MOMENTS OF ORDER STATISTICS FROM A NORMAL POPULATION

by

R. C. Bose and Shanti S. Gupta

University of North Carolina

This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command under contract No. AF 18 (600) - 83.

Institute of Statistics
Mimeograph Series No. 154
July, 1956
1. University of North Carolina, Chapel Hill, N. C.
2. Mathematics Division, Air Force Office of Scientific Research
3. AFOSR-TN-56-501
4. ASTIA AD 110 315
5. MOMENTS OF ORDER STATISTICS FROM A NORMAL POPULATION
6. R. C. Bose and Shanti S. Gupta
7. October, 1956
8. AF 18(600)-83
9. File 3.3
1. Introduction and summary.

Statistics based on ordered observations have been called systematic statistics by Mosteller $\int_{12}^{7}$. They are now being increasingly used in new statistical procedures $\int_{1}^{2}, 3, 4, 7, 8, 13, 14, 15, 16, 17, 20, 22, 7$. The present paper deals with the problem of obtaining the moments of $X_{(k)}$, the $k$-th order statistic for a sample of size $n$ from a normal population $N(0, 1)$. This problem has been considered among others by Cole $\int_{5}^{7}, 7$, Godwin $\int_{6}^{7}$, Hastings, Mosteller, Tukey and Winsor $\int_{9}^{7}$, Jones $\int_{11}^{7}$, Ruben $\int_{18}^{7}$ and Tippett $\int_{23}^{7}$.

It has been shown that $\mu_{t}^{*}(n, k)$, the $t$-th moment of $X_{(k)}$, can be expressed in terms of lower moments of order $t - 21$ ($i = 1, 2, \ldots, t/2$ or $(t - 1)/2$) and the integral

$$
\int_{-\infty}^{+\infty} P_{t+1}(x) e^{-(t+1)x^2/2} dx
$$

where $P_{t+1}(x)$ for $t > 0$, is defined by

$$
P_{t+1}(x) = k \binom{n}{k} \frac{d^t}{d\phi^t} \int \phi^{k-1}(1 - \phi)^{n-k} d\phi
$$

it being understood that in (1.2), $\phi$ is replaced after differentiation by $\phi(x)$, the c. d. f. of $N(0, 1)$. $P_{t}(x)$ is thus a polynomial of degree $(n - t)$ in $\phi(x)$ if $n \leq t$ and is zero if $n > t$. Exact values of all odd order moments can be derived when $n \leq 5$, and the exact values of all even order moments can be derived when $n \leq 6$. Godwin $\int_{6}^{7}$ and Jones $\int_{11}^{7}$ have given tables of exact moments $\mu_{t}^{*}(n, k)$ for $t = 1$ and 2. The corresponding tables for $t = 3$ and 4 are provided.
in this paper. In general the numerical evaluation of the integral (1.1) can be expeditiously done by using the Gauss-Jacobi method of mechanical quadrature \( \int \frac{21}{\gamma} \) based on the zeros and the weight factors of the Hermite-polynomials for which tables have been provided by Salzer, Zucker and Capuano \( \int \frac{19}{\gamma} \). It is believed that the formulae derived here are better suited for numerical computation than those given elsewhere.

2. The function \( P_n(n, k, x) \).

Let \( x_1, x_2, \ldots, x_n \) be \( n \) independent observations from a normal population \( N(0, 1) \) with zero mean and unit variance, and let

\[
(2.1) \quad x_1 \leq x_2 \leq \cdots \leq x_n
\]

be the \( n \) ranked observations among \( x_1, x_2, \ldots, x_n \). Then the cumulative distribution function of \( X_k \), the random variable corresponding to \( x_k \), \( (1 \leq k \leq n) \), is given by

\[
(2.2) \quad P_0(n, k, x) = \text{Prob} \left\{ X_k \leq x \right\}
= \frac{C}{(2\pi)^{1/2}} \int_{-\infty}^{x} \left( \frac{\phi(x)}{\phi(x)} \right)^{k-1} \left( 1 - \frac{\phi(x)}{\phi(x)} \right)^{n-k} e^{-x^2/2} \, dx
\]

where \( \phi(x) \) is defined as

\[
(2.3) \quad \phi(x) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{x} e^{-x^2/2} \, dx
\]

and \( C \) is the constant

\[
(2.4) \quad C = \frac{n!}{(k - 1)! (n - k)!}.
\]
UN CLASSIFIED

Security Information

Bibliographical Control Sheet

1. University of North Carolina, Chapel Hill, N. C.
2. Mathematics Division, Air Force Office of Scientific Research
3. AFOSR-TN-56-501
4. ASTIA AD 110 316
5. MOMENTS OF ORDER STATISTICS FROM A NORMAL POPULATION
6. R. C. Bose and Shanti S. Gupta
7. October, 1956
8. AF 18(600)-83
9. File 3.3
Let us now define the function $P_t(n, k, x)$, which we shall abbreviate to $P_t(x)$ for convenience, by the relation

$$P_{t+1}(x) = (2\pi)^{1/2} e^{x^2/2} \frac{dP_t}{dx}$$

where $P_0(x)$ is given by (2.2). Then

$$P_1(x) = C \int \phi(x) e^{k-1} \int 1 - \phi(x) e^{n-k}$$

It is clear that $P_t(x)$ is a polynomial of degree $n-t$ in $\phi(x)$ if $1 \leq t \leq n$, and is zero for $t > n$. In fact, we can write

$$P_{t+1}(x) = C \frac{d^t}{d\phi^t} \int \phi^{k-1}(1 - \phi)^{n-k}$$

where $\phi$ is replaced by $\phi(x)$ after the differentiation. It follows that for given $t$, $n$, $k$, $\phi_t(x)$ is a bounded function of $x$. The functions $P_2(x)$, $P_3(x)$, $P_4(x)$ and $P_5(x)$ are given below explicitly, where $\phi$ is written for $\phi(x)$.

$$P_2(x) = C \phi^{k-2}(1 - \phi)^{n-k-1} \int (k-1) - (n-1) \phi$$

$$P_3(x) = C \phi^{k-3}(1 - \phi)^{n-k-2} \int (k-1)(k-2) - 2(k-1)(n-2)\phi + (n-1)(n-2)\phi^2$$

$$P_4(x) = C \phi^{k-4}(1 - \phi)^{n-k-3} \int (k-1)(k-2)(k-3) - 3(k-1)(k-2)(n-3)\phi$$

$$+ 3(k-1)(n-2)(n-3)\phi^2 - (n-1)(n-2)(n-3)\phi^3$$

$$P_5(x) = C \phi^{k-5}(1 - \phi)^{n-k-4} \int (k-1)(k-2)(k-3)(k-4) - 4(k-1)(k-2)(k-3)(k-4)\phi$$

$$+ 6(k-1)(k-2)(n-3)(n-4)\phi^2$$

$$- 4(k-1)(n-2)(n-3)(n-4)\phi^3 + (n-1)(n-2)(n-3)(n-4)\phi^4$$

3. A system of differential equations satisfied by $P_0(x)$.

From (2.5)
(3.1) \[ P_1(x) = (2\pi)^{1/2} e^{x^2/2} \frac{dP_0}{dx} \]

(3.2) \[ P_2(x) = (2\pi)^{1/2} e^{x^2/2} \frac{d}{dx} \int (e^x)^{1/2} e^{x^2/2} \frac{dP_0}{dx} \]

\[ = (2\pi) e^{x^2} \int \frac{d^2 P_0}{dx^2} + x \frac{dP_0}{dx} \]

In general let us assume

(3.3) \[ P_t(x) = (2\pi)^{t/2} e^{tx^2/2} \sum_{r=0}^{t-1} g_{r,t}(x) \frac{d^{t-r} P_0}{dy^{t-r}} \]

where

(3.4) \[ g_{0,t}(u) = 0 \]

and \( g_{r,t}(x) \) is a polynomial in \( x \) of the \( r \)-th degree. Differentiating (3.3) and using (2.5), we have

(3.5) \[ P_{t+1}(x) = (2\pi)^{(t+1)/2} e^{(t+1)x^2/2} \sum_{r=0}^{t-1} \left[ g_{r,t}(x) \frac{d^{t-r+1} P_0}{dx^{t-r+1}} + \left\{ txg_{r,t}(x) \right\} \frac{d^{t-r} P_0}{dx^{t-r}} \right] \]

This leads to the recurrence relation

(3.6) \[ g_{r,t+1}(x) = g_{r,t}(x) + \left\{ tx + \frac{d}{dx} \right\} g_{r-1,t}(x) \]

where \( g_{r,t}(x) \) should be interpreted as zero. This together with (3.4) determines all the polynomials \( g_{r,t}(x) \). Starting from

(3.7) \[ g_{0,1}(x) = 1 \]

we can successively calculate
\[ g_{0,2}(x) = 1, \quad g_{1,2}(x) = x \]

\[ g_{0,3}(x) = 1, \quad g_{1,3}(x) = 3x, \quad g_{2,3}(x) = 2x^2 + 1 \]

\[ g_{0,4}(x) = 1, \quad g_{1,4}(x) = 6x, \quad g_{2,4}(x) = 11x^2 + 4, \quad g_{3,4}(x) = 6x^3 + 7x \]

\[ g_{0,5}(x) = 1, \quad g_{1,5}(x) = 10x, \quad g_{2,5}(x) = 35x^2 + 10, \quad g_{3,5}(x) = 50x^3 + 45x \]

\[ g_{4,5}(x) = 24x^4 + 46x^2 + 7. \]

Hence we have the set of equations

\[ \frac{d P_0}{dx} = \frac{1}{(2\pi)^{1/2}} e^{-x^2/2} P_1(x) \]

\[ \frac{d^2 P_0}{dx^2} + x \frac{d P_0}{dx} = \frac{1}{(2\pi)^{1/2}} e^{-x^2} P_2(x) \]

\[ \frac{d^3 P_0}{dx^3} + 3x \frac{d^2 P_0}{dx^2} + (2x^2 + 1) \frac{d P_0}{dx} = \frac{1}{(2\pi)^{3/2}} e^{-3x^2/2} P_3(x) \]

\[ \frac{d^4 P_0}{dx^4} + 6x \frac{d^3 P_0}{dx^3} + (11x^2 + 4) \frac{d^2 P_0}{dx^2} + (6x^3 + 7x) \frac{d P_0}{dx} = \frac{1}{(2\pi)^{1/2}} e^{-2x^2} P_4(x) \]

\[ \frac{d^5 P_0}{dx^5} + 10x \frac{d^4 P_0}{dx^4} + (35x^2 + 10) \frac{d^3 P_0}{dx^3} + (50x^3 + 45x) \frac{d^2 P_0}{dx^2} + (24x^4 + 46x^2 + 7) \frac{d P_0}{dx} = \frac{1}{(2\pi)^{5/2}} e^{-5x^2/2} P_5(x). \]

We can proceed in this manner up to any order but it should be noted that

\[ P_n(x) \]

is a constant and \( P_t(x) = 0 \) if \( t > n \). The general equation is

\[ g_{0,t}(x) \frac{d^t P_0}{dx^t} + g_{1,t}(x) \frac{d^{t-1} P_0}{dx^{t-1}} + \cdots + g_{t-1,t}(x) \frac{d P_0}{dx} = \frac{1}{(2\pi)^{t/2}} e^{-tx^2/2} P_t(x). \]

4. Moments of \( X(k) \).

We shall first prove the following Lemma:
Lemma. If \( \alpha \) and \( r \) are non-negative integers, then

\[
\int_{-\infty}^{+\infty} \alpha \frac{d^{r+1} P_0}{dx^{r+1}} \, dx = (-1)^r \alpha (\alpha - 1) \ldots (\alpha - r + 1) \mu_{\alpha-r} \quad \text{or} \quad 0
\]

according as

\[ r \leq \alpha \quad \text{or} \quad r > \alpha \]

where \( \mu_{\alpha-r} \) is the \((\alpha-r)\)-th order moment of \( X(k) \) about the origin.

It should be noted that by definition

\[
\int_{-\infty}^{+\infty} \alpha \frac{dP_0}{dx} \, dx = \mu_{\alpha}.
\]

From (3.12) and (3.13)

\[
\frac{dP_0}{dx} = \frac{1}{(2\pi)^{1/2}} e^{-x^2/2} P_1(x)
\]

\[
\frac{d^2 P_0}{dx^2} = -\frac{x}{(2\pi)^{1/2}} e^{-x^2/2} P_1(x) + \frac{1}{2\pi} e^{-x^2} P_2(x)
\]

and in general using the system of equations (3.12)-(3.17) we can write

\[
\frac{d^r P_0}{dx^r} = \sum_{i=1}^{r} h_i(x) e^{-ix^2/2} P_1(x)
\]

where \( h_i(x) \) is a polynomial in \( x \). Now

\[
\int_{-\infty}^{+\infty} \alpha \frac{d^{r+1} P_0}{dx^{r+1}} \, dx = \left| \int_{-\infty}^{+\infty} \alpha \frac{d^r P_0}{dx^r} \right| - \int_{-\infty}^{+\infty} \alpha^{r-1} \frac{d^r P_0}{dx^r} \, dx.
\]

Since \( P_t(x) \) for any non-negative integer \( t \) is a bounded function of \( x \), it follows from (4.5) that the first part on the right hand side of (4.6) vanishes.
Repeating this process we get if \( r \leq \alpha \)

\[
\int_{-\infty}^{+\infty} x^{\alpha} \frac{d^{r+1} P_0}{dx^{r+1}} \, dx = (-1)^r \alpha(\alpha-1) \cdots (\alpha-r+1) \int_{-\infty}^{+\infty} x^{\alpha-r} \frac{dP_0}{dx} = (-1)^r \alpha(\alpha-1) \cdots (\alpha-r+1) \mu_{\alpha-r}.
\]

If \( r > \alpha \), we get on repeating the process \( \alpha \) times,

\[
\int_{-\infty}^{+\infty} x^{\alpha} \frac{d^{r+1} P_0}{dx^{r+1}} \, dx = (-1)^\alpha(\alpha-1) \cdots 3 \cdot 2 \cdot 1 \int_{-\infty}^{+\infty} \frac{d^{r-\alpha} P_0}{dx^{r-\alpha}} \, dx
\]

\[
= 0.
\]

This proves the Lemma.

On applying the Lemma and integrating the equations (3.13) ... (3.16) we get

\[(4.7) \quad \mu'_1 = \frac{1}{2\pi} \int_{-\infty}^{+\infty} P_2(x)e^{-x^2} \, dx\]

\[(4.8) \quad -3 + (3 \mu'_1 + 1) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{+\infty} P_3(x)e^{-3x^2/2} \, dx\]

\[(4.9) \quad -2\mu'_1 + (6 \mu'_3 + 7 \mu'_1) = \frac{1}{(2\pi)^{2}} \int_{-\infty}^{+\infty} P_4(x)e^{-2x^2} \, dx\]

\[(4.10) \quad 70 - 150\mu'_2 - 45 + (24 \mu'_4 + 46 \mu'_2 + 7) = \frac{1}{(2\pi)^{5/2}} \int_{-\infty}^{+\infty} P_5(x)e^{-5x^2/2} \, dx.\]

We may write \( \mu'_\alpha(n,k) \) instead of \( \mu'_\alpha \) to denote the fact that we have the \( \alpha \)-th moment about the origin of the \( k \)-th order statistic out of a sample of \( n \) observations from \( N(0,1) \). We then have
\begin{align*}
\mu_1'(n,k) & = \frac{1}{c^n} \int_{-\infty}^{+\infty} P_2(x) e^{-x^2} \, dx \\
\mu_2'(n,k) & = 1 + \frac{1}{2!(2\pi)^{3/2}} \int_{-\infty}^{+\infty} P_3(x) e^{-3x^2/2} \, dx \\
\mu_3'(n,k) & = \frac{5}{2} \mu_1'(n,k) + \frac{1}{3!(2\pi)^{3/2}} \int_{-\infty}^{+\infty} P_4(x) e^{-4x^2/2} \, dx \\
\mu_4'(n,k) & = -\frac{4}{3} + \frac{13}{6} \mu_2'(n,k) + \frac{1}{4!(2\pi)^{5/2}} \int_{-\infty}^{+\infty} P_5(x) e^{-5x^2/2} \, dx
\end{align*}

In general applying the Lemma to (3.17) we can express $u'_t(n,k)$ in terms of lower moments of even (odd) order when $t$ is even (odd) and the integral

\begin{align*}
\int_{-\infty}^{+\infty} P_{t+1}(x) e^{-(t+1)x^2/2} \, dx
\end{align*}

where the polynomials $P_2(x) \ldots P_5(x)$ are given by (2.8) \ldots (2.11). In the particular case when $n = k$, i.e., $x_{(n)}$ is the largest of $x_1, x_2, \ldots, x_n$, $P_t(x)$ assumes the very simple form

\begin{align*}
P_t(x) = n(n-1)(n-2) \ldots (n-t+1) [\phi(x)]^{n-t}.
\end{align*}

Hence we get

\begin{align*}
\mu_1'(n,n) & = \frac{3^n}{2\pi} \int_{-\infty}^{+\infty} [\phi(x)]^{n-2} e^{-x^2} \, dx \\
\mu_2'(n,n) & = 1 + \frac{3^n}{(2\pi)^{3/2}} \int_{-\infty}^{+\infty} [\phi(x)]^{n-3} e^{-3x^2/2} \, dx
\end{align*}
(4.19) \[ \mu_2(n,n) = \frac{5}{2} \mu_1(n,n) + \frac{4(n)}{(2\pi)^{5/2}} \int_{-\infty}^{+\infty} \left[ \phi(x) \right]^{-n-4} e^{-4x^2/2} \, dx \]

(4.20) \[ \mu_3(n,n) = -\frac{4}{3} + \frac{13}{3} \mu_2(n,n) + \frac{5(n)}{(2\pi)^{5/2}} \int_{-\infty}^{+\infty} \left[ \phi(x) \right]^{-n-5} e^{-5x^2/2} \, dx \]

It should be noted that in the formulae (4.17) through (4.20) \[ \left[ \phi(x) \right]^{-n-t} \]

should be interpreted as zero if \( t > n \).

Some integrals of the type occurring in (4.17) through (4.20) have been numerically evaluated by Hojo in [10].

5. Exact values of some moments.

Let

(5.1) \[ I_n(a) = \int_{-\infty}^{+\infty} \left[ \phi(ax) \right]^n e^{-x^2} \, dx \]

then

(5.2) \[ I_0(a) = \frac{1}{2} \]

Now

(5.3) \[ \int_{-\infty}^{+\infty} \left[ \phi(ax) - \frac{1}{2} \right]^{2m+1} e^{-x^2} \, dx = 0, \]

since the integrand is an odd function of \( x \). Hence

(5.4) \[ I_{2m+1}(a) = \sum_{r=1}^{2m+1} (-1)^{r+1} \binom{2m+1}{r} I_{2m-r+1}(2)^r / 2^r. \]

In particular
(5.5) \[ I_1(a) = \frac{1}{2} I_0(a) = \frac{1}{2} \pi^{1/2}, \]

and

(5.6) \[ I_3(a) = \frac{3}{2} I_2(a) - \frac{3}{4} I_1(a) + \frac{1}{6} I_0(a) = \frac{3}{2} I_2(a) - \frac{1}{4} I_0(a). \]

In general, \( I_{2m+1}(a) \) can be expressed as a linear function of \( I_{2m}(a), I_{2m-2}(a), \ldots, I_0(a) \).

Differentiating (5.1) with respect to \( a \), (this is justified in virtue of the uniform convergence of the integrals with respect to \( a \), \(-\infty < a < \infty\), and the continuity of the integrands), we get for \( n = 2 \),

\[ \frac{dI_2}{da} = \frac{1}{\pi^{1/2}} \frac{a}{(a^2 + 2)(a^2 + 1)^{1/2}}, \]

so that

\[ I_2(a) = \frac{1}{\pi^{1/2}} \text{arc tan} (\sqrt{a^2 + 1}). \]

Using (5.6)

\[ I_3(a) = \frac{3}{2\pi^{1/2}} \text{arc tan} (\sqrt{a^2 + 1}) - \frac{\pi^{1/2}}{4}. \]

Since \( P_{t+1}(x) \) is a polynomial in \( \Phi(x) \) of degree \( n-t-1 \), by using (5.2), (5.5), (5.8) and (5.9), we can exactly evaluate (4.15) if \( n \leq t+4 \). Hence we can exactly evaluate \( \mu_t'(n,k) \) for all odd values of \( t \), if \( n \leq 5 \) and all even values of \( t \) if \( n \leq 6 \). Godwin \[ 6 \] and Jones \[ 11 \] have given tables of exact moments \( \mu_t' \) for \( t = 1 \) and 2. The corresponding tables for \( t = 3 \) and 4 are given below.

<table>
<thead>
<tr>
<th>Table I.</th>
<th>( \mu_3'(n,k) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>( k = n )</td>
</tr>
<tr>
<td>2</td>
<td>2A</td>
</tr>
<tr>
<td>3</td>
<td>3A</td>
</tr>
<tr>
<td>4</td>
<td>2B_1 + 2C</td>
</tr>
<tr>
<td>5</td>
<td>-5A + 5B_1 + 5B_2</td>
</tr>
</tbody>
</table>

Here

\[ A = \frac{5}{4\pi^{1/2}} \]

\[ B_1 = \frac{1}{2\pi^{3/2}} \]

\[ B_2 = \frac{15}{2\pi^{3/2}} \]

\[ C = \frac{15}{2\pi^{3/2}} \text{arc tan} (\sqrt{2}) \]
Table II

<table>
<thead>
<tr>
<th>n</th>
<th>3+a</th>
<th>3-2a</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3+2a</td>
</tr>
<tr>
<td>4</td>
<td>3+2a</td>
<td>3-2a</td>
</tr>
<tr>
<td>5</td>
<td>3+b+c</td>
<td>3+10a-4b-4c</td>
</tr>
<tr>
<td>6</td>
<td>3-5b+3b+3c</td>
<td>3+25a-9b-9c</td>
</tr>
</tbody>
</table>

Here

\[ a = \frac{13}{\sqrt{5(2\pi)}} \]
\[ b = \frac{65}{\sqrt{5\pi^2}} \text{ arc tan} \left( \frac{\sqrt{5}}{3} \right) \]
\[ c = \frac{\sqrt{5}}{4\pi^2} . \]

REFERENCES


