TRANSFORMATION OF LIFE-TEST DATA

Norman R. Draper and Irwin Guttman

(received February 29, 1968)

In a recent paper Box and Cox (1964) considered the problem of transforming dependent variables in regression and analysis of variance problems, in order to achieve the usual assumptions of Normality, constant variance and additivity of effects. Here we adopt the same approach to investigate transformations of data which allow the transformed observations to follow a Gamma distribution. A special case of this is the exponential distribution, valuable in life-testing, for which examples are given.

1. The Problem. We shall investigate the following problem. Suppose we have a sample $y_1, y_2, \ldots, y_n$ of observations which it would be convenient to treat as belonging to a Gamma distribution (of order m, known) with unknown parameter θ to be estimated, but we fear that such an assumption may not be valid. Can we find a transformation which can be applied to the y_i which will allow the assumption of a Gamma distribution to be validly applied to the transformed observations? (We shall see that this problem has specific applications, e.g. in the life-testing field.) We shall consider transformations of the type

$$(1.1) \qquad z_i = y_i^{\lambda}, \quad i = 1, 2, \ldots, n, \quad \lambda > 0.$$

(We shall not consider negative values of λ for the following reason. A negative power transformation implies a positive power transformation of the reciprocals of the observations. However taking reciprocals has the effect of merely re-ordering positive observations so that the largest original observation becomes the smallest reciprocal observation and so on,thus reducing to the same formulation as we adopt. Thus any transformation problem involving life-test data can be handled by using positive values of λ only.) We shall estimate an appropriate value for λ in the following way. If the transformed data $z_i = y_i^{\lambda}$, $i = 1, 2, \ldots, n$ follow the Gamma distribution (of order m, known) parameter θ, then the probability distribution function for z_i is

https://doi.org/10.4153/CMB-1968-057-3 Published online by Cambridge University Press

(1.2) $f(z_i) = f(y_i^\lambda) = \{z_i^{m-1} e^{-z_i/\theta}\} / \{\theta^m \Gamma(m)\}, \; z_i > 0, \quad \theta > 0, \; m > 0.$

This implies that the likelihood function for the <u>original</u> y_i is given by

$$(1.3) \qquad |J| \, \theta^{-m} \{\Gamma(m)\}^{-n} \{ \prod_{i=1}^{n} y_i^{\lambda(m-1)} \} \exp \{ - \sum_{i=1}^{n} y_i^\lambda / \theta \}$$

where the Jacobian $|J| = |\partial(y_1^\lambda, y_2^\lambda, \ldots, y_n^\lambda)/\partial(y_1, y_2, \ldots, y_n)|$ is the one necessary to transform from one set of variables to the other. When the family (1.1) is considered this Jacobian has value

$$(1.4) \qquad \lambda^n \prod_{i=1}^{n} y_i^{\lambda-1}$$

as is easily verified. From this point we shall follow the twin paths of likelihood and Bayes, to obtain an estimate for λ, as did Box and Cox (1964).

2. <u>Likelihood approach</u>. Taking the logarithm to the base e of both sides of equation (1.3) with the quantity (1.4) inserted gives the log likelihood

$$(2.1) \qquad L(\lambda, \theta) = n \ln \lambda + (\lambda - 1) \sum_{i=1}^{n} \ln y_i - mn \ln \theta$$

$$- n \ln \Gamma(m) + \lambda(m-1) \sum_{i=1}^{n} \ln y_i - \sum_{i=1}^{n} y_i^\lambda / \theta$$

$$= n \ln \lambda + (\lambda m - 1) \sum_{i=1}^{n} \ln y_i - mn \ln \theta$$

$$- n \ln \Gamma(m) - \sum_{i=1}^{n} y_i^\lambda / \theta .$$

The maximum likelihood estimate $\hat{\theta}$ of θ for a given λ is given by setting $\partial L / \partial \theta = 0$, whence

$$(2.2) \qquad \hat{\theta} = \sum_{i=1}^{n} y_i^\lambda / mn .$$

Thus we can write

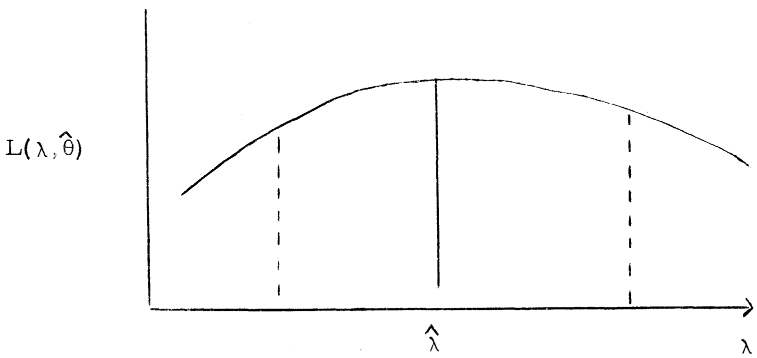
https://doi.org/10.4153/CMB-1968-057-3 Published online by Cambridge University Press

(2.3) $$L(\lambda, \hat{\theta}) = n(m \ln mn - m) + n \ln \lambda$$

$$+ (\lambda m - 1) \sum_{i=1}^{n} \ln y_i - mn \ln \sum_{i=1}^{n} y_i^{\lambda} - n \ln \Gamma(m).$$

The function (2.3) can now be plotted against λ and the maximizing value $\hat{\lambda}$ estimated from the plot (see Figure 1). This value would be used in subsequent analyses to transform the data.

Figure 1. Plot of $L(\lambda, \hat{\theta})$



A confidence interval for λ can be obtained from the equation

(2.4) $$L(\hat{\lambda}, \underline{\hat{\theta}}) - L(\lambda, \hat{\theta}) \leq \frac{1}{2} \chi_1^2 (1 - \alpha)$$

exactly as in Box and Cox (1964). Note that (2.4) is conditional on $\theta = \hat{\theta}$.

3. Bayesian approach. We shall assume that the prior distribution for θ is unaffected by the value of λ; in other words, prior feelings about probable values of θ are insensitive to possible transformations which might be indicated. (Situations might occur in which this would not hold, however.) The parameter θ has a range 0 to ∞ and the character of a variance. Thus it seems reasonable to assume that log θ is locally uniform. If we write $p_o(\lambda)$ for the prior distribution of λ, then the joint prior of λ and θ is

(3.1) $$\theta^{-1} p_o(\lambda) \, d\theta d\lambda$$

Multiplying this by the likelihood (1.3) provides the posterior distribution of λ and θ as

https://doi.org/10.4153/CMB-1968-057-3 Published online by Cambridge University Press

(3.2)

$$p(\lambda,\theta \mid \underline{y}, m) \propto \{\lambda^n \prod_{i=1}^{n} y_i^{m\lambda-1}\} \theta^{-mn-1} \{\Gamma(m)\}^{-n} \exp\{-\sum_{i=1}^{n} y_i^{\lambda} / \theta\} p_o(\lambda).$$

Integrating out for θ gives the marginal posterior for λ

(3.3)
$$p(\lambda \mid \underline{y}, m) \propto \frac{\lambda^n \prod_{i=1}^{n} y_i^{m\lambda-1}}{\{\sum_{i=1}^{n} y_i^{\lambda}\}^{mn}} \cdot p_o(\lambda).$$

Whatever prior ideas about λ are available can be inserted in (3.3). The distribution function (3.3) can then be plotted and λ can be selected as the mode (which we prefer), or the mean of the distribution (preferred by others). This value of λ can then be used to transform the data for subsequent analysis.

The Bayesian approach is identical to the likelihood approach when $p_o(\lambda)$ = constant, and the mode estimate is used. When $p_o(\lambda)$ is not constant, however, a different estimate of λ is obtained. Since $\lambda > 0$, a constant prior does not appear to be reasonable, but in any case it is clear that the Bayesian approach provides a much more flexible solution into which whatever prior ideas are available can be incorporated.

What sort of form should $p_o(\lambda)$ take? The most sensible prior appears to be

(3.4)
$$p_o(\lambda) = 1/\lambda.$$

This prior is reasonable in situations where $\lambda = 1$ is regarded as a central value and where values $\lambda = \lambda_o$ and $\lambda = 1/\lambda_o$ appear equally possible a priori. This implies for example that

$$P(1/\lambda_o \leq \lambda \leq 1) = P(1 \leq \lambda \leq \lambda_o).$$

To construct the posterior distribution in normalised form, the constant of proportionality in (3.3) is needed. In some practical cases, for example when n and/or the y_i are large, computing difficulties arise from (3.3) because numbers too small, or too large, for the computer are obtained. To help avoid this, we can rewrite (3.3) in alternative forms. One example of this is

(3.5) $$p(\lambda|\underline{y}, m) \propto \frac{\lambda^n \prod_{i=1}^{n} w_i^{m\lambda-1}}{\{\sum_{i=1}^{n} w_i^{\lambda}\}^{mn}} p_o(\lambda)$$

where $w_i = ny_i/(y_1+y_2+\ldots+y_n) = y_i/\bar{y}$. Other transformations might be more useful on occasion, depending on the data.

4. Special case: variance estimates. Suppose we have n independent, equal-sized samples each of $(2m+1)$ observations from what are thought to be Normal populations with common variance $\sigma^2 = \theta$ and we wish to make inferences about θ. On the assumption of Normality, the n variance estimates $y_i = s_i^2$, $i = 1, 2, \ldots, n$ each follow a $(\sigma^2 \chi^2_{2m})/2m$ distribution. This means that the variables my_i/θ each follow a Gamma distribution of order m.

Thus, if the original assumption that the n samples (each of $2m+1$ observations) are normal is not satisfied, we can attempt to adjust for this by working with transformed variables $y_i^{\lambda} = s_i^{2\lambda}$ which do satisfy the assumptions, where λ is selected by the procedure of sections 2 and 3. Note that in this case the transformation is made on the variance estimates. An alternative procedure, which involves more computation, is to transform the individual sample observations using the methods and transformation given by Box and Cox (1964).

5. Special case: exponential life testing. Suppose $m = 1$. Then the density (1.2) reduces to the exponential density

(5.1) $$f(z_i) = \theta^{-1} \exp(-z_i/\theta), \qquad z_i > 0 .$$

The assumption that a variable follows the density (5.1) is frequently made in life data work. For example, Maguire, Pearson, and Wynn (1952) examined mine accidents and concluded that the time intervals between accidents appeared to follow an exponential distribution. (Comprehensive bibliographies of life testing and related topics are given by Mendenhall (1958) and Govindarajulu (1964), incidentally.)

One way in which the exponential distribution arises in general is when an event is rare, and the occurrence of events follows a Poisson distribution. Then the distribution of intervals between events is exponential exactly. If the mean number of events, the parameter in the underlying Poisson distribution, is not constant in time however, the distribution of intervals will not have an exact exponential form and the usual estimation of the parameter θ of an assumed exponential distribution, and associated calculations, will be disturbed perhaps seriously. To check on such possibilities, an approach through tests of goodness of fit has been suggested by several authors including Bortkiewicz (1915), Morant (1920), Neyman and Pearson (1928) and Sukhatme (1936). An assumption of exponentiality can also be misplaced when the distribution under study arises from a source not exponentially based.

In such cases we might be interested in finding a transformation of form (1.1) which can be applied to the life-test observations and which will allow the transformed data to be validly treated as exponential observations. The availability of such a procedure would also provide an additional safeguard for the analysis of data usually assumed to be exponential.

If, in a given problem, the transformation $z = y^{\lambda}$ is used to transform observations on y to a form in which they can be treated as exponential observations, and so analysed, it will be necessary to translate conclusions in terms of the original distribution, whatever it may be. Suppose we wish to make an inference about $\gamma = E(y)$.

This parameter can be related to the parameter θ of (5.1) as follows. Since $z = y^{\lambda}$ has the distribution (5.1) and since $dz/dy = \lambda y^{\lambda-1}$, then

(5.2) $$y \sim g(y) \equiv \lambda\theta^{-1} y^{\lambda-1} \exp(-y^{\lambda}/\theta).$$

It can be shown that

(5.3) $$\gamma = E(y) = \int y\, g(y)\, dy = \theta^{1/\lambda}\, \Gamma(1+\lambda^{-1}).$$

Thus estimates and confidence limits for θ can be converted into estimates and confidence limits for γ through the formula (5.3).

6. Examples using half-normal data. To show how the method can be applied to life test data we shall subject, to our analysis, two samples from the half-normal distribution which has a bulk shape something like the exponential distribution. By using data from a specified distribution, we can check the results of our transformed analysis against those obtained from an appropriate analysis of the original data. (In general we would not, of course, know the parent distribution; if we did, transformation would not be necessary.)

The half-normal distribution is obtained by folding over an $N(0, \sigma^2)$ distribution and has probability density function

(6.1) $$f(y, \sigma) = \frac{2^{\frac{1}{2}}}{\sigma\pi^{\frac{1}{2}}} \exp\left\{-\frac{y^2}{2\sigma^2}\right\}, \quad 0 \leq y \leq \infty.$$

It can easily be verified that

$$(6.2) \qquad E(y) = \sigma(2/\pi)^{\frac{1}{2}} = \gamma, \text{ say}; \quad V(y) = \sigma^2(\pi-2)/\pi .$$

If $y_1, y_2, \ldots, y_n$ is a sample from the half-normal distribution, $\bar{y}$ estimates γ and $V(\bar{y}) = \sigma^2(\pi-2)/\pi n = \gamma^2(\pi-2)/2n$. For reasonable $n(> 10$ say) we can assume that, approximately,

$$(6.3) \qquad \bar{y} \sim N(\gamma, \gamma^2(\pi-2)/2n)$$

and so obtain an approximate 100 $(1-\alpha)$ % confidence interval for γ of

$$(6.4) \qquad \gamma_1 \equiv \bar{y}/(1+T) < \gamma < \bar{y}/(1-T) \equiv \gamma_2$$

where $T = z_{1-\frac{1}{2}\alpha}\{(\pi-2)/2n\}^{\frac{1}{2}}$ and $z_{1-\frac{1}{2}\alpha}$ is the $1-\frac{1}{2}\alpha$ point of the $N(0,1)$ distribution.

For a suitable value of λ, selected as described in Sections 2 or 3, the transformed observations $z_i = y_i^\lambda$, $i = 1, 2, \ldots, n$, follow the exponential distribution (5.1), whose mean θ is estimated by $\bar{z}$. Since $\bar{z} \sim \theta\chi^2_{2n}/(2n)$, we can obtain a $(1-\alpha)$ confidence interval for θ from

$$(6.5) \qquad \theta_1 = 2n\,\bar{z}/\chi^2_{2n,\,1-\frac{1}{2}\alpha} < \theta < 2n\,\bar{z}/\chi^2_{2n,\,\frac{1}{2}\alpha} = \theta_2$$

where $\chi^2_{2n,\,1-\alpha}$ is the $(1-\alpha)$ point of the χ^2_{2n} distribution. We now have to find a basis for comparing the results of (6.4) and (6.5). To do this we can apply the formula (5.3) to θ_1, $\bar{z}$, and θ_2 of (6.5) and compare these to $\gamma_1, \bar{y}$, and γ_2 of (6.4) where the agreement, hopefully, will be good. (In a problem where the parent distribution is not known, (6.4) would not, of course, be known and the results of (6.5) converted by (5.3) would be used in its stead.) Since the parent distribution is known in the present case, and is (6.1), we can obtain

$$(6.6) \qquad \theta = E(z) = E(y^\lambda) = \int y^\lambda f(y,\sigma)\,dy$$

$$= 2^{\frac{1}{2}\lambda}\sigma^\lambda \pi^{-\frac{1}{2}}\Gamma\{\tfrac{1}{2}(\lambda+1)\}$$

$$= \gamma^\lambda \pi^{\frac{1}{2}(\lambda-1)}\Gamma\{\tfrac{1}{2}(\lambda+1)\}$$

https://doi.org/10.4153/CMB-1968-057-3 Published online by Cambridge University Press

so that

$$(6.7) \qquad \gamma = \theta^{1/\lambda}/[\pi^{\frac{1}{2}(\lambda-1)} \Gamma\{\tfrac{1}{2}(\lambda+1)\}]^{1/\lambda}$$

$$= \theta^{1/\lambda}/[\pi^{\frac{1}{2}}\{\pi^{-\frac{1}{2}}\Gamma(\tfrac{1}{2}(\lambda+1))\}^{1/\lambda}]$$

$$= \theta^{1/\lambda}\{\pi^{-\frac{1}{2}}[\pi^{-\frac{1}{2}}\Gamma(\tfrac{1}{2}(\lambda+1))]^{-1/\lambda}\}.$$

For our method to be sensible, the conversions (5.3) and (6.7) should be effectively the same when the untransformed data comes from a half-normal parent.

We shall now perform the calculations described above on two sets of half-normal data given (with original sources) in a paper by Daniel (1959) and reproduced in Table 1. In using the half-normal samples we remove those contrast observations, originally recorded, which fall off the half-normal plot and thus do not appear to belong to the underlying half-normal distribution.

Table I: Half-Normal Samples

Sample Reference No.	Observations	$\bar{y}$	n	Source (in Daniel, 1959)
1	2, 9, 14, 14, 32 43, 66, 66, 66, 82 100, 105, 123, 146	62	14	Table 4, page 328 Column a times 100
2	9, 17, 19, 23, 38 53, 55, 116, 116, 127	57.3	10	Table 4, page 328 Column f times 100

For comparison we shall treat these data by both the likelihood method and the Bayesian method with the prior (3.4).

Figure 2 shows four posterior distributions. Two of these (one for eachof samples 1 and 2) arise from the use of a constant prior. This is equivalent to using the likelihood method which we do not use because the posterior distribution provides a better appreciation than does the likelihood function, of other values of λ, apart from the maximising value, which the data regard as plausible. The remaining two posterior distributions of Figure 2 are those which arise from samples 1 and 2 when the prior distribution for λ is proportional to $1/\lambda$.

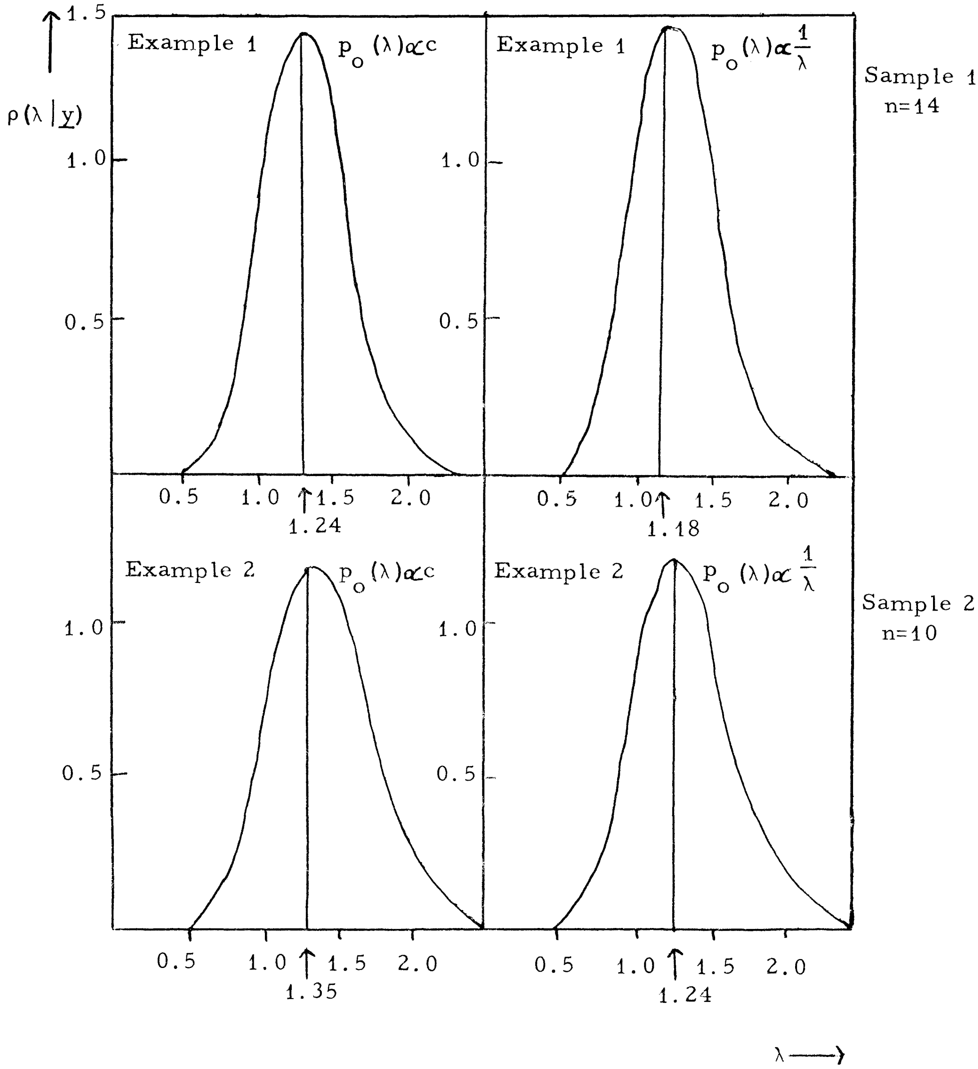


Figure 2. Posterior Distributions, Examples 1 and 2.

Results using the Bayesian method with constant prior. The modal values of the posterior distributions occur at $\hat{\lambda} = 1.24$ (sample No. 1) and $\hat{\lambda} = 1.33$ (sample No. 2). (Note however that because of the small size of the samples, the value $\lambda = 1$ is not an "unreasonable" value when considered in relation to the posterior distribution, though it does not provide the maximum posterior value.) We can immediately compare the two conversion factors which appear in (5.3) and (6.7) as follows

	$\Gamma(1+\lambda^{-1})$	$\pi^{-\frac{1}{2}}\{\pi^{-\frac{1}{2}}\Gamma(\frac{1}{2}(\lambda+1))\}^{-1/\lambda}$
Sample No. 1.	0.933	0.938
Sample No. 2.	0.920	0.912

We see that there is excellent agreement and will use the figures 0.933 and 0.920 from (5.3) for conversion in Samples 1 and 2 respectively on the basis that the factors in (6.7) would not normally be available to us when the parent is unknown.

We now transform the original observations by $z = y^{\lambda}$ and perform the calculations of equations (6.4) and (6.5) converted by (5.3). The results appear in Table 2.

Table 2: Comparison of transformed and untransformed analyses.

Sample No.	λ	Transformed observations y^{λ}	Estimates from (6.4)	Estimates from (6.5) converted by (5.3)
1	1.24	2.36, 15.3, 26.4, 26.4, 73.5, 106.0, 180.4, 180.4, 180.4, 236.1, 302.0, 320.8, 390.4, 482.9	44.4, 62.0, 102.6	42.3, 61.5, 100.2
2	1.33	18.6, 43.3, 50.2, 64.7, 126.2, 196.4, 206.4, 556.9, 556.9, 628.2	39.0, 57.3, 107.8	38.4, 57.4, 99.9

Results using the Bayesian method with $p_o(\lambda) = 1/\lambda$. The maximum values of the posterior distributions occur at $\hat{\lambda} = 1.18$ (Sample No. 1) and $\hat{\lambda} = 1.24$ (Sample No. 2) as shown in Figure 2. (Note however that because of the small size of the samples the value $\lambda = 1$ is not an "unreasonable" value when considered in relation to the posterior distribution though it does not provide the maximum posterior value.) Comparison of the two conversion factors from (5.3) and (6.7) gives

	$\Gamma(1+\lambda^{-1})$	$\pi^{-\frac{1}{2}}\{\pi^{-\frac{1}{2}}\Gamma(\frac{1}{2}(\lambda+1))\}^{-1/\lambda}$
Sample No. 1	0.945	0.953
Sample No. 2	0.933	0.938

Again the agreement is excellent and we use the figures 0.945 and 0.933 from (5.3) for conversion in Samples 1 and 2 respectively. We now transform the original observations by $z = y^{\lambda}$ and perform the calculations of equations (6.4) and (6.5) converted by (5.3). The results appear in Table 3.

Table 3: Comparison of transformed and untransformed analyses

Sample No.	λ	Transformed observations y^{λ}	Estimates from (6.4)	Estimates from (6.5) converted by (5.3)
1	1.18	2.27, 13.4, 22.5, 22.5, 59.7, 84.6, 140.3, 140.3, 140.3, 181.2, 229.1, 242.7, 292.5, 358.1	44.4, 62.0, 102.6	41.5, 61.5, 102.4
2	1.24	15.3, 33.6, 38.5, 48.8, 91.0, 137.4, 143.9, 363.1, 363.1, 406.1	39.0, 57.3, 107.8	37.0, 57.0, 103.1

Examination of Tables 2 and 3 shows that, as far as these examples are concerned, the likelihood (or Bayesian constant prior) approach provides closer agreement at the lower end of the confidence range and the Bayesian approach with $p_o(\lambda) = 1/\lambda$ provides closer agreement at the upper end of the confidence range. In either case agreement is excellent.

7. A further example. We now subject to the same analysis a set of data taken from a paper by Mendenhall and Hader (1958). Page 509 of that paper shows, in Table 3, 107 observations from an unknown distribution about whose mean we wish to make inferences. The data are reproduced in Table 4.

Because the sample is so large the posterior distribution (see Figure 3) is almost completely insensitive to whether the prior distribution for θ is taken to be proportional to a constant (which is equivalent to using the likelihood method) or to $1/\theta$. In both cases the maximum of the posterior is at $\lambda = 1.35$. Note (Figure 3) that in this example the value $\lambda = 1$ lies well in the lower tail of the posterior distribution and is thus regarded by the data as highly implausible.

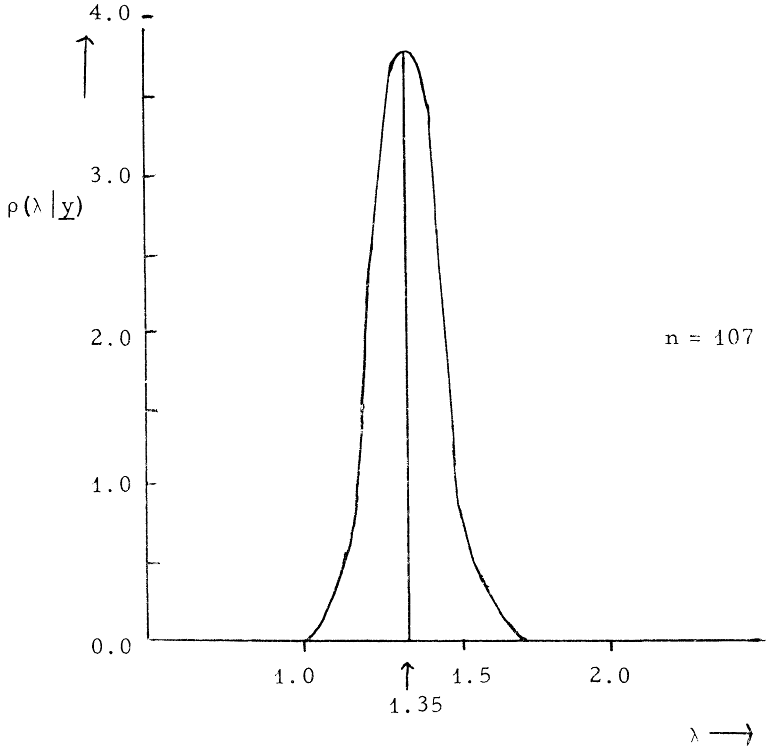


Figure 3. Posterior Distribution, Example 3.

https://doi.org/10.4153/CMB-1968-057-3 Published online by Cambridge University Press

Table 4. Unconfirmed failures.

Hours to failure for ARC-1 VHF radio transmitter receiver

368	136	512	136	472	96	144	112	104	104
344	246	72	80	312	24	128	304	16	320
560	168	120	616	24	176	16	24	32	232
32	112	56	184	40	256	160	456	48	24
200	72	168	288	112	80	584	368	272	208
144	208	114	480	114	392	120	48	104	272
64	112	96	64	360	136	168	176	256	112
104	272	320	8	440	224	280	8	56	216
120	256	104	104	8	304	240	88	248	472
304	88	200	392	168	72	40	88	176	216
152	184	400	424	88	152	184	-	-	-

* Data supplied to the original authors through the courtesy of Dr. G.R. Herd, Aeronautical Radio, Incorporated.

The transformed observations $z_i = y_i^{1.35}$ are shown in Table 5. The mean and 95 % confidence limits from these transformed observations are (1135.07, 1350.14, 1665.77) using (6.5). After conversion by (5.3) we obtain an estimate of the mean of 191 with 95 % confidence limits of (168, 223).

Table 5: 107 Observations after transformation by $y^{1.35}$

16.57	16.57	16.57	42.22	42.22	73.00	73.00
73.00	73.00	107.6	107.6	145.5	145.5	186.1
186.1	229.2	229.2	274.5	274.5	321.7	321.7
321.7	370.9	370.9	421.8	421.8	421.8	421.8
474.3	474.3	528.4	528.4	528.4	528.4	528.4
528.4	583.9	583.9	583.9	583.9	583.9	598.1
598.1	641.0	641.0	641.0	699.3	759.0	759.0
759.0	820.0	820.0	881.8	881.8	945.2	1009
1009	1009	1009	1075	1075	1075	1141
1141	1141	1277	1277	1347	1347	1418
1418	1489	1561	1634	1708	1708	1782
1782	1782	1935	1935	1935	2012	2090
2249	2249	2249	2329	2409	2409	2657
2825	2916	2916	3170	3170	3257	3524
3704	3888	4071	4071	4165	4535	5130
5428	5834	-	-	-	-	-

https://doi.org/10.4153/CMB-1968-057-3 Published online by Cambridge University Press

8. Acknowledgements. We are grateful to Jacobo Sredni who performed the calculations and helped with the figures.

This research was supported by the United States Navy through the Office of Naval Research, under Contract Nonr-1202, Project NR 042 222, and the Wisconsin Alumni Research Fund (WARF).

REFERENCES

L. von Bortkiewicz (1915). Uber die zeitfolge zufalliger ereignisse, Bulletin de L'institut International de Statistique, 20, 2, 30-111 (discussion 20, 1, 87-92).

G.E.P. Box and D.R. Cox (1964). An analysis of transformations, Journal of the Royal Statistical Society, Series B, 26, 211-243.

Cuthbert Daniel (1959). Use of half-normal plots in interpreting factorial two-level experiments. Technometrics, 1, 311-341.

Z. Govindarajulu (1964). A supplement to Mendenhall's bibliography on life testing and related topics. Journal of the American Statistical Association, 59, 1231-1291.

B.A. Maguire, E.S. Pearson and A.H.A. Wynn (1952). The time intervals between industrial accidents, Biometrika, 39, 168-180.

W. Mendenhall and R.J. Hader (1958). Estimation of parameters of mixed exponentially distributed failure time distributions from censored life test data. Biometrika, 45, 504-520.

W. Mendenhall (1958). A bibliography on life testing and related topics. Biometrika, 45, 521-543.

G.M. Morant (1920). On random occurrences in space and time, when followed by a closed interval. Biometrika, 13, 309-337,

J. Neyman and E.S. Pearson (1928). On the use and interpretation of certain test criteria for purposes of statistical inference. Part I. Biometrika, 20A, 175-240.

P.V. Sukatme (1936). On the analysis of k samples from exponential populations with especial reference to the problem of random intervals. Statistical Research Memoirs, 1, 94-112.

University of Wisconsin
Madison