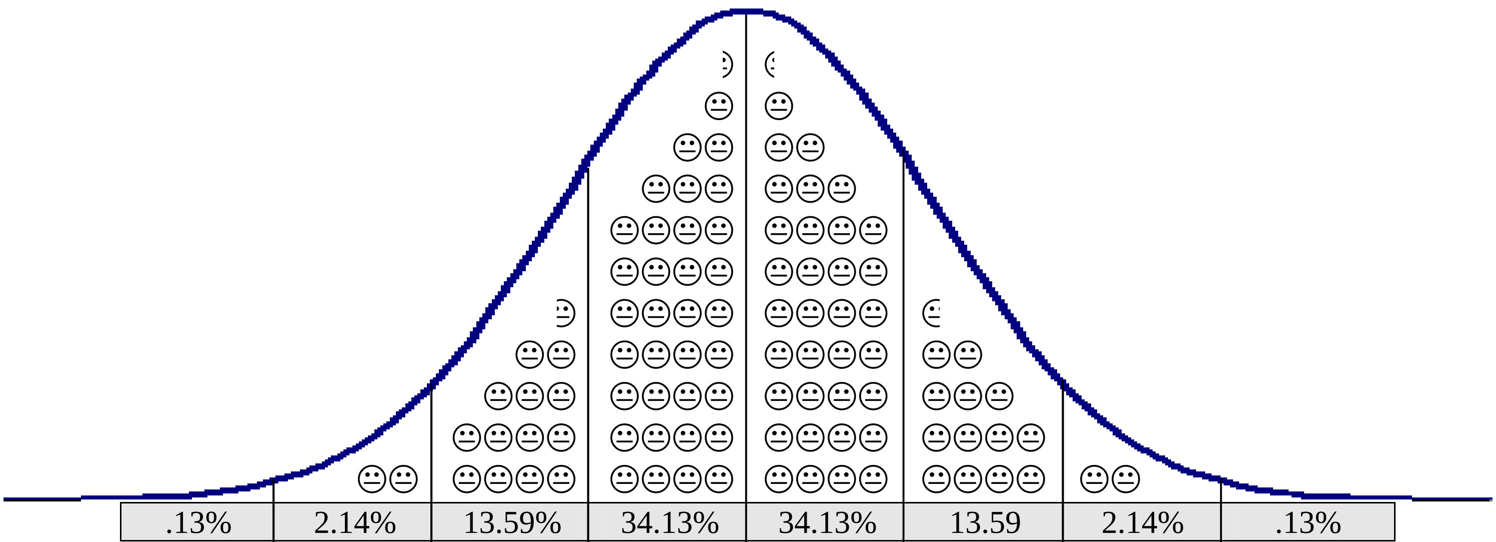
Descriptive Statistics and Psychological Testing

By Stephen E. Brock, Ph.D., NCSP
California State University, Sacramento



	-3sd	-2sd	-1sd	Mean (Median & Mode)	+1sd	+2sd	+3sd	
	Below Average	Low Average	Average			High Average	Above Average	
Z-Score	-3	-2	-1	0	+1	+2	+3	Z-Score
IQ Score	55	70	85	100	115	130	145	IQ Score
T-Score	20	30	40	50	60	70	90	T-Score
Scaled Score	2	4	7	10	13	16	19	Scaled Score
%ile	1	2	16	50	84	98	99	%ile
Rank	1		10 20 30	40 50 60	70 80 90		99	Rank

NOTE: *Z*-scores, IQ scores T-scores, and scaled scores are considered interval scales of measurement. These scores indicate rank and meaningfully reflect relative the distance between scores. Percentiles only indicate ranking, by themselves they do not indicate how far apart scores are.

The Normal Curve

The normal curve is a hypothetical distribution of scores that is widely used in psychological testing. The normal curve is a symmetrical distribution of scores with an equal number of scores above and below the midpoint. Given that the distribution of scores is symmetrical (i.e., an equal number of scores actually are above and below the midpoint) the mean, median, and mode all fall at the same point. Since many psycho-educational measurements (e.g., intelligence and achievement test scores) assume a normal distribution, the concept of the normal curve is very important to school psychologists.

If we divide the distribution up into standard deviations from the midpoint, a specific percentage of scores will lie under each part of the normal curve. As illustrated in the figure above, 34.13% of the scores lie between the mean and 1 standard deviation *above* the mean. This same percentage (34.13%) of scores lies between the mean and 1 standard deviation *below* the mean. Approximately two-thirds of the scores lie within 1 standard deviation of the mean (68.26%), and approximately 95% of the scores lie within 2 standard deviations of the mean. Finally, over 99% of the scores fall within 3 standard deviations of the mean. Thus scores that fall more than 2 standard deviations from the mean are relatively rare (sometime identified as being “clinically significant”).

Standard Deviation

The standard deviation is a measure of the variability of a distribution of test scores. Test developers need to know the standard deviation of the distribution of a tests raw scores before they can standardize these raw scores. Tests that have very little variability (the raw scores are very similar to each other) have small standard deviations, while tests that have significant variability (the raw scores obtained by individuals taking the test are very different from each other) have large standard deviations. The standard deviation of a distribution of raw scores is the square root of the variance. The variance is the sum of the squared raw score values (ΣX^2) minus the square of the sum of all the raw scores $(\Sigma X)^2$ divided by the number of raw scores (N). The resulting figure is then divided by the number of raw scores minus 1 ($N - 1$). This formula is summarized in the following figure:

$$\sqrt{\frac{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}{N - 1}}$$

Standard Scores

When a set of raw scores is converted to standard scores the scores are said to be "standardized." The purpose of standard scores (e.g., *Z*-scores, IQ Scores, *T*-scores, scaled scores) is to transform individual raw scores into a standard form that provides a more meaningful description of the individual scores within the distribution. Raw test data is rarely valuable to the school psychologist. For example, a raw score of 5 on the *Wechsler Intelligence Scale for Children* (WISC) Information subtest may mean different things for different students. A raw score of 5 for a six-year-old will be suggestive of a different level of cognitive functioning than will the same score for a seven-year-old. In addition, a raw score of 5 on one test will not have the same meaning as a raw score of 5 on another test. Thus, the raw scores obtained via psychological tests are most commonly interpreted by reference to norms and by their conversion into some relative reference or "standard" score (a descriptive statistic).

Norms represent the test performance of individuals within a standardization sample. For example, they document how well the standardization sample's six-year-olds did on the *WISC* Information subtest. *Derived scores* are the descriptive statistics used to transform raw test data into a number that more precisely illustrates a student's exact position relative to individuals in the normative group. For example, at age six, a raw score of 5 on the *WISC* Information subtest corresponds to a scaled score of 10. While at age seven, this same raw score corresponds to a scaled score of 6. Derived scores also provide comparable measures that allow direct comparison of a student's performance on different tests. Thus, allowing the school psychologist to identify a relative pattern of unique strengths and weaknesses. For example, a scaled score of 10 on the Information subtest (RS = 5) can be directly compared to a scaled score of 3 on the Coding subtest (RS = 5). Understanding the conversion of raw scores into standard scores, and how they are used to describe a student's performance relative to others (as well as their own unique pattern of strengths and weaknesses) requires knowledge of basic statistical concepts. These concepts underlie the development and utilization of norms. It is critical that school psychologists, who use psychological tests, have a solid understanding of these descriptive statistics.

Z-Scores

Z-Scores are a transformation of individual raw scores into a standard form, where the transformation is based on knowledge about the standardization sample's mean and standard deviation. The formula for computing *Z*-scores is the individual raw score (X) minus the mean of the scores obtained by the standardization sample (M), divided by the standard deviation of scores obtained by the standardization sample (*sd*). *Z*-scores have a mean of 0 and a standard deviation of 1. A score that is one standard deviation below the mean has a Z-score of -1. A score that is at the mean would have a Z-score of 0. The formula for transforming a raw score into a *Z*-score is a follows:

$$\frac{X - M}{sd} = Z$$

Because of the fact that the pulse (+) and minus (-) signs can easily get lost when looking at this type of standard score, *Z*-scores are frequently converted into other types of standard scores. Specifically they are often transformed into Deviation IQ scores, T-scores, and scaled scores.

Deviation IQ Scores

Deviation IQ Scores are a standard score with a mean of 100 and a standard deviation of 15. *Z*-scores can be transformed into Deviations IQ scores by multiplying the given *Z*-score by 15 (the standard deviation of the distribution of Deviation IQ scores), and adding 100 (the mean of the distribution of Deviation IQ scores) to this product. For example, a *Z*-score of –1 equals a Deviation IQ of 85 [100 + 15(-1) = 85]. The formula for transforming a *Z*-score into a Deviation IQ score is a follows:

$$100 + 15(z)$$

If the skills measured by an IQ test are normally distributed, we would expect that two-thirds (68.26%) of the population would have deviation IQ's between 85 and 115. This is considered the normal range. Further, we would expect that 95% of the distribution lies within 2 standard deviations of the mean (that is IQs between 70 and 130). Thus, scores that fall above 130 and below 70 would be considered unusually high and unusually low, as only 5% of the population obtains higher or lower scores.

T-Scores

T-scores are standard scores with a mean of 50 and a standard deviation of 10. *Z*-scores can be transformed into T-scores scores by multiplying the given *Z*-score by 10 (the standard deviation of the distribution of T-scores), and adding 50 (the mean of the distribution of T-scores) to this product. For example, a *Z*-score of –1 equals a Deviation IQ of 40 [50 + 10(-1) = 40]. The formula for transforming *Z*-score into a T-score is a follows:

$$50 + 10(z)$$

If the variable measured by a psychological test is normally distributed, we would expect that two-thirds (68.26%) of the population would obtain scores between 40 and 60. This is considered the normal range. Further, we would expect that 95% of the distribution lies within 2 standard deviations of the mean (that is T-scores between 30 and 70). Thus, scores that fall above 70 or below 30 would be considered unusually high and unusually low, as only 5% of the population obtains higher or lower scores.

Scaled Scores

Scaled scores are standard scores with a mean of 10 and a standard deviation of 3. *Z*-scores can be transformed into scaled scores by multiplying the given *Z*-score by 3 (the standard deviation of the distribution of scaled scores), and adding 10 (the mean of the distribution of scaled scores) to this product. For example, a *Z*-score of –1 equals a scaled of 7 [10 + 3(-1) = 7]. The formula for transforming *Z*-score into a scaled score is a follows:

$$10 + 3(z)$$

If the variable measured by a psychological test is normally distributed, we would expect that two-thirds (68.26%) of the population would obtain scores between 7 and 12. This is considered the normal range. Further, we would expect that 95% of the distribution lies within 2 standard deviations of the mean (that is scaled scores between 4 and 16). Thus, scores that fall above 16 or below 4 would be considered unusually high and unusually low, as only 5% of the population obtains higher or lower scores. As was mentioned earlier, the term “clinically significant” is sometimes used to describe these unusually high or low scores.

Percentile Ranks

The percentile rank reflects the percentage of scores that are lower than an obtained test score. For example, a test result that fell at the 75th percentile rank is higher than that obtained by 74% of the population. In other words, the individual obtaining this test score scored higher than 74% of the individuals in the standardization group.

The median for any set of raw scores is the 50th percentile. That is, 50% of the scores are lower than the median, and 50% of the scores are higher than the median. Typically percentiles are reported as whole numbers so the highest percentile possible would be 99 and the lowest possible would be 1[1].

Another way to think about percentile ranks is that they reflect the percentage of the area underneath the normal curve that is to the left of the given score. For example, a score that is 2 standard deviations below the mean would have a percentile rank of 2 (0.13 + 2.14 = 2.27). In other words, just over 2% of the area underneath the normal curve is to the left of a standard score that is 2 standard deviations below the mean. On the other hand a score that is 2 standard deviations above the mean would have a percentile rank of 98 (0.13 + 2.14 +13.59 + 34.13 + 34.13 + 13.59 = 97.71). In other words, just under 98% of the area underneath the normal curve is to the left of a standard scores that is 2 standard deviations above the mean. The following table illustrates the relationship between specific percentile scores and specific *Z*-scores, Deviation IQ scores, T-scores, and scaled scores.

---

[1] Some test designers have used the concept of extended percentile ranks to make finer divisions for scores at the upper half of the 99th percentile and at the lower half of the 1st percentile (e.g., they may report a given score as falling at the 99.7 percentile rank).

Percentile Rank	*Z-Score*	*Deviation IQ (SD = 15)*	*T-Score*	*Scaled Score*
99	+2.33	135	73	17
98	+2.05	131	71	16
97	+1.88	128	69	
96	+1.75	126	68	
95	+1.64	125	67	15
94	+1.55	123	66	
93	+1.48	122	65	
92	+1.41	121	64	
91	+1.34	120		14
90	+1.28	119	63	
89	+1.22			
88	+1.18	118	62	
87	+1.13	117		
86	+1.08	116	61	
85	+1.04			
84	+0.99	115	60	13
83	+0.95			
82	+0.91	114	59	
81	+0.88	113		
80	+0.84			
79	+0.80	112	58	
78	+0.77			
77	+0.74	111		
76	+0.71		57	
75	+0.67	110		12
74	+0.64			
73	+0.61	109	56	
72	+0.58			
71	+0.55			
70	+0.52	108		
69	+0.49		55	
68	+0.47	107		
67	+0.44			
66	+0.41	106	54	
65	+0.39			
64	+0.36			
63	+0.33	105		11
62	+0.31		53	
61	+0.28	104		
60	+0.25			
59	+0.23			
58	+0.20	103	52	
57	+0.18			
56	+0.15			
55	+0.12	102		
54	+0.10		51	
53	+0.07	101		
52	+0.05			
51	+0.03			
50	0.00	100	50	10

Percentile Rank	*Z Score*	*Deviation IQ (SD = 15)*	*T-Score*	*Scaled Score*
50	0.00	100	50	10
49	-0.03			
48	-0.05			
47	-0.07	99		
46	-0.10		49	
45	-0.12	98		
44	-0.15			
43	-0.18			
42	-0.20	97	48	
41	-0.23			
40	-0.25			
39	-0.28	96		
38	-0.31		47	
37	-0.33	95		9
36	-0.36			
35	-0.39			
34	-0.41	94	46	
33	-0.44			
32	-0.47	93		
31	-0.49		45	
30	-0.52	92		
29	-0.55			
28	-0.58			
27	-0.61	91	44	
26	-0.64			
25	-0.67	90		8
24	-0.71		43	
23	-0.74	89		
22	-0.77			
21	-0.80	88	42	
20	-0.94			
19	-0.88	87		
18	-0.91	86	41	
17	-0.95			
16	-0-99	85	40	7
15	-1.04			
14	-1.08	84	39	
13	-1.13	83		
12	-1.18	82	38	
11	-1.22	82		
10	-1.28	81	37	
9	-1.34	80		6
8	-1.41	79	36	
7	-1.48	78	35	
6	-1.55	77	34	
5	-1.64	75	34	5
4	-1.75	74	33	
3	-1.88	72	31	
2	-2.05	69	30	4
1	-2.33	65	27	3

NOTES: