SOME THEOREMS ON QUADRATIC FORMS AND NORMAL VARIABLES

1. The Multivariate Normal Distribution

The $n \times 1$ vector of random variables, y, is said to be distributed as a multivariate normal with mean vector μ and variance covariance matrix Σ (denoted $y \sim N(\mu,\, \Sigma)$) if the density of y is given by

$$f(y;\, \mu,\, \Sigma) = \frac{e^{-\frac{1}{2}(y-\mu)'\Sigma^{-1}(y-\mu)}}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \tag{1}$$

Consider the special case where $n = 1$: $y = y_1$, $\mu = \mu_1$, $\Sigma = \sigma^2$.

$$\begin{aligned} f(y_1;\, \mu_1, \sigma) &= \frac{e^{-\frac{1}{2}(y_1-\mu_1)\left(\frac{1}{\sigma^2}\right)(y_1-\mu_1)}}{(2\pi)^{\frac{1}{2}}(\sigma^2)^{\frac{1}{2}}} \\ &= \frac{e^{\frac{-(y_1-\mu_1)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \end{aligned} \tag{2}$$

is just the normal density for a single random variable.

2. Theorems on Quadratic Forms in Normal Variables

2.1. **Quadratic Form Theorem 1.**

Theorem 1. *If $y \sim N(\mu_y,\, \Sigma_y)$, then*

$$z = Ay \sim N(\mu_z = A\mu_y;\ \Sigma_z = A\Sigma_y A')$$

where A is a matrix of constants.

2.1.1. *Proof.*

$$\begin{aligned} E(z) &= E(Ay) = AE(y) = A\mu_y \\ \mathrm{var}(z) &= E[(z - E(z))\,(z - E(z))'] \\ &= E[(Ay - A\mu_y)(Ay - A\mu_y)'] \\ &= E[A(y - \mu_y)(y - \mu_y)'A'] \\ &= AE(y - \mu_y)(y - \mu_y)'A' \\ &= A\Sigma_y A' \end{aligned} \tag{3}$$

Date: July 12, 2004.

2.1.2. *Example.* Let $Y_1, \ldots, Y_n$ denote a random sample drawn from $N(\mu, \sigma^2)$. Then

$$Y = \begin{pmatrix} Y_1 \\ \cdot \\ \cdot \\ Y_n \end{pmatrix} \sim N \left[\begin{pmatrix} \mu \\ \cdot \\ \cdot \\ \mu \end{pmatrix}, \begin{pmatrix} \sigma^2 & \ldots & 0 \\ \cdot & \sigma^2 & \cdot \\ \cdot & & \\ 0 & & \sigma^2 \end{pmatrix} \right] \tag{4}$$

Now Theorem 1 implies that:

$$\begin{aligned} \bar{Y} &= \frac{1}{n} Y_1 + \cdots + \frac{1}{n} Y_n \\ &= \left(\frac{1}{n}, \ldots, \frac{1}{n} \right) Y = AY \\ &\sim N(\mu, \sigma^2/n) \quad \text{since} \\ &\left(\frac{1}{n}, \ldots, \frac{1}{n} \right) \begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix} = \mu \quad \text{and} \\ &\left(\frac{1}{n}, \ldots, \frac{1}{n} \right) \sigma^2 I \begin{pmatrix} \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{pmatrix} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \end{aligned} \tag{5}$$

2.2. **Quadratic Form Theorem 2.**

Theorem 2. *Let the* $n \times 1$ *vector* $y \sim N(0, I)$. *Then* $y'y \sim \chi^2(n)$.

Proof: Consider that each y_i is an independent standard normal variable. Write out $y'y$ in summation notation as

$$y'y = \Sigma_{i=1}^n y_i^2 \tag{6}$$

which is the sum of squares of n standard normal variables.

2.3. **Quadratic Form Theorem 3.**

Theorem 3. *If* $y \sim N(0, \sigma^2 I)$ *and* M *is a symmetric idempotent matrix of rank* m *then*

$$\frac{y'My}{\sigma^2} \sim \chi^2(\text{tr } M) \tag{7}$$

Proof: Since M is symmetric it can be diagonalized with an orthogonal matrix Q. This means that

$$Q'MQ = \Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} \tag{8}$$

Furthermore, since M is idempotent all these roots are either zero or one. Thus we can choose Q so that Λ will look like

$$Q'MQ = \Lambda = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \tag{9}$$

The dimension of the identity matrix will be equal to the rank of M, since the number of non-zero roots is the rank of the matrix. Since the sum of the roots is equal to the trace, the dimension is also equal to the trace of M. Now let $v = Q'y$. Compute the moments of

$$\begin{aligned} v &= Q'y \\ E(v) &= Q'E(y) = 0 \\ \text{var}(v) &= Q'\sigma^2 IQ \\ &= \sigma^2 Q'Q = \sigma^2 I \quad \text{since } Q \text{ is orthogonal} \\ \Rightarrow v &\sim N(0,\, \sigma^2 I) \end{aligned} \tag{10}$$

Now consider the distribution of $y'My$ using the transformation v. Since Q is orthogonal, its inverse is equal to its transpose. This means that $y = (Q')^{-1}v = Qv$. Now write the quadratic form as follows

$$\begin{aligned} \frac{y'My}{\sigma^2} &= \frac{v'Q'MQv}{\sigma^2} \\ &= \frac{1}{\sigma^2} v' \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} v \\ &= \frac{1}{\sigma^2} \sum_{i=1}^{\text{tr}\, M} v_i^2 \\ &= \sum_{i=1}^{\text{tr}\, M} \left(\frac{v_i}{\sigma} \right)^2 \end{aligned} \tag{11}$$

This is the sum of squares of $(\text{tr}\, M)$ standard normal variables and so is a χ^2 variable with $\text{tr}\, M$ degrees of freedom.

Corollary: If the $n \times 1$ vector $y \sim N(0, I)$ and the $n \times n$ matrix A is idempotent and of rank m. Then

$$y'Ay \sim \chi^2(m)$$

2.4. **Quadratic Form Theorem 4.**

Theorem 4. *If $y \sim N(0,\, \sigma^2 I)$, M is a symmetric idempotent matrix of order n, and L is a $k \times n$ matrix, then Ly and $y'My$ are independently distributed if $LM = 0$.*

Proof: Define the matrix Q as before so that

$$Q'MQ = \Lambda = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \tag{12}$$

Let r denote the dimension of the identity matrix which is equal to the rank of M. Thus $r = \operatorname{tr} M$.

Let $v = Q'y$ and partition v as follows

$$v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_r \\ v_{r+1} \\ \vdots \\ v_n \end{bmatrix} \tag{13}$$

The number of elements of v_1 is r, while v_2 contains $n - r$ elements. Clearly v_1 and v_2 are independent of each other since they are independent standard normals. What we will show now is that $y'My$ depends only on v_1 and Ly depends only on v_2. Given that the v_i are independent, $y'My$ and Ly will be independent. First use Theorem 3 to note that

$$\begin{aligned} y'My &= v'Q'MQv \\ &= v' \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} v \\ &= v_1' v_1 \end{aligned} \tag{14}$$

Now consider the product of L and Q which we denote C. Partition C as $(C_1,\ C_2)$. C_1 has k rows and r columns. C_2 has k rows and $n - r$ columns. Now consider the following product

$$\begin{aligned} C(Q'MQ) &= LQQ'MQ, \text{ since } C = LQ \\ &= LMQ = 0, \text{ since } LM = 0 \quad \text{by assumption} \end{aligned} \tag{15}$$

Now consider the product of C and the matrix $Q'MQ$

$$\begin{aligned} C(Q'MQ) &= (C_1,\, C_2) \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \\ &= 0 \end{aligned} \tag{16}$$

This of course implies that $C_1 = 0$. This then implies that

$$LQ = C = (0,\, C_2) \tag{17}$$

Now consider Ly. It can be written as

$$\begin{aligned} Ly &= LQQ'y, \text{ since } Q \text{ is orthogonal} \\ &= Cv, \text{ by definition of } C \text{ and } v \\ &= C_2 v_2, \text{ since } C_1 = 0 \end{aligned} \tag{18}$$

Now note that Ly depends only on v_2, and $y'My$ depends only on v_1. But since v_1 and v_2 are independent, so are Ly and $y'My$.

2.5. Quadratic Form Theorem 5.

Theorem 5. *Let the $n \times 1$ vector $y \sim N(0, I)$, let A be an $n \times n$ idempotent matrix of rank m, let B be an $n \times n$ idempotent matrix of rank s, and suppose $BA = 0$. Then $y'Ay$ and $y'By$ are independently distributed χ^2 variables.*

Proof: By Theorem 3 both quadratic forms are distributed as chi-square variables. We need only to demonstrate their independence. Define the matrix Q as before so that

$$Q'AQ = \Lambda = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \tag{19}$$

Let $v = Q'y$ and partition v as

$$v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_r \\ v_{r+1} \\ \vdots \\ v_n \end{bmatrix} \tag{20}$$

Now form the quadratic form $y'Ay$ and note that

$$\begin{aligned} y'Ay &= v'Q'AQv \\ &= v' \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} v \\ &= v_1' v_1 \end{aligned} \tag{21}$$

Now define $G = Q'BQ$. Since B is only considered as part of a quadratic form we may consider that it is symmetric, and thus note that G is also symmetric. Now form the product $G\Lambda = Q'BQQ'AQ$. Since Q is orthogonal its transpose is equal to its inverse and we can write $G\Lambda = Q'BAQ = 0$, since $BA = 0$ by assumption. Now write out this identity in partitioned form as

$$\begin{aligned} G(Q'AQ) &= \begin{pmatrix} G_1 & G_2 \\ G_2' & G_3 \end{pmatrix} \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \\ &= \begin{pmatrix} G_1 & 0 \\ G_2' & 0 \end{pmatrix} = \begin{bmatrix} 0_r & 0 \\ 0 & 0 \end{bmatrix} \end{aligned} \tag{22}$$

where G_1 is $r \times r$, G_2 is $r \times (n-r)$ and G_3 is $(n-r) \times (n-r)$.
This means then that $G_1 = 0_r$ and $G_2 = G_2' = 0$.
This means that G is given by

$$G = \begin{pmatrix} 0 & 0 \\ 0 & G_3 \end{pmatrix} \tag{23}$$

Given this information write the quadratic form in B as

$$\begin{aligned} y'By &= y'Q'QBQQ'y \\ &= v'Gv \\ &= (v_1', v_2') \begin{bmatrix} 0 & 0 \\ 0 & G_3 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\ &= v_2'G_3v_2 \end{aligned} \tag{24}$$

It is now obvious that $y'Ay$ can be written in terms of the first r terms of v, while $y'By$ can be written in terms of the last $n-r$ terms of v. Since the v' s are independent the result follows.

2.6. **Quadratic Form Theorem 6 (Craig's Theorem).**

Theorem 6. *If $y \sim N(\mu, \Omega)$ where Ω is positive definite, then $q_1 = y'Ay$ and $q_2 = y'By$ are independently distributed if $A\Omega B = 0$.*

Proof of sufficiency:
This is just a generalization of Theorem 5. Since Ω is a covariance matrix of full rank it is positive definite and can be factored as $\Omega = TT'$. Therefore the condition $A\Omega B = 0$ can be written $ATT'B = 0$. Now pre-multiply this expression by T' and post-multiply by T to obtain that $T'ATT'BT = 0$. Now define $C = T'AT$ and $K = T'BT$ and note that if $A\Omega B = 0$, then

$$CK = (T'AT)(T'BT) = T'\Omega BT = T'0T = 0 \tag{25}$$

Consequently, due to the symmetry of C and K, we also have

$$0 = 0' = (CK)' = K'C' = KC \tag{26}$$

Thus $CK = 0$ and $KC = 0$ and $KC = CK$. A simultaneous diagonalization theorem in matrix algebra [9, Theorem 4.15, p. 155] says that if $CK = KC$ then there exists an orthogonal matrix Q such that

$$\begin{aligned} Q'CQ &= \begin{bmatrix} D_1 & 0 \\ 0 & 0 \end{bmatrix} \\ Q'KQ &= \begin{bmatrix} 0 & 0 \\ 0 & D_2 \end{bmatrix} \end{aligned} \tag{27}$$

where D_1 is an $n_1 \times n_1$ diagonal matrix and D_2 is an $(n - n_1) \times (n - n_1)$ diagonal matrix. Now define $v = Q'T^{-1}y$. It is then distributed as a normal variable with expected value and variance given by

$$\begin{aligned} E(v) &= Q'T^{-1}\mu \\ \text{var}(v) &= Q'T^{-1}\Omega T^{-1\prime}Q \\ &= Q'T^{-1\prime}TT'T^{-1\prime}Q \\ &= I \end{aligned} \tag{28}$$

Thus the vector v is a vector of independent standard normal variables.

Now consider $q_1 = y'Ay$ in terms of v. First note that $y = TQv$ and that $y' = v'Q'T'$. Now write out $y'Ay$ as follows

$$\begin{aligned} q_1 = y'Ay =& v'Q'T'ATQv \\ &= v'Q'T'(T'^{-1}CT^{-1})TQv \\ &= v'Q'CQv \\ &= v_1'D_1v_1 \end{aligned} \tag{29}$$

Similarly we can define $y'By$ in terms of v as

$$\begin{aligned} q_2 = y'By &= v'Q'T'BTQv \\ &= v'Q'T'(T'^{-1}KT^{-1})TQv \\ &= v'Q'KQv \\ &= v_2'D_2v_2 \end{aligned} \tag{30}$$

Thus $q_1 = y'Ay$ is defined in terms of the first n_1 elements of v, and $q_2 = y'By$ is defined in terms of the last $n - n_1$ elements of v and so they are independent.

The proof of necessity is difficult and has a long history [2], [3].

2.7. **Quadratic Form Theorem 7.**

Theorem 7. *If y is a $n \times 1$ random variable and $y \sim N(\mu,\, \Sigma)$ then*

$$(y-\mu)'\Sigma^{-1}(y-\mu) \sim \chi^2(n)$$

Proof: Let $w = (y-\mu)'\Sigma^{-1}(y-\mu)$. If we can show that $w = z'z$ where z is distributed as $N(0\,, I)$ then the proof is complete. Start by diagonalizing Σ with an orthogonal matrix Q. Since Σ is positive definite all the elements of the diagonal matrix Λ will be positive.

$$Q'\Sigma Q = \Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix} \tag{31}$$

Now let Λ^* be the following matrix defined based on Λ.

$$\Lambda^* = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \frac{1}{\sqrt{\lambda_n}} \end{bmatrix} \tag{32}$$

Now let the matrix $H = Q'\Lambda^* Q$. Obviously H is symmetric. Furthermore

$$\begin{aligned} H'H &= Q'\Lambda^* Q Q'\Lambda^* Q \\ &= Q'\Lambda^{-1} Q \\ &= \Sigma^{-1} \end{aligned} \tag{33}$$

The last equality follows from the definition of $\Sigma = Q\Lambda Q'$ after taking the inverse of both sides remembering that the inverse of an orthogonal matrix is equal to its transpose. Furthermore it is obvious that

$$\begin{aligned} H\Sigma H' &= Q\Lambda^* Q'\Sigma Q\Lambda^* Q' \\ &= Q\Lambda^* Q' Q\Lambda Q' Q\Lambda^* Q' \\ &= I \end{aligned} \tag{34}$$

Now let $\varepsilon = y - \mu$ so that $\varepsilon \sim N(0,\, \Sigma)$. Now consider the distribution of $z = H\varepsilon$. It is a standard normal since

$$
\begin{aligned}
E(z) &= HE(\varepsilon) = 0 \\
\mathrm{var}(z) &= H\,\mathrm{var}(\varepsilon)H' \\
&= H\Sigma H' \\
&= I
\end{aligned}
\tag{35}
$$

Now write w as $w = \varepsilon\Sigma^{-1}\varepsilon$ and see that it is equal to $z'z$ as follows

$$
\begin{aligned}
w &= \varepsilon'\Sigma^{-1}\varepsilon \\
&= \varepsilon' H' H\varepsilon \\
&= (H\varepsilon)'(H\varepsilon) \\
&= z'z
\end{aligned}
\tag{36}
$$

2.8. **Quadratic Form Theorem 8.** Let $y \sim N(0,\, I)$. Let M be a non-random idempotent matrix of dimension $n \times n$ (rank $(M) = r \leq n$). Let A be a non-random matrix such that $AM = 0$. Let $t_1 = My$ and let $t_2 = Ay$. Then t_1 and t_2 are independent random vectors.

Proof: Since M is symmetric and idempotent it can be diagonalized using an orthonormal matrix Q as before.

$$
Q'MQ = \Lambda = \begin{bmatrix} I_{r\times r} & 0_{r\times(n-r)} \\ 0_{(n-r)\times r} & 0_{(n-r)\times(n-r)} \end{bmatrix}
\tag{37}
$$

Further note that since Q is orthogonal that $M = Q\Lambda Q'$. Now partition Q as $Q = (Q_1,\, Q_2)$ where Q_1 is $n \times r$. Now use the fact that Q is orthonormal to obtain the following identities

$$
\begin{aligned}
QQ' &= (Q_1 Q_2)\begin{pmatrix} Q_1' \\ Q_2' \end{pmatrix} \\
&= Q_1Q_1' + Q_2Q_2' = I_n \\
Q'Q &= \begin{pmatrix} Q_1' \\ Q_2' \end{pmatrix}(Q_1 Q_2) = \begin{bmatrix} Q_1'Q_1 & Q_1'Q_2 \\ Q_2'Q_1 & Q_2'Q_2 \end{bmatrix} \\
&= \begin{pmatrix} I_r & 0 \\ 0 & I_{n-r} \end{pmatrix}
\end{aligned}
\tag{38}
$$

Now multiply Λ by Q to obtain

$$
\begin{aligned}
Q\Lambda &= (Q_1 Q_2)\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \\
&= (Q_1\, 0)
\end{aligned}
\tag{39}
$$

Now compute M as

$$
\begin{aligned}
M &= Q\Lambda Q' = (Q_1 Q_2)\begin{pmatrix} Q_1' \\ Q_2' \end{pmatrix} \\
&= Q_1 Q_1'
\end{aligned}
\tag{40}
$$

Now let $z_1 = Q_1' y$ and let $z_2 = Q_2' y$. Note that

$$z = (z_1' \,, z_2') = C' y$$

is a standard normal since $E(x) = 0$ and $\text{var}(z) = CC' = I$. Furthermore z_1 and z_2 are independent. Now consider $t_1 = My$. Rewrite this using (40) as

$$Q_1 Q_1' y = Q_1 z_1$$

Thus t_1 depends only on z_1. Now let the matrix

$$N = I - M = Q_2 Q_2'$$

from (38) and (40). Now notice that

$$AN = A(I - M) = A - AM = A$$

since $AM = 0$. Now consider $t_2 = Ay$. Replace A with AN to obtain

$$
\begin{aligned}
t_2 &= Ay = ANy \\
&= A(Q_2 Q_2')y \\
&= AQ_2(Q_2' y) \\
&= AQ_2 z_2
\end{aligned}
\tag{41}
$$

Now $t_1$depends only on z_1 and t_2 depends only on z_2 and since the zs are independent the ts are also independent.

References

[1] Cramer, H. *Mathematical Methods of Statistics.* Princeton: Princeton University Press, 1946

[2] Driscoll, M.F. and W.R. Grundberg. "A History of the Development of Craig's Theorem." *The American Statistician* 40(1986):65-69

[3] Driscoll, M.F. and B. Krasnicka. "An accessible proof of Craig's theorem in the general case." *The American Statistician* 49(1995):59-61

[4] Goldberger, A.S. *Econometric Theory.* New York: Wiley, 1964

[5] Goldberger, A.S. *A Course in Econometrics.* Cambridge: Harvard University Press, 1991

[6] Hocking, R. R. *The Analysis of Linear Models*. Monterey: Brooks/Cole, 1985

[7] Hocking, R. R. *Methods and Applications of Linear Models*. New York: Wiley, 1996

[8] Rao, C.R. *Linear Statistical Inference and its Applications.* 2nd edition. New York: Wiley, 1973

[9] Schott, J. R. *Matrix Analysis for Statistics*. New York: Wiley, 1997