How Students Understand Statistics: A Prime Time to Implement Projects

Presented at

AMATYC 43rd Annual Conference

San Diego, CA

November 9, 2017

Wendy Pogoda

Hillsborough Community College

Ruskin, FL

wpogoda@hccfl.edu

Dr. Rachel Bates

Redlands Community College

El Reno, Oklahoma

Rachel.Bates@redlandscc.edu

Mathematics Instruction Methods

Mathematics in higher education has long embraced traditional teacher-centered methods of instruction: non-interactive.

Historically, trends in teaching and learning mathematics have cycled through behaviorism and cognitivism. Recent reform efforts that promote conceptual understanding instead of the memorization of isolated facts are rooted in constructivism.

Mathematics Instruction Methods

Despite the persuasive evidence in cognitive science and education research, too many college-level students still experience lecture-based forms of instruction in mathematics related courses.

It is then incumbent upon members of our profession to disseminate education and cognitive research and model the pedagogical strategies that support cognitive construction and align with the reform movements set forth by AMATYC (2006), GAISE (2012), and NCTM (2000).

Student Centered Instruction Methods

Student centered pedagogical approaches can be traced to Piaget and Dewey. Over the decades, multiple student centered pedagogies have been developed and implemented in various disciplines.

Medical and STEM related courses have long used inquiry based learning, problem/project based learning and discovery based learning. Mathematics has been slow to implement student centered instruction methods despite the research and literature.

Outline

- In-class Activities
  - Overview
  - Try It!
  - Student Examples
- Projects
  - Overview
  - Try It!
  - Student Examples
  - Data
- Q&A

Activities

Engage students and help build critical thinking in the classroom with activities

Why do activities?

Traditional passive teaching methods found in lectures provide little time if any, for discussions and dialogue among students themselves and with the instructor. Providing students with opportunities to share ideas, to question their understanding and to find common understandings are essential to learning.

Encouraging discourse among students also assists in developing student's skills to effectively work and collaborate with others.

Activity 1: M&M's

Intro Activity 1 – Data Collection

This activity is a simple, yet effective way to introduce students to the basic concepts they will be encountering in AP Statistics. I recommend doing this on the first or second day of class as an icebreaker activity, but also as a way to gauge your students' understanding of basic statistical concepts. Take the time to discuss the 4 themes of statistics during the activity - Describing Data, Collecting Data, Probability, and Inference - as a way to introduce students to the concepts they'll be learning throughout the year.

Activity 1: Try It!

Data Collection

Welcome to APStats! Throughout the course history, mathematics has been called upon to answer some of the world's most pressing questions, including, but not limited to: "What colors can I expect to find in a bag of milk chocolate m&m's?" Your task is to collect data on the m&m color distribution through a carefully designed and controlled experiment in which you will:

- open a bag,
- count the number of m&m's falling into each color category,
- record the numeric results,
- record your color distribution on the class dotplots, and
- properly dispose of the m&m's by seeing whether or not they really do melt in your mouth, not in your hand.

Be sure to record your data accurately as we will be referring to it throughout the course.

Blue	Brown	Green	Orange	Red	Yellow

Projects

Help your students learn to utilize statistical technology and analyze complex problems with projects!

Why do projects?

Main purpose: To give students an opportunity to work with data in a more realistic setting, so that they can learn how to:

- Use technology
- Explain and present statistical results
- Deal with “messy” data

Six Part Project – Great for Online

Converted online course from 2 unconnected projects using professor given data to 6 connected projects to analyze a single data set the students gather themselves (with student choice of topics).

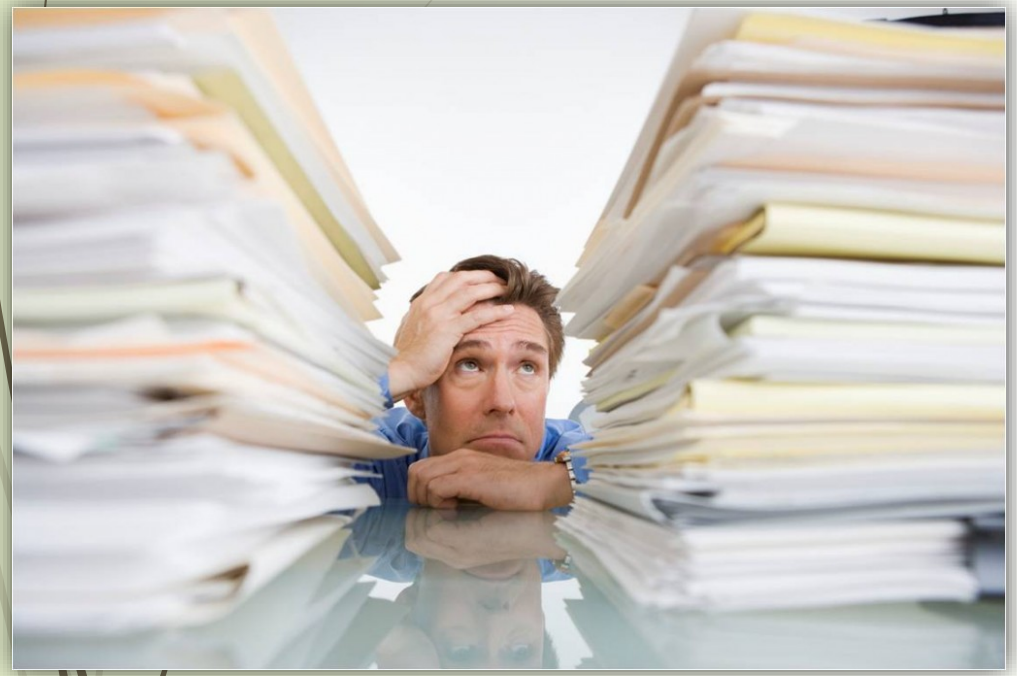
- Part 1: Data gathering plan
- Part 2: Submit your data
- Part 3: Summary statistics
- Part 4: Confidence interval for the population mean
- Part 5: Confidence interval for the population proportion
- Part 6: Simple linear regression

Grading

Part 1 grading: Check if the project is doable, meets the criteria, and is explained correctly. Common mistakes are:

- The variables not matching the experimental unit (e.g. surveying people, yet the variable is "price")
- Confusing the term variable and a summary statistic (e.g. having the variable be the number of females in the sample).
- Missing units of measurement for quantitative variables.
- Qualitative variable being a subject identifier instead of a useful variable (e.g. the variable is the person's name).
- Confusing the sample (group that you intend to include in your study) and the population (group you are ultimately interested in).

Try grading!



Part 3 grading:

- Do the summary statistics seem reasonable?
  - $s \approx \frac{\max - min}{4}$
- Is everything included? (Summary statistics, 1 QN graph, counts for the QL variable, and a QL graph)
- Do the graphs make sense and understandable?

Part 3 Directions

Using the data in Part 2 of this project, attach a document that includes the following:

- For each of your quantitative variables, find the sample size, mean, standard deviation, 5 number summary (min, lower quartile, median, upper quartile, and max). Please use a statistical or spreadsheet program or calculator to calculate the standard deviation. Do NOT calculate this by hand.
- For one, or both, of your quantitative variables, provide an appropriate graph to display this data (either a histogram, box plot, dot plot, or stem & leaf plot). I'd recommend doing this in […statistical software…] and copying the picture over to your document.
- For your qualitative variable, provide the count of each response level (e.g. 20 "yes", 40 "no")
- For your qualitative variable, provide an appropriate graph to display this data (either a bar graph or a pie chart).

Project Part 3 - Example 1 Feedback

This is a good start, but, the mean and standard deviation are wrong for both QN variables. You may wish to use technology to calculate these numbers. Also, there was a typo on the min for the price variable. In (spread sheet program) the formula for finding the mean of numbers in column A would be =average(A:A). Similarly, the formula for standard deviation would be =stdev.s(A:A). Let me know if you need assistance. There are tutorial videos posted as well. The second half of this project part is missing (information on the QL variable). Please include that and fix the errors in the first part and resubmit for full credit on this part. ☺

					Points
Summary Statistics for QN1	Full marks 5 pts	Some values missing or incorrect 4 pts	Multiple Issues 2 pts	No Marks 0 pts	4
Summary Statistics for QN2	Full marks 5 pts	Some values missing or incorrect 4 pts	Multiple Issues 2 pts	No Marks 0 pts	4
QN Graphical display provided	Full marks 5 pts	Yes, but some parts are unclear. 3 pts	No Marks 0 pts		5
Counts given for QL variable	Full marks 5 pts	Some values are missing or incorrect. 3 pts	No Marks 0 pts		0
Graphical display for QL variable	Full marks 5 pts	Yes, but labels are missing or incorrect. 3 pts	No Marks 0 pts		0
				TOTAL	**13**

Project Part 3 - Example 2 Feedback

This is good, except for the "histogram." What you actually provided was a pie chart. To create a histogram in [statistical software] follow the tutorial provided on the main course page. Resubmit with corrections for full credit.

					Points
Summary Statistics for QN1	Full marks 5 pts	Some values missing or incorrect 4 pts	Multiple Issues 2 pts	No Marks 0 pts	5
Summary Statistics for QN2	Full marks 5 pts	Some values missing or incorrect 4 pts	Multiple Issues 2 pts	No Marks 0 pts	5
QN Graphical display provided	Full marks 5 pts	Yes, but some parts are unclear. 3 pts	No Marks 0 pts		3
Counts given for QL variable	Full marks 5 pts	Some values are missing or incorrect. 3 pts	No Marks 0 pts		5
Graphical display for QL variable	Full marks 5 pts	Yes, but labels are missing or incorrect. 3 pts	No Marks 0 pts		5
				TOTAL	**23**

Project Part 3 - Example 3 Feedback

This is good, but there are a few issues. I think you switched both Q1 and Q3 between your variables (i.e. Q1 for age appears to be Q1 for family size). Usually histograms have the bars touching, but the way you have it is okay. Your QL variable is if a family has a pet or not, so use that variable to make the pie chart (pie charts should only be used on QL data). Resubmit with corrections for full credit on this part.

					Points
Summary Statistics for QN1	Full marks 5 pts	Some values missing or incorrect 4 pts	Multiple Issues 2 pts	No Marks 0 pts	4
Summary Statistics for QN2	Full marks 5 pts	Some values missing or incorrect 4 pts	Multiple Issues 2 pts	No Marks 0 pts	4
QN Graphical display provided	Full marks 5 pts	Yes, but some parts are unclear. 3 pts	No Marks 0 pts		5
Counts given for QL variable	Full marks 5 pts	Some values are missing or incorrect. 3 pts	No Marks 0 pts		5
Graphical display for QL variable	Full marks 5 pts	Yes, but labels are missing or incorrect. 3 pts	No Marks 0 pts		0
				TOTAL	**18**

Project Part 3 - Example 4 Feedback

This is good, but there are a few issues. First, the summary statistics for the age variable are not right, and the mean isn't right for the number of flights variable. The minimum age in your sample is 16 (not 17), so the average can't be 7.8 and the median can't be 7.5. Please double check your calculations for both of your variables. Summary statistics and graphs can be found very easily using statistical software, so you may want to try that (there's a tutorial posted on the main course page). You'll want to create a scatter plot for part 6 of the project, but, for now, I'll need you to put in a stem and leaf plot, box plot, or histogram (again, this is easier to do with statistical software). The Stem & Leaf plot shouldn't be used for QL information. To graph the QL variable (gender), you'll want to use a pie chart or bar graph. Make these corrections and resubmit for full credit on this part.

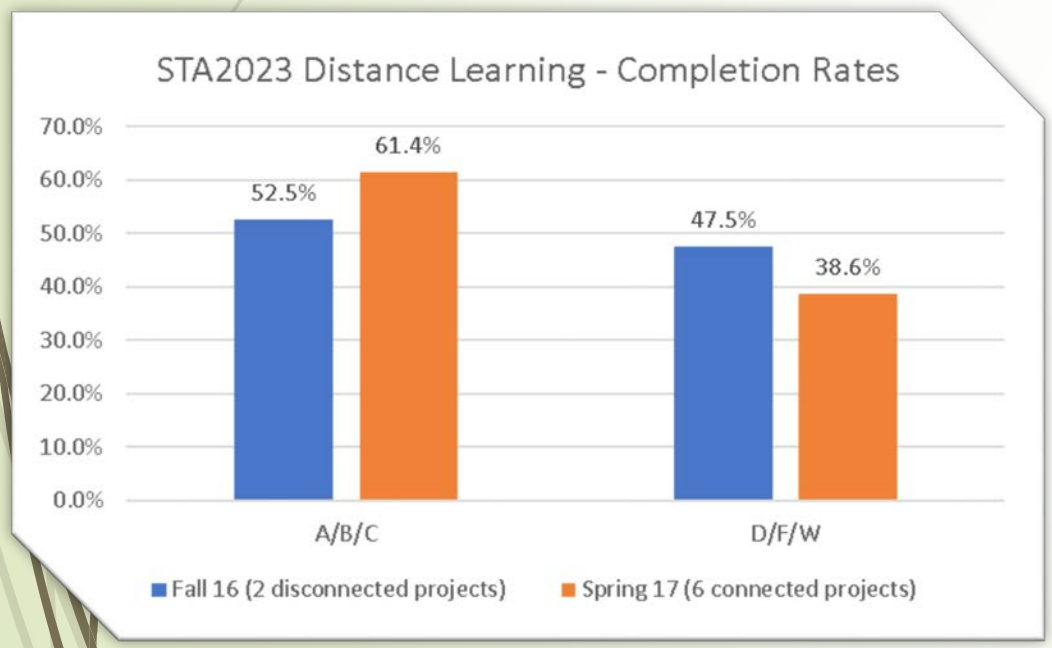
					Points
Summary Statistics for QN1	Full marks 5 pts	Some values missing or incorrect 4 pts	Multiple Issues 2 pts	No Marks 0 pts	2
Summary Statistics for QN2	Full marks 5 pts	Some values missing or incorrect 4 pts	Multiple Issues 2 pts	No Marks 0 pts	2
QN Graphical display provided	Full marks 5 pts	Yes, but some parts are unclear. 3 pts	No Marks 0 pts		3
Counts given for QL variable	Full marks 5 pts	Some values are missing or incorrect. 3 pts	No Marks 0 pts		5
Graphical display for QL variable	Full marks 5 pts	Yes, but labels are missing or incorrect. 3 pts	No Marks 0 pts		0
				TOTAL	**12**

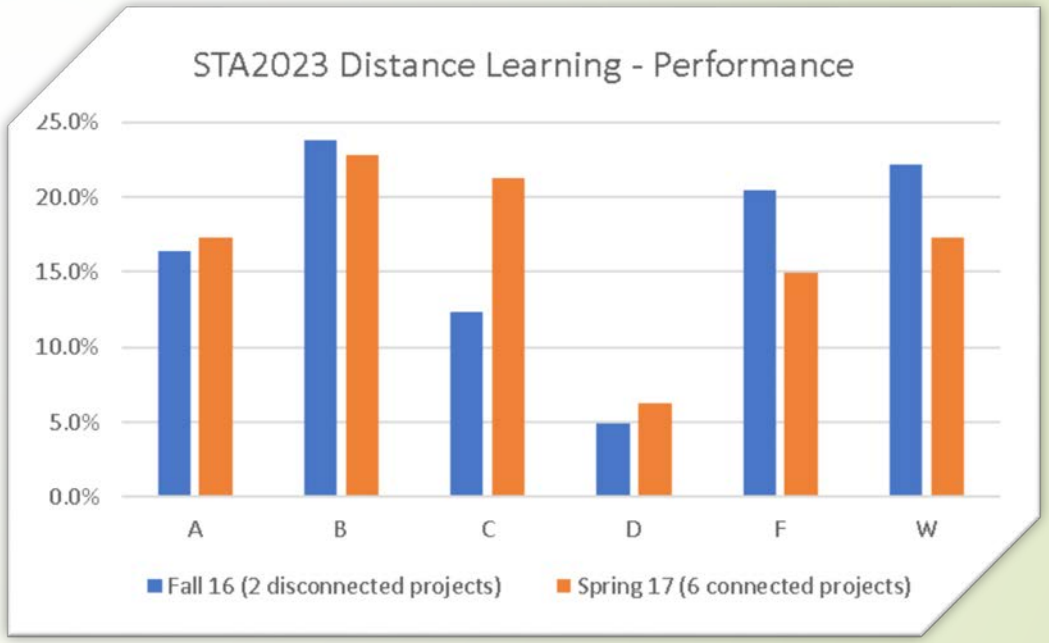
Benefits

- Increase student understanding of complex concepts and vocabulary. Misunderstandings of statistical concepts are clarified early in the semester.
- Increase use and appreciation of statistical software.
- Encourage academic integrity. With every student having a different project, there is a reduced temptation to just copy someone else's projects.
- Provides an opportunity for student-teacher interaction early in the semester.
- Engage students by allowing them to analyze data that interests them.

Benefits (cont.)

- Increase course retention rates!

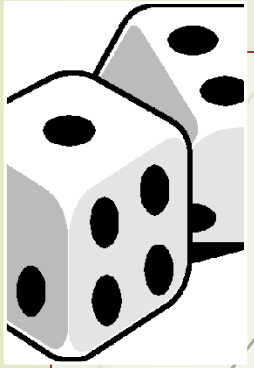




Survey Project

- Creating surveys can be an excellent tool to document what is happening in our communities, help us understand ourselves better, and let others get to know life is like for our people. When we document our experiences, we honor our day-to-day reality as knowledge, while creating a powerful and strategic tool to help bring justice to our communities. Part of oppression is keeping information and knowledge in the hands of the powerful. When we reclaim research, we walk one step closer to achieving liberation.
- This semester, you will be conducting a Survey Project. The project is to write, distribute, and analyze a survey.
- In order to complete this project, you will each need to survey at least 10 people. Start preparing yourself for this process – it means you are going to have to speak to lots of people, some or all of whom might be strangers.

Other Useful/Fun Projects



Probability--great for groups!

Descriptive Statistics with M&M's

Instructor provided published data and have the class interpret the results—leave the data messy!!

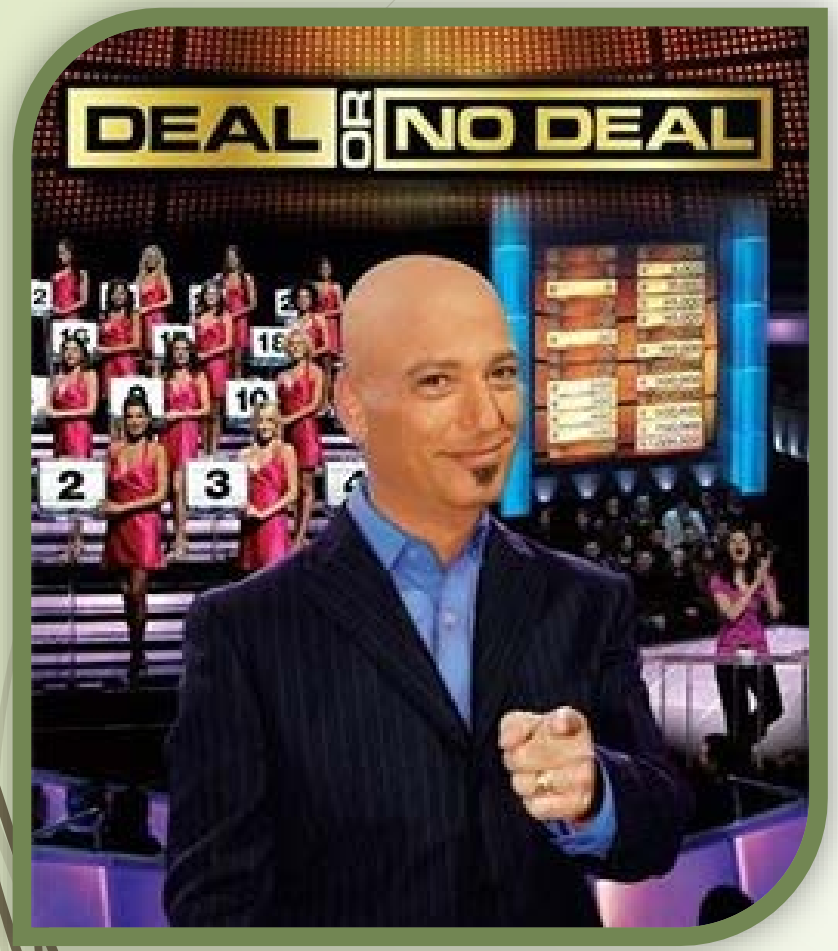
Basic Structure of Probability Project & Creating Your Own

- Each student picks a topic from a list of ideas (they can work in groups or independently). The project list has a dozen probability questions to answer (each student picks one and they may work in a group).
  - Choices allow students to pick a topic that interests them, motivating them to learn, network with students with similar interests, and encourages academic integrity.
  - Semesters where I've given this choice of probability projects, 70% of students said their favorite topic in the course was probability!
- The initial list took approximately a month to create – this list and a guide to creating your own is posted. ☺

Probability Project - Example 1

- **Counting Possible Songs.**
  - Watch the YouTube Video, “Will We Ever Run Out of Music?”
  - How many simple eight beat melodies are possible using only 1 octave with all possible quarter notes over that octave (including sharps/flats and rests)?
  - Optional: To put your answer in perspective, how long would it take you to play all these songs back to back at a rate of 120 beats per minute?

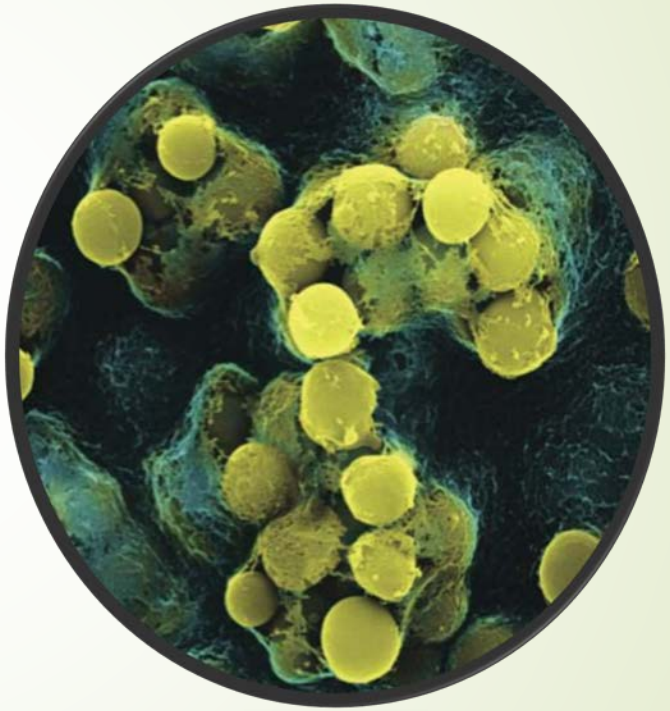
Probability Project - Example 2



- **Deal or No Deal.**
  - Watch “Stephanie Allen Deal or No Deal (parts 1 through 3) on YouTube.
  - Calculate the expected value of the initial case before any other cases are seen. After the first round of cases being revealed, what is the expected value? Was the offer fair? Around the 50 second mark of the third part, she does calculations on the mean and the median. What error did she make? Why did she accept the offer even though she knew it was below the mean? What does this tell you about expected values?

Probability Project - Example 3

- **Medical Testing.**
  - Pharyngeal swabs were taken from 818 patients with suspected streptococcal pharyngitis (strep throat). 171 patients had strep throat. The Strep A Rapid Test Device (SARTD) correctly identified 123 of these 171 specimens as having strep throat.
  - Create a contingency table outlining the results above and then calculate the sensitivity, specificity, positive predictive value, negative predictive value, and overall accuracy of the SARTD. For reference:
    - Sensitivity = P(test + | +)
    - Specificity = P(test - | -)
    - Positive Predictive Value = P(+ | test +)
    - Negative Predictive Value = P(- | test -)
    - Accuracy = P(result is correct)



http://kids.britannica.com

Probability Project #3. Actual Student Project

Medical Testing

How do you know if claims by a medical device manufacturer are misleading or not?

A test was conducted to measure the accuracy of a Strep A Rapid Test Device, also known as SARTD. 818 patients with symptoms of strep throat had pharyngeal swabs taken. The swabs were tested using two methods, with a conventional culture and with a SARTD device.

The results are tabulated in the table on the next slide.

2

SARTD Reliability

	SARTD Positive	SARTD Negative	Total
Culture Positive	123 (True Positive)	48 (False Negative)	171
Culture Negative	37 (False Positive)	610 (True Negative)	647
Total	160	658	818

There are many manufacturers of Strep A Rapid Test Devices. Each manufacturer makes claims regarding the accuracy of their test. These are the results of one manufacturers device compared to the near perfect accuracy of culturing a sample in artificial media and identifying the S. pyogenes microorganism under a microscope.

3

Sensitivity

- Sensitivity is defined as the proportion of positive results that are correct. To find the sensitivity simply divide the number of positive results identified correctly by the total number of cultures that were positive.
- 123/171 = 0.72 or 72%
- 123 true positives out of 171 who are actually infected with the pathogen. 48 cases of streptococcal pharyngitis failed to be identified by SARTD.

Specificity

- Specificity is defined as the proportion of negative results that are correct. To find the specificity simply divide the number of negative results identified correctly by the total number of cultures that were negative.
- 610/647 = 0.94 or 94%
- Notice that sensitivity and specificity are not complementary

5

Positive Predictive Value

- In this example the positive predictive value is the probability that an individual has streptococcal pharyngitis given that they tested positive using the SARTD.
- $P(+ \mid \text{SARTD} +) = 123/160 = 0.77$ or 77%

6

Negative Predictive Value

- In this example the negative predictive value is the probability that an individual does not have streptococcal pharyngitis given that they tested negative using the SARTD.
- $P(- \mid \text{SARTD} -) = 610/647 = 0.93$ or 93%

Conclusion

- Pharyngitis will typically go away on it's own but in some instances the consequences of withholding antibiotics can be severe.
- A physician will typically do a culture on every negative result from a SARTD.

8

Conclusion

- If the manufacturer of a SARTD misleads the physician into thinking that the device is 93% accurate (negative predictive value) this can greatly effect the cost of care.
- The sensitivity of the device is 72% which means that 28% of the SARTD results are false negatives.

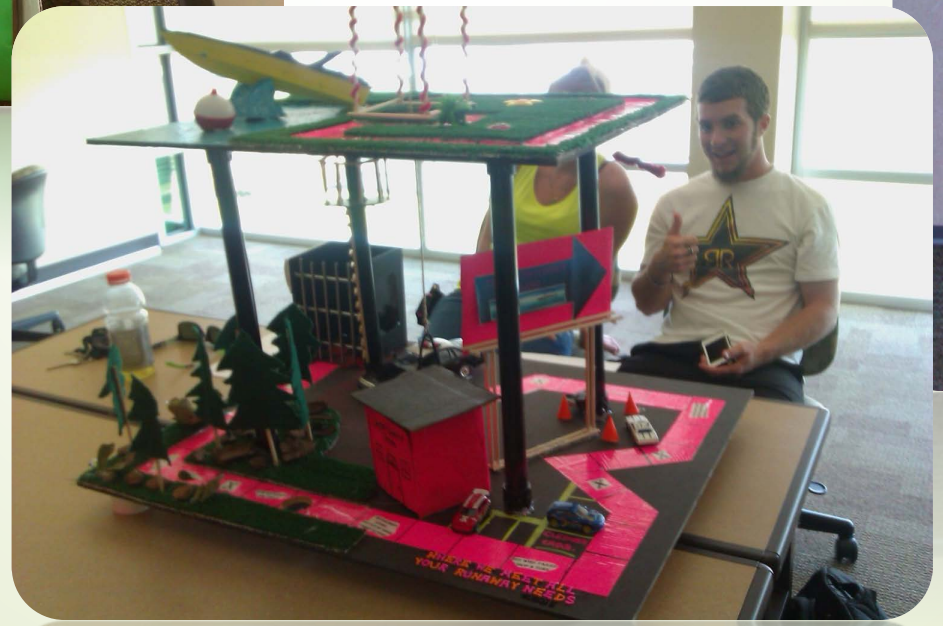
9

- The physician will be most concerned with the sensitivity of the SARTD device because that statistic indicates the number of false negatives that are returned by the device.
- Each true negative result will require a culture regardless of the accuracy of the SARTD.

Probability Project - Example 4

- **Create a gambling/carnival game.**
  - Create a game that someone would play at a carnival, amusement park, or casino.
  - Write a short paper, including:
    - Introduction: Provide an overview of the game (e.g. What type of game is it? Where would you play this game? How much does it cost to play? What are the prizes if you win?)
    - Instructions for the game
    - Game description (what do you need to play the game?)
    - Is this a fair game? Show math calculations for the expected value for winning the game. If the game is not fair, how can you change it to make it fair?

Probability Project #4 - Actual Student Projects



Student Responses

- All of the group and in-class work we did really helped with learning the material better. In lectures I would lose interest very easily and in this class I was able to stay focused. Talking and working as a group helped me learn the material better.
- Before I took this class, I didn't know how to do a lot of the equations and formulas. Nothing really made sense and I was just confused about statistics until I took this class. I know have a better understanding of how the calculations are done and what kind of answer to expect. The hands on things that were brought in gave me a better understanding as well.

Student Responses

- Prior to this course, I had limited understanding of what statistics was and how much it truly applied to my desired career path but now my eyes have been opened to all of its importance. This course was a refreshing reminder that not all math classes are full of brutal number crunching but instead, some actually target the analytical side of your brain.
- I didn't know statistics was so deep. I only thought it was numbers on a page. But now I can see it goes into depth really far. I felt as I learned more from the group activities and projects rather than doing the homework.

Useful Resources for Statistics Activities, Lessons, and Projects

- Our slides, handouts, project directions, etc. will be posted on the AMATYC Conference Proceedings.
- AMATYC Statistics Resource Page: http://www.amatyc.org/?page=StatsResources
  - Joint effort between AMATYC and the American Statistical Association to gather resources useful for community college statistics instructors.
- Academic Databases
- Government Data
  - Bls.gov
  - Census.gov
- Historical Stock Data - Yahoo Finance: https://finance.yahoo.com/
- Google Public Data: https://www.google.com/publicdata/directory

Conclusion

- How students react to and experience pedagogical changes contributes to understanding student learning about the subject area. Additionally, pedagogical changes alter the role of the instructor thereby presenting parallel challenges.
- Most of our students will encounter just one course devoted to the discipline of mathematics. We must prepare our students for a life of learning rather than for simple completion of our course objectives.
- We know that students learn best when they are truly engaged in what they are learning, when they have the opportunity to explore, debates, discuss, examine, defend, and experiment with the concepts and skills they are ready to learn.

Any Questions?

Wendy Pogoda

Hillsborough Community College

Ruskin, FL

wpogoda@hccfl.edu

Dr. Rachel Bates

Redlands Community College

El Reno, Oklahoma

Rachel.Bates@redlandscc.edu