AP® STATISTICS
2008 SCORING GUIDELINES (Form B)

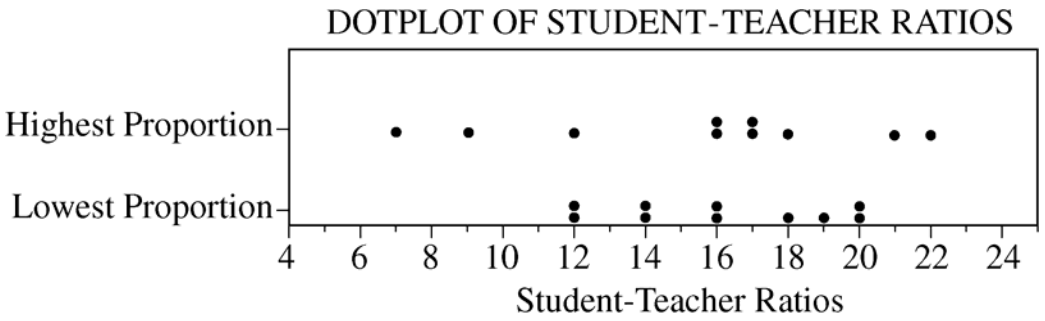
Question 1

Intent of Question

The primary goals of this question were to assess a student's ability to (1) construct a dotplot that can used to compare two distributions and to make appropriate comparisons of the center and spread of the two distributions and (2) understand and explain why in some situations inferential procedures that generalize from samples to larger populations are not appropriate.

Solution

Part (a):



A similarity is that the two distributions are centered in approximately the same place. A difference is that the distribution for the schools with the lowest proportions of students meeting the standards is less variable.

Part (b):

The two groups of schools are not random samples from two populations of interest. One group is the population of all 10 schools with the highest proportion of students meeting the standards, and the other group is the population of all 10 schools with the lowest proportion of students meeting the standards. Therefore, inferential procedures that generalize from samples to larger populations are not appropriate.

Scoring

Part (a) is divided into three sections. Section 1 is related to the dotplots, section 2 is related to the centers of the distributions, and section 3 is related to the spreads. Each is scored as either essentially correct (E) or incorrect (I).

Section 1 is scored as follows:

Essentially correct (E) if the dotplots are given with appropriate labels and scales (minor errors in the placement of the dots are okay).

Incorrect (I) if the labels or scales are missing *OR* if a histogram with dots or a scatterplot is given.

Question 1 (continued)

Section 2 is scored as follows:

Essentially correct (E) if the response states that the two distributions have approximately the same center (no numerical values are required) *OR* if the response states that the means are slightly different.

Incorrect (I) if measures of center are not given *OR* if they are given, but there is no *comparison* of the center of the two distributions.

Section 3 is scored as follows:

Essentially correct (E) if the response states that the spread of the two distributions differs (no numerical values are required).

Incorrect (I) if measures of spread are not given *OR* if they are given, but there is no *comparison* of the spread of the two distributions.

Note: Usually distributions are described in terms of shape, center, and spread. However, with small data sets, shape is difficult to judge, so no comparison of shape is required.

Part (b) is considered as section 4 and is scored as either essentially correct (E), partially correct (P), or incorrect (I).

Section 4 is scored as follows:

Essentially correct (E) if the response states that the data are not samples from some larger population *OR* that they are not random samples but instead are those with the highest and lowest proportions of students meeting a standard, and therefore inference is not appropriate. The response must not include any other reason (such as small sample sizes or the shape of the distribution).

Partially correct (P) if the response states only that the schools are not randomly selected. The response must not include any other reason (such as small sample sizes or the shape of the distribution).

Incorrect (I) if the response states that the sample sizes are too small *OR* if the response refers to the shape of the distribution or to some other criterion.

Note: If there is information relating to part (a) in the explanation in part (b), it may be used in the scoring of part (a).

AP® STATISTICS
2008 SCORING GUIDELINES (Form B)

Question 1 (continued)

4 Complete Response

All four sections essentially correct

3 Substantial Response

Three sections essentially correct

2 Developing Response

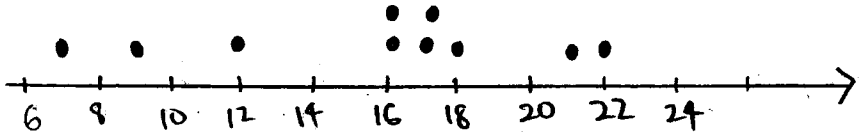
Two sections essentially correct

1 Minimal Response

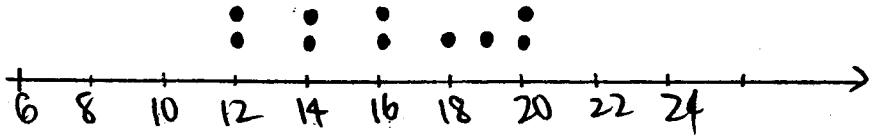
One section essentially correct

/A

(a) Display a dotplot for each group to compare the distribution of student-to-teacher ratios in the top 10 schools with the distribution in the bottom 10 schools. Comment on the similarities and differences between the two distributions.



Dotplot of distribution of student-to-teacher ratios in the top 10 schools (Dotplot A)



Dot plot of distribution of student-to-teacher ratios in the bottom 10 schools. (Dotplot B)

Similarities

Both dotplots have similar centres, with $median_A = 16.5$ and $median_b = 16$.

Differences

Dotplot A has a wider range than dotplot B;
Dotplot A has a larger variance than dotplot B.

(b) Any statistical test that is used to determine whether the mean student-to-teacher ratio is the same for the top 10 schools as it is for the bottom 10 schools would be inappropriate. Explain why in a few sentences.

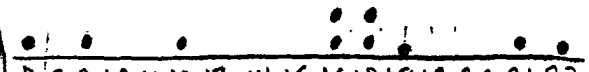
A statistical test is used to ~~dente~~ determine the differences in mean of samples from a large population.
the 10 values of "10 schools with highest proportion of students meeting standards" is the entire population, not an SRS of the population of publicschools in the state. the same goes for the bottom 10 schools. As the SRS condition is not met, any statistical test that is used to determine whether the mean student-to-teacher ratio is the same is not valid.

GO ON TO THE NEXT PAGE.

1B

(a) Display a dotplot for each group to compare the distribution of student-to-teacher ratios in the top 10 schools with the distribution in the bottom 10 schools. Comment on the similarities and differences between the two distributions.

Ratio in Top 10 Schools:

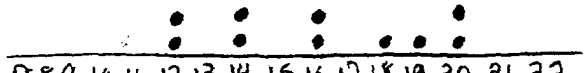


median = middle number

med = $\frac{16+17}{2}$ = 16.5

mean = 15.5

Ratio in Bottom 10 Schools:



med = $\frac{16+16}{2}$ = 16

mean = 16.1

The median student-to-teacher relationship is not much different between the two groups. The top 10 had a median ratio of 16.5, the bottom 10 had a median ratio of 16. While the Bottom 10 have a lower median-ratio, the Top 10 have a lower mean ratio (15.5 for Top 10, 16.1 for bottom 10). This so due to the very low ratios of 7 & 9. The Top 10 distribution also has a larger range (22-7=15) than the Bottom 10 (20-12=8).

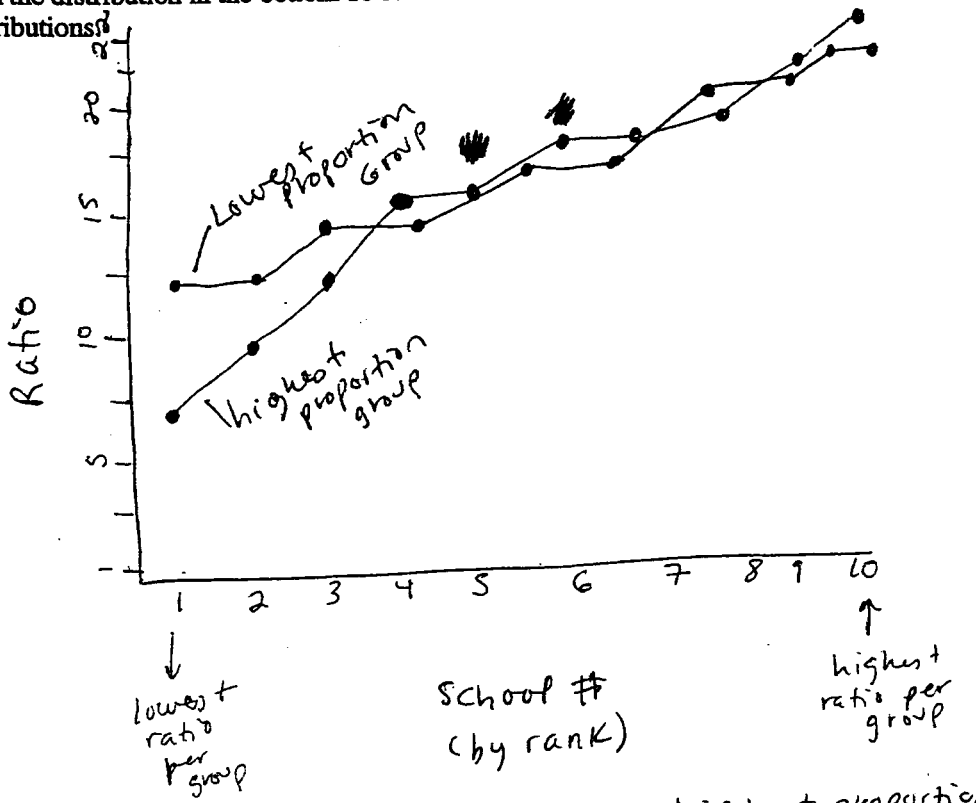
(b) Any statistical test that is used to determine whether the mean student-to-teacher ratio is the same for the top 10 schools as it is for the bottom 10 schools would be inappropriate. Explain why in a few sentences.

Any statistical test you would perform on this would require that the distributions be normal, or have a large enough sample size so that normalcy wasn't a concern (generally $n \geq 40$). These distributions have a sample size of only 10 and are definetly not normal in shape (as evident by the dotplots).

GO ON TO THE NEXT PAGE.

1C

(a) Display a dotplot for each group to compare the distribution of student-to-teacher ratios in the top 10 schools with the distribution in the bottom 10 schools. Comment on the similarities and differences between the two distributions?



The distribution of schools with the highest proportion has a much greater range than that of the schools with the lowest proportion, despite the fact that they both contain a maximum value of at least 20 (as a ratio).

(b) Any statistical test that is used to determine whether the mean student-to-teacher ratio is the same for the top 10 schools as it is for the bottom 10 schools would be inappropriate. Explain why in a few sentences.

The tests that one would think to administer could not be used here because the schools were not chosen randomly. They were selected because of their either poor or excellent performances. In order for this data to meet the assumptions of statistical tests, it would be necessary for the schools to be chosen using a simple random sample, meaning all schools have the same chance of being selected - this would avoid confounding variables.

GO ON TO THE NEXT PAGE.

AP® STATISTICS
2008 SCORING COMMENTARY (Form B)

Question 1

Sample: 1A
Score: 4

In part (a) this outstanding response includes a very readable dotplot for each group, displayed using the same scale and well labeled, which makes it easy to compare the two groups. It states clearly that the groups have "similar" centers and different spreads. (Many responses listed means/medians and standard deviations/ interquartile ranges for each distribution separately but did not compare them for the two distributions and so received no credit for commenting on similarities and differences.) The three sections of part (a) were scored as essentially correct. Part (b) correctly explains that the two groups are really populations, consisting of the "'10 schools with [the] highest proportion of students meeting standards'" and the 10 schools with the lowest proportion of students meeting the standards, rather than a random sample from each of two populations. Part (b), which corresponds to section 4, was scored as essentially correct. The entire answer, based on all four sections, was considered a complete response.

Sample: 1B
Score: 3

Part (a) includes a readable dotplot for each group, displayed on the same scale and well labeled, with appropriate comparisons between the two groups. Using the range as the only measure of variability typically is not optimal, but it is acceptable in this situation because of the absence of outliers. All three sections of part (a) were scored as essentially correct. In part (b) the student incorrectly explains that any statistical test would be inappropriate because it would "require that the distributions be normal" or the sample size be large enough. In fact, if these were random samples from two larger populations, it would be appropriate to perform a two-sample *t*-test with samples of this size, and it would be reasonable to believe that the samples came from a normal distribution. Because only part (b) was scored as incorrect, the overall answer was judged a substantial response.

Sample: 1C
Score: 2

The plot in part (a) is not a dotplot but rather a plot that orders each group of student–teacher ratios from lowest to highest. The choice of plot makes it difficult to compare the two distributions. For example, the response does not include the most important similarity: that the two distributions are centered at about the same place. It was not necessary to include exact values for the measures of center and spread, so the statement about the "greater range" was judged acceptable. The first two sections of part (a) were scored as incorrect, and the last section was scored as essentially correct. The response to part (b) contains the correct reason why a statistical test that generalizes from samples to larger populations is inappropriate. Part (b) was scored as essentially correct. The answer as a whole was deemed a developing response.