EXAM MASTER Academic Manager
Advanced Score Reporting Definitions

Variance: Variance is a measure of the spread of scores for this exam. The larger the variance, the larger the distance of the scores from the group mean

Standard Deviation: Standard Deviation is the square root of the variance for a set of exam. The practical value of understanding the standard deviation of a set of values is in appreciating how much variation there is from the mean. A large standard deviation indicates that the exam scores are far from the mean and a small standard deviation indicates that they are clustered closely around the mean.

Written out the formula for variance is:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2$$

Where N is the number of exams
i is the current exam (or cycle for programming)
x_i is the current exam score
$\overline{x}$ is the mean exam score
σ^2 is the standard deviation squared (variance is the square of the standard deviation and standard deviation is the square root of the variance.)

thus if you have 5 exams graded (75,80,78,87,90) it would work out:

$$\begin{aligned}
\sigma^2 &= (1/5) * ((75-82)^2 + (80-82)^2 + (78-82)^2 + (87-82)^2 + (90-82)^2) \\
&= (1/5) * ((-7)^2 + (-2)^2 + (-4)^2 + (5)^2 + (8)^2) \\
&= (1/5) * (49 + 4 + 16 + 25 + 64) \\
&= (1/5) * (158) \\
&= 31.6
\end{aligned}$$

And in turn the standard deviation for this would be:

$$\begin{aligned}
\sigma &= (31.6)^{1/2} \\
&\approx 5.62
\end{aligned}$$

For larger groups (30 and up) it is better to use N-1 (degrees of freedom) thus

$$\begin{aligned}
\sigma &= (1/4) * (158) \\
&= 39.5
\end{aligned}$$

And in turn the standard deviation for this would be:

$\sigma = (39.5)^{1/2}$
≈ 6.28

Degree of Skewness: A quality of the distribution of a set of exam scores dealing with whether the scores are symmetrically distributed around a central point. (i.e. bell curve) The higher or lower the Degree of Skewness the more off balance the symmetry of the exam is.

$$g_1 = \frac{m_3}{m_2^{3/2}} = \frac{\sqrt{n}\,\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{3/2}},$$

Where N is the number of exams
i is the current exam (or cycle for programming)
x_i is the current exam score
$\bar{x}$ is the mean exam score

thus if you have 5 exams graded (75,80,78,87,90) it would work out:

$$g_1 = \frac{(\sqrt{5})\left((75-82)^3 + (80-82)^3 + (78-82)^3 + (87-82)^3 + (90-82)^3\right)}{\left(\sqrt{\left((75-82)^2 + (80-82)^2 + (78-82)^2 + (87-82)^2 + (90-82)^2\right)}\right)^3}$$

$$= \frac{(\sqrt{5})\left((-7)^3 + (-2)^3 + (-4)^3 + (5)^3 + (8)^3\right)}{\left(\sqrt{\left((-7)^2 + (-2)^2 + (-4)^2 + (5)^2 + (8)^2\right)}\right)^3}$$

$$= \frac{(\sqrt{5})(-343 + -8 + -64 + 125 + 512)}{\left(\sqrt{(49 + 4 + 16 + 25 + 64)}\right)^3}$$

$$= \frac{(\sqrt{5})(222)}{\left(\sqrt{(158)}\right)^3}$$

$$= .24995$$

Degree of Kurtosis: A quality of the distribution of a set of exam scores dealing with the peak of the exam curve symmetry. A high kurtosis exam curve has a sharper "peak" and flatter "tails", while a low kurtosis distribution has a more rounded peak with wider "shoulders".

$$g_2 = \frac{m_4}{m_2^2} - 3 = \frac{n\,\sum_{i=1}^{n}(x_i - \overline{x})^4}{\left(\sum_{i=1}^{n}(x_i - \overline{x})^2\right)^2} - 3$$

Where N is the number of exams

i is the current exam (or cycle for programming)
x_i is the current exam score
$\bar{x}$ is the mean exam score

thus if you have 5 exams graded (75,80,78,87,90) it would work out:

$$g_2 = \frac{5\,(\,(75-82)^4 + (80-82)^4 + (78-82)^4 + (87-82)^4 + (90-82)^4\,)}{(\,(75-82)^2 + (80-82)^2 + (78-82)^2 + (87-82)^2 + (90-82)^2\,)^2} - 3$$

$$= \frac{5\,(\,(-7)^4 + (-2)4 + (-4)^4 + (5)^4 + (8)^4\,)}{(\,(-7)^2 + (-2)^2 + (-4)^2 + (5)^2 + (8)^2\,)^2} - 3$$

$$= \frac{5\,(\,2401 + 16 + 256 + 625 + 4096\,)}{(\,49 + 4 + 16 + 25 + 64\,)^2} - 3$$

$$= \frac{5\,(\,7394)}{158^2} - 3$$

$$= \frac{36970}{24964} - 3$$

$$= -1.5191$$

KR-21 KUDER-RICHARDSON RELIABILITY INDEX is a simple, one step formula which can be applied to the results of any test that has been scored on the basis of the number of correct answers.

$$r_{K-R21} = (k/(k-1))\left[1 - \frac{\bar{x}\;\overline{k - (x)}}{k\;\sigma.\sigma}\right]$$

where $\bar{x}$ = mean score of the test
k = number of items on the test
σ.σ = variance for the test

thus if you have 5 exams graded (75,80,78,87,90) and the exam had 100 questions it would work out:

r^{K-R21} = (100/(100-1)) [1 – ((82(100-82))/ (100(31.6))]
1.0101 [1 – (1476/ 3160)]
1.0101[1 – (0.467)]
1.0101[0.533]
0.538

Point Biserial: A measure of a questions reliability based on a comparison of how well those who answered correctly performed as opposed to those who answered incorrectly. A higher Point Biserial is indicative of a good question.

$$r = \frac{(\bar{X}_1 - \bar{X}_0)\sqrt{p(1-p)}}{S_X}$$

Where

$\bar{X}_0$ is the mean of X when Y=0

$\bar{X}_1$ is the mean of X when Y=1

S_X is the standard deviation of X

p is the proportion of values where Y=1

Y=0 is an incorrect response
Y=1 is a correct response

Thus, in following with our previous examples:

you have 5 exams graded (75,80,78,87,90)
3 students answered this question correctly and their scores were 75,87,90
2 answered incorrectly and their scores were 80, 78
The standard deviation is 5.62

r= $\frac{((75+87+90)/3) - ((80+78)/2)\ \sqrt{((3/5)*(1-(3/5)))}}{5.62}$

r= $\frac{84 - 79\ \sqrt{(.6*.4)}}{5.62}$

r= $\frac{5\ \sqrt{(.24)}}{5.62}$

r= $\frac{5 * 0.489898}{5.62}$

r= $\frac{2.44949}{5.62}$

r= 0.435852

This question would not be very valid as a discriminator.