



Probability

Chetan Chitre

Why Calculate Probability?



- Suppose some legal concern has been expressed that a disproportionate number of managerial people at a company are men. If a worker is randomly selected from the company, what is the probability that the worker is a woman? If a managerial person is randomly selected, what is the probability that the person is a woman? What factors might enter into the apparent discrepancy between probabilities?
- What is the probability of default of a borrower?
- What is the probability that the sales will be higher by 10%, 20% and 25% in the next quarter?
- What is the probability that estimates from the sample data are true estimates?

Probability



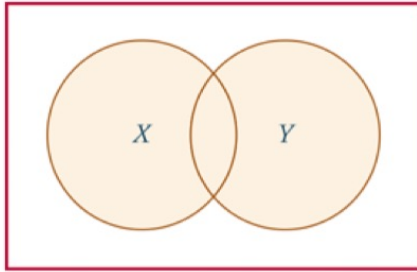
Number of times an event occurred

Total number of opportunities for the event to have occurred

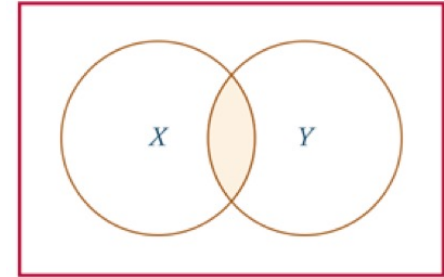
- Experiment: If I roll a pair of dice, what is the probability of getting a sum of 7?
- What is the “Total number of opportunities for the Event to have occurred? (Also known as the “Sample Space”)
- How many “Number of times the event has occurred”?



Set Notations



Union of Sets



Intersection of Sets

- **Mutually Exclusive Events: $P(X \cap Y) = 0$**
 - Probability that a random person in Bangalore works for any IT Company?
 - A manufactured part is either "defective" or "ok"

- **Independent Events: $P(X|Y) = P(X)$ and $P(Y|X) = P(Y)$**

Suppose a supervisor is randomly selecting bolts from a bin that contains 5% defects. If the supervisor samples a defective bolt and returns it to the bin, on the second draw there are still 5% defects in the bin regardless of the fact that the first outcome was a defect.

If the supervisor does not replace the first draw, the second draw is not independent of the first; in this case, fewer than 5% defects remain in the population. Thus, the probability of the second outcome is dependent on the first outcome.



Probability and Set Notations

- **Union Probability: $P(X \cup Y)$**

- Probability that a person owns a Sedan or a SUV.

- **Joint Probability: $P(X \cap Y)$**

- Probability that a person owns a Sedan and a SUV

- **Conditional Probability: $P(X|Y)$**

- Probability that given a person owns Sedan he/she also owns an SUV (prior knowledge of owning a sedan)



National Interiors conducted a survey for the Society of Interior Designers in which workers were asked which changes in office design would increase productivity. Respondents were allowed to answer more than one type of design change. The number one change that 70% of the workers said would increase productivity was reducing noise. In second place was more storage/filing space, selected by 67%. Suppose 56% of all respondents to the survey had said that *both* noise reduction *and* increased storage/filing space would improve productivity. If one of the survey respondents was randomly selected and asked what office design changes would increase worker productivity, what is the probability that this person would select reducing noise *or* more storage/filing space?

$$P(N \cup S) = P(N) + P(S) - P(N \cap S) = 0.7 + 0.67 - 0.56 = 0.81$$



Joint Probability Table

Storage Space

Noise Reduction		Yes	No	Total
	Yes	0.56		0.70
	No			
	Total	0.67		



Union Probability

The client company data from the Business Dilemma reveal that 155 employees worked one of four types of positions. Shown here again is the cross-tabulation table (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. If an employee of the company is selected randomly, what is the probability that the employee is female or a professional worker?

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



Union Probability

Shown here is the cross-tabulation table 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 allowed to select only one locale and one industry type.

	East	West	North	South	
Finance	24	10	8	14	56
Manufacturing	30	6	22	12	70
Communications	28	18	12	16	74
	82	34	42	42	200

- What is the probability that the respondent is from the North?
- What is the probability that the respondent is from the communications industry or from the East?
- What is the probability that the respondent is from the West or from the finance industry?



Joint Probability

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?

$$P (M \cap S) = P (M) \times P (S | M) = 0.57 \times 0.2 = 0.11$$



Joint Probability

Shown here are the cross-tabulation table 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 allowed to select only one locale and one industry type.

	East	West	North	South	
Finance	24	10	8	14	56
Manufacturing	30	6	22	12	70
Communications	28	18	12	16	74
	82	34	42	42	200

- What is the probability that the respondent is from Manufacturing and West?
- What is the probability that the respondent is from the South and works in Finance?
- What is the probability that the respondent is from the Manufacturing and Communication?



Independent Events

Shown here are the cross-tabulation table 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 allowed to select only one locale and one industry type.

	East	West	North	South	
Finance	24	10	8	14	56
Manufacturing	30	6	22	12	70
Communications	28	18	12	16	74
	82	34	42	42	200

- If X and Y are independent events: $P(X \cap Y) = P(X) \cdot P(Y)$
- Find the probability of an executive being from the East and works in Manufacturing. Is the location of the executive being from East independent of the executive being from Manufacturing?



Independent Events

- In a survey of students from section A and B, the number of students passing in Finance, Operations and Marketing is given in the table below:
- Is the student passing in Operations independent of student being from section A?

	A	B	
Finance	8	12	20
Operations	20	30	50
Marketing	6	9	15
	34	51	85



Conditional Probability

National Interiors conducted a survey for the Society of Interior Designers in which workers were asked which changes in office design would increase productivity. Respondents were allowed to answer more than one type of design change. The number one change that 70% of the workers said would increase productivity was reducing noise. In second place was more storage/filing space, selected by 67%. Suppose 56% of all respondents to the survey had said that *both* noise reduction *and* increased storage/filing space would improve productivity. A worker is selected randomly and asked about changes in office design. This worker believes that noise reduction would improve productivity. What is the probability that this worker believes increased storage space would improve productivity?

$$P(S|N) = P(N \cap S) / P(N) = 0.56 / 0.7 = 0.8$$



Conditional Probability

Shown here are the cross-tabulation table 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 allowed to select only one locale and one industry type.

	East	West	North	South	
Finance	24	10	8	14	56
Manufacturing	30	6	22	12	70
Communications	28	18	12	16	74
	82	34	42	42	200

- What is the probability that the executive works in Manufacturing given that he/she is from North?
- What is the probability that the executive is from the West given that he/she works in Communications?
- What is the probability that the executive is from the East given that he/she is from North?



Bayes' Rule

Use of conditional probabilities to allow revision of original probabilities with new information

Machines A, B, and C all produce the same two parts, X and Y. Of all the parts produced, machine A produces 60%, machine B produces 30%, and machine C produces 10%. In addition,

- 40% of the parts made by machine A are part X.
- 50% of the parts made by machine B are part X.
- 70% of the parts made by machine C are part X.

A part produced by this company is randomly sampled what is the probability that its is from A, B, or C?

With the knowledge that it is an X part, revise the probabilities that the part came from machine A, B, or C.



Bayes' Rule

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

	A	B	C	
X	0.24	0.15	0.07	0.46
Y	0.36	0.15	0.03	0.54
	0.60	0.30	0.10	100

	Prior P (M_i)	Conditional P (X M_i)	Joint P (X ∩ M_i)	Posterior P (M_i X)
A	0.6	0.4	0.6 x 0.4 = 0.24	0.24/0.46
B	0.3	0.5	0.15	0.15/0.46
C	0.10	0.70	0.07	0.07/0.46
			0.46	

Counting



Suppose a customer decides to buy a certain brand of new car. Options for the car include two different engines, five different paint colours, and three interior packages. If each of these options is available with each of the others, how many different cars could the customer choose from?

Counting Rule: $(m)(n) = (2)(5)(3) = 30$

Each time a die, which has six sides, is rolled, the outcomes are independent (with replacement) of the previous roll. If a die is rolled three times in succession, how many different outcomes can occur?

Sampling with replacement: $N^n = 6^3 = 216$

Suppose a small law firm has 16 employees and 3 are to be selected randomly to represent the company at the annual meeting of the Industry Association. How many different combinations of lawyers could be sent to the meeting?

Sampling without replacement: ${}_N C_n = N!/n!(N-n)!$ Where $N = 16$ and $n = 3 \rightarrow 16!/3!(16-3)! = 560$