

Determinations Based on Duplication of Readings

Janace A. Speckman

(February 25, 1964)

This paper is concerned with a statistical estimation procedure in which measurements of a quantity are taken until two identical readings are obtained; this duplicated value is then taken as the estimate of the magnitude of the quantity concerned. The properties of this estimation procedure have been investigated numerically, under the assumptions that the individual observations are rounded values of measurements which have a normal distribution, and this estimator is compared with the arithmetic mean of two observations. It is shown that an arithmetic mean of two observations from the rounded distribution is almost always superior to the estimator described above. The exception is where the rounding interval is so wide and the rounding lattice is so advantageously placed that the only real reason for taking repeat measurements would be as a protection against gross errors.

The 1961 *Book of ASTM Standards* [3, p. 1131] gives tentative methods for calibrating upright tanks. These include a proposed method for obtaining circumferential measurements on certain types of tanks which consists, briefly, of wrapping the measuring tape around the tank (at some specified position) with tension applied to the tape, taking a reading to the nearest 0.005 foot, releasing then reapplying the tension, taking another reading, etc., until two identical readings are obtained. The value of the equal readings is then recorded as the circumferential measurement.

The method of estimation used above is familiar to all of us, for it is the method used whenever we count a (finite) number of things. We count them twice and if the two counts agree then we go no further; if they do not agree, then the items are counted again, etc., until two answers agree. The difference between the situations is that in the case of the circumferential measurements we can postulate a continuous distribution underlying the measurement process, so that a recorded observation is a rounded value of a continuous variable, whereas in the case of counting, the distribution of counts is discrete, with incorrect counts corresponding to actual mistakes.

We shall consider here the continuous case; more specifically, we shall look into the statistical properties of such an estimation procedure when the underlying distribution is normal. It will be shown that an arithmetic mean of two observations from the rounded distribution is almost always superior to the estimator described above for the range of cases considered. The exception occurs only when the rounding interval is so wide and the rounding lattice is so advantageously placed that taking repeat measurements serves simply as a protection against gross errors.

Since the placement of the rounding lattice relative to the true value is usually unknown, the occurrence of the exceptional case, when it occurs, is not known. If the rounding interval is wide and the rounding lattice

happens to be disadvantageously placed, estimation by duplication is very much inferior to the arithmetic mean of two observations.

Effect of Grouping

The distribution of a single measurement depends on the width of the rounding interval, and also on the placement of the rounding lattice with respect to the true value. We shall assume that measurements are obtained as rounded values of the (continuous) random variable X , which is normally distributed about the true value μ of the property under consideration, with standard deviation σ . The variable X is rounded to the nearest value X_R in a rounding lattice where the interval, centered on X_R , is of length $2Q\sigma$ and the rounding lattice is placed so that the lower boundary of the interval containing μ is at $\mu + D\sigma$ ($-2Q < D \leq 0$). For simplicity, and with no loss of generality, we let $\mu = 0$ and $\sigma = 1$. Then the distribution of X_R is given by

$$\text{Prob} \{X_R = x_{Ri}\} = \text{Prob} \{x_{Ri} - Q \leq X \leq x_{Ri} + Q\},$$

$$= \int_{x_{Ri}-Q}^{x_{Ri}+Q} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz,$$

where $x_{Ri} = D + (2i + 1)Q$, $i = 0, \pm 1, \pm 2, \dots$. The distribution of X_R will be completely specified when Q and D are given. We note that the distribution of X_R is discrete, that the mean of X_R is not necessarily zero (it is dependent on the position of the rounding lattice), and that the variance of X_R is always greater than the variance of X . (See Eisenhart, Hastay, and Wallis [1],¹ ch. 4). Only a finite number of values of X_R will have probability realistically different from zero (although, theoretically, there would be an infinite number of them). Strictly speaking, the analysis in this paper treats the normal distribution truncated to

¹ Figures in brackets indicate the literature references at the end of this paper.

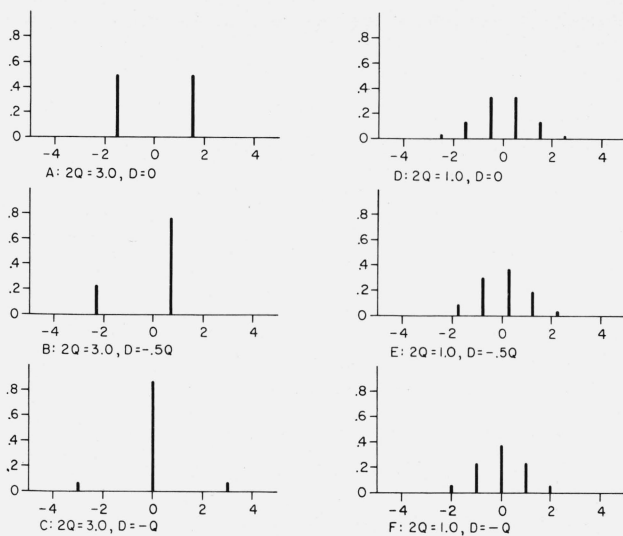


FIGURE 1. Distributions of X_R for selected values of $2Q$ and D .

the interval $(\mu - 5\sigma, \mu + 5\sigma)$. In fact, the numerical results are believed to be correct for the general case.

In studying the properties of this method of estimation by duplication, seven interval lengths were considered: $2Q = 3\sigma, 2\sigma, 1.5\sigma, \sigma, 0.75\sigma, 0.5\sigma$, and 0.25σ ; and 5 positions of the rounding lattice: $D = 0, -0.25Q, -0.5Q, -0.75Q$, and $-Q$. This range of D is sufficient for this study since the distributions of X_R for $-2Q < D < -Q$ are mirror images of the distributions for $-Q < D < 0$. Figure 1 illustrates the nature of the distributions from which observations will be taken. Shown are the "best" (i.e., most advantageous) case, where $D = -Q$, the "worst" case, where $D = 0$, and an intermediate case, $D = -0.5Q$, for $2Q = 3.0$ and $2Q = 1.0$. Table 1 shows the mean and variance of the distribution of X_R and the number, m , of rounding intervals (i.e., the number of values of X_R) necessary to cover the range -5σ to 5σ of the normal curve for all cases considered here.

Distribution of T

Having specified the distribution of a single measurement, we may now turn to the estimator under consideration which will be denoted by T . T is the common value of the first two identical measurements in a sequence of measurements. Obviously the sample size N required to obtain identical measurements, as well as their common value T , is a random variable. N can take on the values $2, 3, \dots$. For the purpose of computation, the number of intervals has been limited to m , so that the range of N becomes $2, 3, \dots, m + 1$. Let

$$p_i = \text{Prob} \{X_R = x_{Ri}\} = \frac{1}{\sqrt{2\pi}} \int_{x_{Ri}-Q}^{x_{Ri}+Q} e^{-z^2/2} dz.$$

Then, the joint probability that $T = x_{Ri} = t_i$ and $N = n$ is given by

$$P \{t_i, n\} = (n-1)! p_i^2 \sum \dots \sum p_j p_k \dots p_l, \quad (1)$$

TABLE 1.—Characteristics of the distribution of X_R : mean, variance, and number of intervals necessary to cover the range -5σ to 5σ of the normal curve

$2Q \backslash D$	0	$-0.25Q$	$-0.5Q$	$-0.75Q$	$-Q$
3.00	0* 2.2986 4	0.0754 2.1318 4	0.1065 1.7380 4	0.0753 1.3569 5	0 1.2027 5
2.00	0 1.3650 6	0.0032 1.3557 6	0.0046 1.3333 6	0.0032 1.3109 6	0 1.3016 5
1.50	0 1.1882 8	0.0001 1.1880 8	0.0001 1.1875 8	0.0001 1.1870 7	0 1.1868 7
1.00	0 1.0833 10	0.0000 1.0833 11	0.0000 1.0833 11	0.0000 1.0833 11	0 1.0833 11
0.75	0 1.0469 14	0.0000 1.0469 14	0.0000 1.0469 14	0.0000 1.0469 15	0 1.0469 15
0.50	0 1.0208 20	0.0000 1.0208 21	0.0000 1.0208 21	0.0000 1.0208 21	0 1.0208 21
0.25	0 1.0052 40	0.0000 1.0052 41	0.0000 1.0052 41	0.0000 1.0052 41	0 1.0052 41

*In each cell of this table, the upper entry is the mean, the middle entry is the variance, and the lower entry is the number of intervals; and "0" means zero exactly, while 0.0000 indicates that the value is zero to at least 4 decimal places.

where the summation is over all $(n-2)$ -fold products such that $j < k < \dots < l$ and $j, k, \dots, l \neq i$. For $n = 2$ and 3 (1) simplifies to

$$P \{t_i, 2\} = p_i^2$$

and

$$P \{t_i, 3\} = 2 p_i^2 (1 - p_i) \quad (1a)$$

By summing the quantities (1) and (1a) over all values of N we obtain the distribution of T independent of the value of N , from which we may deduce some of the properties of this measurement procedure. Figures 2A and 3A show the distribution of T for selected values of $2Q$ and D . The means and variances of T ,

$$E(T) = \sum_{i=1}^m \sum_{n=2}^{m+1} t_i P \{t_i, n\}$$

and

$$\text{Var}(T) = \sum_{i=1}^m \sum_{n=2}^{m+1} t_i^2 P \{t_i, n\} - [E(T)]^2,$$

for all combinations of Q and D considered are given in table 2. (These and other tabled values are believed to be correct to the accuracy given, taking into account errors due to truncation of the normal distribution and due to rounding during calculation.) Note that the $E(T)$, which are biases of T as an estimator of μ are considerably larger than the $E(X_R)$ (c.f., table 1) for large intervals except in the symmetric cases ($D = 0$ and $-Q$) where $E(T)$ and $E(X_R)$ are identically zero.

Expected values of T conditional on $N = n$, $E(T|n)$, and variances of T conditional on the value of N , $\text{Var}(T|n)$, were also calculated, but they are not re-

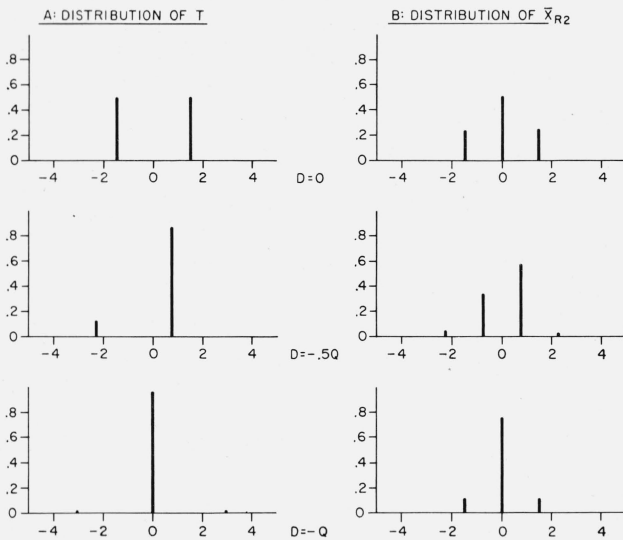


FIGURE 2. Distributions of T and the corresponding distributions of \bar{X}_{R2} for $2Q=3.0$ and selected values of D .

ported here. It turned out that, for the non-symmetric positions of the rounding lattice, the conditional biases, $E(T|n)$, are generally larger than the corresponding expected values of X_R and depend not only on the size and placement of the interval but also on the sample size, n . But more disconcerting is the fact that $\text{Var}(T|n)$ increases rather than decreases as n increases. Thus one would be better off with an estimate obtained with $N=2$ than with a larger N . But there is no control over the sample size since N , too, is a random variable. It can be shown that for the maximum value of N , $N=m+1$, the (conditional) distribution of T is identically the distribution of X_R . Thus an estimate based on the maximum sample size is no better than a single observation from the rounded distribution. While $\text{Prob}\{N=m+1\} \ll 0.0005$, this still is not a happy situation.

TABLE 2. Values of the mean and variance of the (unconditional) distribution of T for selected values of $2Q$ and D

$2Q \backslash D$	0	$-0.25Q$	$-0.5Q$	$-0.75Q$	$-Q$
3.0	0* 2.250	0.259 1.850	0.351 1.049	0.238 0.453	0 0.258
2.0	0 1.038	0.032 0.982	0.046 0.848	0.032 0.717	0 0.664
1.5	0 0.743	-0.002 0.741	-0.002 0.736	-0.002 0.732	0 0.730
1.0	0 0.652	0.000 0.652	0.000 0.652	0.000 0.652	0 0.652
0.75	0 0.618	-0.000 0.618	-0.000 0.618	-0.000 0.618	0 0.618
0.5	0 0.589	0.000 0.589	0.000 0.589	0.000 0.589	0 0.589
0.25	0 0.560	0.000 0.560	0.000 0.560	0.000 0.560	0 0.560

*In each cell of this table, the upper entry is the mean and the lower entry is the variance; and "0" means zero exactly, while 0.000 indicates that the value is zero to at least 3 decimal places.

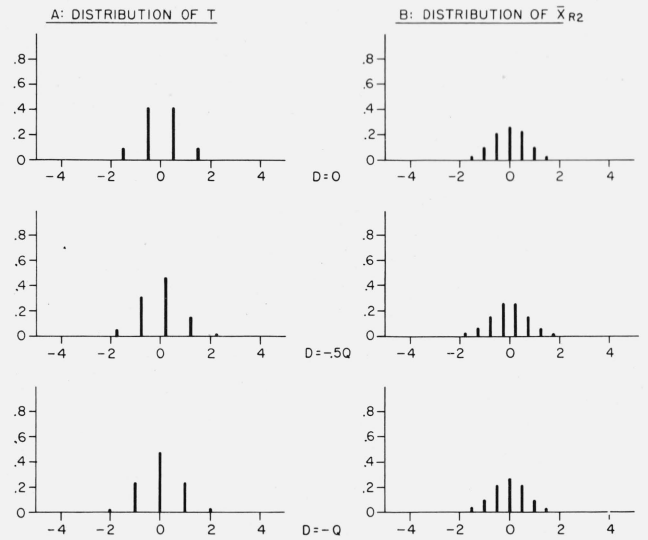


FIGURE 3. Distributions of T and the corresponding distributions of \bar{X}_{R2} for $2Q=1.0$ and selected values of D .

By summing the quantities (1) and (1a) over all values of i we obtain the distribution of N . The probabilities,

$$\text{Prob}\{N=n\} = \sum_{i=1}^m P\{t_i, n\},$$

and the means and variances of N ,

$$E(N) = \sum_{n=2}^{m+1} n \text{Prob}\{N=n\}$$

and

$$\text{Var}(N) = \sum_{n=2}^{m+1} n^2 \text{Prob}\{N=n\} - [E(N)]^2,$$

are given in table 3 for the various distributions of X_R . Probabilities less than 0.0005 are not reported. We see that, for the range of intervals considered here, one would ordinarily expect to take from 3 to 6 observations to obtain an estimate by this procedure. Only when $2Q=3.0$ and the rounding lattice is advantageously placed would one expect to obtain an estimate at $N=2$ —but in these cases the only real reason for taking repeat observations would be as a protection against gross errors.

Comparison of T With \bar{X}_{R2}

Since the arithmetic mean is the best estimator of the mean of a normal distribution, it is reasonable to compare T with this fixed sample size estimator. The arithmetic mean of two observations from the distribution of X_R , \bar{X}_{R2} , is chosen for comparison with T because, as will be seen below it is almost always a better estimator of μ than is T for the range of Q and D considered here.

TABLE 3. Values of Prob $\{N=n\}$, and of the mean and variance of the distribution of N , for selected values of $2Q$ and D

$N \backslash D$	0	$-0.25Q$	$-0.5Q$	$-0.75Q$	$-Q$
$2Q=3.0$					
2.....	0.497	0.537	0.631	0.722	0.760
3.....	.499	.457	.356	.258	.217
4.....	.004	.006	.013	.020	.023
$E(N)$	2.507	2.470	2.382	2.298	2.264
$\text{Var}(N)$	0.258	0.262	0.262	0.249	0.241
$2Q=2.0$					
2.....	0.457	0.465	0.486	0.507	0.516
3.....	.478	.464	.429	.394	.379
4.....	.062	.068	.083	.098	.104
5.....	.003	.003	.002	.001	.001
$E(N)$	2.611	2.608	2.601	2.594	2.591
$\text{Var}(N)$	0.379	0.391	0.418	0.445	0.457
$2Q=1.5$					
2.....	0.384	0.385	0.388	0.390	0.391
3.....	.441	.438	.429	.420	.416
4.....	.153	.157	.165	.174	.177
5.....	.021	.020	.018	.016	.015
$E(N)$	2.812	2.813	2.815	2.817	2.818
$\text{Var}(N)$	0.588	0.590	0.594	0.599	0.600
$2Q=1.0$					
2.....	0.271	0.271	0.271	0.271	0.271
3.....	.373	.373	.372	.372	.372
4.....	.252	.252	.253	.253	.254
5.....	.090	.089	.089	.088	.088
6.....	.014	.014	.014	.015	.015
7.....	.001	.001	.001	.001	.001
$E(N)$	3.206	3.206	3.206	3.206	3.206
$\text{Var}(N)$	0.980	0.979	0.979	0.979	0.979
$2Q=0.75$					
2.....	0.207	0.207	0.207	0.207	0.207
3.....	.315	.315	.315	.315	.315
4.....	.271	.271	.271	.271	.271
5.....	.147	.147	.147	.147	.147
6.....	.050	.050	.050	.050	.050
7.....	.010	.010	.010	.010	.010
8.....	.001	.001	.001	.001	.001
$E(N)$	3.551	3.551	3.551	3.551	3.551
$\text{Var}(N)$	1.389	1.389	1.389	1.389	1.389
$2Q=0.5$					
2.....	0.140	0.140	0.140	0.140	0.140
3.....	.234	.234	.234	.234	.234
4.....	.248	.248	.248	.248	.248
5.....	.193	.193	.193	.193	.193
6.....	.114	.114	.114	.114	.114
7.....	.050	.050	.050	.050	.050
8.....	.016	.016	.016	.016	.016
9.....	.004	.004	.004	.004	.004
10.....	.001	.001	.001	.001	.001
$E(N)$	4.145	4.145	4.145	4.145	4.145
$\text{Var}(N)$	2.248	2.248	2.248	2.248	2.248
$2Q=0.25$					
2.....	0.070	0.070	0.070	0.070	0.070
3.....	.129	.129	.129	.129	.129
4.....	.165	.165	.165	.165	.165
5.....	.172	.172	.172	.172	.172
6.....	.155	.155	.155	.155	.155
7.....	.122	.122	.122	.122	.122
8.....	.085	.085	.085	.085	.085
9.....	.052	.052	.052	.052	.052
10.....	.028	.028	.028	.028	.028
11.....	.013	.013	.013	.013	.013
12.....	.006	.006	.006	.006	.006
13.....	.002	.002	.002	.002	.002
14.....	.001	.001	.001	.001	.001
$E(N)$	5.511	5.511	5.511	5.511	5.511
$\text{Var}(N)$	4.957	4.957	4.957	4.957	4.957

Note that the mean of \bar{X}_{R2} is the same as the mean of X_R and that $\text{Var}(\bar{X}_{R2}) = \text{Var}(X_R)/2$. Figures 2B and 3B show the distributions of \bar{X}_{R2} for $2Q=3.0$ and 1.0 and $D=0, -0.5Q$, and $-Q$. We see that the spacing between possible values of the estimates has decreased to half of the width of the original rounding interval so that if the closest value is not obtained, the size of the miss need not be as large as for T and the probability of a large miss is smaller.

For comparing T with X_{R2} , relative efficiency will be used:

$$\text{Efficiency of } T \text{ relative to } \bar{X}_{R2} = \frac{\text{Var}(X_R)/2 + [E(X_R)]^2}{\text{Var}(T) + [E(T)]^2}$$

(which is usually expressed as a percentage). The relative efficiencies are given in table 4. We see that only in two instances does T show superior behavior over \bar{X}_{R2} — where $2Q=3.0$, $D=-0.75Q$, and $-Q$ — as indicated by relative efficiencies of 134.3 percent and 232.9 percent respectively. The explanation for these high efficiencies lies in the fact that, in such cases, if an estimate is obtained at $N=2$ — which, by reference to table 3, happens more than 70 percent of the time — that estimate is almost sure to be the value in the interval containing $\mu=0$. This leads to a very small variance for the distribution of T conditional on $N=2$ which offsets the larger variances for $N>2$. Actually, this effect is also working at the same positions of the rounding lattice for $2Q=2.0$ and $2Q=1.5$ but not to the same extent. In all these cases, the conditional variance of T for $N=2$ is smaller than the variance of an arithmetic mean of 2 observations, so that it is possible to obtain better results using estimation by duplication, although these circumstances are limited and not within the control of the observer. For $2Q \leq 1.0$, T is never superior to the mean of 2 observations.

TABLE 4. Efficiency of T relative to the arithmetic mean of two observations for selected values of $2Q$ and D

$2Q \backslash D$	0	$-0.25Q$	$-0.5Q$	$-0.75Q$	$-Q$
	%	%	%	%	%
3.0	51.1	55.9	75.1	134.3	232.9
2.0	65.8	69.0	78.4	91.3	98.1
1.5	80.0	80.2	80.6	81.1	81.3
1.0	83.1	83.1	83.1	83.1	83.1
0.75	84.7	84.7	84.7	84.7	84.7
.5	86.7	86.7	86.7	86.7	86.7
.25	89.8	89.8	89.8	89.8	89.8

Comparison of T and \bar{X}_{R2} When the True Value Is Considered To Be a Random Variable

We have seen what results can be expected using estimation by duplication on a single object when the true value of the quantity being estimated is at zero and the rounding lattice is at certain fixed positions relative to zero. If we consider the rounding lattice to be placed at random — which is equivalent to having the rounding lattice centered on zero and assuming that the true values of the quantities are uniformly dis-

tributed between $-Q$ and Q [1, ch 4]—it is apparent that at least for the interval sizes considered here, one would be better off using the mean of two observations as an estimate of the true value rather than the repeated value; for even in the case $2Q=3.0$ one would expect an efficiency of T relative to \bar{X}_{R2} greater than 100 percent only about one-third of the time.

Values for quantities similar to those given in tables 1 through 4 can be calculated under the new assumptions. We note first that T is unbiased under these assumptions, due to symmetry, and that \bar{X}_{R2} is also unbiased for the same reason. Youden, Connor, and Severo [2] have calculated the probability that $N=2$ for intervals of length $2Q=3.0, 2.0,$ and 1.0 . These probabilities are 0.6296, 0.4860, and 0.2709, respectively. Since estimation by duplication yields results superior to the taking of an arithmetic mean only when $N=2$, as indicated by the conditional variance being smaller than the variance of the mean of two observations, we may use these probabilities to obtain an estimate of the probability of obtaining better results using estimation by duplication. The distribution of T conditional on $N=2$ for $2Q=3.0$ and 2.0 has variance smaller than $\text{Var}(X_R)/2$ for almost all values of D between $-0.5Q$ and $-Q$. Thus if $2Q=3.0$, the probability of obtaining better results with T is approximately $(0.6296)(0.5)=0.3148$; and if $2Q=2.0$, the probability is approximately $(0.4860)(0.5)=0.2430$. Since the behavior of T is never better than X_{R2} for $2Q=1.0$, the probability of obtaining better results with T is zero.

While the other quantities, such as the mean and variance of the marginal distribution of T , may be of interest, they are difficult to obtain under the new assumptions to any accuracy and would only point more to the fact that estimation by duplication is not a good estimation procedure to use when the underlying dis-

tribution is normal. Rough estimates of some of the quantities may be obtained by averaging the appropriate values in the tables given.

From the preceding discussion it follows that when the true value of the quantity to be measured is considered to be uniformly distributed in an interval of length $2Q\sigma$ and measurements of that quantity are normally distributed about its true value with standard deviation σ , then for $2Q \leq 3.0$, the probability is at most 0.3148 that estimation by duplication is better than the arithmetic mean of two observations. For $2Q \leq 1.0$ better results can always be obtained with the arithmetic mean of only two observations.

In conclusion it appears that the practice of taking readings until two identical readings are obtained cannot be justified since the average of the first two readings almost always yields a better estimate of the measured quantity.

The author thanks Churchill Eisenhart for suggesting this investigation and for his guidance, and Joseph M. Cameron and Joan R. Rosenblatt for their helpful suggestions for writing this paper.

References

- [1] Eisenhart, C., Hastay, M. W., and Wallis, W. A., ed, Selected Techniques of Statistical Analysis. New York: McGraw-Hill Book Company, Inc., 1947.
- [2] Youden, W. J., Connor, W. S., and Severo, N. C., Measurements made by matching with known standards, *Technometrics* **1** (1959), 101-109.
- [3] 1961 Book of ASTM Standards, pt 7, Petroleum products and lubricants. Philadelphia: American Society for Testing and Materials, 1961.

(Paper 68B2-115)