BULL. AUSTRAL. MATH. SOC. 62H05
VOL. 32 (1985), 103-107. (60E05)

THE MOMENTS OF THE MULTIVARIATE NORMAL

C.S. WITHERS

Explicit expressions are given for the noncentral moments of the multivariate normal. Finding the general moment is shown to be equivalent to finding the general derivative of the density of the multivariate normal, that is to finding an expression for the multivariate Hermite polynomial.

1. Introduction and summary

Expressions are given for the general moment of a p-dimensional normally distributed random variable $X = (X_1, \ldots, X_p)$ with mean μ and covariance $\Sigma = (\sigma_{ij})$.

Two forms of the moment are considered: $EX_{\alpha_1} \ldots X_{\alpha_r}$, where $\alpha_1, \ldots, \alpha_r$ lie in $\{1, \ldots, p\}$, and $EX_1^{\nu_1} \ldots X_p^{\nu_p}$, where $\{\nu_i\}$ lie in $\{0, 1, 2, \ldots\}$. The $\{\alpha_i\}$ need not all be distinct.

In Section 2 we prove our main result:

THEOREM 1.1. *For* $X \sim N_p(\mu, \Sigma)$, *and* $\alpha_1, \ldots, \alpha_r$ *in* $\{1, 2, \ldots, p\}$,

Received 18 December 1984.

(1.1) $$EX_{\alpha_1} \cdots X_{\alpha_r} = \sum_{l+2k=r} \sum^{m} \mu_{a_1} \cdots \mu_{a_l} \sigma_{b_1 b_2} \cdots \sigma_{b_{2k-1} b_{2k}} ,$$

where $\sum^{m}$ *sums over all* $m = r!/\left(l!2^k k!\right)$ *permutations* $\left(a_1 \ldots a_l b_1 b_2 \ldots b_{2k}\right)$ *of* $\left(\alpha_1 \ldots \alpha_r\right)$ *giving distinct terms allowing for the symmetry of* Σ .

EXAMPLE 1.1. $EX_{\alpha_1} X_{\alpha_2} X_{\alpha_3} = \mu_{\alpha_1} \mu_{\alpha_2} \mu_{\alpha_3} + \sum^{3} \mu_{a_1} \sigma_{b_1 b_2}$, where $\sum^{3} \mu_{a_1} \sigma_{b_1 b_2} = \mu_{\alpha_1} \sigma_{\alpha_2 \alpha_3} + \mu_{\alpha_2} \sigma_{\alpha_1 \alpha_3} + \mu_{\alpha_3} \sigma_{\alpha_1 \alpha_2}$. □

Putting $\mu = 0$ it follows that the even moments are given by

COROLLARY 1.1. *For* $X \sim N_p(0, \Sigma)$ *and* $\alpha_1, \ldots, \alpha_{2k}$ *in* $\{1, 2, \ldots, p\}$,

(1.2) $$EX_{\alpha_1} \cdots X_{\alpha_{2k}} = \sum^{m(k)} \sigma_{b_1 b_2} \cdots \sigma_{b_{2k-1} b_{2k}} ,$$

where $\sum^{m(k)}$ *sums over all* $m(k) = (2k)!/\left(2^k k!\right) = 1.3.5 \ldots (2k-1)$ *permutations* $\left(b_1 \ldots b_{2k}\right)$ *of* $\left(\alpha_1 \ldots \alpha_{2k}\right)$ *giving distinct terms.*

EXAMPLE 1.2. If $\mu = 0$,

$$EX_{\alpha_1} \cdots X_{\alpha_6} = \sum^{15} \sigma_{b_1 b_2} \sigma_{b_3 b_4} \sigma_{b_5 b_6} = \sigma_{\alpha_1 \alpha_2} \sigma_{\alpha_3 \alpha_4} \sigma_{\alpha_5 \alpha_6}$$

plus fourteen like terms. □

REMARK 1.1. This corollary was proved by Isserlis [1] by induction. In particular he noted that $E\left(X_1/\sigma_1\right)^{2k} = m(k)$ where $\sigma_i = \sigma_{ii}^{\frac{1}{2}}$ and, for $\nu_1 + \nu_2 = 2k$, $E\left(X_1/\sigma_1\right)^{\nu_1} \left(X_2/\sigma_2\right)^{\nu_2} = \sum_{k \geq 0} \binom{\nu_2}{i} m(i) m(k-i) r^{\nu_2 - 2i} \left(1-r^2\right)^i$ where $r = \sigma_{12} \sigma_1^{-1} \sigma_2^{-1}$, and gave similar expressions for some of the cross-moments of $\left(X_1, X_2, X_3\right)$. □

Also in Section 2 these moments are expressed in terms of the multivariate Hermite polynomial. We use the definition given in Withers [2]: for A a $p \times p$ matrix set

$$(1.3)\quad He^{\alpha_1 \ldots \alpha_r}(x, A) = \exp(Q/2)\left(-D_{\alpha_1}\right) \ldots \left(-D_{\alpha_r}\right)\exp(-Q/2)\ , \quad x \text{ in } R^p\ ,$$

where $Q = x'Ax$ and $D_i = \partial/\partial x_i$.

(Thus if Σ is positive definite (denoted by $\Sigma > 0$) then

$$He^{\alpha_1 \ldots \alpha_r}\left(x, \Sigma^{-1}\right) = \phi_\Sigma(x)^{-1}\left(-D_{\alpha_1}\right) \ldots \left(-D_{\alpha_r}\right)\phi_\Sigma(x)\ , \quad x \text{ in } R^p\ ,$$

where ϕ_Σ is the density of $X - \mu$.)

We prove

THEOREM 1.2. *For* $X \sim N_p(\mu, \Sigma)$,

$$(1.4)\quad EX_{\alpha_1} \ldots X_{\alpha_r} = He^{\alpha_1 \ldots \alpha_r}\left(-\Sigma^{-1}\mu, -\Sigma\right) = (-)^r He^{\alpha_1 \ldots \alpha_r}\left(\Sigma^{-1}\mu, -\Sigma\right)\ .$$

(If Σ is not positive definite this can be interpreted by choosing $\Sigma_\delta > 0$ tending to Σ as $\delta \to 0$; in particular if $\mu = 0$, we may replace $\Sigma^{-1}\mu$ by 0 .)

This shows that *the problem of finding the moments of the multivariate normal is equivalent to the problem of finding the derivatives of the multivariate normal.* The latter are the building blocks of Edgeworth expansions.

The Hermite polynomials have the 'dual' form

$$(1.5)\qquad He_{\nu_1 \ldots \nu_p}(x, A) = \exp(Q/2)\left(-D_1\right)^{\nu_1} \ldots \left(-D_p\right)^{\nu_p}\exp(-Q/2)$$

for x, Q, D as in (1.3) and $\{\nu_i\}$ nonnegative integers.

Hence Theorem 1.2 can be restated as

COROLLARY 1.2. *For* $X \sim N_p(\mu, \Sigma)$,

$$(1.6)\quad EX_1^{\nu_1} \ldots X_p^{\nu_p} = He_{\nu_1 \ldots \nu_p}\left(-\Sigma^{-1}\mu, -\Sigma\right) = (-)^{\nu \cdot} He_{\nu_1 \ldots \nu_p}\left(\Sigma^{-1}\mu, -\Sigma\right)\ ,$$

https://doi.org/10.1017/S000497270000976X Published online by Cambridge University Press

where $\nu. = \nu_1 + \ldots + \nu_p$.

2. Proofs

LEMMA 2.1 (The derivatives of a function of a quadratic). *For* $f : R^p \to R^d$ *a quadratic function,* $g : R^d \to R^e$, $\pi = (\alpha_1, \ldots, \alpha_r)$ *a set of* r *integers in* $\{1, 2, \ldots, p\}$ *and* x *in* R^p *set* $y = f(x)$ *and* $(f)_\pi = (f)_\pi(x) = \partial^r f(x)/\partial x_{\alpha_1} \ldots \partial x_{\alpha_r}$. *Then* $(g \circ f)(x) = g(f(x))$ *has* rth *order derivatives*

$$(g \circ f)_\pi(x) = \sum_{l+2k=r} \sum_i (g)_{i_1 \ldots i_{l+k}}(y) \sum^m (f_{i_1})_{a_1} \ldots (f_{i_l})_{a_l} (f_{i_{l+1}})_{b_1 b_2} \ldots (f_{i_{l+k}})_{b_{2k-1} b_{2k}} ,$$

where $\sum_i$ *sums over* $\{1 \le i_1 \le d, \ldots, 1 \le i_{l+k} \le d\}$ *and* $\sum^m$ *sums over all* $m = r!/(l!2^k k!)$ *permutations* $(a_1 \ldots a_l b_1 \ldots b_{2k})$ *of* $(\alpha_1 \ldots \alpha_r)$ *giving different terms allowing for* $(f)_{ij} = (f)_{ji}$.

Proof. This corresponds to f a quadratic in the more general rule given in Withers [2]. □

Proof of Theorem 1.1. For t in R^p ,

$$E \exp(t'X) = \exp(f) \text{ , where } f = \mu't + t'\Sigma t/2 .$$

Hence

$$EX_{\alpha_1} \ldots X_{\alpha_r} \exp(t'X) = \partial^r \exp(f)/\partial t_{\alpha_1} \ldots \partial t_{\alpha_r} . \tag{2.1}$$

By Lemma 2.1 the right hand side is given by

$$\exp(f) \sum_{l+2k=r} \sum^m I_{l,2k}(a, b)$$

where $a = (a_1 \ldots a_l)$, $b = (b_1 \ldots b_{2k})$,

$$I_{l,2k}(a, b) = f_{a_1} \ldots f_{a_l} f_{b_1 b_2} \ldots f_{b_{2k-1} b_{2k}} ,$$

https://doi.org/10.1017/S000497270000976X Published online by Cambridge University Press

$f_i = \partial f/\partial t_i = \mu_i + (\Sigma t)_i$, $f_{ij} = \partial^2 f/\partial t_i \partial t_j = \sigma_{ij}$, and $\sum^m$ is as for (1.1). Putting $t = 0$ yields (1.1). □

Proof of Theorem 1.2. $2f = R'\Sigma R - c$ where $R = t + \Sigma^{-1}\mu$, $c = \mu'\Sigma^{-1}\mu$. Hence the right hand side of (2.1) is

$$e^f\left(\partial/\partial R_{\alpha_1} \ldots \partial/\partial R_{\alpha_r}\right)\exp(\tfrac{1}{2}R'\Sigma R) = (-)^r e^f He^{\alpha_1 \ldots \alpha_r}(R, -\Sigma)$$

$$= e^f He^{\alpha_1 \ldots \alpha_r}(-R, -\Sigma) .$$

Now put $t = 0$. □

References

[1] L. Isserlis, "On a formula for the product moment coefficient of any order of a normal frequency distribution in any number of variables", *Biometrika* 12 (1918), 134-139.

[2] C.S. Withers, "A chain rule for differentiation giving simple expressions for the multivariate Hermite polynomials", *Bull. Austral. Math. Soc.* 30 (1983), 247-250.

Applied Mathematics Division,
Department of Scientific and Industrial Research,
PO Box 1335,
Wellington,
New Zealand.