

 Open access • Journal Article • DOI:10.1093/BIOMET/78.3.637

## **Bayesian residual analysis in the presence of censoring** — [Source link](#)

Kathryn Chaloner

**Institutions:** University of Minnesota

**Published on:** 01 Sep 1991 - Biometrika (Oxford University Press)

**Topics:** Multivariate normal distribution, Posterior probability, Sampling distribution, Exponential distribution and Outlier

Related papers:

- [A Bayesian approach to outlier detection and residual analysis](#)
- [Bayesian Data Analysis](#)
- [Bayesian analysis of binary and polychotomous response data](#)
- [Model Determination Using Predictive Distributions with Implementation via Sampling-Based Methods](#)
- [Bayesian residual analysis for binary response regression models](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/bayesian-residual-analysis-in-the-presence-of-censoring-ynu2x7871q>

MKC  
9/29/91

**Bayesian residual analysis in the presence of censoring**

By

Kathryn Chaloner

*University of Minnesota*

Technical Report No. 556

School of Statistics

University of Minnesota

December 19, 1990



UNIVERSITY OF MINNESOTA

SCHOOL OF STATISTICS

**Bayesian residual analysis in the presence of censoring**

By

Kathryn Chaloner

*University of Minnesota*

Technical Report No. 556

School of Statistics

University of Minnesota

December 19, 1990

# Bayesian residual analysis in the presence of censoring

KATHRYN CHALONER

*University of Minnesota, St. Paul MN 55108, U.S.A.*

December 19, 1990

## SUMMARY

Using the posterior distribution of the realized error terms for residual analysis in a linear model was advocated in Zellner (1975) and Zellner & Moulton (1985). The same idea was used in Chaloner & Brant (1988) to define outliers and calculate posterior probabilities of observations being outliers. This paper extends the same concept to other models, including regression models for lifetime data, by using an approach similar to that of Cox and Snell (1968). Residual plots are proposed where realized errors are represented by interval estimates. Incorporating censored observations into this framework is straightforward.

*Some key words: Laplace approximation; Lifetime data; Posterior distribution.*

## 1 THE REALIZED ERROR TERMS

Cox & Snell (1968) proposed a general definition of residuals for models where each observation  $y_i$  can be written as  $y_i = g_i(\theta, \epsilon_i)$  with  $\theta$  a vector of unknown parameters and the  $\epsilon_i, i = 1, \dots, n$  a sample of independent and identically distributed random variables from a known distribution. Suppose that the equation  $y_i = g_i(\theta, \epsilon_i)$  has a

unique solution  $\epsilon_i = h_i(y_i, \theta)$ . Cox & Snell define  $\tilde{\epsilon}_i = h_i(y_i, \hat{\theta})$  to be the residuals where  $\hat{\theta}$  are the maximum likelihood estimates of  $\theta$ . In the Bayesian approach used here define  $\epsilon_i = h_i(y_i, \theta), i = 1, \dots, n$  to be the residuals. Each  $\epsilon_i$  is just a function of the unknown parameters and the posterior distribution is therefore straightforward to calculate. The posterior distribution can be examined for indications of possible departures from the assumed model and the presence of outliers.

The posterior distribution of the realized errors is very different from the sampling distribution of their estimates. The posterior distribution represents the uncertainty about functions of the parameters; if the parameters,  $\theta$ , and observations,  $y_i$ , are known then so are the realized errors  $\epsilon_i = h_i(y_i, \theta)$ .

With a large sample size, the posterior distribution of the  $\epsilon_i$  will be, approximately, multivariate normal centred at the posterior mean and with covariance matrix the posterior covariance matrix. An alternative approximation to the posterior mean of the realized errors would be the maximum likelihood estimates  $\tilde{\epsilon}_i$ .

Following Chaloner & Brant (1988) an outlier can be defined to be any observation that has a surprising value of  $\epsilon_i$ . For lifetime data we might perhaps be especially interested in detecting any observations which are unusually large. If the cumulative distribution function of the  $\epsilon_i$  is  $F(\cdot)$  and we have  $n$  observations then we might, for example, choose

$$k = F^{-1}(0.95^{1/n}) \quad (1)$$

and define any observation with  $\epsilon_i > k$  to be an outlier. (We assume that  $g_i(\theta, \epsilon_i)$  is monotonically increasing in  $\epsilon_i$ .) In this way the prior probability of no outliers is 0.95 and outliers are not expected. After observing the data we can calculate the posterior probability that each observation is an outlier and any observation with a posterior probability,  $pr(\epsilon_i > k|y)$ , larger than the prior probability of  $1 - F(k)$  would be suspect. The approach can clearly be adapted to detecting surprisingly short lifetimes, or, alternatively, lifetimes that are either surprisingly short or surprisingly long. Other, related, calculations were demonstrated in Chaloner & Brant (1988) for linear regression models with normal realized errors.

## 2 AN EXAMPLE WITHOUT CENSORING

Consider the data from Feigl & Zelen (1965) which were also used by Cox & Snell (1968). The  $y_i$  are the weeks until death of leukemia patients and the explanatory variable  $x_i$  is the log of the initial white blood cell count divided by 10,000. Only the 17 AG positive patients will be used. The model can be written

$$y_i = e^{\beta_0 + \beta_1 x_i} \epsilon_i,$$

or, equivalently,

$$\log y_i = \beta_0 + \beta_1 x_i + \log \epsilon_i,$$

and so

$$h_i(y_i, \theta) = \frac{y_i}{e^{\beta_0 + \beta_1 x_i}}.$$

The  $\epsilon_i$  for  $i = 1, \dots, n$  are an independent sample from an exponential distribution with mean 1, or, equivalently, the  $\log \epsilon_i$  are an independent sample from a standard extreme value distribution. As the residuals are most usually examined in the untransformed scale in which they have an exponential distribution we compute their posterior mean in that scale. That is, if we denote the data by  $y$ , let

$$\hat{\epsilon}_i = E\left(\frac{y_i}{e^{\beta_0 + \beta_1 x_i}} | y\right).$$

Using the normal approximation to the posterior distribution, we could construct an approximate 95% highest posterior density region for each  $\epsilon_i$  by  $\hat{\epsilon}_i \pm 2\sqrt{\text{var}(\epsilon_i | y)}$ . As the  $\epsilon_i$  are positive quantities, however, intervals calculated this way might contain negative values, and, indeed, for this example they do contain negative values. The normal approximation is therefore used on the posterior distribution of  $\log \epsilon_i$  and the end points exponentiated. Specifically, the highest posterior density region for  $\epsilon_i$  is

$$\exp\{E(\log \epsilon_i | y) \pm 2\sqrt{\text{var}(\log \epsilon_i | y)}\}. \quad (2)$$

The resulting approximate 95% intervals contain only positive values. These intervals are denoted as  $(L_i, U_i)$ . Inspection of the posterior distribution of the  $\epsilon_i$  and  $\log \epsilon_i$  for

this example reveals that the posterior distribution of  $\log \epsilon_i$  is typically less skewed and is closer to normal than that of the  $\epsilon_i$ .

Table 1 gives the data, the posterior means and the calculated intervals under a non-informative prior distribution where  $p(\beta_0, \beta_1)$  is constant on two-dimensional Euclidean space. The posterior means and variances were calculated using the Laplace approximations of Tierney & Kadane (1986) using the software of Tierney (1990, p.70-6). The first order approximations to posterior means and variances using the maximum likelihood estimates and the second derivatives of the log likelihood function could also have been used and, in fact, give very similar results. Figure 1 is a plot of the ordered point estimates,  $\hat{\epsilon}_i$ , and their corresponding interval estimates, plotted against the expected order statistics of a sample from an exponential distribution with mean 1. A straight line through the origin, with slope 1, has been drawn in as this is what is expected under the prior distribution.

A plot of the posterior means alone does not indicate anything other than what would be expected under an exponential distribution. The highest posterior density regions, however, indicate that there is uncertainty associated with some of the the residuals. The  $\epsilon_i$  with the largest  $\hat{\epsilon}_i$  is observation 17 and the interval  $[L_{17}, U_{17}]$  is extremely large. There is great posterior uncertainty about that residual. A plot of the intervals against the explanatory variable indicates that patient 17 also has the largest initial white blood cell count. This patient survived 65 weeks, but two other patients with the same initial white blood cell count survived only 1 week. The augmented residual plot has drawn attention to this somewhat unusual feature.

Outliers can be defined using the method outlined in §1. Choosing  $k$  as in Equation 1 gives  $k = 5.8$ . The Laplace approximation to the marginal posterior distribution  $p(\epsilon_{17}|y)$  can be calculated and a plot of the distribution indicates that there is indeed a non-negligible posterior probability that  $\epsilon_{17}$  is larger than 5.8. As  $p(\log \epsilon_{17}|y)$  is closer to normal than  $p(\epsilon_{17}|y)$  we use a normal approximation to  $p(\log \epsilon_{17}|y)$  to give an approximate probability  $pr(\epsilon_{17} > 5.8|y)$ . The Laplace approximations to the posterior mean and standard deviation of  $\log \epsilon_{17}$  are 1.177 and 0.454 and so  $pr(\epsilon_{17} > 5.8|y)$  is approximately 0.1. This posterior probability is considerably larger

than the prior probability of 0.003. There is, therefore, some evidence that the lifetime of patient 17 might be in the extreme upper tail of the distribution of lifetimes at the value of  $x_{17}$  under the assumed model. Note that observation 17 is also identified as being unusually influential by Cook & Weisberg (1982, p.178-86) using likelihood displacement methods. The analysis of Cox & Snell (1968) did not identify this observation as unusual.

### 3 CENSORED OBSERVATIONS

Now assume that censoring is present. If  $y_i$  is less than or equal to  $c_i$ , the censoring time of the  $i$ th observation, we observe the lifetime  $y_i$ ; if, alternatively,  $y_i$  is greater than  $c_i$  we do not observe the lifetime  $y_i$ , we just know that  $y_i > c_i$ . Units are either therefore observed to fail at time  $y_i$  or are censored at time  $c_i$ . Right censoring of this type is the only kind of censoring discussed here, although similar ideas can be applied to left censoring.

The posterior distribution of  $\theta$  is usually straightforward to calculate numerically. For the uncensored observations the posterior mean and variance of  $h_i(y_i, \theta)$  can be calculated to give point and interval estimates for the realized errors  $\epsilon_i$ , as in §1 and §2. For the censored observations, however, we only know that  $h_i(y_i, \theta) > h_i(c_i, \theta)$  and even if the value of  $\theta$  is known exactly we will always have a distribution on  $h_i(y_i, \theta)$  reflecting uncertainty in the unobserved  $y_i$ .

Let the prior distribution of  $\epsilon_i$  have cumulative distribution function  $F(\cdot)$ , completely specified, and corresponding probability density function  $f(\cdot)$ . Denote all the data,  $y_i$ 's and  $c_i$ 's, by  $y$ , then the posterior distribution of  $\epsilon_i$  conditional on the parameters  $\theta$  and  $y$  is given by

$$p(\epsilon_i | \theta, y) = \begin{cases} f(\epsilon_i)[1 - F\{h_i(c_i, \theta)\}]^{-1} & \text{if } \epsilon_i > h_i(c_i, \theta) \\ 0 & \text{otherwise} \end{cases}$$

which is just the prior distribution truncated at  $h_i(c_i, \theta)$ . The posterior density function of  $\epsilon_i$  is most easily expressed as the following where  $I(\cdot)$  denotes the indicator



function and the integral is over the whole parameter space

$$p(\epsilon_i|y) = \int \frac{f(\epsilon_i)I\{\epsilon_i > h_i(c_i, \theta)\}}{1 - F\{h_i(c_i, \theta)\}} p(\theta|y) d\theta.$$

The posterior mean of  $\epsilon_i$  is the posterior mean of the function  $M(c_i, \theta)$  where

$$M(c_i, \theta) = \int_{h_i(c_i, \theta)}^{\infty} \frac{zf(z)}{1 - F\{h_i(c_i, \theta)\}} dz. \quad (3)$$

This expression often simplifies: for example if  $F(\cdot)$  corresponds to a unit exponential distribution then  $M(c_i, \theta) = h(c_i, \theta) + 1$ . The posterior mean  $\hat{\epsilon}_i = E\{M(c_i, \theta)|y\}$  is a point estimate of  $\epsilon_i$  which can be plotted on a residual plot. (This parallels the suggestion of Lawless, 1982 p. 281 and 320, to use  $M(c_i, \hat{\theta})$  as estimates of residuals for censored data.)

The posterior mean of  $\epsilon_i$  for censored observations is therefore usually easy to calculate. This single number, however, does not adequately reflect the inherent uncertainty due to censoring. Even if  $\theta$  were known exactly we would have a distribution, not a single value, for  $\epsilon_i$ . Suppose  $\theta$  were known, then the value  $h_i(c_i, \theta)$  provides a lower bound and we could choose  $A_i(c_i, \theta)$  such that there would be probability 0.95 that  $\epsilon_i$ , given  $\theta$ , was in  $[h_i(c_i, \theta), A_i(c_i, \theta)]$ . That is  $A_i(c_i, \theta)$  is the 95th percentile of the distribution of  $p(\epsilon_i|\theta, y)$ , specifically:

$$A_i(c_i, \theta) = F^{-1}[0.95 + 0.05F\{h_i(c_i, \theta)\}]. \quad (4)$$

Define the interval  $[\hat{h}_i, \hat{A}_i]$  where  $\hat{h}_i = E\{h(c_i, \theta)|y\}$  and  $\hat{A}_i = E\{A_i(c_i, \theta)\}$ . This interval has been found to be a useful posterior estimate of  $\epsilon_i$ .

In addition to  $[\hat{h}_i, \hat{A}_i]$  a conservative 95% highest posterior density region for  $\epsilon_i$  is also useful. To calculate this interval,  $[L_i, U_i]$ , for a censored observation some further approximations are proposed. First we suggest a method for a lower bound,  $L_i$ . If the posterior distribution of  $h_i(c_i, \theta)$  is approximately normal and we take

$$L_i = E\{h_i(c_i, \theta)|y\} - 2\sqrt{\text{var}\{h_i(c_i, \theta)|y\}} \quad (5)$$

then  $pr\{L_i < h_i(c_i, \theta)|y\}$  is approximately 0.025. As for any value of  $\theta$  we must have  $\epsilon_i > h_i(c_i, \theta)$ , the set of values of  $\theta$  for which  $\epsilon_i < L_i$  must have posterior probability

less than 0.025 and so  $pr\{\epsilon_i < L_i|y\} < 0.025$  and  $L_i$  provides a conservative lower bound for an approximate 95% highest posterior density region. Alternatively, if the posterior distribution of  $\log h_i(c_i, \theta)$  is closer to having a normal posterior distribution we might use

$$L_i = \exp[E\{\log h(c_i, \theta)|y\} - 2\sqrt{\text{var}\{\log h(c_i, \theta)|y\}}]. \quad (6)$$

For the upper end point,  $U_i$ , of the highest posterior density region note that

$$pr(\epsilon_i > u|y) \leq \{1 - F(u)\} \int [1 - F\{h(c_i, \theta)\}]^{-1} p(\theta|y) d\theta.$$

If  $U_i$  is chosen to be

$$U_i = F^{-1}\left(1 - \frac{0.025}{E\{[1 - F\{h_i(c_i, \theta)\}]^{-1}|y\}}\right) \quad (7)$$

we have  $pr(\epsilon_i > U_i|y) \leq 0.025$  and  $U_i$  is a conservative upper bound for a 95% highest posterior density interval.

Both intervals,  $[\hat{h}_i, \hat{A}_i]$  and  $[L_i, U_i]$ , are useful to plot as will be demonstrated with the example to follow. The interval  $[\hat{h}_i, \hat{A}_i]$  is based on the conditional distribution given  $\theta$  and therefore can be thought of as representing an estimate of the uncertainty in  $\epsilon_i$  due to censoring. The interval  $[L_i, U_i]$  is based on the unconditional posterior distribution and is typically a larger interval incorporating all the posterior uncertainty.

## 4 AN EXAMPLE WITH CENSORING

Now consider a data set with censoring. The data set is originally from Glasser (1965) and is reproduced in Lawless (1982, p. 319). The response  $y_i$  is the survival time in days of lung cancer patients. There are two explanatory variables, age ( $x_1$ , in years with the mean age subtracted) and performance status ( $x_2$ ). There are 16 observations, 6 of which are censored. The data set is given in Table 2.

The lifetimes  $y_i$  are taken to have a log-normal distribution and we let

$$\log y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \sigma \epsilon_i$$

and

$$h_i(y_i, \theta) = \sigma^{-1}(\log y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i}).$$

The  $\epsilon_i$ 's are independent and have a standard normal distribution. It seems more natural in this case to work with realized errors which are a sample from a normal distribution rather than a sample from a log-normal distribution as normal residual plots are so familiar. The prior cumulative distribution function of the  $\epsilon_i$  is just the standard normal distribution function,  $\Phi(\cdot)$ .

For the uncensored observations we use the method described in §1 taking  $\hat{\epsilon}_i = E\{h_i(y_i, \theta)\}$  and the highest posterior density region to be  $\hat{\epsilon}_i \pm 2\sqrt{\text{var}\{h_i(y_i, \theta)|y\}}$ . For the censored observations Equation 3 simplifies to give  $\hat{\epsilon}_i$  as the posterior mean of

$$M(c_i, \theta) = \frac{\exp[-\{h_i(c_i, \theta)\}^2/2]}{(2\pi)^{1/2}[1 - \Phi\{h_i(c_i, \theta)\}]}$$

Equation 7 with  $F(\cdot) = \Phi(\cdot)$  gives  $U_i$ , and  $L_i$  is straightforward to calculate using Equation 5. Note also that  $\hat{h}_i = E\{h_i(c_i, \theta)|y\}$  and  $\hat{A}_i$  is the posterior expectation of  $A_i(c_i, \theta)$  given in Equation 4. A non-informative prior distribution is again used, taking  $p(\beta_0, \beta_1, \beta_2, \log\sigma)$  proportional to a constant.

The quantities to be calculated require taking posterior expectations of functions of  $\theta$ . In this example, as in §1, we use Laplace approximations to the posterior expectations, but first order approximations, using the maximum likelihood estimates, could easily have been used instead to give similar results.

The Laplace approximations to the posterior means of  $(\beta_0, \beta_1, \beta_2, \log\sigma)$  are (1.614, -0.0060, 0.1021, -1.033). The  $\hat{\epsilon}_i$ ,  $L_i$  and  $U_i$ , calculated using Laplace approximations to posterior means and variances, are given in Table 2 together with values of  $\hat{h}_i$  and  $\hat{A}_i$  for the censored observations.

Figure 2 is a plot of the residuals and their corresponding interval estimates plotted against case number. A dashed line is used for the intervals  $[\hat{h}_i, \hat{A}_i]$  for the censored observations and solid lines for the remaining part of the intervals  $[L_i, U_i]$ . The posterior means are shown on the plot using a different symbol for the censored and uncensored observations.

The dashed lines represent the inherent uncertainty, due to censoring because  $y_i$  is not observed. The dashed lines are an estimate of an interval, with 95% posterior probability, if  $\theta$  were known. The solid lines, for the remaining part of  $[L_i, U_i]$  represent additional uncertainty due to uncertainty about  $\theta$ . If more observations were taken from the model the uncertainty in  $\theta$  would decrease and the solid part of the interval  $[L_i, U_i]$  become shorter but the dashed lines would have, approximately, constant length.

The plot does not indicate any overwhelming violations of the normality assumption but it does point to two censored observations having interesting features. One is observation 14 which has the largest posterior mean ( $\hat{\epsilon}_{14} = 1.78$ ) and some posterior probability of being quite large. Using the guideline given in §1 and Equation 1 we might take an outlier to be any observation with  $\epsilon_i > 2.73$ . The 95% highest posterior density region for  $\epsilon_{14}$  includes the value 2.73 and so observation 14 may have a small, but non-negligible, probability of being an outlier. The second interesting feature of the plot is the great uncertainty in the residual for observation 12, as the 95% interval is very large. Although there is no indication that  $\epsilon_{12}$  is greater than 2.73, there is some posterior probability that it might be unusually small. Inspection of the data reveals that both observation 12 and 14 have values of the explanatory variables which are extremes. Observation 12 has the largest value of performance status ( $x_2$ ) and observation 14 the smallest. In addition they both have the second largest value of age, ( $x_1$ ). These extreme values will contribute to the large posterior variance of the residual. These two observations are not necessarily outliers but the residual plot indicates that they deserve closer inspection. On closer inspection it is interesting that if observation 14 is deleted the posterior mean of  $\beta_1$  changes from -0.006 to -0.019 and the posterior mean of  $\beta_2$  changes from 0.102 to 0.176. Both these changes are approximately one standard error in magnitude, and, for  $\beta_2$  this change makes the coefficient change from being not significantly different from zero to being significantly different from zero at the usual levels. Deleting observation 12 does not change the estimates very much further. Neither observation 12 or 14 would be noticed in a standard residual plot but the Bayesian residual plot of Figure 2 highlights

them for further inspection.

## Acknowledgement

I am grateful to D. Hawkins, L. Tierney, a referee and the editor for helpful comments. This research was partly supported by a National Science Foundation Grant and by the Deutsche Forschungsgemeinschaft. An earlier version of this paper appeared as report number 176 of the Schwerpunktprogramm der Deutschen Forschungsgemeinschaft: Anwendungsbezogene Optimierung und Steuerung, Universität Augsburg.

## REFERENCES

- CHALONER, K. & BRANT, R. (1988). A Bayesian approach to outlier detection and residual analysis. *Biometrika* 75, 651-9.
- COOK, R. D. & WEISBERG, S. (1982). *Residuals and Influence in Regression*. London: Chapman and Hall.
- COX, D.R. & SNELL, E.J. (1968). A general definition of residuals (with discussion). *J. Roy. Statist. Soc. Ser. B*, 30, 248-275.
- FEIGL, P. & ZELEN, M. (1965). Estimation of exponential probabilities with concomitant information. *Biometrics*, 21, 826-38.
- GLASSER, M. (1965). Regression analysis with dependent variable censored. *Biometrics*, 21, 300-307.
- LAWLESS, J.F. (1982). *Statistical Models and Methods for Lifetime Data*. New York: Wiley.
- NELSON, W. (1982). *Applied Life Data Analysis*. New York: Wiley.
- PRENTICE, R.L. (1973). Exponential survival with censoring and explanatory variables. *Biometrika* 60, 279-288.
- TIERNEY, L. & KADANE, J.B. (1986). Accurate approximations for posterior moments and marginal densities. *J. Am. Statist. Assoc.* 81, 82-86.
- TIERNEY, L. (1990). *LISP-STAT, an object-oriented environment for statistical computing and dynamic graphics*. New York: Wiley.
- ZELLNER, A. (1975). Bayesian analysis of regression error terms. *J. Am. Statist. Assoc.* 70, 138-44.
- ZELLNER, A. & MOULTON, B.R. (1985). Bayesian regression diagnostics with applications to international consumption and income data. *J. Econometrics* 29, 187-211.

Table 1. *Table of  $y_i$ ,  $x_i$ , posterior mean  $\hat{\epsilon}_i$ , and lower and upper bounds for a 95% highest posterior density region,  $L_i$  and  $U_i$ , for Feigl & Zelen data.*

$i$	$y_i$	$x_i$	$\hat{\epsilon}_i$	$L_i$	$U_i$
1	65	-1.470	0.576	0.243	1.176
2	156	-2.590	0.874	0.241	2.332
3	100	-0.844	1.163	0.591	2.082
4	134	-1.347	1.251	0.549	2.486
5	16	-0.511	0.216	0.118	0.366
6	108	0.049	1.890	1.110	3.034
7	121	0	2.069	1.212	3.329
8	4	0.531	0.088	0.051	0.143
9	39	-0.616	0.502	0.269	0.864
10	143	-0.357	2.072	1.166	3.436
11	56	-0.062	0.930	0.542	1.501
12	26	1.163	0.786	0.410	1.380
13	22	1.253	0.696	0.356	1.243
14	1	2.303	0.055	0.020	0.124
15	1	2.303	0.055	0.020	0.124
16	5	1.649	0.194	0.089	0.376
17	65	2.303	3.582	1.309	8.043

Table 2. Table of  $\log y_i$ ,  $\log c_i$ , age,  $x_{2i}$ , posterior mean  $\hat{\epsilon}_i$ , and lower and upper bounds for a 95% highest posterior density region,  $L_i$  and  $U_i$ , with  $\hat{h}_i$  and  $\hat{A}_i$  for the censored observations, for lung cancer patients under the log-normal model.

$i$	$\log y_i$	$\log c_i$	age	$x_{2i}$	$\hat{\epsilon}_i$	$L_i$	$U_i$	$\hat{h}_i$	$\hat{A}_i$
1	1.94		42	4	-0.368	-1.720	0.984		
2	2.23		67	6	0.072	-0.931	1.075		
3	1.94		62	4	-0.087	-1.090	0.916		
4	1.98		52	6	-0.699	-1.377	-0.020		
5	2.23		57	5	0.212	-0.461	0.885		
6	1.59		58	6	-1.488	-2.363	-0.613		
7	2.13		55	6	-0.320	-0.882	0.241		
8	1.80		63	7	-1.228	-2.117	-0.339		
9	2.32		44	5	0.231	-0.910	1.373		
10	1.92		62	7	-0.973	-1.764	-0.182		
11		2.15	51	7	0.4745	-1.412	2.113	-0.612	1.809
12		2.05	64	10	0.213	-3.005	2.017	-1.495	1.708
13		2.48	54	8	0.748	-1.106	2.266	-0.112	1.945
14		2.42	64	3	1.787	-0.158	3.129	1.297	2.622
15		2.56	54	9	0.702	-1.489	2.271	-0.213	1.930
16		2.56	57	9	0.724	-1.413	2.282	-0.171	1.940



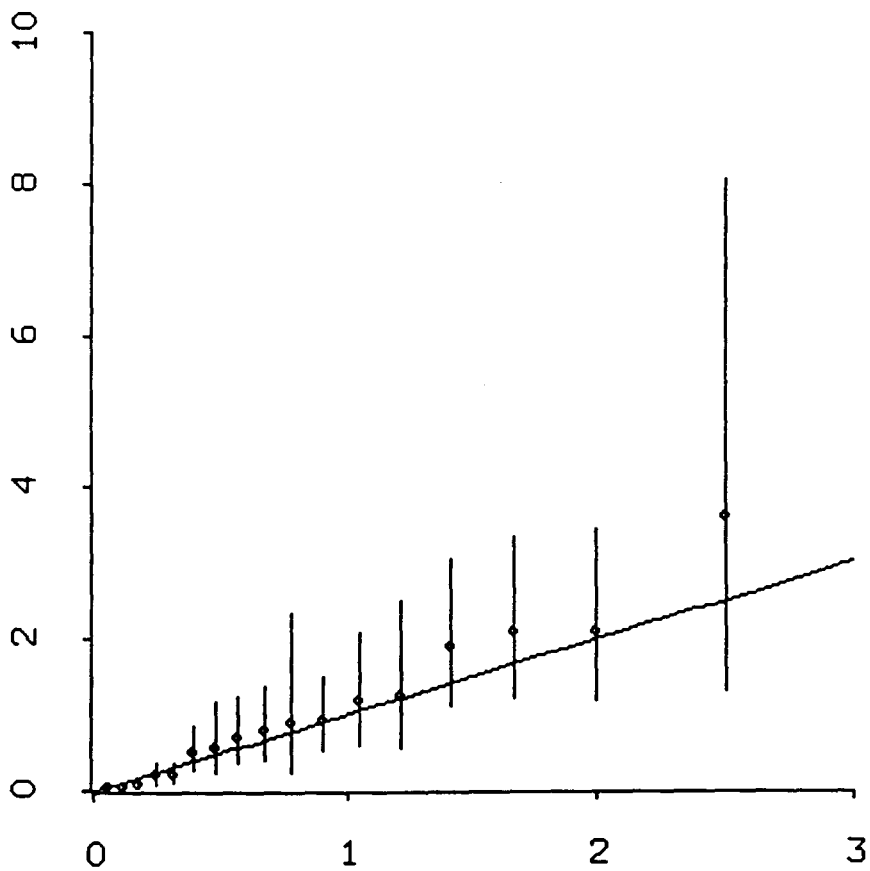


Fig. 1. Plot of estimated residuals and their highest posterior density regions for the Feigl and Zelen data, plotted against the expected order statistics of a standard exponential distribution.

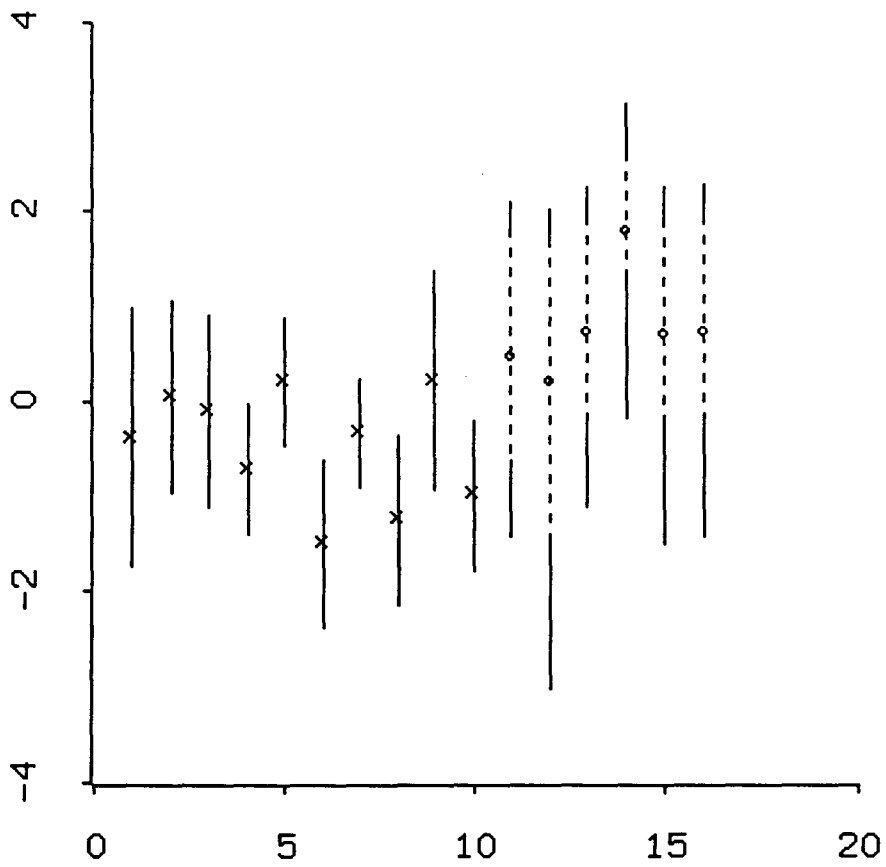


Fig. 2. Plot of the estimated residuals and their highest posterior density regions for the data from Glasser (1965), (x mean for uncensored residual, o mean for censored residual).