THE PROBABILITY THAT GRADES IN AN AVERAGE SIZED CLASS WILL APPROXIMATELY FIT A NORMAL CURVE[1]

R. O. HUTCHINSON

Head of Mathematics Department, Tennessee Polytechnic Institute

During the past two decades numerous studies have been made of the distribution of college grades and evidence of the utter lack of a scientific basis for assigning grades continues to accumulate. Comparative studies of grades have revealed the rather disturbing fact that the chance which a student has of becoming an honor student may be multiplied considerably by merely changing his course, or even by taking the same course under a different instructor.

The remedy for this situation is not so apparent. When we reflect that the accuracy of all measurements in the physical sciences depends upon one's ability to determine when two units are equal, we are brought face to face with the fundamental difficulty in measuring student accomplishment in college. For how can one determine whether a given accomplishment in mathematics is greater or less than a second given accomplishment in history? Obviously not by abstract ideals in the minds of the professors.

In recent years there has been a marked tendency to define college grades in terms of some distribution curve, usually the normal probability curve. For example an A-grade in any subject could be defined as representing the kind of work done by the best 7 per cent of the students taking the course. Similarly other grades could be defined. Some of the most frequently adopted distributions, together with their authors are given in table 1.

The distribution given by Gray is very nearly the average of all the suggested distributions. It approximates very closely the actual distribution of nearly 9000 grades in elementary courses at Harvard as reported by W. T. Foster (1911). This distribution has therefore been adopted in this investigation.

The problem may be stated concretely thus: Assume a group of 10,000 students with abilities distributed according to Gray, namely, 700 A's, 2200 B's, 4200 C's, etc. What distribution may be expected in an average sized class selected at random? Corey (1930) attacked this problem mechanically by taking 100 cards and placing grades on these in accordance with this distribution, shuffling the cards, and dealing them out in piles of 25 cards each. This method, though very direct, has two important limitations:

1. It is not practical to shuffle more than 100 cards and that is too small a number for statistical purposes.

---

[1]Read before the Tennessee Academy of Science at the Cookeville Meeting, April 17, 1936.

TABLE 1

Frequently adopted grade distributions

Author	Grade distribution A	B	C	D	E
Gray	7	22	42	22	7
Rugg	7	24	38	24	7
Corey	8	24	36	24	8
Meyer	3	22	50	22	3
Dearborn	2	23	50	23	2
Cattell	10	20	40	20	10
Mean	6.2	22.5	42.7	22.5	6.2

2. It would be quite laborious to deal out and analyze enough stacks to make the results very reliable.

By use of the theory of probability both these limitations may be removed, for it is easy to deal with an infinitely large number of students and this in turn will give an infinitely large number of classes.

Probability of a Given Number of Students With a Given Grade

Let N be the total number of students

n be the number in one class

A be the number of A-grade students in all classes

a be the number of A-grade students in the given class

B be the number of B-grade students in all classes

b be the number of B-grade students in the given class, etc.

Then the probability that a class of "n" students will contain exactly "a" students of A-ability and (n -a) students of some other ability is

$$p'_a = \frac{{}_A C_a \; {}_{(N-A)} C_{(n-a)}}{{}_N C_n} \qquad (1)$$

But each of these three combinations may be expanded thus:

$${}_A C_a = \frac{A(A-1)\,(A-2)\ldots\ldots(A-\overline{a-1})}{a!}$$

$$\frac{A^a}{a!}\left(1-\frac{1}{A}\right)\left(1-\frac{2}{A}\right)\ldots\ldots\left(1-\frac{a-1}{A}\right).$$

Substituting an expression similar to this for each of the combinations and replacing N^n by $N^a \cdot N^{n-a}$, we get

$$P'_a = \left[(A/N)^a (1 - A/N)^{n-a} \frac{n!}{a!(n-a!)} \right] (H/G), \qquad (2)$$

where

$$H = \left[(1\text{-}1/A)(1\text{-}2/A)\ldots(1 - \frac{a\text{-}1}{A})\right]\left[(1 - \frac{1}{N\text{-}A})\ldots(1 - \frac{n\text{-}a\text{-}1}{N\text{-}A})\right]$$

and

$$G = (1 - 1/N)\,(1 - 2/N)\ldots(1 - \frac{n\text{-}1}{N})$$

If N is now allowed to increase indefinitely, the ratio (H/G) will approach "one" as a limit, but the ratio (A/N) will remain constant. Let this ratio be r_a.

$$\text{Then } p_a = \lim_{N \to \infty} p'_a = \frac{r^a\,(1-r)^{n-a}\,n!}{a!\,(n-a)!} \qquad (3)$$

This equation could be used for calculating p_a for all values of a, but, after one value of p_a is found it is much easier to calculate other values by means of equations (4) and (5) which are derived by substituting (a + 1 and (a — 1) respectively for a in (3) and simplifying.

$$p_{a+1} = p_a \frac{n-a}{a+1}\,\frac{r}{1-r} \qquad (4)$$

$$p_{a-1} = p_a \frac{a}{n-a+1}\,\frac{1-r}{r} \qquad (5)$$

In these equations "a" is the number of students of A-ability in the class of "n" students and r is the ratio (A/N) which we are assuming to be (.07). These same equations may be used for calculating the probability of a class having "b" students of B-ability by merely substituting "b" for "a" and using (B/N) = .22 for r, and similarly for each of the other grades.

Table 2 gives the results of these calculations for a class of 25 students. Opposite each value of a in the first column is a number expressing the probability that a given class of students will have this number of students having the ability listed at the top of the column. For example, in the column headed B or D, the number .1788 is opposite the 6 in the first column. This means that the probability of a class of 25 students containing 6 B-students (exactly) is .1788, or that 17.88 per cent of all classes should contain exactly six B-students. On account of the symmetrical distribution assumed the probabilities for A and for E students is the same and also for B and D students.

From this table it may be observed that 58 per cent of the classes should contain either 1 or 2 A-students and that not more than 3 per

TABLE 2

The probability that a class of twenty-five students will have "a" A-students, and corresponding data for other grades

a	A or E	B or D	C
0	.1629	.0020	.0000
1	.3066	.0141	.00002
2	.2769	.0478	.0002
3	.1598	.1039	.0011
4	.0662	.1605	.0042
5	.0209	.1902	.0129
6	.0052	.1788	.0311
7	.0011	.1369	.0611
8	.0002	.0869	.0996
9	.00003	.0463	.1363
10	----------	.0208	.1579
11	----------	.0080	.1559
12	----------	.0026	.1317
13	----------	.0007	.0954
14	----------	.0002	.0592
15	----------	.00004	.0314
16	----------	---------	.0143
17	----------	---------	.0055
18	----------	---------	.0018
19	----------	---------	.0005
20	----------	---------	.0001

cent of all classes should contain more than 4 A-students. Likewise for E-students. Fifty-three per cent of the classes should contain 4, 5, or 6 B (or D) students; and 58 per cent should contain 9 to 12 C-students. It must not be assumed, however, that more than 50 per cent of the classes should contain 1 or 2 A-students, 1 or 2 E's, 4 to 6 B's, 4 to 6 D's, and 9 to 12 C's, for some of the classes which contain 1 or 2 A's will not contain 4 to 6 B's, etc.

Probabilities of Certain Combinations of Grades

Equation (1) gave the probability that a class would contain exactly "a" students of A-ability and (n - a) of any other ability whatever. It is now desired to find the probability that a class will contain exactly "a" A-students and at the same time contain "b" B-students, "c" C-students, "d" D-students, and "e" E-students, a problem in compound probability. This is given by the equation

$$p' = \frac{{}_{A}C_{a} \cdot {}_{B}C_{b} \; {}_{C}C_{c} \; {}_{D}C_{d} \; {}_{E}C_{e}}{{}_{N}C_{n}}, \qquad (6)$$

By expanding each of these combinations as in (1) and replacing N^n by the product $(N^a N^b N^c N^d N^e)$, this equation reduces to

$$p' = (A/N)^a (B/N)^b (C/N)^c (D/N)^d (E/N)^e \frac{n!}{a!\; b!\; c!\; d!\; e!} \cdot S, \qquad (7)$$

TABLE 2

The probability that a class of twenty-five students will have "a" A-students, and corresponding data for other grades

a	A or E	B or D	C
0	.1629	.0020	.0000
1	.3066	.0141	.00002
2	.2769	.0478	.0002
3	.1598	.1039	.0011
4	.0662	.1605	.0042
5	.0209	.1902	.0129
6	.0052	.1788	.0311
7	.0011	.1369	.0611
8	.0002	.0869	.0996
9	.00003	.0463	.1363
10	----------	.0208	.1579
11	----------	.0080	.1559
12	----------	.0026	.1317
13	----------	.0007	.0954
14	----------	.0002	.0592
15	----------	.00004	.0314
16	----------	---------	.0143
17	----------	---------	.0055
18	----------	---------	.0018
19	----------	---------	.0005
20	----------	---------	.0001

cent of all classes should contain more than 4 A-students. Likewise for E-students. Fifty-three per cent of the classes should contain 4, 5, or 6 B (or D) students; and 58 per cent should contain 9 to 12 C-students. It must not be assumed, however, that more than 50 per cent of the classes should contain 1 or 2 A-students, 1 or 2 E's, 4 to 6 B's, 4 to 6 D's, and 9 to 12 C's, for some of the classes which contain 1 or 2 A's will not contain 4 to 6 B's, etc.

Probabilities of Certain Combinations of Grades

Equation (1) gave the probability that a class would contain exactly "a" students of A-ability and (n - a) of any other ability whatever. It is now desired to find the probability that a class will contain exactly "a" A-students and at the same time contain "b" B-students, "c" C-students, "d" D-students, and "e" E-students, a problem in compound probability. This is given by the equation

$$p' = \frac{{}_{A}C_{a} \; {}_{B}C_{b} \; {}_{C}C_{c} \; {}_{D}C_{d} \; {}_{E}C_{e}}{{}_{N}C_{n}}, \qquad (6)$$

By expanding each of these combinations as in (1) and replacing N^n by the product $(N^a \, N^b \, N^c \, N^d \, N^e)$, this equation reduces to

$$p' = (A/N)^a (B/N)^b (C/N)^c (D/N)^d (E/N)^e \; \frac{n!}{a!\, b!\, c!\, d!\, e!} \cdot S, \qquad (7)$$

where S is an expression of the same general form as (H/G) in (2). As in that case the limit of S is one as "N" increases indefinitely. The ratios $r_a = A/N$, $r_b = B/N$, etc., are defined by our original distribution. Hence for very large values of N, (7) becomes

$$p = (.07)^{a+e}\ (.22)^{b+d}\ (.42)^{c}\ \frac{n!}{a!\ b!\ c!\ d!\ e!} \qquad (8)$$

where p is the limiting value of p′ as "N" increases indefinitely.

For an adjacent distribution in which "b" is increased by 1 and "c" decreased by 1, the probability is given by

$$p_i = p\ (r_b/r_c)\ \frac{c}{b+1} = p\ (22/42)\ \frac{c}{b+1} \qquad (9)$$

Equation (9) is a special case of a general equation for the change in p due to small integral changes in a, b, c, d, and e subject to the condition that the algebraic sum of all such changes must be zero. However, in general, it was found most convenient to allow only two letters to change, each by unity and in opposite directions. Equation (9) or its equivalent was used in calculating most of the probabilities given in Table 3 but (8) was occasionally used as a check.

TABLE 3

The probability of definite combinations of grades in a class of twenty-five

i	a	e	b	d	c	1000Pi	f	1000fPi
1	1	1	6	6	11	3.3869	1	3.3869
2	1	2	5	6	11	3.2331	4	12.9324
3	1	1	5	6	12	3.2331	2	6.4662
4	1	2	6	6	10	3.1043	2	6.2086
5	2	2	5	5	11	3.0861	1	3.0861
6	1	2	5	5	12	3.0861	2	6.1722
7	2	2	5	6	10	2.9632	2	5.9264
8	1	1	5	7	11	2.9031	2	5.8062
9	1	1	5	5	13	2.8487	1	2.8487
10	1	1	6	7	10	2.7880	2	5.5760
11	1	2	5	7	10	2.6608	4	10.6432
12	2	2	6	6	9	2.5869	1	2.5869
13	2	2	4	6	11	2.5717	2	5.1434
14	1	2	4	6	12	2.5717	4	10.2868
15	2	2	4	5	12	2.4548	2	4.9096
16	1	1	4	6	13	2.3739	2	4.7478
17	1	2	6	7	9	2.3233	4	9.2932
18	1	1	4	7	12	2.3093	2	4.6186
19	1	2	4	5	13	2.2660	4	9.0640
20	2	2	5	7	9	2.2173	2	4.4346
21	2	2	4	7	10	2.1166	2	4.2332
22	1	1	7	7	9	2.0862	1	2.0862
23	2	2	4	4	13	1.8021	1	1.8021
24	2	2	6	7	8	1.7425	2	3.4850
25	1	2	7	7	8	1.5647	2	3.1294
26	1	2	4	7	11	1.9237	4	3.6948

The first column in Table 3 gives the number used to specify the particular combination of grades under discussion. The next five columns give the combination for this particular case. The next column gives the probability (multiplied by 1000) that this particular combination will occur. For example, when i is 6 the probability that there will be 1 A-grade, 2 E-grades, 5 B-grades, 5 D-grades, and 12 C-grades is .0030861. Since a and e are everywhere interchangeable, the combination of 2 A-grades, 1 E-grade, and other grades as given in i = 6 is equally probable. Hence the frequency (f) of this combination is given as 2 in the 8th column. The most probable distribution of grades is found to be 1 A, 1 E, 6 B's, 6 D's, and 11 C's, but this combination occurs only three times in a thousand classes. If we were not restricted to integral values, the most probable distribution would occur when

$$
\begin{aligned}
a = e &= .07 \times 25 = 1.75 \\
b = d &= .22 \times 25 = 5.50 \\
\text{and } c &= .42 \times 25 = 10.50
\end{aligned}
$$

As pointed out above it could not be expected that more than about 3 classes in a thousand would be distributed according to the distribution of abilities of the original "N" students. It then seemed desirable to determine how much variation from this assumed distribution should be expected. A 25 per cent deviation from this most probable theoretical distribution given above was arbitrarily chosen for the first specific test.

By taking 25 per cent of 1.75, 5.5, and 10.5 and adding and then subtracting from these numbers it is found that the a and e could not be less than 1.3 nor more than 2.2, that the b and d could not be less than 4.1 nor more than 6.9, and that c could not be less than 7.9 nor more than 13.1. The nearest approach to this that could be found was 1-2 for a and e, 4-7 for b and d, and 8-13 for c. This gives an average deviation of slightly more than 26 per cent. Table 3 contains the probabilities for all possible combinations within this range. Each of these probabilities is multiplied by its frequency (f) and the results added to get the probability that in a class of 25 students the distribution of grades will not deviate by more than 26 per cent from the theoretically most probable distribution. It is thus found that the probability is .1426. This means that if the classes were all selected at random, 14 per cent of the classes would not deviate from the most probable distribution by more than 25 per cent in any grade while 86 per cent would deviate by more than 25 per cent in one or more grades.

A complete study of this problem would necessitate making similar calculations for classes of various sizes and also for deviations of various amounts. Only two such calculations have been carried out to date. For a class of 15 students the probability of a deviation of less than 25 per cent was found to be about .06, that is, for a class of this size, 94 per cent of the classes could be expected to deviate from the most probable distribution by more than 25 per cent. For a class of 25 students the probability of a deviation of less than 50 per cent from the most probable distribution was found to be about .70.

Conclusions

It should be emphasized that this is a mathematical study and that the conclusions are based on very definite assumptions, as mathematical conclusions always are. This analysis makes no attempt to verify or disprove these assumptions. The assumptions are: (1) That ability in every subject is distributed among a very large number of college students according to a normal distribution curve, or, what amounts to the same thing, that the five grades used are defined by the college faculty in such a way that grades in all subjects over a period of a large number of years will fit the normal curve; and (2) that the students in each class are selected entirely at random without prejudice to subject matter, instructor, or schedule.

1. Since in classes from 15 to 25 in size there are not more than 6 to 14 per cent of the classes which, if graded accurately, will come within 25 per cent of the normal curve, the value of the normal curve in assigning grades to a class of this size is practically nil.

2. While it is perfectly possible to have 5 A's in a class of 25 this should not on an average occur oftener than once in every 50 classes. More than half the classes should have either one or two A's. The same is true for the E's (F's in some marking systems). The restrictions on the B's and D's are less marked and the restrictions on the C's is very much less.

3. While this study was confined to the normal curve, it is believed that the results would not differ greatly if the assumed curve were skewed considerably.

4. There is nothing in this study to invalidate the use of some distribution curve as a check on past grades over a period of years. The author believes that such a check is valuable as a guide for the future and for comparing the grading scales of different instructors and different departments.

Bibliography

Foster, W. T. 1911. *Administration of the College Curriculum.* Houghton Mifflin Co., New York.

Corey, S. M. 1930. Use of Normal Curve for Grading in Small Classes. *School and Society,* 21: 514-516.

Rugg, H. O. 1917. *Statistical Methods Applied to Education.* Houghton Mifflin Co., New York.

Meyer, M. F. 1911. *Science,* N. S., 33: 661.

Dearborn, W. 1910. *University of Wisconsin* Bull. No. 368, H. S. Series No. 9.

Cattell, J. M. *Popular Science Monthly,* 66: 367.

Received May 28, 1936.