STA 321
Spring 2014

LECTURE 9
TUESDAY, 18 FEB

Conditional Probability & the Multiplication Rule

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

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

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

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?

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

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

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

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

Law of total probability

- For E and F are evenets,

$$P(E) = P(E \cap F) + P(E \cap F^c)$$

- Example: *A machine produces parts that are either good (90%), slightly defective (2%), or obviously defective (8%). Now assume that a one-year warranty is given for the parts that are shipped to customers. Suppose that a good part fails within the first year with probability 0.01, while a slightly defective part fails within the first year with probability 0.10. What is the probability that a customer receives a part that fails within the first year and therefore is entitled to a warranty replacement?*

Partition

- A collection of events $\{A_1, A_2, \ldots, A_k\}$ to be said a partition of a sample space S if $A_i \cap A_j$ is empty set.

Example: A is any event. Then $\{A, A^c\}$ is a partition.

Example: *A machine produces parts that are either good (90%), slightly defective (2%), or obviously defective (8%).*

Bayes Rule

- For a given partition of S into sets $F_1, \ldots, F_n$, we want to know the probability that some particular case, F_j , occurs, given that some event E occurs. We can compute this easily using the definition

  $P(F_j|E)= P(F_j \cap E)/P(E)$

- This is called Bayes Formula. By applying the Law of Total Probability, we can rewrite the denominator: $P(E) = \Sigma^n_{i=1} P(E|F_i)P(F_i)$.
- Thus, $P(F_j|E)=P(F_j)\ P(F_j|E) / \Sigma^n_{i=1}P(E|F_i)P(F_i)$.

Example

- *Urn 1 contains 5 white balls and 7 black balls. Urn 2 contains 3 whites and 12 black. A fair coin is flipped; if it is Heads, a ball is drawn from Urn 1, and if it is Tails, a ball is drawn from Urn 2. Suppose that this experiment is done and you learn that a white ball was selected. What is the probability that this ball was in fact taken from Urn 2? (i.e., that the coin flip was Tails)*

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

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

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

Motivation: Some parts of the previous example would have been easier with this 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 $P(X \leq 6)$

b.) Graph the cumulative probability distribution of X

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)= $\mu = \sum x_i P(X = x_i)$

- Back to our previous example—what's E(X)?

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

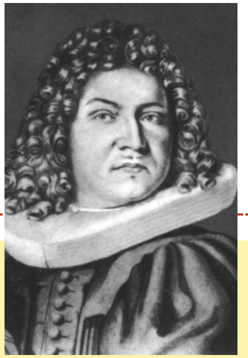
Variance of a Random Variable

- Variance= Var(X) =

$$\sigma^2 = E\left[(X-\mu)^2\right] = \sum (x_i - \mu)^2 \cdot P(X = x_i)$$

- Back to our previous example—what's Var(X)?

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



Bernoulli Trial

- Suppose we have a single random experiment X with two outcomes: "success" and "failure."
- Typically, we denote "success" by the value 1 and "failure" by the value 0.
- It is also customary to label the corresponding probabilities as:

$$P(\text{success}) = P(1) = p \text{ and}$$

$$P(\text{failure}) = P(0) = 1 - p = q$$

- Note: $p + q = 1$

Binomial Distribution I

- Suppose we perform several Bernoulli experiments and they are all independent of each other.
- Let's say we do n of them. The value n is the **number of trials.**
- We will label these n Bernoulli random variables in this manner: $X_1, X_2, \ldots, X_n$
- As before, we will assume that the probability of success in a single trial is p, and that this probability of success doesn't change from trial to trial.

Binomial Distribution II

- Now, we will build a new random variable X using all of these Bernoulli random variables:

$$X = X_1 + X_2 + \cdots + X_n = \sum_{i=1}^{n} X_i$$

- What are the possible outcomes of X?
- What is X counting?
- How can we find $P(X = x)$?

Binomial Distribution III

- We need a quick way to count the number of ways in which k successes can occur in n trials.
- Here's the formula to find this value:

$$\binom{n}{k} = {}_nC_k = \frac{n!}{k!(n-k)!}, \text{ where } n! = n\cdot(n-1)\cdot\dots\cdot 3\cdot 2\cdot 1 \text{ and } 0! \equiv 1$$

- Note: ${}_nC_k$ *is read as "n* choose *k*."

Binomial Distribution IV

- Now, we can write the formula for the binomial distribution:
- The probability of observing *x* successes in *n* independent trials is

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \text{ for } x = 0, 1, \cdots, n$$

under the assumption that the probability of success in a single trial is *p*.

Using Binomial Probabilities

Note: Unlike generic random variables where we would have to be given the probability distribution or calculate it from a frequency distribution, here we can calculate it from a mathematical formula.

Helpful resources (besides your calculator):

- Excel:

Enter	Gives
=BINOMDIST(4,10,0.2,FALSE)	0.08808
=BINOMDIST(4,10,0.2,TRUE)	0.967207

- Table 1, pp. B-1 to B-5 in the back of your book

Binomial Probabilities

We are choosing a random sample of $n = 7$ Lexington residents—our random variable, C = number of Centerpointe supporters in our sample. Suppose, $p = P$ (Centerpointe support) ≈ 0.3. Find the following probabilities:

a) $P(C = 2)$ http://stattrek.com/Tables/Binomial.aspx

b) $P(C < 2)$

c) $P(C \leq 2)$

d) $P(C \geq 2)$

e) $P(1 \leq C \leq 4)$

What is the *expected* number of Centerpointe supporters, μ_C?