On an identity between the Gaussian and Rasch measurement error distributions: making the role of the instrument explicit

D Andrich
Graduate School of Education, The University of Western Australia, 35 Stirling Highway, 6009 Crawley, Western Australia
Email: david.andrich@uwa.edu.au

Abstract. The paper shows an identity between the Gaussian error distribution of replicated measurements and the Rasch model error distribution of inferred, replicated assessments in ordered categories when the thresholds defining the categories are equidistant. The paper concludes by articulating three senses in which the discrete Rasch distribution, in which the role of the instrument is made explicit as it is in modern physics, can be seen as a generalisation of the continuous Gaussian distribution.

The Gaussian is a theoretically derived distribution of random errors of measurement when the same object is measured with the same instrument under constant conditions. It culminated from attempts to justify the mean as the best estimate of the measure when individual measurements produced a distribution [1]. The derivation challenged the best minds in statistics in the 18th Century, for example, Gauss, De Moivre, Lagrange, La Place, and others [1, 2]. A function of the Gaussian is to diagnose whether or not observed, or inferred, replicated measurements approximate it; otherwise, it implies that an unaccounted for factor governs the distribution. The Gaussian, like all random error distributions, is unimodal with smooth transitions between adjacent probabilities - it is strictly log-concave [3].

1. Gauss’s derivation
Among other assumptions, Gauss assumed a symmetric, continuous distribution in which the range of the instrument did not affect the measurements [1]. Gauss appreciated that these assumptions were approximations.

Gauss commented that (1) (the distribution) cannot represent a law of error in full rigor because it assigns probabilities greater than zero to errors outside the range of possible errors, which in practice always has finite limits; that such a feature is unavoidable because one can never assign limits of error with absolute rigor; but this shortcoming is of no importance in the case of (1), because it “decreases so rapidly, when hx has acquired a considerable magnitude, that it can safely be considered as vanishing.” [4, p. 2].

In the quote, \( x = X - \mu \) is a deviation from the mean; in the following, “\( x \)” is a measurement. A complementary discrete Gaussian distribution \( P_{nx} \) over the range of possible measurements \( x = 0,1,2,\ldots,m \) of object \( n \) with measure \( \mu_n \) measured in the units of instrument \( i \) is obtained from
where \( \gamma_{ni} = \sum_{x=0}^{m} \exp\left(\frac{-\left(x - \mu_n\right)^2}{2\sigma_n^2}\right) \) is the ordinate of the continuous Gaussian and \( \sigma_n^2 \) is its variance. The normalisation eliminates \( 1/\sqrt{2\pi\sigma} \), but retains \( \sigma_n^2 \), giving

\[
P_{ni} = \frac{\gamma_{ni}}{\sum_{x=0}^{m} \gamma_{ni}}, \quad x = 0, 1, 2, \ldots, m,
\]

where \( \gamma_{ni} = \sum_{x=0}^{m} \exp\left(\frac{-\left(x - \mu_n\right)^2}{2\sigma_n^2}\right) \) ensures that \( \sum_{x=0}^{m} P_{ni} = 1 \). In assessing whether an empirical discrete distribution approximates the Gaussian, advantage is taken of the virtual identity between \( P_{ni} \) and \( \gamma_{ni} \) when the observed distribution is well within the range of possible measurements, required by Gauss and illustrated below. Although other random error distributions were proposed, the Gaussian is pre-eminent; it is the basis of the t, F and Chi-square distributions. In Gauss’s derivation, the measurements are given and the role of the instrument is not present.

2. Rasch’s probability distribution of discrete measurements

From a starting point entirely different from Gauss’s, the Danish mathematician Georg Rasch derived a general, discrete probability distribution for the assessment of an object with an instrument composed of well-defined ordered categories. More akin to modern physics, the role of the instrument is explicit and equivalent to that of the object. The derivation was based on the requirement that, within a specified frame of reference, the comparison between pairs of objects be invariant with respect to instruments used, and vice versa [5]. Andrich [6] and Stone and Stenner [7] show the centrality of invariance in Rasch’s measurement theory. To realise the invariance in estimation and the interpretation of parameters, two further derivations were necessary [8, 9]. These derivations gave the distribution

\[
P_{ni} = \left[\exp\left\{\frac{1}{2}\left(\frac{\tau_{ix} + x(\beta_n - \delta_i)}{\gamma_{ni}}\right)^2\right\}\right], \quad \gamma_{ni} = \sum_{x=0}^{m} \exp\left\{\frac{-\left(x - \mu_n\right)^2}{2\sigma_n^2}\right\},
\]

where \( P_{ni} \) is the probability of a response \( x, \quad x = 0, 1, 2, \ldots, m \), \( \beta_n \) is the measure of object \( n \), \( \tau_{ix}, \quad x = 0, 1, 2, \ldots, m \) are the \( m \) thresholds of instrument \( i \) which partition the continuum into \( m+1 \) ordered categories, \( \delta_i = \frac{\sum_{x=0}^{m} \tau_{ix}}{m} \) is the mean of the thresholds, and \( \gamma_{ni} \) is a normalising constant. If successive thresholds are equidistant, \( \tau_{i(x+1)} - \tau_{ix} = \Delta_i > 0 \) [9], then

\[
P_{ni} = \left[\exp\left\{\frac{x(m-x)}{2}\Delta_i + x(\beta_n - \delta_i)\right\}\right]/\gamma_{ni}.
\]

In equation (4), which is central to the paper, the value \( x \) can be taken as the measurement of object \( n \) using instrument \( i \). Figure 1 shows the probability response curves for each measurement \( x = 0, 1, 2, \ldots, 10 \) as a function of the measure \( \beta \) for equidistant thresholds, \( \Delta_i = 0.80 \). Then provided there are at least two conformable responses for each object of the set, the estimates \( (\hat{\delta}_i, \hat{\Delta}_i) \) can be obtained independently of any measures of the objects [8]. Inserting these estimates in equation (4) gives the inferred distribution of replicated measurements with instrument \( i \) of any object with measure \( \beta \). If \( \beta_n = \delta_i \), the object is at the centre of the range of the instrument and equation (4) becomes

\[
P_{ni} = \left[\exp\left\{(x(m-x)/2)\Delta_i\right\}\right]/\gamma_{ni}.
\]
Though tedious to show, if the mean $\mu_n = m/2$ is at the centre of the thresholds, then equations (2) and (5) are identical with $\sigma_i^2 = 1/\Delta_j$. However, $\sigma_i^2$ is a property of the instrument’s unit, not merely the variance $\sigma_n^2$ of measurements. In addition, $\Delta_j = h$ in the Eisenhart quote above.

Figure 1. The probability curves for the Rasch distribution: $\delta_i = 0, \Delta_j = 0.8, \sigma_i^2 = 1.25$.

3. Relationship between Rasch’s and Gauss’s distributions
Illustratively, Figure 2 shows two distributions for the instrument in Figure 1 and the maximum measurement is 10. In Panel A, $\beta = 0, E[x] = \mu = m/2 = 5$; in Panel B, $\beta = 3.8, E[x] = \mu = 9$. In Panel A the three distributions are identical; in Panel B, they diverge.

Figure 2. The Gaussian and the Rasch distributions with central and extreme locations.

It is shown readily that if $\delta_i = 0, \Delta_j = 0.8$, and $-2.0 < \beta < 2.0$, i.e $2.5 < \mu < 7.5$, then the three distributions are virtually identical, outside this range they begin to diverge. Figure 3 shows distributions around this margin, with Panel A showing identical distributions, and Panel B divergent ones.
Figure 3. The Gaussian and the Rasch distributions at the margin where they diverge.

The Rasch distribution of equation (4) is a generalisation of the Gaussian in three senses. First, it characterises the instrument’s unit (whose inverse is the variance of the measurements); second, consistent with measurements it is discrete; third, it accounts for the instrument’s operating range. The explicit properties \((\delta, \Delta)\) in equation (4) which govern the measurement \(x\) make the instrument’s role explicit. This distribution may be relevant when the object cannot be aligned experimentally within the range of the instrument and when, as a consequence, the probabilities of extreme scores are not 0.

4. Comment

After the fact it might not seem surprising that when its thresholds are equidistant and the measure of the object is at the mid-point of the range of the instrument, that the Rasch distribution is identical to the discrete Gaussian and that the inverse of its unit is the variance of the complementary latent, continuous Gaussian. On the other hand it also seems remarkable that two such independent mathematical derivations converge to the same equation, but where in the Rasch distribution the role of the instrument and its unit is made explicit. Perhaps unlike the Gaussian, by Gauss’s own admission, the Rasch distribution can be taken to “…represent a law of error in full rigor”.

References

[1] Eisenhart C 1983 Law of error I: development of the concept Encyclopedia of statistical sciences vol 4 ed S Kotz and N L Johnson (Toronto: Wiley) pp 530-47
[2] Stigler S M 1986 The history of statistics: The measurement of uncertainty before 1900 (Cambridge, Massachusetts: The Belknap Press of Harvard University Press)
[3] Keilson J and Gerber H 1971 Some results for discrete unimodality The Journal of the American Statistical Association 66 386-9
[4] Eisenhart C 1983 Law of error II: development of the concept Encyclopedia of statistical sciences vol 4 ed S Kotz and N L Johnson (Toronto: Wiley) pp 547-62
[5] Rasch G 1961 On general laws and the meaning of measurement in psychology Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability IV ed J Neyman (Berkeley CA: University of California) pp 321-34
[6] Andrich D 2018 Advances in social measurement: A Rasch measurement theory Perceived health and adaptation in chronic disease: Stakes and future challenge ed F. Guillemin, A Leplège, S Briançon, E Spitz, and J Coste (CRCS Press, Taylor and Francis) chapter 7 pp 66-91
[7] Stone M H and Stenner A J 2014 Comparison is Key Journal of Applied Measurement 15 26-39
[8] Andersen E B 1977 Sufficient statistics and latent trait models Psychometrika 42 69-81
[9] Andrich D 1978 A rating formulation for ordered response categories Psychometrika 43 561-74