STANDARD SCORES AND THE NORMAL DISTRIBUTION
REVIEW

1. MEASURES OF CENTRAL TENDENCY
   A. MEAN
   B. MEDIAN
   C. MODE

2. MEASURES OF DISPERSIONS OR VARIABILITY
   A. RANGE
   B. DEVIATION FROM THE MEAN
   C. VARIANCE
   D. STANDARD DEVIATION
Test Scores

Suppose you (together with many other students) take tests in three subjects. On each test, the range of possible scores runs from 0 to 100 points. The table below shows your score in each of the three subjects:

- In which subject did you do best?
- In which subject did you do best relative to other students?

The answer to the second question obviously depends on the overall frequency distribution of score.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENG</td>
<td>75</td>
</tr>
<tr>
<td>MATH</td>
<td>65</td>
</tr>
<tr>
<td>SCIE</td>
<td>72</td>
</tr>
</tbody>
</table>
Test Scores (cont.)

• Indeed, the picture looks different when we look at your scores relative to overall distribution of scores and, in particular, to its summary statistics.
• Let’s compare each of your scores to the mean score in each subject. Now what seems to be your strongest subject?

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>SCORE</th>
<th>MEAN SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENG</td>
<td>75</td>
<td>76</td>
</tr>
<tr>
<td>MATH</td>
<td>65</td>
<td>55</td>
</tr>
<tr>
<td>SCI</td>
<td>72</td>
<td>57</td>
</tr>
</tbody>
</table>
While you got the highest score in ENGL, this score was actually slightly below (by 1 point) the mean. In fact, it is likely (but not certain) that you scored in the bottom half of all students taking the test. On the other hand, you scored well above average in both of the other subjects (10 points in Math and 15 in SCI).

Note that the magnitudes that we have just referred to here are your deviations from mean in each subject.

Since your deviation from the mean is greatest with respect to SCI, this may appear to be your strongest subject. But this may not be the case.

### Your Deviations from the Mean

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>SCORE</th>
<th>MEAN SCORE</th>
<th>DEVIATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENG</td>
<td>75</td>
<td>76</td>
<td>-1</td>
</tr>
<tr>
<td>MATH</td>
<td>65</td>
<td>55</td>
<td>+10</td>
</tr>
<tr>
<td>SCI</td>
<td>72</td>
<td>57</td>
<td>+15</td>
</tr>
</tbody>
</table>
Your Deviations from the Mean Compared with Other Deviations from the Mean

• While you have a deviation from the mean in each subject, so does every other student who took the test.

• Consider the distribution of SCI scores. Almost certainly quite a few students scored close to the mean, but probably quite a few others scored well above the mean (like you) and others well below.
  – On the one hand, but if most students scored very close to the mean (so the dispersion in test scores is small),
    • your score of 72 would make you an outlier, scoring higher than almost all other students.
  – On the other hand, if many students scored well above the mean (and — since we know that the sum of all deviations from the mean must sum to zero — many other students scored well below the mean, so the dispersion of test scores is large),
    • your score of 72, while certainly good, would be less outstanding.
Your Deviations Compared with Other Deviations (cont.)

• Thus whether your score is outstanding or merely good depends
  – not just on your score compared with the mean score
  – but also on your deviation from the mean compared with other deviations from the mean, i.e., the dispersion of scores.

• Recall that the standard measure of dispersion — the standard deviation — itself is directly based on the deviations from the mean.

• Recall also that the SD of scores (though precisely defined as the square root of the average of all squared deviations) is approximately the same as (though usually somewhat greater than) the average of the absolute deviations from the mean.
Your Deviations from the Mean Compared with the Standard Deviation from the Mean

Thus, to get a sense of how outstanding your SCI and MATH scores are, we should look at how big your deviation from the mean is compared with the standard (“average”) deviation from the mean, by calculating the ratio of your deviation to the standard deviation.

The result of this calculation is called your standard score.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>SCORE</th>
<th>MEAN SCORE</th>
<th>DEVIATION</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENG</td>
<td>75</td>
<td>76</td>
<td>-1</td>
<td>8</td>
</tr>
<tr>
<td>MATH</td>
<td>65</td>
<td>55</td>
<td>+10</td>
<td>5</td>
</tr>
<tr>
<td>SCI</td>
<td>72</td>
<td>57</td>
<td>+15</td>
<td>15</td>
</tr>
</tbody>
</table>
So in terms of your standard score, i.e., how your deviation from the mean compares with the standard deviation from the mean, it is evident that

- your best performance was actually in MATH (where you scored two standard deviations above the mean),
- compared with POLI (where you scored only one standard deviation above the mean).
- In ENGL you scored 1/8 of a standard deviation below the mean.
Computing a z-score

\[ z = \frac{X - \mu}{\sigma} \quad \text{or} \quad z = \frac{X - \bar{X}}{SD} \]
### Examples of computing z-scores

<table>
<thead>
<tr>
<th>$X$</th>
<th>$\bar{X}$</th>
<th>$X - \bar{X}$</th>
<th>$SD$</th>
<th>$z = \frac{X - \bar{X}}{SD}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1.5</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>-5</td>
<td>4</td>
<td>-1.25</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>.75</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>-4</td>
<td>2</td>
<td>-2</td>
</tr>
</tbody>
</table>
Computing raw scores from z scores

\[ X = z\sigma + \mu \quad or \quad X = zSD + \bar{X} \]

<table>
<thead>
<tr>
<th>( z = \frac{X - \bar{X}}{SD} )</th>
<th>( SD )</th>
<th>( zSD )</th>
<th>( \bar{X} )</th>
<th>( X )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>-2</td>
<td>2</td>
<td>-4</td>
<td>2</td>
<td>-2</td>
</tr>
<tr>
<td>.5</td>
<td>4</td>
<td>2</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>-1</td>
<td>5</td>
<td>-5</td>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>
Other Variants of Standard Scores

• Approximately half of the people who take any test necessarily get negative standard scores.
• This unavoidable arithmetical fact apparently is regarded as demoralizing, so standard scores are commonly converted into so-called *T*-scores, which are all positive. By convention, *T*-scores are calculated by multiplying standard scores by 10 and then adding 50.
• *IQ scores* are also derived from standard scores, calculated by multiplying standard scores by 15 and then adding 100.
• The table below shows how you performed in the three subjects in terms of each of these scoring systems (where, as is conventional, all derived scores have been rounded to the nearest whole point).

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>SCORE</th>
<th>MEAN SCORE</th>
<th>DEVIATION</th>
<th>SD</th>
<th>Z-SCORE</th>
<th>T-SCORE</th>
<th>IQ SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENG</td>
<td>75</td>
<td>76</td>
<td>-1</td>
<td>8</td>
<td>-0.125</td>
<td>49</td>
<td>98</td>
</tr>
<tr>
<td>MATH</td>
<td>65</td>
<td>55</td>
<td>+10</td>
<td>5</td>
<td>+2.0</td>
<td>70</td>
<td>130</td>
</tr>
<tr>
<td>SCI</td>
<td>72</td>
<td>57</td>
<td>+15</td>
<td>15</td>
<td>+1.0</td>
<td>60</td>
<td>115</td>
</tr>
</tbody>
</table>
1. Interpret a z score of 1

2. M = 10, SD = 2, X = 8. Z =?

3. M = 8, SD = 1, z = 3. X =?

4. What is the IQ score for a z score of 1?

• IQ scores are also derived from standard scores, calculated by multiplying standard scores by 15 and then adding 100.
To move from a raw score to a z score, what must we know about the raw score distribution?

- 1 mean and standard deviation
- 2 maximum and minimum
- 3 median and variance
- 4 mode and range
Application

If Judy got a z score of 1.5 on an in-class exam, what can we say about her score relative to others who took the exam?
1. it is above average
2. it is average
3. it is below average
4. it is a ‘B’
Test your mastery of $z$

- If a raw score is 8, the mean is 10 and the standard deviation is 4, what is the $z$-score?
  - 1: -1.0
  - 2: -0.5
  - 3: 0.5
  - 4: 2.0
Test your mastery of $z$ and the normal curve

- If a distribution is normally distributed, about what percent of the scores fall below $+1$ SD?
  - 1: 15
  - 2: 50
  - 3: 85
  - 4: 99
Your Percentile Rank in Each Subject?

- While it is extremely likely that your percentile rank among all students taking each test is highest in MATH and lowest in ENGL, we do not know this for sure in the absence of knowing the full frequency distribution of scores (as opposed to knowing only the two summary statistics: the mean and the SD).

- Much data — particularly including tests scores, many other interval measures, and many types of sample statistics — is (at least approximately) normally distributed.
  - However, a lot of other data (especially ratio measures), such as weight, income (as we have seen), wealth, house prices, and many other ratio variables, is skewed with longer thin tails in the direction of (much) higher values
    - while there is a zero-limit on the minimum value.
The Normal Distribution

- A normal distribution is a continuous frequency density that is a particular type of symmetric bell-shaped curve.

- Because the curve has a single peak and is symmetric about this peak, its mode, median, and mean values coincide at this peak.

- Most observed values lie relatively “close” (in way that is made more specific below) to the center of distribution, and their density falls off on either side of peak.
A Normal Curve
Normal Curve

- The shape of the distribution changes with only two parameters, $\sigma$ and $\mu$, so if we know these, we can determine everything else.
Standard Normal Curve

- **Standard normal curve** has a mean of zero and an SD of 1.

![Standard Normal Curve Diagram]

- Probability (Relative Frequency)
- Scores in standard deviations from mu
- 50 Percent
- 34.13%
- 13.59%
- 2.15%
If $X$ is normally distributed, there will be a correspondence between the standard normal curve and the $z$-score.
Normal curve and z-scores

- We can use the information from the normal curve to estimate percentages from z-scores.

![Standard Normal Curve](image)

- Probability (Relative Frequency)
- Scores in standard deviations from mu
The Mean and SD of the Normal Distribution

• The *mean* of a normal distribution determines its *location* on the horizontal scale. The mean value of the distribution (here equal to the mode) is simply the value (point on the horizontal scale) of the variable that lies under the highest point on the curve.
  – For example, if a constant amount is added to (or subtracted from) every value of the variable, the normal curve slides upwards (or downwards) by that constant amount.

• The *standard deviation* of a normal distribution determines how “spread out” the distribution is.
  – Once the horizontal scale is fixed, if the SD is small, the curve has a high peak with sharp slopes on either side; if the SD is large, the curve it has a low peak with gentle slopes on either side.
Finding the SD of a Normal Curve

- There is a precise connection between the shape of a normal curve and its SD.

- The two points of maximum steepness on either side of the peak are called the *inflection points* of the (normal) curve.

- It turns out (as a mathematical theorem) that horizontal distance from the mean to each inflection point is identical to the standard deviation of the normal curve.
Here is another method for LOOKING AT the magnitude of the SD of a normal distribution.

Put two vertical lines on either side of, and equidistant from, the peak and then draw them apart or bring them closer together (keeping them equidistant from the peak) until it appears that just about two-thirds of the areas under the curve lies in the interval between the two vertical lines.

The horizontal distance from the mean to either line is equal to (a very good approximation of) the standard deviation of the distribution.
The 68%-95%-99.7% Rule

- More generally, we can state what is called the [approximate] 68%-95%-99.7% rule of the normal distribution. The rule is this:
  - about 68% of all observed values lie within one SD of the mean,
  - about 95% lie within two SDs of the mean, and
  - about 99.7% (that is, virtually all) lie within three SDs of the mean.
    - This is why no SAT scores below 200 [3 standard deviations below the mean] or above 800 [3 standard deviations above the mean] are reported.

- And here is another useful rule: in a normal distribution, half the cases have observed values that lie within about 2/3 of the SD of the mean, i.e.,
  - the first and third quartiles lie at just about 2/3 of a SD below and above the mean respectively, so
  - In a normal distribution, the interquartile range is equal to about 1.33 SDs.

- All this is illustrated in the following chart, which shows a standardized normal curve, i.e., a normal curve in which the mean is set at 0 and the SD is set at 1. Put otherwise, the units on the horizontal scale shows standard scores.
FIGURE 1 --- THE NORMAL CURVE (WITH STANDARD SCORES)
Your Percentile Ranks (if Test Scores are Normally Distributed)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Stan. Score</th>
<th>Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENGL</td>
<td>-0.125</td>
<td>45</td>
</tr>
<tr>
<td>MATH</td>
<td>+2.0</td>
<td>97.5</td>
</tr>
<tr>
<td>SCI</td>
<td>+1.0</td>
<td>84</td>
</tr>
</tbody>
</table>