Know Your Position -- How ?

Anne Cherian
Overbrook Elementary

Overview

The purpose of this unit is to expose students to some of the statistical tools that are available for validation and standardization. The emphasis will be on the application of the tools as well as in understanding the significance and interpretation of the outcome in real life situations. Students will learn about statistical distributions such as the binomial and normal distributions with a stress on normal distribution and its significance in the interpretation of standardized test scores. The unit is designed for students taking an Algebra 1,2/ Statistics class or any math class that includes a minimal coverage of probability and statistics. Since the unit is statistics based, it could be used as a model for students who wish to do a math based senior project.

I teach math at Overbrook High School which is a large comprehensive high school and has challenges that most inner city schools have. Majority of the students here, are burdened by socio-economic issues therefore the appreciation and use of mathematical tools is not on their priority list of learning. However, there is a small minority that yearns to climb the ladder of learning, this motivates me to move forward. The age group I teach ranges from sixteen thru eighteen, an age where every word from the media and advertisements is accepted without much questioning and the need in checking its veracity is not important. One of the purposes of this unit is to encourage students to ask the question ‘How do you know ‘ when information is thrown to them and decide for themselves based on facts if there is truth in the statement.

Rationale

Scenario 1: A student has just finished a test and is extremely thrilled that she has done well and expects a very high score probably in the nineties, an A+. It is assumed that the student is not exaggerating but is truthful in stating that most of the questions were not extremely challenging and most of her answers are correct.

Scenario 2: Another student has finished another test but is not quite pleased, he thinks the paper was tough and some of the questions were quite challenging and expects a score in the seventies probably a C.

When the results are published the students were in for quite a surprise. The student in scenario 1 was disappointed when she received only a B- and the student in scenario 2 was ecstatic when he got an A.

Consequently, the exclamation: "How did this happen? " what was the process that caused the change. There are various statistical tools used in assessments and standardization of scores and validating data and information. In this unit students will be exposed to some of the statistical tools and its use in validation and interpretation.

Most students I teach consider mathematics as something they do not have to deal with after high school and I try in many ways to convince them of the misconception they are under. This is another of my attempts in convincing them out of it. Moreover, the mathematics curriculum includes probability, combinations, binomial theorem, distribution and normal distribution at different stages. Since the topics are spread out students are not able to appreciate and see the connection of one with the other, recapitulation is an arduous task and becomes worse when it deals with formulas. Therefore this was one of the primary reasons of taking Pascal's triangle, combinations binomial theorem and the distributions as one complete unit. The constant back and forth reference to the concepts helps in understanding the main idea.

In today's fast pacing world there is no dearth for information. We have information of all sorts thrown to us. Information indeed is power, but this no longer true, as the credibility of the information acquired is sometimes questionable. Media plays a major role as a harbinger of information in most homes. The question is, how do we know for sure the authenticity of the information given to us? Do we use some internal measure to validate it or do we accept it as the truth? When information is impersonal, we tend to use it as conversational material, but when it involves serious matters like job, health, finance or career related topics, we respond to it differently. Similarly, as an educator I need to question my role in encouraging and motivating students to think critically and create a greater awareness that all information is not necessarily the truth, there is the need for validation.

Need for Statistics

It is important for students to truly appreciate the role of statistics and its relevance in today's world. To be able to understand the language of statistics and its interpretation is more important than knowing how to solve it, since appreciation and interest in a topic is by and large the greatest motivating factor in getting to know how to solve it. One way of achieving this is by getting students to predict the outcome in a real life situation. They should be able to distinguish good reasoning from faulty reasoning as it makes them less vulnerable to manipulation. Since statistics provides the tools to react intelligently and make wise choices, it is imperative for students to understand that it is one of the most important subjects in their course of study.

Data and information come to us from virtually every facet of life. Lets take a look at some of the claims we have read somewhere or someplace,

- People tend to be more persuasive when they look others directly in the eye and speak loudly and quickly.
- Women make 75 cents to every dollar a man makes when they work the same job.
- A surprising new study shows that eating egg whites can increase one's lifespan.
- People predict that it is very unlikely there will ever be another baseball player with a batting average over 400.
- There is an 80% chance that in a room full of 30 people that at least two people will share the same birthday.
- 79.48% of all statistics are made up on the spot.
- Almost 85% of lung cancers in men and 45% in women are tobacco-related.

Though all the above examples are diverse, they are statistical in character. They come from psychology, health, law, sports, business, etc. The question is, does the statistical information serve a purpose. One of the purposes of statistics is the ability to make predictions based on previously gathered data. Being able to predict the future not only changes our lifestyle but also helps us to be more effective. With statistics we can determine how we will live a year or ten years from now. Using past data we are able to understand climate changes and global warming and make wise decisions. We can to a certain extent predict hurricanes and tsunamis, though we cannot stop it at least precautions can be taken to avoid disastrous consequences.

An example of a situation where statistics has helped is smoking. Looking at the data gathered in the 90's there is an indication that cigarette smoking in the 10th grade has been slowly declining over the years. From this we can assume that something is being done correctly to bring the statistics down. Another example is that in 1975, more attention was being paid to the spousal homicide rates. Since proper precautions to help lower these rates were taken, it worked. Since 1975, the yearly spousal homicide rate has gone down from 2300 to 800. Statistics helped people to make predictions and help save lives. Statistical methods help examine information. Moreover, statistics is used for decisions when we are faced with uncertainties

Statistics could be defined as a study of data obtained in various ways. The study involves collecting, organizing analyzing and interpreting. There are two aspects in statistics. Data collected needs to be organized in a way best suited for understanding and for these there are various statistical tools such as stem –leaf plot, frequency tables, pie chart and histograms. Interpreting the data is the next aspect. Understanding the relationship between the data, its spread and its significance. For this, some of the tools are the mean, median, mode, variance and standard deviation. Once again, interpretation of the data can be two fold, one where the data relates to the actual fact, it involves the entire population and the analysis is based on the collected data. The other is using a smaller collection of data a prediction or inferences are made about a larger population using probability. I have incorporated both these aspects in this unit for students to get a wholesome understanding on the subject.

Misinterpretation of Statistics

It is true statistics are often presented in an effort to add credibility to an argument or advice. Statistics is definitely a scientific tool to measure validity. But does that indicate that all statistical figures present the whole truth? This famous quote by British Prime Minister Benjamin Disraeli, "There are three kinds of lies - lies, damned lies, and statistics." is interesting, since it is true statistical information can be misleading and the truth can be obscured. But Frederick Mosteller‘s quote "It is easy to lie with statistics but it is easier to lie without them." gives statistics its due. Importance. Since data by itself is meaningless and does not tell a story without interpretation, the importance in interpreting data is crucial. Besides this, the other area open for distortion is choosing the appropriate population for data collection in a statistical study.

Consider the example of the study on the influence of oat bran on cholesterol. When the first study was initiated and the observations made, the outcome was eating oat bran everyday lowered cholesterol levels. The study was well accepted by the scientific community and the media took it a step further, so much so that oats became a staple diet in most American homes and the market value of the Oats industry reached soaring heights. There was truth in the statistical finding and several people benefited from including oats in their daily menu. Interestingly, another study conducted by Swain et al, disagreed with the previous study. In their study however, there was no significant drop in the cholesterol level with the intake of oats. What caused this? Was the earlier study a bunch of lies or did the new study work on a selected population wherein no dramatic changes could be seen in the cholesterol level before and after the inclusion of oats in the diet. Since the article was published in the New England Medical journal it once again caught the attention of the media and educated elite, the result was the oats industry suffered a setback. It was understood later that the group of individuals selected for the study were those with normal cholesterol levels. Can one expect significant changes from a selected group as this?

Students will be exposed to these conflicting views, to make them aware that such things are not uncommon. Therefore the knowledge of statistics is essential lest we become vulnerable to the manipulation of those with vested interests.

The other area of emphasis will be, standardized testing and the importance of knowing their position in terms of academic performance in comparison to a selected population of students. To understand this better, the bell shaped curve or the Gaussian curve commonly called the normal distribution will be the focus of study in this unit. Students will become familiar with the various distributions and in particular the normal curve. The significance of its shape, the statistical relevance of the model and its popularity as a tool in predicting probabilities of events will be dealt within the lesson.

Objective

The objective of this unit is to allow students to explore statistical concepts by way of questioning. The three lesson plans are geared towards making statistics very practical, relevant and applicable to situations and scenarios they can relate to. Students would see the amazing ways the Pascal triangle can be used, the dry, boring binomial theorem springs to life when connected with the Pascal's triangle. They would see that the binomial distribution and normal distribution are not abstract theoretical models but tools that could be used to predict outcomes, validate a claim or answer questions where there is an element of doubt. Most importantly to be confident in understanding and interpreting their own standardized test scores. The entire process is develop logical and analytical skills and to expose students to various tools and its scope with the advantage of technology.

Standards

The lesson in this unit is designed for high school math that incorporates statistics, data analysis and percentages. Students will get a basic understanding of sequences and series, combination, binomial theorem and distributions. They will learn to apply statistical models to real time situations. Students will get to be more familiar with distributions and appreciate the significance and relevance of statistical tools through the procedure of data collection, tabulation, analysis and interpretation. The unit will fulfill the Pennsylvania Academic Standards for Math and Technology listed in the appendix.

Strategies

The strategies I plan to use in the unit are two directional, one is, specifically done by students at an individual level and the other, incorporating it in the classroom.

Student Directed:

Math Journal: Every student will be required to keep a journal that records the date and the lesson they have learned. This helps students to keep track of lessons they have learnt and to make connections when a new lesson is taught. The second strategy I encourage students to use is *Structured Note Taking*. Here students are encouraged to take notes when listening to a presentation. This strategy emphasizes purposeful reading and identifying essential information. Writing while listening helps in sustained focus and retention of concepts. *Concept Maps*: It is an excellent tool for connecting and summarizing ideas around a specific topic. I encourage students to use at the introduction to a lesson and the end of the lesson as it ties up new information with previous information.

Classroom Directed:

Cooperative learning: Students will be work in groups during an activity. This helps in building group dynamic skills besides academics and also encourages peer learning and sharing. *Brainstorming/ Group Discussion* is something they will use after they have finished an activity and analyze the activity done to draw inferences. This activity promotes critical thinking and increases reasoning skills. *Reciprocal Teaching*: Students learn and discover for themselves by posing questions based on the text, summarizing the content and predicting what will be next. The next two strategies make use of technology and I plan to use it more frequently in my classroom as I now have the resources. *Internet based animation*: This is an excellent methodology for abstract concepts. A single picture/animation is worth a thousand words/ expressions. Since the students have access to laptops and a smart board I wish to incorporate it into the lesson and maximize its use.

Classroom Activities

The unit has three lesson plans. The first two covers Pascal's triangle, Binomial theorem and Binomial Probability Distribution and the third lesson on Normal Distribution Theorem. The time frame for the entire unit is five to six classes and each class is for about 50 minutes.

Lesson Plan 1: *Exploring the Binomial Theorem*

This lesson plan is intended as an introduction to Binomial probability distribution. In this lesson students explore the patterns and relationships in Pascal's Triangle through a variety of activities including web activities on real-life applications on Pascal's Triangle. The lesson has 2 parts to it. Part 1 deals with familiarizing and reviewing Pascal's triangle and the combination formula and should take around 2 classes. Part 2 deals with the expansion of the binomial theorem and its application in problem solving, this would take at least 1 class. At the end of the lesson students would be able to complete Pascal's triangle to any number of rows and use it to determine the number of combinations for a given set of parameters. Students will make connections between the combination formula and Pascal's triangle and most importantly to understand and solve problems on the binomial theorem.

Part 1:

As an introduction to the lesson students will be given a handout of the Pascal's triangle (about 8 rows) and will be asked to write what they know about it and if they are able to observe any patterns. It is presumed that students have some basics on series and sequences and how to look for patterns.

Pascal's Triangle

					1					
				1		1				
			1		2		1			
		1		3		3		1		
	1		4		6		4		1	
1		5		10		10		5		1

After having written what they observed or recollect previous knowledge they will answer the following questions.

1. Highlight the following patterns: Natural numbers, Fibonacci numbers and triangle numbers.

2. Using the previous row, can the numbers for the next row of the triangle be found?

3. Find the sum of each row? Is there a pattern in the sums of each row?

4. Do the numbers in each row repeat, is there a pattern?

Since this is a warm up activity, at the end of it there will be a short discussion on the above patterns, this also helps those students who are still trying to complete it. The discussion will lead to students seeing the connection between powers of 2 and the sum of each row, also between the row number and the power.

The next part of the lesson involves tossing coins and making a note of the outcomes. The purpose in this activity is to make a connection with the outcomes of a coin and Pascal's triangle. It can be done with coins and recording the outcomes or verbally listing out the possibilities. If it is an activity, students have to take note of not recording duplicate outcomes/pattern.

A coin is flipped and the outcomes are listed as heads (H) or tails (T). Since the outcomes for a single and two coins take about five to 7 minutes, it can be done as class and the observations made on the worksheet.

Outcome for a single coin

Since the outcome could be just one head or one tail it could be written as follows

1 (H) or 1 (T)

Outcomes for 2 coins

All possibilities for 2 coins will be listed as follows

Coin1	Coin2
H	H
H	T

T H
T T

1 (HH) 2 (HT) 1 (TT)

Outcomes for 3 coins
Similarly all possibilities are listed out

Coin 1 Coin 2 Coin 3

At the end of this, the various combinations from the above experiment will be consolidated and written as follows

------ (HHH) ----- (HHT) -----(HTT) ----- (TTT)

Students are prompted to look for a connection with the numbers on each row of the Pascal's triangle and the outcomes for a single coin, 2 coins and 3 coins. Using Pascal's triangle as a guide they should predict the outcomes when 4 coins or 5 coins are flipped.

The combination trick in Pascal' s triangle

This section is a transition from Pascal's triangle to Binomial theorem. The purpose of this section is for students to figure out the question "How do you know that Pascal's triangle is based on the combination formula ". In addition, students would solve combination problems using the formula or Pascal's triangle. Since the lesson on combination was taught previously, it is assumed students have some knowledge of combinations therefore there will be a quick recapitulation on the formula n!/(n-r)!r!, where r is the number of items taken from n elements.

1. How many combinations are possible for the following
    a. To get 5 heads when 5 coins are tossed
    b. To get 2 heads when 5 coins are tossed
    c. To get 3 heads when 5 coins are tossed
    d. To get 1 tail when 5 heads are tossed
    e. To get 1 head when 5 heads are tossed

2. Do the numbers in the 5^{th} row of the Pascal's triangle match with the answers above. What do you conclude from this?

Students will use the Pascal's triangle to solve combination problems from the website http://mathforum.org/workshops/usi/pascal/pizza_pascal.html
This website allows students to work on different levels. The level chosen here is the intermediate level since it has a problem based on the different combinations called Antonio's Pizza Palace.

For those students who wish to explore further there is another activity called Pascal's petals on the same website.

Part 2 : *Binomial theorem and Pascal's triangle – the connection*

At this juncture since the students have spent some time on the activities and discussions on Pascal's triangle and its connection with the combination formula, it is much easier for them to appreciate the binomial theorem, its expansion and applications. At the end of the lesson they will be able to identify a binomial problem and apply the formula to get the final value. With regard to the expansion, the 3 methods will be explained. Students could choose the method most appropriate – depending on the extent of the expansion

Binomial Theorem: $\sum_{k=0}^{n} {}_nC_k\, x^{n-k}\, y^k$

As an introduction to this section students will use their previous knowledge of FOIL and distributive property to get the expansions of the following expressions

$(x+y)$, $(x+y)^2$, $(x+y)^3$

Students will be asked to write the first 4 rows of Pascal's triangle and compare it with the coefficients of each of the above 3 expansions and make connections.
A short presentation on the binomial theorem, its application and expansion will be given. Students will be made to see the ease in expanding a binomial expression with the help of Pascal's triangle. Some of the questions that will be dealt with at the end of the presentation will be - What does the x and y represent? How do we know that the theorem deals with combinations? Why is the binomial theorem and Pascal's triangle closely related? In conclusion the 3 methods of expansion will be explained using examples.

Method 1 Using the formula on the calculator

Binomial Theorem $(x+y)^n = \sum_{k=0}^{n} {}_nC_k\, x^{n-k}\, y^k$

Method 2: Using Pascal's triangle to get the coefficients

Method 3: When n is large using Pascal's triangle can get too long, in such cases method 3 can be used. For example, the expansion for $(x+y)^6$ is broken down into 3 steps

Step 1 Write the terms beginning from 1, therefore $(x+y)^6$ will have 7 terms

1 2 3 4 5 6 7

Step 2 Write the powers of x in decreasing order and the powers of y in increasing order

x^6y^0 x^5y x^4y^2 x^3y^3 x^2y^4 xy^5 x^0y^6

Step 3 We know the coefficient of the first term is 1. To get the second coefficient multiply the first terms power of x by its coefficient and divide by its term, hence to get the second coefficient, multiply the power of x which is 6 by its coefficient which is 1 and divide by its term which is 1 and that gives 6. Repeat the steps for the remaining terms. Therefore the coefficients for the sixth row will be

1 6 15 20 15 6 1

After the presentation and discussion, students will access the website below and use the animation to see the output for various number of rows and different probabilities. http://www.mathsisfun.com/data/quincunx.html

While viewing this website students discover for themselves the connection between Pascal's triangle and the binomial theorem and why the Galton board is able to bring about the same effect. Students will also take note of the histogram that is generated. This prepares them for the next lesson on the binomial probability curve and the normal curve

Students need to see the connection between the binomial formula and its application in solving real world problems that involve only 2 variables. Some examples of using the binomial theorem are to determine the distribution of answers on a true or false test, to determine the combination of heads and tails when tossing a coin or to determine the possible sequence of boys and girls in a family. In this example the binomial formula is used since the outcome is either a boy or a girl.

Example: There are 7 children in the Jones family. Of the seven children, there are at least 4 girls. How many of the possible groups of boys and girls have at least 4 girls.

Solution: Students should identify it is a binomial problem since there are 2 outcomes. Lets' assume x represents the girls and y the boys. Since the problem asks for at least 4 girls only the coefficient of the terms that have the exponent of x greater than 3 are considered. Hence expanding the eighth row of Pascal's triangle we get

$$1x^7 + 7x^6y + 21x^5y^2 + 35x^4y^3$$

Summing the coefficients gives the maximum possible groups, which is 64.

Class-work/ Homework: As a practice exercise students will be assigned class-work and homework from the text book -Holt Pages 739 -740

Web Assignment: Student will do a web search on Blaise Pascal and his major contributions in mathematics and complete a two- page essay.

Lesson 2: *Binomial Theorem and Binomial Probability Distribution*

At the end of the lesson students will be able to distinguish between a discrete random variable and a continuous random variable and to some extent understand the meaning of randomness. They will learn to solve problems involving both the probability factor and the binomial factor. Students will learn to graph a binomial distribution and interpret the significance of mean, expectation, variance and standard deviation.

The prerequisite for this lesson is basic knowledge of probability, measures of central tendency (mean, median, mode) and measures of dispersion (variance and standard deviation).

Keywords: Discrete random variable, Continuous random variable, binomial probability theorem, probability distributions, binomial probability distribution

To begin with students will be asked to write at least 3 examples where the outcomes from an experiment is a failure or a success. Students will share the information aloud and the examples will be discussed whether or not the examples can be considered as a binomial experiment and the reasons.

The activity described below is to help students get an understanding on the working of the binomial probability theorem. Students will work in pairs on this activity.

Activity: Finding binomial probabilities using the cube

Students will list out the possibilities of getting a 5 when a 6-sided die is rolled thrice. The tree diagram helps students to visually find the probability of getting exactly one 5 from the list of possibilities. The students can do the entire activity on their own or the teacher can model the first half as done below and have the students complete the rest of the activity

Roll1 Roll2 Roll3 Number of 5's Probability

Five 3 1/6 x 1/6 x 1/6 =1/216

Five

Five Other 2 1/6 x 1/6

x 5/6=5/216

Five 2 1/6 x 5/6 x 1/6=5/216

Other

Other 1 1/6 x 5/6 x 5/6=25/216

Other than five

(1,2,3,4 or 6) (This section is for students to complete similar to the one above.)

From the above possibilities the probability of getting exactly one 5 from the first section is 25/216(since there is only one possibility). Similarly students will draw and calculate the probabilities for the different combinations for the second section. This probability will turn out to be 2x25/216(since there are 2 possibilities). The total probability therefore will be 25/216+ 50/216=75/216 which is approximately 0.347.

The Binomial Theorem

The binomial theorem uses the binomial probability formula to find the probability of an event. The formula though looks a little complicated initially for the students, it becomes easier when explained in relation to the above activity since it connects the Pascal's triangle and combination formula

$$_nC_xP^xQ^{n-x} = n!/((n-x)!x!)P^xQ^{n-x}$$

n represents the number of trials(3), x the number of success(1 - getting exactly one five), P the probability of a success(1/6) and Q the probability of getting a failure(5/6). Using the formula the probability is $3!/((3-1)!1!)(1/6)^1(5/6)^2$ which is approximately 0.347. Students should be able to therefore recognize that there are 3 ways to get one 5 when 3 coins are flipped ($_3C_1$ is 3, the probability of success is 1/6 and failure is 5/6). To check their understanding, students could find the outcome from above what would the probability be in getting 2 successes (combinations of 2 five's) or 3 failures (no five's at all)

Example 2: Using the example from lesson 1 but with the probability aspect is a way for students to analyze and connect.

Find the probability that in a family of seven kids there will be different groups with at least 4 girls. Consider the probability of a girl as success. (Hint – the total probability of 4, 5, 6 and 7 girls are to be found since it is at least 4 girls).

Class-work: Problems from the text-book will be assigned for student to practice Holt: pages 745 -747

Binomial Probability Distribution

In this section students will be introduced to the basics of a distribution and in particular probability distribution. Using real life examples they will learn to identify between discrete and continuous variables as this helps to appreciate and see the difference between a binomial probability distribution and a normal distribution (the next lesson in this unit). Since the following facts need to be stressed the lesson will include a presentation for students to take down notes as part of structured note taking.

The binomial probability distribution is a distribution of data based on a binomial experiment hence there are only 2 outcomes success or failure. The main issue in a binomial experiment is to find the probability of successes out of n trials. Just as any distribution has a mean, variance and standard deviation so does the binomial distribution.

The features of the binomial distribution

The mean (μ) is defined as np where n is the number of trials and p the probability of successes. It is also defined as the expected number of successes. The variance is npq where q is the probability of failures, which can be also written as (1-p) and the square root of variance is the standard deviation (σ). The significance of the mean, variance and standard distribution is explained using the 2 examples below. The second example deals with the "how do you know" aspect.

Example 1: In Pittsburgh, PA 56% of the days in a year is cloudy. Find the mean, variance and standard deviation for the number of cloudy days during the month of June. Interpret the results and determine any unusual values

Solution:
N=30, p=.56, q=.44
Mean (μ) = np = 30x.56 =16.8
Variance= npq = 30x.56x.44 = 7.4
SD (σ) = squareroot(7.4) = 2.7

Interpretation:
16.8 days in June are cloudy which is based on the 56% of the cloudy days in a year. The standard deviation is 2.7. Hence values more than 2 standard deviations from the mean are considered unusual. Because 16.8 -2(2.7)=11.4 and 16.8 +2(2.7)= 22.2, a June with less than 12 cloudy days or with more than 22 cloudy days would be considered unusual.

Example 2: A pharmaceutical company claims that a new treatment is successful in reducing fever in more than 60% of the cases. The treatment was tried on 50 randomly selected cases and 18 were successful. Do you think that the company's claim is valid? How do you know? Solution: If the claim is valid, then p (the number of successful cases) has a binomial distribution with n = 50 and has a probability P greater than or equal to 0.60 We will first consider the boundary case, P = 0.6. Is p = 18 a likely outcome from n when P= 0.6.

Solution:

Mean = 50(.6) = 30 this tell us the expected number of successful cases is 30
Variance = 50(.6)(1 - .6) = 12
Standard deviation = squareroot(12) = 3.46

One deviation either way from the mean would be
30 -3.46 = 26.54 or 30 + 3.46 = 33.46

Two deviations either way from the mean would be
26.54 -3.46 = 23.08 or 33.46 + 3.46 = 36.92

Three deviation either way from the mean would be
23.08 – 3.46 =19.62 or 36.92 + 3.46 = 40.38

The observed value, p = 18 is less than 19.62 and not within 3 of its mean, thus 18 is unlikely to be observed when P= 0.6. This gives the students the concept of how far a value is from mean or normal range.

In the above example though the students have not yet been taught z-score this gives them an idea when it is dealt with in the next lesson.

An alternative approach: Using the probability formula

Compute the probability of observing a value as small as or smaller than 18, assuming that P = 0.6 and n=50. If the probability is large, do not doubt the claim. If the probability is small, doubt the claim.

n = 50 and P = 0.6

Using the binomial distribution formula for x = 1 to 18

$_{n}C_{x}P^{x}Q^{n-x} = n!/((n-x)!x!\,)P^{x}Q^{n-x}$

$_{50}C_{18}P^{18}Q^{50-18} = 50!/((50-18)!18!\,)*0.6^{18}*0.4^{32}$

The sum of the probabilities for x= 1 thru 18 = 0.0005145

We thus obtain that P(X = 18) = 0.0005145. The probability is very small. Therefore, we doubt the claim. (Note: It is incorrect to just compute the probability at 18 since that is usually very small if sample size is large

Graphing a binomial distribution

The purpose of this part leans more towards gaining a conceptual understanding of the distribution and what it means when a graph is skewed. The web is ideal to achieve this since the applets visually explain the concepts.

Students will access the websites below to see the graph of a binomial distribution.
http://davidmlane.com/hyperstat/probability.html
http://www.ruf.rice.edu/~lane/stat_sim/binom_demo.html
The applet allows the user to give different values for n and p and it calculates the probability for the specified interval. (The applet works better when n is less than 40). By giving a range of probabilities between 0 and 1, students discover the difference between a right skewed graph and a left skewed graph and apply the interpretations to problem solving

Class Exercise and Homework
Student will work on pages 803, 804 and 80 from the text-book (Holt, Rinehart, Winston

Lesson 3: *Normal Distribution*

At the end of this lesson students will be familiar with the characteristics significance and properties of the normal curve. Students will get to be familiar with the term 'standardizing scores' and 'percentiles' and will calculate the z-scores for a given mean and standard deviation. They will learn to apply the z-score to find the probability of occurrence within one or two standard deviations or within a specified z-score

At this point it is presumed students have some knowledge of mean, median, standard deviation, histogram, probability and percentages.

My focus in this lesson is for students to get a conceptual understanding of normal distribution and its significance in interpreting test scores, therefore the strategy will be reciprocal teaching partnered with technology – internet. Students will work in groups and through the process discover the concept of percentiles and z-scores. The process requires them to connect using previous knowledge and for the computation and graphing they would use technology. The emphasis will not be on the algebraic skills for computation.

To begin with students will be given a data sheet. There would be two sets of data and every student gets to pick one. The two sets of data are weight of a nickel with the frequency and heights of women with the frequency. Using the data each student will individually find the mean, median and mode and record it on the data sheet.

On completion students with the same data will group in three's and there will be a class discussion on the answers each group got after they have discussed the answers among themselves. The second part of the activity is to draw the histogram to get an idea of the shape of the graph that emerges. It can be done on a graph sheet but since they have access to computers they will draw the histogram using the histogram generator on the website.
http://people.hofstra.edu/Stefan_Waner/RealWorld/stats/histogram.html

They will individually enter the data and the frequency and generate the graph. Following this will be a short presentation on the properties, characteristics, significance and parameters of the normal curve making connections with the activity above

Properties of a Normal Distribution

A normal distribution is a continuous probability distribution for a random variable x. The graph of a normal distribution is called the normal curve. A normal distribution has the following properties

a. Mean, median and mode are equal
b. Normal curve is bell shaped and symmetric about the mean
c. Total area under the normal curve is equal to one
d. The normal curve approaches but never touches, the x axis as it extends farther and farther away from the mean.
e. The points at which the curve changes from curving upwards to curving downwards are called inflection point
f. 68% of the area is within 1 standard deviation
g. 95 % of the area is within 2 standard deviations
h. 99% of the area is within 3 standard deviations

The parameters that control a normal curve are the mean and the standard deviation. Mean locates the balance point and the standard deviation determines the extent of the spread. The formula used to generate the normal distribution curve is called the normal density function curve and is given by

$$y = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Though the formula is not used explicitly in this unit its significance in finding the areas under the curve will be mentioned. The significance of the graph of the normal distribution is the portion of the area under the curve and above a given interval represents the probability that a measurement will lie in that interval

After the presentation, students will continue on the activity. They will draw the curve obtained from the histogram and mark along the horizontal axis the point that corresponds to mean, one standard deviation above and below the mean, two standard deviations above and below the mean, three standard deviations above and below the mean and answer the next set of questions

1. What percentage of the data is within one standard deviation of the mean?
2. What percentage of the data is within two standard deviation of the mean?
3. If a student is picked at random what is the probability her height would be within 2 standard deviations of the mean?

Students will be led to connect the answers to the above question with the last 3 properties (68%, 95%) which leads into finding the z score and its the definition as the number of standard deviations the value x is from the mean will be elicited from them

z-score = (mean - value)/ Standard deviation

$= (\mu - x)/\sigma$

The probability for normal distribution will be explained using this website

http://people.hofstra.edu/Stefan_Waner/RealWorld/stats/normaldist.html

A short explanation on how the computer calculates the probability with reference to the calculation of the area under the curve using the probability distribution function formula will be given. In addition a mention will be made on the table of values for the different areas as an alternative to calculating on the TI-83

To check for student understanding of the lesson they will work in groups on the following examples and each group will come forward and share their answers and a discussion following this will help evaluate their understanding

Example 1: Reading test scores

Scores on the verbal section of the SAT are approximately normally distributed with a mean of 500 and standard deviation of 100.
a. What percentage of students score between 400 and 600 on the verbal section of the SAT 1?
b. What percentage of students score over 600 on the verbal section of the SAT 1?
c. What percentage of students score less than 600 on the verbal section of the SAT 1?
d. What does it mean when a student has scored in the 85th percentile, what would be the actual score and what would be the score if the student scored in the 40th percentile?
e. What is the probability that a student selected at random will have a score of 550?
f. How is the area under the bell curve related to standard deviation?
g. What is the significance of the inflection points on the normal curve? Where do they fall?
h. What makes a distribution normal?

Example 2

How does a person know if their LDL cholesterol level is in the normal range or in the moderate risk range or in the high-risk range?

To answer the above question students will access the website http://www.stat.wvu.edu/SRS/Modules/Normal/cholesterol.html
The applet enables students to visually understand probability and percentile in terms of the change in the area under the curve for different values of the upper limit and lower limit. Before the students begin the web activity, there would be a general class discussion on cholesterol levels and related topics such as healthy diets as this helps students appreciate the practicality and significance of the distribution.

Class Exercise
Holt: page 811 – 813

Take Home Assignment
Students will be given a day’s attendance sheet with the time at which each student scanned in for about 50 students (names will be blanked out). They would have to first get the frequency distribution with class limits and class boundaries. The next step

would be to draw the histogram for the above data (the interval is made to be continuous). Calculate the mean and standard deviation and draw inferences from the shape that emerges. Draw the normal curve with z-scores. Find the probability that a student comes 30 minutes late and the probability of a student coming in 20 minutes early.

Annotated Bibliography

Teacher References

Understanding Basic Statistics : Charles Henry Brase, Corrinne Pellillo Brase
(Good book for understanding the concepts and has a variety of problems)

Advanced Mathematical Concepts : Merill Publications
(The binomial theorem using Pascal's triangle is well explained)

Finite Mathematics : Lial, Greenwall and Miller
(*Problems on Normal Approximation to Binomial Distribution*)

Student References

Elementary Statistics: Ron Larson, Betsy Farber
(*Explains most concepts at the student level with graphics*)

Algebra 2 : Holt, Rinehart, Winston
(*Recommended text book followed in class*)

Websites

http://www.mathwords.com/s/scalar.htm :
For math definitions.
http://www.mathsisfun.com/pascals-triangle.html
http://mathforum.org/workshops/usi/pascal/pascal_elemdisc.html
http://www.andrews.edu/~calkins/math/edrm611/edrm04.htm
http://www.intmath.com/Counting-probability/12_Binomial-probability-distributions.php
http://davidmlane.com/hyperstat/probability.html
http://www.ruf.rice.edu/~lane/stat_sim/binom_demo.html
http://wiki.stat.ucla.edu/socr/index.php/SOCR_EduMaterials_Activities_Normal_Probability_examples
http://ptri1.tripod.com/#made
http://www.pbs.org/teachers/math/
http://www.ed.mtu.edu/esmis/id321.htm
http://www.phschool.com/atschool/academy123/html/bbapplet_wl-problem-431949.html(teaching of the binomial probability)
http://strader.cehd.tamu.edu/Mathematics/ eStatistttics/NormalCurve
gives the z-score when the mean and standard deviation are given

Appendix / Standards

Pennsylvania Academic Standards for Mathematics

Students will have the opportunities to develop and use computation concepts, operations and procedures with real numbers in problem-solving situations. Students will use estimation to solve problems for which an exact answer is not needed. In addition, they will construct and apply mathematical models.

Students will use a variety of technological and information resources to gather and synthesize information, and to create and communicate knowledge. They will demonstrate skills for using graphing calculators in a variety of operations.

2.2 Computation and Estimation
Basic functions (+, -, □, ÷), Reasonableness of answers, Calculators

2.3 Measurement and Estimation
Types of measurement (e.g., length, time)
Units and tools of measurement,

2.4 Mathematical Reasoning and Connections
Using inductive and deductive reasoning

2.5 Mathematical Problem Solving and Communication
Problem solving strategies
Representing problems in various ways
Interpreting results

2.6 Statistics and Data Analysis
Collecting and reporting Data (e.g., charts, graphs)
Analyzing data

2.8 Algebra and Functions
Equations, Patterns and functions