

Estimating conditional distributions using a method based on S-distributions: Reference percentile curves for body mass index in Spanish Children

Jaume March¹, Javier Trujillano¹, Merc Tor²
and Albert Sorribas¹

¹Departament de Ciències Mèdiques Bàsiques
Universitat de Lleida
Av. Rovira Roure, 44, 25198-Lleida, Spain

²Servei Municipal de Salut Pública
La Paeria, Lleida, Spain

Correspondence to:

Albert Sorribas
Departament de Ciències Mèdiques Bàsiques
Universitat de Lleida
Av. Rovira Roure, 44, 25198-Lleida, Spain.
Tel: 973 702 406. Fax: 973 702 426.
e-mail: Albert.Sorribas@cmb.udl.es

ABSTRACT: *Background:* Reference intervals are a fundamental tool for characterizing the health status in a given population and play a central role in defining diagnostic values in clinical applications. Estimation of the conditional distribution of a variable, as the body mass index (BMI), is necessary for providing reference values when there is a trend as a function of the covariate.

Subjects and Method: We studied 1453 boys and young between 5 and 16 years old measured in a study carried out in the schools of Lleida (Spain). BMI conditional distributions with age have been derived using a new parametric method based on the one proposed by Sorribas *et al.* [Stat. Med. (2000) 19:697-713]. This method is based on S-distributions as a parametric model for the distribution and uses maximum likelihood estimation of the conditional distribution.

Results: The methods commonly used for estimating reference curves are based on a smoothing of sample quantiles using different techniques.

However, these methods do not provide information on the conditional distribution of the target variable. Our method provides an estimation of such distribution and the corresponding reference curves for the quantiles as a function of a covariate, in our case age.

Conclusions: The suggested methodology provides appropriate reference quantiles for the BMI. Our results allow characterizing the change in distribution within the age range considered. Besides describing a raise in BMI with age, we observe an increase in dispersion around puberty. This must be considered when using BMI as a diagnostic variable.

KEY WORDS: Standard growth curves. Reference intervals. Conditional distribution.

INTRODUCTION

Estimation of percentile values – standard curves – for a variable of interest as a function of a

given covariate is a fundamental tool for defining appropriate reference ranges of the target variable in a given population. In many growth studies, the covariate of interest is age and the dependent variables are either weight, height (Summika *et al.* 2001), or other growth-related variables as bone mineral density (Xu *et al.* 1997). In all cases, the main objective is to obtain smoothed curves of the quantiles. Once obtained, these curves allow to interpolate reference ranges at any desired value of the covariate. The methods currently used in analyzing cross-sectional data provide a non-parametric estimate of the corresponding quantile values. As a result, we obtain smoothed curve computed from the sample quantiles without explicit reference to the variable distribution (Bonellie *et al.* 1996; Cole *et al.* 1992,1998; Healy *et al.* 1998; Pan *et al.* 1990; Royston, 1991; Royston *et al.* 1991; Tango, 1998; Wright *et al.* 1997).

Although these techniques provide useful results, a method for characterizing the conditional distribution of the dependent variable for each value of the covariate would be more appropriate. With such a method we would be able of parametrically describing the trend on the underlying distribution as a function of the covariate. From these results, computation on quantiles would be straightforward. Furthermore, we could use the derived conditional distributions for other purposes. For instance, we could generate random samples of the target population in simulation studies.

From a practical point of view, the estimation of the conditional distribution of a biometric characteristic, like weight, height or concentration of a certain analytic value, could be approached from different perspectives. The most commonly used procedure is to transform the original variable, so that the resulting variable approaches normality. Age-related percentiles on the transformed variable are then obtained by estimating the parameters of the resulting normal distribution. Using these results, age-related percentiles on the original variable are finally computed by an inverse transformation.

Although this procedure produces good results in many cases, it has some potential limitations when

applied to the problem we are facing. First, the same transformation for all the ages can be inappropriate or wrong. In such a case, it may be difficult to derive a consistent transformation and a subsequent statistical model that can be used for interpolating reference intervals for any age. Second, there is no guaranty that a suitable transformation exists for all ages. As a solution, different semi-parametric methods based on a Kernel estimation of the density function have been proposed and they can provide some help. However, their implementation and use are complicated (Rossiter, 1991). Third, independently of the method used, the problem of appropriately describing the observed trend remains. The procedure based on a family of distributions known as S-distribution is an alternative that provides a practical solution to these problems (Voit, 1992; Voit *et al.* 1994, 1995; Balthis *et al.* 1996; Sorribas *et al.* 2000).

The S-distribution is a parametric family of distributions that can be used as a general parametric model for univariate data (Voit, 1992). Within this family, we can obtain a proper S-distribution for the collected data without the need of any assumption on the shape and other characteristics of the unknown underlying distribution. The properties of the S-distribution family assure that the estimated distribution is a valid approximation to the true distribution (Voit, 1992; Sorribas *et al.*, 2002). Besides, as the S-distribution family can accommodate any shape in unimodal distributions, it is particularly appropriated for describing an observed trend as a function of a covariate (Voit *et al.* 1994, 1995; Balthis *et al.* 1996; Sorribas *et al.* 2000). Recently, we developed a method for estimating standard curves based on S distributions in cross sectional studies (Sorribas *et al.* 2000). This method provided excellent results when compared with the usual non-parametric methods.

The goal of the present work is to present an improvement of this methodology by introducing a new strategy of defining the groups of age and by using a novel numerical likelihood estimation method to obtain the S-distribution parameters. We show the utility of this approach in the establishment of the reference values (percentiles) of the

body mass Index (BMI) conditional to age. As a result, we will obtain a parametric model for the conditional distribution of BMI as a function of age. This model can be used to interpolate the distribution of this variable for any age of the studied range and for computing the required quantiles.

DATA SET AND METHOD

Subjects

Male children, with age between 5 and 16 years old, selected from the schools of the city of Lleida (Spain) have been studied. Determinations of weight and height for the BMI calculation were carried out by an experienced team, assuring homogeneous measurement conditions. Each boy's age was calculated from his birth date to the day of measurement. The collection of data was part of a wider cross-sectional study carried out by the Servei de Salut Pública de la Paeria in a public health survey within a joint research project between La Paeria (Lleida's City Council) and the Biostatistics and Biomathematics group of the Departament de Ciències Mèdiques Bàsiques of the Universitat de Lleida (Spain). The data were collected between May of 1999 and May of 2000.

Method: The S-Distribution as a parametric model for univariate and unimodal continuous random variables

Fitting a distribution to data requires knowing the underlying distribution. Otherwise, we should explore a group of distributions, and select the one that provides a better fit to the observed data. Since the choice of any of the possible distributions is somehow arbitrary, we could consider fitting families of statistical distributions that include different particular distributions with specific characteristics as particular cases (Johnson and Kotz, 1970). However, these families are too complicated for their generalized use. Alternatively, the default option consists on adjusting a normal distribution to the data after an appropriate data transformation. The most common method for the reference intervals calculation, consist on using a Box-Cox transformation and then to fit a normal distribution to the

transformed data (Horn *et al.* 1998).

The family known as S-distribution is a practical alternative to obtain a parametric model for the data without the need of assuming a particular distribution. This family is defined by a differential equation where the cumulative distribution function, F , is the independent variable (Voit, 1992):

$$\frac{dF}{dX} = \alpha(F^g - F^h) \quad F(X_0) = F_0 \quad (1)$$

with $\alpha > 0$ and $h > g$. Parameters g and h are related with the shape of the distribution, while α is inversely related with its spread. The initial value is a parameter of localization. In practice, it is common to take the median as X_0 , which corresponds to a value of $F_0 = 0.5$. The mean, variance or any desired moment must be computed by numerical methods as there is not an algebraic expression relating them to the parameters (Voit and Schwacke, 1998). In that sense, interpretation of S-distribution parameters in terms of moments is not as straightforward as in classical distributions. To simplify, we will indicate a certain S-distribution as $S[F_0, X_0, \alpha, g, h]$.

The main advantage of this family of distributions is that it offers a parametric model for unimodal continuous variables. Classical distributions can be accurately represented as S-distributions with well defined parameters. For instance, normal distributions are represented as S-distributions with parameters $g = 0.690923$, $h = 2.885386$ and $\alpha = 0.84307\sigma^{-1}$ (Voit, 1992). Furthermore, S-distributions include distributions that do not correspond to any of the well-known distributions used in statistics. In that sense, the S-distribution provides a general parametric model able of fitting empirical data, even if we ignore the underlying distribution (Voit, 1992, Sorribas *et al.*, 2000, 2002).

An important characteristic of this distribution family is the existence of practically equivalent distributions with different parameters (Sorribas *et al.* 2000). If we consider an S distribution $S[F_0, X_0, \alpha, g, h]$, the density has a maximum at $F_m = \left(\frac{h}{g}\right)^{1/h-g}$ with a value of $f_m = \alpha(F_m^g - F_m^h)$. Then, a practically equivalent S distribution can be ob-

tained by choosing a new value of any of the parameters, say g , and by calculating the new parameters to have the same F_m and f_m values as in the original distribution. This property allows a great flexibility since we can modify the parameters of an estimated distribution obtaining an essentially equivalent distribution that maintains the adjustment to the observed data. This property is fundamental for the method that is presented in this work.

In the next section of this paper we will first outline a maximum-likelihood method (MLE) for obtaining S-distribution parameter estimates. Next, we will introduce a modification of our previous procedure (Sorribas *et al.*, 2000) for obtaining conditional distributions using S-distributions as a parametric model. Then, we will use an example of the body mass index (BMI) data to demonstrate the utility of this method. We will also calculate smooth quantile curves or percentiles based on the S distribution parameter trend obtained with our method.

RESULTS

Fitting an S-distribution to observed data

We can fit an S-distribution to observed data by using a numeric procedure that provides the maximum-likelihood estimation (MLE) of the parameters (Voit, 2000). In the original method, only the shape parameters g and h were obtained by MLE, while estimation of X_0 and α requires a complementary least-squares procedure. Here, we present a new numeric procedure that provides a MLE for all the parameters simultaneously. Briefly, this new method is implemented as follows:

1. Consider the data set: $\{X_1, \dots, X_n\}$, and a set of initial values for the unknown parameters: $\theta_1 = \{F_0, X_0, \alpha, g, h\}_1$. It is convenient to fix a value of $F_0 = 0.5$, so that the starting value X_0 will be the median. With that, there are four parameters to estimate.
2. Compute the corresponding value of $F(X_1) = F_1$ for each data point using the corresponding set of parameters. To do that, we integrate equation (1) from X_0 until the desired X_1 .
3. Once the F_1 values are computed, we can easily calculate the likelihood as:

$$L = \prod_{i=1}^n f(X_1) = \prod_{i=1}^n \alpha (F_i^g - F_i^h)$$

4. To obtain the MLE estimation, run a maximization procedure repeating steps 2 and 3 and changing the parameter values until a maximum for the likelihood is reached. We have used *Mathematica*[®] to implement this procedure through the built-in function *FindMinimum*.

Figure 1 shows different examples to appreciate the performance of this method. In each case, data have been simulated from a given distribution by a random generation procedure. Using the numerical MLE method, we have fitted an *S distribution* to the simulated data, obtaining the distribution indicated by a discontinuous line. In all cases, the method yields appropriate estimates. The flexibility shown by the *S-distribution* to fit data sets with very different shapes is a fundamental property that will provide an appropriate tool for deriving conditional distributions as a function of a covariate.

Estimating conditional distributions

Conditional distributions could be estimated for each age using data corresponding to a given age-group. However, this procedure would provide the best fit in each group, but would provide neither a smooth trend for the whole data set nor a method for interpolating the conditional distribution between ages. In order to obtain an appropriate trend, a method based on *S-distributions* was suggested (Voit *et al.* 1994, 1995; Sorribas *et al.* 2000). This procedure consists on estimating a tendency in the *S-distribution* parameters, so that we can obtain interpolation functions for each parameter. Once obtained, it is immediate to compute the corresponding conditional distribution for any value of the covariate. However, although this approach provides good results in the situations tested (Sorribas *et al.* 2000), there are some steps that can be modified to obtain better results.

In this paper, we introduce two improvements on the original method. First, we modify the procedure for defining the groups of age. Second, we use the

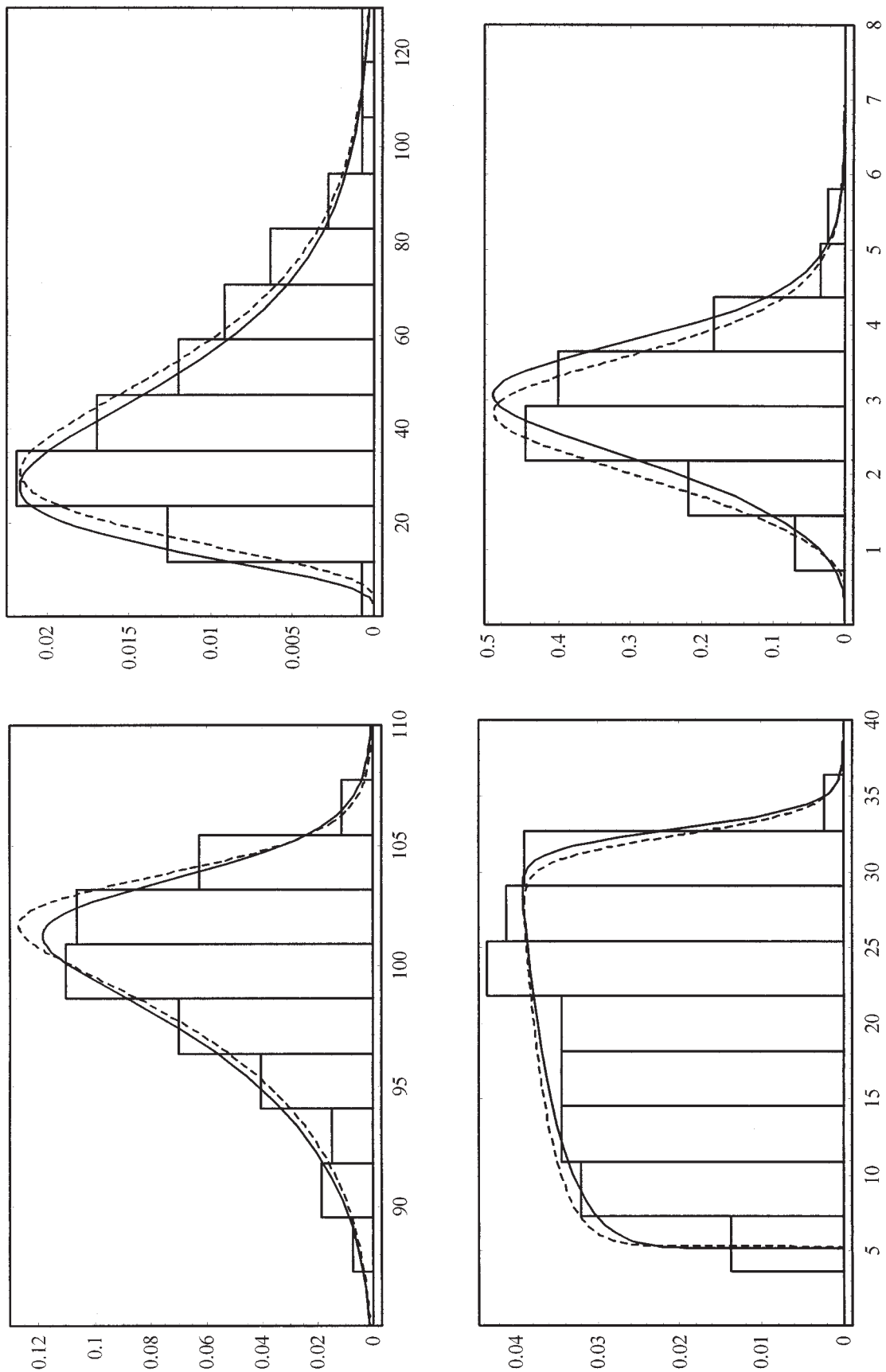


Figure 1. Maximum-likelihood estimation of S-distribution parameters from simulated data. Data are randomly generated from a given distribution (continuous line). The fitted S-distribution (discontinuous line) is shown for each data set.

MLE method indicated above for parameter estimation. Third, we use the property of equivalence between different S-distributions for obtaining the final trend in parameters.

As a first step, we select the age groups assuring that intra-group dispersion doesn't exceed a certain level. To achieve this, the first group of age is selected beginning with the youngest subject, and includes subjects that surpass their age in less than a given value. This value must be adjusted in practice to assure an adequate group composition. This choice is somehow arbitrary and must be adjusted in a problem dependent basis. The goal is to obtain sufficiently homogeneous groups in terms of the dependent variable maintaining a sufficient number of individuals in each group. In our case, we shall use a limit of 0.55 years as we obtain homogeneous groups and a minimum of 45 subjects in each group. The second group then begins in the 15th individual and ends with the individual that surpasses their age as maximum 0.55 years. The third group begins in the 30th individual and follows the same rule. Each resulting group will be represented by its average age, calculated according to the individuals included in the group. This strategy assures that groups are overlapped with a smooth transition among different groups. Although this strategy increases the number of groups, results are better than defining age groups *a priori*. In Figure 2 the data (2a) and the groups resulting from the referred procedure (2b) are shown.

Once the groups are built, we must estimate the trend in distributions so that a smooth description can be derived. As indicated above, we shall use a modification of the method previously described elsewhere (Sorribas *et al.* 2000). The reader is referred to this paper for further details on the underlying rationale of this method. The steps for obtaining the desired conditional distributions are:

1. Estimate the median of BMI for each group. Then we fit a polynomial function to data, using the mean age of each group as independent variable. The resulting polynomial is used to smooth the value of the median of each age group. These values are used as initial values

of the corresponding *S-distribution* for each group.

2. We estimate the parameters α, g, h by MLE for each age group. With that, we obtain the best *S-distribution* for each group. this estimation must be refined to obtain a smooth change in distribution providing the best results for the whole data set.
3. The estimated α values, calculated in the previous point, are then represented as a function of age and a polynomial is fitted to this data. The resulting polynomial is used to smooth the values of α in each group.
4. The estimated g values, calculated in point 2, are represented as a function of age. A polynomial is fitted to this data. The resulting polynomial is used to smooth the values of g in each group.
5. Using the α, g, h values estimated by MLE as a reference and the new parameter values estimated by the interpolations made in steps 3rd and 4th, new h values are recomputed in each age group, in order to obtain a final distribution similar to that obtained by MLE. With that strategy, we will obtain distributions closed to the optimal MLE with a defined trend for the initial value and the parameters α, g, h .
6. The h values estimated in the previous point are finally represented for each age group. A polynomial is fitted to this data. The resulting polynomial is used to smooth the h values in each group. In this step, we obtain new distributions for each group that differ from the optimal MLE. However, the resulting set of distributions show a defined trend that describes the observed tendency in the whole data set. In this step, we sacrifice some accuracy in each group to obtain a smooth description of the overall trend. This is done by a fine tuning of the distribution of each group to obtain an appropriate correspondence with its neighborhoods.
7. The desired age-specific quantiles are calculated using the resulting *S-distributions* in each age. This distribution is calculated from the interpolation polynomials obtained in the points

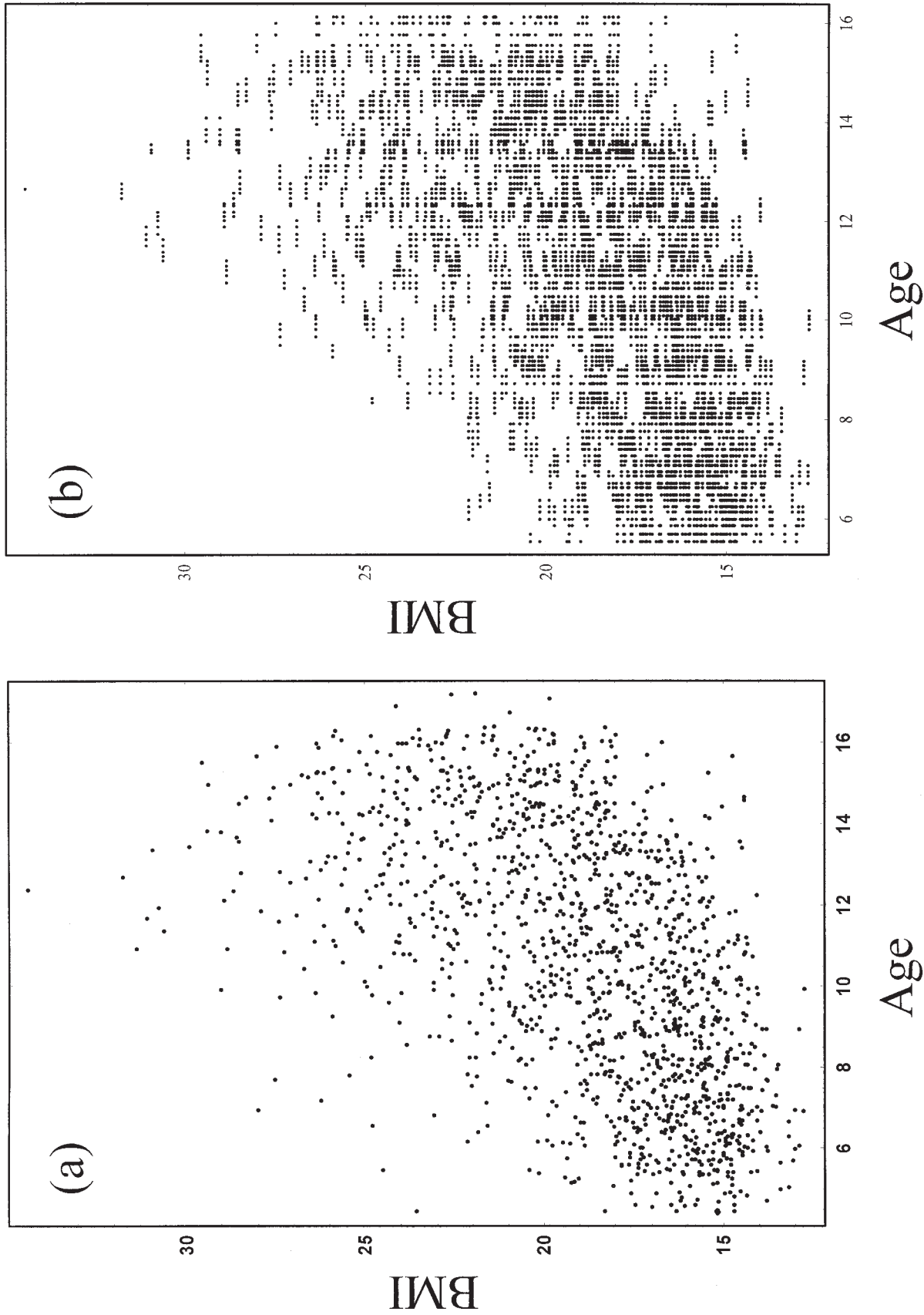


Figure 2. BMI data as a function of age. (a) Sample values. (b) BMI values for each age group computed according to the described procedure (see text).

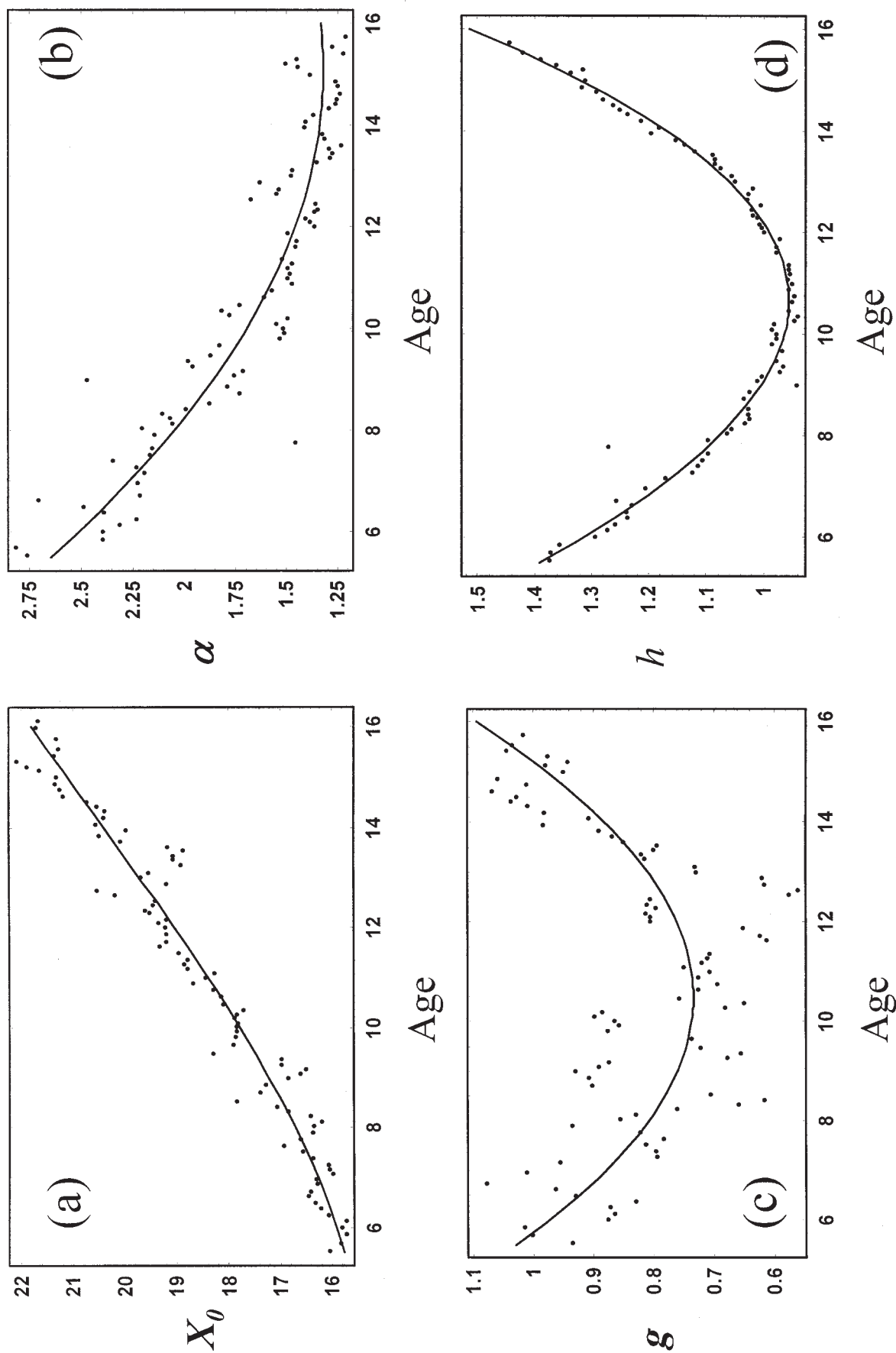


Figure 3. Estimating parameter trends from data. (a) Polynomial interpolation for the observed median (b-c) Polynomial smoothing of the α and g parameter values obtained by MLE on each group of age. (d) Polynomial smoothing for h values (see text for details).

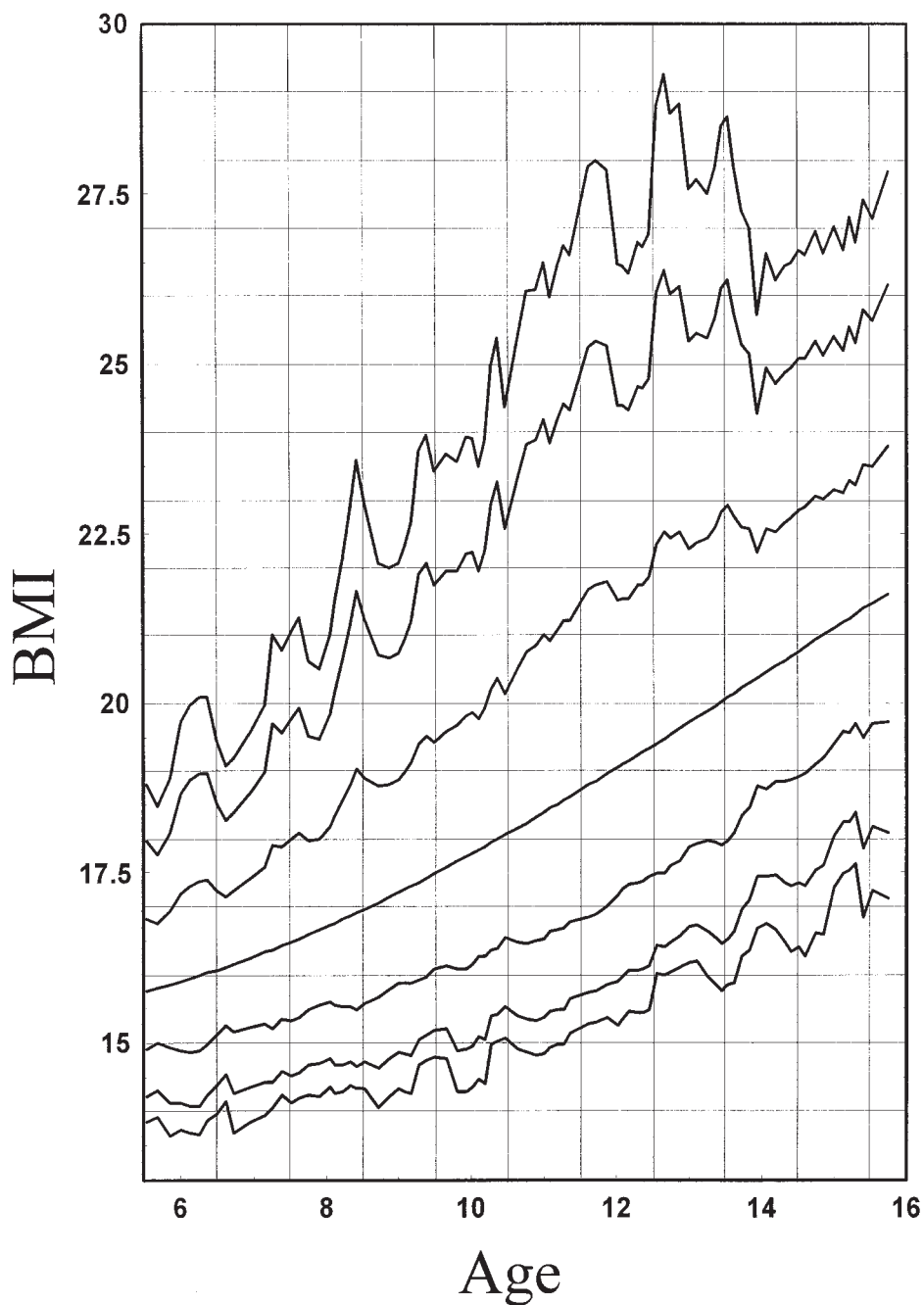


Figure 4. Quantile curves computed from distributions fitted by maximum-likelihood. The fitting procedure of an S-distribution for each group provides a non-smoothed quantile curve that reflects the variability for each sample. Smooth curves require a readjustment of the parameters to account for the age-related trend.

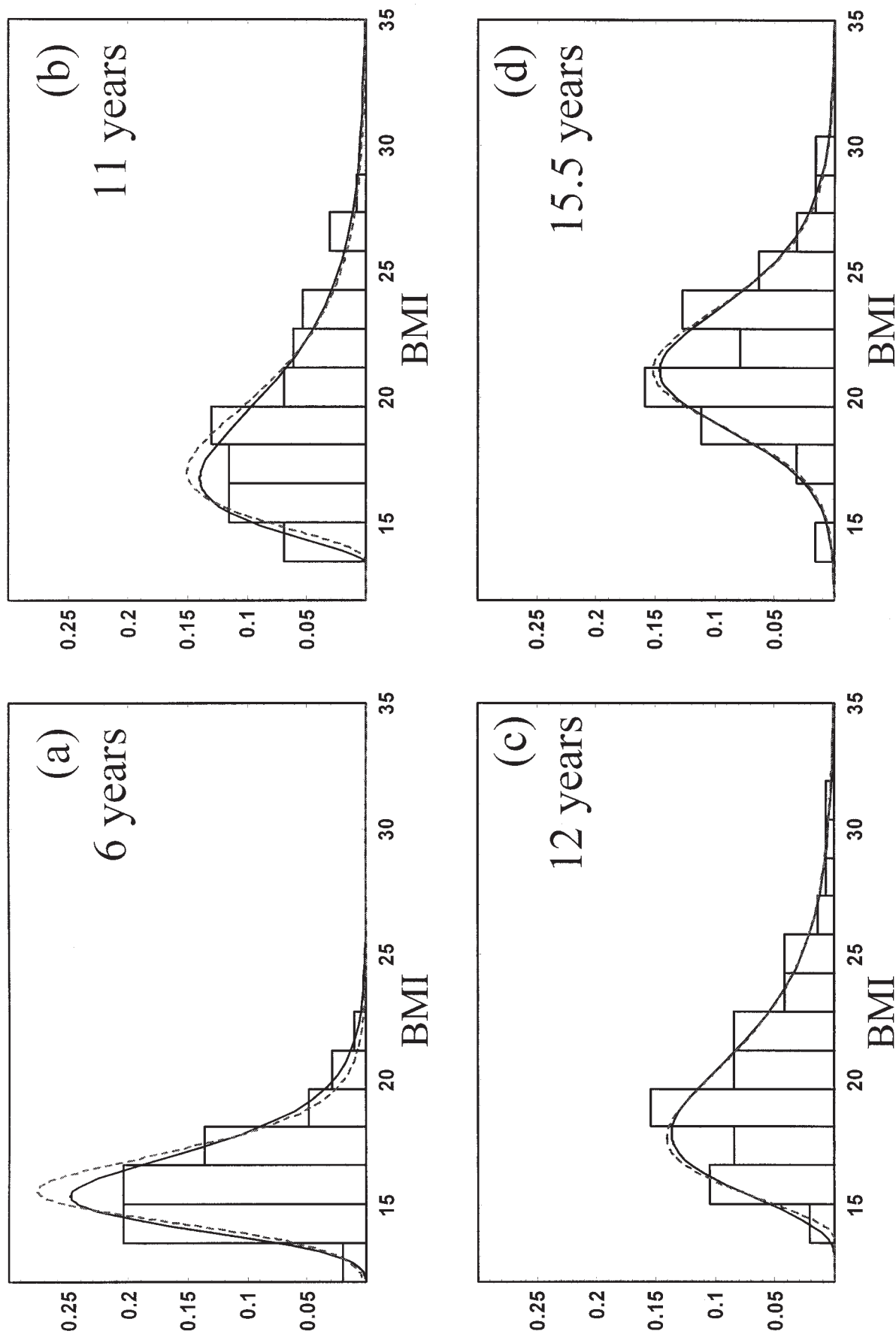


Figure 5. Comparison between S-distributions fitted by maximum-likelihood (continuous curve) and S-distributions interpolated (discontinuous curve). In each case, the corresponding histogram for the original data is shown (a) 6 years; (b) 11 years; (c) 12 years; (d) 15.5 years.

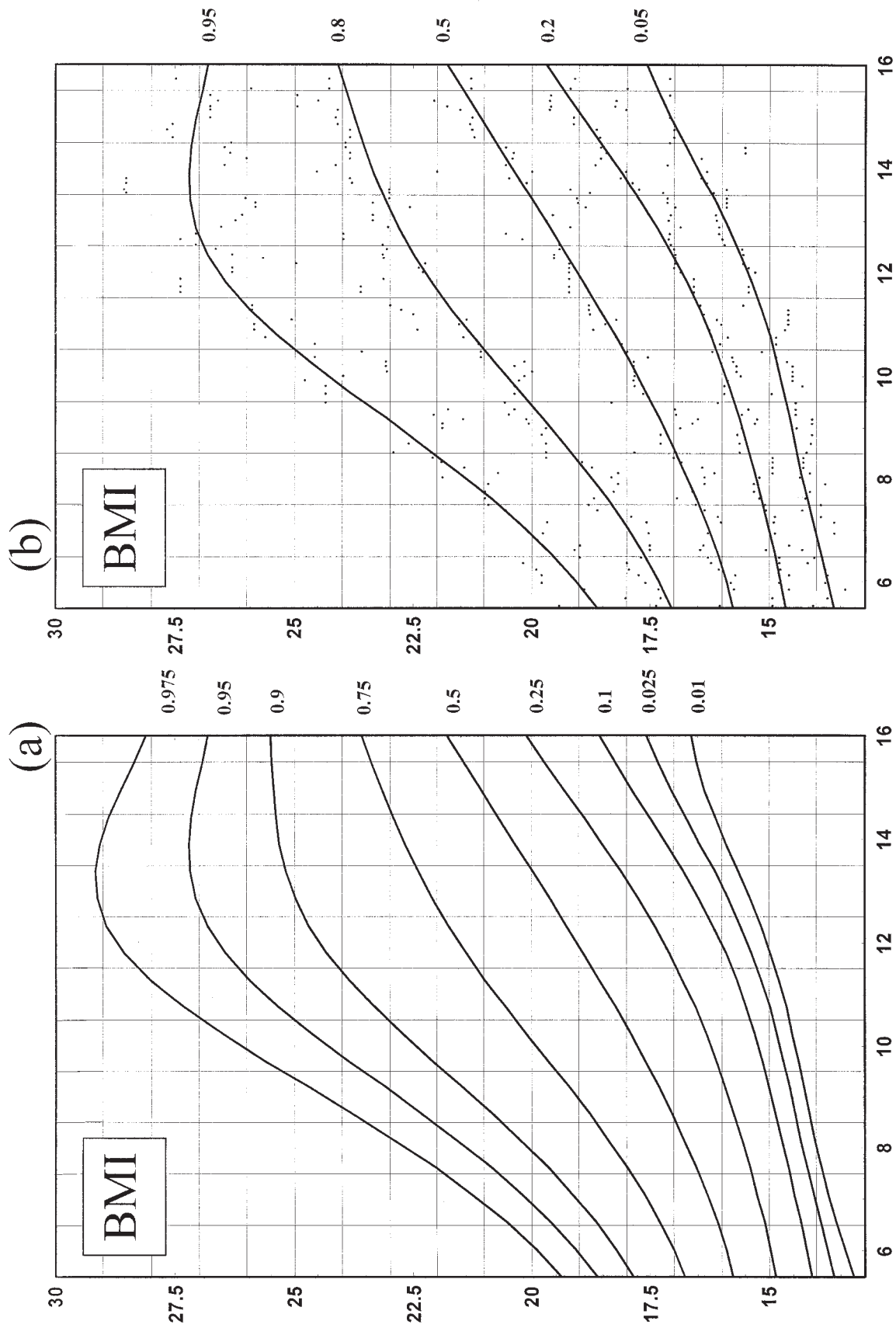


Figure 6. Standard curves of quantiles for BMI data estimated from the S-distribution method. (a) Standard curves. (b) Comparison between standard curves and non-smoothed quantiles. Bold numbers indicate probability values.

1, 3, 4, and 6.

This algorithm was implemented in a *Mathematica*® notebook and the entire process is automatically computed once the age groups are established. We are currently working in a C++ program that will perform all the required computations and that will be available for public use upon request to the authors.

Estimating BMI conditional distribution as a function of age

The previous procedure has been applied to our data set. Following this procedure, the interpolation polynomial for the medians of BMI as a function of age, Figure 3a, was:

$$X_0(x) = 15.978 - 0.407x + 0.0763x^2$$

Using the values of this polynomial, the *S-distribution* parameters, corresponding to the different age groups, were estimated by maximum-likelihood. Figure 4 shows the quantile curves corresponding to this estimate. Obviously, results accurately describe the variation in the different groups without providing a smooth model. In the Figure 5 we can see some examples of the fitted distributions (continuous curve).

In order to smooth the calculated percentiles, the parameters α , g , h were smoothed according to the procedure described above. The conditional distribution of the BMI depending on the age (x) is then obtained as a S-distribution with parameters:

$$\begin{aligned} X_0(x) &= 15.978 - 0.407x + 0.0763x^2 \\ \alpha(x) &= 4.610 - 0.436x + 0.0145x^2 \\ g(x) &= 2.044 - 0.250x + 0.0119x^2 \\ h(x) &= 2.634 - 0.282x + 0.0856x^2 \end{aligned}$$

Figure 3 (b-c-d) shows the results of the smoothing of the parameters for the entire age range.

As we can see in Figure 5, the loss of fitting caused by the smoothing (conditioned distribution: discontinuous curves) is not greatly significant.

Finally, the resulting reference percentiles for the BMI are computed using the interpolation poly-

nomials. These percentiles can be obtained by using the inverse S-distribution:

$$\frac{dX}{dF} = \frac{1}{\alpha(F^g - F^h)} \quad X(F_0) = X_0 \quad (2)$$

and integrating from F_0 until the desired probability value (Voit, 1992). The S-distribution parameters will be selected for each age by using the interpolation polynomials obtained above. The quantile solution derived by Hernández-Bermejo and Sorribas (2001) can also be used in computing the desired quantiles as an alternative to numerically integrating the inverse S-distribution. When applied to the entire age range, we obtain smooth curves as a function of age (Figure 6). As a comparison, these curves are shown together with the sample percentiles calculated on each age group. Our procedure produces an interpolation of these sample percentiles and provides a way for obtaining the conditional distribution of BMI for any value of age within the considered range.

A global assessment of the reference percentiles in these figures reveals an increasing dispersion for values of BMI between 10 and 14 years, dispersion that diminishes in upper groups, from 15 years. These results are consistent with known results on the evolution of BMI around puberty (see for instance Wang *et al.* 1999).

DISCUSSION

The methodology developed in this paper for estimating the trend of change in distribution allows characterizing the variation of the distribution of the target variable in the studied population. When applied to the study of the BMI in a sample of children of different ages, the resulting trend describes the events related to growth between 5 and 16 years, with an increase in dispersion towards obesity values at ages corresponding to pre-puberty and puberty stadiums. This is in accordance with other studies showing the influence of pubertal status on growth-related variables. The observed trend confirms differences between maturation degrees due to hormonal changes and nutritional habits. These variations must be taken into account when using

BMI values in assessing the development status of a child.

The method proposed in this work is of general application for estimating the conditional distributions for variables of clinical or biological interest. This method can be applied to any continuous unimodal variable that shows a trend as a function of a covariate, not necessarily the age. The *S-distribution* based method provides an alternative to using non-parametric methods, such as bootstrap estimation of the conditional percentiles. Previously published results showed that our method produced comparable results to those obtained by smoothing methods on sample percentiles (Sorribas *et al.*, 2000). However, the method developed in this paper provides additional information, as we obtain a parametric conditional distribution that can be used to compute the desired percentiles or for any other purpose. Extension to include multimodal variables is under investigation.

A possible improvement of the procedure proposed in this paper consists on using the results of this method as initial estimates of an algorithm of maximization of the global likelihood. From this initial estimation, the procedure must search for the optimum set of parameters corresponding to the interpolation polynomials used to smooth the parameters as a function of age. A preliminary study of this new strategy indicates that it may provide some improvement of the results of the procedure presented in this paper, although calculation effort is increased considerably.

ACKNOWLEDGEMENTS

This work has been financed by a grant from FIS (00/0235). The authors are grateful to the City Council of Lleida (La Paeria) for allowing the access to the data corresponding to the Joint Project between La Paeria and the University of Lleida (C0160). We also acknowledge the commentaries of an anonymous referee leading to substantial improvement of the original manuscript.

REFERENCES

- BALTHIS, W.L., VOIT, E.O. & MEABURN, G.M. 1996. Setting prediction limits for mercury concentrations in fish having high bioaccumulation potential. *Environmetrics*, 7, 429-439.
- BONELLIE, S.R. & RAAB, G.M. 1996. A comparison of different approaches for fitting percentile curves to birth weight data. *Stat. Med.*, 15, 2657-2667.
- COLE, T.J., FREEMAN, J.V. & PREECE, M.A. 1998. British 1990 growth reference percentiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. *Stat. Med.*, 17, 407-429.
- COLE, T.J. & GREEN, P.J. 1992. Smoothing reference percentile curves: the LMS method and penalized likelihood. *Stat. Med.*, 11, 1305-1319.
- HEALY, M.J., RASBASH, J. & YANG, M. 1988. Distribution-free estimation of age-related percentiles. *Ann. Hum. Biol.*, 15, 17-22.
- HERNÁNDEZ-BERMEJO, B. & SORRIBAS, A. 2001. Analytical quantile solution for the *S-distribution*, random number generation and statistical data modeling. *Biometrical J.* 43, 117-125
- HORN, P.S., PESCE, A.J. & COPELAND, B.E. 1998. A robust approach to reference interval estimation and evaluation. *Clinical Chem.*, 44, 622-631.
- JOHNSON, N.L. & KOTZ, S. 1970. *Continuous univariate distributions. Vol. 1.* Boston MA: Houghton Mifflin.
- PAN, H.Q., GOLDSTEIN H. & YANG Q. 1990. Non-parametric estimation of age-related percentiles over wide age ranges. *Ann. Hum. Biol.*, 17, 475-481.
- ROSSITER, J.E. 1991. Calculating percentile curves using kernel density estimation methods with application to infant kidney lengths. *Stat. Med.*, 10, 1693-1701.
- ROYSTON, P. & MATTHEWS, J.N.S. 1991. Estimation of reference ranges from normal samples. *Stat. Med.*, 10, 691-695.
- ROYSTON, P. 1991. Constructing time-specific reference ranges. *Stat. Med.*, 10, 675-690.
- SORRIBAS, A., MARCH, J. & VOIT, E.O. 2000. Estimating age-related trends in cross-sectional studies using S-distributions. *Stat. Med.*, 19, 697-713.
- SORRIBAS, A., MARCH, J. & TRUJILLANO, J. 2002. A new parametric method based on S-distributions for computing receiver operating characteristic curves for continuous diagnostic tests. *Stat. Med.*, 15, 1213-1235
- SUMIYA, T., NAKAHARA, H. & SHOHOJI, T. 2001. Relationships among biological growth parameters for body weight in Japanese children. *Growth, Dev. & Aging*, 64, 91-112.
- TANGO, T. 1998. Estimation of age-specific reference ranges via smoother AVAS. *Stat. Med.*, 17, 1231-1243.

- VOIT, E.O. 1992. The S-distribution: A tool for approximation and classification of univariate, unimodal probability distributions. *Biometrical J.*, 34, 855-878.
- VOIT, E.O. 2000. A maximum likelihood estimator for shape parameters of S-distributions. *Biometrical J.*, 42, 471-479.
- VOIT, E.O., BALTHIS, W.L. & HOLSER, R.A. 1995. Hierarchical Monte-Carlo modeling with S-distributions: concepts and illustrative analysis of mercury contamination in king mackerel. *Environmental Int.*, 21, 627-635.
- VOIT, E.O., & SCHWACKE L.H. 1998. Scalability properties of the S-distribution, *Biometrical. J.*, 40, 665-684.
- VOIT, EO & YU, S. 1994. The S-distribution. approximation of discrete distributions. *Biometrical J.*, 36, 205-219.
- WANG J., HORLIK, M., THORNTON, J.C., LEVINE L.S., HEYMSFIELD, S.B. & PIERSON R.N. 1999. Correlations between skeletal muscle mass and bone mass in children 6-18 years: influences of sex, ethnicity and pubertal status. *Growth, Dev. & Aging*, 63(3), 99-109.
- WRIGHT, E.M. & ROYSTON, P. 1997. Simplified estimation of age-related reference intervals for skewed data. *Stat. Med.*, 16, 2785-2803.
- XU, S., HUANG, W. & REN, J. 1997. The new model of age-dependent changes in bone mineral density. *Growth, Dev. & Aging*, 61(1), 19-26.