

Plotting on Weibull Paper

Fritz Scholz
Research and Technology
Boeing Information & Support Services

August 23, 1996

Abstract

This report explains Weibull plotting and its rationale. It shows how the two Weibull parameter estimates are easily read off from the Weibull plot. The use of Weibull plotting is introduced first in the context of complete samples and then extended to two common forms of censoring: type I or multiple censoring and type II censoring. Two blank Weibull plotting templates are provided, one for a two cycle \log_{10} scale and the other for three cycle \log_{10} scale on the abscissa. The use of a Weibull plot as a diagnostic tool for checking the Weibull assumption underlying the sample is examined critically. For small sample sizes, answers, if obtained via this tool, should be taken with caution, since the variability of such Weibull samples is still substantial.

1 Introduction

In characterizing the distribution of life lengths or failure times of certain devices one often employs the Weibull distribution. This is mainly due to its weakest link properties, but other reasons are its increasing¹ failure rate with device age and the variety of distribution shapes that the Weibull density offers. The increasing failure rate accounts to some extent for fatigue failures.

Weibull plotting is a graphical method for informally checking on the assumption of Weibull distribution model and also for estimating the two Weibull parameters. The method of Weibull plotting is explained and illustrated both for complete samples of failure times as well as for censored samples. In the latter case we consider either type I or multiply censored data or type II censored data. Type I or multiple censoring occurs typically in field data where either the failure times of devices are observed or their last running time is known, i.e., the last time these devices were checked they were still functioning properly. The term multiple censoring instead of type I censoring is motivated by the multiple time points at which some censoring event takes place and prevents the observation of the complete life time. The censoring time points are extraneous events and independent of the actual device failure times, whether these are observed or not. This is in contrast to type II censoring. This occurs mostly under laboratory conditions when all n devices are put on test at the same time and one observes the failure times of the first r devices, with the remaining $n - r$ devices still running successfully. This kind of censoring is useful when one cannot wait until all devices have failed, but wants to guarantee at least so many failures. Here the censoring time is not independent of the failure times, since it coincides with the r^{th} smallest failure time.

It is assumed that the two-parameter Weibull distribution is a reasonable model for describing the variability in the failure time data. If T represents the generic failure time of a device, then the distribution function of T is given by

$$F_T(t) = P(T \leq t) = 1 - \exp\left(-\left[\frac{t}{\alpha}\right]^\beta\right) \quad \text{for } t \geq 0.$$

The parameter α is called the scale parameter or *characteristic life*. The latter term is motivated by the property $F_T(\alpha) = 1 - \exp(-1) \approx .632$, regardless of the shape parameter β . There are many ways for estimating the parameters α and β from complete or censored data. One of the simplest is through the method of Weibull plotting, which is very popular due to its simplicity, graphical appeal, and its informal check on the Weibull model assumption.

¹for Weibull shape parameter $\beta > 1$

2 Weibull Plotting and its Basis

The basic idea behind Weibull plotting is the relationship between the p -quantiles t_p of the Weibull distribution and p for $0 < p < 1$. The p -quantile t_p is defined by the following property

$$p = F_T(t_p) = P(T \leq t_p) = 1 - \exp\left(-\left[\frac{t_p}{\alpha}\right]^\beta\right)$$

which leads to

$$t_p = \alpha [-\log_e(1-p)]^{1/\beta}$$

or taking decimal logs² on both sides

$$y_p = \log_{10}(t_p) = \log_{10}(\alpha) + \frac{1}{\beta} \log_{10}[-\log_e(1-p)] . \quad (1)$$

Thus $\log_{10}(t_p)$, when plotted against $w(p) = \log_{10}[-\log_e(1-p)]$ should follow a straight line pattern with intercept $\log_{10}(\alpha)$ and slope $1/\beta$. Plotting $w(p)$ against $y_p = \log_{10}(t_p)$, as is usually done in a Weibull plot, one should see the following linear relationship

$$w(p) = \beta [\log_{10}(t_p) - \log_{10}(\alpha)] \quad (2)$$

with slope β and abscissa intercept $\log_{10}(\alpha)$.

In place of the unknown \log_{10} -quantiles $\log_{10}(t_p)$ one uses the corresponding sample quantiles. For a complete sample, T_1, \dots, T_n , these are obtained by ordering these T_i from smallest to largest to get $T_{(1)} \leq \dots \leq T_{(n)}$ and then associate with $p_i = i/(n+1)$ the p_i -quantile estimate or i^{th} sample quantile $T_{(i)}$. These sample quantiles tend to vary around the respective population quantiles t_{p_i} . For large sample sizes and for $p_i = i/(n+1) \approx p$ with $0 < p < 1$ this variation diminishes (i.e., the sample quantile $T_{(i)}$ converges to t_p in a sense not made precise here). For p_i close to 0 or 1 the sample quantiles $T_{(i)}$ may (or may not) exhibit high variability even in large samples. Thus one has to be careful in interpreting extreme sample values in Weibull plots.

The idea of Weibull plotting for a complete sample is to plot $w(p_i) = \log_{10}[-\log_e(1-p_i)]$ against $\log_{10}(T_{(i)})$. Aside from the variation of the $T_{(i)}$ around t_{p_i} one should, according to equation (2), then see a roughly linear pattern.

This plotting is facilitated by Weibull paper with a \log_{10} -transformed abscissa with untransformed labels and a transformed ordinate scale given by $w(p) = \log_{10}[-\log_e(1-p)]$

²The explicit notation \log_{10} and \log_e is used to distinguish decimal and natural logs.

with labels in terms of p . Sometimes this scale is labeled in percent (i.e., in terms of $100p\%$). Two blank samples of such Weibull probability paper are given as the first two figures in Appendix A, although they are not labeled as Figure 1 or Figure 2. Figure 1 has two \log_{10} cycles on the abscissa, facilitating plotting of failure times over two orders of magnitudes and Figure 2 has three cycles on the abscissa, facilitating plotting of failure times over three orders of magnitudes.

For each plotting point $(\log_{10}(T_{(i)}), w(p_i))$ one locates or interpolates the label value of $T_{(i)}$ on the abscissa and the value p_i on the ordinate, i.e., there is no need for the user to perform the transformations $\log_{10}(T_{(i)})$ and $w(p_i) = \log_{10}[-\log_e(1 - p_i)]$. An example of a complete sample plotted on Weibull paper is given in Figure 3 of Appendix A. Figure 3 shows three lines. The dashed line represents the true line corresponding to the Weibull distribution from which the ten values were sampled. The other two lines represent a least squares fit (the formulas for which will be given later) and the other corresponds to maximum likelihood estimates (m.l.e.) of α and β . The process of finding the m.l.e.'s is complicated and usually accomplished through software. Note how susceptible the least squares fit is to the two lower extreme values. This is not surprising since the method of least squares, as applied here, treats all data equally. It does not know that the data come from a Weibull distribution and represent ordered and thus correlated values. The method of maximum likelihood employs the fact that the data come from a Weibull model and knows how to properly weigh the various observations, i.e., stragglers as they show up in Figure 3 will not be given undue influence.

The two blank specimens of Weibull probability paper in Appendix A cover two and three orders of magnitude on the abscissa, namely from 1 to 100 or from 1 to 1000. If the observed life times cover a range from 50 to 4000, one can simply change time units to tens and use the three \log_{10} cycle paper from 5 to 400, which fits. If the ranges are very large, one may have to use Weibull paper covering more orders of magnitude. However, there is a simple transformation device around that difficulty. It is based on the following power transformation property of the Weibull distribution. If $T \sim \mathcal{W}(\alpha, \beta)$ (i.e., T has a Weibull distribution with parameters α and β), then

$$T^a \sim \mathcal{W}(\alpha^a, \beta/a) = \mathcal{W}(\alpha', \beta') .$$

Thus one can always bring the scale of the failure times up or down into the proper range by an appropriate power transformation.

By eye or by more formal least squares methods one can then fit a line through the plotted points and take it as a proxy or estimate of the true line representing the unknown linear relationship (2) between the quantiles y_p and $w(p)$. Here the least squares fit will be

quite sensitive to variability in the extreme sample values. Trying allow for that in fitting by eye will be somewhat subjective.

Since $p = 0.632$ yields $w(p) = 0$ or $\log_{10}(T) - \log_{10}(\alpha) = 0$ one can read off an estimate of α from the abscissa scale where the fitted line intercepts the ordinate level 0.632. For this purpose Weibull paper shows a level line at the ordinate 0.632. On the right side scale of the Weibull probability paper one finds the zero value resulting from the transform $w(0.632) = 0$.

The scale to the left of the ordinate scale runs from zero to ten and is a nomographic device for reading off the estimated shape parameter associated with the line fitted to the plotted data. To obtain it one draws a line parallel to the fitted line going through the solid dot to the right of that shape scale. Appendix A contains several examples illustrating this process.

Some authors suggest to use of $p'_i = (i - .3)/(n + .4)$ in place of p_i to characterize the p for which $T_{(i)}$ is the p -quantile. For large n there is little difference between the two methods and for small n the inherent variability in Weibull samples makes a preference between the two methods somewhat questionable.

3 Plotting with Type I Censoring

Under this type of censoring the censoring times and failure times typically intermingle and it is no longer quite so clear for which quantile to consider each failure time as a quantile estimate. There are various schemes of dealing with this problem. The one presented here is the Herd-Johnson method as given in Nelson (1982).

One orders all observations, censored or uncensored, from smallest to largest and one ranks them in reverse order. If $T_{(1)} \leq \dots \leq T_{(n)}$ are the ordered observations, their corresponding reverse ranks are $(r_1, \dots, r_n) = (n, \dots, 1)$. For the i^{th} failure one computes recursively the “reliability”

$$R_i = \frac{r_{(i)}}{r_{(i)} + 1} R_{i-1}$$

where $R_0 = 1$ is the reliability or survival probability at time 0. Here the distinction of $r_{(i)}$ and r_i is the following. Whereas $r_i = n - i + 1$ represents the reverse rank of the i^{th} ordered observation (censored or uncensored), the notation $r_{(i)}$ is the reverse rank associated with the i^{th} ordered failure time. For example, if the first failure time is preceded by two censoring times, then $r_1 = n$, $r_2 = n - 1$, $r_3 = n - 2 = r_{(1)}$. The sample calculations in Table 1 should clarify this further.

In the above recursive formula for R_i one can view $r_{(i)}/(r_{(i)} + 1)$ as an estimate of the conditional probability of survival at the time of the i^{th} failure, since at the i^{th} failure there are $r_{(i)}$ items, one has just failed and the other $r_{(i)} - 1$ survived. Thus the proportion of survived items is $(r_{(i)} - 1)/r_{(i)}$. Since this can lead to zero values, which cause problems in probability plotting, one modifies this conditional probability estimate to $r_{(i)}/(r_{(i)} + 1)$. R_{i-1} is the estimate of survival prior to i^{th} failure and R_i is the estimate of survival after the i^{th} failure.

As an aside, the proportion $(r_{(i)} - 1)/r_{(i)}$ figures strongly in the definition of the survival function estimate due to Kaplan and Meier (1958). This estimate is defined by

$$R'_i = \frac{r_{(i)} - 1}{r_{(i)}} R'_{i-1}$$

with $R'_0 = 1$. Kaplan and Meier show that the step function with steps $\hat{F}_i = 1 - R'_i$ at the i^{th} failure time is the nonparametric maximum likelihood estimate of the distribution function of failure times without reference to a Weibull model. See Scholz (1980) for a proper definition of maximum likelihood in such wider nonparametric settings.

Returning to the recursion defining R_i , the p_i plotting position for the i^{th} failure time is then taken as $p_i = 1 - R_i$. Note that the index i here counts consecutively through the failure times. No points are plotted corresponding to the censored times. If there are no censoring times, this method reduces to the above method of using $p_i = i/(n + 1)$ in complete samples and can be viewed as a reason for preferring this method over others.

An example calculation is presented in Table 1, which is taken from Nelson (1982). The corresponding Weibull plot is illustrated in Figure 4 of Appendix A. Note that the time units were viewed in tens since the failure data range is $[31.7, 110.0]$ and a two cycle \log_{10} Weibull paper was used.

From the Weibull plot one can read off the following estimates for α and β , namely $\hat{\alpha} \approx 125$ and $\hat{\beta} \approx 1.99$. The line was fitted by the method of least squares. To do so one has to work with the transformed failure times $Y_{[i]} = \log_{10}(T_{[i]})$ and with the corresponding values $w(p_{[i]}) = \log_{10}[-\log_e(1 - p_{[i]})]$. The square brackets around the subscripts indicate that the numbering is consecutive along the k failures. In the example of Table 1 this is along $1, 2, \dots, 7$, since there are $k = 7$ failure times. The least squares calculations use the following formulas

$$\hat{\beta} = \frac{\sum_{i=1}^k w(p_{[i]})(Y_{[i]} - \bar{Y})}{\sum_{i=1}^k (Y_{[i]} - \bar{Y})^2} \quad \text{with} \quad \bar{Y} = \frac{1}{k} \sum_{i=1}^k Y_{[i]}$$

Table 1: Winding Data & Herd-Johnson Calculations

Time	Reverse Rank r_i	Reliability/Survival Probability				Failure Prob. $p_i = 1 - R_i$	
		Cond'l $r_{(i)}/(r_{(i)} + 1)$	×	Previous R_{i-1}	=		Current R_i
31.7	16	(16/17)	×	1.000	=	0.941	.059
39.2	15	(15/16)	×	0.941	=	0.883	.117
57.5	14	(14/15)	×	0.883	=	0.824	.176
65.0 ⁺	13						
65.8	12	(12/13)	×	0.824	=	0.761	.239
70.0	11	(11/12)	×	0.761	=	0.697	.303
75.0 ⁺	10						
75.0 ⁺	9						
87.5 ⁺	8						
88.3 ⁺	7						
94.2 ⁺	6						
101.7 ⁺	5						
105.8	4	(4/5)	×	0.697	=	0.557	.443
109.2 ⁺	3						
110.0	2	(2/3)	×	0.557	=	0.372	.628
130.0 ⁺	1						

+ Running or Censored

and

$$-\hat{\beta} \log_{10}(\hat{\alpha}) = \bar{w} - \hat{\beta} \bar{Y} \quad \text{with} \quad \bar{w} = \frac{1}{k} \sum_{i=1}^k w(p_{[i]}) .$$

The estimates from the least squares calculations are $\hat{\alpha} = 123.3$ and $\hat{\beta} = 1.986$.

Since the variability is in the $Y_{[i]}$ and not in the $w(p_{[i]})$ one may prefer doing the least squares calculations with abscissa and ordinate reversed, i.e., according to the model (1). In that case one obtains

$$\frac{1}{\hat{\beta}} = \frac{\sum_{i=1}^k (w(p_{[i]}) - \bar{w}) Y_{[i]}}{\sum_{i=1}^k (w(p_{[i]}) - \bar{w})^2} \quad \text{with} \quad \bar{w} = \frac{1}{k} \sum_{i=1}^k w(p_{[i]})$$

and

$$\log_{10}(\hat{\alpha}) = \bar{Y} - \bar{w}/\hat{\beta} \quad \text{with} \quad \bar{Y} = \frac{1}{k} \sum_{i=1}^k Y_{[i]} .$$

Although the same symbols $\hat{\alpha}$ and $\hat{\beta}$ were used, their values will not be the same under the two least squares approaches. In fact, for this last method the least squares estimates are $\hat{\alpha} = 120.5$ and $\hat{\beta} = 2.055$. The resulting fitted line is indicated by the dashed line in Figure 4. Also shown is the line resulting from the maximum likelihood estimates with $\hat{\alpha} = 123.2$ and $\hat{\beta} = 2.376$.

4 Plotting with Type II Censoring

Here the observed data consist of $T_{(1)} \leq \dots \leq T_{(r)}$ and all that is known about the other $n - r$ (future) failure times is that they exceed $T_{(r)}$. Since these first r ordered failures are the quantile estimates of t_{p_1}, \dots, t_{p_r} with $p_i = i/(n + 1)$ one can again plot these r points as before on Weibull paper, the only difference being that the $n - r$ censored points are not plotted.

Again one can fit a line by eye or by least squares. However here the low extremes will weigh in more heavily since the upper extremes are missing. As an example, the six lowest values of the complete sample underlying Figure 3 were taken as a type II censored sample and are plotted in Figure 5. As in Figure 3 the true line, the least squares fit line, and the line corresponding to the maximum likelihood estimates are shown. Note that the latter is reasonably close to the true line, whereas the least squares line is led astray substantially.

5 Checking the Weibull Model

According to the motivation behind Weibull plotting one expects to see a roughly linear pattern in the plotted points. The examples shown so far may have put a damper on this, but that is mainly due to the small sample sizes in those examples. Some of the examples were artificially generated from Weibull populations. Thus there is no question about their origin. However, the example given by Nelson is supposed to be real data and its pattern on Weibull paper looks remarkably close to linear. This has been observed in other expositions on Weibull plotting as well and it may raise false expectations in the prospective user. When nonlinear patterns are found (in small samples) the user may then be led on false chase for other causes, such as multiple failure modes. There may well be such multiple modes, but small samples are not a good starting point to look for these.

In order to get a sense for the effect of sample size on the variability in linear Weibull patterns eight Weibull samples each were generated at sample sizes $n = 10$, $n = 30$, $n = 100$ and were plotted on an equivalent of Weibull probability paper (without the grid lines, etc.), see Figures 6-8. Also shown on each plot is the true line corresponding to the Weibull distribution from which the sample was drawn. The sample to sample variability at $n = 10$ is substantial whereas the linearity and proximity to the true line is quite satisfactory at $n = 100$.

References

- [1] Abernethy, R.B., Breneman, J.E., Medlin, C.H., and Reinman, G.L. (1983). *Weibull Analysis Handbook*, Pratt & Whitney Aircraft, Government Products Division, United Technologies, P.O. Box 2691, West Palm Beach, Florida 33402.
- [2] Kaplan, E.L. and Meier, P. (1958). "Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53**, 457-481.
- [3] Nelson, W. 1982. *Applied Life Data Analysis*. John Wiley & Sons.
- [4] Scholz, F.W. (1980). "Towards a unified definition of maximum likelihood." *The Canad. J. of Statist.* **8**, 193-203.

Appendix A

This appendix contains two blank copies of Weibull probability paper (without page number). The first has a two cycle \log_{10} scale on the abscissa and the second has a three cycle \log_{10} scale. This is followed by several illustrations of Weibull paper with plotted samples.

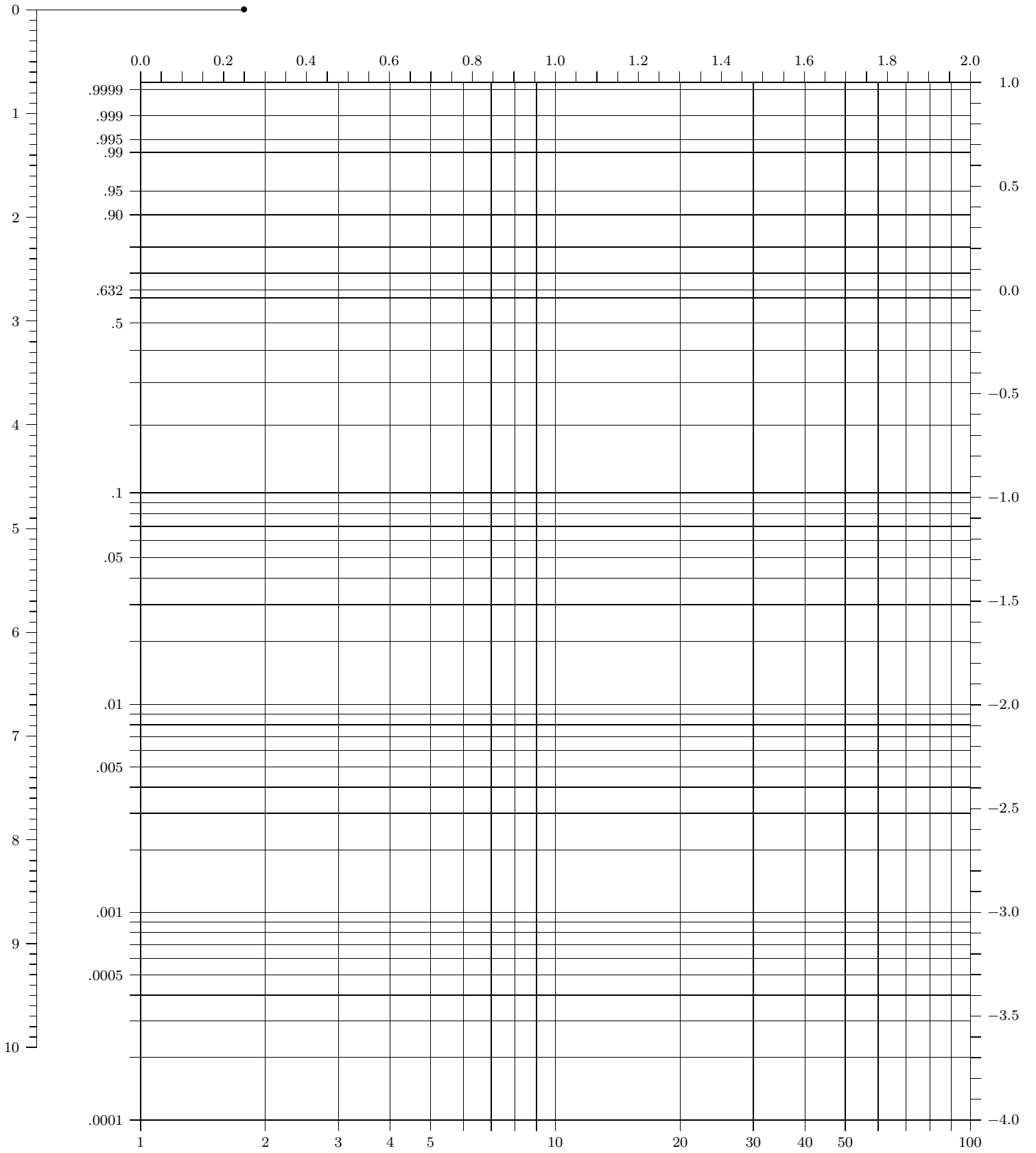
Figure 3 illustrates the plotting of a complete Weibull sample with three line superimposed: 1) the true line representing the sampled Weibull population, 2) a least squares fitted line (which is strongly influenced by stragglers), and 3) a line corresponding to maximum likelihood estimates of the parameters.

Figure 4 illustrates the plotting of a type I censored sample representing some real data taken from Nelson (1982). Aside from the maximum likelihood fitted line there are two types of least squares lines. In one the ordinate is regressed on the abscissa and in the other the abscissa is regressed on the ordinate.

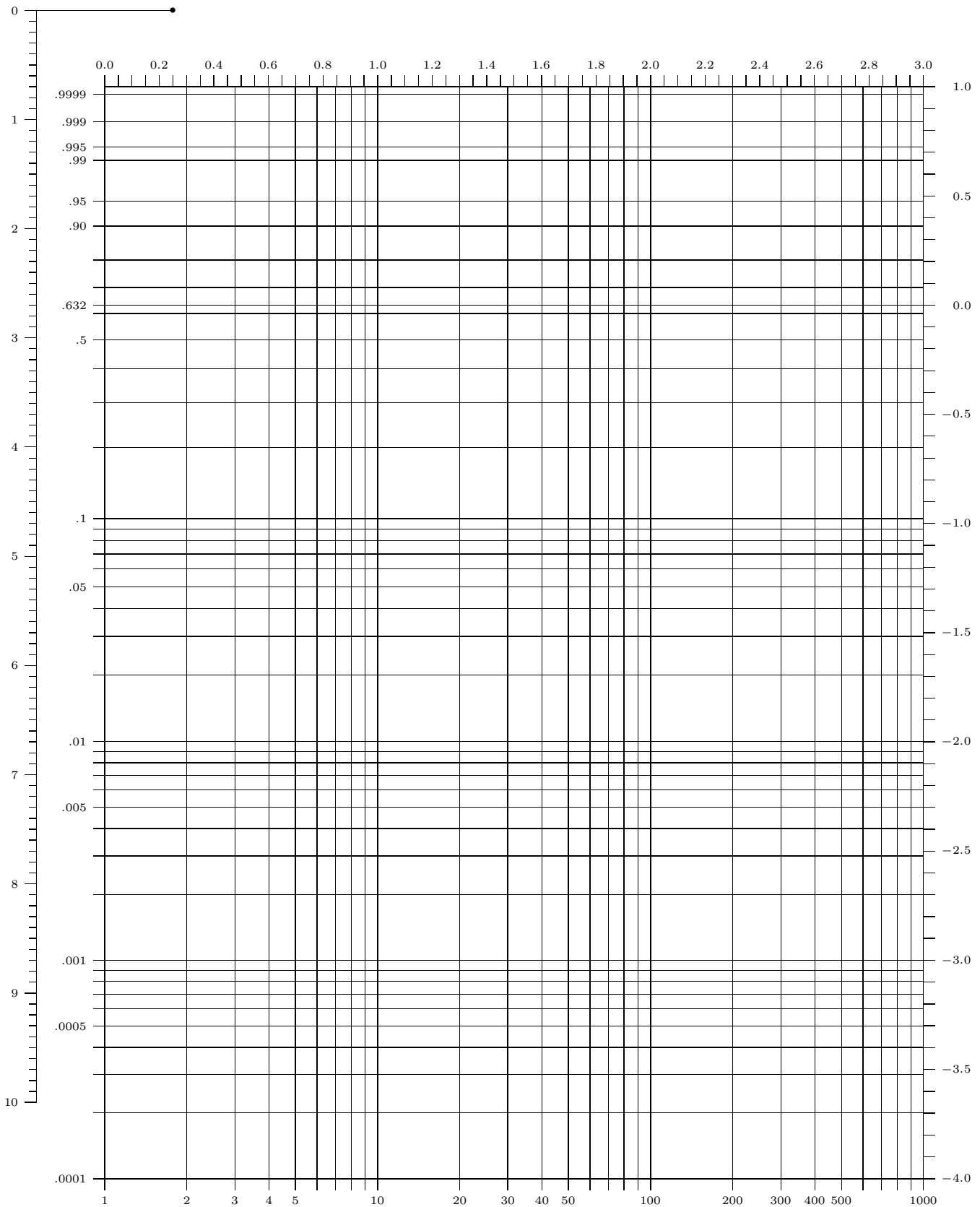
Figure 5 illustrates the plotting of a type II censored sample by taking the six lowest failure times from the complete sample underlying Figure 3. Again the true line, the least squares fitted line and the maximum likelihood fitted lines are shown.

Figure 6-8 show a collection of eight random samples, each taken from a Weibull population. Figure 6 shows those samples of size $n = 10$, and Figure 7 and Figure 8 show similar plots for sample sizes $n = 30$ and $n = 100$, respectively. On each plot the thick represents the true sampled Weibull population and the thin line is the least squares fitted line.

Weibull Probability Paper – 2 Cycle Log



Weibull Probability Paper – 3 Cycle Log



Weibull Probability Paper – 2 Cycle Log

Complete Sample

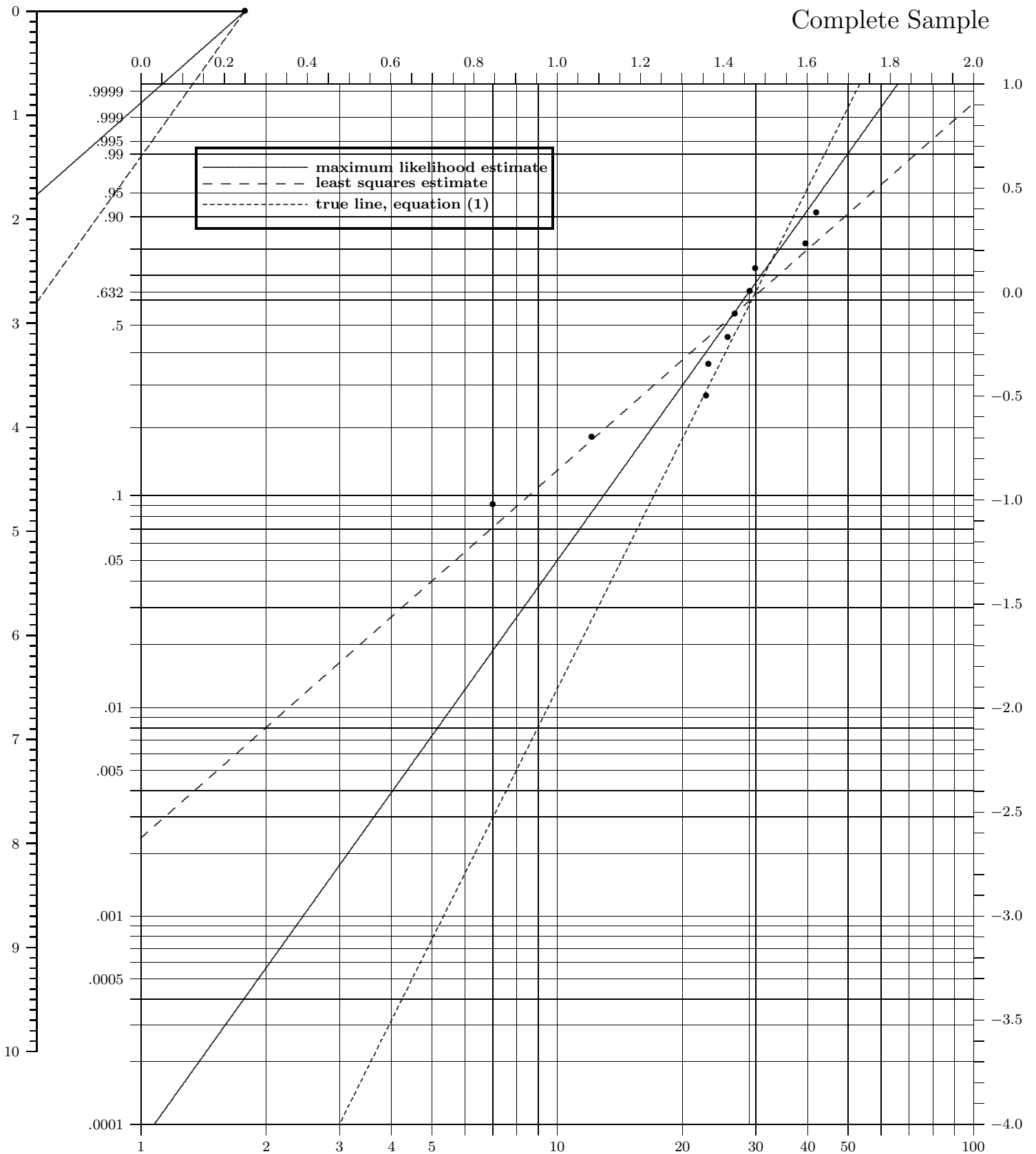


Figure 3

Weibull Probability Paper – 2 Cycle Log

Type I Censored Sample

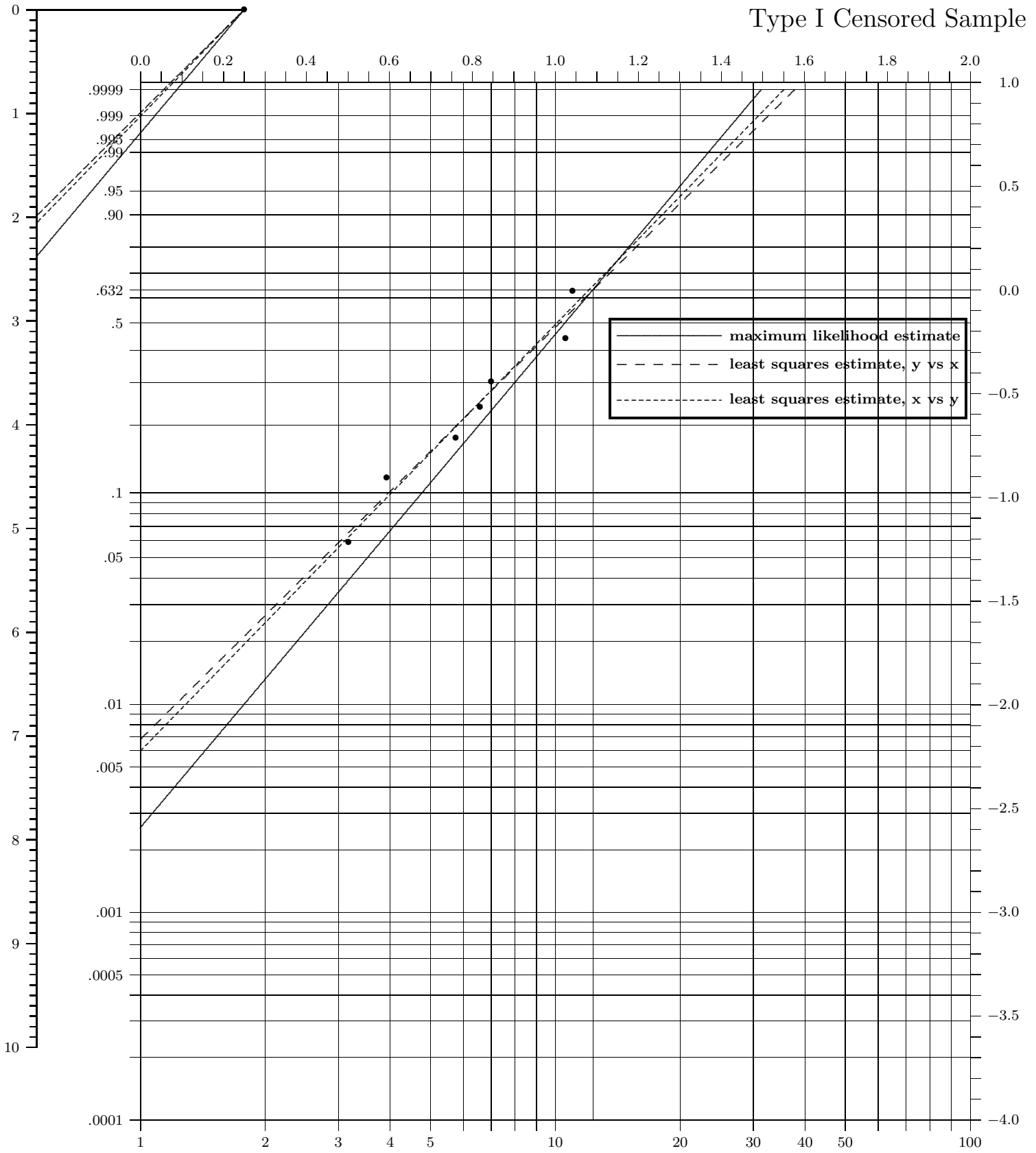


Figure 4

Weibull Probability Paper – 2 Cycle Log

Type II Censored Sample

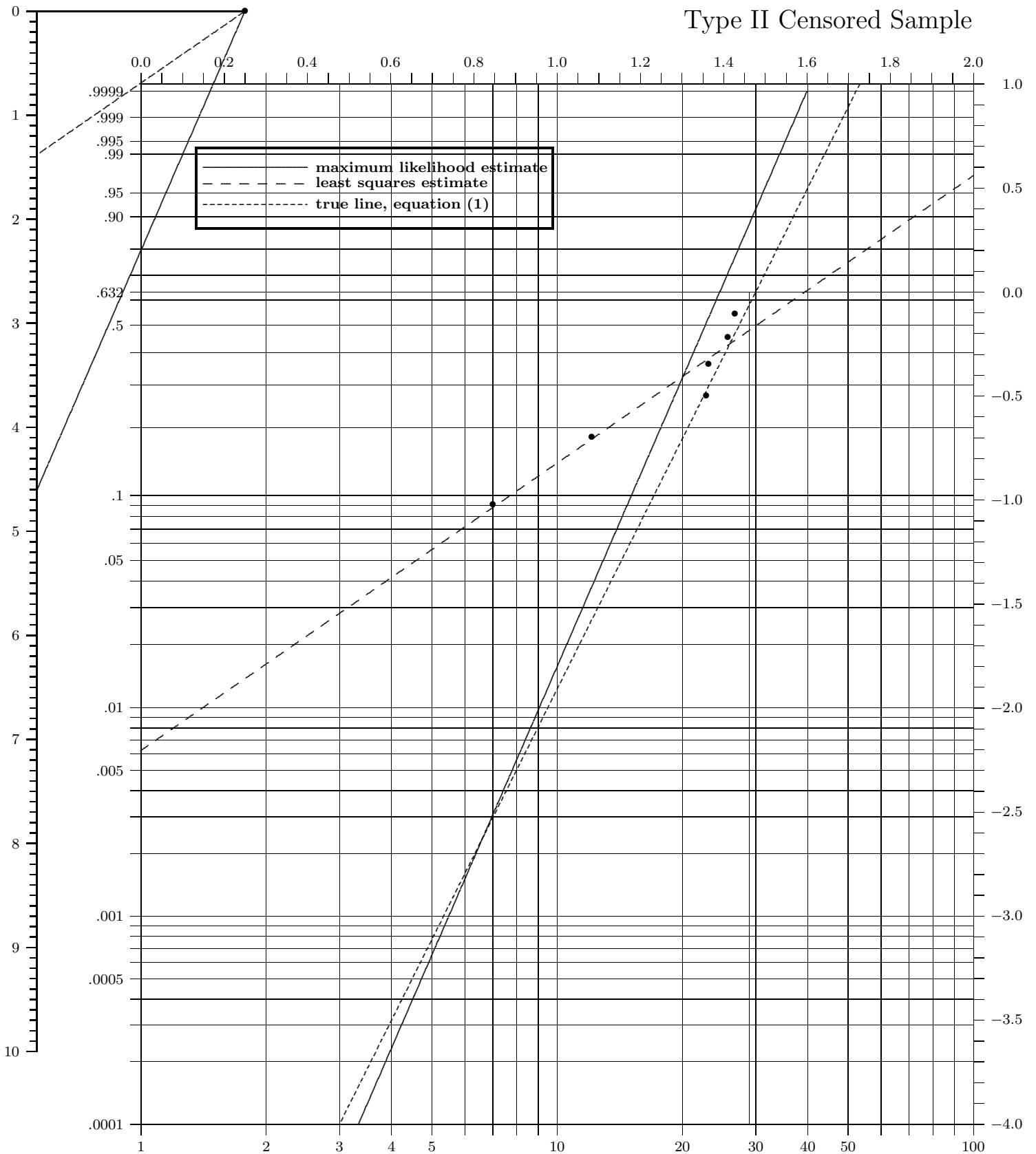


Figure 5

Figure 6: Weibull Probability Plots

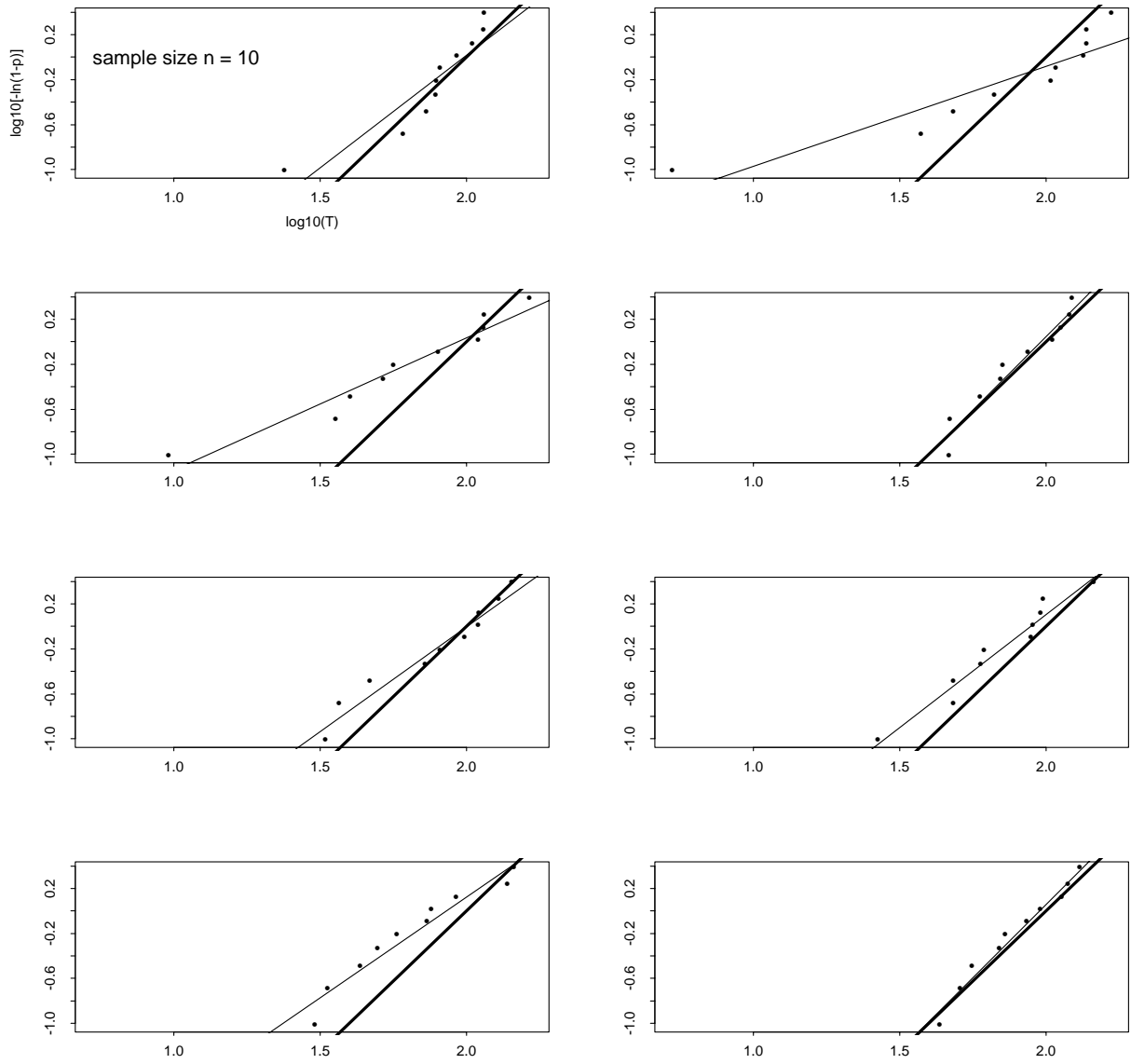


Figure 7: Weibull Probability Plots

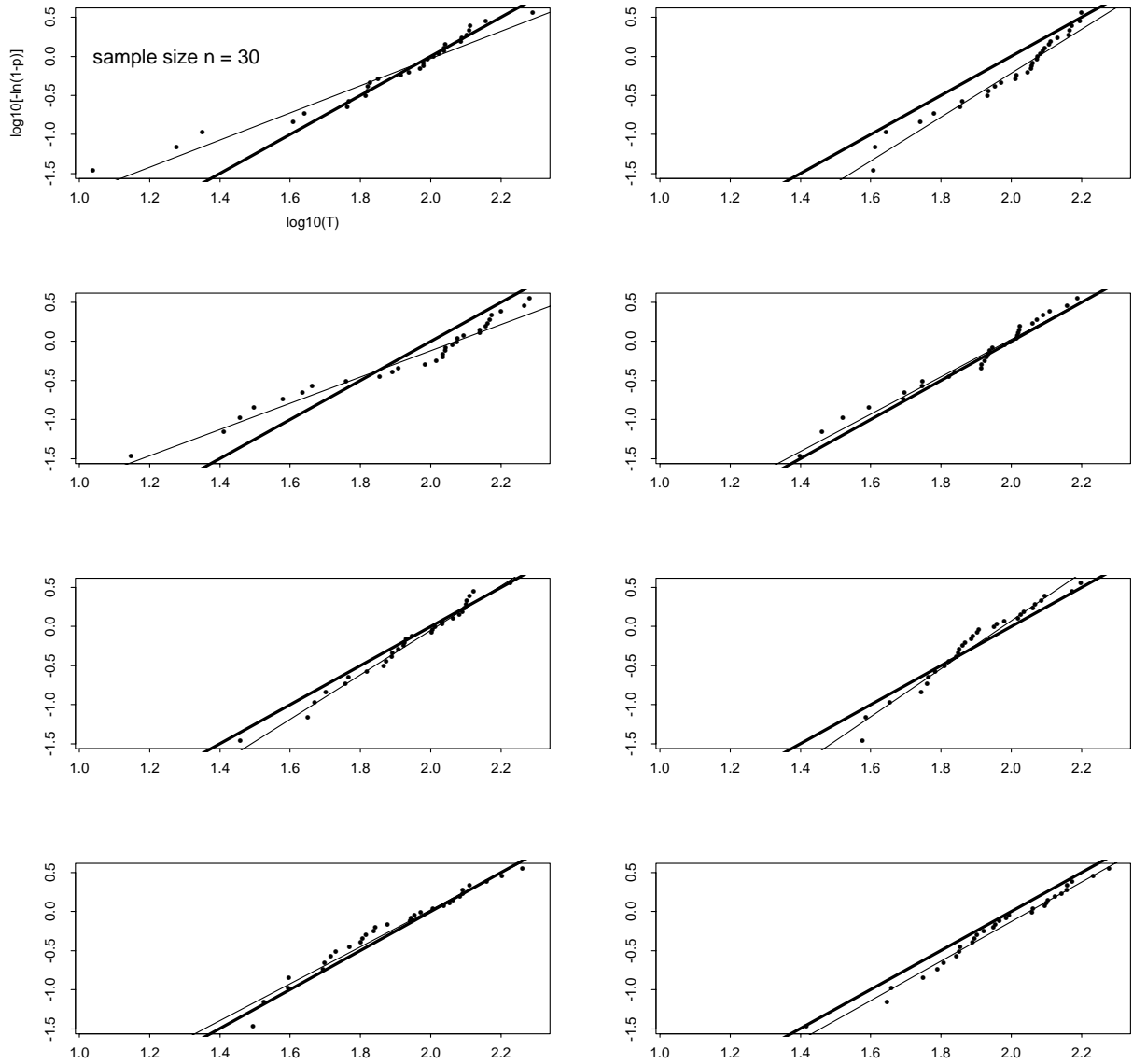


Figure 8: Weibull Probability Plots

