IMPACT OF SKEWING IN BODY WEIGHT UPON BODY MASS INDEX STATISTICS IN A MALE MODEL POPULATION

Boris Neruda

Private Institute for Studies in Science and Medicine, Germany

ABSTRACT

There is insufficient research concerning whether an increase in skewness of the variable body weight W would influence body mass index (BMI) statistics, and hence falsify conclusions. In this study, a model male population was generated. Dividing W by a normally distributed variable produced its skewing to the right. Apart from expected changes in the means, standard deviations, and other statistics of W and BMI, the results revealed that Benn's index and the correlation r_{HBMI} remained unchanged over a large range of distribution asymmetry. Counts of elements that would fall into the overweight category, decreased in favour of the obesity class – an intuitively unexpected result. A verifiable rise in skewness resulted in loss of the normal distribution assumption for W and BMI.

Keywords: *body mass index, body weight, model population, skewing, statistics*

INTRODUCTION

A recent publication has shown that only a single set of data will, as a rule, cope with the prerequisites for a proper usage of the body mass index (BMI) [1]. These requirements are: a high correlation between body weight W and BMI (r_{WBMI}); absence of a correlation between body height H and BMI (r_{HBMI}); and a Benn's index of 2. The latter is obtained as the factor of the logarithm of H in the linear regression between the logarithms of W and H. An increased skewness of variable W, which unlike H is per se always skewed to the right in real populations, would possibly lead to unpredictable results in an experiment comparing two samples with different shapes for the body weight distribution, and hence their corresponding BMIs. Such effects are yet to be considered.

METHODS

A computer program was used to generate normally distributed, but correlated variables, W and H [1]. The input for W was 76.6 ± 12.5, and for H, 172.0 ± 6.8. These figures represent means based on 210,531 male subjects from different populations [2]. Correlation r_{WH} was preset to 0.470. The sample population covered 500,000 elements. This original database was randomly separated into 10 equal groups grp_i (i = 0 to 9), where grp0 served as the base in group comparisons.

Skewing of variable W was achieved by dividing by a normally distributed random variable $V_j = N(100,j)/100$, with j = 0 to 18 (step 2), creating W_i, which then was divided by H_i^2 to yield the corresponding variable BMI_i.

In the individual BMI_i groups, the number of elements and their percentage were calculated for $25 \leq BMI_i \leq 29$, corresponding to overweight, for $30 \leq BMI_i \leq 39$, representing obesity, and for $BMI_i \geq 40$, defined as morbid obesity. Benn's index p was obtained from the regression equation $\log(W_i) = a + p \cdot \log(H_i)$.

All relevant computations were performed using the SPSS and Excel software. Testing for normal distribution was done by the Kolmogorov-Smirnov test (K-S). ANOVA together with post-hoc Tukey-HSD test were used for comparisons between groups.

RESULTS

The computer algorithm produced normally distributed variables W and H as requested, with linear correlation $r_{WH} = 0.483$. Details of the statistics for W_i and BMI_i (i = 0 to 9) are listed in Table 1.

A comparison of groups revealed that groups 0 to 3 formed an individual subgroup, and only groups 4 to 9 displayed a statistical difference from group 0 for both W and BMI. In these groups, excess and kurtosis increased as well. Groups 0 to 2 of W and BMI showed low scores in the K-S test, and figures for asymptotic significance were high, indicating that these groups had normal distribution; all other groups violated this assumption.

The correlation r_{WH} decreased in those groups, where a larger variance of the denominator V_j was used; but r_{HBMI} remained well within non-significantly different zero values. Figures for r_{WBMI} were larger in group categories with higher numbers (i > 3).

Table 1. Simulated body weight W statistics before (Group 0) and after increasing steps of skewing. and resulting BMI statistics

grp	0	1	2	3	4	5	6	7	8	9
W	0	2	4	6	8	10	12	14	16	18
mean	76.6	76.7	76.6	76.8	77.1	77.4	77.8	78.2	78.8	79.2
sd	12.6	12.6	12.9	13.4	14.1	15.0	16.1	17.3	18.9	20.8
CV	16.4	16.4	16.8	17.4	18.3	19.4	20.7	22.2	24.0	26.2
excess	0.002	0.018	0.069	0.115	0.210	0.330	0.493	0.657	0.832	1.060
kurtosis	0.000	0.040	0.068	0.082	0.082	0.288	0.718	1.251	1.749	2.769
K-S	0.457	0.557	1.190	1.778	3.239	4.816	7.161	8.384	10.849	13.123
signif	0.985	0.916	0.118	0.004	0.000	0.000	0.000	0.000	0.000	0.000
Benn	2.058	2.026	2.051	2.074	2.018	2.088	2.060	2.025	2.045	2.060
rHWi	0.483	0.475	0.468	0.459	0.433	0.418	0.380	0.343	0.322	0.303
Tukey-HSD		0.986	1.000	0.624	0.000	0.000	0.000	0.000	0.000	0.000
BMI										
mean	25.8	25.9	25.9	25.9	26.0	26.1	26.3	26.4	26.6	26.7
sd	3.7	3.8	3.9	4.0	4.3	4.6	5.0	5.5	6.1	6.7
CV	14.5	14.5	15.0	15.6	16.6	17.7	19.2	20.9	22.8	25.0
excess	−0.016	0.020	0.080	0.122	0.251	0.364	0.531	0.702	0.883	1.101
kurtosis	0.061	0.086	0.145	0.128	0.229	0.363	0.824	1.316	1.842	2.821
K-S	0.911	0.807	1.635	2.096	3.695	6.012	7.188	9.329	11.623	13.792
signif	0.378	0.532	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000
rHBMIi	0.003	−0.005	0.000	0.007	−0.002	0.009	0.000	−0.010	−0.004	0.004
signif	0.510	0.267	0.999	0.111	0.735	0.036	0.929	0.021	0.383	0.376
rWBMIi	0.875	0.875	0.881	0.889	0.898	0.909	0.922	0.933	0.942	0.951
Tukey-HSD		0.905	1.000	0.292	0.000	0.000	0.000	0.000	0.000	0.000

CV: coefficient of variation

Counting the number of subjects falling into BMI subsets defined as "adipose," "obese," and "morbidly obese," showed that with an increasing group number, the amount of "obese" and "morbidly obese" grew at the expense of the "overweight" ones (Table 2).

Table 2. Number and percentage of elements in different BMI classes and groups
(n= 50.000 in each individual group)

grp	0	1	2	3	4	5	6	7	8	9
BMI										
25–29	24102	23814	23309	22418	21109	19994	18674	17694	16377	15180
%	48.2	47.6	46.6	44.8	42.2	40.0	37.3	35.4	32.8	30.4
30–39	8100	8371	8450	9220	10069	10775	11437	11691	12031	12111
%	16.2	16.7	16.9	18.4	20.1	21.6	22.9	23.4	24.1	24.2
>=40	6	14	27	53	132	285	597	1043	1549	2201
%	0.0	0.0	0.1	0.1	0.3	0.6	1.2	2.1	3.1	4.4
Σ%	64.4	64.4	63.6	63.4	62.6	62.1	61.4	60.9	59.9	59.0

DISCUSSION

Currently, the body mass index is often used as a measure of overweight and obesity. It is defined as the ratio of W with H^2. Ratios of two, even normally distributed variables, however, have the potential to produce skewing of the new ratio variable. Body weight by itself is already skewed to the right in real populations, and the asymmetry is more pronounced in female than in male populations.

This study was designed to show whether additional skewing in W would have an effect upon the resulting BMI. In order to investigate such an effect, computer-generated normally distributed variables W and H were created, which modelled a male population. Restriction to a male population was deliberately chosen for several reasons: first, because it has been shown that a male model population better represents real life data than a female model population [1]; and second, a male model population was deemed sufficient to disclose some initial effects already. A preset correlation between W and H was chosen, because otherwise detecting any changes in Benn's index p due to additional skewing in W would be impossible. Apart from that, a test not shown here from an uncorrelated data-set produced identical results.

From the many options for generating a skewed variable, the ratio of two variables was chosen. The normally distributed variable W was divided by a normally distributed variable V, which had increasing variance values. This method of producing asymmetry of a variable certainly does not reflect the way in which body weight is skewed in real populations. But it should be sufficient as a first approximation to disclose the main effects of the problem under study. However, it must again be stressed that this approach by no means mirrors asymmetry changes in real body weight distributions.

If any skewing to the right had occurred, it would be intuitively expected that the means, standard deviations, and coefficients of variation of W, and likewise of BMI, would also increase. A larger variation coefficient of at least 8% in the confounding variable V (Group 4) was still needed in this setting, however, to generate a statistically significant effect.

The linear correlation between W and BMI increased in higher group numbers. Obviously, as the result of a larger variance in W, and hence in BMI more elements relate to one another. In contrast, r_{HW} decreased significantly at the same time (for $i \geq 4$), probably due to an increase in skewing. This effect is well known for female populations, where r_{HW} is generally lower, and the excess in body weight is larger than in men. Somewhat surprisingly, Benn's index p did not change at all, which means that this index is quite robust to changes in the distribution of body weight.

The result of the Kolmogorov-Smirnov test needs further corroboration. At this stage it would mean that the changes in W and BMI statistics, and the significant difference between groups, can be explained by a simultaneous increase in variance and skewness.

Another noteworthy finding is the number and percentage of the elements that would be rated as overweight and obese. One would expect a different pattern of skewing in body weight; namely, that the number of overweight would increase more than the obese. Yet, the contrary is the case. This behaviour can probably be attributed to the undue use of a ratio as a skewing method.

Further studies using other methods of generating skewness in the variable W are needed to fully understand whether an additional asymmetry would quantitatively falsify results in BMI studies.

REFERENCES

1. Neruda B., Hilger Y. (2010). Impact of the correlation between normally distributed samples of stature and body weight upon body mass index statistics. Papers on Anthropology XIX, 271–278.
2. N. N. (2005). Weight-height relationships and body mass index: some observations from the Diverse Populations Collaboration. Am J Phys Anthropol 128, 1, 220–229.

Address for correspondence:
Boris Neruda
Private Institute for Studies in Science and Medicine,
Eschbergstrasse 28
54585 Esch