MAT1372, Classwork7, Fall2025

3.1 Defining Probability(Conti.)

17. Consider rolling two dice (in red and in white). What are the probabilities of the sum of rolling?

Dice sum	All the combinations (*Red, White*)	Probability
2	(Red=1, white=1)	1/36
3	(1,2) (2,1)	2/36
4	(R=1, W=3) (R=2, W=2) (R=3, W=1)	3/36
5	(1,4) (2,3) (3,2) (4,1)	4/36
6	(1,5) (2,4) (3,3) (4,2) (5,1)	5/36
7	(1,6) (2,5) (3,4) (4,3) (5,2) (6,1)	6/36
8	(2,6) (3,5) (5,3) (6,2)	5/36
9	(3,6) (4,5) (5,4) (6,3)	4/36
10	(4,6) (5,5) (6,4)	3/36
11	(5,6) (6,5)	2/36
12	(6,6)	1/36

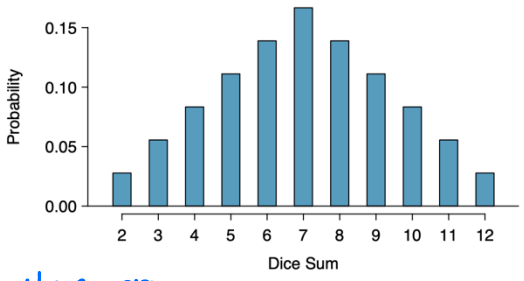
18. RULES FOR PROBABILITY DISTRIBUTIONS:

A probability distribution is a list of possible outcomes with corresponding probabilities that satisfies three rules:

(1) The outcomes listed must be disjoint.

(2) Each probability must be between 0 and 1

(3) The probabilities must total 1.



19. SAMPLE SPACE: All the possible events. For example, there are 11 sums of the two rolling dice,

20. COMPLEMENT:

Let D represent an event. Then the complement of D represents all outcomes in the sample space S that are not in D, which is denoted by D^c

21. Let D = {2, 3} represent the event that the outcome of a die roll is 2 or 3. (a) What is the sample space of a die rolling? (b) What is D^c? (c) What is $P(D^c)$? (d) What is $P(D) + P(D^c)$?

(a) $\{1,2,3,4,5,6\}$

(b) $D^c = \{1,4,5,6\}$

(c) $P(D^c) = P(\{1,4,5,6\}) = P(1)+P(4)+P(5)+P(6) = \frac{4}{6} = \frac{2}{3}$

(d) $P(D) = P(2 \text{ or } 3) = P(2)+P(3) = \frac{2}{6} = \frac{1}{3}$

$P(D)+P(D^c) = \frac{1}{3}+\frac{2}{3} = 1$

22. COMPLEMENT IN PROBABILITY

The complement of event A is denoted by A^c, and A^c represents all outcomes not in A. A and A^c are mathematically related

$P(A)+P(A^c) = 1$, $P(A^c) = 1 - P(A)$

3.2 Conditional Probability

1. The `photo_classify` data set represents a classifier a sample of 1822 photos from a photo sharing website and the given figure is a contingency table about two classifications with these photos.

		truth		
		fashion	not	Total
mach_learn	pred_fashion	197	22	219
	pred_not	112	1491	1603
	Total	309	1513	1822

Figure 3.11: Contingency table summarizing the `photo_classify` data set.

(1) If a photo is actually about fashion, what is the chance the ML classifier correctly identified the photo as being about fashion?

$P(\text{ML is pred fashion given truth is fashion}) = \frac{197}{309}$

(2) We sample a photo from the data set and learn the ML algorithm predicted this photo was not about fashion. What is the probability that it was incorrect and the photo is about fashion?

$P(\text{truth is fashion given ML is pre\_not fashion}) = \frac{112}{1603}$

(3) If a photo was about fashion which is predicted by the machine learning classifier, what is the probability that the photo is actually about fashion?

$P(\text{truth is fashion given ML is pre\_fashion}) = \frac{197}{219}$

2. MARGINAL AND JOINT PROBABILITIES:

If a probability is based on a single variable, it is a marginal probability. The probability of outcomes for two or more variables is called a joint probability.

3. From the table in 1.,

(1) what is the chance of the machine to classify that a photo is about fashion?

$P(\text{ML is pre\_fashion}) = \frac{219}{1822}$ which is a marginal probability.

(2) what is the chance of the machine to classify that a photo is about fashion and it is true?

$P(\text{ML is pre\_fashion and truth is fashion}) = \frac{197}{1822}$ which is a joint probability.

4. What can we observe from the results of questions 1.(3) and 3.(1), (2)?

$$\frac{\#\text{ cases where truth is fashion and ML is pre\_fashion}}{\#\text{ cases where ML is pre\_fashion}} = P(\text{truth is fashion given ML. pre\_fashion})$$

5. CONDITIONAL PROBABILITY:

The conditional probability of outcome A given outcome B is computed as the following

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

($|$: given)