AP Statistics Notes – Unit Five: Randomness and Probability

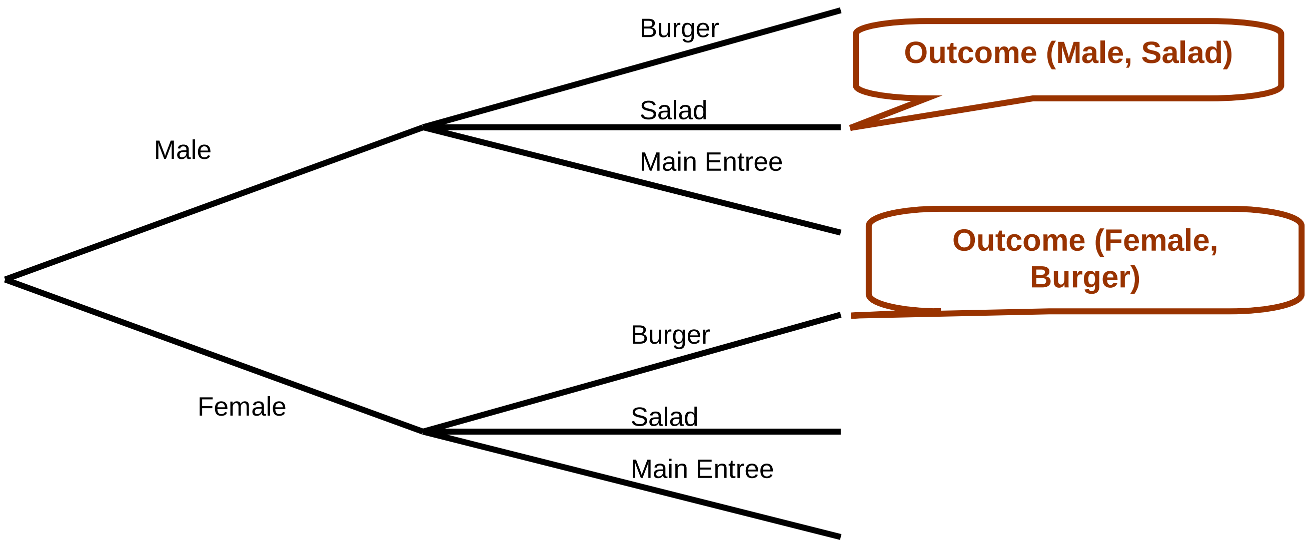
Syllabus Objectives: 3.1 – The student will interpret probability, including the long-term relative frequency distribution. 3.2 – The student will discuss the "Law of Large Numbers" concept.

This unit introduces you to the concept of chance behavior. Probability is the branch of mathematics that describes the pattern of chance outcomes. When we produce data by random sampling or randomized comparative experiments, probability helps us answer the question, "What would happen if we did this many times?" Probability calculations and an understanding of random behavior are the basis for inference, which will be studied in units 8 through 12.

- **The Idea of Probability**
  - **Random Phenomenon** – individual outcomes are uncertain, but there is a regular distribution of outcomes in a large number of repetitions.
  - **Probability** – proportion of times an event occurs in many repeated events of a random phenomenon. Probability can also be thought of as long-term relative frequency.
  - **Independent events (trials)** – outcome of an event (trial) does not influence the outcome of any other event (trial).

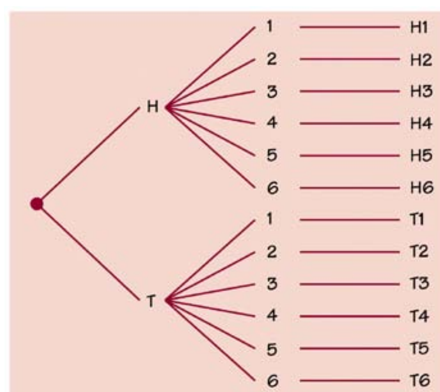
- **Probability Models**
  - **Chance experiment** – any activity or situation in which there is uncertainty about which of two or more possible outcomes will result.
  - **Sample Space** – the collection of all possible outcomes of a chance experiment. Sample spaces can be stated in different ways.
    - **List Ex:** A sample space can be a list of all the possible outcomes. Consider flipping a coin. There are two outcomes. The sample space would be: S = {H, T}. Consider rolling one die. The sample space can be represented as S= {1, 2, 3, 4, 5, 6}.
    - **List Ex:** An experiment is to be performed to study student preferences in the food line in the cafeteria. Specifically, the staff wants to analyze the effect of the student's gender on the preferred food line (burger, salad or main entrée). The sample space consists of the following possible outcomes: 1) A male choosing the burger line. 2) A female choosing the burger line. 3) A male choosing the salad line. 4) A female choosing the salad line. 5) A male choosing the main entrée line. 6) A female choosing the main entrée line. The sample could be represented by using set notation and ordered pairs. S = {(male, burger), (female, burger), (male, salad), (female, salad), (male, main entrée), (female, main entrée)}. This can be simplified to: S = {MB, FB, MS, FS, ME, FE}.

- **Tree diagram –** another way of illustrating a sample space using a picture.



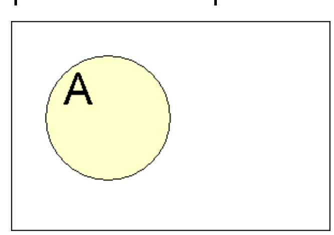
This "tree" has two sets of "branches" corresponding to the two bits of information gathered. To identify any particular outcome of the sample space, you traverse the tree by first selecting a bracket corresponding to gender and then a branch corresponding to the choice of food line.

- **Event** – any collection of outcomes from the sample space of a chance experiment.
- **Simple event** – an event consisting of exactly one outcome.
- If we look at the lunch line example, the event that the student selected is male is given by male = {MB, MS, ME}. The event that the preferred line is the burger line is given by burger = {MB, FB}. The event that the person selected is a female that prefers the salad line is {FS}. This is an example of a simple event.
- **Another tree diagram example:** In this experiment, first the person will flip a coin and then roll a die. The tree illustrates the sample space. There are 12 different outcomes. This leads to the following rule.
- **Multiplication Principle:** If you can do one task in n_1 number of ways and a second task in n_2 number of ways, then both tasks can be done in $n_1 \times n_2$ number of ways. In this example, there are 2 possible results for the coin toss and 6 possible results for the die, giving us 2×6 or 12 different outcomes.

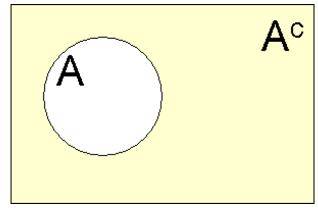


- **Venn Diagram** – an informal picture that is used to identify relationships. The collection of all possible outcomes of a chance experiment are represented as the interior of a rectangle.

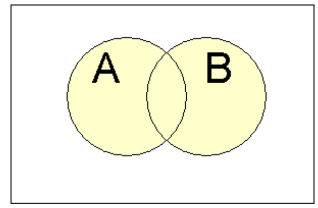
  The rectangle represents the sample space and shaded area represents the event A.



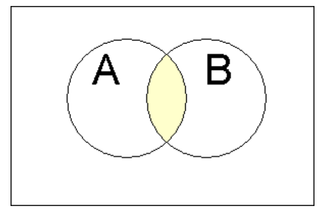
- **Complement** – Let A and B denote two events. The event not A consists of all experimental outcomes that are not in event A. Not A is called the **complement** of A and is denoted by A^c. The shaded area represents not A.



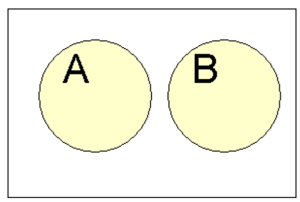
- **Union (or)** - Forming new events: Let A and B denote two events. The event A or B consists of all experimental outcomes that are in at least one of the two events – that is in A or in B or in BOTH of these. A or B is called the **union** of the two events and is denoted $A \cup B$. The shaded area represents $A \cup B$.



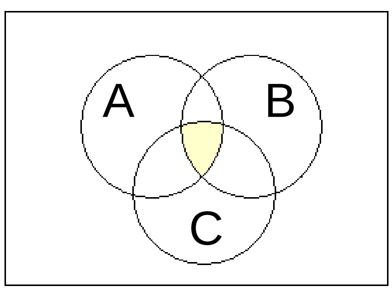
- **Intersection (and)** – Let A and B denote two events. The event A and B consists of all experimental outcomes that are in both of the events A and B. A and B is called the intersection of the two events and is denoted $A \cap B$. The shaded area represents $A \cap B$.



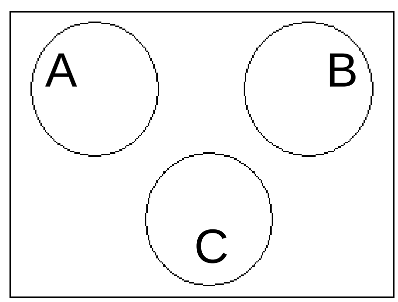
- **Disjoint or Mutually Exclusive** – Two events that have no common outcomes. The two events cannot happen simultaneously. Below, A and B are disjoint events.



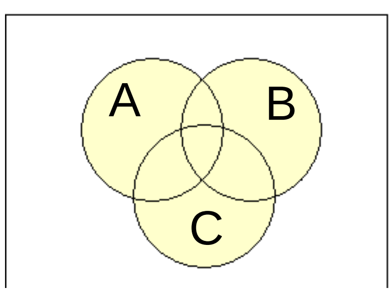
- **More Venn Diagrams** – illustrations showing more than two events.



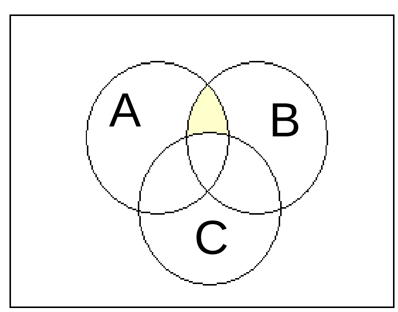
$$A \cap B \cap C$$



A, B & C are Disjoint

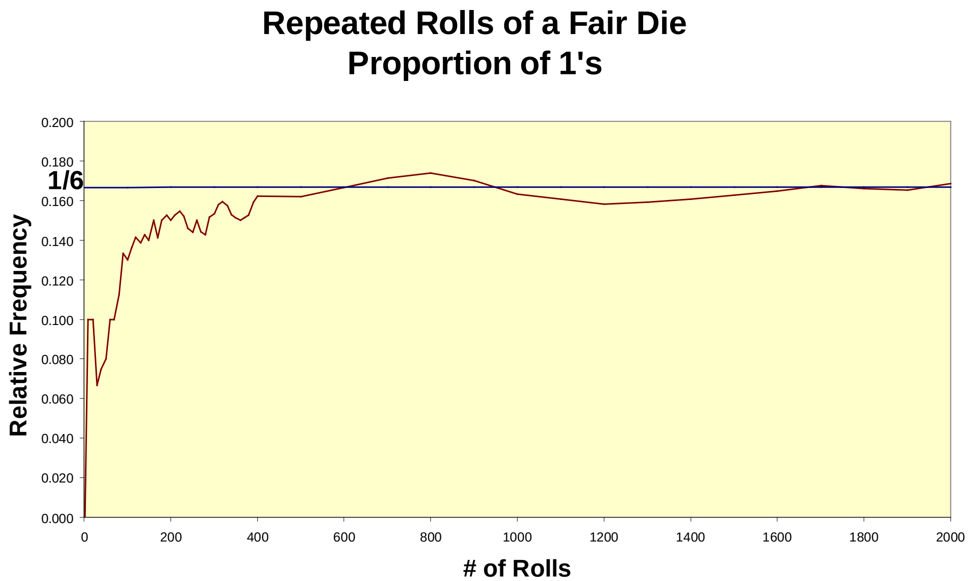


$$A \cup B \cup C$$



$$A \cap B \cap C^c$$

- **Probability**
  - For any event A, the **probability** is $P(A) = \dfrac{\text{number of times A occurs}}{\text{total number of trials}}$.
  - Empirical Approach: The probability of any outcome of a random phenomenon is the proportion of times the outcome would occur in a very long series of repetitions. That is, probability is long-term relative frequency.
    - Consider the chance experiment of rolling a "fair" die. We would like to investigate the probability of getting a "1" for the up face of the die. The die was rolled and after each roll the up face was recorded and then the proportion of times a 1 turned up was calculated and plotted.



This process could be repeated over and over again and the results would be similar. Notice that the proportion of 1's seems to stabilize and in the long run gets closer to the "theoretical" value of 1/6.

- In many "real-life" processes and chance experiments, the probability of a certain outcome or event is unknown, but nevertheless, this probability can be estimated reasonably well from observation. The justification is the **Law of Large Numbers.**

Law of Large Numbers: As the number of repetitions of a chance experiment increases, the chance that the relative frequency of occurrence for an event will differ from the true probability of the event by more than any very small number approaches zero.

Syllabus Objective: 3.7 – The student will create probability distributions by using simulation techniques.

- **Simulation** – The imitation of chance behavior, based on a model that accurately reflects the phenomenon under consideration, is called a **simulation**.
- Simulation steps:
    1. State the problem or describe the random phenomenon.
    2. State the assumptions.
    3. Assign digits to represent outcomes.
    4. Simulate many repetitions (trials).
    5. State your conclusions.
- Calculators and computers are particularly useful for conducting simulations because they can perform many repetitions quickly. Since this cannot be tested on the AP exam, we will not concentrate on this technique. We will simulate using the random digit table, as this can be tested on the exam. Of course coins, dice, cards and other objects can be used for simulation, but again, these cannot be tested.
- **Ex:** A couple plans to have children until they have a girl or until they have four children, whichever comes first. Use a random digit table to simulate this phenomenon.

  Step 1: Looking for a girl or four children.

  Step 2: Assume each child has probability of 0.5 of being a girl or boy and the sexes of successive children are independent.

  Step 3: Assign digits. One digit simulates the sex of one child. 0, 1, 2, 3, 4 = girl and 5, 6, 7, 8, 9 = boy. This gives each event (boy or girl) a 5/10 probability.

  Step 4: To simulate one repetition, read digits from the random digit table until the couple has either a girl or four children. The number of digits needed to simulate one repetition can vary from 1 to 4. We would want to simulate this many, many times. Below is 14 simulations using line 130 of the random digit table.

690	51	64	81	7871	74	0
BBG	BG	BG	BG	BBBG	BG	B
+	+	+	+	+	+	+
951	784	53	4	0	64	8987
BBG	BBG	BG	G	G	BG	BBBB
+	+	+	+	+	+	-

  In these 14 repetitions, a girl was born 13 times.

  Step 5: The estimate of the probability that this strategy will produce a girl is therefore, $\text{estimated probability } = \frac{13}{14} = 0.93$. This is actually very close to the actual probability which is 0.938.

Syllabus Objectives: 3.3 – The student will solve problems using the addition rule, multiplication rule, and conditional probability. 3.4 – The student will determine if two events are independent.

Formal Probability

- **The Basic Probability Rules**
  - The probability P(A) of an event A satisfies $0 \le P(A) \le 1$.
  - If S is the sample space for an experiment, $P(S) = 1$.
  - If two events A and B are disjoint meaning they can never occur simultaneously, then $P(A \text{ or } B) = P(A) + P(B)$

    $P(A \cup B) = P(A) + P(B)$.

    This is the **addition rule** for disjoint events.
    - This rule can be extended for more than two events. If events A, B and C are disjoint in the sense that no two events have any outcomes in common, then $P(A \text{ or } B \text{ or } C) = P(A) + P(B) + P(C)$

      $P(A \cup B \cup C) = P(A) + P(B) + P(C)$.

      This rule extends to any number of disjoint events.
  - For any event A, $P(A) + P(A^c) = 1$ so, $P(A^c) = 1 - P(A)$.

- **Equally likely outcomes** – If a random phenomenon has *k* outcomes that are all equally likely, then each individual outcome has probability, $1/k$. The probability of any event A is

$$P(A) = \frac{\text{count of outcomes in A}}{\text{count of outcomes in S}} = \frac{\text{count of outcomes in A}}{k}$$

  - **Ex:** Roll a fair die. Let A be the event of rolling a one. Let B be the event of rolling an odd number. S = {1, 2, 3, 4, 5, 6}. $P(A) = \frac{1}{6}$ and $P(B) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$. Since obtaining a 1, 3 or 5 (odd numbers) are disjoint, use the addition rule to find the probability.

  - **Ex:** Consider the experiment consisting of randomly picking a card from an ordinary deck of playing cards (52 card deck). Let A stand for the event that the card chosen is a King.

    The sample space is given by S = {A♠, K♠,…,2♠, A♥ , K ♥ , …, 2♥, A♦,…, 2♦, A♣,…, 2♣} and consists of 52 equally likely outcomes.

    The event is given by **A**={K♣, K♦, K♥, K♠} and consists of 4 outcomes, so $\text{P(A)} = \frac{4}{52} = \frac{1}{13} = 0.0769$

  - **Ex:** Consider the experiment consisting of rolling two fair dice and observing the sum of the up faces. Let E stand for the event that the sum is 11. The sample space is given by S = {(1, 1), (1, 2), (1, 3), …(6, 6)} and consists of 36 equally likely outcomes. The event E is given by E = {(5, 6), (6, 5)} and consists of 2 outcomes, so $P(E) = \frac{2}{36} = \frac{1}{18}$.

- **Ex:** Consider the experiment consisting of rolling two fair dice and observing the sum of the up faces. Let E stand for the event that the sum is 11 and Let F stand for the event that the sum is 7. F={(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)}. It has 6 outcomes so $P(F)=\frac{6}{36}$. Since E and F are disjoint events,

  $$P(E \cup F) = P(E) + P(F) = \frac{2}{36} + \frac{6}{36} = \frac{8}{36} = \frac{2}{9}$$

- **More Complex Probability Rules**
  - **Multiplication Rule for Independent Events**
    - Two events are **independent** if knowing that one occurs does not change the probability that the other occurs. Example: Flipping a coin two times. The coin does not have memory and each time you flip the coin, getting heads or tails is equally likely. Getting "head on the first flip" and "head on the second flip" are independent events. Example: If you roll two dice and obtain a sum of 7, the result of that roll has no effect on the next roll, so the two rolls are independent. But if you draw an ace from a deck of cards $P(\text{ace})=\frac{4}{52}$ and without replacing it draw a second card, the probability that the second card is also an ace is $P(\text{ace})=\frac{3}{51}$. These events are NOT independent.
    - If A and B are independent, then $P(A \text{ and } B) = P(A)P(B)$
      $P(A \cap B) = P(A)P(B)$
      This rule ONLY applies to independent events; you cannot use it if events are not independent.
    - Typically independence is assumed for the purposes of calculating probabilities if the samples are being drawn from a large population.
    - Disjoint events cannot be independent. If A and B are disjoint, then the fact that A occurs tells us that B cannot occur (see the Venn Diagram on p. 3). Unlike disjointness of complements, independence cannot be pictured by a Venn diagram, because it involves the probability of the events, rather than just the outcomes that make up the events.

    - **Ex:** We've determined that the probability that we encounter a green light at the corner of College and Main is 0.35, a yellow light is 0.04, and a red light 0.61. Let's think about your morning commute in the week ahead.
      Question: What is the probability you find the light red both Monday and Tuesday? Because the color of the light on Monday doesn't influence the color on Tuesday, these are independent events.

      $$P(\text{red Monday} \cap \text{red Tuesday}) = \text{P(red)} \times \text{P(red)}$$
      $$= (0.61)(0.61) = 0.3721$$

      Question: What's the probability you don't encounter a red light until Wednesday?

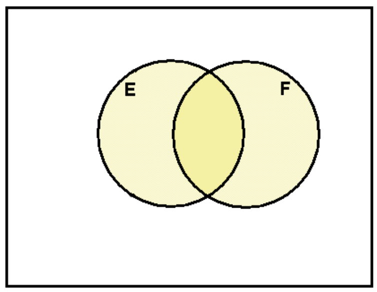
      $$P(\text{not red}) = 1 - P(\text{red}) = 1 - 0.61 = 0.39$$
      $$P(\text{not red Mon.} \cap \text{not red Tues.} \cap \text{red Wed.}) = (0.39)(0.39)(0.61) = 0.092781$$

      Question: What's the probability that you'll have to stop *at least once* during the week? Having to stop at least once means that I have to stop for the light either 1, 2, 3, 4, or 5 times next week. It's easier to think about the complement: never having to stop at a red light. Having to stop at least once means that I didn't make it through the week with NO red lights.

$$P(\text{having to stop at the light at least once in 5 days}) = 1 - P(\text{no red lights for 5 days})$$
$$= 1 - P(\text{not red} \cap \text{not red} \cap \text{not red} \cap \text{not red} \cap \text{not red})$$
$$= 1 - (0.39)(0.39)(0.39)(0.39)(0.39) = 1 - 0.0090 = 0.991$$

There's over a 99% chance I'll hit at least one red light sometime this week!

- **General Addition Rule**
  - If events E and F are not disjoint, they can occur simultaneously. The probability of their union is than *less* than the sum of their probabilities. The outcomes common to both (shown as the shaded region below) are counted twice when we add the probabilities, so we must subtract this probability once.

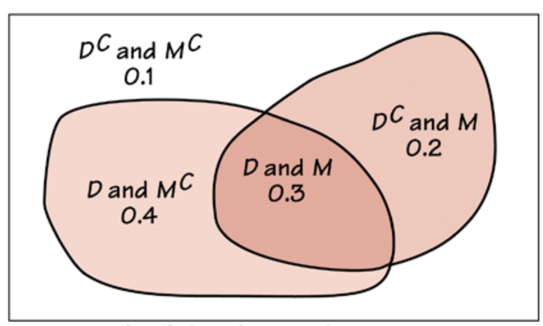


  - For any two events A and B, $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$
    $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - If A and B are disjoint, the event $A \cap B$ that both occur has no outcomes in it. The *empty event* $\varnothing$ is just zero, so the general addition rule just becomes the simple addition rule for disjoint events.
  - **Ex:** Deborah and Matthew are anxiously awaiting word on whether they have been made partners of their law firm. Deborah guess that her probability of making partner is 0.7 and that Matthew's is 0.5. Deborah also guesses that the probability that *both* she and Matthew are made partners is 0.3, then by the general addition rule,

$$P(\text{at least one is promoted}) = P(D \text{ or } M)$$
$$P(D \text{ or } M) = P(D) + P(M) - P(D \text{ and } M)$$
$$= 0.7 + 0.5 - 0.3 = 0.9$$

This can also be illustrated as a Venn diagram,

Note the intersection of the two events is 0.3 and placed in the center of the Venn diagram. We can now obtain the answer of 0.9 by adding the three regions 0.4, 0.3 and 0.2 together. Note also that 0.1 lies outside those events. The probability that *neither* is promoted is 0.1.



D = Deborah is made partner
M = Matthew is made partner

- **Ex:** A survey of job satisfaction of teachers was taken, giving the following results.

		Job Satisfaction		
		Satisfied	Unsatisfied	Total
LEVEL	College	0.095	0.055	0.150
	High School	0.288	0.220	0.508
	Elementary	0.162	0.180	0.658
	Total	0.545	0.455	1.000

If we pick a teacher at random, find the probability that they teach college (*C*) or they are satisfied with their job (*S*). We want to find $P(C \cup S)$. Since these events are NOT disjoint, use the general addition rule.

$P(C) = 0.150$, $P(S) = 0.545$ and $P(C \cap S) = 0.095$ from the table above, so

$$\begin{aligned} P(C \cup S) &= P(C) + P(S) - P(C \cap S) \\ &= 0.150 + 0.545 - 0.095 \\ &= 0.600 \end{aligned}$$

- **Conditional Probability**
  - Conditional probability gives the probability of one event under the condition that we know another event has occurred.
  - The new notation: $P(A \mid B)$. The bar is read "given the information that"
  - Let A and B be two events. The **conditional probability** of the event A given that the event B has occurred is $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$
  - A condition has the effect of reducing the size of the sample space, and therefore, the value of the denominator in a probability fraction. This can be seen in the formula above.

  - **Ex:** A study was performed to look at the relationship between motion sickness and seat positioning in a bus. The following table summarizes the data:

	Seat Position in Bus			
	Front	Middle	Back	Total
Nausea	58	166	193	417
No Nausea	870	1163	806	2839
Total	928	1329	999	3256

Let's use the symbols, N, N^c, F, M, B to stand for the events Nausea, No Nausea, Front, Middle and Back respectively. Computing the probability that an individual in the study gets nausea, we have, $P(N) = \dfrac{417}{3256} = 0.128$.

Other probabilities are easily calculated by dividing the numbers in the cells by 3256 to get

P(N and F)

P(N)

	Seat Position in Bus			
	Front	Middle	Back	Total
Nausea	0.018	0.051	0.059	0.128
No Nausea	0.267	0.357	0.248	0.872
Total	0.285	0.408	0.307	1.000

P(F)

P(M and N^C)

The event “a person got nausea given he/she sat in the front seat” is an example of what is called a conditional probability. Of the 928 people who sat in the front, 58 got nausea so the probability that “a person got nausea given he/she sat in the front seat” is $\frac{58}{928} = 0.0625$. This can also be found using the values on the table $P(N \mid F) = \frac{P(N \text{ and } F)}{P(F)} = \frac{0.018}{0.285} \approx 0.063$. It is not exactly 0.0625 due to roundoff error.

- **Ex:** Back to the survey of job satisfaction of teachers. Find the probability that a teacher chosen at random is a college teacher, given that he or she is satisfied.

		Job Satisfaction		
		Satisfied	Unsatisfied	Total
LEVEL	College	74	43	117
	High School	224	171	395
	Elementary	126	140	266
	Total	424	354	778

$P(S) = \frac{424}{778}$, $P(C \text{ and } S) = \frac{74}{778}$

$$P(C \mid S) = \frac{P(C \text{ and } S)}{P(S)} = \frac{74/778}{424/778} = \frac{74}{424} = 0.175$$

0.175 is the proportion of satisfied that are college teachers.
This is different than the proportion of teachers who are satisfied, given that they are college teachers. $P(S \mid C) = \frac{P(S \cap C)}{P(C)} = \frac{74/778}{117/778} = \frac{74}{117} = 0.632$

0.632 is the proportion of college teachers that are satisfied.

**Note: $P(C \mid S) \neq P(S \mid C)$. This is always true.

- **Ex:** One survey found that 56% of college students live on campus, 62% have a campus meal program, and 42% do both.
  Question: While dining in a campus facility open only to students with meal plans, you meet someone interesting. What is the probability that your new acquaintance lives on campus?
  P(student lives on campus given that the student has a meal plan)

  $$P(\text{on campus} \mid \text{meal plan}) = \frac{P(C \cap M)}{P(M)} = \frac{0.42}{0.62} = 0.677$$

  There is a probability of about 0.677 that a student with a meal plan lives on campus.

- **Ex:** A simple random sample of adults living in a suburb of a large city was selected. The age and annual income of each adult in the sample were recorded. The resulting data are summarized in the table below.

	Annual Income			
Age Category	\$25,000-\$35,000	\$35,001-\$50,000	Over \$50,000	Total
21-30	8	15	27	50
31-45	22	32	35	89
46-60	12	14	27	53
Over 60	5	3	7	15
Total	47	64	96	207

(a) What is the probability that a person chosen at random from those in this sample will be in the 31-45 age category?

$$P(31-45) = \frac{89}{207} \approx 0.43$$

(b) What is the probability that a person chosen at random from those in this sample whose incomes are over \$50,000 will be in the 31-45 age category? Show your work.

$$P(31-45 \mid over\$50,000) = \frac{P(31-45 \cap over\$50,000)}{P(over\$50,000)} = \frac{35/207}{96/207} \approx 0.365$$

(c) Based on your answers to parts (a) and (b), is annual income independent of age category for those in this sample? Explain.

No, these events are dependent. If these two events were independent, then $P(31-45) = P(31-45 \mid over\$50,000)$*, but* $0.43 \neq 0.365$.

(d) What is the probability that a person chosen at random from those in this sample will be in the 46-60 age category or have an income over $50,000?

Need to find the union of these two events. Since these events are NOT disjoint, we need to use the general addition rule.

$$P(46-60 \cup over\$50,000) = P(46-60) + P(over\$50,000) - P(both)$$

$$= \frac{53}{207} + \frac{96}{207} - \frac{27}{207} = \frac{122}{207} \approx 0.589$$

- **General Multiplication Rule for any two events**
  - The joint probability that events A and B both happen can be found by $P(A \text{ and } B) = P(A)P(B \mid A)$. Here, $P(B \mid A)$ is the conditional probability that B occurs, given the information that A occurs.
  - This general rule works for independent AND dependent events. If the two events are independent, then $P(B \mid A) = P(B)$. If A and B are not independent, then they are said to be **dependent** events.
  - **With replacement** – If we select an object and place it back in the sample space, it is said to be sampling with replacement. This is an example of **independent events**.
  - **Without replacement** – If we select an object and DO NOT replace it in the sample space and select another object, this is sampling without replacement. This is an example of **dependent events**.

  - **Ex:** Suppose we are going to select three cards from an ordinary deck of cards. Consider the events: *E* = {event that the first card is a king}, *F* = {event that the second card is a king}, *G* = {event that the third card is a king}. If we select the first card and then place it back in the deck before we select the second, and so on, the sampling will be with replacement.

$$P(E) = P(F) = P(G) = \frac{4}{52}$$

$$P(E \cap F \cap G) = P(E)P(F)P(G)$$

$$= \frac{4}{52}\frac{4}{52}\frac{4}{52} = 0.000455$$

If we select the cards in the usual manner without replacing them in the deck, the sampling will be without replacement.

$$P(E) = \frac{4}{52}, P(F) = \frac{3}{51}, P(G) = \frac{2}{50}$$

$$P(E \cap F \cap G) = P(E)P(F)P(G)$$

$$= \frac{4}{52}\frac{3}{51}\frac{2}{50} = 0.000181$$

*Note: We are still applying the multiplication rule as we did with independent events to find our probability, it's just that with dependent events, we have to "think" about the 2^{nd} and 3^{rd} fraction that we are multiplying, now that the previous events have occurred.

- **Ex:** Police report that 78% of drivers are given a breath test, 36% a blood test, and 22% both tests.
  Define events: *A* = {suspect is given a breath test} *B* = {suspect is given a blood test}. It is known: $P(A) = 0.78, P(B) = 0.36, P(A \cap B) = 0.22$

  Question: Are giving a DWI suspect a blood test and a breath test mutually exclusive? No. Disjoint events cannot BOTH happen at the same time, so check to see if $P(A \cap B) = \varnothing$. Since some suspects are given both tests and the probability is equal to 0.22, they are not mutually exclusive.

  Question: Are the two tests independent?
  Let's make a table.

Breath Test

Blood Test		Yes	No	**Total**
	Yes	0.22	0.14	**0.36**
	No	0.56	0.08	**0.64**
	Total	**0.78**	**0.22**	**1.00**

Does getting a breath test change the probability of getting a blood test? That is, does $P(B \mid A) = P(B)$?

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{0.22}{0.78} \approx 0.28$$

$$P(B) = 0.36 \text{ since } 0.28 \neq 0.36$$

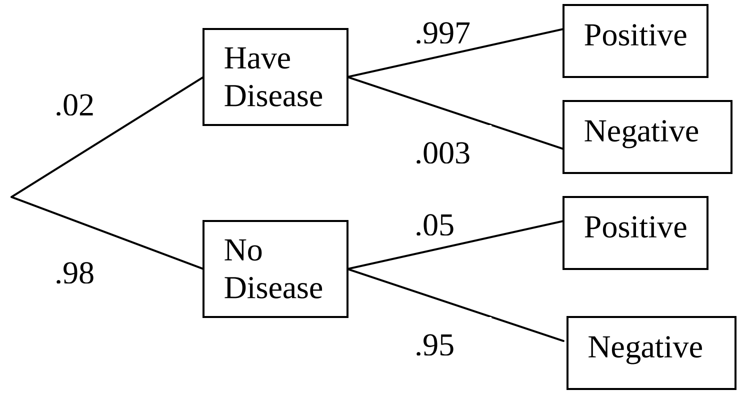
$$P(B \mid A) \neq P(B)$$

Overall, 36% of the drivers get blood tests, but only 28% of those who get a breath test do. Since suspects who get a breath test are less likely to have a blood test, the two events are not independent.

- **Extended Multiplication Rule**
  - Extend the multiplication rule to the probability that all of several events occur. The key is to condition each event on the occurrence of ALL of the preceding events.
  - $P(A \text{ and } B \text{ and } C) = P(A)P(B \mid A)P(C \mid A \text{ and } B)$
  - Best way to figure these types of problems is to use a tree diagram. The tree diagram keeps straight what has occurred and the probabilities of each occurring. Tree diagrams combine the addition and multiplication rules. The multiplication rule says that the probability of reaching the end of any complete branch is the product of the probabilities written on its segments.

  - **Ex:** A laboratory test for the detection of a certain disease gives a positive result 5 percent of the time for people who do not have the disease. The test gives a negative result 0.3 percent of the time for people who have the disease. Large-scale studies have shown that the disease occurs in about 2 percent of the population.

    (a) What is the probability that a person selected at random would test positive for this disease? Show your work.



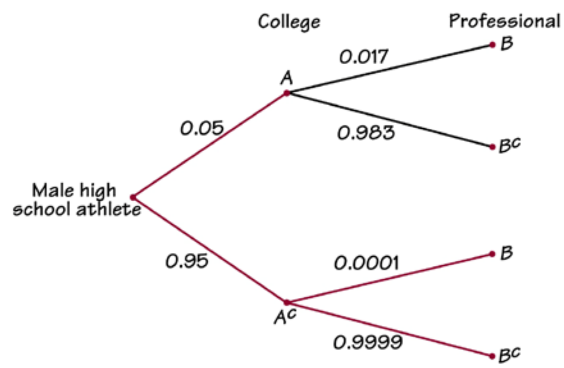
$$P(positive) = (.02)(.997) + (.98)(.05) = 0.06894$$

(b) What is the probability that a person selected at random who tests positive for the disease does not have the disease? Show your work.

Find P(No disease | tests positive)

$$= \frac{P(NoDisease \cap test+)}{P(+)} = \frac{(.98)(.05)}{.06894} = \frac{.049}{.06894} \approx 0.711$$

- **Ex:** Only 5% of male high school basketball, baseball and football players go on to play at the college level. Of these, only 1.7% enter major league professional sports. Of those athletes that do not compete in college, only 0.01% enter professional sports. Below is a tree diagram illustrating these percents.



Question: What is the probability that a male high school athlete will go on to professional sports? To find P(B), we multiply the branches along the two that reach B and add them together.

$$P(B) = (0.05)(0.017) + (0.95)(0.0001) = 0.00085 + 0.000095 = 0.000945$$

About 9 high school athletes out of 10,000 will play professional sports.

Question: What proportion of professional athletes competed in college? This is a conditional probability. It is the conditional probability $P(A \mid B)$ where *B* = {competes professionally} and *A* = {competes in college}.

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{.00085}{.000945} = 0.8995$$

Almost 90% of professional athletes competed in college.