# ON THE DISTRIBUTIONS OF MIDRANGE AND SEMI-RANGE IN SAMPLES FROM A NORMAL POPULATION

K. C. S. Pillai

*University of Travancore, Trivandrum.*

**1. Summary.** In this paper the simultaneous distribution of midrange and semi-range has been obtained and used to derive the distributions of midrange and semi-range in samples taken from a normal population.

**2. Introduction.** The concept of ordering a sample has given rise to innumerable problems for statistical investigation. Several authors have contributed to the study of ordered individuals and, in particular, to the study of extreme individuals, their sum and difference in samples from a normal population. L. H. C. Tippett [1] has studied the first four moments of the range and has tabled the mean-range for sample size ranging from two to thousand. Student [2] has determined the nature of the distribution of range for particular sample sizes by purely empirical methods. T. Hojo [3] has compared the standard error of midrange to that of median and mean in normal samples. E. S. Pearson and H. O. Hartley [4] have tabled the values of the probability integral of range for sample size up to twenty. E. J. Gumbel [5], [6], [7] has established the independence of the extreme values in large samples from population of unlimited range and obtained the distributions of range and midrange. The asymptotic distribution of range has also been investigated by G. Elfving [8]. J. F. Daly [9] has devised a $t$-test adopting range in place of standard deviation in Student's $t$ and in a modified $t$-test E. Lord [10] has used range instead of standard deviation. An extension to two populations of an analogue of Student's $t$-test using the sample range has been worked out by John E. Walsh [11]. S. S. Wilks [12] has given a complete and detailed account of the researches on order statistics and also a number of suggestions regarding possibilities of utilising order statistics in statistical inference. In this paper the distribution of midrange has been developed as a series and a method of evaluating the probability integral for semi-range based on an infinite series expansion for the normal probability integral has been suggested.

**3. Distributions of midrange and semi-range.** Let

$$x_1 \leq x_2 \cdots \leq x_n$$

be an ordered sample from a normal population with zero mean and unit standard deviation. Then the joint distribution of $x_1$ and $x_n$, the lowest and highest values respectively, is given by [13],

$$p(x_1, x_n) = [n(n-1)/2\pi]\left[\int_{x_1}^{x_n} e^{-t^2/2}\, dt/\sqrt{2\pi}\right]^{n-2} e^{-(x_1^2+x_n^2)/2}. \tag{1}$$

Let

$$M = (x_1 + x_n)/2$$

and

$$W = (x_n - x_1)/2.$$

$M$ is the midrange and $W$ is the semi-range of the sample. From (1) the simultaneous distribution of $M$ and $W$ reduces to

$$p(M, W) = [n(n-1)/\pi]e^{-(M^2+W^2)}\left[\int_{M-W}^{M+W} e^{-t^2/2}\, dt/\sqrt{2\pi}\right]^{n-2}. \tag{2}$$

It has been shown [14] that if

$$F(M, W) = \left[\int_{M-W}^{M+W} e^{-t^2/2}\, dt/\sqrt{2\pi}\right]^{k}, \tag{3}$$

$$F(M, W) = e^{-k(M^2+W^2)/2}[A_0^{(k)} + A_1^{(k)}M^2 + \cdots + A_i^{(k)}M^{2i} + \cdots], \tag{4}$$

where $A_i^{(k)}$ coefficient is given by

$$\begin{aligned} 2iA_i^{(k)} = kA_{i-1}^{(k)} - k\sqrt{2/\pi}[A_{i-1}^{(k-1)}W \\ + A_{i-2}^{(k-1)}W^3/\Gamma(4) + \cdots + A_0^{(k-1)}W^{2i-1}/\Gamma(2i)]. \end{aligned} \tag{5}$$

Using expansion (4) equation (2) reduces to

$$p(M, W) = [n(n-1)/\pi]e^{-n(M^2+W^2)/2} \sum_{i=0}^{\infty} A_i^{(n-2)} M^{2i}. \tag{6}$$

It is evident that the $A$'s involve terms of the form

$$[\phi(W)]^s W^q e^{-mW^2/2}$$

where $s$, $q$, $m$ are positive integers and

$$\phi(W) = \sqrt{2/\pi}\int_0^W e^{-t^2/2}\, dt.$$

Integrating (6) with respect to $W$

$$p(M) = [n(n-1)/\pi]e^{-nM^2/2} \sum_{i=0}^{\infty} B_i M^{2i} \tag{7}$$

where

$$B_0 = \sqrt{\pi/2}\, I(n-2, 0, 2), \tag{8}$$

$$B_1 = [(n-2)/2][\sqrt{\pi/2}\, I(n-2, 0, 2) - I(n-3, 1, 3)], \tag{9}$$

$$B_2 = [(n-2)/2^2\Gamma(3)][\sqrt{\pi/2}\,(n-2)I(n-2, 0, 2)$$

$$- (2n - 5)I(n - 3, 1, 3) - (1/3)I(n - 3, 3, 3) + \sqrt{2/\pi}\,(n - 3)I(n - 4, 2, 4)] \tag{10}$$

where

$$I(s, q, m) = \sqrt{2/\pi} \int_0^\infty [\phi(x)]^s x^q e^{-mx^2/2}\, dx. \tag{11}$$

Using the method of integration by parts, the evaluation of $I(s, q, m)$ can be reduced ultimately to that of $I(p, 0, r)$ and this function for different values of $p$ and $r$ is given in Table I.

TABLE I

*Values of Integrals $I(p, 0, r)$*[1]

| $p$ | $r$ = 2 | 4 | 6 | 8 |
|---|---|---|---|---|
| 1 | 0.277,063,21 | 0.147,583,62 | 0.100,735,97 | 0.076,490,19 |
| 2 | 0.152,980,4 | 0.064,094,20 | 0.037,255,93 | 0.025,060,53 |
| 3 | 0.098,373 | 0.033,453,6 | 0.016,808,71 | |
| 4 | 0.069,10 | 0.019,535,1 | 0.008,589,57 | |
| 5 | 0.051,44 | 0.012,325,5 | | |
| 6 | 0.039,90 | 0.008,223,9 | | |
| 7 | 0.031,94 | | | |
| 8 | 0.026,17 | | | |

The first five $B$ Coefficients for $n$ ranging from 3 to 10 are tabled below.

TABLE II

*Values of $B$ Coefficients.*

| $n$ | $B_0$ | $B_1$ | $B_2$ | $B_3$ | $B_4$ |
|---|---|---|---|---|---|
| 3 | 0.347,247,25 | 0.040,642,87 | 0.002,772,90 | 0.000,133,80 | 0.000,005,00 |
| 4 | 0.191,732 | 0.058,751 | 0.010,906 | 0.001,460 | 0.000,153 |
| 5 | 0.123,292 | 0.067,184 | 0.021,526 | 0.004,988 | 0.000,909 |
| 6 | 0.086,60 | 0.070,93 | 0.033,23 | 0.011,20 | 0.002,97 |
| 7 | 0.064,47 | 0.072,20 | 0.045,65 | 0.020,28 | 0.007,14 |
| 8 | 0.050,01 | 0.072,09 | 0.057,22 | 0.032,21 | 0.014,59 |
| 9 | 0.040,03 | 0.071,27 | 0.068,95 | 0.047,01 | 0.024,98 |
| 10 | 0.032,80 | 0.069,97 | 0.080,31 | 0.064,66 | 0.040,51 |

[1] The integrals have been evaluated by using (14).

The accuracy obtained by keeping the first five terms in $p(M)$ may be judged from the following values of the total probability calculated for small values of $n$.

TABLE III.

*Total probability keeping the first five terms in $p(M)$*

| Size of sample | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|
| Total probability | 0.999,998 | 0.999,92 | 0.999,56 | 0.998,8 | 0.997,8 |

Integrating (6) with respect to $M$, $p(W)$ may be obtained. But $p(W)$ involves integral $\phi(W)$ and to evaluate the integral probability of $W$ expansions for $\phi(W)$ and its powers have to be developed.

Since $$\phi(W) = \sqrt{2/\pi}\int_0^W e^{-t^2/2}\,dt = \sqrt{2/\pi}W(1 - W^2/6 + \cdots),$$

a convenient expansion is given by

$$\sqrt{2/\pi}\int_0^W e^{-t^2/2}\,dt = \sqrt{2/\pi}\,We^{-W^2/6}(1 + a_2W^4 + \cdots + a_iW^{2i} + \cdots) \tag{12}$$

where $a_i$ follows the recurrence relation

$$3(2i + 1)a_i - a_{i-1} = (-1)^i/3^{i-1}\Gamma(i + 1), \tag{13}$$

as may be seen by differentiating (12) with respect to $W$ and equating the coefficient of $W^{2i}$ on both sides. Again

$$[\phi(W)]^j = (2/\pi)^{j/2}e^{-jW^2/6}W^jS^j \tag{14}$$

where

$$S = 1 + a_2W^4 + a_3W^6 + \cdots + a_iW^{2i} + \cdots \tag{15}$$

and

$$S^j = 1 + K_2^{(j)}W^4 + K_3^{(j)}W^6 + \cdots \tag{16}$$

where

$$K_i^{(j)} = \sum_{s=1}^{i} {}^jC_s\, s!\,a_1^{s_1}a_2^{s_2}\cdots a_i^{s_i}/s_1!\,s_2!\cdots s_i! \tag{17}$$

and

$$\begin{aligned} s_1 + 2s_2 + \cdots + is_i &= i,\\ s_1 + s_2 + \cdots + s_i &= s. \end{aligned} \tag{17a}$$

Clearly $a_i = K_i^{(1)}$. In evaluating the $K_i^{(j)}$'s summation with respect to $s$ is first performed, the values of $s_1, s_2, \cdots, s_i$ being obtained so as to satisfy the relations (17a); and thereafter the values of the a's are substituted. It may be noted that $a_1 = 0$. The $K$ coefficients for $j$ up to 8 and $i$ up to 13 are given below.

TABLE IV

$K_i^{(j)}$ *Coefficients.*

| $j$ | $i$ | | | |
|---|---|---|---|---|
| | 2 | 3 | 4 | 5 |
| 1 | $0.011,111,11$ | $-0.0^{3}35,273,369$ | $0.0^{4}44,091,711$ | $-0.0^{5}17,814,833$ |
| 2 | $0.022,222,22$ | $-0.0^{3}70,546,737$ | $0.0^{3}21,164,021$ | $-0.0^{4}11,401,493$ |
| 3 | $0.033,333,33$ | $-0.0^{2}10,582,011$ | $0.0^{3}50,264,550$ | $-0.0^{4}28,860,029$ |
| 4 | $0.044,444,44$ | $-0.0^{2}14,109,348$ | $0.0^{3}91,710,758$ | $-0.0^{4}54,157,091$ |
| 5 | $0.055,555,56$ | $-0.0^{2}17,636,684$ | $0.0^{2}14,550,265$ | $-0.0^{4}87,292,680$ |
| 6 | $0.066,666,67$ | $-0.0^{2}21,164,021$ | $0.0^{2}21,164,021$ | $-0.0^{3}12,826,680$ |
| 7 | $0.077,777,78$ | $-0.0^{2}24,691,358$ | $0.0^{2}29,012,346$ | $-0.0^{3}17,707,944$ |
| 8 | $0.088,888,89$ | $-0.0^{2}28,218,695$ | $0.0^{2}38,095,238$ | $-0.0^{3}23,373,061$ |

| $j$ | $i$ | | | |
|---|---|---|---|---|
| | 6 | 7 | 8 | 9 |
| 1 | $0.0^{6}10,087,459$ | $-0.0^{8}38,065,882$ | $0.0^{9}14,772,299$ | $-0.0^{11}47,770,889$ |
| 2 | $0.0^{5}13,059,860$ | $-0.0^{7}78,306,957$ | $0.0^{8}57,379,607$ | $-0.0^{9}32,240,604$ |
| 3 | $0.0^{5}49,870,764$ | $-0.0^{6}35,414,321$ | $0.0^{7}37,246,865$ | $-0.0^{8}26,934,251$ |
| 4 | $0.0^{4}12,515,888$ | $-0.0^{6}96,195,746$ | $0.0^{6}13,039,809$ | $-0.0^{7}10,793,811$ |
| 5 | $0.0^{4}25,264,163$ | $-0.0^{5}20,323,918$ | $0.0^{6}33,614,797$ | $-0.0^{7}30,234,979$ |
| 6 | $0.0^{4}44,603,642$ | $-0.0^{5}36,960,883$ | $0.0^{6}72,070,037$ | $-0.0^{7}68,563,784$ |
| 7 | $0.0^{4}71,905,926$ | $-0.0^{5}60,836,892$ | $0.0^{5}13,654,992$ | $-0.0^{6}13,526,252$ |
| 8 | $0.0^{3}10,854,319$ | $-0.0^{5}93,258,365$ | $0.0^{5}23,672,301$ | $-0.0^{6}24,174,891$ |

| $j$ | $i$ | | | |
|---|---|---|---|---|
| | 10 | 11 | 12 | 13 |
| 1 | $0.0^{12}14,640,444$ | $-0.0^{14}40,268,872$ | $0.0^{15}10,359,029$ | $-0.0^{17}24,535,539$ |
| 2 | $0.0^{10}18,330,114$ | $-0.0^{12}91,351,579$ | $0.0^{13}43,595,840$ | $-0.0^{14}19,132,452$ |
| 3 | $0.0^{9}21,506,514$ | $-0.0^{10}14,469,203$ | $0.0^{12}96,661,910$ | $-0.0^{13}58,727,628$ |
| 4 | $0.0^{8}10,849,591$ | $-0.0^{10}87,178,260$ | $0.0^{11}72,767,557$ | $-0.0^{12}54,213,617$ |
| 5 | $0.0^{8}36,260,639$ | $-0.0^{9}32,719,538$ | $0.0^{10}32,219,900$ | $-0.0^{11}27,049,719$ |
| 6 | $0.0^{8}95,092,297$ | $-0.0^{9}93,120,388$ | $0.0^{9}10,472,881$ | $-0.0^{11}96,020,717$ |
| 7 | $0.0^{7}21,247,442$ | $-0.0^{8}22,112,968$ | $0.0^{9}27,825,332$ | $-0.0^{10}27,369,553$ |
| 8 | $0.0^{7}42,365,199$ | $-0.0^{8}46,218,579$ | $0.0^{9}64,147,144$ | $-0.0^{10}66,862,484$ |

Using (12) the probability integral for $W$ can be evaluated with the help of tables of Incomplete Gamma Functions.

## REFERENCES

[1] L. H. C. TIPPET, "On the extreme individuals and the range of samples taken from a normal population," *Biometrika*, Vol. 17 (1925),pp. 364–387.

[2] STUDENT, "Errors in routine analysis," *Biometrika*, Vol. 19 (1927), pp. 151–164.

[3] T. HOJO, "Distribution of the median, quartiles and interquartile distance in samples from a normal population," *Biometrika*, Vol. 23 (1931), pp. 315–360.

[4] E. S. PEARSON AND H. O. HARTLEY, "The probability integral of the range in samples of $n$ observations from a normal population," *Biometrika*, Vol. 32 (1942), pp. pp. 301–310.

[5] E. J. GUMBEL, "Ranges and midranges," *The Annals of Math. Stat.*, Vol. 15 (1944), No. 4, pp. 414–422.

[6] E. J. GUMBEL, "On the independence of the extremes in a sample," *The Annals of Math. Stat.*, Vol. 17 (1946), No. 1, pp. 78–80.

[7] E. J. GUMBEL, "The distribution of the range," *The Annals of Math. Stat.*, Vol. 18 (1947), No. 3, pp. 384–412.

[8] G. ELFVING, "The asymptotical distribution of range in samples from a normal population," *Biometrika*, Vol. 34, (1947), pp. 111–119.

[9] J. F. DALY, "On the use of the sample range in an analogue of Student's $t$-test," *The Annals of Math. Stat.*, Vol. 17 (1946), No. 1, pp. 71–74.

[10] E. LORD, "The use of range in place of standard deviation in the $t$-test," *Biometrika*, Vol. 34 (1947), pp. 41–67.

[11] J. E. WALSH, "An extension to two populations of an analogue of Student's $t$-test using the sample range," *Annals of Math. Stat.*, Vol. 18 (1947), No. 2, pp. 280–285.

[12] S. S. WILKS, "Order Statistics," *Am. Math. Soc. Bull.*, Vol. 54 (1948), pp. 6–50.

[13] S. S. WILKS, *Mathematical Statistics*, Princeton University Press, Princeton, 1943, p. 90.

[14] K. C. S. PILLAI, "A note on ordered samples," *Sankhya*, Vol. 8 (1948), Part 4, pp. 375–380.