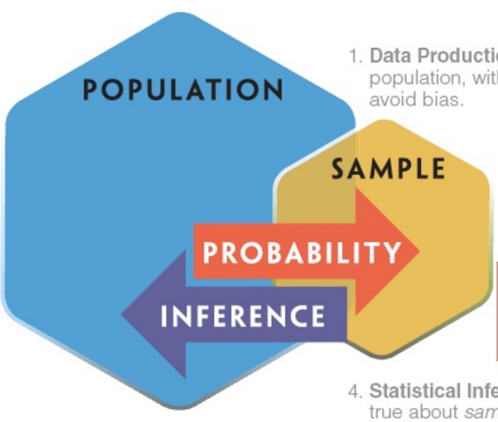
Lecture 9: Chapter 6, Sections 1-2
Finding Probabilities:
Beginning Rules

- Probability: Definition and Notation
- Basic Rules
- Independence; Sampling With Replacement
- General "Or" Rule; General "And" Rule
- More about Conditional Prob; 2 Types of Error

Looking Back: *Review*

- **4 Stages of Statistics**
  - Data Production (discussed in Lectures 1-3)
  - Displaying and Summarizing (Lectures 3-8)
  - Probability
    - Finding Probabilities
    - Random Variables
    - Sampling Distributions
  - Statistical Inference

Four Processes of Statistics



1. **Data Production:** Take sample data from the population, with sampling and study designs that avoid bias.
2. **Displaying and Summarizing:** Use appropriate displays and summaries of the sample data, according to variable types and roles.
3. **Probability:** Assume we know what's true for the *population;* how should **random** *samples* behave?
4. **Statistical Inference:** Assume we only know what's true about *sampled* values of a single variable or relationship; what can we **infer** about the larger *population?*

Definitions

- **Statistics**
  - Science of producing, summarizing, drawing conclusions from data
  - Summaries about sample data
- **Probability**
  - Science dealing with random behavior
  - Chance of happening

Example: *Ways to Determine a Probability*

- **Background**: Some probability statements:
  - Probability of randomly chosen card a heart is 0.25
  - Probability of randomly chosen student in a class getting A is 0.25, according to the professor.
  - Probability of candidate being elected, according to an editorial, is 0.25.
- **Question:** Are these probabilities all determined in the same way?
- **Response:**

Definition

- **Probability:** chance of an event occurring, determined as the
  - Proportion of equally likely outcomes comprising the event; or
  - Proportion of outcomes observed in the long run that comprised the event; or
  - Likelihood of occurring, assessed subjectively.

Example: *Three Ways to Determine a Probability*

- **Background:** Probabilities can be determined as a
  - Proportion of equally likely outcomes; or
  - Proportion of long-run outcomes observed; or
  - Subjective likelihood of occurring.
- **Question**: How was each of these determined?
  1. Probability of randomly chosen card a heart is 0.25.
  2. Probability of randomly chosen student in a class getting A is 0.25, according to the professor.
  3. Probability of candidate being elected, according to an editorial, is 0.25.
- **Response:**
  1. **Card a heart?**
  2. **Student gets an A?**
  3. **Candidate elected?**

Looking Ahead: *Probabilities can be based on opinions, as long as we obey Rules.*

Practice: 6.1 p.236

Notation

Use capital letters to denote events in probability: H=event of getting a heart.

Use "P( )" to denote probability of event: P(H)= probability of getting a heart.

Use "not A" to denote the event that an event A does not occur.

Basic Probability Rules

To stress how intuitive the basic rules are, we

- Begin with example for which we can intuit the solution.
- State general rule based on solution.
- Apply rule to solve a second example.

Looking Ahead: *This process will be used to establish all the rules needed to understand behavior of* ***random variables*** *in general,* ***sampling distributions*** *in particular, so we have theory needed to perform* ***inference****.*

Example: *Intuiting Permissible Probabilities Rule*

- **Background**: A six-sided die is rolled once.
- **Questions:** What is the probability of getting a nine? What is the probability of getting a number less than nine?
- **Responses:** P(N)=__________________

  P(L)=__________________

Permissible Probabilities Rule

The probability of an impossible event is 0, the probability of a certain event is 1, and all probabilities must be between 0 and 1.

Example: *Applying Permissible Probabilities Rule*

- **Background**: Consider the values -1, -0.1, 0.1, 10.
- **Question:** Which of these are legitimate probabilities?
- **Response:**

Practice: 6.3 p.237

Example: *Intuiting Sum-to-One Rule*

- **Background**: Students' year is classified as being 1st, 2nd, 3rd, 4th, or Other.
- **Question:** What do we get if we sum the probabilities of a randomly chosen student's year being 1st, 2nd, 3rd, 4th, and Other?
- **Response:**

Sum-to-One Rule

The sum of probabilities of all possible outcomes in a random process must be 1.

Example: *Applying Sum-to-One Rule*

- **Background**: A survey allows for three possible responses: yes, no, or unsure. We let P(Y), P(N), and P(U) denote the probabilities of a randomly chosen respondent answering yes, no, and unsure, respectively.
- **Question:** What must be true about the probabilities P(Y), P(N), and P(U)?
- **Response:**

Practice: 6.5 p.237

Example: *Intuiting "Not" Rule*

- **Background**: A statistics professor reports that the probability of a randomly chosen student getting an A is 0.25.
- **Question:** What is the probability of *not* getting an A?
- **Response:**

Looking Back: *Alternatively, since A and not A are the only possibilities, according to the Sum-to-One Rule, we must have 0.25+P(not A)=1, so P(not A)=1-0.25=0.75.*

“Not” Rule

For any event A, P(not A)=1-P(A).
Or, we can write P(A)=1-P(not A).

Example: *Applying “Not” Rule*

- **Background**: The probability of a randomly chosen American household owning at least one TV set is 0.96.
- **Question:** What is the probability of not owning any TV set?
- **Response:** P(not TV)=

Practice: 6.6 p.237

Example: *Intuiting Non-Overlapping “Or” Rule*

- **Background**: A statistics professor reports that the probability of a randomly chosen student in her class getting an A is 0.25, and the probability of getting a B is 0.30.
- **Question:** What is the probability of getting an A or a B?
- **Response:**

Example: *When Probabilities Can’t Simply be Added*

- **Background**: A statistics professor reports that the probability of a randomly chosen student in her class getting an A is P(A)=0.25, and the probability of being a female is P(F)=0.60.
- **Question:** What is the probability of getting an A or being a female?
- **Response:**

Practice: 6.12 p.237

Definition; Non-Overlapping "Or" Rule

For some pairs of events, if one occurs, the other cannot, and vice versa. We can say they are **non-overlapping**, the same as **disjoint** or **mutually exclusive**.

For any two non-overlapping events A and B, P(A or B)=P(A)+P(B).

Note 1: *Events "female" and "getting an A"* ***do*** *overlap →Rule does* ***not*** *apply.*

Note 2: *The word "or" entails addition.*

Example: *Applying Non-Overlapping "Or" Rule*

- **Background**: Assuming adult male foot lengths have mean 11 and standard deviation 1.5, if we randomly sample 100 adult males, the probability of their sample mean being less than 10.7 is 0.025; probability of being greater than 11.3 is also 0.025.
- **Question:** What is the probability of sample mean foot length being less than 10.7 or greater than 11.3?
- **Response:**

Practice: 6.30 p.258

Example: *Intuiting Independent "And" Rule*

- **Background**: A balanced coin is tossed twice.
- **Question:** What is the probability of both the first and the second toss resulting in tails?
- **Response:**

Looking Back: *Alternatively, since there are 4 equally likely outcomes HH, HT, TH, TT, we know each has probability ___________.*

Example: *When Probabilities Can't Simply be Multiplied*

- **Background**: In a child's pocket are 2 quarters and 2 nickels. He randomly picks a coin, does *not* replace it, and picks another.
- **Question:** What is the probability of the first and the second coins both being quarters?
- **Response:**

Practice: 6.10c p.237

Definitions

For some pairs of events, whether or not one occurs impacts the probability of the other occurring, and vice versa: the events are said to be **dependent**.

If two events are **independent**, whether or not one occurs has no effect on the probability of the other occurring.

Independent "And" Rule; Replacement?

For any two independent events A and B,

P(A and B)=P(A) × P(B).

Note: The word "and" entails multiplication.

- Sampling *with replacement* is associated with events being *independent*.
- Sampling *without replacement* is associated with events being *dependent*.

Example: *Applying Independent "And" Rule*

- **Background**: We're interested in the probability of getting a female and then a male…
- **Questions:** What is the probability if we pick…
  1. 2 people with replacement from a household where 3 of 5 (that is, 0.6) are female?
  2. 2 people without replacement from household where 3 of 5 (that is, 0.6) are female?
  3. 2 people without replacement from a large university where 0.6 are female?
- **Responses:**
  1. 2 from 5 with replacement:
  2. 2 from 5 without replacement:
  3. 2 from 1,000's without replacement:

 Practice: 6.8 p.237

Approximate Independence when Sampling Without Replacement

Rule of Thumb: When sampling without replacement, events are approximately independent if the population is at least 10 times the sample size.

Looking Ahead: *Because almost all real-life sampling is* ***without*** *replacement, we need to check routinely if population is at least 10n.*

Probability of Occurring At Least Once

To find the probability of occurring at least once, we can apply the "Not" Rule to the probability of not occurring at all.

Example: *Probability of Occurring At Least Once*

- **Background**:Probability of heads in coin toss is 0.5.
- **Question:** What is probability of at least one head in 10 tosses?
- **Response:**

> ***Looking Back:*** *Theoretically, we could have used the "Or" Rule, adding the probabilities of all the possible ways to get at least one heads. However, there are over 1,000 ways altogether!*

 Practice: 6.11b p.237

Basic Probability Rules (*Review)*

Non-Overlapping "Or" Rule: For any two ***non-overlapping*** events A and B,

P(A or B)=P(A)+P(B).

Independent "And" Rule: For any two ***independent*** events A and B,

P(A and B)=P(A)×P(B).

More General Probability Rules

- Need "Or" Rule that applies even if events ***overlap***.
- Need "And" Rule that applies even if events are ***dependent***.
- Consult two-way table to consider combinations of events when more than one variable is involved.

Example: *Parts of Table Showing "Or" or "And"*

- **Background**: Professor notes gender (female or male) and grade (A or not A) for students in class.
- **Questions:** Shade parts of a two-way table showing…

1. Students who are female *and* get an A?

	A	not A	Total
Female			
Male			
Total			

2. Students who are female *or* get an A?

	A	not A	Total
Female			
Male			
Total			

- **Responses:** 1. Shade ______________________________

  2. Shade ________________________________

Practice: 6.14a,c p.252

Example: *Intuiting General "Or" Rule*

- **Background**: Professor reports: probability of getting an A is 0.25; probability of being female is 0.60. Probability of both is 0.15.
- **Question:** What is the probability of being a female or getting an A?
- **Response:**

Practice: 6.14d-h p.252

Example: *Intuiting General "Or" Rule*

- **Response:** Illustration with two-way tables:

	A	not A	Total
Female	0.60		
Male			
Total			

+

	A	not A	Total
Female	0.25		
Male			
Total			

−

	A	not A	Total
Female	0.15		
Male			
Total			

=

	A	not A	Total
Female	P(Female or A) =0.60+0.25-0.15 =0.70		
Male			
Total			

Practice: 6.14d-h p.252

General "Or" Rule (General Addition Rule)

For *any* two events A and B,

P(A or B) = P(A) + P(B) - P(A and B).

=0 if no overlap

A Closer Look: *In general, the word "or" in probability entails addition.*

Example: *Applying General "Or" Rule*

- **Background**: For households in a certain district in Benin, the probability of having a bank account is 0.08. The probability of owning a latrine is 0.15. The probability of having a bank account *and* a latrine is 0.03.
- **Question:** What is the probability of having a bank account *or* a latrine? What's the probability of having neither?
- **Response:**

Probability Rules (*Review)*

Non-Overlapping "Or" Rule: For any two ***non-overlapping*** events A and B,
P(A or B)=P(A)+P(B).

Independent "And" Rule: For any two ***independent*** events A and B,
P(A and B)=P(A)×P(B).

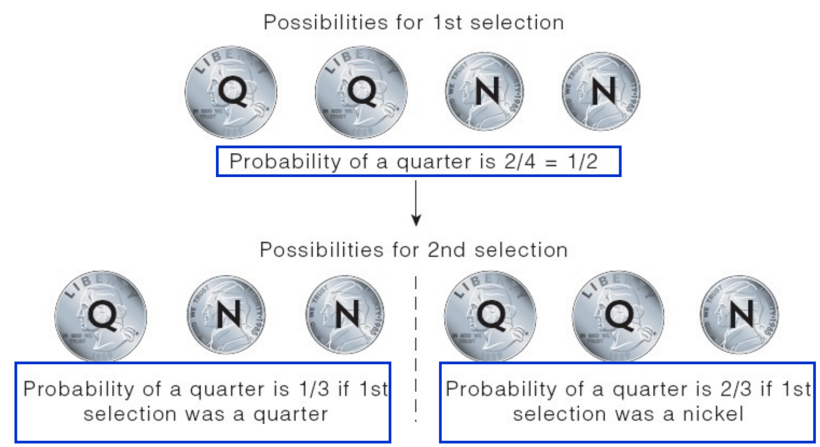
General "Or" Rule: For any two events A and B,
P(A or B)=P(A)+P(B)-P(A and B).

Need "And" Rule that applies even if events are ***dependent***.

Example: *When Probabilities Can't Simply be Multiplied (Review)*

- **Background**: In a child's pocket are 2 quarters and 2 nickels. He randomly picks a coin, does *not* replace it, and picks another.
- **Question:** What is the probability of the first and the second coins both being quarters?
- **Response:** To find the probability of the first *and* the second coin being quarters, we can't multiply 0.5 by 0.5 because after the first coin has been removed, the probability of the second coin being a quarter is *not* 0.5: it is 1/3 if the first coin was a quarter, 2/3 if the first was a nickel.

Example: *When Probabilities Can't Simply be Multiplied*



Definition and Notation

Conditional Probability of a second event, given a first event, is the probability of the second event occurring, assuming that the first event has occurred.

P(B given A) denotes the conditional probability of event B occurring, given that event A has occurred.

Looking Ahead: *Conditional probabilities help us handle dependent events.*

Example: *Intuiting the General "And" Rule*

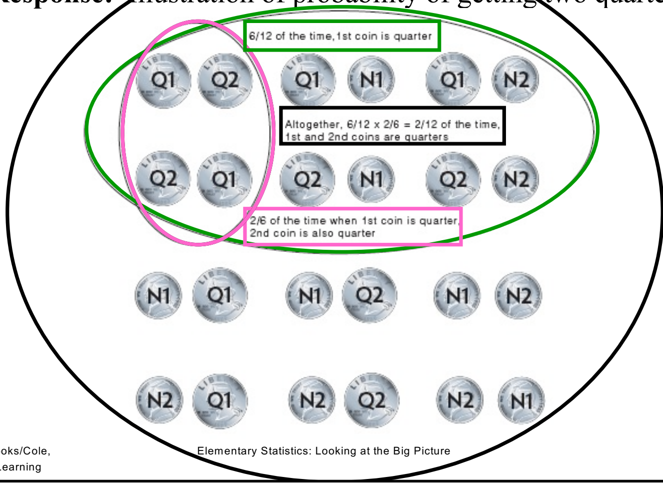
- **Background**: In a child's pocket are 2 quarters and 2 nickels. He randomly picks a coin, does *not* replace it, and picks another.
- **Question:** What is the probability that the first *and* the second coin are quarters?
- **Response:** probability of first a quarter (____), times (conditional) probability that second is a quarter, **given** first was a quarter (____):

____

Practice: 6.16a p.253

Example: *Intuiting the General "And" Rule*

- **Response:** Illustration of probability of getting two quarters:



Example: *Intuiting General "And" Rule with Two-Way Table*

- **Background**: Surveyed students classified by sex and ears pierced or not.

	Ears Pierced	Ears Not Pierced	Total
Female	270	30	300
Male	20	180	200
Total	290	210	500

- **Question:** What are the following probabilities?
  - Probability of being male
  - Probability of having ears pierced, given a student is male
  - Probability of being male and having ears pierced
- **Response:**
  - P(M) =
  - P(E given M) =
  - P(M and E)=

Practice: 6.14g,i,b p.252

General "And" Rule (General Multiplication Rule)

For *any* two events A and B,

P(A and B)=P(A)× P(B given A)

=*P(B) if A and B are independent*

A Closer Look: *In general, the word "and" in probability entails multiplication.*

Example: *Applying General "And" Rule*

- **Background**: Studies suggest lie detector tests are "well below perfection", 80% of the time concluding someone is a spy when he actually is, 16% of the time concluding someone is a spy when he isn't. Assume 10 of 10,000 govt. employees are spies.
- **Question:** What are the following probabilities?
  - Probability of being a spy and being detected as one
  - Probability of *not* being a spy but "detected" as one
  - Overall probability of a positive lie detector test
- **Response:** First "translate" to probability notation:

  P(D given S)=_____; P(D given not S)=_____; P(S)=_____; P(not S)=_____
  - P(S and D) = ____________
  - P(not S and D) =____________
  - *6.48* P(D) = P(S and D or not S and D) =__________________

 Practice: 6.16a,b,d p.253

Example: *"Or" Probability as Weighted Average of Conditional Probabilities*

- **Background**: Studies suggest lie detector tests are "well below perfection", 80% of the time concluding someone is a spy when he actually is, 16% of the time concluding someone is a spy when he isn't. Assume 10 of 10,000 govt. employees are spies.
- **Question:** Should we expect the overall probability of being "detected" as a spy, P(D), to be closer to P(D given S)=0.80 or to P(D given not S)=0.16?
- **Response:** Expect P(D) closer to ____________________ because ___________________________

  (In fact, P(D) =0.16064.)

 Practice: 6.16c p.253

General "And" Rule Leads to Rule of Conditional Probability

Recall: For *any* two events A and B,

P(A and B)=P(A)×P(B given A)

Rearrange to form ***Rule of Conditional Probability:***

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

Example: *Applying Rule of Conditional Probability*

- **Background**: For the lie detector problem, we have
  - Probability of being a spy: P(S)=0.001
  - Probability of spies being detected: P(D given S)=0.80
  - Probability of non-spies detected: P(D given not S)=0.16
  - Probability of being a spy and detected: P(D and S)=0.0008
  - Overall probability of positive lie detector: P(D)=0.16064
- **Question:** If the lie-detector indicates an employee is a spy, what is the probability that he actually is one?
- **Response:** P(S given D) =

Note: *P(S given D) is very different from P(D given S).*

A Closer Look: *Bayes Theorem uses conditional probabilities to find probability of earlier event, given later event is known to occur.*

 Practice: 6.16e,f p.253

Two Types of Error in Lie Detector Test

1st Type of Error: Conclude employee is a spy when he/she actually is not.

2nd Type of Error: Conclude employee is not a spy when he/she actually is.

Example: *Two Types of Error in Lie Detector Test*

- **Background**: For the lie detector problem, we have
  - Probability of spies being detected: P(D given S)=0.80
  - Probability of non-spies detected: P(D given not S)=0.16
- **Questions:**
  - What is probability of 1st type of error (conclude employee is spy when he/she actually is not)?
  - What is probability of 2nd type of error (conclude employee is not a spy when he/she actually is)?
- **Responses:**
  - 1st type:
  - 2nd type:

Lecture Summary
(Finding Probabilities; Probability Rules)

- Probability: definitions and notation
- Rules
  - Permissible probabilities
  - "Sum-to-One" Rule
  - "Not" Rule
  - "Or" Rule (for non-overlapping events)
  - "And" Rule (for independent events)
- Independence and sampling with replacement
- More Rules
  - General "Or" Rule (events may overlap)

Lecture Summary

(Finding Probabilities; More General Rules)

- General "And" Rule
- More about Conditional Probabilities
- Two Types of Error