STANDARD SCORES AND THE NORMAL DISTRIBUTION

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, for each of the three subjects: your score, the mean of all scores, and the standard deviation of all scores.

Subject	*Your Score*	*Mean Score*	*SD of Scores*
ENGL	75	76	8
MATH	65	55	5
POLI	72	57	15

Looking only at your own scores, it appears that you did best in ENGL, almost as well in POLI, and much the worst in MATH. But this may be only because the MATH test was the hardest and the ENGL test the easiest. It is likely that your actual grade (as opposed to your raw score) on each test will depend on the overall distribution of scores on that test, as well as your own score. One way to derive grades in this manner is by means of *standard scores*.

This entails looking at your scores on each test *in relation to the overall frequency distribution of scores* on each test. The table above shows two *summary statistics* pertaining to the distribution of student scores — namely, its *mean* and *standard deviation*. Looking at your score on each test relative to the mean score on that test changes our sense of which subjects you did best and worst in. While you got the highest score in ENGL, this score was actually slightly below average (precisely, it was 1 point below the mean). Indeed, it is likely (but not certain) that you scored in the bottom half of all students taking the test. (It is not certain because we don't know what the *median* score was, and it is the median, not the mean, that defines the cutting point between the top half and bottom half of scores. However, for reasons to be discussed later, it is likely that mean and median test scores are just about the same.) On the other hand, you scored well above average in both of the other subjects (precisely, 10 points above the mean in MATH and 15 points above the mean in POLI).

Note that the magnitudes that we have just referred to here are your *deviations from mean* (as discussed in the two previous handouts) in each subject.

Subject	*Your Score*	*Mean Score*	*Your Deviation from the Mean*
ENGL	75	76	−1
MATH	65	55	+10
POLI	72	57	+15

But while you have a deviation from the mean in each subject, so does every other student who took the test. Consider the distribution of POLI scores. Almost certainly quite a few students scored close the mean (had small positive or negative deviations from the mean), but probably quite a few others scored well above the mean (like you) or well below. If most students scored very close to the mean (so the *dispersion* in test scores is low), a 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 is zero — many other students scored well below the mean), a score of 72, while certainly good, would not be outstanding.

Thus whether your score is outstanding or merely good depends also on the dispersion of scores. Recall from Handout #7 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 the average of all the absolute deviations.

Thus, to get a sense of how outstanding your POLI 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*.

Subject	*Your Dev. from the Mean*	*SD from the Mean*	*Ratio*	*Your Standard Score*
ENGL	−1	8	−1 / 8	= −0.125
MATH	+10	5	+10 / 5	= +2.0
POLI	+15	15	+15 / 15	= +1.0

We respect to 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) and ENGL (where you scored 1/8 of a standard deviation below the mean).

Approximately half of the people who take any test necessarily get negative standard scores. This unavoidable arithmetical fact apparently is regarded as excessively 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. In turn, *SAT scores* are T-scores multiplied by 10. *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).

Subject	*Standard Score*	*T-score*	*SAT Score*	*IQ Score*
ENGL	−0.125	49	488	98
MATH	+2.0	70	700	130
POLI	+1.0	60	600	115

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 the 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 *normally distributed* (or at least approximately so). (However, much other data — particularly including income, wealth, house prices, and many other ratio variables — are — at least slightly and sometimes very strongly — skewed with long thin tails in the direction of higher values.)

A *normal distribution* (or "bell curve") 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, its mode, median, and mean are all the same. 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 graph of a normal curve appears in Figure 1.

The *mean* of a normal distribution determines its *location* on the horizontal scale. The mean value of the distribution (being identical 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. (See **The Normal Curve** applet discussed below.)

There is a precise connection between the shape of a normal curve and its SD. Consider the normal curve in Figure 1. Imagine that it represents the cross-section of a mountain you must traverse, first climbing and then descending as you move from left to right. At first the trail is almost level but it gets steeper as you proceed. At some point the degree of steepness reaches a maximum and thereafter the steepness of the trail declines until it levels out at the top of the mountain. The trail then follows a mirror-image pattern on the descent.

The two points of maximum steepness on either side of the peak are called the *inflection points* of the (normal) curve. Since the curve is symmetric, the two inflection points are equidistant from the peak (mean). 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. However, the normal curve is virtually straight in the vicinity of its inflection points, so it can be hard to determine very precisely by eyeball methods where the inflection point is and thus what the magnitude of the SD is.

Here is another method for eyeballing 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 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.

More generally, we can state what is called the (approximate) *68%-95%-99.7% rule of the normal distribution*. The rule is this: (i) about 68% of all observed values lie within one SD of the mean, (ii) about 95% lie within two SDs of the mean, and (iii) 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.) (For a clear "picture" of this rule, click on **The Normal Curve: The 68% / 95% / 99.7% Rule** under **Online Statistical Demonstrations** on the course website.) And here is another useful rule: in a normal distribution, half the cases have observed values that lie within about ⅔ of a SD of the mean.

See Figure 1, 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 are simply *standard scores*.

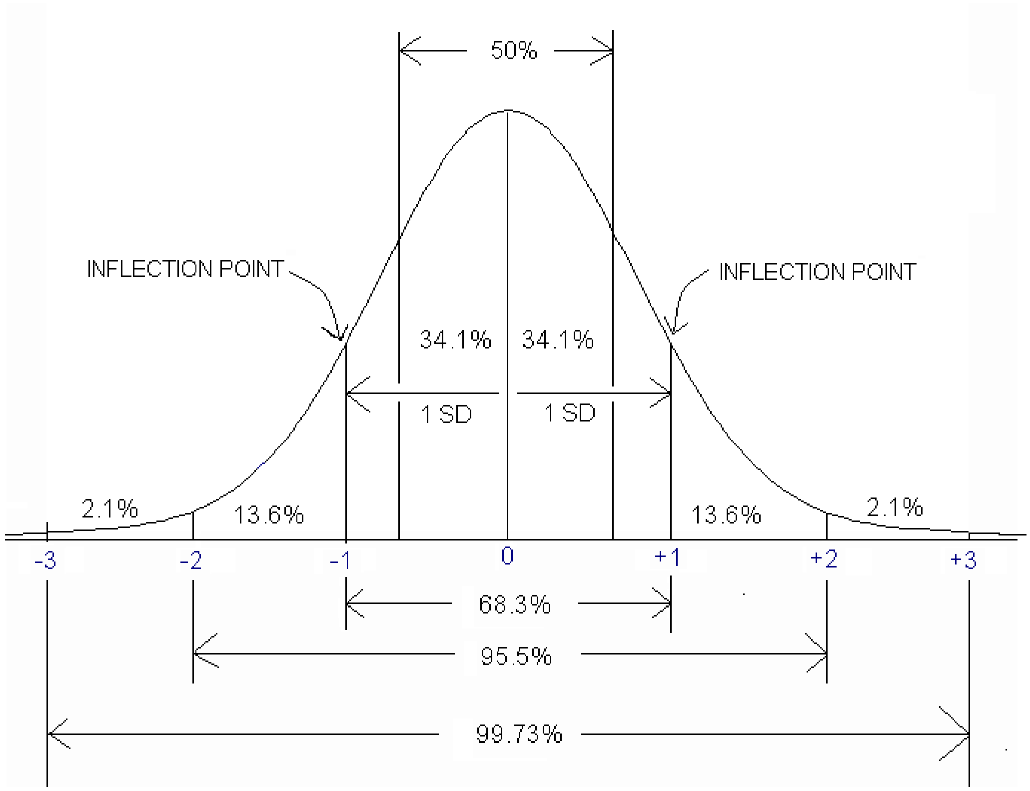


FIGURE 1 --- THE NORMAL CURVE

If test scores are normally distributed, we know from the 68-95-99.7% and 50% rules the percentile ranks associated with the following standard scores:

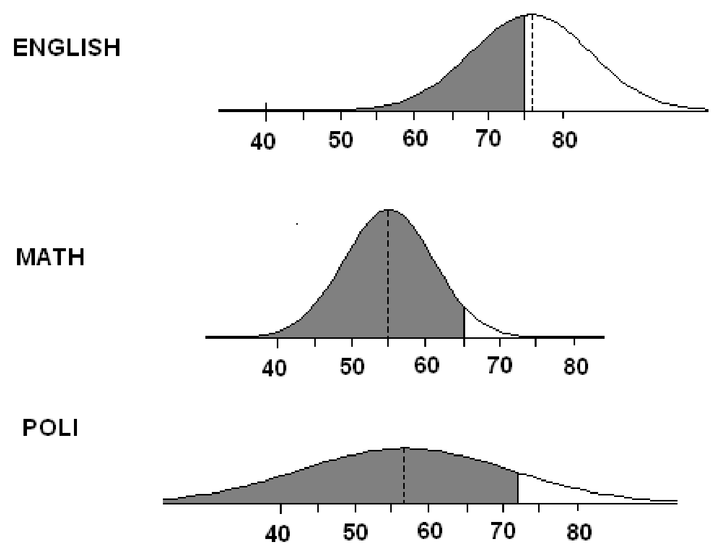
If your <u>Standard Score</u> is	*then your <u>Percentile</u> rank is about*
+3	0.15
+2	2.5
+1	16
+0.67	25
0	50
−0.67	75
−1	84
−2	97.5
−3	99.85

(Remember that if you score at the P^{th} percentile, P % of the cases have lower scores than you and 100 − P % have higher scores than you.) From tables (found in appendices to statistics texts) or computer programs (see below), one can determine the percentile rank of associated with any standard score in a normal distribution. Here are your percentiles in the three tests. (Your POLI and MATH percentiles can be determined by the 68-95-99.7% rule; I got your ENGL percentile from the "statistical applet" noted below.)

Subject	*Standard Score*	*Percentile*
ENGL	−0.125	45
MATH	+2.0	97.5
POLI	+1.0	84

Figure 2 shows where you stand in each subject, assuming that the distribution of scores in each subject is (approximately) normal.

Figure 2



It is worthwhile explicitly to highlight one characteristic of the normal distribution: most cases are "packed" into a relatively narrow interval quite close to the mean. In this range of "mediocrity" (literally, in the vicinity of the median), a small change in a score can produce a big change in percentile rankings. For example, to get a score of 460 when you first take the SAT and then get a score of 540 when you take it a second time is nice but not spectacular improvement (80 points), but it jumps you from the 33rd percentile to the 67th (i.e., it jumps you over one-third of all SAT takers). But to jump above the remaining third of SAT takers still above you (i.e., to the 99th percentile or better), your score would have to go from 540 to 800 (260 points).

You might check out these three addition links (under "On-Line Statistical Demonstrations") on the course website.

(a) Click on **<u>Statistical "Applets"</u>** and then on Normal density calculator. You will find a standardized normal curve on which you can perform exactly the exercise described on the previous page ("Put two vertical lines symmetrically on either side of the peak . . . "). You can also adjust the mean and SD of the normal curve. However, these adjustments do not change the location and shape of curve as it appears on the screen — rather they change the calibration of the horizontal scale.

(b) Click on **The Normal Curve**. Here you will find another normal curve that allows you to adjust the mean and SD of the normal curve so that the curve itself changes location and shape (while the calibration of the horizontal axis remains fixed).

(c) Many probability distributions related to sampling are approximately normal. For example, suppose we have a large population in which 50% of the units are "red cards" (vs. "black cards") or "left" (vs. "right") or "approve of the President's performance" (vs. "disapprove") or whatever. This 50% is the population parameter of interest. Now we take repeated simple random samples of size $n = 8$ from this population and calculate the appropriate statistic for each sample. From earlier discussions we know that: (a) *on average* the samples will give the correct statistic of 50% , but (b) because of sampling errors, most individual samples will not give the correct statistic or even come very close (the margin of error is about $100\%/\sqrt{n} \approx 35\%$). More specifically, each sample can produce one of the nine possible sample statistics shown just below. Moreover, we can calculate the probability of getting each statistic (or its expected frequency in a large number of samples). These calculations are shown, where $n\ C\ m$ means "the number of ways out choosing m things out of a set of n things" and is equal to $n!/[m!\,(n-m)!]$, where $k!$ (*k factorial*) $= 1 \times 2 \times 3 \times \ldots \times (k-1) \times k$.

Sample Statistic	*Probability or Expected Frequency*
0/8 = 0%	$.5^8 \times (8\ C\ 0) = .5^8 \times 1 = .004$
1/8 = 12.5%	$.5^8 \times (8\ C\ 1) = .5^8 \times 8 = .031$
2/8 = 25%	$.5^8 \times (8\ C\ 2) = .5^8 \times 28 = .109$
3/8 = 37.5%	$.5^8 \times (8\ C\ 3) = .5^8 \times 56 = .219$
4/8 = 50%	$.5^8 \times (8\ C\ 4) = .5^8 \times 70 = .273$
5/8 = 62.5%	$.5^8 \times (8\ C\ 5) = .5^8 \times 56 = .219$
6/8 = 75%	$.5^8 \times (8\ C\ 6) = .5^8 \times 28 = .109$
7/8 = 87.5%	$.5^8 \times (8\ C\ 7) = .5^8 \times 8 = .031$
8/8 = 100%	$.5^8 \times (8\ C\ 8) = .5^8 \times 1 = \underline{.004}$
	1.000

As we can see, the expected frequency distribution is "normal looking." If you click on **Probability and the Normal Distribution**, you can see such a sampling experiment take place on the screen and you can observe how the actual frequency distribution of sample statics assumes a more and more normal shape as the number of sample increases. (The experiment as described entails running balls through a pinball machine in which there are eight levels of pins and at each level the ball goes either left or right with equal probability. This generates exactly the same probabilities as the sampling experiment described above.)