

العنوان:	Is Grading On The Normal Curue Efficient ? A Test Of Reliability
المصدر:	مجلة علوم وفنون - دراسات وبحوث - مصر
المؤلف الرئيسي:	Abd Almegeed, Samir Mohamed
المجلد/العدد:	مج1, ع2
محكمة:	نعم
التاريخ الميلادي:	1989
الشهر:	ابريل
الصفحات:	35 - 40
رقم MD:	657595
نوع المحتوى:	بحوث ومقالات
قواعد المعلومات:	HumanIndex
مواضيع:	الكفاءة العلمية ، الاختبارات و المقاييس التربوية
رابط:	http://search.mandumah.com/Record/657595

IS GRADING ON THE NORMAL CURVE EFFICIENT ? A TEST OF RELIABILITY

Dr. Samir Mohamed Abdel-Megeed
Assistant Professor
Mathematics Department
Helwan University
Cairo - Egypt

Abstract

In assigning a grade, an instructor rates a student's performance on a scale whose points are labelled from A to F. The use of grading on 5 points (A,B,C,D and F) or 11 points (A,A⁻,B⁺,B,B⁻,C⁺,C,C⁻,D⁺,D and F) are compared with the usage of cut-off points according to percentage system and according to the normal curve. The Kuder-Richardson Formula 21 for Reliability in five cases argued that the usage of the normal curve is not efficient.

1. Introduction:

As we know in studying the effectiveness of different instructional programmes or practices or systems, it is possible that the outcomes measure may be used to assist in determining the utilization of this practice, programme or system. The score received on a good achievement test is meant to reflect the actual knowledge of an individual or group. This score gives us an indication of what a student has achieved. At the same time it gives the student a grade he can use to calculate his grade-point average (GPA). In assigning a grade

and knowing that this grade is going to affect the student's GPA, an instructor usually rates a student's performance on a scale from zero to one hundred. He then aggregates the students into groups according to their marks and assigns a letter grade for each group. This means that the recording of our variable (marks) is used to represent a continuous trait. However, by assigning a letter grade, this variable cannot be considered as strictly continuous but is broken into discrete sections instead.

The number of sections most universities practice could be five or eleven. If we use the five sections, rating a student's performance will be on a scale whose points are labelled from A to F (A,B,C,D, and F) as well as using the eleven sections the scale will also vary from A to F. However, it contains eleven points (A,A⁻,B⁺,B,B⁻,C⁺C,C⁻,D⁺,D, and F). Which is the better way ? What are the cut-off points we are going to use to distinguish the grades from each other ? Should these cut-off points be determined according to percentage system or the normal curve? Finally, is grading on the normal curve efficient ?

2. Methodology:

The problem as outlined in the introduction can easily be solved by designing an experiment to study the difference between the two systems of grading (percentage system or normal curve). In each system we use a scale of five points and another scale of eleven points. Statistically we will compare the reliability of the four ways with the reliability of the original score.

To conduct this research we picked the results of 54 students in a test of 100 items. The appropriate measure of reliability in this case (single test) is The Kuder-Richardson Formula 21 which yields a reasonably close but conservative approximation to coefficient alpha (a) where a is the estimate of reliability. The Kuder-Richardson Formula 21 is symbolized by r_{11} and stated as follows:

$$r_{11} = \frac{n}{n-1} [1 - \frac{m(1 - m/n)}{S^2}]$$

where

- r_{11} is an estimate of reliability
- n is the number of items in the test
- s is the standard deviation of the test, and
- M is the means score of the group

After calculating r_{11} from the original scores we assigned a letter grade for each student according to his mark in the following two schemes using percentage system:

Scheme I:

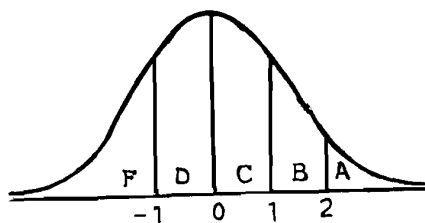
Score	Letter grade	Points
100-85	A	4
84-70	B	3
69-58	C	2
57-50	D	1
49- 0	F	0

Scheme II:

Score	Letter grade	Points
100-93	A	4
92-85	A ⁻	3.7
84-80	B ⁺	3.3
79-72	B	3
71-70	B ⁻	2.7
69-62	C ⁺	2.3
61-60	C	2
59-58	C ⁻	1.7
57-53	D ⁺	1.3
52-50	D	1
49- 0	F	0

Then the z-value for the original scores has been calculated and a letter grade is assigned for each student according to his standard score in Schmes III and IV as follows:

Scheme III:



In this scheme we assigned a letter grade according to the normal curve. We considered that any student whose mark is lower than the means with one standard deviation or more should fail. The scheme can be outlined as follows: (this applies to scheme IV as well).

Table 1 Points of the Schemes

Score Schemes	I	II	Z	III	IV	Score	I	II	z	III	IV
95	4	4	2.00	4	4	59	2	1.7	0	2	2
93	4	4	1.89	3	3.7	59	2	1.7	0	2	2
92	4	3.7	1.83	3	3.7	58	2	1.7	-.05	1	1.7
86	4	3.7	1.50	3	3.3	55	1	1.3	-.22	1	1.7
84	3	3.3	1.39	3	3.3	55	1	1.3	-.22	1	1.7
82	3	3.3	1.28	3	3	55	1	1.3	-.22	1	1.7
82	3	3.3	1.28	3	3	54	1	1.3	-.28	1	1.7
80	3	3.3	1.17	3	3	54	1	1.3	-.28	1	1.7
76	3	3	.95	2	2.7	54	1	1.3	-.28	1	1.7
74	3	3	.83	2	2.7	54	1	1.3	-.28	1	1.7
72	3	3	.72	2	2.7	53	1	1.3	-.33	1	1.7
71	3	2.7	.67	2	2.7	52	1	1	-.39	1	1.3
71	3	2.7	.67	2	2.7	52	1	1	-.39	1	1.3
70	3	2.7	.61	2	2.3	52	1	1	-.39	1	1.3
70	3	2.7	.61	2	2.3	51	1	1	-.44	1	1.3
67	2	2.3	.45	2	2.3	50	1	1	-.50	1	1.3
67	2	2.3	.45	2	2.3	50	1	1	-.50	1	1.3
65	2	2.3	.33	2	2.3	47	0	0	-.67	1	1
63	2	2.3	.22	2	2	45	0	0	-.78	1	1
63	2	2.3	.22	2	2	43	0	0	-.89	1	1
62	2	2.3	.17	2	2	40	0	0	-1.05	0	0
62	2	2.3	.17	2	2	36	0	0	-1.28	0	0
61	2	2	.11	2	2	35	0	0	-1.33	0	0
61	2	2	.11	2	2	28	0	0	-1.72	0	0
60	2	2	.06	2	2	22	0	0	-2.05	0	0
60	2	2	.06	2	2	13	0	0	-2.55	0	0
59	2	1.7	0	2	2	11	0	0	-2.76	0	0

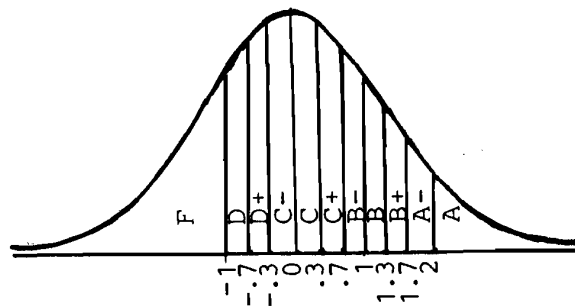
Summary Statistics

Table 2

Original Score	Scheme I	Scheme II	Scheme III	Scheme IV
58,98	1.72	1.77	1.57	1.85
18.01	1.20	1.18	.94	1
324.36	1.45	1.39	.89	1
.9347	.432	.826	-.1	.70

Standard score	Letter grade	Points
2 or more	A	4
1 and less than 2	B	3
0 and less than 1	C	2
-1 and less than 0	D	1
less than -1	F	0

Scheme IV



Standard score	Letter grade	Points
1.99 or more	A	4
1.66-1.98	A ⁻	3.7
1.33-1.65	B ⁺	3.3
0.99-1.32	B	3
0.66-0.98	B ⁻	2.7
0.33-0.65	C ⁺	2.3
0 -0.32	C	2
(-.33)-(-.01)	D ⁻	1.7
(-.66)-(-.34)	D ⁺	1.3
(-1) -(-.67)	D	1
less than (-1)	F	0

3. Results:

The results are shown in the following tables:

* in calculating r_{11} for schemes I and III to apply the Kuder-Richardson Formual 21 we had to consider $n=4$. We also considered $n=10$ for schemes II and IV. Moreover, in schemes II and IV in order to consider $n=10$ and get an appropriate estimate for s we used a linear transformation to convert the points 4,3,7,3,3, to 10,9,8 and then calculator r_{11} using the new values i.e. we used the new values means and standard deviation (the correlation coefficient between the old points and the new ones is .9996).

4. Discussion and Conclusion:

In order to answer the previous research questions we used the coefficient alpha (α) as an estimate for reliability which gives us the accuracy of the measurement. The reliability of the original scores was .9347 which indicates that this test was a very good instrument for measuring the students performance. The scheme whose reliability is closer to the original one is the best scheme. From the summary statistics it is obvious that Scheme II (grading with the usage of percentage system on eleven points) gives the highest reliability (.826). This assures that the use of (+ & -) with the letter grade is more accurate than using only letter grades without plus or minus whether the cut-off points were determined according to the percentage system or the normal curve. Also, grading on the normal curve does not give higher reliability whether we use 5-points scale or 11-points scale which means that the usage of the normal curve in grading is not efficient.

References:

- NNNALLY. Psychometric Theory. McGraw-Hill; 2nd., 1978.
- SUSLOW, S. A report on an interinstitutional survey of undergraduate scholastic grading, 1960's to 1970's. Office of Institutional Research, University of California, Berkeley; 1976.
- THORNDIKE & HAGEN. Measurement and Evaluation in Psychology and Education. John Wiley & Sons Inc., 4th ed., 1977, p. 83.
- WHITE, E.M. Sometimes an A is really and F. Chronicle of Higher Education; Feb. 1975.