STATISTICS IN MEDICINE
Statist. Med. 2002; **21**:1213–1235 (DOI: 10.1002/sim.1086)

A new parametric method based on *S*-distributions for computing receiver operating characteristic curves for continuous diagnostic tests

Albert Sorribas*,†, Jaume March and Javier Trujillano

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

SUMMARY

Receiver operating characteristic (ROC) curves provides a method for evaluating the performance of a diagnostic test. These curves represent the true positive ratio, that is, the true positives among those affected by the disease, as a function of the false positive ratio, that is, the false positives among the healthy, corresponding to each possible value of the diagnostic variable. When the diagnostic variable is continuous, the corresponding ROC curve is also continuous. However, estimation of such curve through the analysis of sample data yields a step-line, unless some assumption is made on the underlying distribution of the considered variable. Since the actual distribution of the diagnostic test is seldom known, it is difficult to select an appropriate distribution for practical use. Data transformation may offer a solution but also may introduce a distortion on the evaluation of the diagnostic test. In this paper we show that the distribution family known as the *S*-distribution can be used to solve this problem. The *S*-distribution is defined as a differential equation in which the dependent variable is the cumulative. This special form provides a highly flexible family of distributions that can be used as models for unknown distributions. It has been shown that classical statistical distributions can be represented accurately as *S*-distributions and that they occur in a definite subspace of the parameter space corresponding to the whole *S*-distribution family. Consequently, many other distributional forms that do not correspond to known distributions are provided by the *S*-distribution. This property can be used to model observed data for unknown distributions and is very useful in constructing parametric ROC curves in those cases. After fitting an *S*-distribution to the observed samples of diseased and healthy populations, ROC curve computation is straightforward. A ROC curve can be considered as the solution of a differential equation in which the dependent variable is the ratio of true positives and the independent variable is the ratio of false positives. This equation can be easily obtained from the *S*-distributions fitted to observed data. Using these results, we can compute pointwise confidence bands for the ROC curve and the corresponding area under the curve. We shall compare this approach with the empirical and the *binormal* methods for estimating a ROC curve to show that the *S*-distribution based method is a useful parametric procedure. Copyright © 2002 John Wiley & Sons, Ltd.

KEY WORDS: ROC curves; diagnostic test; *S*-distribution; *S*-system; power-law formalism

* Correspondence to: Albert Sorribas, Departament de Ciències Mèdiques Bàsiques, Universitat de Lleida, Av. Rovira Roure 44, 25198-Lleida, Spain
† E-mail: albert.sorribas@cmb.udl.es

Contract/grant sponsor: FIS, Spain; contract/grant number: 00/0235
Contract/grant sponsor: La Paeria(Lleida); contract/grant number: X0148

Received June 2000
Accepted July 2001

1. INTRODUCTION

Receiver operating characteristic (ROC) curves are widely used in clinical practice to assess the performance of a diagnostic test [1–3]. A ROC curve represents the joint values of the true positive ratio TPR (*sensitivity*) and false positive ratio FPR (1-*specificity*) for each value of the diagnostic variable. In some cases, the diagnostic variable is a discrete score and the resulting ROC curve is a step-line. In other cases, the diagnostic variable is continuous and would correspond to a continuous ROC curve [3–6].

In the case of a continuous diagnostic test one may ask if it would be possible to obtain a continuous ROC curve from the sample data. An obvious method involves the assumption of a given distribution accounting for the statistical behaviour of the test variable in the healthy and diseased populations. The most popular alternative is considering normal or log-normal distributions [6–8]. This strategy yields appropriate results so long as the selected distribution is appropriate [6, 7]. In practice (see for instance Sorribas *et al.* [9]), the search for an appropriate distribution is a subjective and uncertain process. This is a potential drawback for this approach, unless enough information exists on the underlying distribution.

To overcome these problems, different non-parametric methods have been suggested. The most popular one is based on a step-by-step computation of the TPR and FPR on the observed test values [1, 2]. In such cases, the resulting curve is a step-line that approximates the theoretical ROC curve. Using this methodology, the area under the empirical ROC curve (AUC) can be calculated as the Mann–Whitney version of the two-sample rank-sum statistic of Wilcoxon [10, 11].

An alternative approach is based on a *kernel* estimation of the density function. The ROC curve is computed using the estimated densities. The algorithm known as ROC&ROL can be used for performing the required computations [12]. In this method, no assumption is needed for the underlying distributions and a smooth ROC curve can be derived from the kernel estimates of the corresponding densities. A minor disadvantage of this approach is the need to select an appropriate kernel function and the corresponding bandwidth for an accurate result. With an appropriate choice, the resulting ROC curve is a good estimation of the actual one.

The *binormal* method [4, 5, 13] provides another interesting solution. This method is based on the assumption that the ROC curve has a binormal form. This is equivalent to assuming that some monotonic transformation of the test variable produces a normal distribution for both populations. ROC curve estimation is obtained by a maximum likelihood procedure on the transformed z-scores. The program ROCKIT implements a quasi-ML procedure (algorithm LABROC 5) for computing the required ROC curve using this approach [5]. The binormal approach produces appropriate ROC curves in many applications, although it may be inaccurate in some special situations [14].

Although all those methods are appropriate in most cases, it would be interesting to investigate if a more general approach can be defined. As a goal, such a method should skip the assumption of a particular distribution, it should solve the potential problems related to non-parametric and semi-parametric methods, and it should provide a continuous ROC curve.

A solution would be to use a sufficiently general family of distributions that could account for the actual unknown distributions of the diagnostic test in healthy and diseased

populations. Ideally, such a family should include as particular cases as many classical distributions as possible. Besides, it would be desirable that this family could account for distributions that do not correspond to any of the classical ones. Although some of such families actually exist [15–17], they are often too complicated for practical use. An exception may be the *S*-distribution family. The *S*-distribution is defined as a differential equation in which the cumulative is the dependent variable [18]. This provides a useful family of distributions that contains most continuous unimodal distributions as special cases. Furthermore, known distributions occur in definite regions of the *S*-distribution parameter space but they do not fill all the theoretically possible combinations [18]. Therefore, *S*-distributions provide an infinity of new models different from the classical distributions. This possibility makes the *S*-distribution an interesting framework for data analysis [9, 19, 20]. We shall see that it is also an appropriate tool for ROC curve estimation.

In this paper we shall present the application of the *S*-distribution to the computation of ROC curves showing that no previous assumption, other than being unimodal and continuous, is needed on the underlying distributions. First, we shall briefly review the idea of *S*-distribution and the method for deriving an appropriate *S*-distribution for a given data set. Second, we shall discuss how to compute a ROC curve and its corresponding pointwise confidence bands using the fitted *S*-distributions. Furthermore, we shall see that computation of the area under the ROC curve (AUC) is straightforward using *S*-distributions. By simulation experiments we shall compare our method with the usual empiric (non-parametric) and binormal (semi-parametric) approaches. Finally, we present a couple of application examples using real data.

2. ROC CURVE COMPUTATION USING *S*-DISTRIBUTIONS

2.1. Motivation

Diagnostic tests based on continuous clinical variables must deal with a great variety of distributional forms. It is common that diseased people show highly asymmetrical distributions, while healthy people tend to follow more symmetrical distributions. Also, it is common to observe distributions with long tails to the left and a sharp right tail close to zero. Real examples of these situations are illustrated in Figures 13 and 14. Here, we wish to use clinical measurements taken on patients during the first 24 hours of their admission to intensive care units (ICU) to predict subsequent events. In Figure 13 worst mean arterial pressure is used to predict the future need for inotropic agents, and in Figure 14 urinary output is used to predict acute renal failure.

Although these situations can be studied either by the non-parametric or by the semi-parametric methods, a parametric description would provide a model for the test variable in both populations. Such a model can facilitate some computations and it can be used for different purposes. For instance, a parametric model can be used to generate random samples that can be used in simulation studies. Also, a parametric model can be useful for discussing the calibration and discrimination of the diagnostic test.

With this in mind, we shall present a method for computing a ROC curve based on the *S*-distribution. The *S*-distribution is a parametric family that can provide an accurate representation of most unimodal statistical distributions. We shall first briefly review the use of

the S-distribution for data representation. Then we will discuss the ROC curve computation from the estimated S-distributions.

2.2. Data representation using S-distributions

The S-distribution [18, 21, 22] is defined in terms of a differential equation in which the cumulative, F, is the dependent variable [18]:

$$\frac{\mathrm{d}F}{\mathrm{d}X} = \alpha(F^g - F^h), \quad F(X_0) = F_0, \ \ \alpha > 0, \ \ h > g \tag{1}$$

The density, PDF, is thus defined as a function of the cumulative, CDF (F in (1)), and the random variable does not appear explicitly. For a given value of X_i, the corresponding CDF, is obtained by numerically integrating (1) from X_0 to X_i [18]. Parameters g and h are responsible for the shape of the S-distribution, while parameter α is inversely related to the variance of the variable. The initial condition $F(X_0) = F_0$ determines the location of the resulting distribution. For a given value of X, its PDF can be computed as $f(X) = \alpha(F(X)^g - F(X)^h)$. For simplicity, an S-distribution with parameters α, g and h, and initial condition $X(F_0) = X_0$, will be indicated as $\mathrm{S}[F_0, X_0, \alpha, g, h]$. Technical details for computing quantiles are discussed in the Appendix.

As defined, the S-distribution can accurately represent most of the known unimodal statistical distributions. Each *classical* statistical distribution (that is, Normal, Weibull, central and non-central Student's t, etc.) corresponds to a point in the S-distribution parameter space and the resulting representation preserves the limit relationships between variables [18]. Besides this property, the high flexibility of the S-distribution results in a whole new scenario in which the *classical* distributions are embedded in a more general framework where other distributions are possible. This makes the S-distribution a good candidate for data modelling [9, 19, 20, 22].

S-distributions can be fitted to observed data to obtain an estimation of the underlying distribution. Different methods for estimating the S-distribution parameters are available in the literature [9, 19, 20, 23]. For practical purposes, we will use the two-step procedure suggested by Sorribas *et al.* [9]. According to this method, we organize the data of a sample of size n in a histogram in which the total area equals one. The height of the histogram at each class is taken as an estimate of a PDF value and the corresponding CDF are computed by adding the PDF values. Taking $f_i = \mathrm{PDF}_i$ and $F_i = \mathrm{CDF}_i$ we can fit

$$f_i = \alpha(F_i^g - F_i^h) \quad i = 1, \ldots, n \tag{2}$$

using a non-linear fit procedure. Selection of a barwidth is not critical at this step, since the resulting histogram is used only for obtaining initial estimates for the parameters. In practice, we can use the barwidth according to the procedure suggested in reference [12]. From this step, we obtain a first estimation of α, g and h. Once these parameters are computed, we refine the estimation by using the S-distribution differential equation (1) and a least-squares procedure that minimizes the sum of squares between the observed and predicted quantile values. At this step, the previously estimated g and h are taken as fixed values and new values for α and the initial condition X_0 are obtained. For simplicity, a value of $F_0 = 0.5$ is taken so that X_0 is an estimation of the median. Details of this procedure and examples of its performance can be found in reference [9]. Alternatively, the shape parameters g and h of

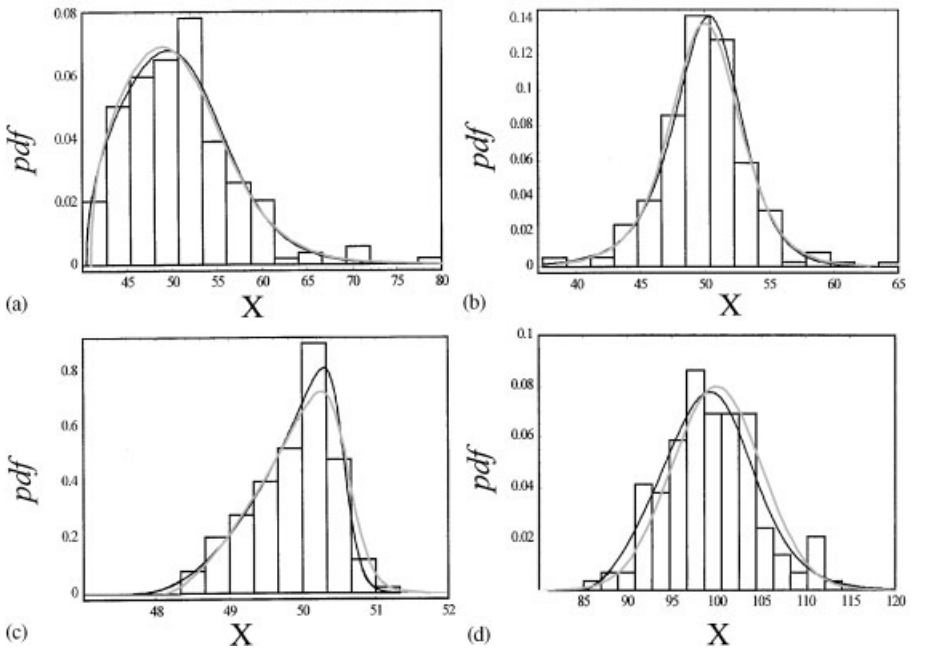

Figure 1. Fitting observed data with *S*-distributions. Data generated from different distributions are fitted by an *S*-distribution. In each case, the sample size is 200. The black line indicates the fitted distribution and the grey line indicates the actual distribution from which data is simulated. (*a*) Original distribution S[0.5, 50, 0.1, 0.3, 3], fitted distribution S[0.5, 50.12, 0.097, 0.32, 3.35]. (*b*) Original distribution S[0.5, 50, 1, 1.2, 1.75], fitted distribution S[0.5, 50.12, 0.625, 1.14, 2.13]. (*c*) Original distribution S[0.5, 50, 1, 0.6, 7], fitted distribution S[0.5, 49.91, 1.07, 0.67, 9.28]. (*d*) Original distribution Normal[100, 5], fitted distribution S[0.5, 99.15, 0.15, 0.64, 2.85].

the *S*-distribution can be obtained by using the corresponding maximum likelihood estimator [23]. In such case, α and the initial condition can be obtained using the same procedure stated above.

Using samples of moderate size, the estimation procedure yields, for each case, an *S*-distribution that closely approximates the original distribution (see examples and discussion on the performance of the *S*-distribution for data representation in references [9, 19, 20, 23]). As an example, we generate four random data sets from different distributions (Figure 1). In the considered examples, only one is generated from a classical distribution (Figure 1(*d*)). All the others are generated from *S*-distributions with no correspondence to known distributions (Figure 1((*a*), (*b*) and (*c*)). In either case, the obtained *S*-distributions closely resemble the actual ones. For practical purposes they can be used as parametric models for each of the unknown underlying distributions.

Once the corresponding *S*-distribution parameters are estimated on a given sample, computation of CDF and quantiles is straightforward (see Appendix). As we shall see, this makes the *S*-distribution specially suited for ROC curve estimation from observed data in continuous

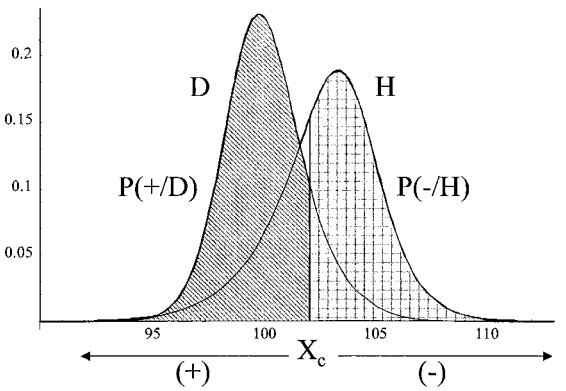

Figure 2. Continuous diagnostic test. The statistical behaviour of X in the diseased (D) and healthy (H) populations can be represented by continuous statistical distributions. From these distributions a continuous diagnostic test can be constructed and the corresponding sensitivity and specificity for a diagnostic point X_c can be easily computed.

diagnostic tests. The main advantage of this approach is that the resulting S-distribution is a parametric model that facilitates further computations.

2.3. ROC curve computation from S-distributions

Let us suppose that in the population suffering the disease (D) the statistical behaviour of a given characteristic can be represented by a random variable with S-distribution:

$$\mathrm{S}[F_{0\mathrm{D}}, X_{0\mathrm{D}}, \alpha_{\mathrm{D}}, g_{\mathrm{D}}, h_{\mathrm{D}}] \tag{3}$$

Let us consider that the same characteristic can be represented by an S-distribution:

$$S[F_{0\mathrm{H}}, X_{0\mathrm{H}}, \alpha_{\mathrm{H}}, g_{\mathrm{H}}, h_{\mathrm{H}}] \tag{4}$$

on the healthy (H) population. For convenience, let us take $F_{0\mathrm{D}}$ and $F_{0\mathrm{H}}$ equal to 0.5, so that the initial condition corresponds to the median of each population. Without lack of generality, we shall consider $X_{0\mathrm{D}} \leqslant X_{0\mathrm{H}}$. As illustration of a generic case, a typical situation is shown in Figure 2.

For a given diagnostic point X_c, an individual is classified as positive ($+$) if $X < X_c$. On the contrary, if $X \geqslant X_c$ the individual is classified as negative ($-$). In selecting an appropriate X_c, two properties shall be taken into account: (i) sensitivity, defined as the probability of a correct diagnosis on a diseased individual, and (ii) specificity, defined as the probability of a correct diagnosis on a healthy individual. According to this definition, in Figure 2 sensitivity can be computed as

$$\text{sensitivity} = P(+|D) = P(X < X_c|D) = F_{\mathrm{D}}(X_c) \tag{5}$$

that is, it can be interpreted as the corresponding CDF for X_c on the diseased population. Similarly, specificity can be computed as

$$\text{specificity} = P(-|H) = P(X > X_c|H) = 1 - F_{\mathrm{H}}(X_c) \tag{6}$$

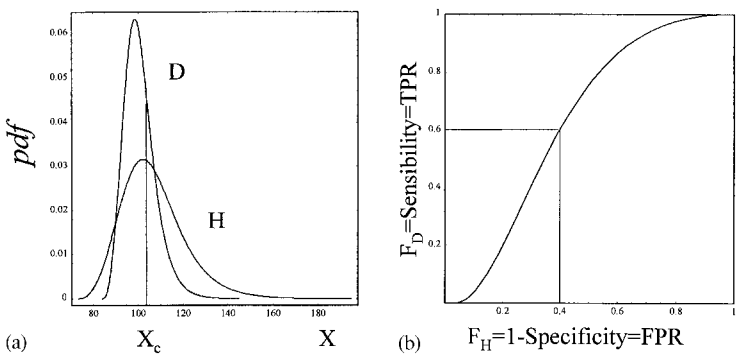

Figure 3. Receiver operating characteristic (ROC) curve. The diagnostic variable follows an *S*-distribution both in the diseased and healthy populations: (*a*) probability density functions; (*b*) ROC curve. The ROC curve can be seen as a trajectory that is a solution of the differential equation dF_D/dF_H (see text).

that is, it can be interpreted as the corresponding 1 – CDF value for X_c on the healthy population. If the diseased population has a distribution shifted to the right of that of the healthy population then sensitivity will correspond to $1 - F_D(X_c)$ and specificity to $F_H(X_c)$. As indicated before, for simplicity we shall consider the case of a diseased population with a distribution shifted to the left of that of the healthy population.

Then, for a given situation in which diseased and healthy populations are represented by *S*-distributions, computation of sensitivity and specificity is straightforward by any of the available techniques [18, 22]. All we need to do is select a given X_c and compute its corresponding CDF. Since sensitivity and specificity are dependent on the selected value of X, it is convenient to represent the change in these quantities as a function of the diagnostic point X_c. This representation should help evaluate the utility of the considered variable as a basis for a correct diagnosis. Receiver operating characteristic (ROC) curves represent the probability of a true positive within the diseased (true positive ratio, TPR) as a function of the false positive among the healthy (false positive ratio, FPR)(Figure 3). In the case we are considering, TPR corresponds to F_D and FPR to F_H. Then, according to (5) and (6), the ROC curve represents F_D as a function of F_H, and can be interpreted as the solution of the differential equation

$$\frac{dF_D}{dF_H} = \phi(F_D, F_H) \tag{7}$$

If the random variable in the diseased and healthy populations is represented as an *S*-distribution, the ROC curve can be obtained from (1) by taking

$$\frac{dF_D}{dX} = \alpha_D(F_D^{g_D} - F_D^{h_D}) \quad F_D(X_{0D}) = 0.5 \tag{8}$$

$$\frac{dF_H}{dX} = \alpha_H(F_H^{g_H} - F_H^{h_H}) \quad F_H(X_{0H}) = 0.5 \tag{9}$$

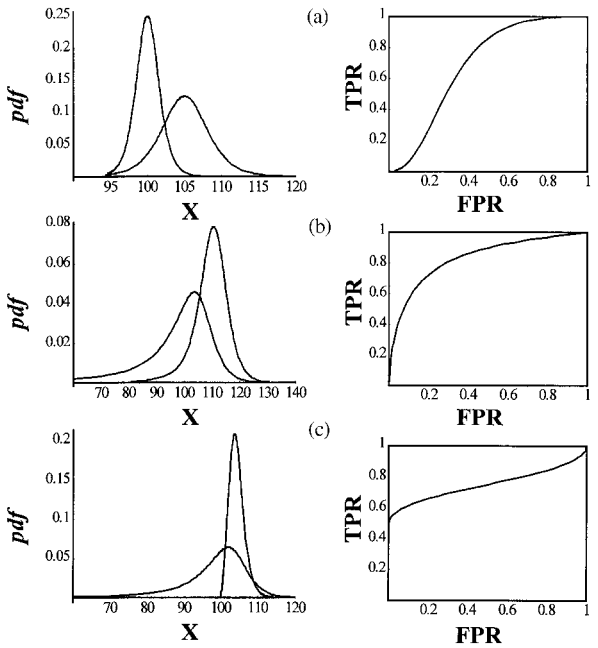

Figure 4. Examples of ROC curve computation from *S*-distributions. For each example, the figure on the left represents the PDF for each population. The figure on the right represents the corresponding ROC curve. (*a*) Diseased population S[0.5, 100.0, 1.0, 1.0, 2.0], healthy population S[0.5, 105.0, 0.5, 1, 2], AUC 0.686199. (*b*) Diseased population S[0.5, 100.0, 0.2, 1.7, 3.2], healthy population S[0.5, 110.0, 0.5, 1.3, 2.0], AUC 0.831128. (*c*) Diseased population S[0.5, 100.0, 0.2, 1.3, 3.2], healthy population S[0.5, 104.0, 0.5, 0.6, 2.0], AUC 0.749436.

and obtain $\phi(F_D, F_H)$ in (7) by dividing (8) and (9). After that, equation (7) can be written as

$$\frac{dF_D}{dF_H} = \frac{\alpha_D(F_D^{g_D} - F_D^{h_D})}{\alpha_H(F_H^{g_H} - F_H^{h_H})} \tag{10}$$

with appropriate initial conditions (see Appendix). These initial conditions can be obtained after selecting a starting value X_s for X, and computing the corresponding $F_D(X_s)$ and $F_H(X_s)$. These values are then used as starting values for integrating (10) and compute the corresponding ROC curve using $F_H = 1$ as a final value of the integration procedure. Although this approach requires numerical integration, it can be easily performed by using any standard mathematical package like *Mathematica* or *Mat-Lab*. Alternatively, all computations can be performed in PLAS, a computer software specially designed to simulate power-law models [24]. We have developed a *Mathematica* package that can perform all the required computations automatically. All results in this paper are obtained using this package. Examples of ROC curve computation using *S*-distributions as models for the underlying statistical distributions are shown in Figure 4.

2.4. Pointwise confidence bands for the ROC curve

Confidence bands for ROC curves can be obtained in different ways [2, 12, 25–29]. The methodology suggested by Zou *et al.* [12] is particularly suited for computing pointwise confidence bands for the ROC curve estimated by using the *S*-distribution method. We shall briefly reproduce this method here to show how it can be adapted to the case of *S*-distribution based computations. Given a value $p = \text{FPR}$, the goal is constructing a confidence interval for the $\hat{q} = \text{TPR}$ obtained in the ROC curve. As suggested by Zou *et al.*, it is convenient to use a logit transformation and compute the corresponding confidence interval in logit-space. Once obtained, an inversion of the logit transformation yields the required confidence interval for $\hat{q}$. If $\hat{v} = \text{logit}(\hat{q}) = \log(1/(1-\hat{q}))$, an estimation of the variance of $\hat{v}$ can be obtained as

$$s^2(\hat{v}) \approx \frac{\frac{1}{m}\hat{\beta}(t)^2\, p(1-p) + \frac{1}{n}\hat{q}(1-\hat{q})}{\hat{q}(1-\hat{q})^2} \tag{11}$$

In this equation, $\hat{\beta}(t) = \hat{f}_{\text{D}}(t)/\hat{f}_{\text{H}}(t)$ and $\hat{q} = \hat{F}_{\text{D}}(t)$ with $t = \hat{F}_{\text{H}}^{-1}(p)$. Since the PDF for a given value t of X, in the case of an *S*-distribution, corresponds to $\alpha(F(t)^g - F(t)^h)$, $\hat{\beta}(t)$ can be expressed in our case as

$$\hat{\beta}(t) = \frac{\alpha_{\text{D}}(\hat{q}^{g_{\text{D}}} - \hat{q}^{h_{\text{D}}})}{\alpha_{\text{H}}(p^{g_{\text{H}}} - p^{h_{\text{H}}})} \tag{12}$$

where the corresponding parameters are those obtained by fitting an *S*-distribution to the corresponding samples of n diseased and m healthy subjects. For a confidence $(1-\alpha)$, the corresponding interval for $\hat{v}$ is

$$\hat{v} \pm z_{1-\alpha/2}\, s(\hat{v}) \tag{13}$$

The desired confidence interval for $\hat{q}$ is obtained as $\text{logit}^{-1}(\hat{v} \pm z_{1-\alpha/2}\ s(\hat{v}))$. The resulting interval is adequate except for extreme values of p. The denominator of (11) becomes zero when $p \to 1 \Rightarrow \hat{q} \to 0$ and when $p \to 0 \Rightarrow \hat{q} \to 1$. In those cases, $s^2(\hat{v}) \to \infty$ (see examples of this situation in Figures 7–13.

2.5. Examples of ROC curve estimation using the S-distribution method

The performance of this method can be tested with some examples. Results in Figure 5 illustrate the different steps of our method. In this example, we compare the resulting ROC curves obtained with the empirical and the *S*-distribution methods on simulated data from known distributions. First, we define a distribution for the test variable in the diseased and healthy populations. Then, random samples are obtained for each case and the corresponding *S*-distributions are fitted (Figure 5(*a*) and (*b*)). A ROC curve is then computed using these estimated *S*-distributions as parametric models. Finally, the confidence bands are obtained. The estimated ROC curve using the *S*-distribution method is compared with the ROC curve obtained by the non-parametric procedure. It can be seen that the *S*-distribution result is close to the actual ROC curve and that the confidence bands contains the actual curve.

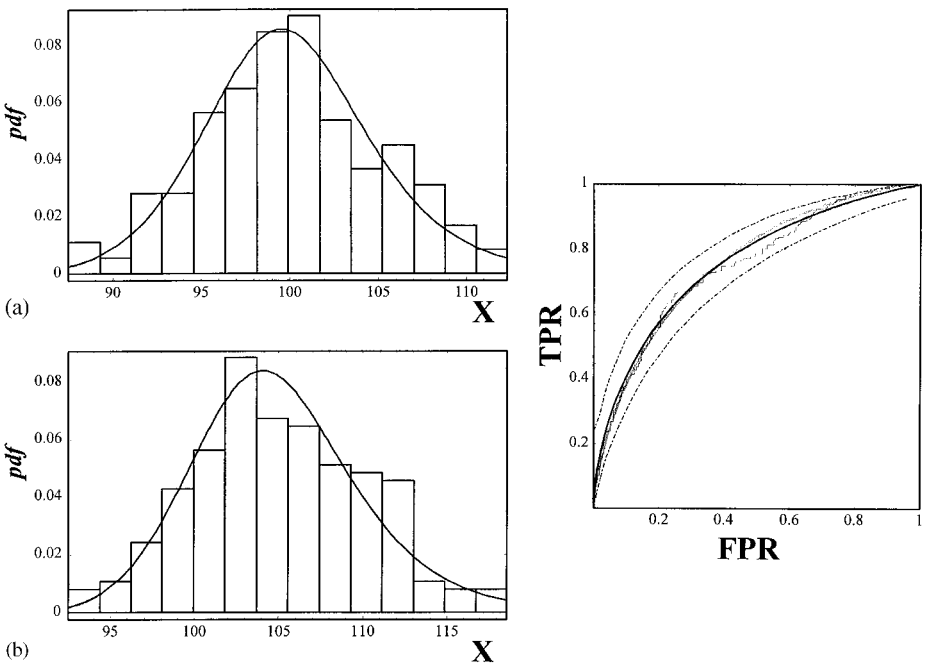

Figure 5. ROC curve estimation from observed data. Samples are obtained from normal distributions: (*a*) diseased population N(100,5), sample size $n = 200$; (*b*) healthy population N(105,5), sample size $n = 200$. In (*a*) and (*b*) the data histograms are compared to the fitted *S*-distributions. The resulting *S*-distributions are: (*a*) S[0.5, 99.8796, 0.28489, 0.84811, 1.9705]; (*b*) S[0.5, 104.634, 0.283396, 0.805393, 1.84707]; (*c*) The theoretical ROC curve (grey line) is compared with the empirical ROC curve (step line) and with the *S*-distribution ROC curve (black line). The dashed lines indicate the 95 per cent confidence bands computed as indicated in the text.

In a second example, we select two normal distributions as underlying distributions for the two populations and generate a set of different simulations. Representative results are shown in Figure 6. The results obtained show that the confidence bands and the ROC curve are meaningful when compared to the actual ROC curve. Also, it can be appreciated that the *S*-distribution approach yields a smooth ROC curve that is more realistic than the empirical step-line. Since we use samples of size = 50 for both populations, data variability is an important factor to be considered. As it appears in Figure 6, in some of the data sets data variability results in inaccurate ROC curves for both the empirical and the *S*-distribution methods. The area under the ROC curve (AUC) is an indicator of the test performance and it can be used to compare the results of different methods on the same data set. In the next section, we shall discuss how to compute the AUC using the *S*-distribution approach and then we shall compare this method with the empirical approach. Additional examples of the performance of the *S*-distribution method for computing ROC curves are provided in the simulation study (see below).

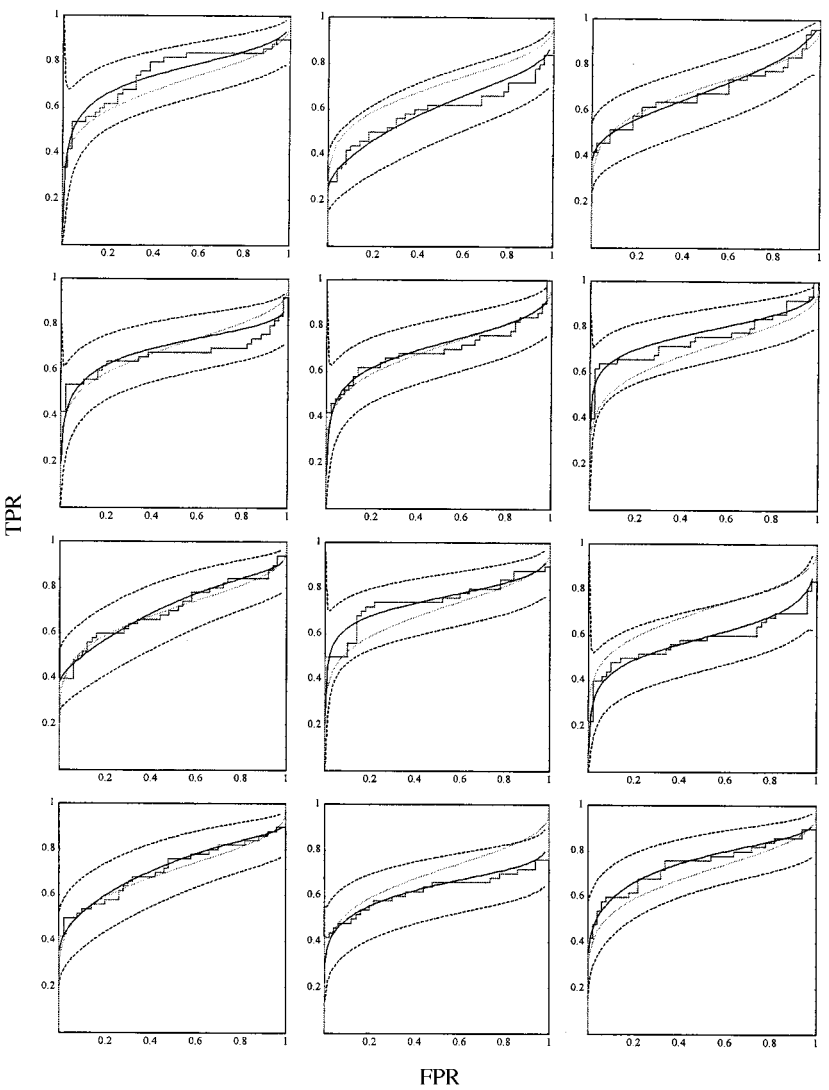

Figure 6. Comparison of the empirical method and the *S*-distribution method for computing ROC curves. Simulated samples of size $n = 50$ are obtained for diseased and healthy populations. The reference distributions are: disease N(100,18); healthy N(110,7). The computed ROC curves and the corresponding confidence bands are indicated as in Figure 5. The theoretical ROC curve (grey line) is compared with the empirical ROC curve (step line) and with the *S*-distribution ROC curve (black line). The dashed lines indicate the 95 per cent confidence bands computed as indicated in the text.

2.6. Computation of the area under the ROC curve using the S-distribution method

The area under the ROC curve, AUC, represents the probability of correctly classifying one diseased and one healthy individual. According to this interpretation, the resulting AUC for

a given situation can be used as a measure for evaluating the performance of the studied variable as a basis for diagnosis. An AUC close to 1 identifies a variable that will produce few misclassifications. When both populations have the same distribution for the considered variable no discrimination is possible and we obtain an AUC of 0.5. Between both extreme situations, the different values allow ranking different variables according to their performance in classifying diseased and healthy individuals. Although the AUC criteria may be arguable as a method for assessing the performance of a diagnostic test in all cases, it is widely used in clinical practice and it is an important aspect of ROC curves.

The AUC can be computed easily as the integral of the ROC curve for F_H between 0 and 1. In our case, this may be obtained by adding the equation

$$\frac{\mathrm{d}AUC}{\mathrm{d}F_H} = F_D, \quad AUC[F_H(0)] = F_D(0) \tag{14}$$

to (10).

3. SIMULATION STUDY

3.1. Performance of the S-distribution method compared to the non-parametric and binormal methods

The performance of the *S*-distribution method for computing ROC curves can be analysed by a simulation study. We shall select different distributions as true distributions and generate random data for each population. Then, each sample will be fitted by an *S*-distribution and the corresponding ROC curve and AUC will be computed using the procedure discussed in this paper. As a reference, the empirical non-parametric method will be used to compute a step-line approach to the ROC curve. The corresponding AUC is computed as the Mann–Witney version of the two-sample rank-sum statistic of Wilcoxon [10, 11].

A comparison of results using both the *S*-distribution method and the empirical method is presented in Figures 7–10. In these examples, we show that both methods yield comparable results in computing the AUC. In each example, estimated AUC values are compared for a set of 200 data sets. ROC curves estimated using the *S*-distribution method and their confidence bands are included for some representative simulations. According to this study, we may conclude that the *S*-distribution method produces AUC values that are equivalent to those resulting from the empirical method. Since the *S*-distribution method produces a continuous ROC curve and a meaningful confidence band, these results suggest that the *S*-distribution method can be used instead of the empirical method.

The binormal method can also be used to obtain a continuous ROC curve. We have compared this method with the *S*-distribution method in different situations. Figures 11 and 12 present some results in which the binormal method is not appropriate [14]. In those special cases, the empirical and the *S*-distribution methods produce reasonable ROC curves, while the binormal method fails to provide a meaningful result. In the example of Figure 12, the binor-

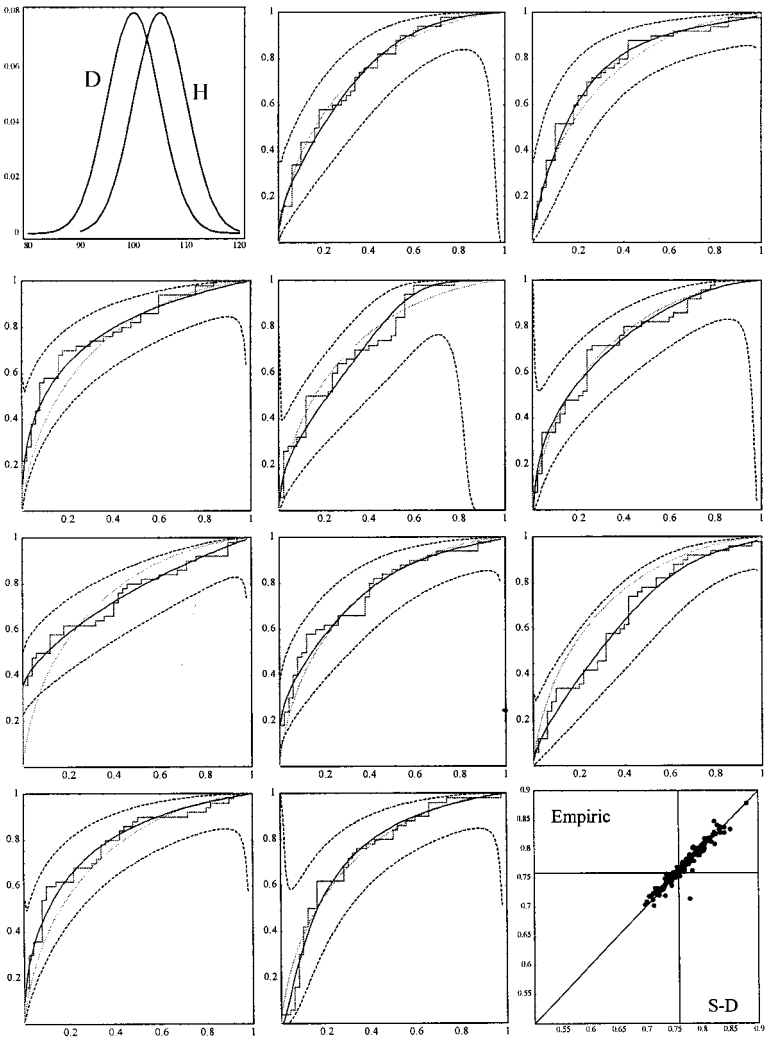

Figure 7. Simulation study. The empirical (non-parametric) and the *S*-distribution methods are compared using random samples of a given situation in which the underlying distributions are: diseased population (D) N(100,5); healthy population (H) N(105,5). The top left figure indicates the two distributions. In the figure at the bottom right corner, the AUC results of 200 simulations with sample size of $n = 50$ for each population are presented. Points indicate the AUC computed with the *S*-distribution (*S*-D) and the empirical non-parametric (Empiric) methods. The diagonal line represents equality between both methods. The vertical and horizontal line indicates the actual value of the AUC computed from the theoretical distributions ($\mathrm{AUC} = 0.759$). Results of 10 simulations out the 200 are included as examples of the results. The labelling of the axes is the same as in preceeding figures and is omitted for clarity.

mal method produces almost equivalent AUC values when compared with the other methods. However, a close look at the ROC curves show that this approach is inadequate in those cases.

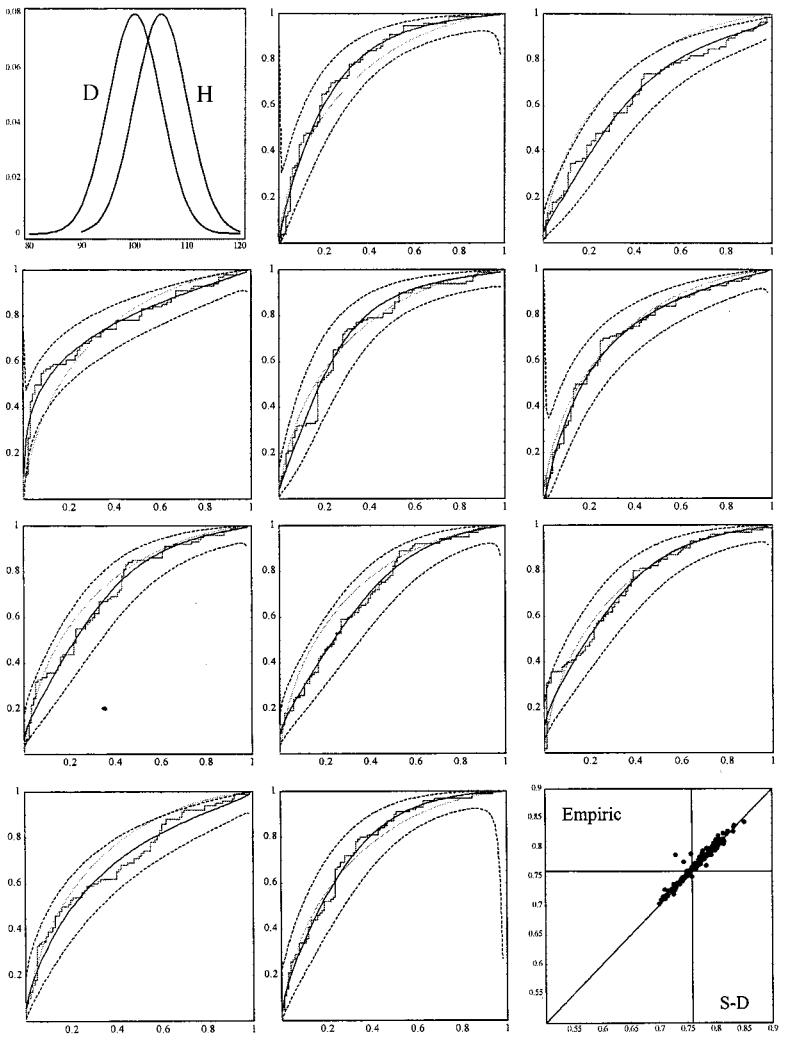

Figure 8. Simulation study. The same conditions as those of simulations in Figure 7 are used. In this case, the sample size is $n = 100$ for each population.

3.2. Application of the S-distribution ROC computation method to actual data

As a final example, we have applied the *S*-distribution ROC computation method to a set of data collected on a database of the Intensive Care Unit (ICU) at the Hospital Arnau de Vilanova (University of Lleida, Spain). A total of 585 patients (1996–1998 period) that satisfy the inclusion criteria (stay longer than 72 hours) are studied.

First we evaluate the use of the worst mean arterial pressure (MAP mmHg) during the first 24 hours of admission into ICU as a diagnostic test to identify those patients that would

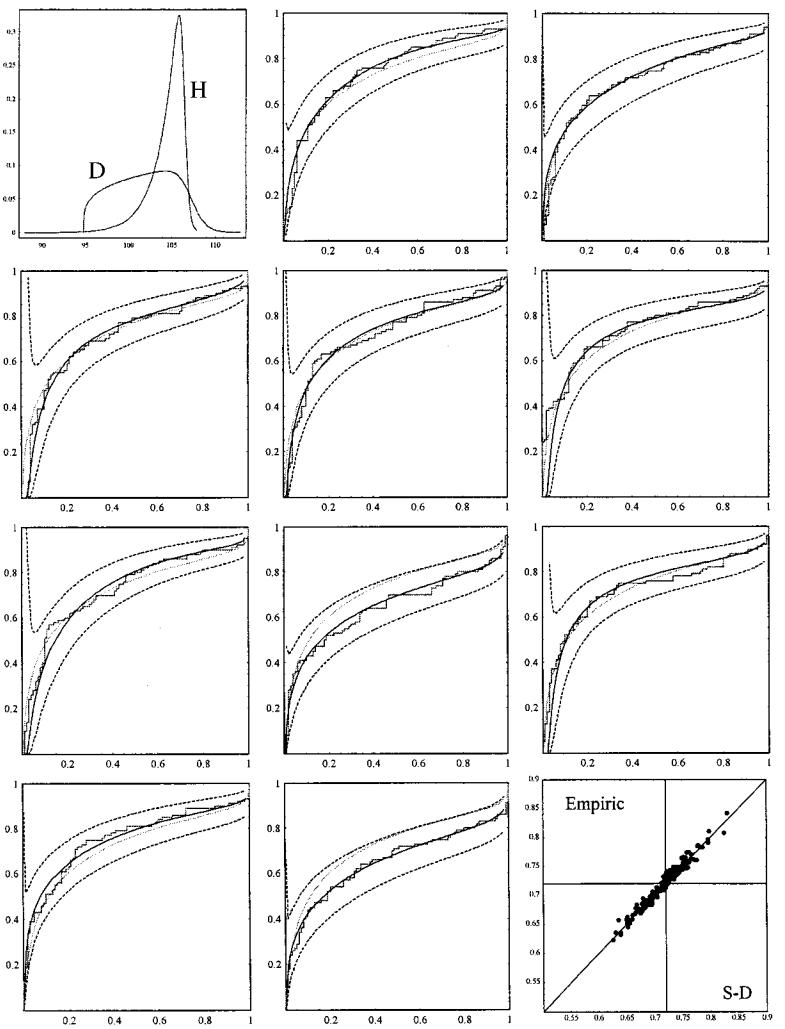

Figure 9. Simulation study. The underlying distributions are: diseased population (D) S[0.5, 102, 0.1, 0.2, 12]; healthy population (H) S[0.5, 105, 0.5, 1, 8]. With these distributions the value of the AUC is 0.721. The simulation study is made with a sample size of $n = 100$ for each population. Results are indicated as in Figure 7.

need inotropic agent therapy during the next 48 hours. In the considered sample, 174 patients of the 585 needed inotropic therapy (disease group). The resulting *S*-distributions for each subpopulation and the corresponding ROC curve are shown in Figure 13. Data from each subpopulation are well described by the fitted *S*-distributions, yielding a smooth ROC curve with an AUC of 0.826.

As a second example, we evaluate the performance of the urinary output (cc/day) of the first 24 hours of admission in ICU as a diagnostic test for acute renal failure in the next 48 hours. In the studied sample, 105 patients of the 585 had renal failure within this period (disease group).

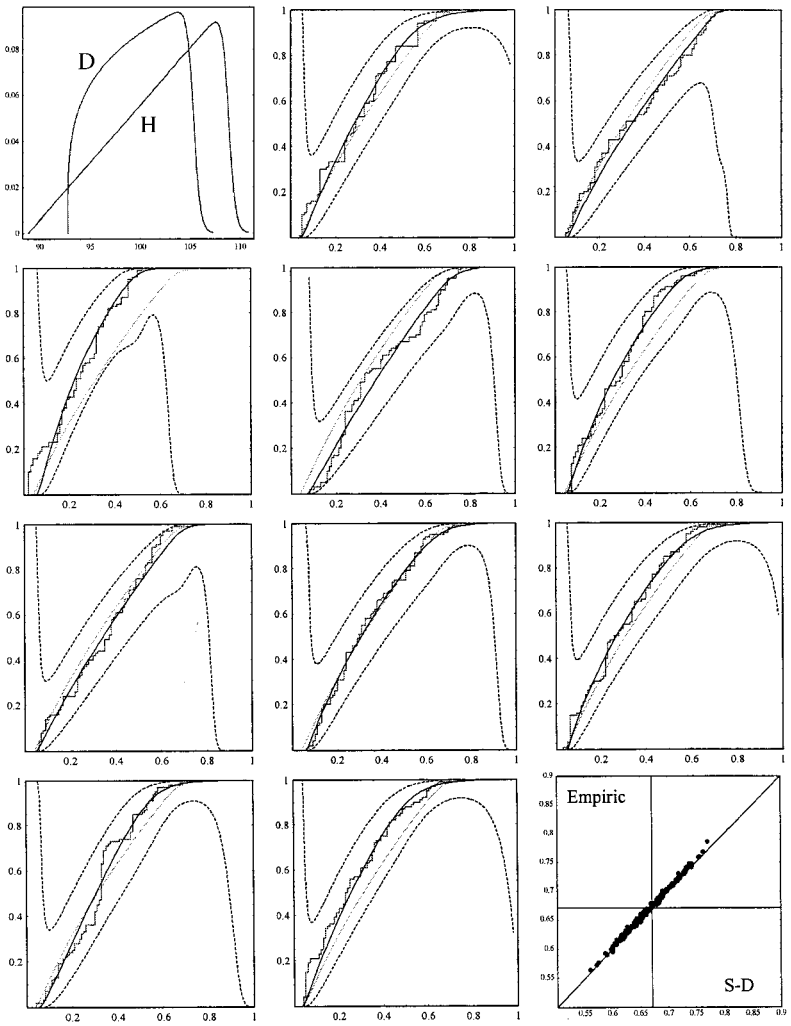

Figure 10. Simulation study. Diseased population (D) S[0.5, 100, 0.1, 0.2, 30]; healthy population (H) S[0.5, 103, 0.1, 0.5, 30]. With these distributions the value of the AUC is 0.671. The simulation study is made with a sample size of $n = 100$ for each population. Results are indicated as in Figure 7.

The resulting S-distributions for each subpopulation and the corresponding ROC curve are shown in Figure 14. In that case, the disease group presents a highly asymmetric distribution as a consequence of patients with low urine output. S-distributions can accurately represent this situation and compute the resulting ROC curve. The situation presented in this example, with a highly asymmetric distribution starting at zero, is representative of the potential of this technique. Such a situation arises for many clinical parameters and may be difficult to model using traditional distributions. In that case, the S-distribution is a good alternative.

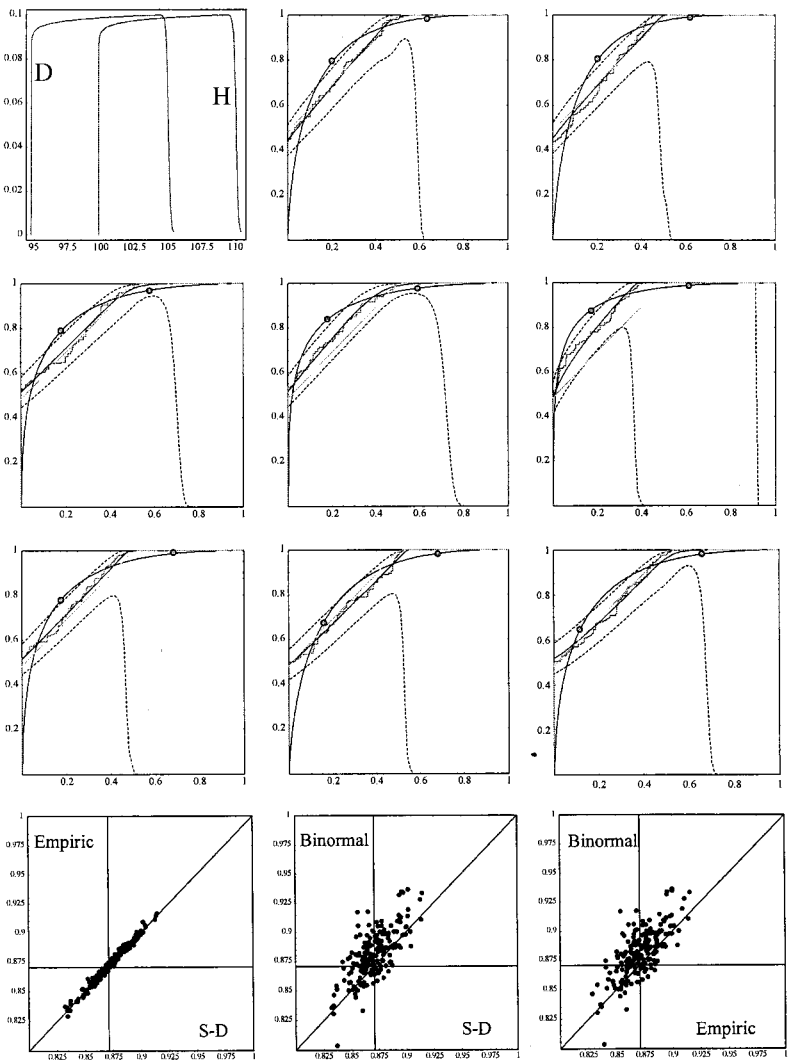

Figure 11. Simulation study. The empirical (non-parametric), binormal (semi-parametric) and the *S*-distribution methods for computing ROC curves are compared using random samples of a given situation in which the underlying distributions are: diseased population (D) S[0.5, 100, 0.1, 0.02, 120]; healthy population (H) S[0.5, 105, 0.1, 0.02, 120]. The top left figure indicates the two distributions. In the figures at the bottom, the AUC results of 200 simulations with sample size of $n = 100$ for each population are presented. Points indicate the AUC computed with the *S*-distribution (*S*-D), empirical non-parametric and binormal methods. The diagonal line represents equality between methods. The vertical and horizontal line indicates the actual value of the AUC computed from the theoretical distributions (AUC = 0.871). Results of eight simulations out the 200 are included as examples of the results. Results are indicated as in Figure 7. Dots are used to identify the ROC curve obtained with the binormal method.

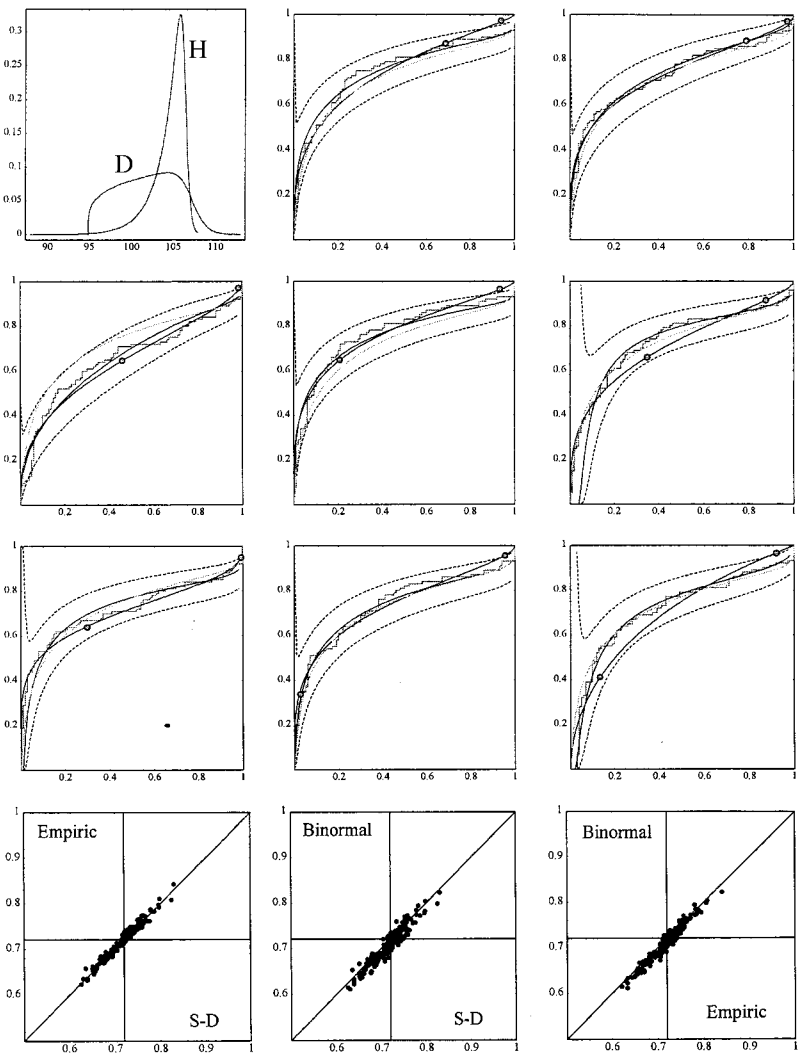

Figure 12. Simulation study. The empirical (non-parametric), binormal (semi-parametric) and the *S*-distribution methods for computing ROC curves are compared using random samples of a given situation in which the underlying distributions are: diseased population (D) S[0.5, 102, 0.1, 0.2, 12]; healthy population (H) S[0.5, 105, 0.5, 1, 8]. The corresponding AUC is 0.721. Sample size for simulations is $n = 100$ for each population. Results are indicated as in Figure 11.

4. CONCLUDING REMARKS

Although other existing methods provide a useful solution to estimating the ROC curve, ROC computation using the method suggested in this paper has several advantages. First, it provides an estimation of the underlying continuous distribution. This estimation, an *S*-

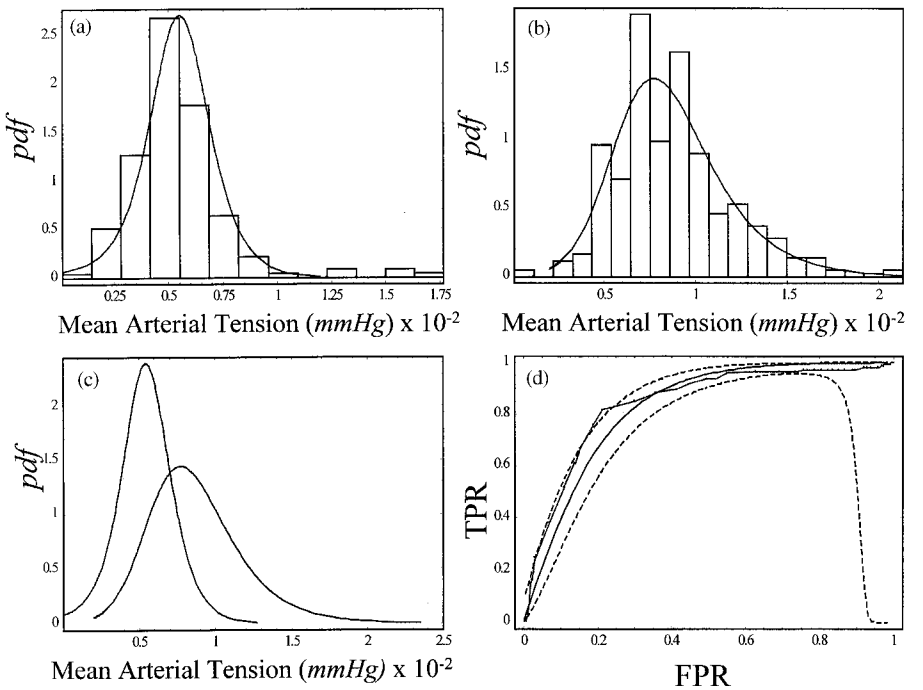

Figure 13. Worst mean arterial pressure (MAP) as a diagnostic test for the need of inotropic therapy. (*a*) Patients that needed inotropic therapy. Estimated *S*-distribution S[0.5, 0.83, 42.62, 1.085, 1.189]. (*b*) Patients that did not need inotropic therapy. Estimated *S*-distribution S[0.01, 0.01, 73.78, 1.4025, 1.549]. (*c*) Comparison of adjusted distributions. (*d*) *S*-distributions ROC curve (continuous line) and 95 per cent confidence bands (dashed lines) are shown. The step line corresponds to the empirical ROC curve.

distribution, can be used easily for computing quantiles and estimating the corresponding ROC curve. Second, the resulting ROC curve is smooth and its computation is straightforward from the estimated distributions. This method takes advantage of the *S*-distribution ability for data representation, yielding a useful ROC curve. Our simulation results show that the *S*-distribution method produces accurate ROC curves when compared to the expected ones. A complementary advantage of deriving a continuous ROC curve is the possibility of computing the TPR for each value of FPR. With that, *S*-distribution ROC curves can be used for evaluating any subset of values of the diagnostic test variable. This possibility is an advantage over using non-parametric approaches.

In this paper we have developed the basic methodology for applying the *S*-distribution to the computation of ROC curves, including confidence intervals and the computation of the area under the curve. Besides these results, there are some complementary improvements that can be easily incorporated. For instance, improved confidence intervals for the ROC curve and the corresponding AUC could be computed using resampling techniques [26, 27]. The AUC for two diagnostic tests can then be compared either using the resulting confidence intervals or by an appropriate modification of the existing methods [10, 11]. These possibilities will be developed in a forthcoming paper.

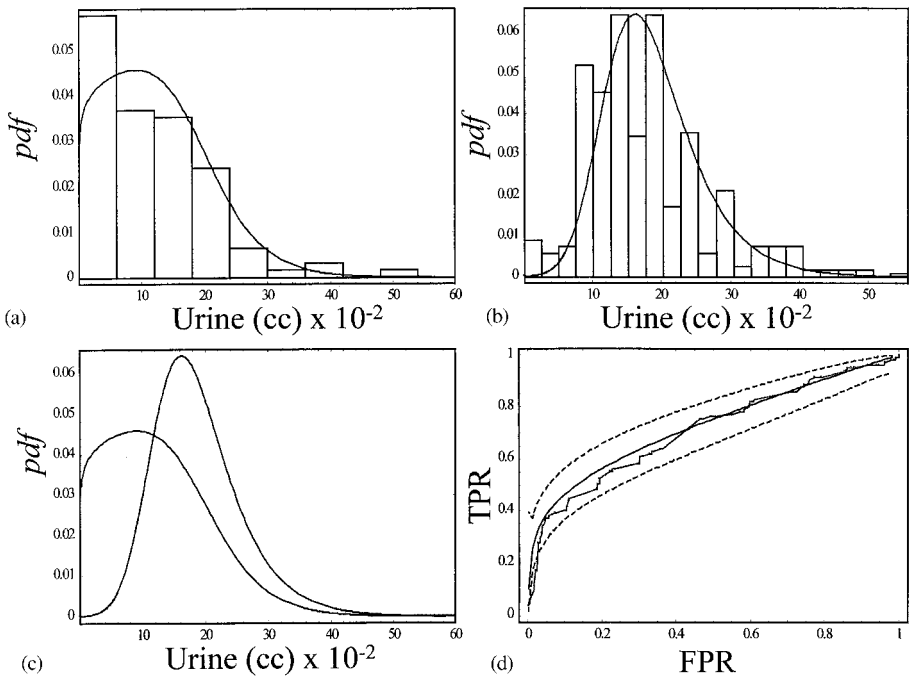

Figure 14. Urine output at 24 hours as a diagnostic test for acute renal failure. (*a*) Patients with acute renal failure. Estimated *S*-distribution S[0.0001, 0.0001, 0.051, 0.091, 4.09]. (*b*) Patients without acute renal failure. Estimated *S*-distribution S[0.0001, 0.0001, 1.061, 0.9941, 1.173]. (*c*) Comparison of adjusted distributions. (*d*) *S*-distributions ROC curve (continuous line) and 95 per cent confidence bands (dashed lines) are shown. The step line corresponds to the empirical ROC curve.

As a conclusion, we have shown that the *S*-distribution provides a parametric model for computing ROC curves based on a continuous diagnostic test and that this approach is a valid option for computing such curves.

APPENDIX

A1. Quantile computations in S-distributions

A quantile equation for an *S*-distribution can be obtained as [18]

$$\frac{\mathrm{d}X}{\mathrm{d}F} = \frac{1}{\alpha(F^g - F^h)} \quad X(F_0) = X_0 \tag{A1}$$

This equation can be used for computing the quantile corresponding to a given probability p. The desired quantile is obtained after numerical integration of (A1) from F_0 to p.

The analytical solution for the quantile equation [22] is an alternative way of computing a given quantile. In the general case, for a given value of CDF $= p$, this solution is

$$X_p = X_0 + \frac{p^\lambda}{\alpha\lambda}\left(1 + \frac{\lambda p^\gamma}{\gamma}\Phi[p^\gamma, 1, 1+\lambda/\gamma]\right) - \frac{F_0^\lambda}{\alpha\lambda}\left(1 + \frac{\lambda F_0^\gamma}{\gamma}\Phi[F_0^\gamma, 1, 1+\lambda/\gamma]\right) \tag{A2}$$

with $\gamma = h - g$ and $\lambda = 1 - g$. Φ is the Lerch's trascendent function [30], defined as

$$\Phi[z, s, v] = \sum_{n=0}^{\infty} \frac{z^n}{(v+n)^s}, \quad |z| < 1, \quad v \neq 0, -1, -2, \ldots \tag{A3}$$

The quantile solution indicated in (A2) can take particular forms for some special cases in the parametric space. These special cases are discussed elsewhere [22]. In practice, (A2) applies to most relevant cases. Using the analytical quantile solution (A2) it was shown that an *S*-distribution with $g < 1$ has no infinite left tail. That is, the equation $F(X) = 0$ has a finite solution. This property is very important for a correct computation of ROC curves using *S*-distributions (see below). For practical purposes, equation (A2) will be used only to compute values of zero-quantiles in those cases in which $g < 1$. These zero-quantiles are required for a correct computation of initial conditions for ROC curve computation. This result is easily obtained using *Mathematica* or equivalent programs [22]. Equation (A1) will be used for computing the required quantiles once the initial conditions are obtained.

A2. Computation of initial conditions

A critical step for correctly computing the ROC curve using the *S*-distribution method is to determine appropriate starting points for integrating (10). We shall discuss this problem in some detail. In practice, the values of g_D and g_H are critical for an appropriate selection, since an *S*-distribution with $g < 1$ has no infinite left tail [22]. If for a given probability p, we denote the quantile solution as $F^{-1}(p) = X_p$, then $g < 1$ requires that the equation $F^{-1}(0) = X_p$ has a solution different from $-\infty$. Considering this result, different cases must be taken into account:

1. *Case I:* $g_D < 1$, $g_H < 1$. In this case, we have two zero-quantile values $F_D^{-1}(0) = X_D(0)$ and $F_H^{-1}(0) = X_H(0)$. Then:
   (a) if $X_D(0) \leqslant X_H(0)$ (Figure A1(a), the starting values are $F_H = 0$ and $F_D(0) = F_D^{-1}(X_H(0))$ (case I(a));
   (b) if $X_D(0) > X_H(0)$ (Figure A1(b), we must first compute $p_s = F_H(X_D(0))$. Then the starting values are $F_H = p_s$ and $F_D(p_s) = X_D(0)$ (case I(b)).
2. *Case II:* $g_D < 1$, $g_H \geqslant 1$. Since in this case a solution exists for $F_D^{-1}(0) = X_D(0)$ (Figure A1(c), the starting values are computed as in case I(b).
3. *Case III:* $g_D \geqslant 1$, $g_H < 1$. In this case (Figure A1(d), the starting values are computed as in case I(a).
4. *Case IV:* $g_D \geqslant 1$. $g_H \geqslant 1$. In this case (Figure A1(e)), both distributions have an infinite left tail. Since no zero-quantile can be obtained, the solution must be computed by selecting a sufficiently low value of p_s and computing the starting values as in case I(a). If the left tail is more heavy for the healthy group (Figure A1(f)), it may be more convenient to select the start point as $p_s = F(F_D^{-1}(p_\varepsilon))$, where p_ε is a sufficiently small value. Then the starting values are computed as in case I(b).

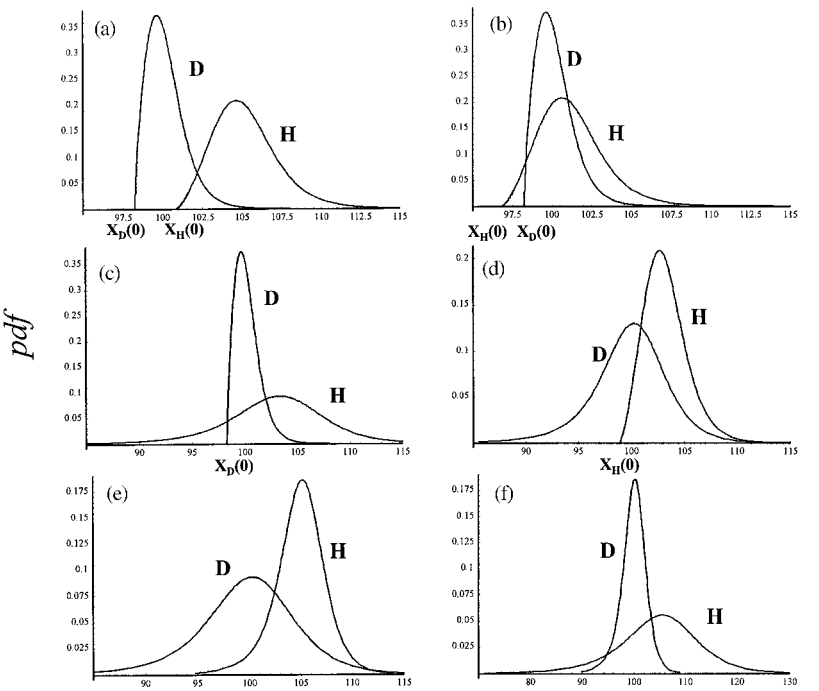

Figure A1. Initial values for ROC curve computation using S-distributions. The choice of an appropriate initial value for computing the corresponding ROC curve is a critical step. This choice depends on the value of g_D and g_H and on the relative situation of the hypothetical zero-quantiles: (a) case I: $g_D < 1$, $g_H < 1$ and $X_D(0) < X_H(0)$; (b) case I: $g_D < 1$, $g_H < 1$ and $X_D(0) > X_H(0)$; (c) case II: $g_D < 1$, $g_H \geqslant 1$ and $X_D(0) > X_H(0)$; (d) case III: $g_D \geqslant 1$ and $g_H < 1$; (e) case IV: $g_D \geqslant 1$ and $g_H \geqslant 1$; (f) case IV: $g_D \geqslant 1$ and $g_H \geqslant 1$. See text for details on the initial values.

As illustration, different examples of ROC curves computed using this strategy are shown in Figure 4. For practical purposes, all these computations can be performed automatically. At this stage, we have defined a package in *Mathematica*. In the future, we shall develop a Visual C++ program for computing ROC curves using this approach.

ACKNOWLEDGEMENTS

We thank Dr E.O.Voit for his many suggestions leading to an improvement in this work and Dr B. Hernández-Bermejo for useful comments and a critical review of this paper. We also thank two anonymous referees for positive criticism. The authors acknowledge financial support from the FIS of Spain (grant 00/0235) and a grant from La Paeria (Lleida) (X0148).

REFERENCES

1. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry* 1993; **39**:561–577.
2. Campbell G. General methodology I: Advances in statistical methodology for the evaluation of diagnostic and laboratory tests. *Statistics in Medicine* 1994; **13**:499–508.

3. Zou KH, Tempany CM, Fielding JR, Silverman SG. Original smooth receiver operating characteristic curve estimation from continuous data: statistical methods for analyzing the predictive value of espiral CT of ureteral stones. *Academic Radiology* 1998; **5**:680–687.
4. Hanley JA. The use of the binormal model for parametric ROC analysis of quantitative diagnostic tests. *Statistics in Medicine* 1996; **15**:1575–1585.
5. Metz CE, Herman BA, Shen JH. Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. *Statistics in Medicine* 1998; **17**:1033–1053.
6. Hajian-Tilaki KO, Hanley JA, Joseph L, Collet JP. A comparison of parametric and nonparametric approaches to ROC analysis of quantitative diagnostic tests. *Medical Decision Making* 1997; **17**:94–102.
7. Walsh SJ. Limitations to the robustness of binormal ROC curves: effects of model misspecification and location of decision thresholds on bias, precision, size and power. *Statistics in Medicine* 1997; **16**:669–679.
8. Tosteson AN, Begg CB. A general regression methodology for ROC curve estimation. *Medical Decision Making* 1988; **8**:207–215.
9. Sorribas A, March J, Voit EO. Estimating age-related trends in cross-sectional studies using *S*-distributions. *Statistics in Medicine* 2000; **19**:697–713.
10. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic ROC curve. *Radiology* 1982; **143**:29–36.
11. Hanley JA, McNeil BJ. A method of comparing areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983; **148**:839–843.
12. Zou KH, Hall WJ, Shapiro DE. Smooth non-parametric receiver operating characteristic (ROC) curves for continuous diagnostic tests. *Statistics in Medicine* 1997; **16**:2143–2156.
13. Metz CE. ROC methodology in radiology imaging. *Investigation in Radiology* 1986; **21**:720–723.
14. Hanley JA. The robustness of the 'Binormal' assumptions used in firring roc curves. *Medical Decision Making* 1988; **8**:197–203.
15. Johnson NL, Kotz S. *Continuous Univariate Distributions-1.* Houghton Mifflin: Boston MA, 1970.
16. Savageau MA. A suprasystem of probability distributions. *Biometrical Journal* 1982; **24**:323–330.
17. Voit EO, Rust PF. Invited tutorial: *S*-system analysis of continuous univariate probability distributions. *Journal of Statistical Computing and Simulation* 1992; **42**:187–249.
18. Voit EO. The *S*-distribution: a tool for approximation and classification of univariate, unimodal probability distributions. *Biometrical Journal* 1992; **7**:855–878.
19. Balthis WL, Voit EO, Meaburn GM. Setting prediction limits for mercury concentrations in fish having high bioaccumulation potential. *Environmetrics* 1996; **7**:429–439.
20. Voit EO, Balthis WL, Holser RA. Hierarchical Monte Carlo modeling with *S*-distributions: concepts and illustrative analysis of mercury contamination in king mackerel. *Environment International* 1995; **21**:627–635.
21. Voit EO, Yu S. The *S*-distribution: approximation of discrete distributions. *Biometrical Journal* 1994; **36**:205–219.
22. Hernández-Bermejo B, Sorribas A. Analytical quantile solution for the *S*-distribution: application to distribution design and data fitting. *Biometrical Journal* 2001; **43**:1007–1025.
23. Voit EO. A maximum likelihood estimator for shape parameters of *S*-distributions. *Biometrical Journal* 2000; **42**:471–479.
24. Ferreira A. PLAS: Power-Law modeling and simulation. At http://correio.cc.fc.ul.pt/aenf/plas.html.
25. Ma G, Hall WJ. Confidence bands for receiver operating characteristic curves. *Medical Decision Making* 1993; **13**:191–197.
26. Moosman D. Resampling techniques in the analysis of non-binormal ROC data. *Medical Decision Making* 1995; **15**:358–366.
27. Platt RW, Hanley JA, Yang H. Bootstrap confidence intervals for the sensitivity of a quantitative diagnostic test. *Statistics in Medicine* 2000; **19**:313–322.
28. Schafer H. Efficient confidence bounds for ROC curves. *Statistics in Medicine* 1994; **13**:1551–1561.
29. Jensen K, Muller HH, Schafer H. Regional confidence bands for ROC curves. *Statistics in Medicine* 2000; **19**:493–509.
30. Erdélyi A, Magnus W, Oberhettinger F, Tricomi FG. *Higher Transcendental Functions*, Vol. I. McGraw-Hill: New York, 1953.

