ASA AMERICAN STATISTICAL ASSOCIATION

NCTM NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS

The Statistics Teacher Network

www.amstat.org/education/stn/index.html

Number 71 | **ASA/NCTM Joint Committee** on the **Curriculum** in **Statistics** and **Probability** | Winter/Spring **2007**

Poster and Project Competitions Need Entries

Introduce students to the world of statistics through the 2007 poster and project competitions. The time has come to submit entries for the American Statistical Association Poster Competition and Project Competition. The competitions, now in their 18th and 21st years, respectively, offer opportunities for students to formulate questions and gather and display statistical data while drawing conclusions from that data. Posters are judged in four grade-level categories—K–3, 4–6, 7–9, and 10–12—while projects are judged in three categories—4–6, 7–9, and 10–12. The deadline for both competitions is April 1, 2007. Winners will be recognized with trophies, cash prizes, certificates, ribbons, and calculators (donated by Texas Instruments). ASA members or representatives will personally present the prizes to the winners at their schools.

For more information and entry forms, visit the K–12 area of *www.amstat.org/education*.

Update on GAISE

From the ASA/NCTM Joint Committee

The entire STN winter 2006 issue number 68 was devoted to highlighting a hallmark document entitled GAISE – Guidelines for Assessment and Instruction in Statistics Education, A Pre-K–12 Curriculum Framework. The foundation for the Framework rested on the National Council of Teachers of Mathematics' Principles and Standards for School Mathematics (2000). The Framework is intended to support and complement the objectives of the NCTM PSSM, not to supplant them.

The ASA/NCTM Joint Committee reviewed the GAISE document with the authors and other interested statistics educators. As a result, an updated version has been produced. A bound volume is available through the ASA online marketplace as well as at *www.amstat.org/education/gaise/*.

GAISE was written to help educators work toward developing statistically literate citizens who can use statistics to make reasoned judgments, evaluate quantitative information, and value the role of statistics in everyday life. It was also written to provide stakeholders such as writers of state standards, writers of assessment items, educators at teacher preparation programs, curriculum directors, and Pre-K–12 teachers with guidance in developing standards in data analysis and statistics as part of the Pre-K–12 mathematics curriculum.

It is a must read for all who are involved with statistics education at the school level.

Also In This Issue…

K-12

THIS ARTICLE IS RATED

I have helped grade (along with several hundred other teachers) Advanced Placement Statistics exams. If you are an elementary or middle school teacher, I hope you will read this article because it contains suggestions about how K-8 teachers can contribute to improving the understanding of probability and data analysis of younger children. This article ties lessons I learned about AP statistics to ideas and skills for teachers in all grades. I have also recently reviewed items from high school "exit tests." Many of the skills taught in an AP Statistics course are also needed to pass these exams. Read on.

Beth Lazerick, STN editor

What I Learned Grading AP Statistics Exams

The Inside Story!

Understand, Explain, Communicate

The following is problem 2 from the free-response portion of the 2005 AP Statistics Exam. For the complete problem and solutions, visit the College Board website at http://apcentral.collegeboard.com.

2. Let the random variable X represent the number of telephone lines in use by the technical support center of a software manufacturer at noon each day. The probability distribution of X is shown in the table below.

x	0	1	2	3	4	5
$p(x)$	0.35	0.20	0.15	0.15	0.10	0.05

(a) Calculate the expected value (the mean) of X.

(b) Using past records, the staff at the technical support center randomly selected 20 days and found that an average of 1.25 telephone lines were in use at noon on those days. The staff proposes to select another random sample of 1,000 days and compute the average number of telephone lines that were in use at noon on those days. How do you expect the average from this new sample to compare to that of the first sample? Justify your response.

(c) The median of a random variable is defined as any value x such that $P(X \leq x) \geq 0.5$ and $P(X \geq x) \geq 0.5$. For the probability distribution shown in the table above, determine the median of X.

(d) In a sentence or two, comment on the relationship between the mean and the median relative to the shape of this distribution.

To successfully respond to this question, students must

- Understand grouped data
- Write no "naked"answers
- Explain how to obtain a result, and
- Communicate conclusions appropriately

1. Understand grouped data

In this problem students had to understand that the data represent the number of telephone lines in use at noon and the probability that each number of lines will be in use at noon. The expected value is $\sum x_i \bullet p_i$. Students must understand that expected value is just an average.

2. Write no naked answers and explain how you have come to this conclusion

On the exam, the answer to part (a) of this problem is 1.6 telephone lines. Many students just write "1.6" without the units. Understandably, students are not doing this level of work in the early grades, but they can be held accountable for attaching appropriate units to answers to problems or exercises.

3. Explain how to obtain a result

To answer part (c) students had to calculate the median number of lines in use at noon. In this problem the correct answer is "The median number of lines in use at noon is 1 line. This is the median because at least 50% of the data falls at or above one line and at least 50% falls at or below one line." Encourage younger children to justify their responses to problems and to defend their ideas.

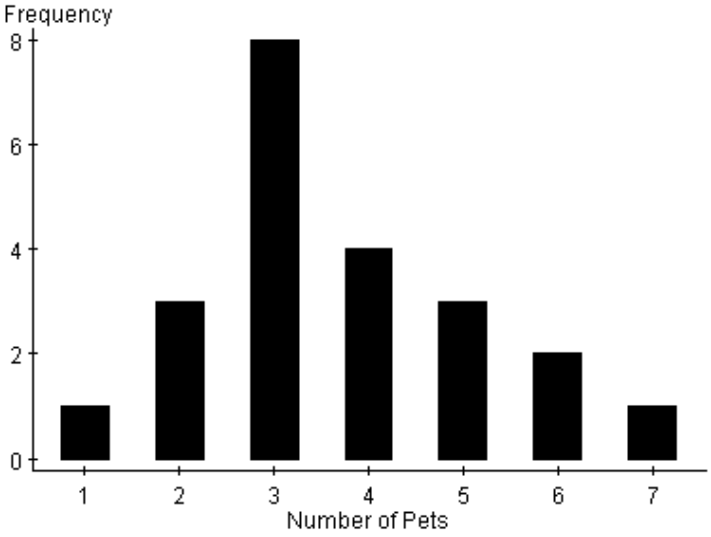
4. Communicate conclusions appropriately

To answer part (d) a student must retrieve knowledge about the relationship of the mean to the median in skewed data and respond that the mean in larger than the median, as is typical of a right-skewed distribution.

Early Years Extensions and Implications

Just how does an AP Statistics problem relate to the early years for students? Consider this scenario: Have students create a picture graph illustrating the number of pets each student has at home. (Students will probably need to discuss how to count fish!) Teachers may want to create this graph on the whiteboard using a sticky-note for each student's response. That way you can change a student from one column to another easily. *For example:*

Number of Pets



1. Understand grouped data

Discuss with students how many total pets the class as a whole owns. Calculate the mean, median, and modal numbers as well as the range of number of pets if your students are ready for these ideas. The youngest students can certainly find the "middle" number if they string out the sticky-notes from least to greatest and look for the middle person. The mean number can be estimated by "evening" off the stacks and seeing what number comes up. The range or spread can be easily calculated. The objective with this activity is to "see" the meanings of these statistical terms.

2. Write no "naked" answers

As students answer questions about the bar graph be sure that they use "descriptors" with each response. "The median number of pets is three pets," not just the number/numeral "3."

3. Explain how to obtain a result and

4. Communicate conclusions appropriately

When responding to the question, "What is the median number of pets owned by the students in the class?" Students should answer, "The median number of pets is three pets because at least half of the students own at least three pets and at least half own three or fewer pets."

Middle Years Extensions and Implications

Older students might relate to graphing the number of music CDs they own or the number of pairs of shoes a female member of their family possesses! Discuss calculating the range, mean, median, and mode of the number of CDs students own from a split stem-and-leaf plot as follows:

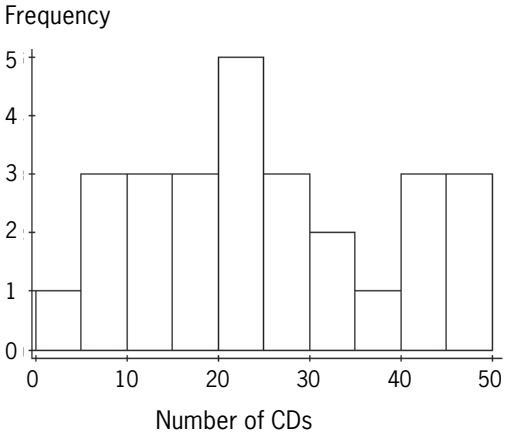
Stem-and-Leaf: Number of CDs

0	4
0	899
1	123
1	567
2	01112
2	589
3	12
3	7
4	013
4	679

Range = 49-4 = 45
Mean = (4+8+9+...+46+47+49)/27 = 24.7
Median = 21
Mode = 21

1. Understand grouped data

When you have data in the following form without knowledge of the actual values,



you cannot calculate the exact range because you do not know the minimum and maximum values. You can calculate an approximate value of the mean for this grouped data by using modified summation notation after discussing what the "middle" of each group would be. For instance, the middle of the first bar is 2. 5 CDs. The following calculation estimates the mean for this grouped data:

Mean ≈ [1(2.5) + 3(7.5) + 3(12.5) + 3(17.5) + 5(22.5) + 3(27.5) + 2(32.5) + 1(37.5) + 3(42.5) + 3(47.5)]/27

2. Write no "naked" answers

Again, be certain that students answer using appropriate units.

3. Explain how to obtain a result, and
4. Communicate conclusions appropriately

Students need to explain how they determined each step. Teachers who are used to using journals or writing about mathematics will find this routine. If students are not used to writing about mathematics, teachers may want to enlist the help of the English departments to encourage students to write coherently and to use appropriate descriptive language. Teachers will need to grade communication routinely in order to emphasize the idea that they value it.

Compare Does Not Mean Describe!

The following is a portion from problem #1 from free-response portion of the May 2004 AP Statistics Exam.

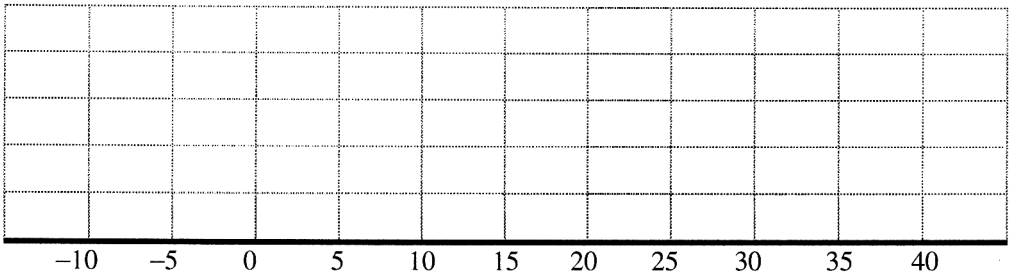
1. A consumer advocate conducted a test of two popular gasoline additives, A and B. There are claims that the use of either of these additives will increase gasoline mileage in cars. A random sample of 30 cars was selected. Each car was filled with gasoline and the cars were run under the same driving conditions until the gas tanks were empty. The distance traveled was recorded for each car.

   Additive A was randomly assigned to 15 of the cars and additive B was randomly assigned to the other 15 cars. The gas tank of each car was filled with gasoline and the assigned additive. The cars were again run under the same driving conditions until the tanks were empty. The distance traveled was recorded and the difference in the distance with the additive minus the distance without the additive for each car was calculated.

   The following table summarizes the calculated differences. Note that negative values indicate less distance was traveled with the additive than without the additive.

Additive	Values Below Q_1	Q_1	Median	Q_3	Values Above Q_3
A	−10, −8, −2	1	3	4	5, 7, 9
B	−5, −3, −3	−2	1	25	35, 37, 40

(a) On the grid below, display parallel boxplots (showing outliers, if any) of the differences of the two additives.



To successfully respond to this question, students must

- Write using appropriate words of **comparison**
- Label all graphs appropriately
- Understand the difference between **data** and **statistics**
- Define how outliers and other important numbers are calculated

1. Write using appropriate words of comparison

If the question asks students to compare two sets of data, a graph or any other two things, they must write about the data with words like, "The median of one distribution is greater than the other. The second set of data is skewed to the right, while the other is symmetric." Compare ALL given sets of data using

Shape (symmetric or skewed)
Outliers (name them or state there are none)
Measure of center (mean or median and possibly mode), and
Spread (standard deviation, variance, interquartile range or range)

The following acronym is often used to remind students to do this:

SOCS: Shape-Outliers-Center-Spread

2. Label all graphs appropriately

Students must label each box-plot with the correct identifier "A" or "B." Students should also always give a title to

their graphs and identify the units for the horizontal axis, in this case "difference in distance traveled with additive A minus additive B."

3. Understand the difference between data and statistics

In this problem, the values for the first and third quartile and the median are statistics. Because, for instance the "3" for the median of A could actually be between "real" values of 2 and 4, this "3" is a statistic. However the –10, –8, and –2 are actual readings or data. Many students who answered this question simply added the nine give numbers and divided by nine to get a faulty average. Be sure that your students know the difference between data and statistics.

4. Define how outliers and other important numbers are calculated

Students should write down how they determine which values fall outside the overall pattern of the data. Many texts indicate an obervation is an outlier if greater than **Q3** plus **1.5 (IQR)** or if less than **Q1** minus **1.5 (IQR)**.

Early and Middle Years Implications and Extensions

1. Write using appropriate words of comparison

Have students write at their own level a description of themselves. This can be done as a language arts lesson. Alternately have them describe one of the original 13 Colonies of the USA or the voyage of one of the early explorers. Then have them each compare himself or herself or a colony or an explorer to another similar description. Emphasize use of comparative words such as "taller than," "farther than," "more than," etc. Continue this type of language when the class explores graphing in mathematics or during any other mathematics lesson.

2. Label all graphs appropriately

Include labeling graphs with titles and units or descriptors on each axis as part of every graphing assignment.

3. Understand the difference between data and statistics

When students begin to explore the ideas of mean and median, be sure that they notice that sometimes there is no "middle" number and that they are taking an average of two pieces of "real" information. When they find the average of their spelling test scores, this number is usually not one of their actual scores. Use the words data and statistics as students become increasingly aware of terminology.

4. Define how important numbers are calculated

Students should begin to articulate how statistics are calculated. The use of appropriate terminology, "I added up all my spelling test scores and divided by the number of scores I had, to get my average," is a start. "The median score was the average number between the third and fourth scores out of my six scores ordered in increasing value," will help students begin to talk about statistics appropriately.

Insist on Correct Notation and Close Reading of Problems and Examples

The following is a portion of problem #6 from the free-response portion of the 2006 AP Statistics Exam.

6. A manufacturer of thermostats is concerned that the readings of its thermostats have become less reliable (more variable). In the past, the variance has been 1.52 degrees Fahrenheit (F) squared. A random sample of 10 recently manufactured thermostats was selected and placed in a room that was maintained at 68°F. The readings for those 10 thermostats are given in the table below.

Thermostat	1	2	3	4	5	6	7	8	9	10
Temperature (°F)	66.8	67.8	70.6	69.3	65.9	66.2	68.1	68.6	67.9	67.2

(a) State the null and alternative hypotheses that the manufacturer is interested in testing.

It can be shown that if the population of thermostat temperatures is normally distributed, the sampling distribution of $\frac{(n-1)s^2}{\sigma^2}$ follows a chi-square distribution with $n-1$ degrees of freedom.

For this question, students must
• Understand and use appropriate notation
• Recognize that if a particular phenomenon "follows" the Chi Square distribution that the task is NOT necessarily a typical Chi Square question
• Answer each question completely and in the manner described

1. Understand and use appropriate notation

In this questions students should immediately recognize s^2 is the symbol for sample variance. They need to calculate the variance from the given data set. Since standard deviation and variance are treated early in the course, repeated reference to this notation is likely.

2. Recognize that if a particular phenomenon "follows" the Chi Square distribution that the task is NOT necessarily a typical Chi Square question

Students are rarely introduced to the idea that a phenomenon or statistical analysis may "act" like something they are familiar to but not be that thing. Many students just assumed that since the data followed a Chi Square distribution, they should do the Chi Square test for which they are most familiar. In this case the data is interpreted using the Chi Square table of values but not using the formula for the Chi Square test that is introduced in the AP statistics curriculum.

Early and Middle Years Extensions and Implications

1. Understand and use appropriate notation

In language arts and writing classes students learn to capitalize the first word in a sentence and place a period at the end of a sentence. This attention to detail should extend to mathematics as well. When students begin to look at perimeter and learn that

P = 2 length + 2 width and later that P = 2*l* + 2*w*, stress that the "P" is a capital and the *l* and w are lower case and in script to eliminate the confusion between "1" and "*l*" (numeral "one" and the letter "*l*").

Teachers should insist upon correct notation at all times. For $A = \pi r^2$, the "A" should be capitalized and the "r" lowercase. No exceptions should be made when this is taught. (In statistics, Greek letters are usually used for information about a population and English letters are usually used for information gathered from a survey or experiment about a subset of the population.)

2. Answer each question completely and in the manner described

Teachers frequently ask two-part questions and questions in which interpretation is needed.

The length of a rectangle is 20 feet. The perimeter of this rectangle is 50 feet. What is the area of the rectangle? Show all work.

For this question a diagram could be drawn, the width of the rectangle determined and then the area calculated. All parts should be evaluated for the top score for the problem.

The ideas presented in this article are not intended to be comprehensive. Hopefully they will give rise to an understanding that the seeds to the successful completion of the exit goals for data analysis and probability for our high school exams are planted in the elementary and middle school classrooms. High school teachers usually have concepts of probability and statistics embedded in their curriculums, although these concepts are often ignored because of "time" constraints.

How do you incorporate these ideas into your teaching? Let the editors know and we would love to publish your great ideas and activities.

Meeting Within a Meeting

ASA Program for Math/Science Teachers

Based on the Utah Secondary Core Curriculum Standards for Statistics and Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report A Pre-K–12 Curriculum Framework

Sponsored by the American Statistical Association
2007 Joint Statistical Meetings

Date
Monday, July 30, 2007

Place
Salt Palace Convention Center, Salt Lake City, Utah

Audience
Middle School/Junior High Math and Science Teachers

Submit applications, letter of intent, and letter from principal/district supervisor

Review and acceptance of applications to attend MWM **will begin March 1, 2007** and will continue until the course is filled.

See *www.amstat.org/education/mwm* to apply online.

Objectives
Enhance understanding and teaching of statistics within the math/science curriculum based on conceptual understanding, active learning, real-world data, and appropriate technology

Content
Based on standards, assessments, and textbooks

Consistent with GAISE recommendations *(www.amstat.org/education/gaise),* NCTM *Principles and Standards for School Mathematics,* and *Curriculum Focal Points for Prekindergarten through Grade 8 Mathematics (www.nctm.org/focalpoints)*

Format
One-day meeting

Activity-based sessions, including lesson plan development

Lunch with Utah ASA Chapter members and dinner with career panel discussion

Optional activities on Tuesday (Statistics Education Sessions, Poster Sessions, JSM Exhibit Hall)

Provided
Registration cost

Meals and refreshments

Lodging Monday night for local teachers

Lodging Sunday and Monday nights for teachers from outside the Salt Lake City area

Handouts

Certificate of participation from the ASA

Possible professional development half-credit for 7–13 contact hours

Follow-up
Program evaluation for future follow-up workshops

Development of listserv for participants

ASA Chapters network with local teachers to organize learning communities

Contact
Rebecca Nichols, *rebecca@amstat.org;* (703) 684-1221, Ext. 1877
See *www.amstat.org/education/mwm* for more information and to apply online.

Dear Subscribers,

We are looking for good articles for the next issue of *STN*. We may want to highlight successful classroom practices. If you have had a great experience teaching probability and/or data analysis, let us hear from you.

To subscribe to the *STN Newsletter*, please send your email and snail mail addresses to *Rebecca@amstat.org*. ALL NEW SUBSCRIBERS will receive the newsletter electronically unless they specify otherwise. IF YOU ARE CURRENTLY RECEIVING A PAPER COPY AND ARE WILLING TO RECEIVE YOUR NEWSLETTER ELECTRONICALLY, please let Rebecca know by sending her your current address and email address also. In this way we can expand the subscriber base and decrease costs.

Beth Lazerick and Murray Siegel would like to "retire" and we are looking for one or two people to become the new editor or co-editors to take STN into a new era. If interested, please contact Jerry Moreno at *moreno@jcu.edu* or 216-397-4681.

Thanks,

Beth Lazerick, *stneditor@aol.com*
Murray Siegel, *siegel@gssm.k12.sc.us*

Statistics Teacher Network
c/o American Statistical Association
732 North Washington Street
Alexandria, VA 22314-1943
USA

Let us know your Zip +4!

Non-Profit
Organization
U.S. Postage
PAID
Permit No. 361
Alexandria, VA