Answers to Prompts for Forum 2a: Evans Basic Statistics: Chapters 1-4

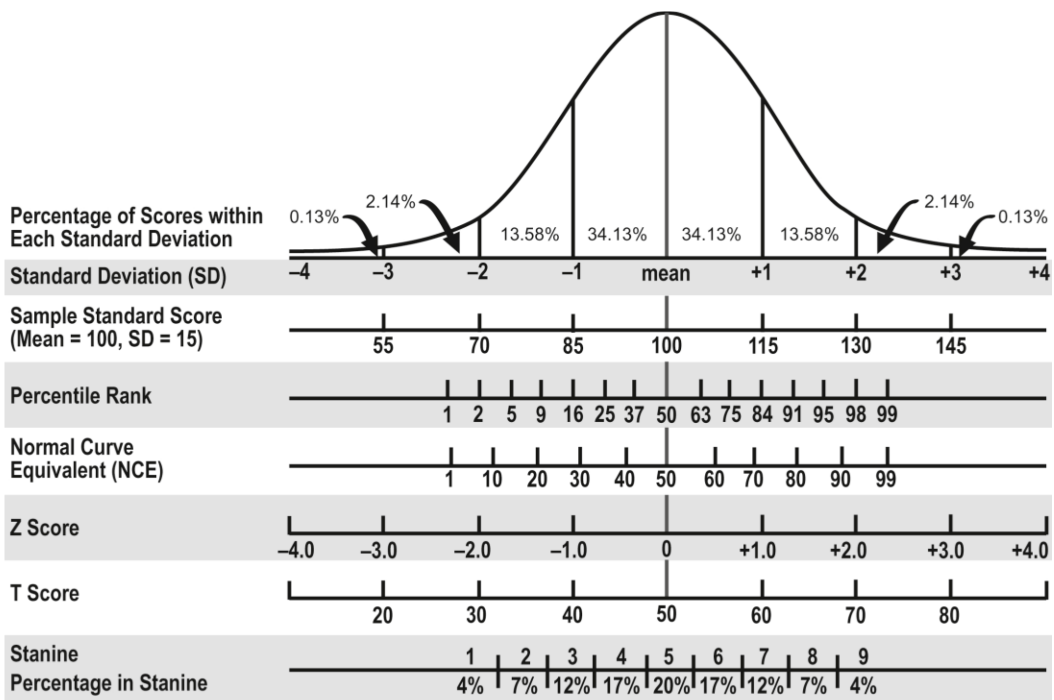
Here are a few additional exercises (and some additional instruction) to go with Chapters 1-4 (actually 2-3) in Evans' Basic Statistics Web Site. Please complete the exercises and, **briefly**, provide answers and comments on Forum 2a. You can work in groups on these exercises.

Exercise 1: When and why would you prefer the **median** over the **mean**?

When reporting statistics summarizing such phenomena as national achievement levels(or trends) districts commonly (or should commonly) use **median** national percentiles instead of **mean** national percentiles. Using the mean in these cases would have been incorrect. Can you explain why?

There are many situations where using the median as a measure of central tendency would be preferable to using the mean. One such situation, in which the median is *always* preferable is when the data are *ordinal* (i.e., ranked), as is the case with percentiles (%tiles). Using the mean with ranked data is inappropriate since the intervals between adjacent ranks are not equal.

The figure below displays a graphic of the normal curve along with several scales, including the percentile rank scale, associated with the curve.



Look at the percentile rank (PR) scale. Notice that the PRs toward the middle of the scale tend to bunch up, or be close together, while those toward the ends of the scale are further apart (there is little separation among PRs in the neighborhood of 50, yet there is clear separation PR between 1 and 25, and 75 and 99).

There are other situations where the median would be preferred to the mean. One of the more obvious situations is when the data are highly skewed, i.e., when there are extreme values (or *outliers*) on one side or the other of the score distribution. For instance, suppose we took a sample of CEOs from Charlotte and wanted to compute their average annual salary (including bonuses). While it is likely that nearly all the CEOs in the sample would have high salaries, only a comparative few, like the CEO of the Bank of America, would have extremely high salaries. So computing the mean CEO salary, while including the few having multi-million dollar salaries, would be misleading since the salaries of those individuals would pull the mean to the right (i.e., toward the high end.) Instead, a better measure of the average salary of CEOs would be the median salary.

Exercise 2: Can you explain, in simple terms, what the **standard deviation** tells us?

In Evan's Lesson 2 you were introduced to the **standard deviation**. Here is another, layman's definition of the standard deviation.

The standard deviation describes how far, *on average*, a score drawn at random from a given distribution of scores will *deviate* from the *mean* of that distribution.

Does this definition help? How do you interpret it?

We need to be careful about over-interpreting the standard deviation. It is just one measure of variability that helps describe a distribution of scores and is usually reported along with the mean (when the median is used as the measure of central tendency, the appropriate measure of variability is the *range* (highest score minus the lowest score.)

The larger the standard deviation, the more spread-out the distribution of scores (hence, the further away from the mean a score drawn at random is likely to deviate.)

Exercise 3: Interpret and use **standard scores**.

In Evans' discussion of the **normal curve** (or, more appropriately, normal distribution) he did not talk about **standard scores** (or ***z* scores**), even though this is where they would normally be introduced. So, here, I will remedy that omission.

Standard scores (*z* scores) play an important role in measurement and statistics. Given a person's *z* score, we can easily determine that individual's percentile rank. The *z* score also helps us determine how likely or unlikely (rare) an individual's score is in a particular population. First, however, let's see how we compute a *z* score.

This is done using the following simple forumula:

$$z = \frac{(X - M)}{SD},$$

where X is the score to be converted to a z score, M is the mean of the distribution (of scores), and SD is the standard deviation of the distribution of scores.

For instance, suppose X = 37, M = 32, and SD = 5. Then,

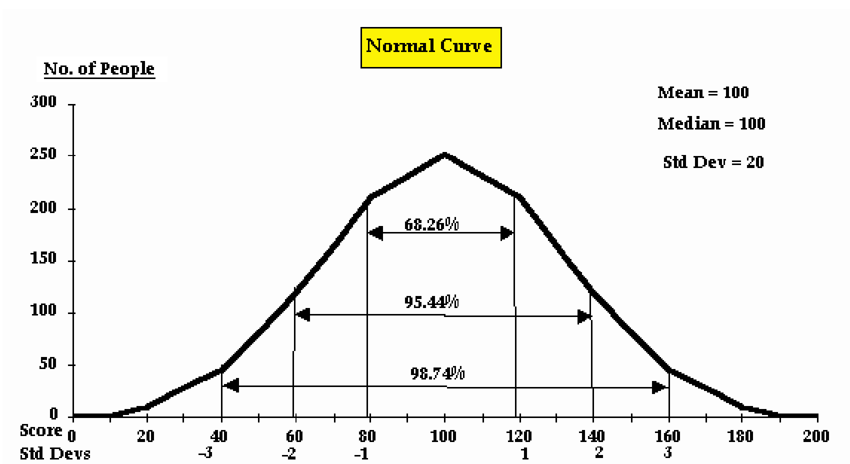
$$z = \frac{(37 - 32)}{5} = 1.$$

As another example, in the same distribution, X = 27. In this case,

$$z = \frac{(27 - 32)}{5} = -1.$$

What do these z scores tell us?

Look at Evans' Figure 6, where he shows the percent of cases lying within various standard deviations of the mean:



The bottom row of the figure, the one labeled Std Devs, are also called z scores. So we see immediately, that a z score tells us how far, in standard deviation units, the corresponding scores (X) is from the mean. In our case, the first z score (1) is one standard deviation above the mean. Look at the figure. What percentage of the cases is below a z score of 1? Realizing that 50% of the cases are below a z score of zero (the median of the distribution) we can determine that half the cases in the interval described by -1 and +1 z (or Std Dev) equals roughly 34% (68%/2) . So, the percent of cases below a z score of +1 is 50% + 34% = 84%. Since a percentile rank is defined as the percent of cases below a given raw score, X, we conclude that a score of 37, in our distribution, is equivalent to the 84th %tile. Similarly, a score of 27, which in our distribution is equivalent to a z score of -1 is at the 16th %tile (I'll leave it to you to figure this out.)

For this prompt, use the distribution having M = 32 and SD = 5 to compute the %tiles for the following scores (rough approximations are OK): 20, 25, 30, 35, 40, and 45. Which of these scores would you consider to be rare in the distribution?

I'm sorry, but I kind of misled you, somewhat, about percentiles. Actually, percentiles (%tiles) are the score-equivalents of percentile ranks (PRs). For example, in the exercise as score of 25, if you computed it correctly, is found to be equivalent to the 8th PR. The score, 25, itself is the *percentile* corresponding to the 8th PR. Unfortunately, the terms, percentiles and percentile ranks, tend to be used interchangeably in much of the literature. I will try to keep the terms distinct, but as you have seen already, I too occasionally mix them up.

There are actually *two* definitions for percentile ranks for a given score, X:, (1) the percent of cases that fall *below* X, and (2) the percent of cases that fall *at or below* X. The first definition is the one that is typically used.

To compute the PR using definition (1), use the equation,

$$100 \times \left(\frac{Number\ below\ X}{Number\ of\ Caes} \right)$$

To compute the PR using definition (2), use the equation,

$$100 \times \frac{(Number\ below\ X + Number\ equal\ to\ X)}{Number\ of\ Caes}$$

Determining which scores should be considered rare depends. Some would consider rare any score that occurs 5% or less of the time; others would consider rare any score that occurs 1% or less of the time. Using the first determination, scores of 20, 40, and 45 would be called rare; using the second determination, only scores of 20 and 45 would be considered rare.

It is customary to report percentile ranks in whole numbers. Hence, using definition 1 the corresponding PRs are.

20—1st PR,
25—8th PR,
30—34th PR,
35—73^{ed} PR,
40—95th PR,
45—99th PR*.

*Note: typically, a PR of 1 is never reported. PR typically are reported in the range 1 to 99. Hence, even though a score of 45 equates to a *z*-score of 2.6, which equates to a PR of 99.53—which when rounded is 100—we still report it as a PR of 99.

Also, using the same distribution of scores, figure out the following:

What percent of scores fall between 25 and 35; between 35 and 40; and between 40 and 45?

Suppose that you consider "rare" to be a score that is more than 2 standard deviations from the mean. Then the scores of 20 (z=-2.4) and 45 (z=2.6) are "rare."

I would expect 65% of scores to be between 25 and 35,
22% between 35 and 40, and
5% between 40 and 45.

Exercise 4: Use a *t* test to test the difference mean achievement tests scores for two *independent* groups.

In lesson 3, Evans showed you how to compute two *t* test for testing the hypothesis that the means of two groups are equal. The first *t* test was for testing the difference in means for two *independent* samples (i.e., samples where what effects one sample has not effect on the other sample.) The other *t* test is for two *dependent* samples (i.e., samples where the score for one member of a pair is correlated with the score for the other member of the pair.) Some examples of dependent samples include pretest-posttest scores, husband's and wife's scores on a political survey, and scores in samples where pairs of cases have been matched.

Below is a table of test scores for two samples of students: males and females.

Males	Females
35	24
27	33
31	37
20	29
29	24
32	30
30	33
22	37

24	22
38	29
	31
	26

Compute a *t* test to determine which group has the highest average test score.

Using Evans, the Excel worksheet I provided, or another procedure (such as Vassar Stats), we compute the following:

	Males	Females
N	10	12
Means	28.8	29.6
SD	5.67	4.73

Mean difference: -.78 [i.e., 2(8.8-29.6)]
Standard Error of the Mean difference: 2.215
t: -.352
degrees of freedom: n1 + n2 -2 = 20
p > .05 (not significant)

Based on these results (since the probability of the obtained *t* statistic is not less than .05) we conclude that there is no difference between the average scores of males and females.

Because the difference is NOT statistically significant we cannot talk about the difference in means as though they are real (remember the inference is to the *population* not to the *samples*). Hence, it is incorrect to conclude that the females had a higher average score, even though the *sample* of females had a higher average score.

Exercise 5: Suppose you test the Null Hypothesis that the means of two groups receiving different treatments are equal. You compute a two-tailed *t* test, with α = .05, obtain a statistically significant result, and REJECT the Null Hypothesis. What is the probability that you have committed a TYPE I error. Explain your reasoning.

A Type I Error occurs whenever a TRUE null hypothesis is REJECTED; i.e., when the null hypothesis should NOT have been rejected. The probably of a Type I error (which can occur only when the null hypothesis IS rejected is equal to *α,* in this case, .05).