

TECHNICAL REPORT

Generalized lambda distribution for flexibly testing differences beyond the mean in the distribution of a dependent variable such as body mass index

K Ejima^{1,2,3}, G Pavea^{3,4}, P Li⁵ and DB Allison^{1,3,5,6}

BACKGROUND/OBJECTIVES: Conventional statistical methods often test for group differences in a single parameter of a distribution, usually the conditional mean (for example, differences in mean body mass index (BMI; kg m^{-2}) by education category) under specific distributional assumptions. However, parameters other than the mean may of be interest, and the distributional assumptions of conventional statistical methods may be violated in some situations.

SUBJECTS/METHODS: We describe an application of the generalized lambda distribution (GLD), a flexible distribution that can be used to model continuous outcomes, and simultaneously describe a likelihood ratio test for differences in multiple distribution parameters, including measures of central tendency, dispersion, asymmetry and steepness. We demonstrate the value of our approach by testing for differences in multiple parameters of the BMI distribution by education category using the Health and Retirement Study data set.

RESULTS: Our proposed method indicated that at least one parameter of the BMI distribution differed by education category in both the complete data set ($N = 13\,571$) ($P < 0.001$) and a randomly resampled data set ($N = 300$ from each category) to assess the method under circumstances of lesser power ($P = 0.044$). Similar method using normal distribution alternative to GLD indicated the significant difference among the complete data set ($P < 0.001$) but not in the smaller randomly resampled data set ($P = 0.968$). Moreover, the proposed method allowed us to specify which parameters of the BMI distribution significantly differed by education category for both the complete and the random subsample, respectively.

CONCLUSIONS: Our method provides a flexible statistical approach to compare the entire distribution of variables of interest, which can be a supplement to conventional approaches that frequently require unmet assumptions and focus only on a single parameter of distribution.

International Journal of Obesity advance online publication, 28 November 2017; doi:10.1038/ijo.2017.262

INTRODUCTION

The aim of many statistical tests in life science research is to identify differences in outcomes (that is, dependent variables (DVs)) as a function of one or more independent variables, with the t-test and analysis of variance being prototypical examples. These tests, and many others, test for differences in only one parameter of the conditional distribution of the DV: the mean. Yet the distribution of DVs may be defined by more than the mean; distributions can also differ in their median, range, s.d. and other aspects. An incomplete examination of the parameters in which the distribution of a DV, such as body mass index (BMI), might differ between populations may result in an incomplete characterization of the DV, burden of obesity, for example, and missed opportunities for understanding and intervention.

Our proposed method provides a more comprehensive test of group differences in the distribution of continuous outcomes such as BMI, including measures of location and dispersion. In the following sections, we (1) provide further motivation for the value of statistical tests that examine group differences beyond the

conditional mean; (2) describe our method, based on the generalized lambda distribution (GLD), and (3) apply the proposed method in an analysis of the association between BMI and education using data from the Health and Retirement Study (HRS).

More than the mean

Good empirical and theoretical reasons exist to examine group differences in the distributions of DVs beyond the conditional mean, and BMI is among those variables. First, the distribution of BMI is increasingly skewed, so the measures of central tendency, such as mean, may change little when the distribution changes at the tails.¹ Komlos and Brabec² examined trends in BMI deciles between 1882 and 1986 among US adults and found that the 90th percentile of BMI increased between 18 and 22 units while the lowest decile increased by only 1–3 units.

We are not the first to note this problem and methods that are not limited to the comparison of conditional means, in particular quantile regression.^{3,4} The ability to test for differences other than the conditional mean may be particularly valuable when the effect

¹Office of Energetics, School of Health Professions, University of Alabama at Birmingham, Birmingham, AL, USA; ²Institute of Industrial Science, The University of Tokyo, Tokyo, Japan; ³Nutrition Obesity Research Center, University of Alabama at Birmingham, Birmingham, AL, USA; ⁴Department of Health Behavior, University of Alabama at Birmingham, Birmingham, AL, USA; ⁵Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL, USA and ⁶Department of Nutrition Sciences, University of Alabama at Birmingham, Birmingham, AL, USA. Correspondence: Dr DB Allison, Department of Nutrition Sciences, University of Alabama at Birmingham, 1705 University Boulevard, SHPB 677, Birmingham, AL 35294, USA.

E-mail: dallison@uab.edu

Received 15 May 2017; revised 14 September 2017; accepted 27 September 2017; accepted article preview online 30 October 2017

of a treatment or variable is expected to vary by level of BMI. For example, researchers interested in the effect of breastfeeding on childhood BMI may hypothesize that breastfeeding is associated with increased BMI among children who would otherwise be underweight but with decreased BMI among children who would otherwise be overweight, thereby reducing variance in BMI while having little effect on measures of central tendency. This effect pattern, while theoretically plausible and having important public health implications, might otherwise be missed by statistical methods that focus only on the mean. Indeed, Beyerlein *et al.*⁵ found that breastfeeding was associated with a reduced BMI among children at higher BMI percentiles and an increased BMI among children at lower BMI percentiles, whereas breastfeeding had no association with BMI in children with BMI percentiles of 0.4 and 0.8 in the linear regression model.

Education and BMI: an example

We now turn our attention to the BMI–education association and briefly explain its importance in the literature. The inverse association between education and BMI is well documented in multiple races/ethnicities in both genders, especially among white women in high Gross Domestic Product countries.^{6–11} However, many of these studies used *t*-tests, linear regression or other methods that tested for an association between education and mean BMI. Recent research suggested that the focus only on mean differences in BMI by educational category might overlook important nuances in the relationship. For example, Joliffe¹² found that income was associated with the dispersion of BMI but not the mean BMI. In this case, a focus on mean differences would have missed these findings.

Expanding the reach of one's analysis across the entire distribution and for multiple parameters of interest encourages one to think more critically about the distribution that best fits the data and the assumptions under conventional approaches. Methods limited to an analysis of the conditional mean may overlook these and other reasons why the distribution of BMI differs by educational group and how it differs. In the following section, we apply the proposed method to test the difference in the distribution of BMI by education levels.

MATERIALS AND METHODS

Data

Demographic and anthropometric data used in the illustrative example come from the HRS and the Research and Development Corporation HRS data set, a cleaned version of the raw data funded by the National Institute on Aging and the Social Security Administration.¹³ In the present study, we used the data of participants who completed the enhanced face-to-face interview in 2006 or 2008 with complete measurements on height and weight. We carried out two analyses: one using the complete HRS sample meeting the above criteria (henceforth 'complete sample', $N=13\,571$) and the other using a randomly resampled data set ($N=300$ each from the individuals who graduated and not graduated high school, for a total N of 600) to examine outcomes in a smaller data set offering less statistical power.

Generalized lambda distribution

Normal and gamma distributions are widely used to fit a probabilistic distribution to an empirical distribution. However, probability distributions are often unable to fully characterize an empirical distribution because of restrictions in their shape. The GLD is defined by four interpretable parameters: $\tilde{\mu}$, $\tilde{\sigma}$, χ and ξ , indicating median, interquartile range, asymmetry and steepness, respectively, so it can describe a broad types of shapes, including the normal distribution as a special case.¹⁴ Given its flexibility, the GLD has been widely used in finance, meteorology and other fields in which data distributes frequently with a heavy tail.^{14,15}

It has also been suggested instead of a normal distribution in multiple imputation.¹⁶

Methods to test for differences in outcome distribution

In the following example, X indicates educational attainment, with 0 indicating less than a high school degree and 1 indicating at least a high school degree. Y denotes BMI. We first assume that the distribution of BMI follows a GLD for both categories of education but allow each distribution to vary in its parameterization, defined as follows:

$$\begin{aligned} Y_{X=0} &\sim \text{GLD}(\tilde{\mu}_0, \tilde{\sigma}_0, \chi_0, \xi_0), \\ Y_{X=1} &\sim \text{GLD}(\tilde{\mu}_1, \tilde{\sigma}_1, \chi_1, \xi_1). \end{aligned}$$

Having allowed the parameterization of BMI to vary by educational group, we can then use likelihood ratio test (LRT) to test whether the parameterizations of BMI differed between groups.¹⁷ The LRT is appropriate when testing nested models, as was the case for this example. The likelihood under the null hypothesis, L_0 , is the likelihood given that the parameterization of the GLD for BMI is the same across educational groups:

$$\tilde{\mu}_0 = \tilde{\mu}_1, \tilde{\sigma}_0 = \tilde{\sigma}_1, \chi_0 = \chi_1, \xi_0 = \xi_1.$$

Under the null hypothesis, the data are assumed to be sampled from one GLD, thus the number of parameters is four. The likelihood under the alternative hypothesis, L_1 , is the likelihood given that at least one of the parameters differs between groups, and the data from each group are assumed to be sampled from two different GLD parametrizations, requiring eight parameters in total (four parameters per GLD). The test statistic, $-2\log(L_0/L_1)$, asymptotically follows the χ^2 distribution with degrees of freedom equal to the difference in the number of parameters, which is four. To perform the statistical analysis, we used the statistical computing software R 3.3.1 and its package 'gldist'. The package allowed us to fit the GLD to empirical data and estimate the parameters of the GLD by maximizing the likelihood. We confirmed that the method of moments¹⁸ provided highly similar estimates. Maximum likelihood estimation was employed because of its asymptotic properties and the LRT used in our method.

Methods to infer association using linear regression with GLD

Data are often multivariate in obesity research and innumerable studies to infer associations between a treatment and BMI have been conducted. In many of these studies, multiple linear regression is used, which allows researchers to statistically adjust for potential confounders and other covariates. In general, ordinary linear regression assumes that residuals follow a normal distribution with a mean of zero. For the cases in which residuals do not follow a normal distribution, Su¹⁹ proposed a linear regression model in which residuals follow a GLD with a mean of zero. To demonstrate this method, we used BMI as an outcome and education level as an independent variable of interest, adjusting for sex and age. The regression analysis was performed only for the resampled data set using R package 'GLDreg'. All statistical code is available in Supplementary Appendix.

RESULTS

Figure 1 shows the histogram of the BMI distributions among individuals who did not graduate high school overlaid with the distribution of BMI among those who graduated high school. In both data sets, the distribution of BMI was skewed to the right for both education groups, and the GLD curves captured this aspect (Table 1 presents estimated parameters including a parameter of steepness, χ is >0 for all groups, which corresponds to a right skewed distribution). Table 2 presents the descriptive statistics of the study sample. The difference in mean BMI by education group was statistically significant in the complete sample ($P=0.001$ from *t*-test and $P<0.001$ from Wilcoxon rank-sum test) but not in the smaller resampled subsample ($P=0.866/0.645$).

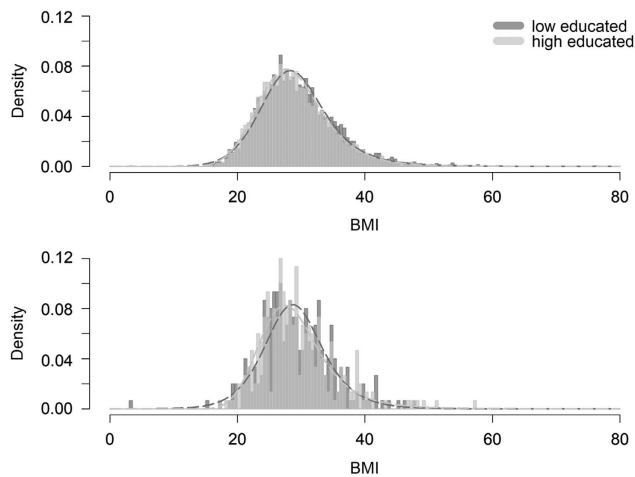


Figure 1. Distribution of BMI by education. The dashed lines show the estimated distributions. Upper panel: complete data set; lower panel: a resampled subsample.

Table 1. Testing the differences in each parameter of the BMI distribution between the groups

Parameter	Did not graduate high school ^a	Graduated high school ^a	P-value	Adjusted P-value ^b
<i>Original</i>				
Median	28.89 (0.11)	28.43 (0.06)	< 0.001*	< 0.001*
Interquartile range	7.25 (0.13)	7.15 (0.06)	0.466	1.000
Asymmetry	0.21 (0.02)	0.26 (0.01)	0.036*	0.146
Steepness	0.5 (0.01)	0.45 (0.01)	0.001*	0.002*
<i>Resampled</i>				
Median	28.88 (0.31)	28.43 (0.31)	0.316	1.000
Interquartile range	6.63 (0.36)	6.81 (0.4)	0.739	1.000
Asymmetry	0.09 (0.07)	0.34 (0.05)	0.006*	0.023*
Steepness	0.52 (0.05)	0.46 (0.06)	0.445	1.000

* $P < 0.05$. ^aS.e. in parentheses. ^bAdjusted by Bonferroni correction (multiplied P-values by 4).

Our proposed method suggested that at least one of the parameters of the GLD for BMI was significantly different by education category ($P < 0.001$ for the complete data set and $P = 0.025$ for the resampled data set). We performed the same analysis using a normal distribution, finding a significant difference in mean BMI by education group in the complete data set ($P < 0.001$) but not in the smaller resampled subsample ($P = 0.612$).

On the basis of these results, researchers could conduct additional *post hoc* tests to identify the specific parameters that differed between groups by modifying the alternative hypothesis.

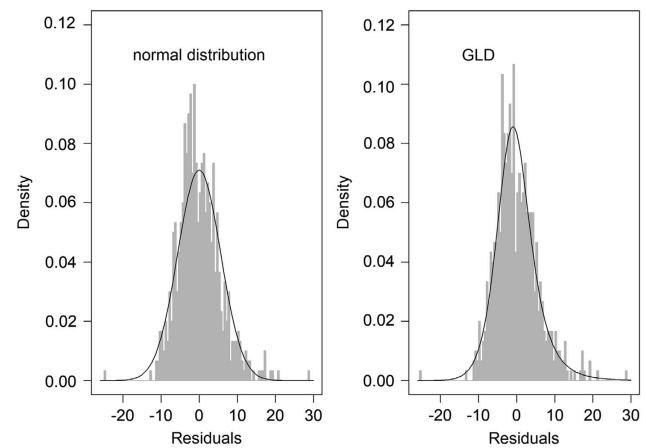


Figure 2. Distribution of residuals. Right panel: normal distribution; left panel: GLD. The lines show the fitted distributions.

For example, one could test whether the median BMI, $\tilde{\mu}$, differs by education category. The corresponding null hypothesis is $\tilde{\mu}_0 = \tilde{\mu}_1$, and the alternative hypothesis would thus be specified with the following constraints: $\tilde{\mu}_0 \neq \tilde{\mu}_1$. Table 1 summarizes the results (significance levels adjusted by Bonferroni correction). In the complete sample, we found evidence of a significant difference in the median ($P < 0.001$) and steepness ($P = 0.002$), which implies that median of BMI distribution of the low education category (did not graduate high school) is higher and the distribution has sharper inclination compared with the high education category (graduated high school). Interpretation might be easier for the resampled data set. A significant difference in asymmetry was suggested ($P = 0.023$), which implies the BMI distribution of the low education category is more right skewed (having a long right tail, which is visually observed in the lower panel in Figure 1). The estimated coefficients from linear regressions using normal distribution and GLD were similar in direction. However, as the distribution of residuals is right skewed (Figure 2), GLD fits better than normal distribution (Akaike Information Criteria (AIC)²⁰: 3736 vs 3780 (a model with lower AIC fits the data better)).

DISCUSSION

Many methods commonly used to test for group differences in a variable of interest test for differences in only one aspect of the distribution, often the mean. However, distributions can vary on many other parameters, and a narrow focus on the mean or median can lead researchers to overlook important distinctions. The method we have proposed allows testing of whether at least one parameter differs between distributions and, if so, specifying which parameters are different. We have demonstrated the potential value of this method in an analysis of the association between BMI and education. Our results indicated a significant difference in at least one parameter of the BMI distributions between education groups in both the complete sample and the resampled data. Substantively, the results from our analysis

Table 2. Baseline statistics of the population used in the analysis, by education group

	BMI, kg m ⁻² (mean \pm s.d.)		P-value	
	Did not graduate high school (N = 2679/300)	Graduated high school (N = 10 892/300)	t-Test	Wilcoxon
Original	29.49 \pm 6.10	29.07 \pm 5.86	0.001*	< 0.001*
Resampled	29.17 \pm 5.54	29.25 \pm 5.86	0.866	0.645

* $P < 0.05$.

suggest a more nuanced relationship between education and weight than is typically assumed in the literature. In the complete sample, in addition to median, the 'steepness' of the BMI distribution among individuals without a high school degree was significantly higher than the steepness of the BMI distribution among individuals with a high school degree, indicating that having less than a high school education is associated with more extreme values in BMI at both tails of the distribution. We performed the same test but using a normal distribution. Both methods (methods using a normal distribution and a GLD) returned significant results for complete data set. However, for the resampled data set, a significant difference in the distributions was detected from the method using a GLD but not from the method using a normal distribution, which suggests that GLD testing under some circumstances has higher power to detect the significant difference between distributions.

Researchers have been aware for some time of the limitations of using only the mean to examine the relationship between variables, and quantile regression⁴ has increased in popularity because it allows researchers to evaluate associations between the conditional median or other quantiles of a response variable and exposures. Such evaluation could potentially detect differences in distributions of variables even when there is a weak to null relationship between their conditional means. An important limitation of quantile regression, however, and one addressed by our proposed method, is that it is unable to directly test differences in the parameters of the distribution other than the selected quantiles and can exacerbate multiple testing concerns if many quantiles are examined. By contrast, GLD allows researchers to test for differences in four prespecified parameters. An LRT can test for differences in the distribution as a whole or in a specific parameter. Recently, researchers demonstrated the feasibility and usefulness of fitting a nonconventional flexible distribution, a transformed beta distribution, to an empirical BMI distribution by using statistics commonly available in surveillance data.²¹ However, the parameters of the model are difficult to interpret. In contrast, the parameters of the GLD are easier to interpret and can be confirmed visually from a figure, another rationale for having employed this distribution in our proposed method and illustrative example.

Owing to the frequency of missing data in obesity-related research, researchers must still take care to handle missing data appropriately and that the valid use of methods to account for missingness (for example, multiple imputation) are still contingent upon requisite assumptions being met—the use of a GLD for the analysis of data with missingness at the analysis stage does not change anything in this regard.

In conclusion, we illustrated a new statistical method that can be used to make more comprehensive comparisons of distributions of continuous variables, such as BMI. Use of this model allows rich descriptions of the effects of (and associations with) treatments and other variables with obesity.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Dr Jasmin Divers for reviewing our initial manuscript and providing constructive comments. We also thank Jennifer Holmes for editing our manuscript.

This study was supported in part by NIH grants P30DK056336 and UL1TR001417, Japan Society for Promotion of Science (JSPS) grant KAKENHI 15J00009 and Grant-in-Aid for Epidemiological Research from St Luke's International University. The opinions expressed are those of the authors and do not necessarily represent those of the NIH or any other organization.

AUTHOR CONTRIBUTIONS

KE, GP and PL prepared the original draft. KE and PL led the data analysis. GP conducted the literature search and managed the data. DBA conceived and supervised the project.

REFERENCES

- 1 Peeters A, Gearon E, Backholer K, Carstensen B. Trends in the skewness of the body mass index distribution among urban Australian adults, 1980 to 2007. *Ann Epidemiol* 2015; **25**: 26–33.
- 2 Komlos J, Brabec M. The trend of BMI values of US adults by deciles, birth cohorts 1882–1986 stratified by gender and ethnicity. *Econ Hum Biol* 2011; **9**: 234–250.
- 3 Beyerlein A. Quantile regression-opportunities and challenges from a user's perspective. *Am J Epidemiol* 2014; **180**: 330–331.
- 4 Redden DT, Fernandez JR, Allison DB. A simple significance test for quantile regression. *Stat Med* 2004; **23**: 2587–2597.
- 5 Beyerlein A, Toschke AM, von Kries R. Breastfeeding and childhood obesity: shift of the entire BMI distribution or only the upper parts? *Obesity (Silver Spring)* 2008; **16**: 2730–2733.
- 6 Garn SM, Sullivan TV, Hawthorne VM. Educational level, fatness, and fatness differences between husbands and wives. *Am J Clin Nutr* 1989; **50**: 740–745.
- 7 Tavani A, Negri E, La Vecchia C. Determinants of body mass index: a study from northern Italy. *Int J Obes Relat Metab Disord* 1994; **18**: 497–502.
- 8 Hermann S, Rohrmann S, Linseisen J, May AM, Kunst A, Besson H et al. The association of education with body mass index and waist circumference in the EPIC-PANACEA study. *BMC Public Health* 2011; **11**: 169.
- 9 Zhang Q, Wang Y. Trends in the association between obesity and socioeconomic status in US adults: 1971 to 2000. *Obes Res* 2004; **12**: 1622–1632.
- 10 Malarious A, Seidell JC, Sans S, Tuomilehto J, Kuulasmaa K. Educational level, relative body weight, and changes in their association over 10 years: an international perspective from the WHO MONICA Project. *Am J Public Health* 2000; **90**: 1260–1268.
- 11 Mujahid MS, Diez Roux AV, Borrell LN, Nieto FJ. Cross-sectional and longitudinal associations of BMI with socioeconomic characteristics. *Obes Res* 2005; **13**: 1412–1421.
- 12 Jolliffe D. Overweight and poor? On the relationship between income and the body mass index. *Econ Hum Biol* 2011; **9**: 342–355.
- 13 RAND. *RAND HRS Data, Version M*. RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration: Santa Monica, CA, USA 2013.
- 14 Chalabi Y, Scott DJ, Wuertz D. Flexible distribution modeling with the generalized lambda distribution. *MPRA* 2012; 43333. Available at: <https://mpra.ub.uni-muenchen.de/43333/>.
- 15 Ozturk A, Dale R. A study of fitting the generalized lambda distribution to solar radiation data. *J Appl Meteorol* 1982; **21**: 10.
- 16 Demirtas H. Multiple imputation under the generalized lambda distribution. *J Biopharm Stat* 2009; **19**: 77–89.
- 17 Neyman J, Pearson ES. On the problem of the most efficient tests of statistical hypotheses. *Philos Trans R Soc Lond A* 1933; **231**: 289–337.
- 18 Pearson K. Method of moments and method of maximum likelihood. *Biometrika* 1936; **28**: 34–59.
- 19 Su S. Fitting flexible parametric regression models with GLDreg in R. *J Mod Appl Stat Methods* 2016; **15**: 46.
- 20 Akaike H. Information theory and an extension of the maximum likelihood principle. In: Kotz S, Johnson NL, eds. *Breakthroughs in Statistics: Foundations and Basic Theory*. Springer: New York, NY, USA, 1992; 610–624.
- 21 Ng M, Liu P, Thomson B, Murray CJ. A novel method for estimating distributions of body mass index. *Popul Health Metr* 2016; **14**: 6.

Supplementary Information accompanies this paper on International Journal of Obesity website (<http://www.nature.com/ijo>)