

S1211Q Introduction to Statistics

Lecture 5

Wei Wang

July 9, 2012

About Midterm 1

- ▶ Practice questions are given.
- ▶ Two-sided A4 sheet, should not contain any specific or worked-out examples.
- ▶ Bring a calculator, you only need really dumb one.
- ▶ You won't need distribution tables.

Let's go over a hw question together

Ex 2.38 A box in a certain supply room contains four 40-W light-bulbs, five 60-W bulbs, and six 75-W bulbs. Suppose that three bulbs are randomly selected.

Part a

- ▶ What is the probability that exactly two of the selected bulbs are rated 75-W?

Part a

- ▶ What is the probability that exactly two of the selected bulbs are rated 75-W?

- ▶ The probability is given by this ratio

$$\frac{\text{total number of selections that have two 75-W bulbs}}{\text{total number of possible selections}}$$

Part b

- ▶ What is the probability that all three of the selected bulbs have the same rating?

Part b

- ▶ What is the probability that all three of the selected bulbs have the same rating?
- ▶ Decompose the target event into three disjoint events.

Part c

- ▶ What is the probability that one bulb of each type is selected?

Part d

- ▶ Bulbs are to be selected until a 75-W bulb is found. What is the probability of having to exam at least six bulbs?

Part d

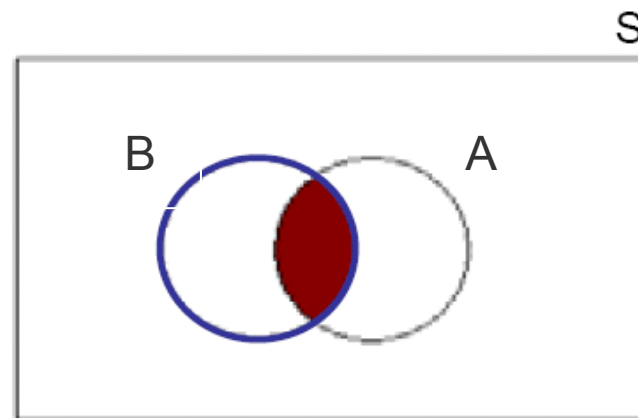
- ▶ Bulbs are to be selected until a 75-W bulb is found. What is the probability of having to exam at least six bulbs?
- ▶ We have to exam at least six bulbs if and only if the first five selections don't contain 75-W bulbs. So this question is equivalent to: What is the probability that no 75-W bulbs are selected if five bulbs are randomly selected?

Conditional Probabilities

- If we are given information that an event A has occurred, then the probability of another event B occurring may change.
- We use $P(A|B)$ to represent the conditional probability of A given that the event B has occurred. B is the “conditioning event”.
- Definition: for any two events A and B with $P(B)>0$, the conditional probability of A given B has occurred is defined by

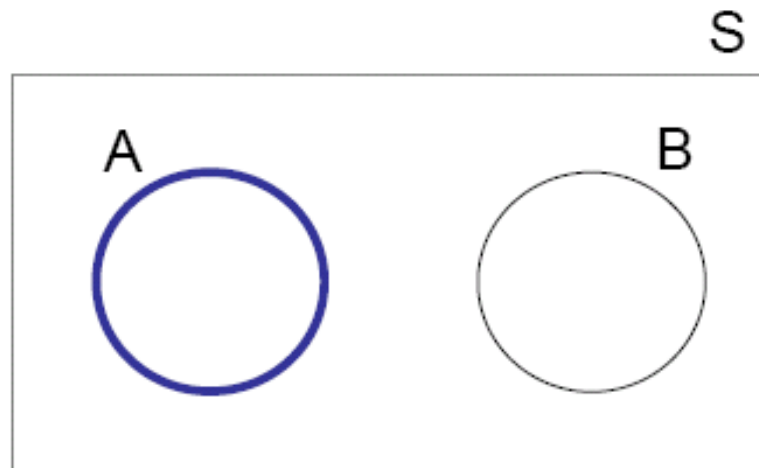
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Proportion of outcomes in B that are also in A.



Revisit disjoint events

- If A and B are disjoint then $P(A \cap B) = 0$.
- Hence, $P(A|B) = 0$. If B occurs, then A cannot occur.



Example

Ex. 1,000 voters were asked about their party affiliation and position on death penalty.

	Favor	Oppose	Total
Republican	260	40	300
Democrat	120	240	360
Other	240	100	340
Total	620	380	1000

R = A randomly chosen voter is a Republican.

D = A randomly chosen voter is a Democrat.

DP = A randomly chosen voter favors the death penalty.

What is the probability that a randomly chosen voter favors the death penalty given he is a Republican? ($26/30 = 0.867$)

What is the probability that a voter is a Democrat given he favors the death penalty? ($12/62 = 0.194$)

General Product Rule

- Conditional probability can be used to calculate the intersection of two events.

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

- Two events are **independent**, if the occurrence of B does not affect the probability of A.
- Two events A and B are **independent** if $P(A|B) = P(A)$ and are dependent otherwise.

Ex. (Death penalty example). $P(D|DP) = 0.194$ and $P(D) = 0.36$. Thus, DP and D are not independent.

- Independent events vs. disjoint events.

Example

Ex. A box contains 8 blue balls and 4 red balls. We draw two balls from the box **without replacement**. What is the probability that both are red?

A = first ball is red.

B = second ball is red.

$$\begin{aligned} P(\text{both balls are red}) &= P(A \cap B) = P(A) P(B|A) \\ &= 4/12 * 3/11 \\ &= 1/11 \end{aligned}$$

More general “**multiplication rule**”: $P(A \cap B \cap C) = P(C|A \cap B) P(B|A) P(A)$

Question: Draw three balls without replacement, what is the probability that all are red?

Tree diagram

- Often a problem involves several stages of using the **multiplication and addition rules**.
- A **tree diagram** is a useful way of keeping track of the multiplication and addition rules.
- It helps us think about conditional probabilities by showing the sequences of events as paths that resemble the branches of a tree.

Tree example

Ex. A company purchases 60% of its parts from vendor A and the rest from vendor B. Past experience indicates that 10% of vendor A's parts are defective, while 20% of vendor B's parts are defective.

A = part from vendor A

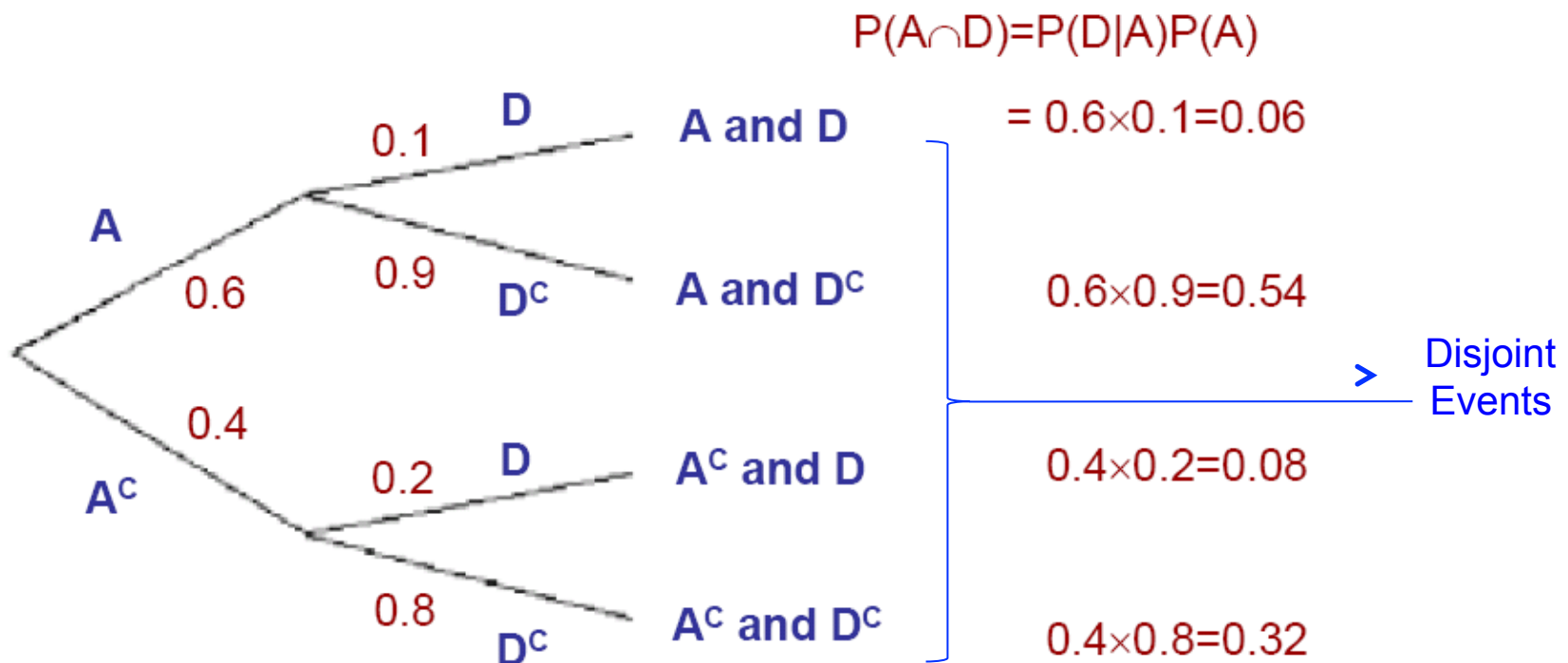
$$P(A) = 0.6$$

$$P(A^c) = 0.4$$

D = part defective

$$P(D|A) = 0.1$$

$$P(D|A^c) = 0.2$$



Example cont.

- A part is selected at random, what is the probability that it's defective?

$$\begin{aligned}P(D) &= P(D|A) P(A) + P(D|A^c) P(A^c) \\ &= 0.06 + 0.08 = 0.14\end{aligned}$$

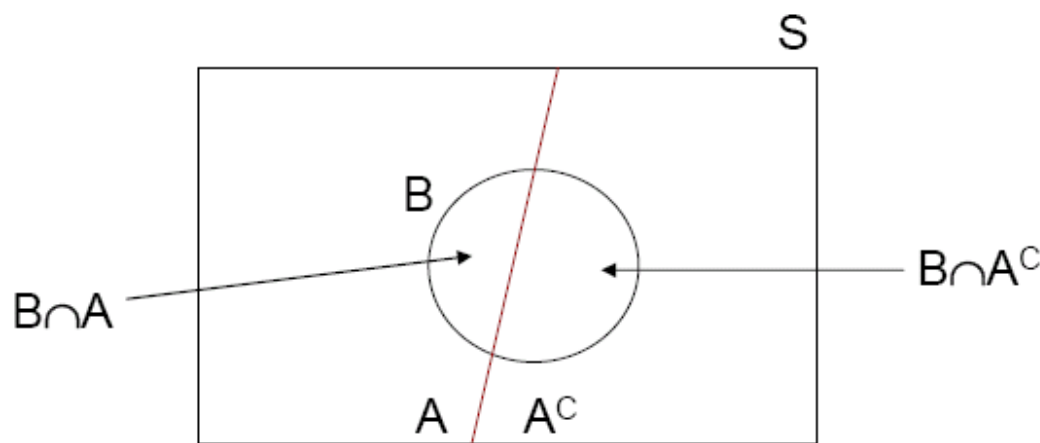
- Suppose a part is defective, what is the probability that it's from vendor A?

$$P(A|D) = \frac{P(A \cap D)}{P(D)} = \frac{0.06}{0.14} = \frac{3}{7}$$

- In these two examples, we used **rule of total probability** and **Bayes rule**.

Motivations

- We can write any event B as the union of the portion contained in another event A and the portion contained in A^c , i.e., $B = (B \cap A) \cup (B \cap A^c)$.



- Because A and A^c are disjoint, so are $B \cap A$ and $B \cap A^c$.

The Law of Total Probability

- The **Law of Total Probability** states, Let A_1, \dots, A_k be mutually exclusive and *exhaustive* events. Then for any other event B .

$$\begin{aligned} P(B) &= P(B|A_1)P(A_1) + \dots + P(B|A_k)P(A_k) \\ &= \sum P(B|A_i)P(A_i) \end{aligned}$$

- A_1, \dots, A_k are *exhaustive*, if one A_i must occur, so that $A_1 \cup \dots \cup A_k = S$.
- Proof: when $k=2$,

$$\begin{aligned} P(B) &= P((B \cap A) \cup (B \cap A^c)) \\ &= P(B \cap A) + P(B \cap A^c) \\ &= P(B|A)P(A) + P(B|A^c)P(A^c) \end{aligned}$$

Bayes Theorem

- With the help of the Law of Total Probability, we can state the Bayes Rule, which says, let A_1, \dots, A_k be a collection of k mutually exclusive and exhaustive events with *prior* probabilities $P(A_i)$ ($i=1, \dots, k$). Then for any other event B for which $P(B) > 0$, the *posterior* probability of A_j given that B has occurred is,

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^k P(B|A_i)P(A_i)}$$

- When $k=2$, we have,

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

- Bayes Rule can be used to “*reverse*” the probability from the conditional probability that was originally given, or *to find the cause given the result*.

Bayes Theorem Example

- ▶ One percent of all individuals in a certain population are carriers of a particular disease. A diagnostic test for this disease has a 90% detection rate for carriers and a 5% detection rate for noncarriers. If a person is tested positive, what's the probability that this person is a carrier?

Bayes Theorem Example

- ▶ One percent of all individuals in a certain population are carriers of a particular disease. A diagnostic test for this disease has a 90% detection rate for carriers and a 5% detection rate for noncarriers. If a person is tested positive, what's the probability that this person is a carrier?
- ▶ $P(\text{is a carrier} | \text{tested positive}) = \frac{P(\text{tested positive} | \text{is a carrier})P(\text{is a carrier})}{P(\text{positive} | \text{is a carrier})P(\text{is a carrier}) + P(\text{positive} | \text{is not a carrier})P(\text{is not a carrier})}$

Independence

- Two events A and B are **independent**, *if and only if*

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

- Recall that independence implies:

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

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

Therefore, by the multiplication rule for $P(A \cap B)$, we have

$$\begin{aligned} P(A \cap B) &= P(A|B) \cdot P(B) \\ &= P(B|A) \cdot P(A) \\ &= P(A) \cdot P(B) \end{aligned}$$

Multiple Events

- Events A_1, \dots, A_n are **mutually independent** if for every k ($k = 2, 3, \dots, n$) and every subset of indices i_1, i_2, \dots, i_k ,

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) P(A_{i_2}) \dots P(A_{i_k}).$$

- Independence is **very very important!**

Example

Ex. You recently bought a new set of tires from a manufacturer who just announced a recall because 2% of that particular brand were defective. What is the probability that at least one of your tires is defective? You may assume that the tires are defective independently of one another.

$$P(\text{at least one defective tire}) = 1 - P(\text{no defective tire})$$

Let A_i = tire i is not defective

$$P(A_i) = 1 - 0.02 = 0.98$$

$$\begin{aligned} P(\text{no defective tire}) &= P(A_1 \cap A_2 \cap A_3 \cap A_4) \\ &= P(A_1) P(A_2) P(A_3) P(A_4) = (0.98)^4 \end{aligned}$$

$$P(\text{at least one defective tire}) = 1 - (0.98)^4 = 0.0776$$