ISBN 0-85066-354-7

# MULTIDIMENSIONAL GROWTH STANDARDS

100

Harvey Goldstein

University of London Institute of Education

## 1 Univariate Growth Standards

All norms or standards involve the concept of an 'atypical' value which one may detect, and hence possibly study further an individual having such a value. In children, because measurements change with age, it is sensible to provide age related norms, and for many measurements there are also important differences between distinct sub-populations or groups of children for example boys and girls, so that it may also be sensible and efficient to prepare separate norms for each group. (See Goldstein and Tanner, 1980, for a general discussion of sub-population standards). In nearly all norms in current use, and certainly those concerned with body measurements, atypical individuals are taken to be those in the 'tails' of the distribution. A common convention is to take regions including the upper and lower 3% of the population values as the 'atypical' regions.

According to Tanner (1981), in his comprehensive history of human growth studies, the first simple cross-sectional percentile growth standards appeared at the end of the 19th century and velocity standards during the 1930's. It was not until the 1950's, however, that Tanner (1952) himself first provided a satisfactory formalisation of growth norms and discussed, among other matters, so called conditional norms such as standards of weight for height, and more recently Cameron (1980) has developed conditional longitudional standards for height to replace traditional velocity standards. In such standards the distribution of a meassurement at one age or occasion is studied for given



Human Growth - A Multidisciplinary Review, edited by A. Demirjian

Pages 97 - 108

values of the measurement at a past occasion usually one year previously and by computing the regression of the second occasion measurement on the first, conditional standards can then be constructed in a straightforward manner. The choice of a one-year interval is made in order to avoid possible seasonal effects. Both the velocity and the conditional standards use pairs of measurements, but nevertheless are unidimensional in the sense that the distribution of a single-valued function of the two measurements is used to provide the atypical region. Thus, however useful such standards may be, they will not necessarily summarise all the useful information in the underlying 2-dimensional distribution. For example, an individual may be 'typical' for height as judged by the conditional standards for height at age 7.0 given height at age 6.0 but at the same time also be 'atypical' for height at age 7.0. Thus, for example, suppose a girl at age 6.0 years living in Britain has a height of 99.0 cm, and a height of 109.5 cm at age 7.0 years. Her conditional percentile is above the 25th percentile indicating satisfactory progress, yet she lies below the 3rd percentile at both ages. Hence, in considering a pair of measurements we need to recognise that any unidimensional summary loses information which could be useful. The question arises therefore as to how to construct 2-dimensional (and indeed multi-dimensional)growth standards.

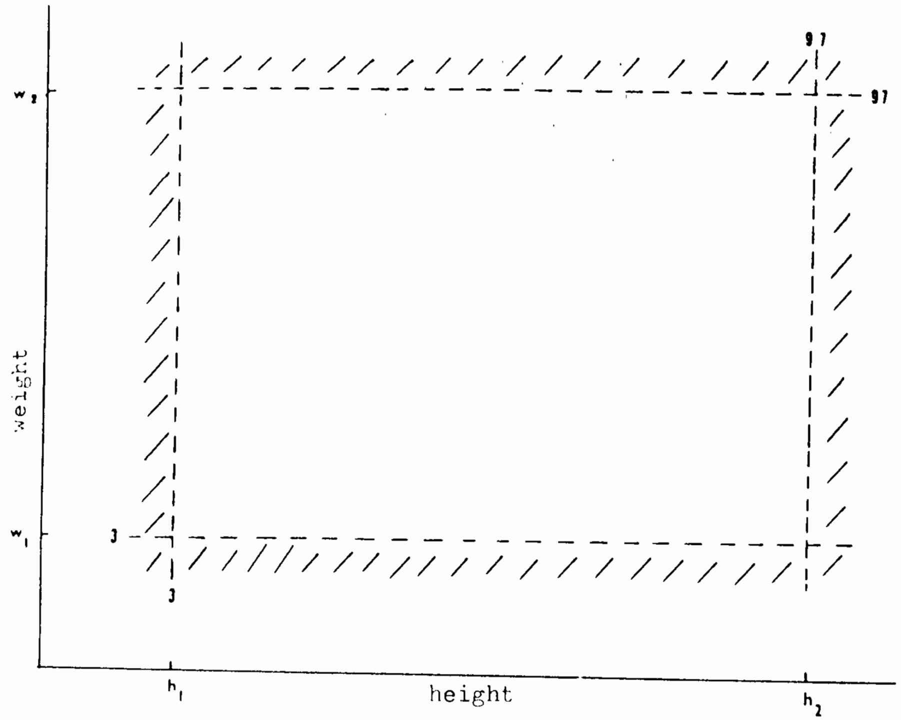
## 2 Two-dimensional Standards

Consider Figure 1. This shows two measurements, and I use height and weight to illustrate the method, with lines drawn to define, separately for each measurement, atypical regions. Thus individuals with values outside the interval $(h_1, h_2)$ are defined as 'atypical' for height and likewise those outside $(w_1, w_2)$ are 'atypical' for weight. For simplicity these regions are assumed to be symmetric. If each 'atypical' region includes 6% of the population values, then the shaded area, i.e. the region which inlcudes individuals who are atypical on at least one measurement, clearly includes more than 6% of the population. If the distribution is in fact bivariate normal this region actually will include 11% of all individuals. If we focus on the bivariate distribution and classify values as atypical if they lie outside a chosen region of the height-weight plane then it is



clear that in some sense they should be those that lie far from the mode of the distribution.

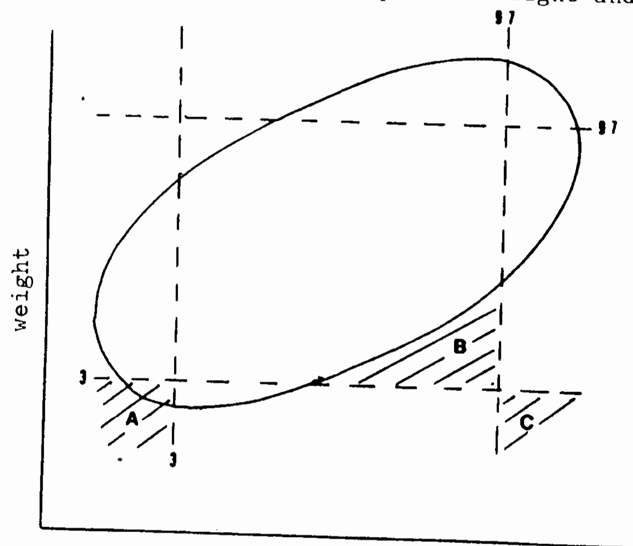
Figure 1 Unidimensional height and weight percentiles




Consider therefore Figure 2, which is one approach which has been used to define the shape of a 2-dimensional atypical region (Defrise-Gussenhoven, 1954 gives one of the earliest applications). The ellipse is a so-called equiprobability ellipse, in this case based on a bivariate normal distribution, and designed to 'exclude' 6% of the population. It has several drawbacks, however. For example, the shaded region labelled B contains values which are atypical according to the 2-dimensional ellipse, but which are typical according to both unidimensional standards and there are also some values, in the region labelled A, which are atypical according to both unidimensional standards but typical according to the 2-dimensional standard. (This is also known, in a slightly different context, as Rao's paradox, see Healy (1969).) The undesirability which stems from having values in both regions A & B could be avoided either by increasing the size of the ellipse to enclose the typical region defined by the unidimensional standards or by shrinking it to lie wholly within this region. In the first case, the % lying in the atypical region would become very small and in the second case rather large, so that neither procedure is really a satisfactory solution. Clearly this situation is undesirable and some other approach is needed.

If we study Figure 2 we see that an obvious deficiency of the ellipse is that it pays no attention to the direction of atypicality. There are infinitely many possible ways of drawing a closed figure to exclude a given percentage of the population and there seems to be no clear a priori reason for preferring an equiprobability ellipse over any other shape. It is worth noting that we are not discussing here the easier problem of defining an atypical region when there is an outcome variable and where we are searching for a function of the two variables in order to predict the outcome variable. (See for example, Hellier and Goldstein, 1979). A fundamental characteristic of growth standards is that they are general purpose indicators or screening instruments, rather than concerned with the prediction of specific outcomes.

Figure 2 Equiprobability 6% ellipse for height and weight



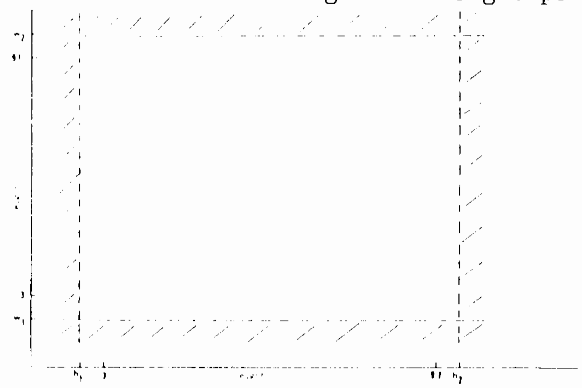



## 3 Consistent Two-dimensional Standards

Consider Figure 3 which is like Figure 1, but with $(h_1, h_2)$ and $(w_1, w_2)$ extended, so that the shaded area now contains 6% of the population, and where in terms of unidimensional standards, each separate atypical region will now include less than 6% of values.

This diagram has the following properties. First, any individual judged atypical by the 2-dimensional standards is also judged to be atypical by at least one (6%) unidimensional standard. Secondly there are also individuals who may be judged atypical by one or both unidimensional standards but who are typical according to the 2-dimensional standard. This, however, is quite reasonable since if we use both unidimensional standards at 6% the extra information provided by a second measurement will increase the total number of atypical values and if we wish to keep the % of typical values constant at 6%, then we must allow that only the more extreme individuals will now be detected. From the point of view of any individual, she now has to exhibit a more extreme value on at least one measurement to count as 'atypical' which is clearly desirable, since otherwise as the number of measurements increases, the use of separate unidimensional 6% regions will increase the chance of at least one value being atypical, leading in the limit to the unsurprising but useless conclusion that every individual is atypical in some respect. This point is worth emphasising especially in the context of multiple screening projects when the use of conventional atypical judgements for each measurement in the screening battery will lead to too many 'extreme' cases.

Figure 3 Two-dimensional height and weight percentiles



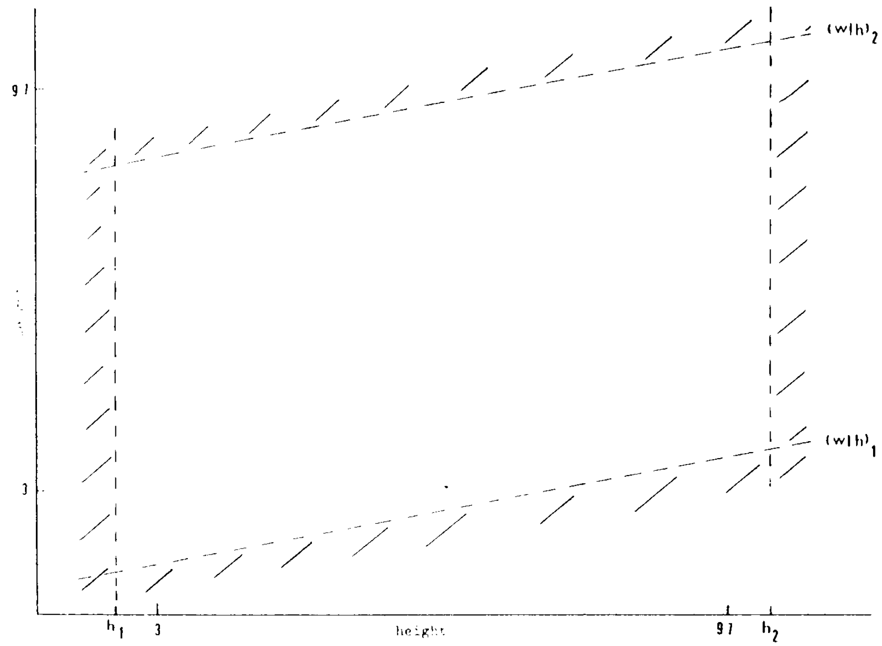
If we adopt the procedure implicit in Figure w, there is now a consistency of judgements in going from 2-dimensional to unidimensional standards, although we still have to decide on the relative weights to be given to each measurement, that is, the lengths of $(h_1, h_2)$ and $(w_1, w_2)$, consistent with an overall 6% atypical region. (Note that 1-sided or asymetrical regions can be incorporated into the argument without difficulty). For example, we can move $h_1$ and $h_2$ further apart and hence attach less importance to height in judging atypicality. The relative lengths of $(h_1, h_2)$ and $(w_1, w_2)$ will depend on substantive considerations which I will not discuss in detail, but, for example, we should want to study the cross hatched area if we are interested, say, in malnutrition. In the limit, if we became uninterested in height, we would then be back with the unidimensional region for weight. Thus, for the 2-dimensional standard we need one parameter in addition to that which specifies the size of the total region, namely the ration of the size of the uni-dimensional height and weight regions, reflecting our judgement as to Figure 4 shows how a 2-dimensional region can be constructed using a combination of an unconditional and a conditional standard, say height, and weight-for-height, to detect an overall 6% of cases. Such a standard might be particularly useful for detecting cases of malnutrition (the cross-hatched area)when the principal interest centres on weight for height, but it is desirable also to detect very short children. Indeed, this is often what is done in practice and the present method provides a formal framework for quantifying such a procedure. We also note that a 2-dimensional standard can be viewed as the successive application of unidimensional standards (of appropriate sizes). Thus a judgement based on the second measurement would be used only if the individual is 'typical' on the first measurement and this is clearly equivalent to our definition of a consistent 2-dimensional standard. Finally we may construct a combination of two conditional standards and this is illustrated in the examples below.

The extension of all these approaches to the multi-dimensional case is in principle straightforward. In practice, however, apart from the difficult computational problems this will raise, we suspect that their usefulness will be limited. In any case, it would seem sensible to gain experience with the use of two-dimensional standards before attempting to estimate 3 or 4 dimensional ones.



The extension of all these approaches to the multi-dimensional case is in principle straightforward. In practice, however, apart from the difficult computational problems this will raise, we suspect that their usefulness will be limited. In any case, it would seem sensible to gain experience with the use of two-dimensional standards before attempting to estimate 3 or 4 dimensional ones.

Figure 4 Two-dimensional height and weight-for-height percentiles




## 4 Examples

Figure 5 compares three 2-dimensional regions, an equi-probability ellipse, a rectangular region and a double conditional region using data on adult measurements of leg length and sitting height taken from Harrison and Marshall (1970), each measurement having the same weighting. The conditional percentile used here assumes a linear relationship between measurements and where this is not the case, an appropriate higher order curve should be fitted.

Figure 5 Equiprobability ellipse, leg length and sitting height and sitting height for leg length 5% regions; Adults (r=0.50)

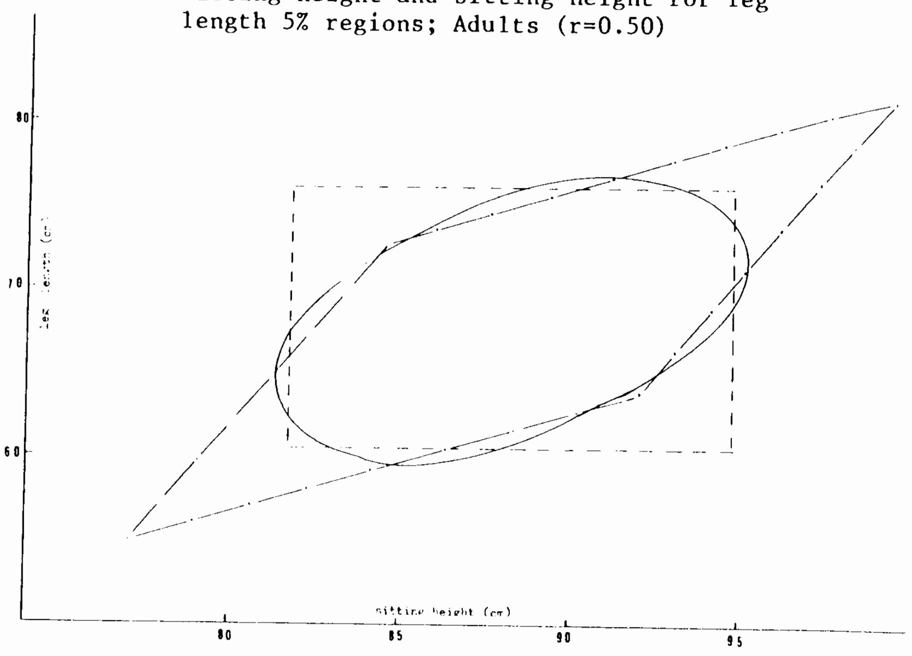
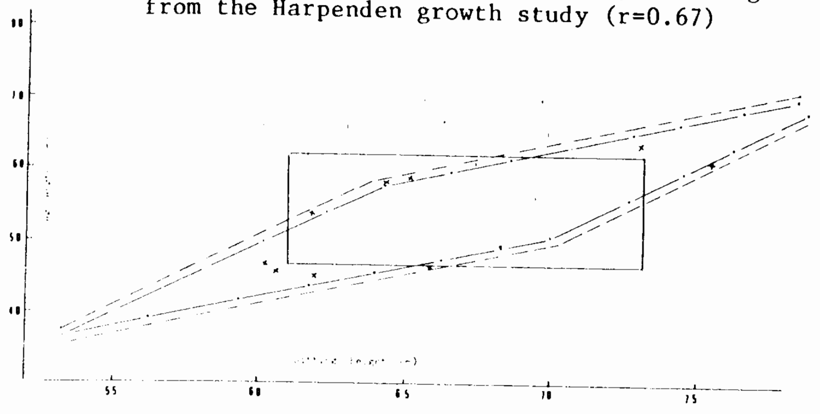



Figure 6 uses leg length and sitting height data (with equal proportions of atypical values) on 100 seven year old girls and gives a 6% rectangular region and a 6% and 2% double conditional regions with 9 atypical cases plotted. Since the double conditional standard defines individuals who are atypical either on leg length for sitting height or sitting height for leg length, we may regard such cases as having atypical shape. Furthermore, since the parallelograms defining different size regions have the same orientation, and one which defines a given atypical region therefore includes any other defining a larger atypical region, each individual will lie on one and only one parallelogram. Thus, each individual can be assigned a percentile value which is simply the size of the corresponding atypical region, and this gives a measure of shape which seems reasonable and useful. If we wish we can give equivalent standard normal scores to produce a standardised shape variable, and further define a direction to the measure with respect to the axes.



Figure 6 Leg length and sitting height 6% region, leg length for sitting height and sitting height for leg length 6% and 2% regions; 7 year-old girls from the Harpenden growth study (r=0.67)




## 5 Conclusion

I have defined a consistent multidimensional growth standard as one where, for a given size of the region judged to be atypical, an individual judged to be atypical would also be judged as atypical by at least one of the constituent unidimensional standards. Furthermore an individual judged to be typical according to the bivariate standard, is not necessarily typical by any of the univariate standards.

To construct a multidimensional standard we need both to specify the size of the atypical region and the relative proportions of values defined as typical by the constituent measurements.

For the 2-dimensional case I have discussed combinations of unconditional and conditional unidimensional standards and indicated potential uses. A frequent application is where a sequential judgement is made; typicality first being assessed on the basis of a single unidimensional standard, say for height. This will screen out a number of abnormally short (or tall) individuals. For the remainder, a conditional standard, say of weight for height, is used to screen out further abnormally light (or heavy) individuals, with a predefined percentage of atypical individuals being screened out overall. It should be noted, however, that the order of such judgements is immaterial. The important feature is the requirement that the overall size of the atypical region is maintained at a specific value.

As with unidimensional standards, 2-dimensional ones lend themselves readily to visual presentation, but this will not be true for higher-dimensional standards. In such cases, reliance will have to be placed in the use of appropriate computer software to construct atypical regions and to identify atypical individuals. Experience in the construction and use of 2-dimensional standards will provide pointers to the design of such systems and to requirements for the kinds of data most suitable for these purposes.

Finally, having decided upon the definition of an atypical region, it can also be viewed as a definition of an individual's shape based on the measurements chosen, and this idea seems worthy of further study.

Acknowledgements

I am indebted to Darrell Bock, Michael Healy, Roland Hauspie and Jim Tanner for their suggestions and comments.



REFERENCES

Cameron, N., 1980 Conditional Standards for growth in height of British Children from 5.0 to 15.99 years of age. Annals of Human Biology 7, 331-338

Defrise-Gussenhoven 1954 Croissance et Debilite Brussels, Memoire no 128, Institut Royal des Sciences Naturelles de Belgique.

Goldstein, H., 1984 Current developments in the design and analysis of Growth studies. In: Borns et al (eds) Human growth and development, Plenum, New York

Goldstein, H., and Tanner, J. M., 1980 Ecological considerations in the creation and use of child growth standards, Lancet, i, 582-5

Harrison, J. M., and Marshall, W. A., 1970 Normal Standards for the relationships between the lengths of Limbs and of limb segments in young British Women: A Photogrammetric Study. Human Biology, 47, 90-104

Hellier, J. L., and Goldstein, H., 1979, The use of birthweight and gestation to assess perinatal mortality risk, Journal of Epidemiology and Community Health, 33 (3), 183-185

Healy, M. J. R., 1969 Rao's Paradox concerning Multivariate tests of significance. Biometrics 25, 411-413

Tanner, J. M., 1952 The assessment of growth and development in children, Archives of Diseases in Childhood, 27, 10-33

Tanner, J. M., 1981 A history of the study of human growth, Cambridge University Press, Cambridge

## APPENDIX

It is convenient to assume a bivariate normal distribution for the two measurements.

Let $a$ be the chosen proportion of atypical values, then

$$1 - a = \int_{h_1}^{h_2} \int_{W_1}^{W_2} \phi\ (x,y)\ dx\ dy$$

Let $b$ be the ratio of the proportions found to be atypical for the corresponding unidimensional standards. Then,

$$b = \frac{1 - \int_{h_1}^{h_2} \phi\ (x)}{1 - \int_{W_1}^{W_2} \phi\ (y)}$$

For given values of $a$ and the correlation between $x$ and $y$ the value of $b$ will determine the values of $h_1$, $h_2$, $w_1$, $w_2$. In practice of course a bivariate normal[1] distribution may not be an adequate approximation and we will want to use an empirical approximation to the bivariate frequency surface. In the unidimensional case (Goldstein 1984), we can do this by smoothing the cumulative frequency plot of the measurements and we can extend this to the bivariate case by choosing suitable 2-dimensional smoothing functions.