Technical Whitepaper 3

August 2003, Revised June 2011

Percentiles and Percentile Ranks

Confused or what?

Advanced Projects R&D Ltd.
T +64-9-238-6336
M +64-21-415625
W www.pbarrett.net
E aprdev@pbarrett.net
pbar088

Percentiles and textbook definitions – confused or what?

Take the following definitions …

HyperStat Online:
A percentile rank is the proportion of scores in a distribution that a specific score is greater than or equal to. For instance, if you received a score of 95 on a math test and this score was greater than or equal to the scores of 88% of the students taking the test, then your percentile rank would be 88. You would be in the 88th percentile.
http://davidmlane.com/hyperstat/glossary.html

Hinkle, D., Wiersma, W., & Jurs, S. (1994). Applied statistics for the behavioral sciences. (3rd ed.). Boston: Houghton Mifflin Company.(p. 49-50)
A percentile is the point in a distribution at or below which a given percentage of scores is found. For example, the 28th percentile of a distribution of scores is the point at or below which 28% of the scores fall.

Monroe County School District, Florida, US
The percentile is a point on a scale of scores at or below which a given percent of the cases falls. For example, a child who scores at the 42 percentile, *is doing as well as, or better than*, 42 percent of the students who took the same test.

Wisconsin Department of Public Instruction
The State Percentile is a ranking of an individual student's results relative to the results of other test takers in the state.For example, a student who scored at the 82nd percentile had a score that was equal to or better than 82% of the scores of all students in Wisconsin who took the same test.
https://dpi.wi.gov/sites/default/files/imce/assessment/pdf/Users_Guide_to_Interpreting_Reports_2017-18.pdf (see page 9)

Moore, D.S. and McCabe, G.P. (1993) Introduction to the Practice of Statistics 2nd Edition. New York: W.H. Freeman and Company (p. 40)
The *pth* percentile of the distribution is the value such that *p* percent of the observations fall at or below it.

Hays, W.L. (1994) Statistics, 5th Edition. Florida: Harcourt Brace. (p. 194)
In any frequency distribution of numerical scores, the percentile rank of any specific value *x* is the percentage of the total cases that fall at or below *x* in value.

Kiess, H.O. (1996) Statistical Concepts for the Behavioral Sciences. London: Allyn and Bacon (p. 46)
A percentile is the score at or below which a specified percentage of scores in a distribution falls.

STATISTICA 6 (Statsoft Inc.)
The *percentile* (this term was first used by Galton, 1885a) of a distribution of values is a number x_p such that a percentage p of the population values are **less than or equal to** x_p. For example, the 25th *percentile* (also referred to as the .25 quantile or lower quartile) of a variable is a value (x_p) such that 25% (p) of the **values of the variable fall below** that value.

Howell, D. (1989) Fundamental Statistics for the Behavioral Sciences. 2nd Edition. Boston: PWS-Kent Publishing. (p. 36)
A percentile is the point on a scale at or below which a given percentage of the scores fall.
*contrast this with the definition by Howell (2002) on page 4 …

Contrast the above with the following:

Bartram, D. and Lindley, P.A. (1994) BPS Level A Open Learning Training Manual: Scaling Norms and Standardization, Module 2, part 1. London: BPS Publications (p.17)
The proportion of people scoring less than a particular score is called the percentile rank of the score. More commonly we refer to this as just the percentile.

Crocker, L., & Algina, J. (1986). Introduction to Classical and Modern Test Theory. New York: Holt, Rinehart and Winston. (p. 439)
Loosely speaking, the percentile rank corresponding to a particular raw score is interpreted as the percentage of examinees in the norm group who scored below the score of interest.

*Testing And Assessment: An Employer's Guide To Good Practices. A document by the U.S. Department of Labor Employment and Training Administration (O*Net, 2000)*
Percentile score: The score on a test below which a given percentage of scores fall. For example, a score at the 65th percentile is equal to or higher than the scores obtained by 65% of the people who took the test. https://www.onetcenter.org/dl_files/empTestAsse.pdf (Appendix B, p. 79, B-3)

Pagano, R.R. (1994) Understanding Statistics in the Behavioral Sciences. 4th Edition. New York: West Publishing Company. (p. 44)
A percentile or percentile point is the value on the measurement scale below which a specified percentage of the scores in a distribution fall.

Kline, P. (2000) A Psychometrics Primer. London; Free Association Books. (p. 41) and Kline, P. (2000) A Handbook of Psychological Testing. London: Routledge. (p. 59)
A percentile is defined as the score below which a given proportion of the normative group falls.

Ferguson, G.A. and Takane, Y. (1989) Statistical Analysis in Psychology and Education 6th Edition. New York: McGraw-Hill (p. 482)
If k percent of the members of a sample have scores less than a particular value, that value is the k^{th} percentile point.

Rosenthal, R. and Rosnow, R. (1991) Essentials of Behavioral research: Methods and Data Analysis 2nd Edition. New York: McGraw-Hill. (p. 625)
A percentile is the location of a score in a distribution defining the point below which a given percentage of the cases fall. E.g. a score at the 90th percentile falls at a point such that 90 percent of the scores fall at or below that score.

Cronbach, L.J. (1990) Essentials of Psychological Testing 5th Edition. New York: Harper Collins. (p. 109-110).
"Tony stands third out of 40 on Test A, tenth on test B". Because ranks depend upon the number of persons in the group, we have difficulty when group size changes. Therefore ranks are changed to percentile scores. A percentile rank tells what proportion of the group falls below this person.

Howell, D.C. (2002) Statistical Methods for Psychology 5th Edition. Duxbury Press. (p. 62)
Finally, most of you have had experience with percentiles, which are values that divide the distribution into hundredths. Thus the 81st percentile is that point on the distribution below which 81% of the scores lie.

Glass, G.V. and Hopkins, K.D. (1996) Statistical Methods in Education and Psychology, 3rd Edition. London: Allyn and Bacon. (p. 25)

Percentiles are points in a distribution below which a given *p* percent of the cases lie.

Fisher, L.D. and van Belle, G. (1993) Biostatistics: a methodology for the Health Sciences. New York: Wiley. (Wiley Series in Probability and Mathematical Statistics) (p. 51)
The 25th percentile is that value of a variable such that 25% of the observations are less than that value, and 75% of the observations are greater.

Armitage, P. and Berry, G. (1994) Statistical Methods in Medical Research, 3rd edition. London: Blackwell Science. (p. 34)
The value below which P% of the values fall is called the P^{th} percentile

SPSS Inc. (version 10.05)
Percentiles are values that divide cases according to values below which certain percentages of cases fall. For example, the median is the 50% percentile, the value below which 50% of the cases fall.

The Original 2003 explanation

So, what exactly is it?

A percentile is the point in a distribution **at or below** which a given percentage of scores is found

-or-

The value **below** which P% of the values fall is called the P^{th} percentile

Answer:

In fact, both definitions are correct. What is at fault is the lack of clarity in some cases over what constitutes a "score". Let's use the median to exemplify what's going on.

All authors invariably refer to an observed frequency distribution which is referred to a continuous value, real-number distribution like the Normal Distribution. Further, examples will be given in terms of the median value for a set of scores, which is that number above and below which 50% of the scores in a distribution lie. In short, the 50th percentile. If you recall, the calculation for the median for an odd-numbered set of *ordered* scores is the middle value. So, if there are 5 *ordered* scores, the median is the 3rd score in the series. If it is an equal number of scores (say 4), then the median is the average of the 2nd and 3rd score. Note carefully, this score is sometimes not defined when using integer test scores e.g. take four scores on a test which is scored out of 10, in integer units … 2, 4, 5, 9. The median of these scores is (4+5)/2 = 4.5. This is the 50th percentile score – yet no-one can ever obtain it as the test scores are always 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

So, the **most correct** definition for a percentile is, given this example is:

The value **below which** P% of the values fall is called the P^{th} percentile

as this is the score below which 50% of the observations will lie. And nobody can equal it.

But, now take the scores 2, 4, 5, 8, 9. The median is 5. This is an attainable score. What do we say if someone scores a 5? You guessed it … *the person scores **at** the 50th percentile – attaining a median score*. So the definition that now looks most appropriate in this case is:

A percentile is the point in a distribution **at or below which** a given percentage of scores is found.

So how can both be correct – yet seem to be more appropriate under different conditions? The clue is spread throughout the various texts quoted above. The test score, although in many cases an integer value, is in fact deemed a point-estimate of a hypothetical interval of continuous real-value number scores. So, a test score of 4 is actually considered to be a point-estimate of scores that can range from 3.5 through to 4.4999999999999999999999999999999999999. Therefore, when computing the median of 2, 4, 5, 9 as (4+5)/2 = 4.5, we are in fact computing an average of 4.4999999999999999999999999999999999999 + 4.5 = 4.5 (rounded). The first number is the upper bound of the point-estimate 4.0. The second number is the lower bound of the point-estimate 5.0.

Now take the example 2, 4, 5, 8, 9. The median is 5. But, the upper bound of this number is 5.4999999999999999999999999999999. It is a verbal "shorthand" that states that 5 is the median – in fact the upper bound of the median is 5.49999 etc (note it could also be as low as 4.5 given the definition of a point-estimate number).

So, we have to be very careful with our terminology of what a "score" is actually said to represent. If we are referring to observed, integer-value scores, without any regard to the hypothetical score intervals, then to find the percentile of a distribution of scores requires finding that single observed score which cleanly separates the scores above and below it into an integer percentile. i.e the score value below which 33% of the scores lie, and above which 67% of the score lie. This one score will be the 33rd percentile. However, unless we have extremely large samples of scores (in the thousands), **and** a test score range of exactly 0 to 100 in unit (=1) steps, this is never likely to happen. So, the most efficient way of always being able to compute an exact percentile score is by using a standard formula to calculate any required percentile for any frequency distribution of scores. What this requires however is that we taken into account the upper and lower bound for every integer score – assuming that each exact integer score is actually the middle score of an interval extending 0.5 either side … in which an infinity of continuous, real-valued scores can be theoretically "observed" (which begs the question "how!!?!").

The formula is:

$$P_i = ll + \left(\frac{np - cf}{f_i} \right) \cdot w$$

where

$P_i =$ the i^{th} percentile

$ll =$ the exact lower limit of the interval containing the percentile point

$n =$ the total number of scores

$p =$ the proportion corresponding to the desired percentile

$cf =$ the cumulative frequency of scores below the interval containing the percentile point

$f_i =$ the frequency of scores in the interval containing the i^{th} percentile point

$w =$ the width of the class interval

Let's take an example of some test scores … the EPQR Extraversion scale, with a 0-23 test score range…

Category	Frequency table: LONG_E (EPQR100M.STA) Count	Cumulative Count	Percent	Cumulative Percent
0	9	9	1.475410	1.4754
1	12	21	1.967213	3.4426
2	13	34	2.131148	5.5738
3	17	51	2.786885	8.3607
4	16	67	2.622951	10.9836
5	12	79	1.967213	12.9508
6	15	94	2.459016	15.4098
7	16	110	2.622951	18.0328
8	22	132	3.606557	21.6393
9	26	158	4.262295	25.9016
10	32	190	5.245902	31.1475
11	31	221	5.081967	36.2295
12	36	257	5.901639	42.1311
13	31	288	5.081967	47.2131
14	29	317	4.754098	51.9672
15	33	350	5.409836	57.3770
16	39	389	6.393443	63.7705
17	35	424	5.737705	69.5082
18	29	453	4.754098	74.2623
19	31	484	5.081967	79.3443
20	34	518	5.573770	84.9180
21	39	557	6.393443	91.3115
22	33	590	5.409836	96.7213
23	20	610	3.278689	100.0000
Missing	0	610	0.000000	100.0000

What would be the 75th percentile score – that score below which 75% of the sample score? Well, we can see from the above table that it must be between 18 and 19 … as this is where between 74.26% and 79.34% of the sample scores are found. Applying the formula …
Our scores in this case are single values – no range at all. So, our class intervals are in fact the scores themselves. E.g. 0-0, 1-1, 2-2, 3-3 etc. The exact limits however correspond to ±0.5 around each class interval boundary score – the 0, 1, 2, 3, 4, 5, 6 etc. So, our exact limits are:

0 = -0.5 to +0.5
1 = +0.5 to +1.5
2 = +1.5 to +2.5
3 = +2.5 to +3.5
etc.

Let's re-label the table to correspond with our notation in the formula …

Score	Frequency table: LONG_E (EPQR100M.STA) Exact Limits	Midpoint	f	cf Count	Percent	Cumulative Percent
0	-0.5 to 0.5	0	9	9	1.475410	1.4754
1	0.5 to 1.5	1	12	21	1.967213	3.4426
2	1.5 to 2.5	2	13	34	2.131148	5.5738
3	2.5 to 3.5	3	17	51	2.786885	8.3607
4	3.5 to 4.5	4	16	67	2.622951	10.9836
5	4.5 to 5.5	5	12	79	1.967213	12.9508
6	5.5 to 6.5	6	15	94	2.459016	15.4098
7	6.5 to 7.5	7	16	110	2.622951	18.0328
8	7.5 to 8.5	8	22	132	3.606557	21.6393
9	8.5 to 9.5	9	26	158	4.262295	25.9016
10	9.5 to 10.5	10	32	190	5.245902	31.1475
11	10.5 to 11.5	11	31	221	5.081967	36.2295
12	11.5 to 12.5	12	36	257	5.901639	42.1311
13	12.5 to 13.5	13	31	288	5.081967	47.2131
14	13.5 to 14.5	14	29	317	4.754098	51.9672
15	14.5 to 15.5	15	33	350	5.409836	57.3770
16	15.5 to 16.5	16	39	389	6.393443	63.7705
17	16.5 to 17.5	17	35	424	5.737705	69.5082
18	17.5 to 18.5	18	29	453	4.754098	74.2623
19	18.5 to 19.5	19	31	484	5.081967	79.3443
20	19.5 to 20.5	20	34	518	5.573770	84.9180
21	20.5 to 21.5	21	39	557	6.393443	91.3115
22	21.5 to 22.5	22	33	590	5.409836	96.7213
23	22.5 to 23.5	23	20	610	3.278689	100.0000
Missing			0	610	0.000000	100.0000

where

$P_{75} =$ the 75^{th} percentile

$ll = 18.5 =$ the exact lower limit of the interval containing the percentile point

$n = 610 =$ the total number of scores

$p = 0.75 =$ the proportion corresponding to the desired percentile (note this is nothing more than the percentile expressed as a proportion $(75 \div 100)$

$cf = 453 =$ the cumulative frequency of scores below the interval containing the percentile point

$f_i = 31 =$ the frequency of scores in the interval containing the i^{th} percentile point

$w = 1.0 =$ the width of the class interval

feeding these values into the formula we obtain …

$$P_i = ll + \left(\frac{np - cf}{f_i} \right) \cdot w$$

$$P_{75} = 18.5 + \left(\frac{610 \cdot 0.75 - 453}{31} \right) \cdot 1.0$$

$$P_{75} = 18.5 + \left(\frac{457.5 - 453}{31} \right) \cdot 1.0$$

$$P_{75} = 18.5 + 0.1452 \cdot 1.0 = 18.645$$

So, the 75th percentile is a score of 18.645. This is the score at which 75% of observations will be observed to be below this score. BUT – the score is unattainable as this is an integer scored test. What we actually observe is that 74.26% scores will lie at or below 18, with 79.34% of scores at 19 or below. IF we want to use exact percentiles – then we have to accept that our scores are estimates of hypothetical real-valued continuous numbers, hence a score of 18.645 is perfectly valid under these conditions, and the definition of a percentile is most correctly defined as **the value below which P% of the values fall**.

However, what would the 76th percentile look like …

$$P_i = ll + \left(\frac{np - cf}{f_i} \right) \cdot w$$

$$P_{76} = 18.5 + \left(\frac{610 \cdot 0.76 - 453}{31} \right) \cdot 1.0$$

$$P_{76} = 18.5 + \left(\frac{463.6 - 453}{31} \right) \cdot 1.0$$

$$P_{76} = 18.5 + 0.3419 \cdot 1.0 = 18.84$$

note, all figures remain the same except for the proportion – which changes from 0.75 to 0.76.

So, when an individual scores 19 on a test, what do we conclude?

Here we need to compute the **percentile rank** of the score – which is just the reverse of computing the score for a particular percentile. Now we know the score (=19), but need to compute the percentile for it …

The formula is:

$$PR_x = \left[\frac{\left(cf + \left(\frac{x - ll}{w} \right) \cdot f_i \right)}{n} \right] \cdot 100.0$$

where

$PR_x =$ the percentile rank of score x

$ll =$ the exact lower limit of the interval containing the score x

$n =$ the total number of scores

$cf =$ the cumulative frequency of scores below the interval containing the score x

$f_i =$ the frequency of scores in the interval containing x

$w =$ the width of the class interval

So, for a score of 19, the exact percentile rank is:

$$PR_x = \left[\frac{\left(cf + \left(\frac{x - ll}{w} \right) \cdot f_i \right)}{n} \right] \cdot 100.0$$

$$PR_{19} = \left[\frac{\left(453 + \left(\frac{19 - 18.5}{1.0} \right) \cdot 31 \right)}{610} \right] \cdot 100.0$$

$$PR_{19} = \left[\frac{(453 + 0.5 \cdot 31)}{610} \right] \cdot 100.0$$

$$PR_{19} = 76.80\%$$

a score of 19 is at the 76.8th percentile – the score at which 76.80% of scores will be found to be below this score.

BUT …

All the above is standard fare – and is highly confusing given that only integer value scores can ever be observed. What we know from our observed frequency distribution table is that 79.3443% of individuals scored 19 or below.

But, using actual scores means that only certain % values can be provided – based upon the exact number of frequencies observed for each score. So, there can be no 75th percentile for our observed frequency distribution – only a 74.26th or 79.34th percentile. So …

➊ If you want to assign exact percentile ranks to scores, then you must use the formulae above and assume each integer score is actually a point-estimate from an interval of possible scores. Here, the definition of a percentile is the value **below** which P% of the values fall.

➋ Alternatively, if you simply prefer to state the frequency of people who score at or below an observed test score, then you use the actual frequencies of scores in your normative data. Here, the percentile is the point **at or below** which a given percentage of scores is observed.

The first statement assumes that scores are theoretically continuous but can only be observed as integers; the second assumes scores are simply discrete integers.

June, 2011: A Reality Check

While the above may explain the apparent contradictory definition of a percentile, I think the "continuous underlying score" assumption is just mad, **unless** the test scores really are continuous, real-valued entities, or when the test score range exceeds the range of integer percentiles (so it is possible to have a fractional percentile for an observed fractional score).

But, for the vast majority of applications where scores may wish to be expressed as percentiles within psychological testing, scores usually range between 0 and 100 or less, with an equal interval of 1 between scores. *Indeed, many test scores only range between 0 and 20, which begs the question why use percentiles at all as they will always be "lumpy" [too many percentiles (101) for the actual realizable scores (21)].*

The "continuous" assumption calculation only uses half the frequencies at an observed score. So for the maximum possible score, only half the observed frequencies are used to express the final percentile. Quite simply, it is factually incorrect to state the percentile for this score is the value below which n% of the norm-group score, because half the people scoring the maximum observed score are included in the cumulative proportion. Alternatively, it is also wrong to state the percentile is the score at or below which n% of the group score, because half the people scoring at the percentile score are missing from the calculation.

To correctly express a percentile for a discrete integer score range whose values lie inside a 0 to 100 range, we need to compute the cumulative frequency distribution from the observed frequencies at each integer score, given an integer score is a discrete entity. This distribution is what you find if you computed a standard frequency distribution for a set of scores in say SPSS, STATISTICA, or another stats package. The percentiles of this distribution range between a potential minimum value of 0 (if there are no observed frequencies at 0), and ALWAYS 100.

The interpretation of these percentiles is very simple, accurate, and very clear.

Given person X with percentile rank of say 80:

They score **at or above** n% of the group.
Or
n% of the group score **at or below** a percentile rank of 80.

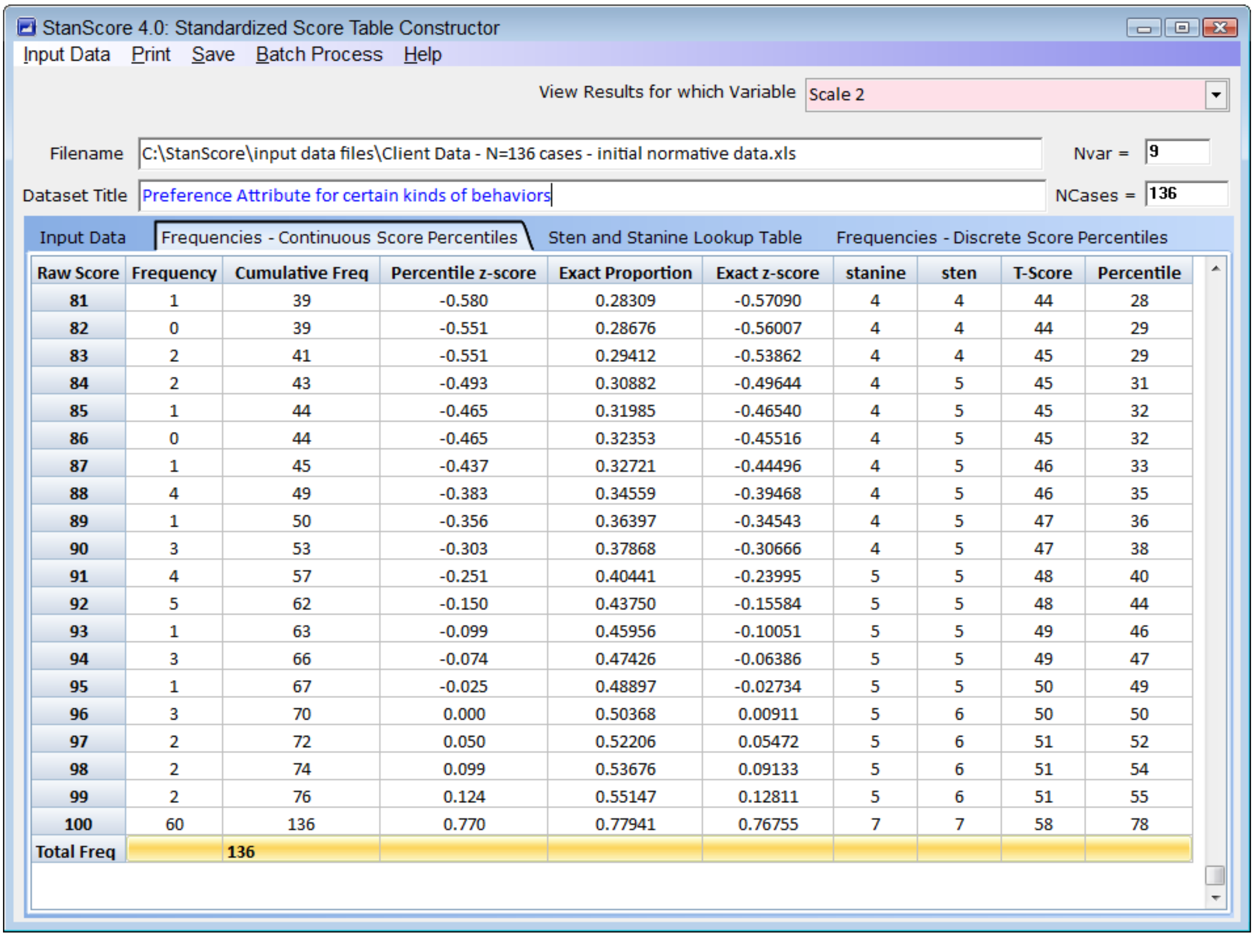
A person receiving the maximum score would always have a percentile rank of 100, which would mean they score at or above 100% of the norm group.

No fuss. Just absolute clarity.

Another more compelling practical reason NOT to use the textbook formulae:

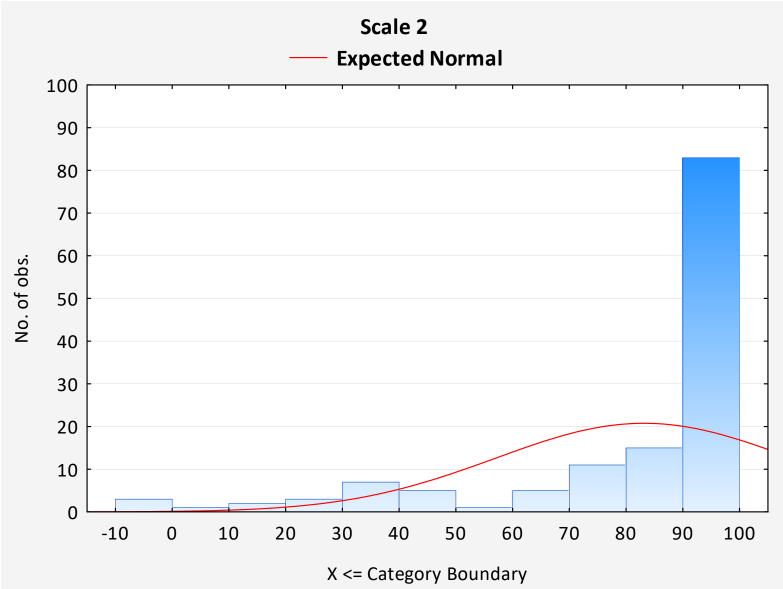
And this is what drove me to rethink entirely the received wisdom about percentiles. I was examining some data using a scale to assess a preference attribute which behaves more as a "threshold" assessment rather than something continuously distributed over the possible range. That is, the majority of people are expected to respond in the upper quadrant of the scale, and at the maximum possible score. The scale score range is actually 0 to 100, in integers, with an equal interval of 1. The client wanted to report percentile scores, interpreted against a norm group consisting of 136 people. Several scales require such normative data to be computed.

I ran the analysis using Stanscore 3.1, and was faced with a problem that the maximum score would be reported as a percentile of as low as 78 on one scale.



Raw Score	Frequency	Cumulative Freq	Percentile z-score	Exact Proportion	Exact z-score	stanine	sten	T-Score	Percentile
81	1	39	-0.580	0.28309	-0.57090	4	4	44	28
82	0	39	-0.551	0.28676	-0.56007	4	4	44	29
83	2	41	-0.551	0.29412	-0.53862	4	4	45	29
84	2	43	-0.493	0.30882	-0.49644	4	5	45	31
85	1	44	-0.465	0.31985	-0.46540	4	5	45	32
86	0	44	-0.465	0.32353	-0.45516	4	5	45	32
87	1	45	-0.437	0.32721	-0.44496	4	5	46	33
88	4	49	-0.383	0.34559	-0.39468	4	5	46	35
89	1	50	-0.356	0.36397	-0.34543	4	5	47	36
90	3	53	-0.303	0.37868	-0.30666	4	5	47	38
91	4	57	-0.251	0.40441	-0.23995	5	5	48	40
92	5	62	-0.150	0.43750	-0.15584	5	5	48	44
93	1	63	-0.099	0.45956	-0.10051	5	5	49	46
94	3	66	-0.074	0.47426	-0.06386	5	5	49	47
95	1	67	-0.025	0.48897	-0.02734	5	5	50	49
96	3	70	0.000	0.50368	0.00911	5	6	50	50
97	2	72	0.050	0.52206	0.05472	5	6	51	52
98	2	74	0.099	0.53676	0.09133	5	6	51	54
99	2	76	0.124	0.55147	0.12811	5	6	51	55
100	60	136	0.770	0.77941	0.76755	7	7	58	78
Total Freq		136							

That's just not sensible. Think about what you would say to a client who might naturally ask on seeing someone with a 78th percentile, "what is the raw score associated with the 100th percentile". What do you answer – some gobbledook about hypothetical continuous scores and an unrealizable raw score being associated with the missing 100th percentile? Or you tell them they don't understand how to correctly interpret a percentile? The problem is being caused by a score distribution which is not only non-normal … but has nearly half the scores at maximum:



So, if we now use Stanscore 4, my program for computing percentiles, and other transformed score lookup tables, and use the discrete score frequencies themselves to compute percentiles, we see:

StanScore 4.0: Standardized Score Table Constructor

Input Data Print Save Batch Process Help

View Results for which Variable: Scale 2

Filename: C:\StanScore\input data files\Client Data - N=136 cases - initial normative data.xls Nvar = 9

Dataset Title: Preference Attribute for certain kinds of behaviors NCases = 136

Input Data | Frequencies - Continuous Score Percentiles | Sten and Stanine Lookup Table | Frequencies - Discrete Score Percentiles

Raw Score	Frequency	Cumulative Freq.	Proportion	Cumulative Propn.	z-score	stanine	sten	T-Score	Percentile
81	1	39	0.007	0.28676	-0.56007	4	4	44	29
82	0	39	0.000	0.28676	-0.56007	4	4	44	29
83	2	41	0.015	0.30147	-0.51741	4	4	45	30
84	2	43	0.015	0.31618	-0.47569	4	5	45	32
85	1	44	0.007	0.32353	-0.45516	4	5	45	32
86	0	44	0.000	0.32353	-0.45516	4	5	45	32
87	1	45	0.007	0.33088	-0.43482	4	5	46	33
88	4	49	0.029	0.36029	-0.35520	4	5	46	36
89	1	50	0.007	0.36765	-0.33568	4	5	47	37
90	3	53	0.022	0.38971	-0.27790	4	5	47	39
91	4	57	0.029	0.41912	-0.20237	5	5	48	42
92	5	62	0.037	0.45588	-0.10970	5	5	49	46
93	1	63	0.007	0.46324	-0.09133	5	5	49	46
94	3	66	0.022	0.48529	-0.03646	5	5	50	49
95	1	67	0.007	0.49265	-0.01822	5	5	50	49
96	3	70	0.022	0.51471	0.03646	5	6	50	51
97	2	72	0.015	0.52941	0.07301	5	6	51	53
98	2	74	0.015	0.54412	0.10970	5	6	51	54
99	2	76	0.015	0.55882	0.14658	5	6	51	56
100	60	136	0.441	1.00000	4.00000	7	7	58	100
Total Freq		136							

Now we get sensible percentiles – lumpy as you'd expect from that score distribution, but at least the 100th percentile does accord to the maximum score on the test.

I wouldn't recommend the client uses norms at all for this kind of test because how can you offer feedback on how they might improve to someone who scores at the 56th percentile, when their raw score is 99 out of 100! That's the problem using a relative scoring scheme on data which is not near-normally distributed.

What test publishers and consultants fail to convey in their training to HR and others is that relative scoring can sometimes yield peculiar "norm tables" when the scores are not distributed near-normally around a central mean value. The more severe the departure from normality, the more problems occur.

Assuming that all attributes required to be "normed" will be distributed normally in any sub-group is somewhat adventurous. Fortunately, some test publishers do make their raw-score to norm-score lookup tables available to purchasers of their tests, or at least provide the discrete score histograms which allow a user to see if any "peculiarities" might be likely to occur.

Personally, I don't like working with percentiles as they really are just ranked values of raw scores. It is actually the raw score which carries the seriously important information for this particular test (and any test) because scoring low really does mean something (*or should mean something*) very different to scoring high. In short, it is the attribute meaning itself which carries the interpretation of the score magnitude, not some relative version of the score.

You now know why I had misgivings about that assumption of a continuous, underlying, score range – and what it did to the expression of a percentile score.

And, beware, mine is an extremely simple-minded view of norms and percentiles. I do not assume norm groups are some random sample from a hypothetical population; I just take them "as they are" – a sample of a group of people who will be used as a comparison group for another individual's scores. Whether or not that group can be considered a representative sampling of some "target" group population is a matter for sometimes deep consideration. I don't try and estimate the hypothetical error around each percentile given my norm-group sample size, or estimate what a real-valued score might be from the integer representation.

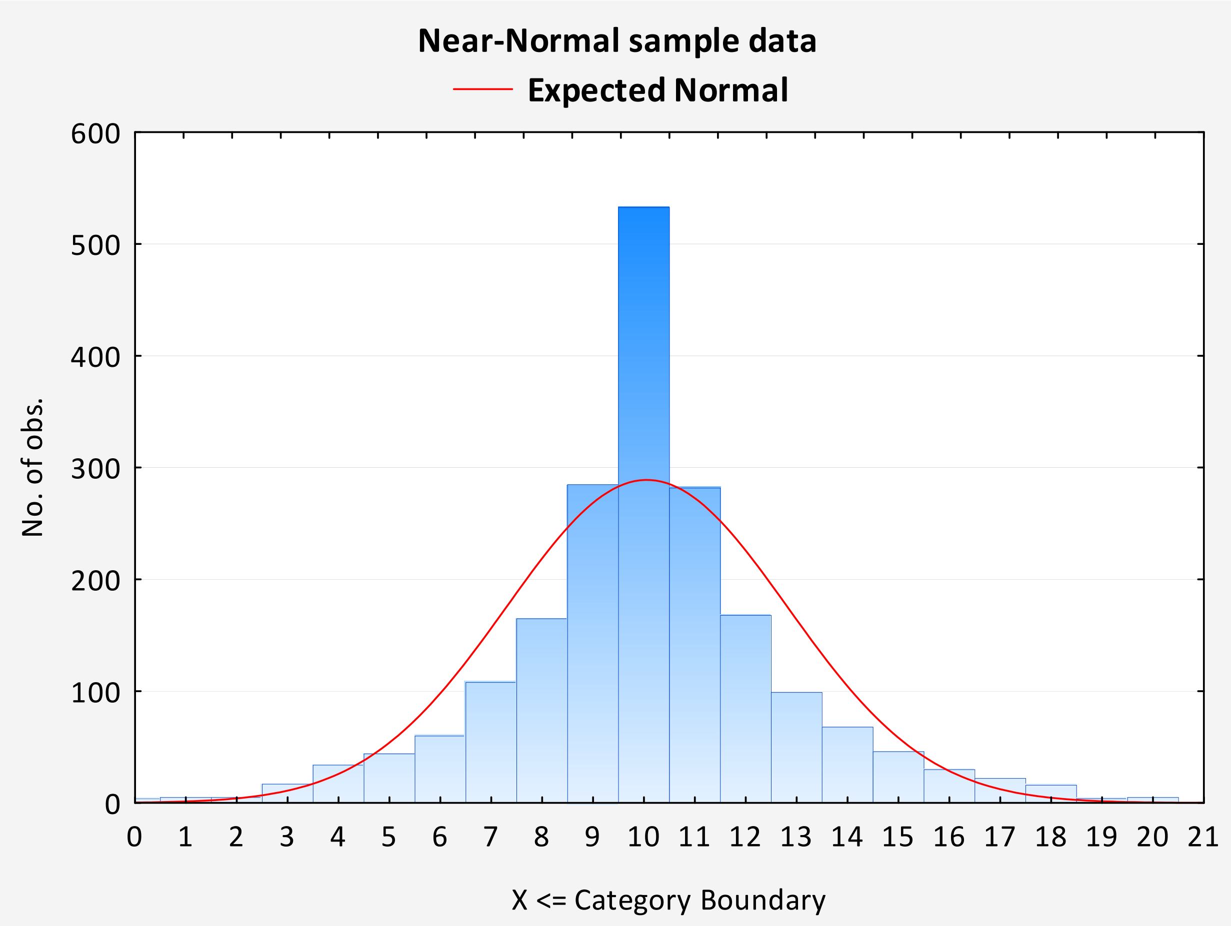
The approach to estimating percentiles which incorporate statistical sampling error as part of their calculation can be found in John Crawford and colleagues' work – an excellent paper outlining the logic, algorithms, and results can be downloaded from John's website:

#160 Crawford, J. R., Cayley, C., Wilson, P. H., Lovibond, P. F., & Hartley, C. (2011). Percentile norms and accompanying interval estimates from an Australian general adult population sample for self-report mood scales (BAI, BDI, CRSD, CES-D, DASS, DASS-21, STAI-X, STAI-Y, SRDS, and SRAS). The Australian Psychologist, 46, 3-14.

See also #151: Crawford, J. R., Garthwaite, P. H., & Slick, D. J. (2009). On percentile norms in neuropsychology: Proposed reporting standards and methods for quantifying the uncertainty over the percentile ranks of test scores. The Clinical Neuropsychologist, 23, 1173-1195.

That "normality" assumption

As a score distribution approaches the typical "bell-shape" of a normal distribution, so do the differences between the computation of percentiles become negligible. For example, here is some data sampled from a perfect normal distribution, truncated integer scores between 0 and 20, 2000 cases, designed to possess a mean of 10 and SD of near 3. …



The Stanscore-4 Continuous-Score and Discrete Score percentile tables are:

Continuous-Score assumption percentiles

Filename C:\StanScore\input data files\Truncated 0-20 integer score near-normal data, 2000 cases.xls Nvar = 1

Dataset Title Optional Data Title, type something here if you wish NCases = 2000

Input Data | Frequencies - Continuous Score Percentiles | Sten and Stanine Lookup Table | Frequencies - Discrete Score Percentiles

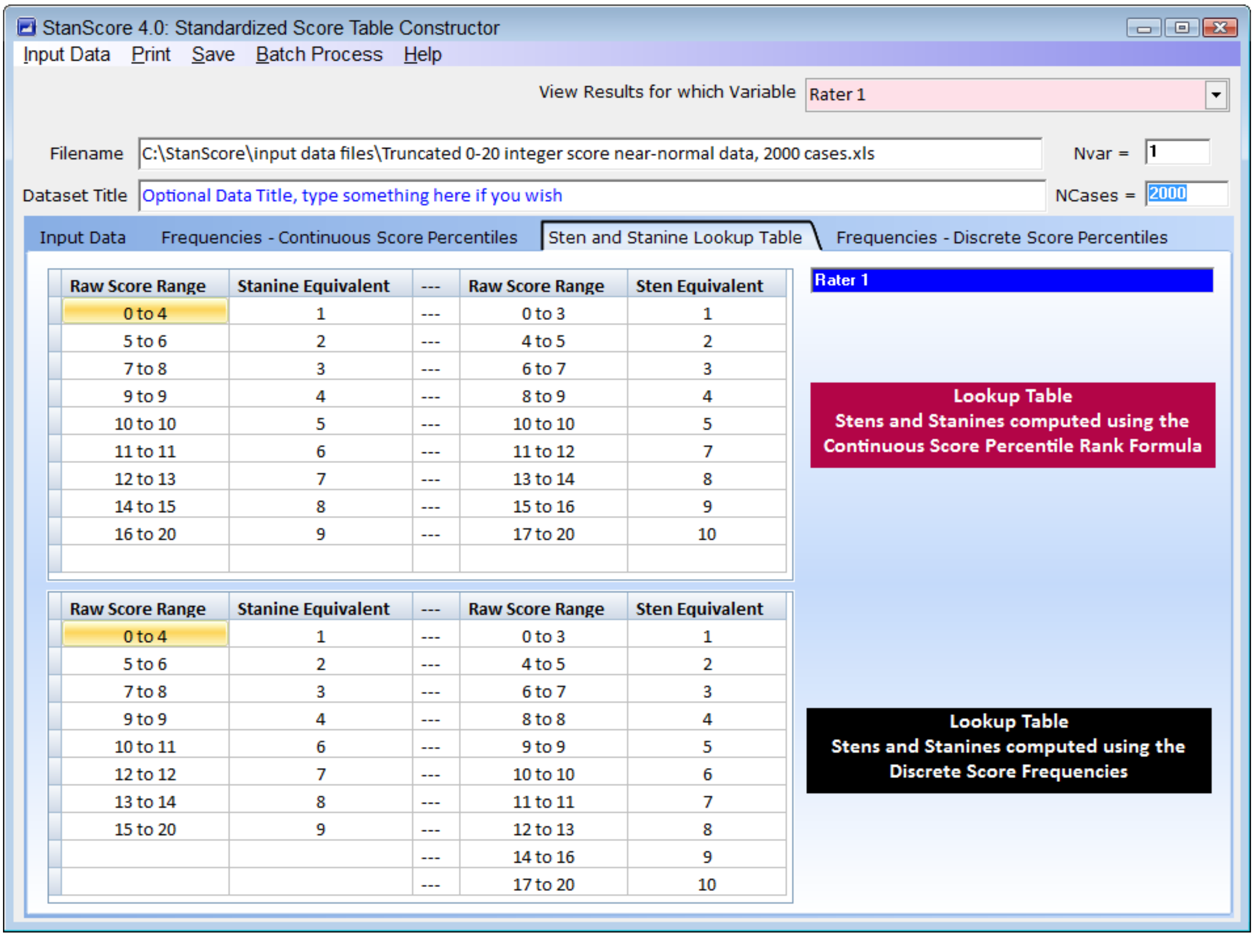
Raw Score	Frequency	Cumulative Freq	Percentile z-score	Exact Proportion	Exact z-score	stanine	sten	T-Score	Percentile
0	4	4	-4.000	0.00100	-3.09262	1	1	19	0
1	5	9	-4.000	0.00325	-2.72340	1	1	23	0
2	5	14	-2.328	0.00575	-2.52879	1	1	25	1
3	17	31	-2.328	0.01125	-2.28303	1	1	27	1
4	34	65	-2.054	0.02400	-1.97790	1	2	30	2
5	44	109	-1.751	0.04350	-1.71127	2	2	33	4
6	60	169	-1.475	0.06950	-1.47868	2	3	35	7
7	108	277	-1.225	0.11150	-1.21698	3	3	38	11
8	165	442	-0.913	0.17975	-0.91393	3	4	41	18
9	285	727	-0.551	0.29225	-0.54404	4	4	45	29
10	533	1260	0.000	0.49675	-0.00805	5	5	50	50
11	282	1542	0.522	0.70050	0.52307	6	7	55	70
12	168	1710	0.875	0.81300	0.88656	7	7	59	81
13	99	1809	1.173	0.87975	1.17200	7	8	62	88
14	68	1877	1.404	0.92150	1.41420	8	8	64	92
15	46	1923	1.644	0.95000	1.64449	8	9	66	95
16	30	1953	1.881	0.96900	1.86654	9	9	69	97
17	22	1975	2.054	0.98200	2.09774	9	10	71	98
18	16	1991	2.328	0.99150	2.38813	9	10	74	99
19	4	1995	4.000	0.99650	2.69877	9	10	77	100
20	5	2000	4.000	0.99875	3.02567	9	10	80	100
Total Freq		2000							

Discrete Score percentiles

Input Data | Frequencies - Continuous Score Percentiles | Sten and Stanine Lookup Table | Frequencies - Discrete Score Percentiles

Raw Score	Frequency	Cumulative Freq.	Proportion	Cumulative Propn.	z-score	stanine	sten	T-Score	Percentile
0	4	4	0.002	0.00200	-2.88033	1	1	21	0
1	5	9	0.003	0.00450	-2.61386	1	1	24	0
2	5	14	0.003	0.00700	-2.45881	1	1	25	1
3	17	31	0.009	0.01550	-2.15802	1	1	28	2
4	34	65	0.017	0.03250	-1.84545	1	2	32	3
5	44	109	0.022	0.05450	-1.60222	2	2	34	5
6	60	169	0.030	0.08450	-1.37427	2	3	36	8
7	108	277	0.054	0.13850	-1.08511	3	3	39	14
8	165	442	0.083	0.22100	-0.76616	3	4	42	22
9	285	727	0.143	0.36350	-0.34668	4	5	47	36
10	533	1260	0.267	0.63000	0.32947	6	6	53	63
11	282	1542	0.141	0.77100	0.73945	6	7	57	77
12	168	1710	0.084	0.85500	1.05607	7	8	61	85
13	99	1809	0.050	0.90450	1.30628	8	8	63	90
14	68	1877	0.034	0.93850	1.54165	8	9	65	94
15	46	1923	0.023	0.96150	1.76835	9	9	68	96
16	30	1953	0.015	0.97650	1.98685	9	9	70	98
17	22	1975	0.011	0.98750	2.24253	9	10	72	99
18	16	1991	0.008	0.99550	2.61386	9	10	76	100
19	4	1995	0.002	0.99750	2.80911	9	10	78	100
20	5	2000	0.003	1.00000	4.00000	9	10	80	100
Total Freq		2000							

The Sten-Stanine Score tables are almost identical, reflecting the difference in calculation of the cumulative score proportions (only half the frequencies used at any raw score), prior to computing a percentile.



Raw Score Range	Stanine Equivalent	---	Raw Score Range	Sten Equivalent
0 to 4	1	---	0 to 3	1
5 to 6	2	---	4 to 5	2
7 to 8	3	---	6 to 7	3
9 to 9	4	---	8 to 9	4
10 to 10	5	---	10 to 10	5
11 to 11	6	---	11 to 12	7
12 to 13	7	---	13 to 14	8
14 to 15	8	---	15 to 16	9
16 to 20	9	---	17 to 20	10

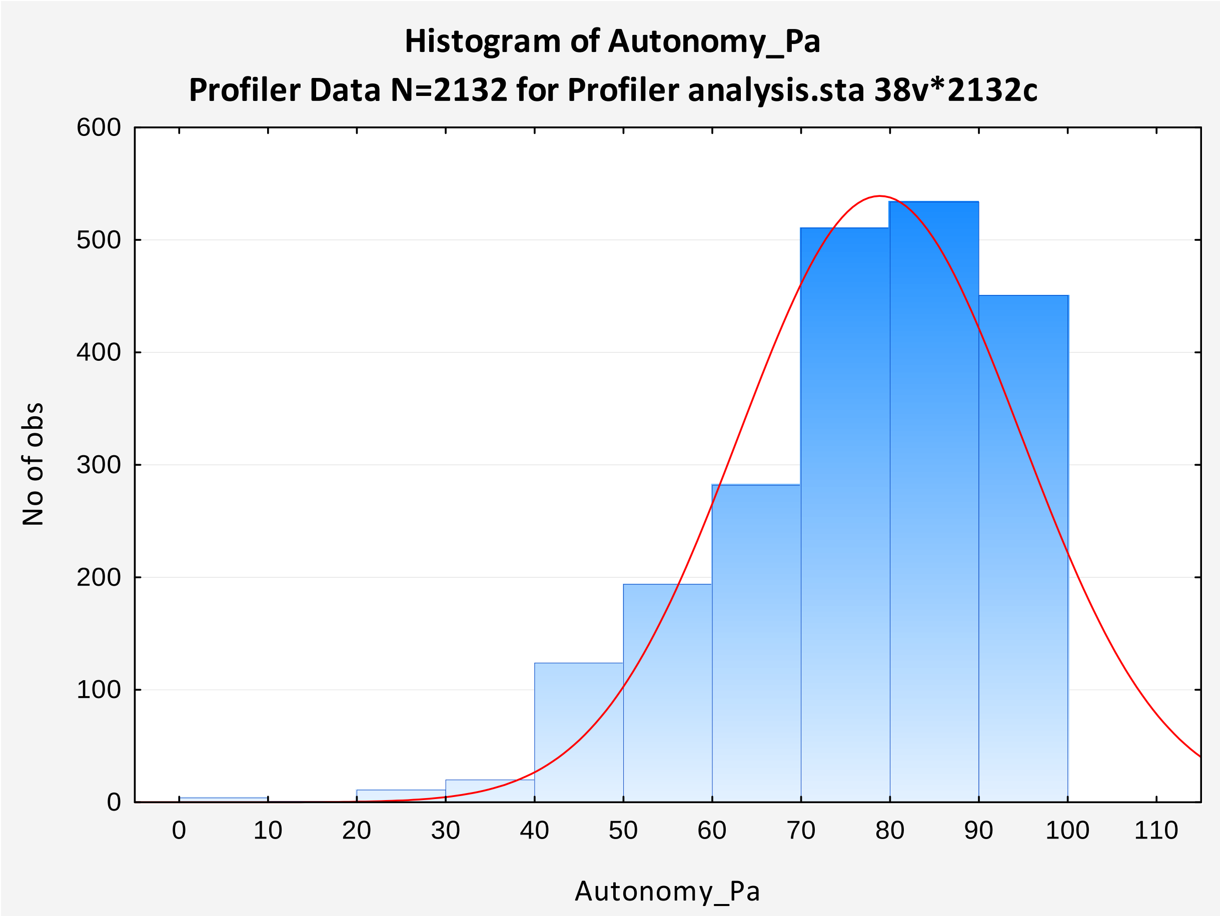
Raw Score Range	Stanine Equivalent	---	Raw Score Range	Sten Equivalent
0 to 4	1	---	0 to 3	1
5 to 6	2	---	4 to 5	2
7 to 8	3	---	6 to 7	3
9 to 9	4	---	8 to 8	4
10 to 11	6	---	9 to 9	5
12 to 12	7	---	10 to 10	6
13 to 14	8	---	11 to 11	7
15 to 20	9	---	12 to 13	8
		---	14 to 16	9
		---	17 to 20	10

BUT … **look at the percentiles around the mean raw score** – highlighted within the red box on each screenshot on the previous page .. you can see what the “normalization” of the continuous-score formula is doing – it’s *flattening out* the actual raw score distribution (where too many scores for a normal distribution are occurring with a score of 10 – the mean score).

That’s a heads-up as to what the percentile formula is doing – essentially normalizing a distribution of scores given a set of observed frequencies.

Your frequencies at each percentile are no longer what you observe, but a transformed version of them. Is that what you really want to know?

Contrast these data with those from the Mariner7 graphical profiler – which assessed preferences for a work activities over a 0 to 100 integer score range:



The Stanscore-4 Continuous-Score and Discrete Score percentile tables are:

Continuous-Score assumption percentiles

Filename C:\StanScore\input data files\Preference data.xls Nvar = 10

Dataset Title Optional Data Title, type something here if you wish NCases = 2132

Input Data | Frequencies - Continuous Score Percentiles | Sten and Stanine Lookup Table | Frequencies - Discrete Score Percentiles

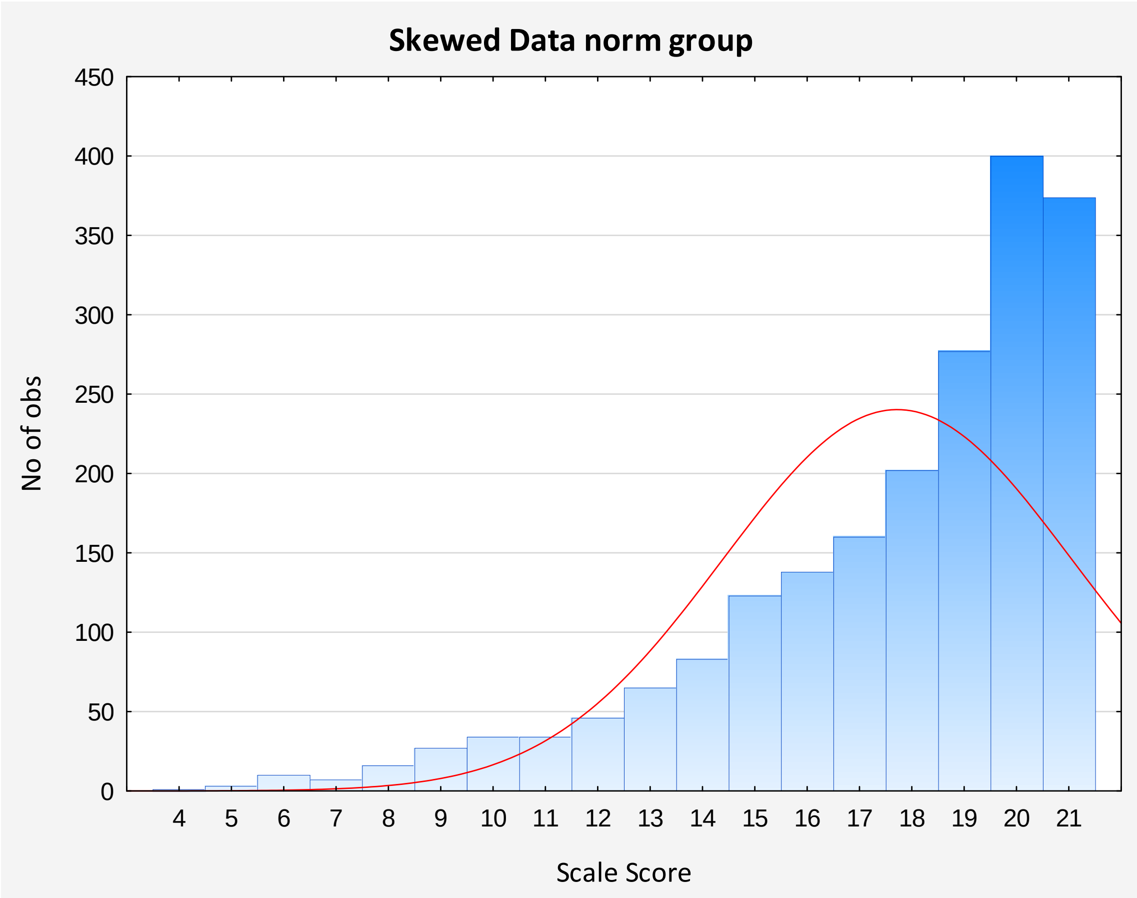
Raw Score	Frequency	Cumulative Freq	Percentile z-score	Exact Proportion	Exact z-score	stanine	sten	T-Score	Percentile
81	36	1183	0.124	0.54644	0.11550	5	6	51	55
82	38	1221	0.150	0.56379	0.15909	5	6	52	56
83	22	1243	0.200	0.57786	0.19469	5	6	52	58
84	22	1265	0.226	0.58818	0.22097	5	6	52	59
85	83	1348	0.277	0.61280	0.28443	6	6	53	61
86	20	1368	0.356	0.63696	0.34790	6	6	53	64
87	22	1390	0.383	0.64681	0.37420	6	6	54	65
88	26	1416	0.410	0.65807	0.40459	6	6	54	66
89	36	1452	0.437	0.67261	0.44445	6	6	54	67
90	229	1681	0.610	0.73476	0.62448	6	7	56	73
91	44	1725	0.839	0.79878	0.83473	7	7	58	80
92	38	1763	0.913	0.81801	0.90541	7	7	59	82
93	29	1792	0.952	0.83372	0.96672	7	7	60	83
94	24	1816	1.034	0.84615	1.01793	7	8	60	85
95	44	1860	1.078	0.86210	1.08784	7	8	61	86
96	8	1868	1.125	0.87430	1.14513	7	8	61	87
97	9	1877	1.173	0.87828	1.16469	7	8	62	88
98	1	1878	1.173	0.88063	1.17641	7	8	62	88
99	8	1886	1.173	0.88274	1.18710	7	8	62	88
100	246	2132	1.554	0.94231	1.57388	8	9	66	94
Total Freq		**2132**							

Discrete Score percentiles

Input Data | Frequencies - Continuous Score Percentiles | Sten and Stanine Lookup Table | Frequencies - Discrete Score Percentiles

Raw Score	Frequency	Cumulative Freq.	Proportion	Cumulative Propn.	z-score	stanine	sten	T-Score	Percentile
81	36	1183	0.017	0.55488	0.13666	5	6	51	55
82	38	1221	0.018	0.57270	0.18160	5	6	52	57
83	22	1243	0.010	0.58302	0.20781	5	6	52	58
84	22	1265	0.010	0.59334	0.23418	5	6	52	59
85	83	1348	0.039	0.63227	0.33547	6	6	53	63
86	20	1368	0.009	0.64165	0.36039	6	6	54	64
87	22	1390	0.010	0.65197	0.38808	6	6	54	65
88	26	1416	0.012	0.66417	0.42122	6	6	54	66
89	36	1452	0.017	0.68105	0.46792	6	6	55	68
90	229	1681	0.107	0.78846	0.79849	7	7	58	79
91	44	1725	0.021	0.80910	0.87211	7	7	59	81
92	38	1763	0.018	0.82692	0.93974	7	7	59	83
93	29	1792	0.014	0.84053	0.99442	7	7	60	84
94	24	1816	0.011	0.85178	1.04202	7	8	60	85
95	44	1860	0.021	0.87242	1.13607	7	8	61	87
96	8	1868	0.004	0.87617	1.15428	7	8	62	88
97	9	1877	0.004	0.88039	1.17523	7	8	62	88
98	1	1878	0.000	0.88086	1.17759	7	8	62	88
99	8	1886	0.004	0.88462	1.19671	7	8	62	88
100	246	2132	0.115	1.00000	4.00000	8	9	66	100
Total Freq		**2132**							

Still pretty similar except for the maximum score … now let's skew things some more …



The Stanscore-4 Continuous-Score and Discrete Score percentile tables are on the next page.

You can see there is now greater divergence.

It's all about checking the shape of the frequency distribution and the terminal frequency at the maximum observed score.

By the way, the descriptive statistics for these data are:

	Descriptive Statistics (Profiler Data N=2132 for Profiler analysis.sta)					
Variable	Valid N	Mean	Median	Std.Dev.	Skewness	Kurtosis
Rater 1	2000	17.64900	19.00000	3.320740	-1.25492	1.194375

In the end, dealing with truncated, small-range, non-normal, integer score distributions, is always going to be awkward when it comes to computing percentiles.

Continuous-Score assumption percentiles

Filename C:\StanScore\input data files\Beta distributed data n=2000 cases.xls Nvar = 1

Dataset Title Optional Data Title, type something here if you wish NCases = 2000

Input Data | Frequencies - Continuous Score Percentiles | Sten and Stanine Lookup Table | Frequencies - Discrete Score Percentiles

Raw Score	Frequency	Cumulative Freq	Percentile z-score	Exact Proportion	Exact z-score	stanine	sten	T-Score	Percentile
0	0	0	-4.000	0.00000	-4.00000	1	1	10	0
1	0	0	-4.000	0.00000	-4.00000	1	1	10	0
2	0	0	-4.000	0.00000	-4.00000	1	1	10	0
3	0	0	-4.000	0.00000	-4.00000	1	1	10	0
4	1	1	-4.000	0.00025	-3.48342	1	1	15	0
5	3	4	-4.000	0.00125	-3.02567	1	1	20	0
6	10	14	-4.000	0.00450	-2.61386	1	1	24	0
7	7	21	-2.328	0.00875	-2.37743	1	1	26	1
8	16	37	-2.328	0.01450	-2.18449	1	1	28	1
9	27	64	-1.881	0.02525	-1.95618	1	2	30	3
10	34	98	-1.751	0.04050	-1.74483	2	2	33	4
11	34	132	-1.554	0.05750	-1.57555	2	2	34	6
12	46	178	-1.404	0.07750	-1.42108	2	3	36	8
13	65	243	-1.225	0.10525	-1.25068	2	3	37	11
14	83	326	-1.078	0.14225	-1.06825	3	3	39	14
15	123	449	-0.875	0.19375	-0.86166	3	4	41	19
16	138	587	-0.641	0.25900	-0.64366	4	4	44	26
17	160	747	-0.437	0.33350	-0.42762	4	5	46	33
18	202	949	-0.200	0.42400	-0.18996	5	5	48	42
19	277	1226	0.099	0.54375	0.10878	5	6	51	54
20	400	1626	0.551	0.71300	0.55938	6	7	56	71
21	374	2000	1.340	0.90650	1.31819	8	8	63	91
Total Freq		2000							

Discrete Score percentiles

Input Data | Frequencies - Continuous Score Percentiles | Sten and Stanine Lookup Table | Frequencies - Discrete Score Percentiles

Raw Score	Frequency	Cumulative Freq.	Proportion	Cumulative Propn.	z-score	stanine	sten	T-Score	Percentile
0	0	0	0.000	0.00000	-4.00000	1	1	10	0
1	0	0	0.000	0.00000	-4.00000	1	1	10	0
2	0	0	0.000	0.00000	-4.00000	1	1	10	0
3	0	0	0.000	0.00000	-4.00000	1	1	10	0
4	1	1	0.001	0.00050	-3.29308	1	1	17	0
5	3	4	0.002	0.00200	-2.88033	1	1	21	0
6	10	14	0.005	0.00700	-2.45881	1	1	25	1
7	7	21	0.004	0.01050	-2.30925	1	1	27	1
8	16	37	0.008	0.01850	-2.08655	1	1	29	2
9	27	64	0.014	0.03200	-1.85239	1	2	31	3
10	34	98	0.017	0.04900	-1.65429	2	2	33	5
11	34	132	0.017	0.06600	-1.50550	2	2	35	7
12	46	178	0.023	0.08900	-1.34571	2	3	37	9
13	65	243	0.033	0.12150	-1.16577	3	3	38	12
14	83	326	0.042	0.16300	-0.97996	3	4	40	16
15	123	449	0.062	0.22450	-0.75441	3	4	42	22
16	138	587	0.069	0.29350	-0.54041	4	4	45	29
17	160	747	0.080	0.37350	-0.32024	4	5	47	37
18	202	949	0.101	0.47450	-0.06328	5	5	49	47
19	277	1226	0.139	0.61300	0.28493	6	6	53	61
20	400	1626	0.200	0.81300	0.88656	7	7	59	81
21	374	2000	0.187	1.00000	4.00000	8	8	63	100
Total Freq		2000							