2.11 Other Distributional Results and Quadratic Forms

Definition: A random variable U is said to have a non-central chi square distribution with ν degrees of freedom and non-centrality parameter δ if it can be expressed as $U = \sum_{i=1}^{\nu} Z_i^2$ where $(Z_1, \ldots, Z_\nu)$ are independent Gaussian random variables with unit variance and $\delta = \left[\sum_{i=1}^{\nu} E(Z_i)^2\right]^{\frac{1}{2}}$. We write

$$U \sim \chi^2(\nu, \delta)$$

Note that an ordinary or central chi square distribution is the special case of the non-central distribution with $\delta = 0$. Note also that there are several ways that the non-centrality parameter is defined in the literature. Here, it is defined as the length of the mean vector of the Z_i's. Other authors define a non-centrality parameter, λ, as the square of the length (i.e. $\lambda = \delta^2$) or as one half times the square of the length (i.e. $\lambda = \frac{\delta^2}{2}$).

One can always choose $E(Z_1) = \delta$ and $E(Z_i) = 0$ for $2 \leq i \leq \nu$ so that U may be represented as the sum of a chi square variable with one degree of freedom and non-centrality parameter δ and an independent central chi square with $\nu - 1$ degrees of freedom. From this representation, it is easy to derive the first two moments of U:

$$\begin{aligned} E(U) &= \nu + \delta^2 \\ \text{var}\,(U) &= 2\nu + 4\delta^2 \end{aligned}$$

Also if $U_1 \sim \chi^2(\nu_1, \delta_1)$, and $U_2 \sim \chi^2(\nu_2, \delta_2)$, with U_1, U_2 independent, we have $U_1 + U_2 \sim \chi^2(\nu_1 + \nu_2, (\delta_1^2 + \delta_2^2)^{\frac{1}{2}})$

Definition: A random variable **V** is said to have a non central F-distribution with ν_1 and ν_2 degrees of freedom and non-centrality parameter δ if it may be written as

$$V = \frac{\frac{U_1}{\nu_1}}{\frac{U_2}{\nu_2}}$$

where U_1 and U_2 are independent random variables with $U_1 \sim \chi^2(\nu_1, \delta^2)$ and $U_2 \sim \chi^2(\nu_2, 0)$. We write $V \sim F(\nu_1, \nu_2, \delta)$.

2.11.1 Derivations

Let $Z_1, Z_2, \ldots, Z_p$ be independent $N(0,1)$ then

$$f_{\mathbf{Z}}(\mathbf{z}) = (2\pi)^{-\frac{p}{2}} \exp\left\{-\frac{1}{2}\mathbf{z}^T\mathbf{z}\right\}$$

and if $\mathbf{Y} = \boldsymbol{\mu} + \mathbf{BZ}$ where $\mathbf{BB}^T = \mathbf{V}$ then $\mathbf{Y}$ has a MVN $(\boldsymbol{\mu}, \mathbf{V})$ distribution. The expected value of $\mathbf{Y}$ is $\boldsymbol{\mu}$ and the variance covariance matrix of $\mathbf{Y}$ is $\mathbf{V}$. The density function of $\mathbf{Y}$ is given by

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-\frac{p}{2}}[\det(\mathbf{V})]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^T\mathbf{V}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right\}$$

where we assume that $\mathbf{V}$ is of rank p.

From the definition the m.g.f. of $\mathbf{Z}$ is

$$M_{\mathbf{Z}}(\mathbf{t}) = \exp\left\{\frac{1}{2}\mathbf{t}^T\mathbf{t}\right\}$$

and the m.g.f. of $\mathbf{Y}$ is

$$M_{\mathbf{Y}}(\mathbf{t}) = \exp\left\{\mathbf{t}^T\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}^T\mathbf{V}\mathbf{t}\right\}$$

Again using the definition we have that

$$\int (2\pi)^{-\frac{p}{2}}[\det(\mathbf{V})]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^T\mathbf{V}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right\}\mathbf{dy} = 1$$

and hence

$$\int \exp\left\{-\frac{1}{2}\mathbf{y}^T\mathbf{V}^{-1}\mathbf{y} + \boldsymbol{\mu}^T\mathbf{V}^{-1}y\right\}\mathbf{dy} = (2\pi)^{\frac{p}{2}}[\det(\mathbf{V})]^{\frac{1}{2}} \exp\left\{\frac{1}{2}\boldsymbol{\mu}^T\mathbf{V}^{-1}\boldsymbol{\mu}\right\}$$

We thus have the fundamental identity

$$\int \exp\left\{-\frac{1}{2}\mathbf{y}^T\boldsymbol{\Sigma}^{-1}\mathbf{y} + \boldsymbol{\ell}^T\mathbf{y}\right\}\mathbf{dy} = (2\pi)^{\frac{p}{2}}[\det(\boldsymbol{\Sigma})]^{\frac{1}{2}} \exp\left\{\frac{1}{2}\boldsymbol{\ell}^T\boldsymbol{\Sigma}\boldsymbol{\ell}\right\}$$

which is valid for any positive definite matrix $\boldsymbol{\Sigma}$ and any $\boldsymbol{\ell}$.

Now consider the joint m.g.f. of

$$\mathbf{W} = \mathbf{a} + \mathbf{LY},\ Q_1 = \mathbf{Y}^T\mathbf{B}_1\mathbf{Y} \text{ and } Q_2 = \mathbf{Y}^T\mathbf{B}_2\mathbf{Y}$$

where $\mathbf{L}$ is $r \times p$ of rank r and $\mathbf{B}_1$ and $\mathbf{B}_2$ are symmetric non negative definite matrices of ranks r_1 and r_2 respectively. By definition the joint m.g.f. is given by

$$\int \exp\left\{\mathbf{t}^T(\mathbf{Ly}+\mathbf{a}) + \theta_1\mathbf{y}^T\mathbf{B}_1\mathbf{y} + \theta_2\mathbf{y}^T\mathbf{B}_2\mathbf{y}\right\} f_{\mathbf{Y}}(\mathbf{y})\mathbf{dy}$$

which is equal to

$$\left[(2\pi)^{-\frac{p}{2}}[\det(\mathbf{V})]^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\boldsymbol{\mu}^T\mathbf{V}^{-1}\boldsymbol{\mu})+\mathbf{t}^T\mathbf{a}\right\}\right]$$
$$\times\left[\int\exp\left\{(\mathbf{t}^T\mathbf{L}+\boldsymbol{\mu}^T\mathbf{V}^{-1})\mathbf{y}-\frac{1}{2}\mathbf{y}^T[\mathbf{V}^{-1}-2\theta_1\mathbf{B}_1-2\theta_2\mathbf{B}_2]\mathbf{y}\right\}\mathbf{dy}\right]$$

Using the fundamental identity the joint m.g.f. $M_{\mathbf{W},Q_1,Q_2}(\mathbf{t},\theta_1,\theta_2)$is given by

$$\left[[\det(V)]^{-\frac{1}{2}}[\det(\mathbf{V}^{-1}-2\theta_1\mathbf{B}_1-2\theta_2\mathbf{B}_2)]^{-\frac{1}{2}}\right]$$
$$\times\left[\exp\left\{-\tfrac{1}{2}\boldsymbol{\mu}^T\mathbf{V}^{-1}\boldsymbol{\mu}+\mathbf{t}^T\mathbf{a}\right\}\right]$$
$$\times\left[\exp\left\{\tfrac{1}{2}(\mathbf{t}^T\mathbf{L}+\boldsymbol{\mu}^T\mathbf{V}^{-1})[\mathbf{V}^{-1}-2\theta_1\mathbf{B}_1-2\theta_2\mathbf{B}_2]^{-1}(\mathbf{L}^T\mathbf{t}+\mathbf{V}^{-1}\boldsymbol{\mu})\right\}\right]$$

Now note that

$$\begin{aligned}
\det(\mathbf{V}^{-1}-2\theta_1\mathbf{B}_1-2\theta_2\mathbf{B}_2) &= \det(\mathbf{V}^{-1})\det(\mathbf{I}-2\theta_1\mathbf{B}_1\mathbf{V}-2\theta_2\mathbf{B}_2\mathbf{V}) \\
\left[\mathbf{V}^{-1}-2\theta_1\mathbf{B}_1-2\theta_2\mathbf{B}_2\right]^{-1} &= \mathbf{V}[\mathbf{I}-2\theta_1\mathbf{B}_1\mathbf{V}-2\theta_2\mathbf{B}_2\mathbf{V}]^{-1} \\
[\mathbf{I}-2\theta_1\mathbf{B}_1\mathbf{V}-2\theta_2\mathbf{B}_2\mathbf{V}]^{-1} &= \sum_{r=0}^{\infty}2^r(\theta_1\mathbf{B}_1\mathbf{V}+\theta_2\mathbf{B}_2\mathbf{V})^r
\end{aligned}$$

The last identity is called the Von-Neuman expansion and is valid if the largest characteristic root of the matrix $\theta_1\mathbf{B}_1\mathbf{V} + \theta_2\mathbf{B}_2$ is less than one. Since θ_1 and θ_2 are arbitrary positive numbers this condition can be assumed.

Using these identities we have that

$$(\mathbf{t}^T\mathbf{L} + \boldsymbol{\mu}^T\mathbf{V}^{-1})[\mathbf{V}^{-1} - 2\theta_1\mathbf{B}_1 - 2\theta_2\mathbf{B}_2]^{-1}(\mathbf{L}^T\mathbf{t} + \mathbf{V}^{-1}\boldsymbol{\mu})$$

is given by

$$\begin{gathered}(\mathbf{t}^T\mathbf{L} + \boldsymbol{\mu}^T\mathbf{V}^{-1})\mathbf{V}(\mathbf{L}^T\mathbf{t} + \mathbf{V}^{-1}\boldsymbol{\mu}) + \\ (\mathbf{t}^T\mathbf{L}\mathbf{V} + \boldsymbol{\mu}^T)\left[\sum_{r=1}^{\infty} 2^r(\theta_1\mathbf{B}_1\mathbf{V} + \theta_2\mathbf{B}_2\mathbf{V})^r\right](\mathbf{L}^T\mathbf{t} + \mathbf{V}^{-1}\boldsymbol{\mu})\end{gathered}$$

The joint m.g.f. of $\mathbf{W}$, Q_1 and Q_2 is thus

$$\begin{gathered}[\det(\mathbf{I} - 2\theta_1\mathbf{B}_1\mathbf{V} - 2\theta_2\mathbf{B}_2\mathbf{V})]^{-\frac{1}{2}}\left[\exp\left\{\mathbf{t}^T(\mathbf{a} + \mathbf{L}\boldsymbol{\mu}) + \tfrac{1}{2}\mathbf{t}^T(\mathbf{L}\mathbf{V}\mathbf{L}^T)t\right\}\right] \\ \times \exp\left\{\tfrac{1}{2}(\mathbf{t}^T\mathbf{L}\mathbf{V} + \boldsymbol{\mu}^T)\left[\sum_{r=1}^{\infty} 2^r(\theta_1\mathbf{B}_1\mathbf{V} + \theta_2\mathbf{B}_2\mathbf{V})^r\right](\mathbf{L}^T\mathbf{t} + \mathbf{V}^{-1}\boldsymbol{\mu})\right\}\end{gathered}$$

2.11.2 Results of Importance in Linear Models

(1) If θ_1 and θ_2 are both 0 the m.g.f. of $\mathbf{W}$ is

$$\exp\left\{\mathbf{t}^T(\mathbf{a}+\mathbf{L}\boldsymbol{\mu}) + \frac{1}{2}\mathbf{t}^T(\mathbf{LVL}^T)\mathbf{t}\right\}$$

which shows that $\mathbf{W}$ is MVN $(\mathbf{a}+\mathbf{L}\boldsymbol{\mu}, \mathbf{LVL}^T)$.

(2) If $\mathbf{t} = \mathbf{0}$ then the joint m.g.f. of Q_1 and Q_2 is given by

$$[\det(\mathbf{I} - 2\theta_1\mathbf{B}_1\mathbf{V} - 2\theta_2\mathbf{B}_2\mathbf{V})]^{-\frac{1}{2}}$$
$$\exp\left\{\tfrac{1}{2}\boldsymbol{\mu}^T\left[\textstyle\sum_{r=1}^{\infty} 2^r(\theta_1\mathbf{B}_1\mathbf{V} + \theta_2\mathbf{B}_2\mathbf{V})^r\right]\mathbf{V}^{-1}\boldsymbol{\mu}\right\}$$

(3) If $\mathbf{t} = \mathbf{0}$ and $\theta_2 = 0$ the m.g.f. of Q_1 is

$$[\det(\mathbf{I} - 2\theta_1\mathbf{B}_1\mathbf{V})]^{-\frac{1}{2}} \exp\left\{\frac{1}{2}\boldsymbol{\mu}^T\left[\sum_{r=1}^{\infty} 2^r(\theta_1\mathbf{B}_1\mathbf{V})^r\right]\mathbf{V}^{-1}\mu\right\}$$

Similarly the m.g.f. of Q_2 is given by

$$[\det(\mathbf{I} - 2\theta_2\mathbf{B}_2\mathbf{V})]^{-\frac{1}{2}} \exp\left\{\frac{1}{2}\boldsymbol{\mu}^T\left[\sum_{r=1}^{\infty} 2^r(\theta_2\mathbf{B}_2\mathbf{V})^r\right]\mathbf{V}^{-1}\boldsymbol{\mu}\right\}$$

(4) If $\mathbf{B}_1\mathbf{V}\mathbf{B}_2 = \mathbf{0}$ then

$$\begin{aligned}
[\det(\mathbf{I} - 2\theta_1\mathbf{B}_1\mathbf{V})]\,[\det(\mathbf{I} - 2\theta_2\mathbf{B}_2\mathbf{V})] &= [\det(\mathbf{I} - 2\theta_1\mathbf{B}_1\mathbf{V} - 2\theta_2\mathbf{B}_2\mathbf{V})] \\
(\theta_1\mathbf{B}_1\mathbf{V} + \theta_2\mathbf{B}_2\mathbf{V})^r &= (\theta_1\mathbf{B}_1\mathbf{V})^r + (\theta_2\mathbf{B}_2\mathbf{V})^r
\end{aligned}$$

Thus if $\mathbf{B}_1\mathbf{V}\mathbf{B}_2 = \mathbf{0}$ then Q_1 and Q_2 are independent.

(5) If $\mathbf{LVB} = \mathbf{0}$ then $\mathbf{W}$ and $Q = \mathbf{Y}^T\mathbf{BY}$ are independent.

(6) If $(\mathbf{BV})^2 = \mathbf{BV}$ then

$$\det(\mathbf{I} - 2\theta\mathbf{BV}) = (1 - 2\theta)^q \quad \text{where} \quad q = \text{rank}\,(\mathbf{B})$$

It follows that the m.g.f. of Q is given by

$$\begin{aligned} M_Q(\theta) &= (1-2\theta)^{-\frac{q}{2}} \exp\left\{\frac{1}{2}\boldsymbol{\mu}^T \left[\sum_{r=1}^{\infty}(2\theta)^r\right](\mathbf{BV})\mathbf{V}^{-1}\boldsymbol{\mu}\right\} \\ &= (1-2\theta)^{-\frac{q}{2}} \exp\left\{\frac{1}{2}\boldsymbol{\mu}^T \left[\frac{1}{1-2\theta} - 1\right]\mathbf{B}\boldsymbol{\mu}\right\} \\ &= (1-2\theta)^{-\frac{q}{2}} \exp\left\{\left[\frac{\theta}{1-2\theta}\right]\boldsymbol{\mu}^T\mathbf{B}\boldsymbol{\mu}\right\} \end{aligned}$$

which is the m.g.f. of a non-central chi-square distribution.

(7) If $(\mathbf{BV})^2 = \mathbf{BV}$ and $\mathbf{B}\boldsymbol{\mu} = \mathbf{0}$ then the m.g.f. of $Q = \mathbf{Y}^T\mathbf{BY}$ is

$$(1-2\theta)^{-\frac{q}{2}}$$

so that Q is chi-square with q degrees of freedom.

(8) Results (4), (5), (6) and (7) are necessary and sufficient i.e. if and only if results.