \le F(x) \le 1$.
2. If $x_1 < x_2$, then $F(x_1) \le F(x_2)$
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 $P(X \le 6)$.
b.) Graph the cumulative probability distribution function.
c.) Find $P(X > 6)$.

Expected Value of a Random Variable

The Expected Value, or mean, of a random variable, X, is

Mean = E(X)= $m = \sum x_i P(X = x_i)$

Example

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

What is E(X)?

Expected Profit Example

- The L. M. Corporation purchases old, run-down buildings, remodels them, and then sells them.
- The corporation has the opportunity to purchase a building for $100,000, which will cost $60,000 to remodel.
- The manager thinks there is a 50% chance that the building will sell for $120,000, yielding a $40,000 loss, a 20% chance that the building will sell for $180,000, yielding a $20,000 profit, and a 30% chance that the building will sell for $230,000, yielding a $70,000 profit.
- What is the expected profit?

Variance of a Random Variable

- Variance= Var(X)

$$= E(X - m)^2 = s^2 = \sum (x_i - m)^2 P(X = x_i)$$

Example

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

What is Var(X)?