REFERENCES

1. Gentzen, Gerhardt, *Untersuchungen über das logische Schliessen*, Mathematische Zeitschrift, 39 (1934-35), 176-210, 405-431.
2. Hilbert, D., and Ackerman, W., *Principles of Mathematical Logic*, Chelsea Publishing Company, New York, 1950.
3. Kleene, S. C., *Mathematical Logic*, John Wiley & Sons, Inc., New York, 1967.
4. Mendelson, Elliott, *Introduction to Mathematical Logic*, D. Van Nostrand Company, Inc., New York, 1964.
5. Nicod, J. G., *A Reduction in the Number of Primitive Propositions of Logic*, Proc. Camb. Phil. Soc., 19 (1917), 32-41.
6. Rosser, J. B., *Logic for Mathematicians*, New York, 1953.
7. Whitehead, A. N., and Russell, B., *Principia Mathematica*, Vol. I, University Press, Cambridge, 1925.

MATCHING PRIZE FUND

If your chapter presents awards for outstanding mathematical papers or student achievement in mathematics, you may apply to the National Office to match the amount spent by your chapter. For example, $30 of awards can result in the chapter receiving $15 reimbursement from the National Office. These funds may also be used for the rental of mathematical films. Write to the National Office for more details.

REGIONAL MEETINGS OF MAA

Many regional meetings of the Mathematical Association regularly have sessions for undergraduate papers. If two or more colleges and at least one local chapter help sponsor or participate in such undergraduate sessions, financial help is available up to $50 for one local chapter to defray postage and other expenses. Send requests to:

R. V. Andree
Secretary-Treasurer, Pi Mu Epsilon
601 Elm Avenue, Room 423
The University of Oklahoma
Norman, Oklahoma 73069

A RELATIONSHIP BETWEEN APPROXIMATION THEORY AND STATISTICAL MEASUREMENTS

by Herbert L. Dershem
Hope College

1. Introduction.

There are many statistics which are used to measure central tendency and dispersion of a random variable. Each of several measures of central tendency is shown to be a best approximation to the variable over the set of all constant functions in a specified norm. The error of this best approximation serves as a corresponding measure of dispersion.

Hamming ([1], pp. 224-226) has noted this relationship between best approximation and measurements of central tendency. His work is extended here to include additional norms, measurements of dispersion, continuous distributions, and measurements which result from transformations of best approximations.

2. Discrete Random Variables.

Given a discrete random variable x which can take on any of a set of data values $\{x_i\}_{i=1}^{M}$, each value x_i occurring with respective probability f_i, and a given function norm $\|\ \|$, we define a *measurement of central tendency* to be the constant value c^* which, when substituted for c, minimizes

$$\|x - c\|.$$

This is the best approximation to x over the set of constant functions where x is the function defined on the discrete domain $\{1,2,\cdots,M\}$ by $x(i) = x_i$ and c is the constant function defined on the same domain by $c(i) = c$. We examine a number of choices of norm and show that these choices all correspond to commonly used measurements of central tendency. The error in the best approximation, $\|x - c^*\|$, is defined as the *corresponding measurement of dispersion.*

The discrete L_2 norm with weight function f is defined by

$$\|u\|_{2,f} = \left\{\sum_{i=1}^{M} f_i u_i^2\right\}^{1/2}.$$

The measurement of central tendency associated with this norm is found by minimizing the function defined by

$$\phi(c) = \|x - c\|_{2,f} = \left\{\sum_{i=1}^{M} f_i(x_i - c)^2\right\}^{1/2} .$$

The minimizing value of c, which we call $E(x)$ (the expected value of x), is found by setting the first derivative of ϕ with respect to c equal to zero and solving the resulting equation for c^* to obtain

$$c^* = E(x) = \frac{\sum_{i=1}^{M} f_i x_i}{\sum_{i=1}^{M} f_i} = \sum_{i=1}^{M} f_i x_i .$$

We have used the fact that $\Sigma f_i = 1$, that is, the sum of the probabilities is 1. This measurement is well known as the arithmetic mean or expected value of the random variable. The corresponding measurement of dispersion is the error of this expected value, given by

$$\|x - E(x)\|_{2,f} = \left\{\sum_{i=1}^{M} f_i(x_i - E(x))^2\right\}^{1/2}$$

which is the standard deviation of the random variable.

The discrete L_1 norm with weight function f is defined by

$$\|u\|_{1,f} = \sum_{i=1}^{M} f_i|u_i| .$$

We show that the median is the measurement of central tendency associated with this norm. For ease of notation we define $S(a,b) = \{i \mid x_i \in (a,b)\}$ with similar definitions for semi-open intervals. If m is the median of the random variable defined by x and f, and ε is any positive number, then

$$\phi(m + \varepsilon) - \phi(m) = \|x - (m + \varepsilon)\|_{1,f} - \|x - m\|_{1,f}$$

$$= \sum_{S(-\infty,m]} f_i(m + \varepsilon - x_i) + \sum_{S(m,\, m+\varepsilon]} f_i(m + \varepsilon - x_i)$$

$$+ \sum_{S(m+\varepsilon,\, \infty)} f_i(x_i - m + \varepsilon) - \sum_{S(-\infty,m]} f_i(m - x_i)$$

$$- \sum_{S(m,\, m+\varepsilon]} f_i(x_i - m) - \sum_{S(m+\varepsilon,\, \infty)} f_i(x_i - m)$$

$$= \varepsilon \sum_{S(-\infty,m]} f_i + \sum_{S(m,\, m+\varepsilon]} f_i(2m - 2x_i + \varepsilon) - \varepsilon \sum_{S(m+\varepsilon,\, \infty)} f_i .$$

But for x_i in the interval $(m, m + \varepsilon]$,

$$2m - 2x_i + \varepsilon \geq -\varepsilon .$$

Therefore,

$$\phi(m + \varepsilon) - \phi(m) \geq \varepsilon \sum_{S(-\infty,m]} f_i - \varepsilon \sum_{S(m,\infty)} f_i \geq 0$$

where the last inequality follows from the definition of the median. A similar argument can be used to show that $\phi(m - \varepsilon) - \phi(m) \geq 0$. Hence, m is a minimum of ϕ and a measurement of central tendency with respect to the discrete weighted L_1 norm. In the case where the median is not one of the data values, this best approximation is not unique. The corresponding measurement of dispersion is

$$\|x - m\|_{1,f} = \sum_{i=1}^{M} f_i|x_i - m| ,$$

the mean deviation.

Another common measurement of central tendency is derived from the discrete L_∞ norm which is defined by

$$\|u\|_\infty = \max_{1\leq i\leq M} |u_i| .$$

In this case we wish to find the value of c which minimizes

$$\|x - c\|_\infty = \max_{1\leq i\leq M} |x_i - c| .$$

It is easy to determine that the minimizing value must be located midway between the maximum and minimum values which x can take on, a value known as the *midrange*. If we denote the midrange by m_r, the corresponding measurement of dispersion is

$$\|x - m_r\|_\infty = \max_{1\leq i\leq M} |x_i - m_r| ,$$

which is half of the range.

We define one final discrete norm by

$$\|u\|_{m,f} = \sum_{i=1}^{M} f_i[1 - \delta(u_i)] ,$$

where

$$\delta(x) = \begin{cases} 0 \text{ if } x \neq 0 \\ 1 \text{ if } x = 0 . \end{cases}$$

The value of $\|x - c\|_{m,f}$ is the probability that the random variable is not equal to c, and any value of c which minimizes this is called a *mode*. The corresponding measurement of dispersion is the probability that the random variable is not equal to the mode.

3. *Statistics which are Transformations of Best Approximations.*

A number of measurements of central tendency do not arise directly as best approximations in some norm, but can be found as the transformation of a best approximation. Such measures are defined as follows.

Choose a norm $\| \ \|$ and a transformation function θ. Find c^* such that c^* is the best approximation to $\theta(x)$ in the given norm, that is,

$$\|\theta(x) - c^*\| \leq \|\theta(x) - c\| \text{ for all } c \in (-\infty,\infty).$$

Then a measurement of central tendency is given by $\theta^{-1}(c^*)$.

If the discrete L_2 norm is chosen, then the measurement of central tendency is

$$\theta^{-1}\left(\sum_{i=1}^{M} f_i\theta(x_i)\right) .$$

The geometric mean, harmonic mean, and root-mean-square are examples of this type of measurement when $\theta(x)$ equals $\log x$, $1/x$, and x^2, respectively, and the data values x_i are such that the appropriate function θ is defined.

4. *Continuous Random Variables.*

The above results can be extended to continuous random variables by replacing the discrete norms by the corresponding continuous norms. We now assume that we have a continuous random variable with probability density function $f(x)$. Therefore, f is such that $\int_{-\infty}^{\infty} f(x)dx = 1$ and $f(x) \geq 0$ for all x.

The best approximation in the continuous L_2 norm with weight function f is the constant c which minimizes

$$\phi(c) = \|x - c\|_{2,f} = \left\{\int_{-\infty}^{\infty} f(x)(x - c)^2 dx\right\}^{1/2} .$$

By differentiation, the minimum of ϕ is found to be

$$c^* = \int_{-\infty}^{\infty} xf(x)dx .$$

This value of c^* is the expectation of a continuous random variable with probability density function f. The corresponding measurement of dispersion is

$$\|x - c^*\|_{2,f} = \left\{\int_{-\infty}^{\infty} f(x)(x - c^*)^2 dx\right\}^{1/2} ,$$

the square root of the variance of the random variable.

If we consider the continuous L_1 norm with weight function f, then, since f is always nonnegative, we have

$$\phi(c) = \|x - c\|_{1,f} = \int_{-\infty}^{\infty} f(x)|x - c|dx$$

$$= \int_{-\infty}^{c} f(x)(c - x)dx + \int_{c}^{\infty} f(x)(x - c)dx ,$$

and

$$\phi'(c) = \int_{-\infty}^{c} f(x)dx - \int_{c}^{\infty} f(x)dx$$

$$= \int_{-\infty}^{c} f(x)dx - \int_{-\infty}^{\infty} f(x)dx + \int_{-\infty}^{c} f(x)dx$$

$$= 2\int_{-\infty}^{c} f(x)dx - 1 .$$

Setting $\phi'(c)$ to zero, we find that ϕ is minimized when $\int_{-\infty}^{c} f(x)dx = 1/2$, that is, at that value of c for which the probability is exactly one-half that $x < c$. This is the natural extension of the median to a continuous distribution.

Unless $f(x)$ is zero everywhere outside of some bounded interval, there is no continuous extension of the midrange. If f is zero outside the interval (a,b) and positive somewhere in every neighborhood of a and b, then the continuous L_∞ norm defined by

$$\phi(c) = \|x - c\|_\infty = \sup_{a<x<b} |x - c|$$

has as its minimizing value and measurement of central tendency

$$c^* = (a + b)/2$$

which is the middle value of the interval in which f is non-zero. The corresponding measurement of dispersion is

$$\phi(c^*) = (b - a)/2 .$$

The continuous extension of the mode is found by minimizing

$$\phi(c) = \|x - c\|_{m,f} = 1 - \int_{-\infty}^{\infty} f(x)\delta(x)dx = 1 - f(c)$$

where δ is the Dirac delta function ([2], p. 6). The function ϕ is minimized at those values where f attains its maximum value. If m_0 is such a value, then the corresponding measurement of dispersion is $1 - f(m_0)$.

REFERENCES

1. Hamming, R. W., *Numerical Methods for Scientists and Engineers*, McGraw-Hill, New York, 1962.
2. Stakgold, I., *Boundary Value Problems of Mathematical Physics*, *Volume I*, Macmillan, New York, 1967.

1975 NATIONAL MEETING IN KALAMAZOO

There is still time for local chapters to be making plans for the national meeting at Western Michigan University in Kalamazoo, Michigan in conjunction with the Mathematical Association of America. Plan now to send your best undergraduate speaker or delegate (or both) to that meeting. Travel money for *one* approved speaker or delegate is available from National. Send requests and proposed papers to:

R. V. Andree
Secretary-Treasurer, Pi Mu Epsilon
601 Elm Avenue, Room 423
The University of Oklahoma
Norman, Oklahoma 73069

GENERALIZING BINARY OPERATIONS

by Dennis C. Smolarski
St. Louis University

Most day to day calculations take place within the field of real numbers with the two binary operations of addition and multiplication. In this field, these two operations are definitionally independent of one another. However, if we approach binary operations from a different point of view, e.g. that of recursive formulae, we can develop multiplication from addition by use of the concept of repeated addition. Along similar lines, we can develop exponentiation from multiplication by repeated multiplication. The next logical step would be to try to develop another binary operation based on repeated exponentiation.

Professor D. F. Borrow of the University of Georgia in the *American Mathematical Monthly*, 43 (1936), p. 150, developed some theorems and a notation for repeated exponentiation. As Σ is used for summation and Π is used for products, he used E for repeated exponentiation. The development of a "fourth operation" would depend on all the indexed terms of E being equal, similar to what is necessary in developing multiplication and exponentiation itself.

In order to clarify relations and notations, let us look at addition, multiplication, exponentiation, and a projected new fourth operation in terms of functions and recursive formulae. Let

$$f_1(n,m) = n + m$$

$$f_2(n,m) = n \cdot m$$

and

$$f_3(n,m) = n^m .$$

We know the following:

$$n \cdot m = n + [n \cdot (m - 1)] = \sum_1^m n_i \qquad (\text{where all } n_i = n)$$

and

$$n^m = n \cdot [n^{(m-1)}] = \prod_1^m n_i \qquad (\text{where all } n_i = n) .$$