An Easy Parameter Estimation by the EM Algorithm in the New Up-and-down Method

Y. Komori and H. Hirose
Kyushu Institute of Technology
Fukuoka, Japan

ABSTRACT

Flashover voltage estimation has been carried out conventionally with the up-and-down method by Dixon and Mood. Recently, Hirose and Kato proposed a new version of the method to change the way of analyzing data. Although the method has better properties than the conventional one, it requires solving maximum likelihood equations. In this paper we reduce the troublesome task (*e.g.* implementation and the selection of a proper initial value) attendant on the requirement by using the expectation-maximization (EM) algorithm that gives a useful iterative formula to solve the equations. The iterative formula almost always can give the solutions of the likelihood equations because of its excellent global convergence of the EM algorithm.

1 INTRODUCTION

THE flashover voltage in gaseous electrical insulation is considered to obey a normal distribution. As an experimental means to investigate the statistical property, the up-and-down method is used [1]. In this paper we deal especially with the up-and-down 50% disruptive discharge test. IEC [1] mentions that the 50% discharge voltage U_{50} can be obtained by carrying out some simple arithmetical operations with the data that up-and-down experiments give. Here, the formula used to evaluate U_{50} essentially is equivalent to the approximate expression to a maximum likelihood estimator (MLE) given by Dixon and Mood. Strictly speaking, the estimates of the mean and the standard deviation that Dixon's formula gives are obtained by setting up the likelihood equations corresponding to the up-and-down data, simplifying the equations, and solving them. One of these estimates, *i.e.* the estimate of the mean parameter, is the very same estimate as U_{50}. We deal with the maximum likelihood (ML) method relating to conventional data. We call this the 'conventional method'. However, note that the results we will obtain here also remain correct in U_{50} although we do not deal with U_{50} directly.

The conventional method has three restrictions. First, the up-and-down distance has to be chosen to be approximately equal to the standard deviation to be estimated. If the up-and-down distance is not so chosen, the standard errors of the estimators may happen to become larger. Second, to obtain accurate estimation in practice, the sample size is recommended to be $\gtrsim 40$. However, it occurs that the number is less in some cases. This is mainly because the time for 1-shot sampling depends on insulation materials in an experiment, and will be very long for some types of material. Lastly, the number of voltage stress levels that simultaneously include one or more disruptive discharges and withstands has to be >1, otherwise, the maximum likelihood estimate does not exist.

Recently, the shape of imposed voltage may be observed by using a fast oscilloscope, and the values of flashover voltages may be seen. Hirose and Kato [4] remarked on this point and proposed a new version of the up-and-down method, in which new up-and-down data were used and the maximum likelihood method was adopted. We call this the 'new method' in the sequel. The new method can relax the restrictions mentioned above. First, the up-and-down distance does not necessarily have to be chosen to be approximately equal to the standard deviation. If only the up-and-down distance is 4× less than the standard deviation, the standard errors of the estimators scarcely depend on it. Second, the standard errors of the estimators by the new method are smaller than those by the conventional method. In other words, if both errors for the mean are almost the same, the new method requires only three fourths of the sample size that the conventional method requires. Similarly, if both errors for the standard deviation are almost the same, the new method requires only one fourth of the sample size that the conventional method requires. Moreover, even in the case mentioned in the last of the previous paragraph, it hardly happens that the maximum likelihood estimates do not exist. This is because the way of analyzing data is different from that in the conventional method.

As stated, the new method has three main advantages over the conventional one. However, the new method does not have a very simple formula to obtain the estimates as the conventional one has. The simple formula of the conventional method allows experimenters to obtain

1070-9878/0/ $3.00 © 2000 IEEE

the estimates of the mean and the standard deviation by the four simple arithmetical operations (+,−,×, /). This is one of the main reasons that the conventional method has been used up to now. Therefore, not only good estimation (*e.g.* smaller standard errors) but also tractability in practice are required of the estimation methods. In this paper we aim at making the new method tractable by applying the expectation-maximization (EM) algorithm [6] to it.

In the next Section we introduce the comparison of the new and conventional methods. In Section 3 we show a concrete formula of the method with EM algorithm. In the last Section we state a merit of the EM algorithm and some comments on the new method.

2 ASYMPTOTIC ERRORS OF THE TWO METHODS

In this Section, we introduce the new up-and-down data and the maximum likelihood method for them, and compare the asymptotic errors of the new and conventional methods.

2.1 THE NEW UP-AND-DOWN METHOD

First, we explain the new method.

2.1.1 THE WAY OF OBTAINING DATA IS AS FOLLOWS.

1. Set the initial stress voltage v_1 for some voltage guessed as the mean of flashover voltage. Determine the up-and-down distance d suitably. Here, note that the value is not necessarily close to the standard deviation of flashover voltage, unlike the case of the conventional one. It is required to provide a measuring device to record the shape of the imposed voltage.
2. In the same fashion as the conventional method, impose the impulse voltage that takes v_1 as the peak voltage on a sample.
3. Look at the shape of the impulse voltage and observe whether the disruptive discharge occurred or not. If it occurred, record the voltage V_1 at which breakdown started, and set $v_2 = v_1 - d$. If not, record v_1 and set $v_2 = v_1 + d$.
4. Take v_2 as the new stress voltage and test the next sample in a similar way.
5. Repeat the above for the number of times decided in advance.

Concerning (3), we give an example of the imposed voltage in Figure 1. The Figure shows that the imposed voltage falls in the process of going up. This is because the discharge occurs before the imposed voltage reaches the peak. Here, note that part of graph closer to the bottom indicates the voltage is higher. Such discharges are called breakdown at wave head. We restrict ourselves to such cases in this paper.

2.1.2 THE WAY OF ANALYZING DATA

Use likelihood method to estimate the mean and the standard deviation parameters. Let the mean be μ, the standard deviation σ. Then the likelihood function is defined as

$$L(\mu,\sigma) \stackrel{\text{def}}{=} \prod_{i=1}^{n} f(V_i;\mu,\sigma)^{I(i)} \left(1 - F(v_i;\mu,\sigma)\right)^{1-I(i)} \quad (1)$$

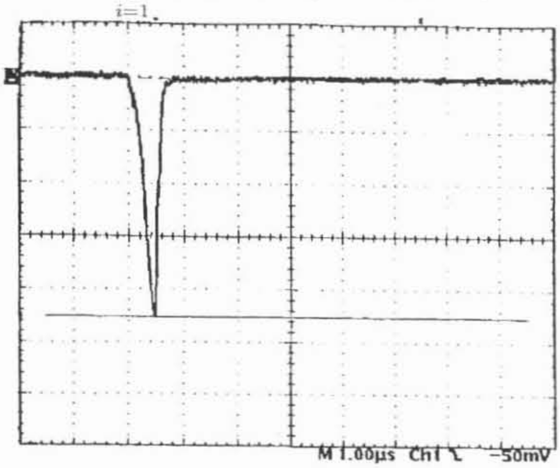

Figure 1. Breakdown at wave head.

where n stands for the sample size

$$\begin{aligned} f(v;\mu,\sigma) &\stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(v-\mu)^2}{2\sigma^2}\right\} \\ F(v;\mu,\sigma) &\stackrel{\text{def}}{=} \int_{-\infty}^{v} f(t)dt \end{aligned} \quad (2)$$

$$I(i) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{disruptive discharge occurs on trial } i \\ 0 & \text{disruptive discharge does not occur on trial } i \end{cases} \quad (3)$$

After all, the thing we have to do is to find a point $(\hat{\mu}, \hat{\sigma})$ where L has a relative maximum.

On the other hand, the likelihood function for conventional up-and-down data is

$$\tilde{L}(\mu,\sigma) \stackrel{\text{def}}{=} \prod_{i=1}^{n} F(v_i;\mu,\sigma)^{I(i)} \left(1 - F(v_i;\mu,\sigma)\right)^{1-I(i)} \quad (4)$$

Note that the difference between $L(\mu,\sigma)$ and $\tilde{L}(\mu,\sigma)$ is only the difference between $f(V_i;\mu,\sigma)$ and $F(v_i;\mu,\sigma)$ when $I(i) = 1$.

2.2 ERROR ANALYSIS

Using Fisher information matrices, we work on the theoretical error analysis of the estimators for both the up-and-down methods.

2.2.1 PRELIMINARY

We introduce a symbol and a function to make the analysis easier. Let u_0 be the initial imposed voltage and define $u_k \stackrel{\text{def}}{=} u_0 + kd$. Then, the i-th imposed voltage may take one of the voltages u_k with

$$k = \begin{cases} \pm1, \pm3, \ldots, \pm i & i \quad \text{odd} \\ 0, \pm2, \pm4, \ldots, \pm i & i \quad \text{even} \end{cases} \quad (5)$$

Next, define the indicator function

$$\delta(i,k) \stackrel{\text{def}}{=} \begin{cases} 1 & v_i = u_k \\ 0 & v_i \neq u_k \end{cases} \quad (6)$$

Once i and k have been given, $\delta(i,k)$ could become a random variable. Then, the probability function may be calculated with the recurrence relations

$$
\begin{aligned}
P[\delta(0,0)=1] &= 1 \\
P[\delta(l,-l)=1] &= P[\delta(l-1,l+1)=1]\times p_{-l+1} \\
P[\delta(l,l)=1] &= P[\delta(l-1,l-1)=1]\times q_{l-1} \\
P[\delta(l,k)=1] &= P[\delta(l-1,k-1)=1]\times q_{k-1} \\
&\quad + P[\delta(l-1,k+1)=1]\times p_{k+1} \\
k &= -l+2,-l+4,\dots,l-2
\end{aligned} \tag{7}
$$

where

$$p_l \stackrel{\text{def}}{=} F(u_l;\mu,\sigma) \qquad q_l \stackrel{\text{def}}{=} 1-p_l \tag{8}$$

However, for notational simplicity, we also introduce the symbols

$$x_k \stackrel{\text{def}}{=} \frac{u_k-\mu}{\sigma} \qquad z_k \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}x_k^2\right] \tag{9}$$

and $(B)_{ij}$, which means the i-th row and the j-th column element of matrix B.

2.2.2 THE CONVENTIONAL DATA CASE

From (2)

$$
\begin{aligned}
\ln \tilde{L}(\mu,\sigma) = \sum_{i=1}^{n}\sum_{k=-i}^{i} &\{I(i)\delta(i,k)\ln F(u_k;\mu,\sigma) \\
&+(1-I(i))\delta(i,k)\ln(1-F(u_k;\mu,\sigma))\}
\end{aligned} \tag{10}
$$

Therefore,

$$
\begin{aligned}
&E\left[\frac{\partial^2}{\partial\mu^2}\ln\tilde{L}(\mu,\sigma)\right] = \\
&\quad \sum_{i=1}^{n}\sum_{k=-i}^{i}\left\{E[I(i)\delta(i,k)]\frac{\partial^2}{\partial\mu^2}\ln F(u_k;\mu,\sigma)\right. \\
&\quad \left. +E[(1-I(i))\delta(i,k)]\frac{\partial^2}{\partial\mu^2}\ln(1-F(u_k;\mu,\sigma))\right\}
\end{aligned} \tag{11}
$$

Here,

$$
\begin{aligned}
E[I(i)\delta(i,k)] &= P[I(i)=1,\ \delta(i,k)=1] \\
&= P[I(i)=1|\delta(i,k)=1]P[\delta(i,k)=1] \\
&= p_k\times P[\delta(i,k)=1] \\
E[(1-I(i))\delta(i,k)] &= q_k\times P[\delta(i,k)=1]
\end{aligned} \tag{12}
$$

Moreover,

$$
\begin{aligned}
\frac{\partial^2}{\partial\mu^2}\ln F(u_k;\mu,\sigma) &= -\frac{1}{\sigma^2 p_k}\left(z_k x_k + \frac{z_k^2}{p_k}\right) \\
\frac{\partial^2}{\partial\mu^2}\ln(1-F(u_k;\mu,\sigma)) &= \frac{1}{\sigma^2 q_k}\left(z_k x_k - \frac{z_k^2}{q_k}\right)
\end{aligned} \tag{13}
$$

Therefore,

$$E\left[\frac{\partial^2}{\partial\mu^2}\ln\tilde{L}\right] = -\frac{1}{\sigma^2}\sum_{i=1}^{n}\sum_{k=-i}^{i}P[\delta(i,j)=1]\frac{z_k^2}{p_k q_k} \tag{14}$$

By similar calculations, we obtain the Fisher information matrix

$$
\begin{aligned}
\tilde{I}_n &\stackrel{\text{def}}{=} -\begin{bmatrix} E\left[\frac{\partial^2}{\partial\mu^2}\ln\tilde{L}\right] & E\left[\frac{\partial^2}{\partial\mu\partial\sigma}\ln\tilde{L}\right] \\ E\left[\frac{\partial^2}{\partial\sigma\partial\mu}\ln\tilde{L}\right] & E\left[\frac{\partial^2}{\partial\sigma^2}\ln\tilde{L}\right]\end{bmatrix} \\
&= \frac{1}{\sigma^2}\sum_{i=1}^{n}\sum_{k=-i}^{i}P[\delta(i,k)=1]\frac{z_k^2}{p_k q_k}\begin{bmatrix}1 & x_k \\ x_k & x_k^2\end{bmatrix}
\end{aligned} \tag{15}
$$

As the theoretical unit errors concerning with μ and σ, we define

$$
\begin{aligned}
\tilde{e}(\mu) &\stackrel{\text{def}}{=} \sqrt{n\left(\tilde{I}_n^{-1}\right)_{11}} \\
\tilde{e}(\sigma) &\stackrel{\text{def}}{=} \sqrt{n\left(\tilde{I}_n^{-1}\right)_{22}}
\end{aligned} \tag{16}
$$

respectively.

2.2.3 THE NEW DATA CASE

From (1)

$$
\begin{aligned}
\ln L(\mu,\sigma) = \sum_{i=1}^{n}\sum_{k=-i}^{i}&\Big\{I(i)\delta(i,k)\ln f(v_i;\mu,\sigma) \\
&+(1-I(i))\delta(i,k)\ln(1-F(v_i;\mu,\sigma))\Big\}
\end{aligned} \tag{17}
$$

Then, the Fisher information matrix is

$$
\begin{aligned}
I_n &\stackrel{\text{def}}{=} -\begin{bmatrix} E\left[\frac{\partial^2}{\partial\mu^2}\ln L\right] & E\left[\frac{\partial^2}{\partial\mu\partial\sigma}\ln L\right] \\ E\left[\frac{\partial^2}{\partial\sigma\partial\mu}\ln L\right] & E\left[\frac{\partial^2}{\partial\sigma^2}\ln L\right]\end{bmatrix} \\
&= \frac{1}{\sigma^2}\sum_{i=1}^{n}\sum_{k=-i}^{i}P[\delta(i,k)=1]\times \\
&\begin{bmatrix} p_k - z_k x_k + \frac{z_k^2}{q_k} & -z_k - z_k x_k^2 + \frac{z_k^2 x_k}{q_k} \\ -z_k - z_k x_k^2 + \frac{z_k^2 x_k}{q_k} & 2p_k - x_k z_k - z_k x_k^3 + \frac{z_k^2 x_k^2}{q_k}\end{bmatrix}
\end{aligned} \tag{18}
$$

Similarly to the conventional data case, we define the unit errors

$$
\begin{aligned}
e(\mu) &\stackrel{\text{def}}{=} \sqrt{n\left(I_n^{-1}\right)_{11}} \\
e(\sigma) &\stackrel{\text{def}}{=} \sqrt{n\left(I_n^{-1}\right)_{22}}
\end{aligned} \tag{19}
$$

We show in Figure 2 the tendencies of $\tilde{e}(\mu)$, $\tilde{e}(\sigma)$, $e(\mu)$, and $e(\sigma)$ *vs.* d/σ. In the Figure the solid lines correspond to those for the case $u_0=\mu$ and the dotted lines correspond to those for $u_0=\mu\pm\frac{d}{2}$.

Figure 2 shows the tendencies of the unit errors for $n=40$. There is no difference between the errors computed here and the asymptotic errors that Dixon and Mood computed approximately. This is because if only $d/\sigma\ge 0.4$, each $P[\delta(n,k)=1]$ $(k=-n,-n+2,\dots,n)$ remains stationary for all the n which are greater than 20 and are odd numbers, and so is each for all the n which are >20 and are even numbers, moreover, even if $d/\sigma=0.2$, similar things hold for all $n>40$.

From Figure 2, we can summarize our observation as follows.

1. When $0.2\le d/\sigma\le 4$, $1.25\le e(\mu)\le 1.35$ and $0.9\le e(\sigma)\le 1.1$. Comparing them with $\tilde{e}(\mu)$ or $\tilde{e}(\sigma)$, we can mention that they are small and they seem to be independent of the change of d/σ.
2. Because $\tilde{e}(\sigma)=1.96$ and $e(\sigma)=1.01$ when $d/\sigma=1$, we obtain $e(\sigma)/\tilde{e}(\sigma)\approx 0.5$. Hence, in the estimation of σ, the use of the new method leads to the incredible decrease of the sample size up to one fourth of the sample size required by the conventional method, provided that the both methods achieve almost the same magnitude of the estimation errors. In the estimation of μ, because that $e(\mu)/\tilde{e}(\mu)\approx 0.88$, such a thing as above holds with three fourths of the sample size required by the conventional method.

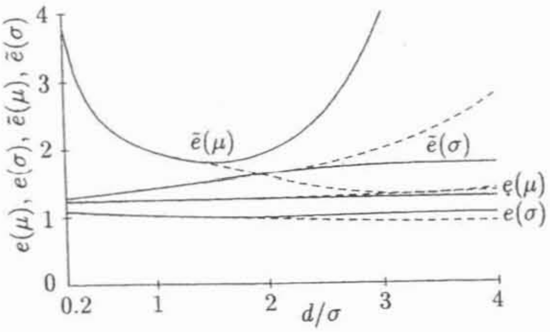

Figure 2. The unit error of the estimates when $n = 40$.

3 SIMPLE ITERATIVE FORMULA FOR THE NEW METHOD

3.1 COMPUTING PROCEDURE

We explain the application of the EM-algorithm to the new method. The EM-algorithm is an iterative method to obtain a MLE for each parameter. The algorithm has two steps, expectation step (E-step) and maximization step (M-step) in an iterative computation. In the sequel we suppose to treat the calculation concerned with only $L(\mu, \sigma)$. Let $(\hat{\mu}^{(k)}, \hat{\sigma}^{(k)})$ be the approximation to MLE in the kth iterative computation. When $(\hat{\mu}^{(k)}, \hat{\sigma}^{(k)})$ are given, E-step requires the calculation of

$$Q(\mu, \sigma; \hat{\mu}^{(k)}, \hat{\sigma}^{(k)}) \stackrel{\text{def}}{=} -n \ln(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \Big\{ I(i)(V_i - \mu)^2 + (1 - I(i)) E_{\hat{\mu}^{(k)}, \hat{\sigma}^{(k)}} \left[(X - \mu)^2 | X > v_i \right] \Big\} \tag{20}$$

in the $k+1$st iteration. Here, $E_{\hat{\mu}^{(k)}, \hat{\sigma}^{(k)}}$ stands for the expectation operator in which $f(v; \hat{\mu}^{(k)}, \hat{\sigma}^{(k)})$ is used for $f(v; \mu, \sigma)$.

The M-step requires to seek a point $(\hat{\mu}^{(k+1)}, \hat{\sigma}^{(k+1)})$ that $Q(\mu, \sigma; \hat{\mu}^{(k)}, \hat{\sigma}^{(k)})$ becomes maximum. Therefore, we obtain

$$\hat{\mu}^{(k+1)} = \frac{1}{n} \sum_{i=1}^{n} \left[I(i) V_i + (1 - I(i)) \left\{ \hat{\mu}^{(k)} + A_i(\hat{\mu}^{(k)}, \hat{\sigma}^{(k)}) \right\} \right] \tag{21}$$

$$\hat{\sigma}^{(k+1)} = \Bigg[\frac{1}{n} \sum_{i=1}^{n} \Big\{ I(i)(V_i - \hat{\mu}^{(k+1)})^2 + (1 - I(i)) \times \Big((\hat{\mu}^{(k)} - \hat{\mu}^{(k+1)})^2 + \big(\sigma^{(k)}\big)^2 \Big) + (1 - I(i)) \times (v_i + \hat{\mu}^{(k)} - 2\hat{\mu}^{(k+1)}) A_i(\hat{\mu}^{(k)}, \hat{\sigma}^{(k)}) \Big\} \Bigg]^{1/2} \tag{22}$$

where

$$A_i(\mu, \sigma) \stackrel{\text{def}}{=} \sigma^2 \frac{f(v_i; \mu, \sigma)}{1 - F(v_i; \mu, \sigma)} \tag{23}$$

Let $\hat{\mu}$ be the MLE of μ, $\hat{\sigma}$ the MLE of σ. If $\hat{\mu}^{(k)} \to \hat{\mu}$ and $\hat{\sigma}^{(k)} \to \hat{\sigma}$ as $k \to \infty$, we can express the components of the observed information matrix as

$$\begin{aligned}
-\left(\frac{\partial^2 \ln L(\mu, \sigma)}{\partial \mu^2} \right)_{(\mu,\sigma)=(\hat{\mu},\hat{\sigma})} &= \frac{1}{\hat{\sigma}^4} \sum_{i=1}^{n} \Big\{ I(i)(V_i - \hat{\mu})^2 + (1 - I(i)) A_i(\hat{\mu}, \hat{\sigma}) \Big\} \\
-\left(\frac{\partial^2 \ln L(\mu, \sigma)}{\partial \mu \partial \sigma} \right)_{(\mu,\sigma)=(\hat{\mu},\hat{\sigma})} &= -\frac{1}{\hat{\sigma}^5} \sum_{i=1}^{n} (1 - I(i)) \Big\{ (V_i - \hat{\mu})^2 + \hat{\sigma}^2 - (V_i - \hat{\mu}) A_i(\hat{\mu}, \hat{\sigma}) \Big\} \times A_i(\hat{\mu}, \hat{\sigma}) \\
-\left(\frac{\partial^2 \ln L(\mu, \sigma)}{\partial \sigma^2} \right)_{(\mu,\sigma)=(\hat{\mu},\hat{\sigma})} &= \sum_{i=1}^{n} \Bigg[\frac{2I(i)}{\hat{\sigma}^2} - \frac{1 - I(i)}{\hat{\sigma}^6} \times \Big\{ (V_i - \hat{\mu})^2 + \hat{\sigma}^2 - (V_i - \hat{\mu}) \times A_i(\hat{\mu}, \hat{\sigma}) \Big\} (V_i - \hat{\mu}) A_i(\hat{\mu}, \hat{\sigma}) \Bigg]
\end{aligned} \tag{24}$$

In (3), (4) and (5), we have to evaluate $A_i(\mu, \sigma)$.

The next formula ([2], p. 932) is useful because $F(v; \mu, \sigma) = \Phi((v - \mu)/\sigma)$.

$$\begin{aligned}
\Phi(x) &\stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{ -\frac{t^2}{2} \right\} dt \\
&\approx 1 - \frac{1}{2} \left\{ 1 + \sum_{i=1}^{6} d_i x^i \right\}^{-16}
\end{aligned} \tag{25}$$

where $d_1 = 0.0498673470$, $d_2 = 0.0211410061$, $d_3 = 0.0032776263$, $d_4 = 0.0000380036$, $d_5 = 0.0000488906$, $d_6 = 0.0000053830$.

3.2 NUMERICAL EXAMPLE

We give a numerical example below, see Table 1.

Then, from (3), (4), and (5)

$$\hat{\mu}^{(k+1)} = \frac{1}{30} \left[\sum_{\substack{i=1 \\ I(i)=1}}^{30} V_i + \sum_{\substack{i=1 \\ I(i)=0}}^{30} \left\{ \hat{\mu}^{(k)} + A_i(\hat{\mu}^{(k)}, \hat{\sigma}^{(k)}) \right\} \right] \tag{26}$$

$$\hat{\sigma}^{(k+1)} = \Bigg[\frac{1}{30} \Bigg\{ \sum_{\substack{i=1 \\ I(i)=1}}^{30} (V_i - \hat{\mu}^{(k+1)})^2 + \sum_{\substack{i=1 \\ I(i)=0}}^{30} \Big\{ (\hat{\mu}^{(k)} - \hat{\mu}^{(k+1)})^2 + \big(\sigma^{(k)}\big)^2 \Big(v_i + \hat{\mu}^{(k)} - 2\hat{\mu}^{(k+1)} \Big) A_i(\hat{\mu}^{(k)}, \hat{\sigma}^{(k)}) \Big\} \Bigg\} \Bigg]^{1/2} \tag{27}$$

Table 1. An example of new up-and-down data.

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$I(i)$	0	0	1	0	1	0	1	1	1	0	0	0	1	1	0
V_i			40.7		35.1		40.6	39.4	37.0				39.9	40.6	
v_i	39.0	41.0		41.0		41.0				37.0	39.0	41.0			39.0
i	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
$I(i)$	1	0	0	1	1	1	0	0	0	1	1	1	0	0	1
V_i	40.4			40.4	40.5	34.8				35.8	39.4	38.9			40.8
v_i		39.0	41.0				37.0	39.0	41.0				37.0	39.0	

$$-\left(\frac{\partial^2 \ln L(\mu,\sigma)}{\partial \mu^2}\right)_{(\mu,\sigma)=(\hat{\mu},\hat{\sigma})} = \frac{1}{\hat{\sigma}^4}\left\{\sum_{\substack{i=1\\I(i)=1}}^{30}(V_i-\hat{\mu})^2+\sum_{\substack{i=1\\I(i)=0}}^{30}A_i(\hat{\mu},\hat{\sigma})\right\} \quad (28)$$

$$-\left(\frac{\partial^2 \ln L(\mu,\sigma)}{\partial \mu \partial \sigma}\right)_{(\mu,\sigma)=(\hat{\mu},\hat{\sigma})} = -\frac{1}{\hat{\sigma}^5}\sum_{\substack{i=1\\I(i)=0}}^{30}\left\{(V_i-\hat{\mu})^2+\hat{\sigma}^2-(V_i-\hat{\mu})A_i(\hat{\mu},\hat{\sigma})\right\}A_i(\hat{\mu},\hat{\sigma}) \quad (29)$$

$$-\left(\frac{\partial^2 \ln L(\mu,\sigma)}{\partial \sigma^2}\right)_{(\mu,\sigma)=(\hat{\mu},\hat{\sigma})} = \frac{30}{\hat{\sigma}^2} - \frac{1}{\hat{\sigma}^6}\sum_{\substack{i=1\\I(i)=0}}^{30}\left\{(V_i-\hat{\mu})^2+\hat{\sigma}^2-(V_i-\hat{\mu})A_i(\hat{\mu},\hat{\sigma})\right\}\times (V_i-\hat{\mu})A_i(\hat{\mu},\hat{\sigma}) \quad (30)$$

Using these, we obtain

$$(\hat{\mu},\hat{\sigma}) = (40.4,\ 2.48) \quad (31)$$

and

$$-\left(\frac{\partial^2 \ln L(\mu,\sigma)}{\partial \mu^2}\right) = 3.72 \quad (32)$$

$$-\left(\frac{\partial^2 \ln L(\mu,\sigma)}{\partial \mu \partial \sigma}\right) = -1.77 \quad (33)$$

and

$$-\left(\frac{\partial^2 \ln L(\mu,\sigma)}{\partial \sigma^2}\right) = 5.50 \quad (34)$$

on the point. Therefore, the asymptotic standard errors of $\hat{\mu}$ and $\hat{\sigma}$ are 0.563 and 0.463, respectively.

4 SUMMARY AND REMARKS

In this paper we perform the error analysis of the new and conventional up-and-down methods and show the following on the new method.

1. The estimation errors are seldom affected by the magnitude of the up-and-down distance.
2. The estimation errors are still smaller than those of the conventional method. The analysis shown here is about only the case in which the sample size equals 40. However, note that there is no difference between the tendencies of the analyzed errors and those of the asymptotic errors that have sufficiently large sample size.

By using the EM-algorithm, we give the simple iterative scheme to obtain the MLE of the mean and the standard deviation parameters in the new up-and-down method. This fact may help experimenters use the new up-and-down method, which has much better properties than the conventional one.

Because the EM-algorithm has excellent global convergence properties and almost always gives a local maximizer, the iterative formulas we give almost always provide the MLE of the mean and standard deviation parameters. Furthermore, the formulas are very simple. In addition, we also give the expressions to obtain the observed information matrix.

This paper concerns only the case in which all the breakdowns in an experiment occur before the imposed voltages come up to the expected peak voltages. Therefore, the result here can not be used directly if some breakdown occurs around the tail of the waveform. In such a case [5] gives a further consideration.

REFERENCES

[1] International Electrotechnical Commission 60-1 HV test techniques. 1989.

[2] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, 1970.

[3] W. J. Dixon and A. M. Mood, "A method for obtaining and analyzing sensitivity data", J. Amer. Statist. Assoc., Vol. 43, pp. 109–126, 1948.

[4] H. Hirose and K. Kato, "A new analysis of the up-and-down method", Trans. IEE of Japan, Vol. 118-A, (**10**), pp. 1087–1093, 1998 (in Japanese).

[5] H. Hirose and Y. Komori, "A remark on the new up-and-down method analysis", Trans. IEE of Japan, Vol. 119, (**4**), pp. 527–528, 1999 (in Japanese).

[6] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. John Wiley & Sons, New York, 1997.

Manuscript was received on 30 August 1999, in revised form 1 June 2000.