Stat 343, Tue 8-Sep-2020 -- Tue 8-Sep-2020
Probability and Statistics
Fall 2020

Tuesday, September 8th 2020

Wk 2, Tu
Topic:: TBD
Topic:: Read FASt 2.2
Due::   PS02 at 6 pm

Warmup: How many different 8-pizza orders are possible from the choices of cheese, pepperoni, sausage, and veggie pizzas?

$$ {}_{11}C_3 = \binom{11}{3} = \frac{11!}{8!\,3!} = 165 $$

Notes:

- This can also be viewed as the problem of counting the nunmber of ways to partition the number 8 into sums of nonnegative integers, or

$$8 = c + p + s + v.$$

- It is as if we are sampling (making a random draw) of pizzas in which
  - order doesn't matter, only the count of cheese, etc. pizzas, and
  - we are sampling with replacement.

$n + k - 1$
$= 8 + 4 - 1 = 11$

c c | p | s s s | v v

11 slots

| | s s | v v v v v v

In summary: Counting samples of size k drawn from a set of size n:

	with replacement	without replacement
order matters	n^k	$\boxed{\frac{n!}{(n-k)!}}$
order doesn't matter	$\binom{n+k-1}{n}$	$\binom{n}{k}$

$k-1$?

$${}_nC_k = \frac{n!}{(n-k)!\,k!}$$

Note
$$\binom{n+k-1}{k-1} = \binom{n+k-1}{n}$$

$$= {}_nC_{n-k}$$

Conditional probability 2.2

Let S be the sample space for some random experiment and A, B be two events of interest. The **conditional probability** of A given B, denoted as $P(A \mid B)$, is defined as

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

Note: This definition holds up only when $P(B) \neq 0$. The concept is undefined if $P(B) = 0$.

Example 2.2.15. Assume that a randomly-selected child is equally-likely to be a girl or a boy. Given that a family with two children has a boy, what is the probability that the other child is a girl? Note that, given our assumption, the sample space consists of four equally likely outcomes:

$$GG, \quad \boxed{GB, \quad BG, \quad BB.}$$

Our events can be described by

$$A = \text{at least one girl}, \qquad B = \text{at least one boy}.$$

$$\boxed{P(A \mid B) = \frac{P(A \cap B)}{P(B)}} = \frac{2/4}{3/4} = \frac{2}{3}$$

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

Example 2.2.16. Suppose students in a grade school express preferences for color as follows:

	Blue	Other
Girls	7	9
Boys	11	8

Suppose a child is selected at random. Consider these events:

A: the child's favorite color is blue
B: the child is a boy
G: the child is a girl

Determine the values of

1. $P(A)$ $= 18/35$
2. $P(A^c)$ $= 1 - P(A) = 17/35$
3. $P(A \mid B)$ $= \frac{11}{19} = \frac{P(A \cap B)}{P(B)} = \frac{11/35}{19/35}$
4. $P(A^c \mid B)$ $= 8/19 = 1 - P(A \mid B)$
5. $P(B \mid A)$ $= 11/18$
6. $P(B \mid A^c)$ $= 8/17$
7. $P(A \mid G)$ $= 7/16$
8. $P(B \mid G)$ $= 0$

Lemma 1: If A, B are events with nonzero probabilities, then

$$P(A \cap B) \;=\; \boxed{P(A)P(B \mid A) \;=\; P(B)\boxed{P(A \mid B).}}$$

Theorem 1 (Bayes' Thm): If $P(A \mid B)$ and $P(B \mid A)$ are both defined, then

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

The idea that wouldn't die

Example 2.2.18. Find the probability of a flush.

A = get a card

B = 2^{nd} card is of same suit

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

$$= \left(\tfrac{12}{51}\right)(1)$$

$$P(\text{flush}) = (1)\left(\tfrac{12}{51}\right)\left(\tfrac{11}{50}\right)\left(\tfrac{10}{49}\right)\left(\tfrac{9}{48}\right)$$

Note the similarity to our approach (yesterday) for finding the probability that, among 20 people, no two have the same birthday.

Example 2.2.19. Find the probability of a large straight in Yahtzee. (in one roll)

A = none of 5 dice match each other

1, 2, 4, 5, 6 (3 left out)

3, 1, 2, 5 6 (4 left out)

neither is a straight

⋮

B = the number left out is a 1 or a 6

$$P(\text{straight}) = P(A \cap B) = P(B \mid A)\,P(A)$$

$$= \left(\frac{{}_6C_5}{6^5}\right)\left(\frac{2}{6}\right)$$

Example 2.2.22. Suppose a lab test correctly identifies diseased people 98% of the time and correctly identifies healthy people 99% of the time. Furthermore assume that in a certain population, one person in 1000 has the disease. If a random person is tested and the test comes back positive, what is the probability that the person has the disease?

Let events be labeled:

$$\begin{aligned} D &= \text{"a person has the disease"} \\ H &= \text{"a person is healthy"} \\ + &= \text{"the test comes back positive"} \\ - &= \text{"the test is negative"} \end{aligned}$$

Note $P(+ \mid D)$ is called the **sensitivity** of the test, while $P(- \mid H)$ is called the **specificity**.

Probability trees

Independent events

Suppose A, B are events with $P(B) \neq 0$. If

$$P(A) \;=\; P(A \,|\, B),$$

then we say these A and B are **independent**.

How many ways to choose 3 committee members from 11 people $_{11}C_3$

$=$ How many ways to choose 8 non-committee members from 11 people