

Survey Solutions: Anthropometric Z-Score Functions

Sergiy Radyakin

March 25, 2017

1 Purpose

Anthropometric measurements like height or weight play an important role in socio-economic studies, as they are commonly used as a measure of child development or malnutrition. It is thus very important to obtain their accurate measurements during the data collection process. The functions discussed here help validate the anthropometric measurements to prevent gross errors.

Anthropometric Z-score functions serve as a way to measure the distance between an individual person's indicator and the corresponding average in the reference population. For example, when collecting height measurements of children, one may be interested to make sure the values are not unplausible (not too high, and not too low). Because the children grow with age, simple bounds, which can be implemented with

```
height.InRange(L,H)
```

function with two constant height brackets L and H , are not a very good solution. We would like to have L and H adequate for the corresponding group of children. Hence the Z-scores take into account the age and also whether the child is a boy or a girl:

```
ZScore.Lhfa(ageMonths,isBoy,height).InRange(-3,3)
```

Here, one relies on the height-for-age Z-score function to return a standardized distance between the actual height of the child (*height*), and the average for the group of children of the specified gender (*isBoy*) and age (*ageMonths*). Having obtained the standardized difference, it is possible to specify a constant interval (in this example [-3;3]), which determines how much of a deviation from the mean one wants to allow, before an action is taken by the CAPI system: e.g. an answer is considered invalid.

The motivation for this is the so called *68-95-99.7 rule*, which states that for a normal distribution the values less than one standard deviation away from the mean account for 68.27% of the set; while two standard deviations from the mean account for 95.45%; and three standard deviations account for 99.73%.

In practice this means that this is a very rare occasion when we observe valid measurements outside of [-3;3] range (if we believe the corresponding measure is normally distributed).

2 Source of information

Anthropometric Z-score functions use the reference tables published by the WHO. These tables are compiled based on extensive research, and are made available to the public via the WHO website:

<http://www.who.int/childgrowth/en/>

These tables are widely used in the applied research, and are implemented in the corresponding form in popular statistical packages: SPSS, Stata, and others.

Where we felt necessary the age series were combined together to have a longer series (this means that when calling the function, one usually doesn't want to specifically identify whether the height of the child should be measured standing or laying, while different methods were used for different age groups).

3 Implemented anthropometric Z-score functions

Function	Indicator	Age range
ZScore.Acfa	arm circumference-for-age	3 months to 5 years.
ZScore.Bmifa	BMI-for-age	birth to 5 years.
ZScore.Hcfa	head circumference-for-age	birth to 5 years.
ZScore.Lhfa	length/height-for-age	birth to 5 years.
ZScore.Ssfa	subscapular skinfold-for-age	3 months to 5 years.
ZScore.Tcfa	triceps skinfold-for-age	3 months to 5 years.
ZScore.Wfa	weight-for-age	birth to 5 years.
ZScore.Wfl	weight-for-length	birth to 2 years.
ZScore.Wfh	weight-for-height	2 to 5 years.

Note also the sequence of the arguments when calling the Z-score functions:

```
ZScore.Acfa(ageMonths, isBoy, measurement)
ZScore.Bmifa(ageMonths, isBoy, measurement)
ZScore.Bmifa(ageMonths, isBoy, weight, height)
ZScore.Hcfa(ageMonths, isBoy, measurement)
ZScore.Lhfa(ageMonths, isBoy, measurement)
ZScore.Ssfa(ageMonths, isBoy, measurement)
ZScore.Tsfa(ageMonths, isBoy, measurement)
ZScore.Wfa(ageMonths, isBoy, measurement)
```

4 How to use Z-score functions

Since all of the S-score functions rely on the age and gender of the child, these information should be collected prior to the collection of the height, weight, or other anthropometric measurement. This simply means that the age and gender questions should be placed early in the interview. They don't have to immediately precede the anthropometric measurements, and can even be in a different section of the questionnaire.

Here is a typical validation expression for the *height* question:

```
ZScore.Lhfa(age,gender==1,height).InRange(-3,3)
```

Here *height* is the name of the height variable, *gender* is a categorical variable with categories 1=Boy; 2=Girl; and *age* is the variable corresponding to the question of the child's age in months.

Note that the *age* variable is always specified in months in every function here that requires an age parameter.

Note also that the gender specification is of boolean type, and if one has it as categorical, it should be converted to a boolean type, as shown. The logical value *true* should correspond to the boys, and the logical value *false* should correspond to the girls.

Be alert for the units of measurement. All units are metric (cm, kg) not imperial (inches, pounds). If the questionnaire is collecting the information using non-metric units, they need to be converted to metric units prior to the call to the Z-score functions. For example, for imperial units the following conversion factors can be used:

1 inch = 2.54 centimeters,

1 pound = 0.453592 kilogram.

So if the child's height was measured in inches, the following expression should be used to validate it:

```
ZScore.Lhfa(age,gender==1,height*2.54).InRange(-3,3)
```

And if the child's height was measured in feet and inches, the following expression should be used to validate it:

```
ZScore.Lhfa(age,gender==1,(feet*12+inches)*2.54).InRange(-3,3)
```

For every measurement the designer of the questionnaire should pick a tolerance level, which determines the threshold when e.g. the error message is displayed. In this example it is picked as 3 standard deviations from the mean. One can pick a different value, especially if there is expectation that the population under study may be substantially different from the reference population analyzed by WHO to construct the reference tables.

If the measurement yields a Z-score outside of the specified permitted range there is a good chance that it is a wrong measurement, and the interviewer needs to repeat the measurement procedure. This could also be a signal that the other variables (age or gender) were registered incorrectly. Finally, this could also mean the event of observing the child with extreme measurement. Note that the validation in Survey Solutions is not blocking: even if the validation check has produced an error message, the interviewer can still submit the data for supervisor approval.

Note that if any of the arguments is missing an exception is thrown by the Z-score function. An exception is also thrown if the age is out of the defined range, or if the measurement is less than or equal to zero.

One can use a check with a Z-score function to trigger other questions, rather than showing an error message. For example, a picture question can be triggered if the measurement is too large or too small, or a repetitive measurement using a different instrument, or just a text question to the interviewer to provide comments.

In addition to Z-scores the *Bmi(weight,height)* function allows to compute the Body-Mass-Index based on height in meters and weight in kilograms. The *Bmi()* function is not using reference tables, but instead the following formula:

$$BMI = \frac{W}{H^2}$$

where *W* is weight in kilograms and *H* is height in meters.

The *Bmifa* function can take as an argument the body-mass-index if directly asked in the interview, or as a more common case, the result of the *Bmi()* function from weight and height, or directly the weight and height:

```
ZScore.Bmifa(age,gender==1,b).InRange(-3,3)
ZScore.Bmifa(age,gender==1,Bmi(weight,height)).InRange(-3,3)
ZScore.Bmifa(age,gender==1,weight,height).InRange(-3,3)
```

5 Practice Exercises

1. Can a result of a Z-score function be negative? If yes, what will be the interpretation of it?
2. How should the following validation condition be revised if the interviewers are getting too many error notifications?

```
ZScore.Wfa(age,gender==1,weight).InRange(-1,1)
```

How do you adjust the same if you have a prior knowledge that the population under study is somewhat overweight?

3. Create a questionnaire, which collects demographic information about the family and anthropometric information for children under 5 y.o. For children with signs of obesity collect dietary information.
4. Create a questionnaire which collects height and weight of small children in cm and kg, and asks for the gender of the child, but not age. Verify the weight corresponds to the height. Explain how the verification works in the absense of the age information.
5. Create a questionnaire that asks for two measurements of the same indicator (say, weight). Show an error message if these measurements differ by more than 5% or by more than 1 S.D.
6. The questionnaire collects measurements of weight of children in *stones*, *pounds*, and *ounces*. If the child's weight is extreme, the interviewer should take a picture of the child, and take a picture of the instrument's scale when the child is being weighted. Use reference sources if you are not familiar with the non-standard units of weight.