

EXPANSIONS FOR REPEATED INTEGRALS OF PRODUCTS WITH APPLICATIONS TO THE MULTIVARIATE NORMAL

CHRISTOPHER S. WITHERS¹ AND SARALEES NADARAJAH²

Abstract. We extend Leibniz' rule for repeated derivatives of a product to multivariate integrals of a product. As an application we obtain expansions for $P(a < Y < b)$ for $Y \sim N_p(0, V)$ and for repeated integrals of the density of Y . When $V^{-1}y > 0$ in R^3 the expansion for $P(Y < y)$ reduces to one given by [H. Ruben *J. Res. Nat. Bureau Stand. B* **68** (1964) 3–11]. in terms of the moments of $N_p(0, V^{-1})$. This is shown to be a special case of an expansion in terms of the multivariate Hermite polynomials. These are given explicitly.

Mathematics Subject Classification. 60E05, 62H05.

Received July 19, 2009. Revised 1st February, 2010.

1. INTRODUCTION AND SUMMARY

Expansions of the multivariate normal density are commonly used in statistical theory and its applications. One most common use is to compute multinormal probabilities, for example, rectangular and orthant multinormal probabilities arising in the multivariate probit model, the multivariate ordinal response model, multivariate paired comparisons and Thurstonian models for rankings; such probabilities also arise in almost every area of science, engineering and medicine. Many of the known procedures for computing multinormal probabilities are based on series truncations. Another common use is to find the approximate distribution of a specified statistic. The exact distribution of a statistic is often too complicated even when the data are multinormal. Examples include distributions of the eigenvalues and eigenvectors of the sample covariance matrix for a normal population.

The calculation of multinormal probabilities is one reason why we need expansions for repeated integrals of products of univariate and multivariate normal densities. The need for such expansions arises in many other areas too. We mention five applications:

- in the calculation of moments of truncated normal distribution [2];
- in the expression of the non-central t density [2];
- in the posterior distribution of a Poisson variate with chi-squared prior for the squared mean parameter of the Poisson variate [2];
- in the calculation of shape distributions [3,8];

Keywords and phrases. Asymptotic expansion, Leibniz' rule, repeated integrals of products, multivariate Hermite polynomials, multivariate normal.

¹ Applied Mathematics Group Industrial Research Limited Lower Hutt, New Zealand

² School of Mathematics University of Manchester Manchester M13 9PL, UK. mbbssn2@manchester.ac.uk

- in directional statistics with respect to computing the mean resultant length of the spherically projected normal distribution [10].

Each of these applications involves explicitly repeated integrals of products of univariate and multivariate normal densities. We refer the readers to the cited references for further details.

There is a large literature on expansions for the multivariate normal distribution. We refer the readers to Tong [14], von Rosen ([15], Sects. 3 and 4), Kotz *et al.* ([6], Chap. 45), Kotz and Nadarajah ([7], Chap. 6), Kollo and von Rosen ([5], Chap. 1, p. 152) and Kollo and von Rosen ([5], Chap. 3) for comprehensive reviews. Further motivation for expansions of the multivariate normal can be found in these references.

The aim of this paper is to provide expansions based on an extension of Leibniz' rule. The results are organised as follows. In Section 2 we extend Leibniz' rule for differentiating a product to obtain an expansion for repeated integrals of a product of q functions $\phi_1(y) \dots \phi_q(y)$ from C^p to C , the complex numbers, in terms of the repeated integrals of ϕ_1 and the derivatives of ϕ_2, \dots, ϕ_q . For example, if $q = 2$ and $\phi_1(y) = 1$ and ϕ_2 is the density of a random variable Y in R^p for R , the reals, this gives 2^p Taylor series expansions for $P(x < Y < y)$ in powers of $y - x$. These expansions are given in Section 3 together with expansions about zero. In Section 4 we take $\phi_1(y) = \exp(y'z)$ and $\phi_2 = \phi_V$, the density of $Y \sim N_p(0, V)$. We obtain expansions in inverse powers of $z = V^{-1}\mu$ for the 2^p probabilities $P(Y_j > \text{ or } < \mu_j, 1 \leq j \leq p)$ in terms of the multivariate Hermite polynomials and the moments of $N_p(0, V^{-1})$. These polynomials and moments are given in the appendix. The special case $P(Y < \mu)$ for $z > 0$ was obtained by [11].

We use the following notation: $N = \{0, 1, 2, \dots\}$, $Z = \{\dots, -1, 0, 1, \dots\}$, \overline{C} is the closure of C , that is, it includes all points at ∞ . For y in C and λ in N ,

$$(y)_\lambda = y(y - 1) \dots (y - \lambda + 1) = \Gamma(y + 1)/\Gamma(y - \lambda + 1)$$

and $\binom{y}{\lambda} = (y)_\lambda/\lambda!$ the binomial coefficient, so

$$\binom{-y}{\lambda} = (-1)^\lambda \binom{y - 1 + \lambda}{\lambda}. \tag{1.1}$$

Fix y_0 in \overline{C}^p . For y, z in \overline{C}^p , λ in N^p and v in Z^p set

$$y^z = \prod_{j=1}^p y_j^{z_j}, (y)_\lambda = \prod_{j=1}^p (y_j)_{\lambda_j}, \binom{y}{\lambda} = \prod_{j=1}^p \binom{y_j}{\lambda_j},$$

$$\lambda! = \prod_{j=1}^p \lambda_j!, \Gamma(y) = \prod_{j=1}^p \Gamma(y_j), \partial_y^v = \prod_{j=1}^p \partial_{y_j}^{v_j},$$

where $\partial_y = \partial/\partial y$, $\partial_y^{-1} = \int_{y_0}^y dy = (-1)^p \int_y^{y_0} dy$, so $\partial_y^{-1} = \int_{y_0}^{y_j} dy_j$, $y \times z = (y_1 z_1, \dots, y_p z_p)'$, $\text{Re}z = (\text{Re}z_1, \dots, \text{Re}z_p)'$, $|v| = \sum_1^p v_j$ and $(-)^V = (-1)^{|V|}$. In N^p , 1 is the vector of 1's and 0 is the vector of 0's. For $\phi : C^p \rightarrow C$ a given function,

$$H_\lambda(y) = \phi(y)^{-1} (-\partial_y)^\lambda \phi(y). \tag{1.2}$$

For y, z in R^p , $\text{sign}y = (\text{sign}y_1, \dots, \text{sign}y_p)'$, $m(y, z) = (m(y_1, z_1), \dots, m(y_p, z_p))'$ for $m = \text{min or max}$, $y < z$ means $y_j < z_j$ for $1 \leq j \leq p$, $y \leq z$ means $y_j \leq z_j$ for $1 \leq j \leq p$, \sum_λ sums over λ in N^p . We shall also denote $\phi_j^{(\lambda)}(y) = \partial_y^\lambda \phi_j(y)$ for $j = 1, 2, \dots, q$.

2. LEIBNIZ' RULE FOR INTEGRALS

Theorem 2.1 extends Leibniz' rule for the product of q functions.

Theorem 2.1. *Leibniz' rule for differentiating a product can be written*

$$\partial_y^v \phi_1(y)\phi_2(y) = \sum_{\lambda} \binom{v}{\lambda} \phi_1^{(v-\lambda)}(y)\phi_2^{(\lambda)}(y) \tag{2.1}$$

for v in N^p , where $\{\phi_n\}$ are functions from C^p to C and $y \in C^p$. Repeated application of (2.1) implies for v in N^p

$$\partial_y^v \prod_{n=1}^q \phi_n(y) = \sum_{\lambda} v! \prod_{n=1}^q \phi_n^{(\lambda_n)}(y)/\lambda_n!, \tag{2.2}$$

or equivalently

$$\partial_y^v \prod_{n=1}^q \phi_n(y) = \sum_{\lambda_2} \cdots \sum_{\lambda_q} (v)_{\lambda_1} \phi_1^{(\lambda_1)}(y) \prod_{n=2}^q \phi_n^{(\lambda_n)}(y)/\lambda_n!, \tag{2.3}$$

where \sum_{λ} sums over $\lambda_1, \dots, \lambda_q \in N^p$ such that $v = \sum_{n=1}^q \lambda_n$. Both (2.1) and (2.3) hold as asymptotic expansions for v in Z^p provided that for $\{\lambda_n\}$ of (2.3) with $\binom{v}{\lambda} \neq 0$ and $1 \leq j \leq p$,

$$\text{if } v_j < 0 \text{ then } \prod_{n=1}^q \phi_n^{(\lambda_n)}(y) \rightarrow 0 \tag{2.4}$$

as $y_j \rightarrow y_{0j}$.

Proof. Suppose first $p = 1$. Integrating by parts if $\phi_2(y)\partial_y^{-1}\phi_1(y) \rightarrow 0$ as $y \rightarrow y_0$ then $\partial_y^{-1}\phi_1(y)\phi_2(y) = \phi_2(y)\partial_y^{-1}\phi_1(y) - \partial_y^{-1}\{\phi_2^{(1)}(y)\partial_y^{-1}\phi_1(y)\}$ which we write as $I = A + I B$ on $\phi_1 \phi_2$, where $I_y = \partial_y^{-1}$, $A_y = \partial_{1y}^{-1}$, $B_y = -\partial_{2y}\partial_{1y}^{-1}$ and ∂_{ny} is ∂_y acting on ϕ_n only. So, if $\phi_2^{(n-1)}(y)\partial_y^{-n}\phi_1(y) \rightarrow 0$ as $y \rightarrow y_0$ for $1 \leq n \leq m$ then

$$\begin{aligned} I &= A \sum_{n=0}^{m-1} B^n + I B^m \text{ on } \phi_1\phi_2 \\ &= A(1 - B)^{-1} \text{ if } m = \infty \text{ and } I B^m \phi_1\phi_2 \rightarrow 0 \text{ as } m \rightarrow \infty. \end{aligned}$$

So, for $v = -n < 0$, since A and B commute, (2.4) at $q = 2$ implies

$$\partial^v \phi_1\phi_2 = I^n \phi_1\phi_2 = A^n (1 - B)^{-n} \phi_1\phi_2 = A^n \sum_{\lambda} \binom{-n}{\lambda} (-B)^\lambda \phi_1\phi_2,$$

which is equal to the right hand side of (2.1) at y . So, (2.1) holds for v in Z^p provided that for $1 \leq j \leq n$, $(I B^m)^j \phi_1\phi_2 \rightarrow 0$ as $m \rightarrow \infty$. If we only seek asymptotic expansions and are not primarily interested in convergence, we may ignore this condition.

For $p \geq 1$, (2.1) now follows from $\partial_y^v = \prod_{j=1}^p \partial_{y_j}^{v_j}$. Replacing ϕ_2 by $\phi_2 \cdots \phi_q$ and applying (2.2) with $v = \lambda$, we obtain (2.3) for v in Z^p . □

Note 2.1. By reversal of order in integration for $m \geq 1$ in N^p ,

$$\partial_y^{-m} \phi_1(y) = \int_{y_0}^y (y-x)^{m-1} \phi_1(x) dx / (m-1)! \tag{2.5}$$

or equivalently $(-\partial_y)^{-m} \phi_1(y) = \int_y^{y_0} (x-y)^{m-1} \phi_1(x) dx / (m-1)!$ so (2.1) with $v = -n$ and (1.1) imply for $n \geq 0$ in N^p ,

$$\begin{aligned} n! \left(\int_{y_0}^y dy \right)^{n+1} \phi_1(y) \phi(y) &= \int_{y_0}^y (y-x)^n \phi_1(x) \phi(x) dx \\ &= \sum_{\lambda} (-)^{|\lambda|} \phi^{(\lambda)}(y) (\lambda!)^{-1} \int_{y_0}^y (y-x)^{n+\lambda} \phi_1(x) dx \\ &= \phi(y) \sum_{\lambda} H_{\lambda}(y) (\lambda!)^{-1} \int_{y_0}^y (y-x)^{n+\lambda} \phi_1(x) dx \end{aligned}$$

for $H_{\lambda}(y)$ of (1.2).

Note 2.2. It should be possible to extend (2.1) and (2.3) to v in C^p using

$$\phi_1^{(v)}(y) = (2\pi i)^{-p} \Gamma(v+1) \int (y-x)^{-v-1} \phi_1(x) dx,$$

where x_j is integrated along a contour in C around y_j . For v_j in $\{-1, -2, \dots\}$, one would use (2.5) instead.

Note 2.3. Now suppose that the derivatives of $\phi_2, \dots, \phi_n, \{\phi_n^{(\lambda)}(y), n \geq 2, \lambda$ in $N^p\}$, are each bounded near $y = y_0$. Then (2.4) holds if for $\lambda \leq v$ and $1 \leq j \leq p$

$$\text{if } \binom{v}{\lambda} \neq 0 \text{ and } v_j < 0 \text{ then } \phi_1^{(\lambda)}(y) \rightarrow 0 \text{ as } y \rightarrow y_0. \tag{2.6}$$

This holds if $v \leq -1$ and the right hand side of (2.5) is bounded near $y = y_0$ for $m \geq 1$. For v in Z^p set $J = \{1 \leq j \leq p : v_j < 0\}$ and $K = \{1 \leq j \leq p : v_j \geq 0\}$. Then $n_j = -\lambda_j - 1 \geq 0$ for j in J if $\lambda_j \geq v_j$ by (2.5). So,

$$\{\lambda_j \leq v_j < 0, j \in J\} \text{ and } \{\lambda_j > 0, j \in K\} \tag{2.7}$$

implies

$$\phi_1^{(\lambda)}(y) = \prod_{j \in J} \int_{y_0}^{y_j} (y_j - x_j)^{n_j} \phi_1^{(\lambda_K)}(x) dx_j / n_j!, \tag{2.8}$$

where $x_k = y_k$ for k in K and $\phi_1^{(\lambda_K)}(x) = \prod_{k \in K} \partial_{x_k}^{\lambda_k} \phi_1(x)$. So, (2.6) holds if (2.7) implies that the right hand side of (2.8) is bounded near $y = y_0$. This illustrates the purely notational difficulty of applying (2.1) for v in Z^p as opposed to the cases $v \geq 0$ or $v \leq -1$.

We now set $q = 2, \phi = \phi_2$ and fix z in C^p .

Example 2.1. Take $\phi_1(y) = 1$ and y_0 in C^p . By (2.5), $\partial_y^{-m}1 = (y - y_0)^m/m!$ for $m \in N^p$, so by (2.1), for n in N^p

$$\begin{aligned} \partial_y^{-n} \phi(y) &= \left(\int_{y_0}^y dy \right)^n \phi(y) = \sum_{\lambda} \binom{-n}{\lambda} \phi^{(\lambda)}(y) (y - y_0)^{n+\lambda} / (n + \lambda)! \\ &= (n - 1)!^{-1} \phi(y) \sum_{\lambda} H_{\lambda}(y) (y - y_0)^{n+\lambda} (n + \lambda)^{-1} / \lambda! \end{aligned} \tag{2.9}$$

if also $n \geq 1$. If we compare this with

$$\partial_y^{-n} \phi(y) = \left(\int_{y_0}^y dy \right)^n \phi(y) = \sum_{\lambda} \phi(y_0) (y - y_0)^{\lambda+n} / (\lambda + n)!$$

obtained by expanding $\phi(y)$ about y_0 and multiplying by ∂_y^n , we obtain the identity

$$\sum_{j=0}^{\lambda} (-)^j / \{(n + j)!(\lambda - j)!\} = \sum_{k=0}^{\lambda} (-)^{\lambda-k} / \{k!(\lambda + n - k)!\} = 1 / \{(\lambda + n)\lambda!(n - 1)!\}$$

which we have not managed to prove directly. To state the extension of (2.1) to general v in Z^p , one uses

$$\phi_1^{(v-\lambda)}(y) = I(\lambda_k = v_k \text{ for } k \in K) \prod_{j \in J} (y_j - y_{0j})^{m_j} / m_j!,$$

where now $m = \lambda - v$, $J = \{1 \leq j \leq p : m_j > 0\}$ and $K = \{1 \leq j \leq p : m_j \leq 0\}$.

Example 2.2. Take $\phi_1(y) = \exp(y'z)$, where $z_1, \dots, z_p \neq 0$. For $p = 1$, $n \in Z$ and $-\infty \leq y_0 \text{ Re}z < \infty$,

$$\partial_y^{-n} \exp(yz) = z^{-n} \exp(yz) A_n(z(y - y_0)),$$

where

$$A_n(h) = \begin{cases} 1 - \exp(-h) \sum_{j=0}^n h^j / j!, & \text{for } n > 0 \text{ and } h \text{ finite,} \\ 1, & \text{for } n \leq 0 \text{ or } \text{Re}h = \infty. \end{cases}$$

So, for $p \geq 1$, n in Z^p and $-\infty \leq y_0 \times \text{Re}z < \infty$,

$$\partial_y^n \exp(y'z) = z^{-n} \exp(y'z) A_n(z \times (y - y_0)),$$

where

$$A_n(h) = \prod_{j=1}^p A_{n_j}(h_j). \tag{2.10}$$

So,

$$\begin{aligned} \left(\int_{y_0}^y dy \right)^n \phi(y) \exp(y'z) &= \exp(y'z) \sum_{\lambda} \binom{-n}{\lambda} \psi^{(\lambda)}(y) z^{-n-\lambda} A_{n+\lambda}(z \times (y - y_0)) \\ &= \phi(y) \exp(y'z) \sum_{\lambda} \binom{n-1+\lambda}{\lambda} H_{\lambda}(y) z^{-n-\lambda} A_{n+\lambda}(z \times (y - y_0)) \end{aligned} \tag{2.11}$$

or equivalently, for n in Z^p

$$\partial_y^n \phi(y) \exp(y'z) = \exp(y'z) \sum \binom{n}{\lambda} \phi^{(\lambda)}(y) z^{n-\lambda} A_{-n+\lambda}(z \times (y - y_0)).$$

So, for finite h

$$A_n(h) = O(\exp(-h)h^{n+1}) \begin{cases} = 0, & \text{if } h = 0, \\ \rightarrow 0, & \text{as } h \rightarrow \infty \text{ in } C^p, \\ = O(h^{n+1}), & \text{as } n \rightarrow \infty \text{ in } N^p, \end{cases}$$

and $A_n(h)$ is not defined if for some j , $1 \leq j \leq p$, $n_j > 0$ and $\text{Re}h_j = -\infty$. In particular, for n and v in N^p and $y_0 = -\infty \times \text{Re}z$,

$$\left(\int_{y_0}^y dy \right)^n \phi(y) \exp(y'z) = \sum_{\lambda} \binom{-n}{\lambda} \phi^{(\lambda)}(y) z^{-n-\lambda} \exp(y'z) \tag{2.12}$$

and

$$\begin{aligned} \left(\int_{y_0}^y dy \right)^n (-\partial_y)^\gamma \phi(y) \exp(y'z) &= \exp(y'z) \sum_{\lambda} (-)^\lambda \binom{-n}{\lambda} z^{-n-\lambda} (-\partial_y)^{\gamma+\lambda} \phi(y) \\ &= \exp(y'z) \sum_{\lambda} z^{-1-\lambda} (-\partial_y)^{\gamma+\lambda} \phi(y) \text{ if } n = 1. \end{aligned}$$

Further formulae may be obtained by transforming (2.12). For example, multiplying (2.12) by ∂_z^m for m in N^p ,

$$\begin{aligned} \left(\int_{y_0}^y dy \right)^n \phi(y) \exp(y'z) y^m &= \sum_{\lambda} \binom{-n}{\lambda} \phi^{(\lambda)}(y) \sum_k z^{-n-\lambda+m-2k} (-n-\lambda)_k \exp(y'z) \\ &= \phi(y) \exp(y'z) z^{m-n} \sum_j z^{-j} c_{jmn}(y), \end{aligned} \tag{2.13}$$

where

$$c_{jmn}(y) = \sum_k (-)^k \binom{m}{k} \binom{n-1+j-k}{j-k} H_{j-2k}(y) / (j-2k)! = c_j$$

say. So, $c_0 = 1$ and if $p = 1$ then

$$\begin{aligned} c_1 &= nH_1, \\ c_2 &= \binom{n+1}{2} H_2/2 - mn, \\ c_3 &= \binom{n+2}{3} H_3/3! - m \binom{n+1}{2} H_1, \\ c_4 &= \binom{n+3}{4} H_4/4! - m \binom{n+2}{3} H_2/2 + \binom{m}{2} \binom{n+1}{2}, \end{aligned}$$

where $H_r = H_r(y)$.

Note 2.4. Note that (2.12) and (2.13) remain good as asymptotic expansions in inverse powers of z for y_0 finite if $\text{sign}(y - y_0) = \text{sign}z$. For, if $p = 1$ and $h = z(y - y_0)$ then $z^{-n} \exp(yz)(1 - A_n(h))$ is exponentially small as $|z| \rightarrow \infty$ for fixed $y - y_0 \neq 0$.

Example 2.3. Take $\phi_1(y) = y^z$ and y_0 finite. Set

$$B_v(y, z) = \partial_y^v y^z = \prod_{j=1}^p \partial_{y_j}^{v_j} y_j^{z_j} = \prod_{j=1}^p B_{v_j}(y_j, z_j).$$

So, $B_v(y, z) = y^{z-v}(z)_v$ if $v \geq 0$ and if $n = -v > 1$ then

$$\begin{aligned} B_v(y, z) &= y^{z+n}/(z+n)_n - \sum_{k=1}^n y_0^{z+k}(y - y_0)^{n-k}/\{(z+k)_k(n-k)!\}, \\ \partial_y^v y^z \phi(y) &= \sum_{\lambda} \binom{v}{\lambda} B_{v-\lambda}(y, z) \phi^{(\lambda)}(y). \end{aligned}$$

Now fix v_0 in N^p and take $\phi_2(y) = (-\partial_y)^{v_0} \phi(y)$. Then

$$\partial_y^v \{y^z (-\partial_y)^{v_0} \phi(y)\} = \sum_{\lambda} (-)^{\lambda} \binom{v}{\lambda} B_{v-\lambda}(y, z) H_{v_0+\lambda}(y) \phi(y)$$

for $H_{\lambda}(y)$ of (1.2).

3. MULTIVARIATE RECTANGLE PROBABILITIES

Theorem 3.1 provides various Taylor series expansions for $P(a < Y < b)$ in powers of $b - a$, for a random variable Y in R^p .

Theorem 3.1. Suppose Y is a random variable in R^p with density ϕ and distribution Φ . For $a < b$ in R^p ,

$$P(a < Y < b) = (-)^q G(y, y - y_0), \tag{3.1}$$

where

$$q = \sum_{j=1}^p I(y_j < y_{0j}), \quad a = \min(y_0, y), \quad b = \max(y_0, y) \tag{3.2}$$

and

$$G(y, h) = \phi(y) \sum_{\lambda} H_{\lambda}(y) h^{\lambda+1}/(\lambda + 1)! \tag{3.3}$$

for $H_{\lambda}(y)$ of (1.2).

Proof. The Taylor series expansion in (3.1) follows by (2.9) with $n = 1$. □

There are 2^p choices for (y_0) since $(y_{0j}, y_j) = (a_j, b_j)$ or (b_j, a_j) . Corollary 3.1 considers some of these choices.

Corollary 3.1. *If $(y_0, y) = (a, b)$ then (3.1) gives*

$$P(a < Y < b) = \phi(b) \sum_{\lambda} H_{\lambda}(b)(b - a)^{\lambda+1}/(\lambda + 1)!$$

If $(y_0, y) = (b, a)$ then (3.1) gives

$$P(a < Y < b) = \phi(a) \sum_{\lambda} H_{\lambda}(a)(a - b)^{\lambda+1}/(\lambda + 1)!$$

Note 3.1. If Y is symmetric, that is if $P(Y < y) = P(-Y < y)$ or $\phi(-y) = \phi(y)$, then $(-)^{\lambda}H_{\lambda}(-y) = H_{\lambda}(y)$ and for $|\lambda|$ odd $H_{\lambda}(0) = 0$ which makes expansions in $\{H_{\lambda}(0)\}$ attractive using

$$\sum_{\lambda} c(\lambda)H_{\lambda}(0) = \sum_{r=0}^{\infty} \sum_{\lambda}^{2r} c(\lambda)H_{\lambda}(0).$$

If also $Y \sim N_p(0, V)$ then by Theorem A.2 in the appendix

$$H_{\lambda}(0) = E(iW)^{\lambda}, \tag{3.4}$$

where $W \sim N_p(0, V^{-1})$ and $i = \sqrt{-1}$.

Corollaries 3.2 and 3.3 considers a rectangular decomposition, where a and b satisfy

$$a_j b_j = 0 \tag{3.5}$$

for $1 \leq j \leq p$. Any rectangle in R^p can be decomposed as sums and differences of 2^p rectangles satisfying (3.5), since for $p = 1$, $[a, b] = [a, 0] + [0, b]$ if $a < 0 < b$, $[a, b] = [0, b] - (0, a)$ if $0 < a < b$ and $[a, b] = [a, 0] - (b, 0)$ if $a < b < 0$. The corollaries use this fact to establish representations for $P(a < Y < b)$.

Corollary 3.2. *Suppose (3.5) holds. Then, taking $y = 0$ and $y_0 = a_j$ if $a_j \neq 0$ and $y_0 = b_j$ if $b_j \neq 0$, (3.1) gives*

$$P(a < Y < b) = (-)^{p-Q}G(0, y_0) = (-)^{p-Q}\phi(0) \sum_{\lambda} H_{\lambda}(0)y_0^{\lambda+1}/(\lambda + 1)!, \tag{3.6}$$

where $Q = \sum\{I(a_j > 0) + I(b_j > 0)\}$.

Corollary 3.3. *Under the assumptions of Corollary 3.2,*

$$P(a < Y < b) = \prod_{j=1}^p (B_j - A_j) \tag{3.7}$$

with $X_1 \dots X_p$ replaced by $F(x) = G(0, x)$, where $x_j = a_j$ for $X_j = A_j$ and $x_j = b_j$ for $X_j = B_j$.

Proof. First suppose $0 < a < b$. Then $[a, b] = \prod_1^p \{[0, b_j] - [0, a_j]\} = \prod_1^p (\overline{B_j} - \overline{A_j})$ say, so $P(a < Y < b) = \sum (-)^R P(0 < Y < B)$ summed over the 2^p possibilities $B_j = a_j$ or b_j , where R is the number of $\{a_j\}$ in B . So, (3.6) holds with $Q = p$ and so (3.7) holds. Now suppose $a_1 < 0 < b_1$ and $0 < a_j < b_j$ for $2 \leq j \leq p$. Then

$$[a, b] = ([a_1, 0] + (0, b_1)) \times \prod_2^p \{[0, b_j] - [0, a_j]\}$$

so $P(a < Y < b) = \sum (-)^S P(A < Y < B)$ summed over the 2^p possibilities $(A_1, B_1) = (a_1, 0)$ or $(0, b_1)$ and $(A_j, B_j) = (0, a_j)$ or $(0, b_j)$ for $2 \leq j \leq p$, and S is the number of $(A_j, B_j) = (0, a_j)$ for $j \geq 2$. The choice $(A_1, B_1) = (0, b_1)$ gives $S = R$ and the same contribution as for $0 < a < b$. A typical term with $(A_1, B_1) = (a_1, 0)$ is

$$(a_1, 0) \times \prod_2^{S+1} (0, a_j) \times \prod_{S+2}^p (0, b_j).$$

This gives $y_{0j} = a_j$ for $j \leq S + 1$ and $y_{0j} = b_j$ for $j > S + 1$, so $p - Q = 1$ in (3.6). This change in sign cancels with the change in sign from replacing $+[a_1, 0]$ in the previous case by $-[0, a_1]$. So, (3.7) still holds. A similar cancelation of sign change takes place if $a_1 < b_1 < 0$. (For example, if $p = 1$ then $P(a < Y < b) = P(a < Y < 0) - P(b < Y < 0)$, $P(a < Y < 0) = (-)^{1-Q} F(a)$ with $Q = 0$ and $P(b < Y < 0) = (-)^{1-Q} F(b)$ with $Q = 0$.) So, (3.6) holds for $a_1 < b_1$. Similarly, it holds for $a_j < b_j$, $2 \leq j \leq p$. This completes the proof of (3.7). \square

Example 3.1. For $p = 1, 2, 3$, (3.7) reduces to

$$P(a < Y < b) = F(b) - F(a),$$

$$P(a < Y < b) = F(b) - F(b_1, a_2) - F(a_1, b_2) + F(a),$$

and

$$\begin{aligned} P(a < Y < b) &= F(b) - F(b_1, b_2, a_3) - F(b_1, a_2, a_3) - F(a_1, b_2, b_3) + F(a_1, a_2, b_3) \\ &\quad + F(a_1, b_2, a_3) + F(b_1, a_2, a_3) - F(a), \end{aligned}$$

respectively.

4. MULTIVARIATE NORMAL PROBABILITIES

Theorem 4.1 provides expansions for $P(a < Y < b)$, $Y \sim N_p(0, V)$ in terms of multivariate Hermite polynomials.

Theorem 4.1. *Suppose $Y \sim N_p(0, V)$ with $\det V \neq 0$, density ϕ and distribution Φ . Fix z in R^p with $z_1 \dots z_p \neq 0$ and set $\mu = Vz$. Then*

$$P(a < Y < b) = (-)^q \phi(y - \mu) \sum_{\lambda} H_{\lambda}(y) z^{-1-\lambda} A_{1+\lambda}(h) = (-)^q F(y, y_0, z) \tag{4.1}$$

say for q of (3.2), for A_n of (2.10) and for $h = z \times (y - y_0)$ in $(-\infty, \infty]$, where $a + \mu = \min(y_0, y)$ and $b + \mu = \max(y_0, y)$. Furthermore, $\{H_{\lambda}(y)\}$ are the multivariate Hermite polynomials given explicitly in the appendix.

Proof. Note $\phi(y) \exp(y'z) = \phi(y - \mu)\alpha$, where $\alpha = \exp(\mu'V^{-1}\mu) = \exp(z'Vz)$. Dividing (2.11) by α gives for n in Z^p and $y_0 - \mu = y - \mu_0$

$$\begin{aligned} \left(\int_{y_0}^y dy \right)^n \phi(y - \mu) &= \left(- \int_{\mu_0}^{\mu} d\mu \right)^n \phi(\mu - y) = \left(\int_{\mu}^{\mu_0} d\mu \right)^n \phi(\mu - y) \\ &= \phi(y - \mu) \sum_{\lambda} \binom{n-1+\lambda}{\lambda} H_{\lambda}(y) z^{-n-\lambda} A_{n+\lambda}(h). \end{aligned}$$

Set $n = 1$ to obtain the result. \square

As in Section 3 we can choose (y_0, y) to satisfy $a + \mu = \min(y_0, y)$ and $b + \mu = \max(y_0, y)$ in 2^p ways so that (4.1) gives 2^p expansions. Two examples of this are illustrated in Corollary 4.1.

Corollary 4.1. *If $(y_0, y) = (a + \mu, b + \mu)$ then (4.1) gives*

$$P(a < Y < b) = \phi(b) \sum_{\lambda} H_{\lambda}(b + Vz) z^{-1-\lambda} A_{1+\lambda}((b-a) \times z).$$

If $(y_0, y) = (b + \mu, a + \mu)$ then (4.1) gives

$$P(a < Y < b) = (-)^q \phi(a) \sum_{\lambda} H_{\lambda}(a + Vz) z^{-1-\lambda} A_{1+\lambda}((a-b) \times z).$$

We now apply the rectangle decomposition of Section 3 with a and b replaced by $a + \mu$ and $b + \mu$, respectively. That is, we decompose $[a, b]$ into sums and differences of 2^p rectangles $[a, b]$ such that $(a_j + \mu_j)(b_j + \mu_j) = 0$ for $1 \leq j \leq p$, and see how Corollaries 3.2 and 3.3 change.

Corollary 4.2. *Suppose $(a_j + \mu_j)(b_j + \mu_j) = 0$ for $1 \leq j \leq p$. Taking $y = 0$ and $y_{0j} = a_j + \mu_j$ if $a_j + \mu_j \neq 0$ and $y_{0j} = b_j + \mu_j$ if $b_j + \mu_j \neq 0$, we see (4.1) gives*

$$P(a < Y < b) = (-)^Q F(0, y_0, z),$$

where

$$Q = \sum_{j=1}^p \{I(a_j + \mu_j > 0) + I(b_j + \mu_j > 0)\}.$$

Corollary 4.3. *Under the assumptions of Corollary 4.2,*

$$\begin{aligned} P(a < Y < b) &= (-)^q \phi(\mu) \sum_{\lambda} H_{\lambda}(0) z^{-1-\lambda} A_{1+\lambda}(a, b) \\ &= (-)^q \phi(\mu) \sum_{r=0}^{\infty} (-)^r \sum_{\lambda}^{2r} E W^{\lambda} z^{-1-\lambda} A_{1+\lambda}(a, b), \end{aligned} \quad (4.2)$$

where $W \sim N(0, V^{-1})$ and

$$A_{1+\lambda}(a, b) = \prod_{j=1}^p (B_j - A_j)$$

at $X_1 \dots X_p = A_{1+\lambda}(-z \times (x + \mu))$ for $x_j = a_j$ if $X_j = A_j$ and $x_j = b_j$ if $X_j = B_j$.

Proof. Note from (4.1) that

$$F(x) = (-)^q F(0, x + Vz, z) = (-)^q \phi(\mu) \sum_{r=0}^{\infty} (-)^r \sum_{\lambda}^{2r} E W^{\lambda} z^{-1-\lambda} A_{1+\lambda}(h)$$

provided $h = -z \times (x + \mu)$ is in $(-\infty, \infty]$. So, (3.7) can be written in the form (4.2). \square

Example 4.1. For $p = 1$, $A_{1+\lambda}(a, b) = A_{1+\lambda}(-bz - Vz^2) - A_{1+\lambda}(-az - Vz^2)$. For $p = 2$, $A_{1+\lambda}(a, b) = \sum_1^4 (-)^j A_{1+\lambda}(-z \times (x_j + \mu))$, where $x_2 = a$, $x_4 = b$, $x_1 = (b_1, a_2)'$ and $x_3 = (a_1, b_2)'$.

Note 4.1. If $b - a$ is small, it may be more accurate to use one of the 2^p expansions (3.1) or (4.1) rather than (3.7) with $F(x) = G(0, x)$ or (4.2).

The expansion (4.1) and derived expressions such as (4.2) are remarkable in that z is *any* element in R^p with non-zero components. The proviso to this is that $h > -\infty$ in $A_{1+\lambda}(h)$. So, in (4.2) $z_j > 0$ if $a_j = -\infty$ and $z_j < 0$ if $b_j = \infty$. See Corollary 4.4 for examples. Some further particular cases of (4.1) are considered by Corollary 4.5.

Corollary 4.4. Taking $a = -\infty$ in (4.2), for $z > 0$

$$\Phi(b) = P(Y < b) = (-)^q \phi(\mu) \sum_{\lambda} H_{\lambda}(0) z^{-1-\lambda} A_{1+\lambda}(-\infty, b),$$

where

$$\begin{aligned} A_{1+\lambda}(-\infty, b) &= A_{1+\lambda}(-bz - Vz^2) - 1 \text{ if } p = 1, \\ &= 1 + A_{1+\lambda}(-z \times (b + \mu)) - \sum_{j=1}^2 A_{1+\lambda}(-z_j(b_j + \mu_j)) \text{ if } p = 2 \end{aligned} \tag{4.3}$$

and so on, while for $z < 0$

$$\Phi(-a) = P(a < Y) = (-)^q \phi(\mu) \sum_{\lambda} H_{\lambda}(0) z^{-1-\lambda} A_{1+\lambda}(a, \infty),$$

where $A_{1+\lambda}(a, \infty) = 1 - A_{1+\lambda}(-az - Vz^2)$ if $p = 1$, $A_{1+\lambda}(a, \infty)$ is equal to the right hand side of (4.3) at $b = a$ if $p = 2$, and so on. In particular, for $p = V = 1$ we have by (3.4)

$$\Phi(x) = \phi(x) z^{-1} \sum_{r=0}^{\infty} (-z^2)^{-r} c_r A_{1+2r}^* \tag{4.4}$$

where

$$c_r = E(N(0, 1))^{2r} = (2r)! 2^{-r} r! = 1.3 \dots (2r - 1) \tag{4.5}$$

and

$$\begin{aligned} A_n^* &= \begin{cases} 1 - A_n(-xz - z^2), & \text{if } z > 0, \\ A_n(xz - z^2) - 1, & \text{if } z < 0 \end{cases} \\ &= \{1 - A_n(-x|z| - z^2)\} \text{sign}z, \end{aligned}$$

so the right hand side of (4.4) is symmetric about $z = 0$. For $z = -x$ this gives the well known expansion

$$\Phi(x) = \phi(x) |x|^{-1} \sum_{r=0}^{\infty} (-x^2)^{-r} c_r \tag{4.6}$$

for $x < 0$.

Note 4.2. The relative error in (4.6) is $\sim 10\%$ at $x = -3$ and $\sim 1\%$ at $x = -6$ if we truncate when

$$|(r + 1)\text{st term}| > \theta |r\text{th term}|, \text{ where } \theta = 1 \text{ or } 0.5. \tag{4.7}$$

Table 1 gives the relative error of (4.4) for selected values of $\Phi(x)$ and $z + x$ truncating in the same way. The x values are taken from Table 1.2 of [9]. The values of $\Phi(x)$ were computed using NAG.

TABLE 1. Relative error of (4.4) truncating as in (4.7).

			$\theta = 1$		$\theta = 0.5$	
x	$\Phi(x)$	$z + x$	Relative error	r	Relative error	r
-1.28155	0.1	-0.1	-53%	1	-53%	1
		0	-46%	1	-46%	1
		0.1	-48%	1	-48%	1
-2.32635	0.01	-0.1	35%	2	17%	1
		0	-6%	3	-7%	1
		0.1	-30%	3	-31%	1
-3.09023	0.001	-0.1	35%	4	36%	2
		0	-0.66%	5	1.1%	2
		0.1	-34%	5	-34%	3
-3.71902	0.0001	-0.1	40%	7	39%	3
		0	-0.0075%	7	-0.19%	3
		0.1	-40%	7	-40%	4
-4.74342	10^{-6}	-0.1	50%	11	50%	5
		0	-0.0001%	11	0.0003%	6
		0.1	-50%	12	-50%	6
-5.99781	10^{-9}	-0.1	62%	17	62%	9
		0	0.000001%	18	-0.00001%	9
		0.1	-62%	19	-62%	9

Corollary 4.5. Putting $y_0 = -\infty \times z$ and $x = y - \mu$ in (4.1) gives

$$P((Y - x) \times \text{sign} z < 0) = (-)^q \phi(x) \sum_{\lambda} H_{\lambda}(x + Vz) z^{-1-\lambda}, \tag{4.8}$$

where $q = \sum_{j=1}^p I(z_j < 0)$. In particular,

$$\Phi(x) = P(Y < x) = \phi(x) \sum_{\lambda} H_{\lambda}(x + Vz) z^{-1-\lambda} \text{ if } z > 0$$

and since $-Y$ and Y have the same distribution,

$$\Phi(-x) = P(Y > x) = (-)^q \phi(x) \sum_{\lambda} H_{\lambda}(x + Vz) z^{-1-\lambda} \text{ if } z < 0.$$

These can also be written as

$$\Phi(y - Vz) / \phi(y - Vz) = \sum_{\lambda} H_{\lambda}(y) z^{-1-\lambda} \text{ if } z > 0$$

and

$$\Phi(-y + Vz) / \phi(-y + Vz) = (-)^q \sum_{\lambda} H_{\lambda}(y) z^{-1-\lambda} \text{ if } z < 0,$$

respectively. Putting $z = -V^{-1}x$ in (4.8) gives

$$(-)^q P((Y - x) \times \text{sign } z < 0) = \phi(x) \sum_{\lambda} H_{\lambda}(0) z^{-1-\lambda} = \phi(x) \sum_{r=0}^{\infty} (-)^r \sum_{\lambda}^{2r} EW^{\lambda} z^{-1-\lambda} \tag{4.9}$$

by (3.4). In particular, if $z = V^{-1}\mu > 0$,

$$\Phi(-\mu) = \phi(\mu) \sum_{\lambda} H_{\lambda}(0) z^{-1-\lambda}, \tag{4.10}$$

an expansion given by [11], while if $z = V^{-1}\mu < 0$,

$$\Phi(\mu) = (-)^q \phi(\mu) \sum_{\lambda} H_{\lambda}(0) z^{-1-\lambda}.$$

All of the expansions in Corollary 4.5 reduce to (4.6) if $p = V = 1$. The expansions with $y = 0$ can be used to give upper and lower bounds: see Savage (1962) for some examples using the terms $r = 0, 1$ in (4.9).

Note 4.3. Suppose that

$$z_j \begin{cases} > 0, & \text{for } j \text{ in } J, \\ < 0, & \text{for } j \text{ in } K, \end{cases} \tag{4.11}$$

where $J \cup K = \{1, \dots, p\}$. Set

$$(Y_j^*, x_j^*, z_j^*) = \begin{cases} (Y_j, x_j, z_j), & \text{for } j \text{ in } J, \\ -(Y_j, x_j, z_j), & \text{for } j \text{ in } K \end{cases}$$

and $V^* = \text{cov}Y^*$. Then the left hand side of (4.8) is $P(Y^* < x^*)$, $(Vz)^* = V^*z^*$, $\phi_V = \phi$ satisfies $\phi_V(x) = \phi_{V^*}(x^*)$ and $H_{\lambda}(x, V) = H_{\lambda}(x)$ satisfies $(-)^{\lambda^*} H_{\lambda}(x^*, V^*) = H_{\lambda}(x, V)$, where $\lambda^* = \sum_{j=1}^p \lambda_j I(z_j < 0)$. So, (4.8) gives: for $z^* > 0$

$$P(Y^* < x^*) = \phi_{V^*}(x^*) \sum_{\lambda} H_{\lambda}(x^* + V^*z^*, V^*) (z^*)^{-1-\lambda}$$

which is just the $*$ version of (4.8) for $z > 0$. Conversely, we could have obtained (4.8) from the case $z > 0$.

Steck [13] commented that the condition $z > 0$ in (4.10) ‘‘appears to be a severe one’’. We now show how to remove this condition in order to find an expansion for $\Phi(x)$ for almost any x in R^p .

Note 4.4. Suppose $z = -V^{-1}x$ satisfies $z_1 \dots z_p \neq 0$. Define J, K by (4.11). Let q be the number of elements in K . Decompose $(-\infty, x)$ as $\prod_{j \in J} (-\infty, x_j) \times \prod_{k \in K} (-\infty, x_k)$. Then

$$\Phi(x) = P(Y < x) = C \prod_{k \in K} (B_k - A_k), \tag{4.12}$$

where

$$C \prod_{k \in K} X_k = P(Y_j < x_j \text{ for } j \in J, Y_k > x_k^* \text{ for } k \in K)$$

and $x_k^* = -\infty$ if $X_k = B_k$ and $x_k^* = x_k$ if $X_k = A_k$. The last of the 2^q terms in (4.12) is

$$\begin{aligned} (-)^q C \prod_{k \in K} A_k &= (-)^q P(Y_j < x_j \text{ for } j \in J, Y_k > x_k \text{ for } k \in K) \\ &= \phi(x) \sum_{\lambda} H_{\lambda}(0) z^{-1-\lambda} \text{ by (4.9).} \end{aligned}$$

The other $2^q - 1$ terms are of the form $\pm P(Y_j^* > \text{ or } < x_j^*, 1 \leq j \leq r)$, where $r < p$ and $Y^* \sim N_r(0, V^*)$ say. These can all be dealt with in the same way. So, by repeated application of (4.9) the dimension is reduced from p to 1.

Example 4.2. Suppose $p = 2$ and $z_1 < 0 < z_2$. Then $\Phi(x) = P(Y_2 < x_2) - P(Y_1 > x_1, Y_2 < x_2)$ is the sum of $\Phi_1(x_2 V_{22}^{-1/2})$ and the right hand side of (4.9), where Φ_1 is Φ for $N(0, 1)$.

Example 4.3. Suppose $p = 2$ and $z < 0$. Then $\Phi(x)$ is the sum of $1 - P_1 - P_2$ and the right hand side of (4.9), where $P_j = P(Y_j > x_j) = \Phi_1(-x_j V_{jj}^{-1/2})$.

Example 4.4. Suppose $p = 3$ and $z_1 < 0 < z_2, z_3$. Then $\Phi(x)$ is the sum of $P(Y_2 < x_2, Y_3 < x_3)$ and the right hand side of (4.9).

Example 4.5. Suppose $p = 3$ and $z_1, z_2 < 0 < z_3$. Then $\Phi(x)$ is the sum of $P(Y_3 < x_3) - \sum_1^2 P_j$ and the right hand side of (4.9), where $P_j = P(Y_j > x_j, Y_3 < x_3)$.

We end this section with some explicit formulas for $\partial_y^{-n} \phi(y)$ and $\partial_y^n \Phi(y)$.

Example 4.6. Suppose $p = V = 1$. Then

$$\begin{aligned} \int_{y_0}^y (y-x)\phi(x)dx &= \left(\int_{y_0}^y dy \right)^2 \phi(y) \\ &= \phi(y) - \phi(y_0) + y\{\Phi(y) - \Phi(y_0)\} \\ &= \phi(y) + y\Phi(y) \text{ if } y_0 = -\infty. \end{aligned}$$

For, the left and right hand sides are equal at $y = y_0$ and so are their partial derivatives with respect to y . Similarly,

$$\begin{aligned} 2 \left(\int_{y_0}^y dy \right)^3 \phi(y) &= y\phi(y) - 2y\phi(y_0) + y_0\phi(y_0) + (1+y^2)\{\Phi(y) - \Phi(y_0)\} \\ &= y\phi(y) + (1+y^2)\Phi(y) \text{ if } y_0 = -\infty. \end{aligned}$$

More generally, for $r \geq 0$

$$\begin{aligned} r! \left(\int_{y_0}^y dy \right)^{r+1} \phi(y) &= a_{r-1}(y)\{\phi(y) - \phi(y_0)\} + b_{r-1}(y, y_0)\phi(y_0) \\ &\quad + H_r^+(y)\{\Phi(y) - \Phi(y_0)\} \\ &= a_{r-1}(y)\phi(y) + H_r^+(y)\Phi(y) \text{ if } y_0 = -\infty, \end{aligned} \tag{4.13}$$

where $H_r^+(y)$ is $H_r(y)$ with all signs made positive:

$$H_r^+(y) = \sum \{y^{r-2j} (r)_{2j} 2^{-j} / j! : 0 \leq j \leq r/2\}$$

and a_j and b_j are polynomials of degree j . For, differentiating (4.13) gives

$$rH_{r-1}^+(y) = \partial_y H_r^+(y), \quad ra_{r-2}(y) = (\partial_y - y)a_{r-1}(y) + H_r^+(y),$$

and

$$-ra_{r-2}(y) + rb_{r-2}(y, y_0) = \partial_y(-a_{r-1}(y) + b_{r-1}(y, y_0)).$$

Now apply induction. The last equation gives the recursive formula for b_r :

$$b_r(y, y_0) = (r + 1) \int_{y_0}^y \{b_{r-1}(y, y_0) - a_{r-1}(y)\} dy + a_r(y) - a_r(y_0).$$

It would be interesting to know if (4.13) can be extended to $p > 1$.

Example 4.7. Set $\Phi_V(y) = \Phi(y)$, $\phi_V(y) = \phi(y)$, partition $y' = (y', y'_0)$, and similarly V is a 2×2 block matrix. Set $M_{12} = V_{12}V_{22}^{-1}$ and $V_{1.2} = V_{11} - V_{12}V_{22}^{-1}V_{21}$. Since $Y_1/Y_2 \sim \mathcal{N}(M_{12}Y_2, V_{1.2})$,

$$\Phi_V(y) = \int^{y_2} dy_2 \phi_{V_{22}}(y_2) \Phi_{V_{1.2}}(y_1 - M_{12}y_2).$$

Now suppose $p = 2$. Then

$$\partial_{y_2} \Phi_V(y) = \phi_{V_{22}}(y_2) \Phi_{V_{1.2}}(y_1 - M_{12}y_2),$$

$$\partial_{y_2}^2 \Phi_V(y) = \phi'_{V_{22}}(y_2) \Phi_{V_{1.2}}(y_1 - M_{12}y_2) - M_{12} \phi_{V_{22}}(y_2) \phi_{V_{1.2}}(y_1 - M_{12}y_2),$$

and so on.

Acknowledgements. The authors would like to thank the Editors-in-Chief, the Associate Editor and the two referees for carefully reading the paper and for their comments which greatly improved the paper.

A. THE MULTIVARIATE HERMITE POLYNOMIALS

Suppose $Y \sim N_p(0, V)$, and $\det V \neq 0$. Then $\phi(y) = (2\pi)^{-p/2}(\det V)^{-1/2} \exp(-y'V^{-1}y/2)$ and $\{H_\lambda(y)\}$ are the multivariate Hermite polynomials. They can be obtained iteratively from

$$H_{\lambda+e_j}(y) = (x_j + \partial_{y_j})H_\lambda(y), \quad H_0(y) = 1, \tag{A.1}$$

where e_j is the j th unit vector in R^p and $x = V^{-1}y$. However, an explicit form is easily given by using the dual form

$$\begin{aligned} H^{\alpha_1 \dots \alpha_r}(y) &= \phi(y)^{-1} D_{\alpha_1} \dots D_{\alpha_r} \phi(y) \text{ for } D = -\partial/\partial y \\ &= H_\lambda(y) \text{ for } |\lambda| = r \text{ and } \lambda_j \text{ the number of } j\text{'s in } \{\alpha_1, \dots, \alpha_r\}. \end{aligned}$$

Set $x = V^{-1}y$ and $(V^{ij}) = V^{-1}$. By (A.1),

$$H^{\alpha_1 \dots \alpha_r}(y) = (x_{\alpha_r} + D_{\alpha_r})H^{\alpha_1 \dots \alpha_{r-1}}(y),$$

so

$$\begin{aligned} H^{\alpha_1}(y) &= x_{\alpha_1}, \quad H^{\alpha_1\alpha_2}(y) = x_{\alpha_1}x_{\alpha_2} - V^{\alpha_1\alpha_2}, \\ H^{\alpha_1\alpha_2\alpha_3}(y) &= x_{\alpha_1}x_{\alpha_2}x_{\alpha_3} - \sum^3 x_{\alpha_1}V^{\alpha_2\alpha_3}, \\ H^{\alpha_1\dots\alpha_4}(y) &= x_{\alpha_1}\dots x_{\alpha_4} - \sum^6 x_{\alpha_1}x_{\alpha_2}V^{\alpha_3\alpha_4} + \sum^3 V^{\alpha_1\alpha_2}V^{\alpha_3\alpha_4}, \end{aligned}$$

where \sum^j sums over all j permutations of $\alpha_1 \dots \alpha_r$ giving distinct terms. Comparing these with the univariate Hermite polynomials:

$$H_r(x) = \sum \{(-1)^{j-1} R(r, j) x^{r-2j} : 0 \leq j \leq r/2\},$$

where $R(r, j) = (r)_{2j} 2^{-j} / j!$, we see the following result.

Theorem A.1. *With the notation as above,*

$$H^{\alpha_1\dots\alpha_r}(y) = \sum \{(-1)^{j-1} \sum^{R(r,j)} x_{\alpha_1} \dots x_{\alpha_{r-2j}} V^{\beta_1\beta_2} \dots V^{\beta_{2j-1}\beta_{2j}} : 0 \leq j \leq r/2\},$$

where $\beta_k = \alpha_{r-2j+k}$. Setting $y = 0$ gives

$$(-1)^r H^{\alpha_1\dots\alpha_{2r}}(0) = \sum^{c_r} V^{\alpha_1\alpha_2} \dots V^{\alpha_{2r-1}\alpha_{2r}} \tag{A.2}$$

for $R(2r, r) = c_r$ of (4.5).

Barndorff-Nielsen and Pederson [1] gave expressions for $H^{\alpha_1\dots\alpha_r}(y)$ for $p = 2$ and $r \leq 6$ without recognising the general form (A.2). For a formal proof of (A.2) see Withers [16].

Theorem A.2. *If $W \sim N_p(0, V^{-1})$ then $H_\lambda(y) = E(iW)^\lambda \exp(iy'W)$, $H_\lambda(0) = E(iW)^\lambda$, $H^{\alpha_1\dots\alpha_r}(y) = E(iW_{\alpha_1}) \dots (iW_{\alpha_r}) \exp(iy'W)$ and $H^{\alpha_1\dots\alpha_r}(0) = i^r EW_{\alpha_1} \dots W_{\alpha_r}$.*

Proof. Simply note that $\phi(0)^{-1}\phi(y) = E \exp(iy'W)$ for $i = \sqrt{-1}$. □

It follows from Theorem A.2 that $EW_{\alpha_1} \dots W_{\alpha_{2r}}$ is equal to the right hand side of (A.2), as given by Corollary 1.1 of Withers [17]. Some of Withers [17] results are duplicated in Holmquist [4].

Theorem A.3. *Suppose $W \sim N_p(0, V^{-1})$. Set $\mu = Vz$. For $n \in R^p$, $\phi(y - \mu)^{-1} \partial_y^n \phi(y - \mu) = E(z - iW)^n \exp(iy'W)$. Specialising to n in N^p , $H_n(\mu) = E(z - iW)^n$.*

Proof. By (2.11) for $y_0 = -\infty \times z$,

$$\begin{aligned} \phi(y - \mu)^{-1} \partial_y^n \phi(y - \mu) &= z^n \sum_\lambda \binom{n}{\lambda} H_\lambda(y) (-z)^{-\lambda} \\ &= z^n \sum_\lambda \binom{n}{\lambda} E(-iW \times z^{-1})^\lambda \exp(iy'W) \\ &= z^n E(1 - iW \times z^{-1})^n \exp(iy'W) \\ &= E(z - iW)^n \exp(iy'W). \end{aligned}$$

The result follows. □

We now show how to convert a sum \sum_{λ}^r to its dual.

Note A.5. Let $\sum_{\alpha}^{p,r}$ denote summation over $\alpha_1, \dots, \alpha_r$ in $\{1, \dots, p\}$. Then

$$\sum_{\alpha}^{p,r} y_{\alpha_1} \dots y_{\alpha_r} = (1'y)^r = r! \sum_{\lambda}^r y^{\lambda} / \lambda!.$$

So, we can write

$$\begin{aligned} & z^n \sum_{\lambda} \binom{n}{\lambda} H_{\lambda}(y) (-z)^{-\lambda} \\ = & z^n \sum_{r=0}^{\infty} \sum_{\lambda}^r (n)_{\lambda} H_{\lambda}(y) (-z)^{-\lambda} / \lambda! \\ = & z^n \sum_{r=0}^{\infty} (r!)^{-1} \sum_{\alpha}^{p,r} (n)_{\lambda_{\alpha}} H^{\alpha_1 \dots \alpha_r}(y) (-z_{\alpha_1})^{-1} \dots (-z_{\alpha_r})^{-1}, \end{aligned}$$

where $\lambda_{\alpha} = e_{\alpha_1} + \dots + e_{\alpha_r}$, and e_j is as in (A.1). More generally,

$$\sum_{\alpha}^{p,r} f(\alpha) y_{\alpha_1} \dots y_{\alpha_r} = r! \sum_{\lambda}^r \bar{f}(\lambda) y^{\lambda} / \lambda!$$

if either $f(\alpha) = \bar{f}(\lambda_{\alpha})$ or $f(\alpha)$ is symmetric and $\bar{f}(\lambda) = f(\alpha_{\lambda})$, where $\alpha_{\lambda} = (1^{\lambda_1}, \dots, p^{\lambda_p})'$ and $j^{\lambda_j} = (j, \dots, j)$ has dimension λ_j . For example, we can write G in (3.3), (3.6) as

$$\begin{aligned} G(y, h) &= h^1 \phi(y) \sum_{r=0}^{\infty} \sum_{\lambda}^r H_{\lambda}(y) h^{\lambda} / \{(\lambda + 1)^1 \lambda!\} \\ &= h^1 \phi(y) \sum_{r=0}^{\infty} (r!)^{-1} \sum_{\alpha}^{p,r} H^{\alpha_1 \dots \alpha_r}(y) h_{\alpha_1} \dots h_{\alpha_r} (\lambda_{\alpha} + 1)^{-1} \end{aligned}$$

and Ruben’s expansion, (4.10), as:

$$\Phi(-\mu) = \phi(\mu) z^{-1} \sum_{r=0}^{\infty} (-)^r (r!)^{-1} \sum_{\alpha}^{p,2r} \lambda_{\alpha}! E W_{\alpha_1} \dots W_{\alpha_{2r}} z_{\alpha_1}^{-1} \dots z_{\alpha_{2r}}^{-1}$$

for $z = V^{-1}\mu > 0$.

REFERENCES

[1] O. Barndorff-Nielsen and B.V. Pederson, The bivariate Hermite polynomials up to order six. *Scand. J. Stat.* **6** (1978) 127–128.
 [2] R.A. Fisher, *Introduction of “Table of Hh functions”, of Airey (1931)*, xxvi–xxxvii, Mathematical Tables, 2nd edition 1946, 3th edition 1951. British Association for the Advancement of Science, London (1931), Vol. 1,
 [3] C.R. Goodall and K.V. Mardia, A geometric derivation of the shape density. *Adv. Appl. Prob.* **23** (1991) 496–514.
 [4] B. Holmquist, Moments and cumulants of the multivariate normal distribution. *Stoch. Anal. Appl.* **6** (1988) 273–278.
 [5] T. Kollo and D. von Rosen, *Advanced Multivariate Statistics with Matrices*. Springer, New York (2005).
 [6] S. Kotz, N. Balakrishnan and N.L. Johnson, *Continuous Multivariate Distributions*. 2nd edition, Wiley, New York (2000) Vol. 1.

- [7] S. Kotz and S. Nadarajah, *Multivariate t Distributions and Their Applications*. Cambridge University Press, Cambridge (2004).
- [8] K.V. Mardia, Fisher's repeated normal integral function and shape distributions. *J. Appl. Stat.* **25** (1998) 231–235.
- [9] D.B. Owen, *Handbook of Statistical Tables*. Addison Wesley, Reading, Massachusetts (1962).
- [10] B. Presnell and P. Rumcheva, The mean resultant length of the spherically projected normal distribution. *Stat. Prob. Lett.* **78** (2008) 557–563.
- [11] H. Ruben, An asymptotic expansion for the multivariate normal distribution and Mills ratio. *J. Res. Nat. Bureau Stand. B* **68** (1964) 3–11.
- [12] R. Savage, Mills ratio for multivariate normal distributions. *Journal of Research of the National Bureau of Standards B* **66** (1962) 93–96.
- [13] G.P. Steck, Lower bounds for the multivariate normal Mills ratio. *Ann. Prob.* **7** (1979) 547–551.
- [14] Y.L. Tong, *The Multivariate Normal Distribution*. Springer Verlag, New York (1990).
- [15] D. von Rosen, Influential observations in multivariate linear models. *Scand. J. Stat.* **22** (1995) 207–222.
- [16] C.S. Withers, A chain rule for differentiation with applications to multivariate Hermite polynomials. *Bull. Aust. Math. Soc.* **30** (1984) 247–250.
- [17] C.S. Withers, The moments of the multivariate normal. *Bull. Aust. Math. Soc.* **32** (1985) 103–108.