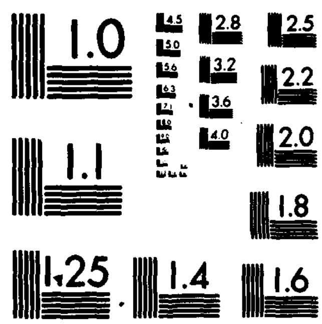
AD-A082 921 STANFORD UNIV CA DEPT OF STATISTICS F/G 12/1
APPROXIMATING CONDITIONAL MOMENTS OF THE MULTIVARIATE NORMAL DI--ETC(U)
DEC 79 J G DEKEN N00014-75-C-0442
UNCLASSIFIED TR-41 NL

END
DATE
FILMED
5-80
DTIC



MICROCOPY RESOLUTION TEST CHART

(12) LEVEL II

ADA082921

APPROXIMATING CONDITIONAL MOMENTS OF THE MULTIVARIATE NORMAL DISTRIBUTION

BY

JOSEPH G. DEKEN

TECHNICAL REPORT NO. 41
DECEMBER 1979

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)
OFFICE OF NAVAL RESEARCH
THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

APPROXIMATING CONDITIONAL MOMENTS OF THE
MULTIVARIATE NORMAL DISTRIBUTION

by

JOSEPH G. DEKEN

Technical Report No. 41

December 1979

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)
OFFICE OF NAVAL RESEARCH

Theodore W. Anderson, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

APPROXIMATING CONDITIONAL MOMENTS OF THE MULTIVARIATE

NORMAL DISTRIBUTION

By

Joseph G. Deken

ABSTRACT

A practical method for computing the conditional expectation of a polynomial in the components of a multivariate normal random variable $\underset{\sim}{X}$, when $\underset{\sim}{X}$ is restricted to a subset of $\mathbb{R}^p$, is given. This method makes the application of certain missing data techniques possible in cases where repeated numerical integration is not feasible.

Key words: Multiple integration, numerical integration, multivariate normal distribution, EM estimation.

APPROXIMATING CONDITIONAL MOMENTS OF THE MULTIVARIATE NORMAL DISTRIBUTION

By

Joseph G. Deken

1. Introduction.

The conditional moments such as $E(X_j^k)$ of a multivariate normal random variable $\underset{\sim}{X} = (X_1, X_2, \ldots, X_p)$, when $\underset{\sim}{X}$ is restricted to a subset $A \subset \mathbb{R}^p$, are not readily obtained numerically, since the required integration in p-dimensions is time-consuming except for very small p. These conditional moments are of interest, for example in the derivation of E-M estimates in missing data problems (Dempster, Laird and Rubin, 1977). We present here an efficient approximation scheme for these moments, which makes the computation practical for moderately large p.

For convenience of description, we restrict attention to sets A of the form $I_1 \times I_2 \times \cdots \times I_p$, where all the I_j are intervals, but the approach is more general, as indicated below. We start by observing that the approximations to the ratio

$$E_k = \frac{\int_{s-t}^{s+t} x^k e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx}{\int_{s-t}^{s+t} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx}$$

obtained by the first terms in a Taylor series around $t = 0$ may be written as polynomials in μ:

$$E_k \approx \sum_{\alpha=0}^{m} c_{k\alpha}(s,t,\sigma^2)\mu^{\alpha} .$$

The conditional expectation of a polynomial $P(X) = a_0+a_1X+a_2X^2+\cdots+a_nX^n$ in the normal random variable X is thus approximated by replacing P by a polynomial $Q(\mu)$ in the mean of X, where the transformation $T(s,t,\sigma^2)$: $P \to Q$ is linear. Since the transformation V: $\underset{\sim}{c} \to \underset{\sim}{b}$ defined by

$$\sum_{\alpha=0}^{m} c_{\alpha}(g+hX)^{\alpha} = \sum_{\beta=0}^{m} b_{\beta}X^{\beta}$$

is also linear, and the conditional mean of X_p given $X_1,\ldots,X_{p-1}$ is of the form $(g+hX_{p-1})$ where g is linear in $X_1,\ldots,X_{p-2}$, any approximation to EX_p^k which is a polynomial in the mean of X_p will produce a polynomial in $(X_1,\ldots,X_{p-1})$ when X_p is conditional on $X_1,\ldots,X_{p-1}$. The conditional expectation in $\mathbb{R}^p$ may thus be accomplished in $2p$ steps, by applying T and then V in succession p times, to obtain

$$E(\cdot|X_1\cdots X_{p-1}) \approx E(\cdot|X_1\ldots X_{p-2}) \approx \cdots \approx E(\cdot\ |X_1) \approx \sum_{\alpha=0}^{N} g_{\alpha}\mu_1^{\alpha} .$$

Computationally, this process requires $2p$ matrix multiplications and is thus practical for moderate p.

2. An Example.

The following simple example will serve to establish some ideas and notation. Let

$$I_k(\sigma^2,\mu,s,t) = \int_{s-t}^{s+t} x^k e^{-\frac{(\mu-x)^2}{2\sigma^2}}\, dx \;.$$

We are concerned with the conditional expectation I_k/I_0 of X^k, where X is a normal random variable. The sum of the first terms of a Taylor series for this ratio about $t = 0$ is of the form

$$(1) \qquad \sum_{\alpha=0}^{q} c^{*}_{k\alpha}(\sigma^2,\mu,s)t^{2\alpha} \;,$$

but also may be represented by identifying the coefficients of μ^α in (1) as

$$\sum_{\alpha=0}^{2q} c_\alpha(\sigma^2,s,t)\mu^\alpha \;.$$

For example, (letting $v = \sigma^2$)

$$I_1/I_0 = s + \frac{(\mu-s)t^2}{3v} + \frac{2v(s-\mu) + (s-\mu)^3}{45v^3}\, t^4 + \ldots$$

$$= s\left(1 + \frac{-15t^2v^2+2t^4v+s^2t^4}{45v^3}\right) + \left(\frac{15t^2v^2-2t^4v-3s^2t^4}{45v^3}\right)\mu + \left(\frac{st^4}{15v^3}\right)\mu^2 - \left(\frac{t^4}{45v^3}\right)\mu^4 + \ldots$$

3. Multiple Integration.

If (X_1, X_2) are bivariate normal, the conditional distribution of X_2, given X_1, is $\mathcal{N}(\mu_2 + b_{21}(X_1 - \mu_1), \sigma^2_{2.1})$. The approximation procedure given above yields (writing $\underset{\sim}{p}$ for the vector of coefficients of P, and with C the matrix whose j,kth element is the coefficient of μ^k in the approximation of I_j/I_0):

$$E(P(X_2)|X_2 \epsilon (s_2 - t_2, s_2 + t_2), X_1) \doteq (\underset{\sim}{p}C(s_2, t_2, \sigma^2_{2.1}))(\mu_2 + b_{21}(X_1 - \mu_1)) .$$

This last is of course a polynomial in X_1, with coefficients

$$q_\alpha := b_{21}^\alpha \sum_{\beta \geq \alpha} (\underset{\sim}{p}C(s_2, t_2, \sigma^2_{2.1}))_\beta \binom{\beta}{\alpha} (\mu_2 - b_{21}\mu_1)^{\beta - \alpha} ,$$

so that

$$E(P(X_2)|X_2 \epsilon (s_2 - t_2, s_2 + t_2),\ X_1 \epsilon (s_1 - t_1, s_1 + t_1)) =$$
$$(\underset{\sim}{q}C(s_1, t_1, \sigma_1^2))(\mu_1) .$$

Other cases (higher dimensions, multivariate polynomials) are treated similarly, but it will be best to introduce some notation at this point.

If $\underset{\sim}{p}$ is a vector of coefficients, then

<u>SUBS</u>$(a, b, \underset{\sim}{p})$ is a vector of coefficients,

$$(\mathrm{SUBS}(a, b, \underset{\sim}{p}))_j := b^j \sum_{k \geq j} \binom{k}{j} a^{k-j} .$$

(i.e. $\mathrm{SUBS}(a, b, \underset{\sim}{p})$ is the vector of coefficients of y^j when $a + by$ is substituted for x in $p(x)$).

With the aid of a symbolic computation system such as MIT's MACSYMA, the coefficients of the transformation

$$X^k \to \sum_{\alpha} c_{k\alpha} \mu^{\alpha} ,$$

as functions $c_{k\alpha}(v,s,t)$ may be computed for much larger values of k and α then would be practical by hand. (A partial table of values $c_{k\alpha}$ is given in ap endix A of this paper. Card versions consisting of FORTRAN assignment statements " c (I,J) = ... " may be obtained on request from the author.) As subsequent examples show, the approximation of X_p^k in the box $I_1 \times I_2 \times \cdots I_p$ involves $c_{m\alpha}$ for higher values of m than k, depending on the order of the underlying Taylor series. Fortunately, the approximation based on only a few terms is quite accurate, as evidenced by the following example: Since the integral

$$\int_{s-t}^{s+t} x e^{-\frac{x^2}{2}} dx = e^{-\frac{(s-t)^2}{2}} - e^{-\frac{(s+t)^2}{2}} ,$$

the Taylor series

$$\frac{I_1(1,0,s,t)}{I_0(1,0,s,t)} \sim S(1 - \frac{t^2}{3} + \frac{2+s^2}{45} t^4)$$

gives

$$I_0(1,0,s,t) = \Phi(s+t) - \Phi(s-t) \approx \frac{e^{-\frac{(s-t)^2}{2}} - e^{-\frac{(s+t)^2}{2}}}{\sqrt{2\pi}\, s(1 - \frac{t^2}{3} + \frac{2+s^2}{45} t^4)} .$$

The approximate values obtained for $\Phi(x) = .5 + \Phi(x) - \Phi(0)$ by this method agree with tabled values for four decimal places for $0 < x < 1.1$.

$\underline{INT}(s,t,v,\underset{\sim}{p})$ is a map from the scalars s,t,v and the vector p to the vector $\underset{\sim}{p}C(s,t,v)$, i.e. matrix postmultiplication, where we take

$$I_k/I_0 \sim \sum_{\alpha=0}^{m} c_{\alpha j}\mu^j \text{ as defining } C.$$

We will also consider $\underline{INT}$ as a function, so that if $\underline{INT} = (a_0,a_1,\ldots,a_m)$, then $(\underline{INT})(x) = a_0+a_1X+\cdots+a_mX^m$.

With the above notation, we can define multiple integrals in a fairly compact fashion. The multivariate normal distribution may be parameterized by a vector of means $(\mu_1,\mu_2,\ldots,\mu_p)$, the conditional variances $(V_1,\ldots,V_p)$ defined by $V_1 = \mathrm{Var}(X_1)$, $V_j = \mathrm{Var}(X_j|X_1,\ldots,X_{j-1}), j > 1$, and the regression coefficients $b_{21},b_{32},b_{31},\ldots,b_{p,p-1},\ldots,b_{p1}$ defined by

$$E(X_j|X_1,\ldots,X_{j-1}) = \sum_{i=1}^{j-1} b_{ji}X_i .$$

One-dimensional integration is done directly:

$$(\underset{\sim}{p},s,t,\mu,v) \leftarrow (\mathrm{INT}(s,t,v,\underset{\sim}{p}))(\mu) .$$

Further integrals $(I2,I3,\ldots)$ in higher dimensions are defined recursively, e.g. in two dimensions:

$$I2(s_1,t_1,s_2,t_2,\mu_1,\mu_2,v_1,v_2,b_{21},\underset{\sim}{p}) :=$$

$$(\mathrm{INT}(s_1,t_1,v_1,\mathrm{SUBS})\mu_2\text{-}b_{21}\mu_1,b_{21},\mathrm{INT}(s_2,t_2,v_2,\underset{\sim}{p})))) ,$$

and from this we obtain the two-dimensional integral as

$$(I2(s_1,t_1,s_2,t_2,\mu_1,\mu_2,v_1,v_2,b_{21},\underset{\sim}{p}))(\mu_1) .$$

In general, IK = INT(...SUBS(...I(K-1))). We have restricted attention here to sets of the form $I_1 \times I_2 \times \cdots I_p$ where the I_j are intervals. For more general sets A, a similar scheme would work provided the *section of* $X_i \in A$ *given* $X_1 \cdots X_{i-1}$ is an interval *with* endpoints which are polynomials in $X_1, \cdots, X_{i-1}$.

4. A Numeric Implementation.

The following example shows how the three-dimensional conditional expectation approximation may be defined, using only vector and matrix arithmetic with numeric arguments. The polynomial to be approximated is presumed to be of degree at most two, and perhaps bivariate in X_2 and X_3. (This case is general enough for all the means, squares, and products $EX_1, EX_2, \ldots EX_1^2, EX_2^2, \ldots EX_1X_2, \ldots EX_2X_3$). In fact, we describe here the approximate values of $E(X^3)$, $E(X_3^2)$, $E(X_2X_3)$, since other expectations reduce, by interchange of variables, to these three. The approximation used will include powers of t up to t^4 inclusive, i.e.

$$EX^j \approx \sum_{\alpha=0}^{3} c_{j\alpha}(s,t,v)\mu^{\alpha} .$$

The functions $\{c_{j\alpha}\}_{\alpha=0}^{3}\ {}_{j=1}^{6}$ may be found in appendix A).

To approximate $E(X_3)$, the integration on X_3 is first carried out, yielding

$$\sum_{\alpha=0}^{3} c_{1\alpha}(s_3,t_3,v_3)(\mu_3-\mu_1 b_{31}-\mu_2 b_{32}+X_1 b_{31}+X_2 b_{32})^{\alpha}$$

$$= \sum_{\alpha=0}^{3} c_{1\alpha}(s_3,t_3,v_3) \sum_{\beta=0}^{\alpha} \sum_{\gamma=0}^{\alpha-\beta} \binom{\alpha}{\beta\gamma} (b_{31}X_1)^{\beta}(b_{32}X_2)^{\gamma}(\mu_3-\mu_1 b_{31}-\mu_2 b_{32})^{\alpha-(\beta+\gamma)}$$

$$\sum_{\alpha=0}^{3} \sum_{\beta=0}^{3-\alpha} \left(\sum_{\gamma=\alpha+\beta}^{3} \binom{\gamma}{\alpha,\beta} c_{1\gamma}(s_3,t_3,v_3)(\mu_3-\mu_1 b_{31}-\mu_2 b_{32})^{\gamma-(\alpha+\beta)}\right) b_{31}^{\alpha} b_{32}^{\beta} \cdot X_1^{\alpha} X_2^{\beta} .$$

The above expression is of the form $\sum_{\alpha=0}^{3} \sum_{\beta=0}^{3-\alpha} r_{\alpha\beta} X_1^{\alpha} X_2^{\beta}$, where

$$r_{\alpha\beta} = b_{31}^{\alpha} b_{32}^{\beta} \sum_{\gamma=\alpha+\beta}^{3} \binom{\gamma}{\alpha,\beta} c_{1\gamma}(s_3,t_3,v_3) \cdot (\mu_3-\mu_1 b_{31}-\mu_2 b_{32})^{\gamma-(\alpha+\beta)} .$$

Integration on X_2 then yields:

$$\sum_{\alpha=0}^{3} \sum_{\beta=0}^{3-\alpha} r_{\alpha\beta} X_1^{\alpha} \left(\sum_{\gamma=0}^{3} c_{\beta\gamma}(s_2,t_2,v_2) \cdot (\mu_2-\mu_1 b_{21}+b_{21}X_1)^{\gamma} \right)$$

$$= \sum_{\alpha=0}^{6} s_{\alpha} X_1^{\alpha} , \quad \text{where} \quad s_{\alpha} = \sum_{\delta=(\alpha-2)\vee 0}^{\alpha\wedge 3} \sum_{\beta=0}^{3-\delta} r_{\delta\beta} \sum_{\gamma=\alpha-\delta}^{3} \binom{\gamma}{\alpha-\delta} b_{21}^{\alpha-\delta} \cdot$$

$$c_{\beta\gamma}(s_2,t_2,v_2) \cdot (\mu_2-\mu_1 b_{21})^{\gamma-(\alpha-\delta)} ,$$

and the final form is

$$\sum_{\alpha=0}^{6} s_{\alpha} \sum_{\beta=0}^{3} c_{\alpha\beta}(s_1,t_1,v_1)\mu_1^{\beta} .$$

The approximation of X_3^2 is carried out in the same fashion:

$$EX_3^2 = \sum_{\alpha=0}^{6} s'_{\alpha} \sum_{\beta=0}^{3} c_{\alpha\beta}(s_1,t_1,v_1)\mu_1^{\beta} ,$$

$$s'_{\alpha} = \sum_{\delta=(\alpha-3)\vee 0}^{\alpha\wedge 3} \sum_{\beta=0}^{3-\delta} r'_{\delta\beta} \sum_{\gamma=\alpha-\delta}^{3} \binom{\gamma}{\alpha-\delta} b_{21}^{\alpha-\delta} c_{\beta\gamma}(s_2,t_2,v_2) \cdot$$

$$(\mu_2-\mu_1 b_{21})^{\gamma-(\alpha-\delta)} ,$$

$$r'_{\alpha\beta} = b_{31}^{\alpha} b_{32}^{\beta} \sum_{\gamma=\alpha+\beta}^{3} \binom{\gamma}{\alpha\beta} c_{2\gamma}(s_3,t_3,v_3)(\mu_3-\mu_1 b_{31}-\mu_2 b_{32})^{\gamma-(\alpha+\beta)} .$$

(Note that the only change is the substitution of $c_{2\gamma}(s_3,t_3,v_3)$ for $c_{1\gamma}(s_3,t_3,v_3)$).

To approximate X_2X_3, the first approximate integration is exactly as in approximating X_3, but the resulting sum is

$$X_2\left(\sum_{\alpha=0}^{3}\sum_{\beta=0}^{3-\alpha} r_{\alpha\beta}X_1^{\alpha}X_2^{\beta}\right) ,$$

i.e. we have $r''_{\alpha\beta} = r_{\alpha(\beta-1)}$, $\beta = 1,\ldots,4-\alpha$. A second integration gives

$$\sum_{\alpha=0}^{3}\sum_{\beta=1}^{4-\alpha} r''_{\alpha\beta}X_1^{\alpha}\left(\sum_{\gamma=0}^{3} c_{\beta\gamma}(s_2,t_2,v_2)(\mu_2-b_{21}\mu_1+b_{21}X_1)^{\gamma}\right) ,$$

so the final result is

$$\sum_{\alpha=0}^{6} s''_{\alpha} \sum_{\beta=0}^{3} c_{\alpha\beta}(s_1,t_1,v_1)\mu_1^{\beta} ,$$

where

$$s''_{\alpha} = \sum_{\delta=(\alpha-3)\vee 0}^{\alpha\wedge 3}\ \sum_{\beta=1}^{4-\delta} r_{\delta\beta} \sum_{\gamma=\alpha-\delta}^{3} \binom{\gamma}{\alpha-\delta} b_{21}^{\alpha-\delta}\, c_{\beta\gamma}(s_2,t_2,v_1)(\mu_2-\mu_1 b_{21})^{\gamma-(\alpha-\delta)} .$$

<u>Higher Order Approximation, Higher Dimensions.</u>

In general, if $EX \doteq \sum_{\alpha=0}^{N} c_{1\alpha}\mu^{\alpha}$ is taken as the basic approximation, the three-dimensional integral will be of the form

$$\sum_{\alpha=0}^{2N} s_{\alpha} \sum_{\beta=0}^{N} c_{\alpha\beta}(s_1,t_1,v_1)\mu_1^{\beta} ,$$

where

$$s_{\alpha} = \sum_{\alpha}\sum_{\beta} s_{\alpha\beta}^{(2)} ,\quad s_{\alpha\beta}^{(2)} = \sum\sum\sum s_{\alpha\beta\gamma}^{(3)} ,\ldots,\ s_{\alpha_1\alpha_2\ldots\alpha_{p-1}}^{(p-1)} .$$

Thus, increasing accuracy is computationally expensive, and increasing dimension more so, but nonetheless the process should compare favorably with numerical integration, as borne out by some preliminary comparisons, where Romberg integration was about 10x slower.

Appendix A: Values of c_{jk}

$$\begin{cases} \text{Interval } (s-t, s+t) \\ X \sim \mathcal{N}(\mu, \sigma^2) \\ w = \dfrac{1}{\sigma^2} = \dfrac{1}{v} \end{cases}$$

$$\boxed{j = 1}$$

k=0

$$-\frac{s(t^2w(t^2w(w(2t^2(s^2w(s^2w+4)+1)-21s^2)-42)+315-945)}{945}$$

k=1

$$\frac{t^2w(w(2t^2(s^2w(5s^2w+12)+1)-63s^2)-42+315)}{945}$$

k=2

$$-\frac{st^4w^3(4t^2w(5s^2w+6)-63)}{945}$$

k=3

$$\frac{t^4w^3(4t^2w(5s^2w+2)-21)}{945}$$

k=4

$$-\frac{2st^6w^5}{189}$$

k=5

$$\frac{2t^6w^5}{945}$$

$$\boxed{j = 2}$$

k=0

$$-(t^2(2w(t^2(w(t^2(2s^2w(s^2w+5)+2)-1)-21s^2(s^2w+3))+21)+315s^2)-315)$$
$$-945s^2)/945$$

k=1

$$\frac{2st^2w(t^2w(w(2t^2(s^2w+3)(5s^2w+1)-63s^2)-84)+315)}{945}$$

k=2

$$-\frac{2t^4w^2(w(2t^2(2s^2w(5s^2w+9)+1)-63s^2)-21)}{945}$$

k=3

$$\frac{2st^4w^3(4t^2w(5s^2w+4)-21)}{945}$$

k=4

$$-\frac{4t^6w^4(5s^2w+1)}{945}$$

k=5

$$\frac{4st^6w^5}{945}$$

$$\boxed{j = 3}$$

k=0

$$-\ s(t^2(w(t^2(w(t^2(s^2w(2s^2w(s^2w+6)+3)-14)-21s^2(s^2w+4))$$

$$+\ 105)+315s^2)-315)-315s^2)/315$$

k=1

$$t^2w(t^2(w(t^2(s^2w(10s^2w(s^2w+4)+1)-12)-63s^2(s^2w+2))+63)$$

$$+\ 315s^2)/315$$

k=2

$$-\ \frac{st^4w^2(w(t^2(2s^2w+5)(10s^2w-1)-63s^2)-42)}{315}$$

k=3

$$\frac{t^4w^3(t^2(4s^2w(5s^2w+6)-3)-21s^2)}{315}$$

k=4

$$-\ \frac{2st^6w^4(5s^2w+2)}{315}$$

k=5

$$\frac{2s^2t^6w^5}{315}$$

$\boxed{j = 4}$

k=0

$$-(8s^8t^6w^5+56s^6t^6w^4-4s^4t^6w^3-84s^6t^4w^3-192s^2t^6w^2$$

$$-420s^4t^4w^2+36t^6w+1008s^2t^4w+1260s^4t^2w-189t^4-1890s^2t^2$$

$$-945s^4)/945$$

k=1

$$4st^2w(t^2(w(t^2(s^2w(2s^2w(5s^2w+24)-13)-54)$$

$$-21s^2(3s^2w+8))+189)+315s^2)/915$$

k=2

$$-\frac{4t^4w^2(t^2(s^2w(20s^2w(s^2w+3)-21)-9)-63s^2(s^2w+1))}{945}$$

k=3

$$\frac{4st^4w^3(t^2(4s^2w(5s^2w+8)-9)-21s^2)}{945}$$

k=4

$$-\frac{8s^2t^6w^4(5s^2w+3)}{945}$$

k=5

$$\frac{8s^3t^6w^5}{945}$$

$$\boxed{j = 5}$$

k=0

$$-s(2s^8t^6w^5+16s^6t^6w^4-8s^4t^6w^3-21s^6t^4w^3-112s^2t^6w^2$$
$$-126s^4t^4w^2+63t^6w+462s^2t^4w+315s^4t^2w-189t^4-630s^2t^2$$
$$-189s^4)/189$$

k=1

$$t^2w(10s^8t^4w^4+56s^6t^4w^3-36s^4t^4w^2-63s^6t^2w^2-144s^2t^4w$$
$$-210s^4t^2w+27t^4+378s^2t^2+315s^4)/189$$

k=2

$$-\frac{st^4w^2(2t^2(s^2w(2s^2w(5s^2w+18)-23)-18)-21s^2(3s^2w+4))}{189}$$

k=3

$$\frac{s^2t^4w^3(2t^2(10s^2w(s^2w+2)-9)-21s^2)}{189}$$

k=4

$$-\frac{2s^3t^6w^4(5s^2w+4)}{189}$$

k=5

$$\frac{2s^4t^6w^5}{189}$$

$$\boxed{j = 6}$$

k=0

$$- (4s^{10}t^6w^5+36s^8t^6w^4-36s^6t^6w^3-42s^8t^4w^3-430s^4t^6w^2$$
$$- 294s^6t^4w^2+450s^2t^6w+1470s^4t^4w+630s^6t^2w-45t^6-945s^2t^4$$
$$- 1575s^4t^2-315s^6)/315$$

k=1

$$2st^2w(10s^8t^4w^4+64s^6t^4w^3-68s^4t^4w^2-63\ s^6t^2w^2-300s^2t^4w$$
$$- 252s^4t^2w+135t^4+630s^2t^2+315s^4)/315$$

k=2

$$- 2s^2t^4w^2(2t^2(2s^2w(s^2w+5)(5s^2w-4)-45)-21s^2(3s^2w+5))/315$$

k=3

$$\frac{2s^3t^4w^3(2t^2(2s^2w(5s^2w+12)-15)-21s^2)}{315}$$

k=4

$$-\frac{4s^4t^6w^4(s^2w+1)}{63}$$

k=5

$$\frac{4s^5t^6w^5}{315}$$

$$\boxed{j = 7}$$

k=0

$$- s(2s^{10}t^6w^5+20s^8t^6w^4-31s^6t^6w^3-21s^8t^4w^3-366s^4t^6w^2$$

$$- 168s^6t^4w^2+585s^2t^6w+1071s^4t^4w+315s^6t^2w-135t^6$$

$$- 945s^2t^4-945s^4t^2-135s^6)/135$$

k=1

$$s^2t^2w(10s^8t^4w^4+72s^6t^4w^3-109s^4t^4w^2-63s^6t^2w^2-540s^2t^4w$$

$$- 294s^4t^2w+405t^4+945s^2t^2+315s^4)/315$$

k=2

$$-\frac{s^3t^4w^2(t^2(s^2w(4s^2w(5s^2w+24)-123)-180)-63s^2(s^2w+2))}{135}$$

k=3

$$\frac{s^4t^4w^3(t^2(4s^2w(5s^2w+14)-45)-21s^2)}{135}$$

k=3

$$-\frac{2s^5t^6w^4(5s^2w+6)}{135}$$

k=5

$$\frac{2s^6t^6w^5}{135}$$

$$\boxed{j = 8}$$

k=0

$$- s^2(16s^{10}t^6w^5+176s^8t^6w^4-376s^6t^6w^3-168s^8t^4w^3$$

$$- 4592s^4t^6w^2-1512s^6t^4w^2+10080s^2t^6w+11760s^4t^4w$$

$$+ 2520s^6t^2w-3780t^6-13230s^2t^4-8820s^4t^2-945s^6)/945$$

k=1

$$8s^3t^2w(10s^8t^4w^4+80s^6t^4w^3-159s^4t^4w^2-68s^6t^2w^2$$

$$-882s^2t^4w-336s^4t^2w+945t^4+1323s^2t^2+315s^4)/945$$

k=2

$$8s^4t^4w^2(t^2(s^2w(4s^2w(5s^2w+27)-175)-315)-21s^2(3s^2w+7))/945$$

k=3

$$\frac{8s^5t^4w^3(t^2(4s^2w(5s^2w+16)-63)-21s^2}{945}$$

k=4

$$-\frac{16s^6t^6w^4(5s^2w+7)}{945}$$

k=5

$$\frac{16s^7t^6w^5}{945}$$

$\boxed{j = 9}$

k=0

$$- s^3(2s^{10}t^6w^5+24s^8t^6w^4-66s^6t^6w^3-21s^8t^4w^3-848s^4t^6w^2$$
$$-210s^6t^4w^2+2394s^2t^6w+1932s^4t^4w+315s^6t^2w-1260t^2$$
$$-2646s^2t^4-1260s^4t^2-105s^6)/105$$

k=1

$$s^4t^2w(10s^8t^4w^4+88s^6t^4w^3-218s^4t^4w^2-63s^6t^2w^2-1344s^2t^4w$$
$$-387s^4t^2w+1890t^4+1764s^2t^2+315s^4)/105$$

k=2

$$-\frac{s^5t^4w^2(4t^2(s^2w(5s^2w(s^2w+6)-59)-126)-21s^2(3s^2w+8))}{105}$$

k=3

$$\frac{s^6t^4w^3(4t^2(s^2w(5s^2w+18)-21)-21s^2)}{105}$$

k=4

$$-\frac{2s^7t^6w^4(5s^2w+8)}{105}$$

k=5

$$\frac{2s^8t^6w^5}{105}$$

$$\boxed{j = 10}$$

k=0

$$- s^4(4s^{10}t^6w^5+52s^8t^6w^4-176s^6t^6w^3-42s^8t^4w^3-2394s^4t^6w^2$$
$$- 462s^6t^4w^2+8316s^2t^6w+4914s^4t^4w+630s^6t^2w-5670t^6$$
$$- 7938s^2t^4-2835s^4t^2-189s^6)/189$$

k=1

$$2s^5t^2w(10s^8t^4w^4+96s^6t^4w^3-286s^4t^4w^2-63s^6t^2w^2$$
$$- 1944s^2t^4w-420s^4t^2w+3402t^4+2268s^2t^2+315s^4)/189$$

k=2

$$2s^6t^4w^2(2t^2(s^2w(2s^2w(5s^2w+33)-153)-378)-63s^2(s^2w+3))/189$$

k=3

$$\frac{2s^7t^4w^3(4t^2(5s^2w(s^2w+4)-27)-21s^2)}{189}$$

k=4

$$-\frac{4s^8t^6w^4(5s^2w+9)}{189}$$

k=5

$$\frac{4s^9t^6w^5}{189}$$

Acknowledgement

The author would like to thank the Mathlab Group at the Laboratory for Computer Science at M. I. T. for the use of the MACSYMA symbolic computation system.

TECHNICAL REPORTS

OFFICE OF NAVAL RESEARCH CONTRACT N00014-67-A-0112-0030 (NR-042-034)

1. "Confidence Limits for the Expected Value of an Arbitrary Bounded Random Variable with a Continuous Distribution Function," T. W. Anderson, October 1, 1969.

2. "Efficient Estimation of Regression Coefficients in Time Series," T. W. Anderson, October 1, 1970.

3. "Determining the Appropriate Sample Size for Confidence Limits for a Proportion," T. W. Anderson and H. Burstein, October 15, 1970.

4. "Some General Results on Time-Ordered Classification," D. V. Hinkley, July 30, 1971.

5. "Tests for Randomness of Directions against Equatorial and Bimodal Alternatives," T. W. Anderson and M. A. Stephens, August 30, 1971.

6. "Estimation of Covariance Matrices with Linear Structure and Moving Average Processes of Finite Order," T. W. Anderson, October 29, 1971.

7. "The Stationarity of an Estimated Autoregressive Process," T. W. Anderson, November 15, 1971.

8. "On the Inverse of Some Covariance Matrices of Toeplitz Type," Raul Pedro Mentz, July 12, 1972.

9. "An Asymptotic Expansion of the Distribution of "Studentized" Classification Statistics," T. W. Anderson, September 10, 1972.

10. "Asymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions," T. W. Anderson, September 28, 1972.

11. "Population Mixing Models and Clustering Algorithms," Stanley L. Sclove, February 1, 1973.

12. "Asymptotic Properties and Computation of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance," John James Miller, November 21, 1973.

13. "Maximum Likelihood Estimation in the Birth-and-Death Process," Niels Keiding, November 28, 1973.

14. "Random Orthogonal Set Functions and Stochastic Models for the Gravity Potential of the Earth," Steffen L. Lauritzen, December 27, 1973.

15. "Maximum Likelihood Estimation of Parameters of an Autoregressive Process with Moving Average Residuals and Other Covariance Matrices with Linear Structure," T. W. Anderson, December, 1973.

16. "Note on a Case-Study in Box-Jenkins Seasonal Forecasting of Time series," Steffen L. Lauritzen, April, 1974.

TECHNICAL REPORTS (continued)

17. "General Exponential Models for Discrete Observations," Steffen L. Lauritzen, May, 1974.

18. "On the Interrelationships among Sufficiency, Total Sufficiency and Some Related Concepts," Steffen L. Lauritzen, June, 1974.

19. "Statistical Inference for Multiply Truncated Power Series Distributions," T. Cacoullos, September 30, 1974.

Office of Naval Research Contract N00014-75-C-0442 (NR-042-034)

20. "Estimation by Maximum Likelihood in Autoregressive Moving Average Models in the Time and Frequency Domains," T. W. Anderson, June 1975.

21. "Asymptotic Properties of Some Estimators in Moving Average Models," Raul Pedro Mentz, September 8, 1975.

22. "On a Spectral Estimate Obtained by an Autoregressive Model Fitting," Mituaki Huzii, February 1976.

23. "Estimating Means when Some Observations are Classified by Linear Discriminant Function," Chien-Pai Han, April 1976.

24. "Panels and Time Series Analysis: Markov Chains and Autoregressive Processes," T. W. Anderson, July 1976.

25. "Repeated Measurements on Autoregressive Processes," T. W. Anderson, September 1976.

26. "The Recurrence Classification of Risk and Storage Processes," J. Michael Harrison and Sidney I. Resnick, September 1976.

27. "The Generalized Variance of a Stationary Autoregressive Process," T. W. Anderson and Raul P.Mentz, October 1976.

28. "Estimation of the Parameters of Finite Location and Scale Mixtures," Javad Behboodian, October 1976.

29. "Identification of Parameters by the Distribution of a Maximum Random Variable," T. W. Anderson and S.G. Ghurye, November 1976.

30. "Discrimination Between Stationary Guassian Processes, Large Sample Results," Will Gersch, January 1977.

31. "Principal Components in the Nonnormal Case: The Test for Sphericity," Christine M. Waternaux, October 1977.

32. "Nonnegative Definiteness of the Estimated Dispersion Matrix in a Multivariate Linear Model," F. Pukelsheim and George P.H. Styan, May 1978.

TECHNICAL REPORTS (continued)

33. "Canonical Correlations with Respect to a Complex Structure," Steen A. Andersson, July 1978.

34. "An Extremal Problem for Positive Definite Matrices," T.W. Anderson and I. Olkin, July 1978.

35. " Maximum likelihood Estimation for Vector Autoregressive Moving Average Models," T. W. Anderson, July 1978.

36. "Maximum likelihood Estimation of the Covariances of the Vector Moving Average Models in the Time and Frequency Domains," F. Ahrabi, August 1978.

37. "Efficient Estimation of a Model with an Autoregressive Signal with White Noise," Y. Hosoya, March 1979.

38. "Maximum Likelihood Estimation of the Parameters of a Multivariate Normal Distribution,"T.W. Anderson and I. Olkin, July 1979.

39. "Maximum Likelihood Estimation of the Autoregressive Coefficients and Moving Average Covariances of Vector Autoregressive Moving Average Models," Fereydoon Ahrabi, August 1979.

40. "Smoothness Priors and the Distributed Lag Estimator," Hirotugu Akaike, August, 1979.

41. "Approximating Conditional Moments of the Multivariate Normal Distribution," Joseph G. Deken, December 1979.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 41
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): (6) APPROXIMATING CONDITIONAL MOMENTS OF THE MULTIVARIATE NORMAL DISTRIBUTION
5. TYPE OF REPORT & PERIOD COVERED: (9) Technical Rep[redacted]t
6. ~~PERFORMING ORG. REPORT NUMBER~~
7. AUTHOR(s): (10) JOSEPH G. DEKEN
8. CONTRACT OR GRANT NUMBER(s): (15) N00014-75-C-0442
9. PERFORMING ORGANIZATION NAME AND ADDRESS:
   Department of Statistics
   Stanford University
   Stanford, California 94305
   (12) 29
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: (NR-042-034)
11. CONTROLLING OFFICE NAME AND ADDRESS:
    Office of Naval Research
    Statistics & Probability Program Code 436
    Arlington, Virginia 22217
    (11)
12. ~~REPORT DATE~~: Dec[redacted]79
13. ~~NUMBER OF PAGES~~: 22
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): (14) TR-41
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report):

    APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number):

    Multiple integration, numerical integration, multivariate normal distribution, EM estimation.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

A practical method for computing the conditional expectation of a polynomial in the components of a multivariate normal random variable X, when X is restricted to a subset of R^p, is given. This method makes the application of certain missing data techniques possible in cases where repeated numerical integration is not feasible.

R superscript p

DD FORM 1473, 1 JAN 73 — EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

332580 alt