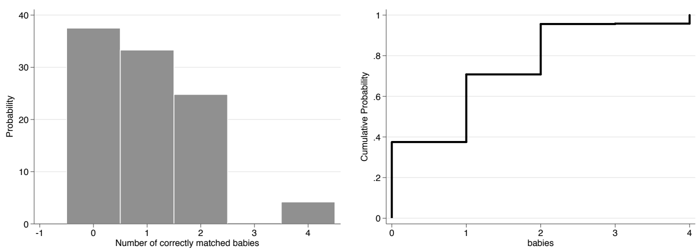
Activity #4: Discrete Random Variables, Conditional Probability, and Bayes' Theorem

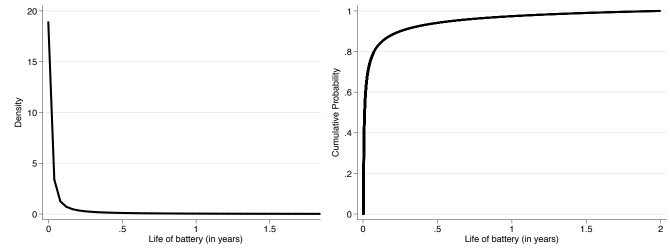
In Activity 3, we found probability models for the number of babies we correctly returned at random to 4 mothers:

The number of babies correctly matched with their mothers.

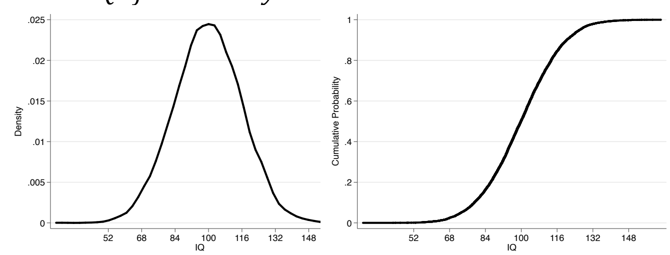
Matched Babies	Probability
0	0.375
1	0.333
2	0.250
3	0.000
4	0.042



The lifespan of a randomly selected battery (in years).



The IQ of a randomly selected American adult.



In scenarios like returning babies, we can easily count and list all possible outcomes (or values the variable can take).
We say, therefore, that these variables are discrete random variables.

Discrete Random Variable: A variable that can take a finite (or countably infinite) number of possible values

Probability models for discrete variables can be displayed by graphing probability mass functions or cumulative distributions.

Probability Mass Function: Displays the probability of each possible outcome , P(X = x)
Properties: a) For each outcome, 0 ≤ P(X=x) ≤ 1
b) The sum of all probabilities in the sample space is 1.00

Cumulative function: Displays the probability of each outcome and all preceding outcomes, P(X ≤ x)
Properties: a) It is monotonically increasing
b) It always starts at 0 and ends at 1.0

For other variables, such as those displayed on the top-right of the page, it's impossible to count and list all possible outcomes. These variables are continuous random variables.

Continuous Random Variable: A variable that can take on an infinite number of possible values within a region

We can display probability models for continuous variables through probability density functions or cumulative distributions.

1) Classify each of the following variables as either **discrete** or **continuous**:

________ Number of children in a randomly chosen family ________ Age (in years) of a randomly selected SAU student

________ A random person's height (rounded to the nearest inch) ________ Number of grains of sand on the beach

Over the next several weeks, we'll work with discrete (and then continuous) variables in an attempt to:

a) Display a probability model (via a graph or formula)
b) Calculate the expected value and variance
c) Calculate probabilities of interest

To do that, we need to learn a few more probability rules and formulas.

2) Suppose 80% of students applying to St. Ambrose are accepted. If we randomly select a student who applies to SAU next year, what's our best estimate of the probability that the student is accepted? Suppose we find out the student earned straight As in high school and scored a 36 on the ACT. Would that change your estimate for the probability that the student is accepted?

3) Suppose you roll two 6-sided dice. What's the probability that the first die shows a 6?

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

4) Suppose you roll two 6-sided dice and find their sum is 7. What's the probability that the first die shows a 6?

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

5) Suppose you roll two 6-sided dice and find their sum is 5. What's the probability that the first die shows a 6? Why does this answer differ from the previous question?

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

6) Suppose you roll two 6-sided dice and find their sum is 5. What's the probability that the first die shows a 2?

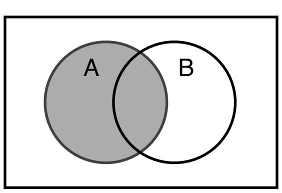
11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

The **conditional probability** of an event is the revised probability we estimate when we have additional information

If we are given (or we assume) information about event B, then the revised probability of event A is written as:

$$P(A|B) \text{ = the probability of A given B}$$

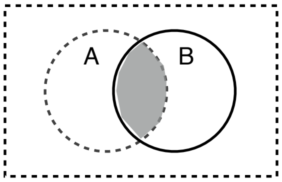
Suppose we're interested in the probability of event A occurring. In a Venn Diagram, we would display this as:



An estimate of this probability, $P(A) = \dfrac{\text{\# of outcomes in A}}{\text{\# of possible outcomes in the sample space}}$

Now suppose we know event B has happened. Given event B, what's the conditional probability of event A?

Since we already know event B has happened, that limits our sample space (the total number of possible outcomes).



An estimate of this probability, P(A | B) = $\dfrac{\text{\# of outcomes in A and B}}{\text{\# of possible outcomes in event B}}$

> The **conditional probability** of event A, given event B has occurred, is estimated by: $P\left(A\middle|B\right) = \dfrac{P\left(A \cap B\right)}{P\left(B\right)}$

Note that we still do not have a formula to calculate the numerator of this formula

7) In Europe, 88% of all households have a television. 51% of all households have a television and a VCR. What's the probability that a household with a television has a VCR?

	TV	No TV	Total
VCR			
No VCR			
Total			100

8) According to flightstats.com, 86.55% of flights out of Salt Lake City International Airport departed on-time in 2011. 72.68% of flights departed on-time and arrived on-time. Given a flight departs on-time, what's the probability it will arrive on-time? Use the formula to solve this problem and check your answer with relative frequencies.

	Arrive on	Arrive off	Total
Depart on			
Depart off			
Total			100

> The **conditional probability** of event A, given event B has occurred, is estimated by: $P\left(A\middle|B\right) = \dfrac{P\left(A \cap B\right)}{P\left(B\right)}$
>
> We can manipulate this equation to get the **General Multiplication Rule**: $P\left(A \cap B\right) = P\left(A\middle|B\right)P\left(B\right)$
>
> We can also note that: $P\left(A \cap B\right) = P\left(A\middle|B\right)P\left(B\right) = P\left(B \cap A\right) = P\left(B\middle|A\right)P\left(A\right)$

9) A new test is developed to detect a disease that inflicts 10% of the population. The test has a 97% chance of correctly identifying the disease in an infected individual. Suppose you take this test and you do not know if you are infected. What's the probability that you are infected and the test will correctly detect your infection?

			Total
Total			100

10) One hundred adults were surveyed to determine their opinions regarding loan debt burdens of college students. The proportion of responses, separated by adults with and without children in college, are displayed below:

	Too much loan debt	About right	Too little loan debt	Total
Has child in college	0.20	0.09	0.01	0.30
Does not have child in college	0.41	0.21	0.08	0.70
Total	0.61	0.30	0.09	1.00

Estimate the following probabilities:

a) P(does not have child in college **|** too little loan debt) = ______________________________

b) P(too little loan debt **|** does not have child in college) = ______________________________

c) P(does not have child in college AND too little loan debt) = ______________________________

d) P(does not have child in college AND too little loan debt)' = ______________________________

e) P(does not have child in college OR too little loan debt) = ______________________________

f) P(does not have child in college OR too little loan debt)' = ______________________________

11) From a Life Table, one finds that 89.835% of females can expect to live to age 60, while 57.062% can expect to live to age 80. Given that a woman has reached the age of 60, what's the probability that she will live to age 80? (Hint: Start by pretending you have a population of 100,000 women and determine how many are expected to live to each age).

			Total
Total			100