IIT ROORKEE

FREE ONLINE EDUCATION swayam शिक्षित भारत, उन्नत भारत

NPTEL

NPTEL ONLINE CERTIFICATION COURSE

Lecture 7: Introduction to Probability-II

Dr. A. Ramesh

Department of Management Studies

Problem

- A company data reveal that 155 employees worked one of four types of positions.
- Shown here again is the raw values matrix (also called a contingency table) with the frequency counts for each category and for subtotals and totals containing a breakdown of these employees by type of position and by sex.

Contingency Table

COMPANY HUMAN RESOURCE DATA

		Sex: Male	Sex: Female	
Type of Position	Managerial	8	3	11
	Professional	31	13	44
	Technical	52	17	69
	Clerical	9	22	31
		100	55	155

Solution

- If an employee of the company is selected randomly, what is the probability that the employee is female or a professional worker?

$$P(\mathrm{F} \cup \mathrm{P}) = P(\mathrm{F}) + P(\mathrm{P}) - P(\mathrm{F} \cap \mathrm{P})$$

$$P(\mathrm{F} \cup \mathrm{P}) = .355 + .284 - .084 = .555.$$

Problem

- Shown here are the raw values matrix and corresponding probability matrix for the results of a national survey of 200 executives who were asked to identify the geographic locale of their company and their company's industry type.
- The executives were only allowed to select one locale and one industry type.

RAW VALUES MATRIX

		Geographic Location				
		Northeast D	Southeast E	Midwest F	West G	
Industry Type	Finance A	24	10	8	14	56
	Manufacturing B	30	6	22	12	70
	Communications C	28	18	12	16	74
		82	34	42	42	200

Questions

a. What is the probability that the respondent is from the Midwest (F)?

b. What is the probability that the respondent is from the communications industry (C) or from the Northeast (D)?

c. What is the probability that the respondent is from the Southeast (E) or from the finance industry (A)?

PROBABILITY MATRIX

Industry Type		Geographic Location: Northeast D	Southeast E	Midwest F	West G	
	Finance A	.12	.05	.04	.07	.28
	Manufacturing B	.15	.03	.11	.06	.35
	Communications C	.14	.09	.06	.08	.37
		.41	.17	.21	.21	1.00

Mutually Exclusive Events

Type of Position	Gender: Male	Female	Total
Managerial	8	3	11
Professional	31	13	44
Technical	52	17	69
Clerical	9	22	31
Total	100	55	155

$$P(T \cup C) = P(T) + P(C)$$

$$= \frac{69}{155} + \frac{31}{155}$$

$$= .645$$

Mutually Exclusive Events

Type of Position	Gender: Male	Female	Total
Managerial	8	3	11
Professional	31	13	44
Technical	52	17	69
Clerical	9	22	31
Total	100	55	155

$$P(P \cup C) = P(P) + P(C)$$

$$= \frac{44}{155} + \frac{31}{155}$$

$$= .484$$

Law of Multiplication

$$\boxed{P(X \cap Y) = P(X) \cdot P(Y|X) = P(Y) \cdot P(X|Y)}$$

Problem

- A company has 140 employees, of which 30 are supervisors.
- Eighty of the employees are married, and 20% of the married employees are supervisors.
- If a company employee is randomly selected, what is the probability that the employee is married and is a supervisor?

		Married		
		Y	N	Sub total
Supervisor	Y	0.1143		30
	N			110
	Sub total	80	60	140

$$P(M) = \frac{80}{140} = 0.5714$$

$$P(S|M) = 0.20$$

$$P(M \cap S) = P(M) \cdot P(S|M)$$

$$= (0.5714)(0.20) = 0.1143$$

Law of Multiplication

Probability Matrix of Employees

Supervisor	Married: Yes	Married: No	Total
Yes	.1143	.1000	.2143
No	.4571	.3286	.7857
Total	.5714	.4286	1.00

$$P(\bar{S}) = 1 - P(S)$$
$$= 1 - 0.2143 = 0.7857$$
$$P(\bar{M} \cap \bar{S}) = P(\bar{S}) - P(M \cap \bar{S})$$
$$= 0.7857 - 0.4571 = 0.3286$$

$$P(M \cap \bar{S}) = P(M) - P(M \cap S)$$
$$= 0.5714 - 0.1143 = 0.4571$$
$$P(\bar{M} \cap S) = P(S) - P(M \cap S)$$
$$= 0.2143 - 0.1143 = 0.1000$$
$$P(\bar{M}) = 1 - P(M)$$
$$= 1 - 0.5714 = 0.4286$$

Special Law of Multiplication for Independent Events

- General Law

$$P(X \cap Y) = P(X) \cdot P(Y \mid X) = P(Y) \cdot P(X \mid Y)$$

- Special Law

If events X and Y are independent,

$P(X) = P(X \mid Y)$, and $P(Y) = P(Y \mid X)$.

Consequently,

$$P(X \cap Y) = P(X) \cdot P(Y)$$

Law of Conditional Probability

- The conditional probability of X given Y is the joint probability of X and Y divided by the marginal probability of Y.

$$\boxed{P(X|Y) = \frac{P(X \cap Y)}{P(Y)} = \frac{P(Y|X) \cdot P(X)}{P(Y)}}$$

Conditional Probability

- A conditional probability is the probability of one event, given that another event has occurred:

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

→ The conditional probability of A given that B has occurred

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

→ The conditional probability of B given that A has occurred

Where P(A and B) = joint probability of A and B

P(A) = marginal probability of A

P(B) = marginal probability of B

Computing Conditional Probability

- Of the cars on a used car lot, 70% have air conditioning (AC) and 40% have a CD player (CD). 20% of the cars have both.
- What is the probability that a car has a CD player, given that it has AC ?
- We want to find P(CD | AC).

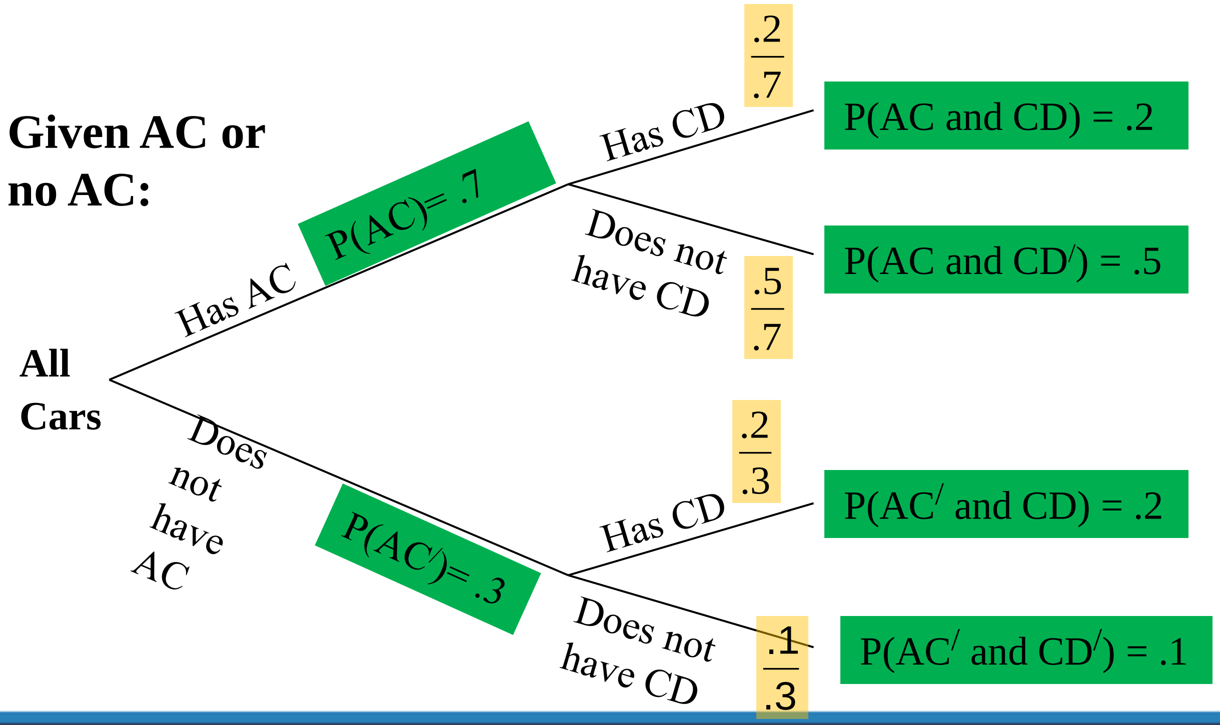
Computing Conditional Probability

	CD	No CD	Total
AC	0.2	0.5	0.7
No AC	0.2	0.1	0.3
Total	0.4	0.6	1.0

$$P(CD \mid AC) = \frac{P(CD \text{ and } AC)}{P(AC)} = \frac{.2}{.7} = .2857$$

Given AC, we only consider the top row (70% of the cars). Of these, 20% have a CD player. 20% of 70% is about 28.57%.

Computing Conditional Probability: Decision Trees



Independent Events

- If X and Y are independent events, the occurrence of Y does not affect the probability of X occurring.
- If X and Y are independent events, the occurrence of X does not affect the probability of Y occurring.

If X and Y are independent events,

$$P(X|Y) = P(X), \text{ and}$$

$$P(Y|X) = P(Y).$$

Statistical Independence

- Two events are **independent** if and only if:

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

- Events A and B are independent when the probability of one event is not affected by the other event

Independent Events Demonstration

	Geographic Location				
	Northeast D	Southeast E	Midwest F	West G	
Finance A	.12	.05	.04	.07	.28
Manufacturing B	.15	.03	.11	.06	.35
Communications C	.14	.09	.06	.08	.37
	.41	.17	.21	.21	1.00

Test the matrix for the 200 executive responses to determine whether industry type is independent of geographic location.

Independent Events Demonstration Contd...

$$P(A|G) = \frac{P(A \cap G)}{P(G)} = \frac{0.07}{0.21} = 0.33 \qquad P(A) = 0.28$$

$$P(A|G) = 0.33 \neq P(A) = 0.28$$

Independent Events

	D	E	
A	8	12	20
B	20	30	50
C	6	9	15
	34	51	85

$$P(A|D) = \frac{8}{34} = .2353$$

$$P(A) = \frac{20}{85} = .2353$$

$$P(A|D) = P(A) = 0.2353$$

Revision of Probabilities: Bayes' Rule

- An extension to the conditional law of probabilities
- Enables revision of original probabilities with new information

$$P(X_i|Y) = \frac{P(Y|X_i)P(X_i)}{P(Y|X_1)P(X_1) + P(Y|X_2)P(X_2) + \cdots P(Y|X_n)P(X_n)}$$

Problem

- A particular type of printer ribbon is produced by only two companies, Alamo Ribbon Company and South Jersey Products.
- Suppose Alamo produces 65% of the ribbons and that South Jersey produces 35%.
- Eight percent of the ribbons produced by Alamo are defective and 12% of the South Jersey ribbons are defective
- A customer purchases a new ribbon. What is the probability that Alamo produced the ribbon? What is the probability that South Jersey produced the ribbon?

Revision of Probabilities with Bayes' Rule: Ribbon Problem

$$P(Alamo) = 0.65$$

$$P(SouthJersey) = 0.35$$

$$P(d|Alamo) = 0.08$$

$$P(d|SouthJersey) = 0.12$$

$$P(Alamo|d) = \frac{P(d|Alamo)\cdot P(Alamo)}{P(d|Alamo)\cdot P(Alamo) + P(d|SouthJersey)\cdot P(SouthJersey)}$$

$$= \frac{(0.08)(0.65)}{(0.08)(0.65)+(0.12)(0.35)} = 0.553$$

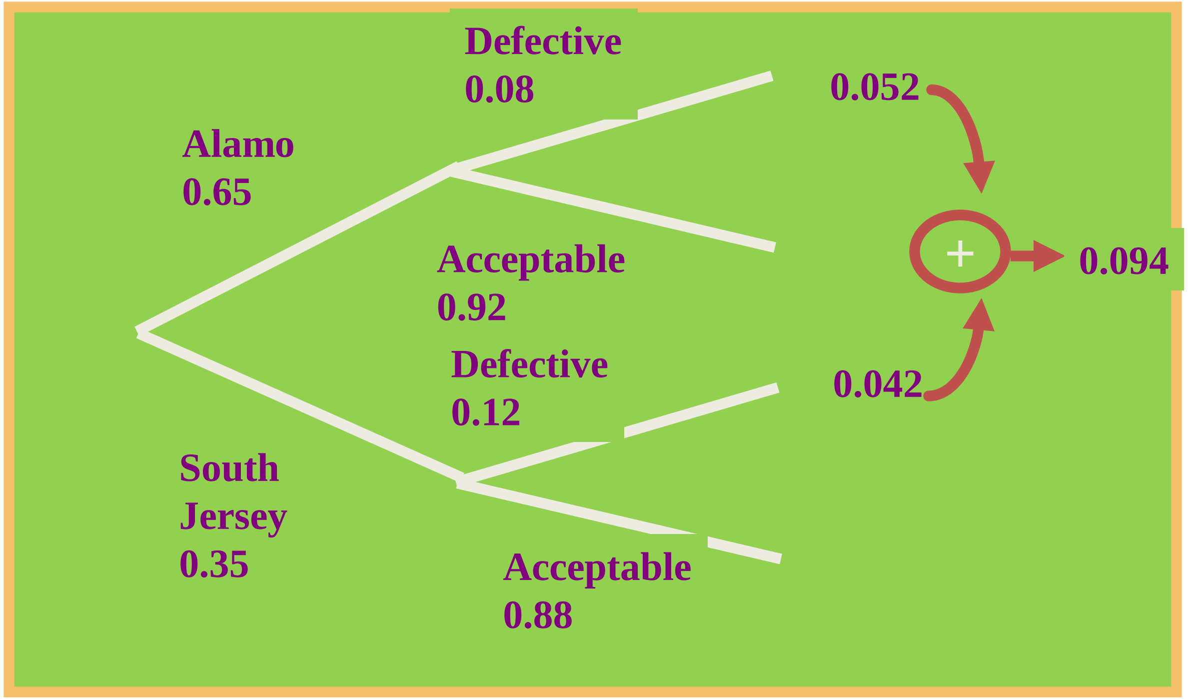
$$P(SouthJersey|d) = \frac{P(d|SouthJersey)\cdot P(SouthJersey)}{P(d|Alamo)\cdot P(Alamo) + P(d|SouthJersey)\cdot P(SouthJersey)}$$

$$= \frac{(0.12)(0.35)}{(0.08)(0.65)+(0.12)(0.35)} = 0.447$$

Revision of Probabilities with Bayes' Rule: Ribbon Problem

Event	Prior Probability $P(E_i)$	Conditional Probability $P(d\|E_i)$	Joint Probability $P(E_i \cap d)$	Revised Probability $P(E_i\|d)$
Alamo	0.65	0.08	0.052	$\frac{0.052}{0.094}$ =0.553
South Jersey	0.35	0.12	0.042 0.094	$\frac{0.042}{0.094}$ =0.447

Revision of Probabilities with Bayes' Rule: Ribbon Problem



THANK YOU