Calculations for the Normal Family

Transformations

Suppose that $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ form a random sample.

(a) By the univariate transformation result

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

(b) By the multivariate transformation result

$$T = \frac{\overline{X} - \mu}{s/\sqrt{n}} \sim \text{Student}(n-1)$$

as T can be written $T = Z/\sqrt{V/\nu}$, where $Z \sim \mathcal{N}(0, 1)$ and $V \sim \chi^2_\nu$ are independent random variables defined by

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \qquad V = \frac{(n-1)s^2}{\sigma^2}$$

Transformations (cont.)

(c) **Fisher-F distribution**: By the multivariate transformation result, if $V_1 \sim \chi^2_{k_1}$ and $V_2 \sim \chi^2_{k_2}$ are independent random variables, then

$$Q = \frac{V_1/k_1}{V_2/k_2}$$

has pdf

$$f_Q(x) = \frac{\Gamma\left(\frac{k_1+k_2}{2}\right)}{\Gamma\left(\frac{k_1}{2}\right)\Gamma\left(\frac{k_2}{2}\right)} \left(\frac{k_1}{k_2}\right)^{k_1/2} x^{k_1/2-1} \left(1+\frac{k_1}{k_2}x\right)^{-(k_1+k_2)/2}$$

and zero otherwise. We say that $Q \sim \text{Fisher}(k_1, k_2)$.

Transformations (cont.)

Now, if

- $X_1, \ldots, X_{n_X} \sim \mathcal{N}(\mu_X, \sigma_X^2)$
- $Y_1, \ldots, Y_{n_Y} \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$

are independent random samples from different distributions, then as

$$\frac{(n_X - 1)s_X^2}{\sigma_X^2} \sim \chi^2_{n_X - 1} \qquad \frac{(n_Y - 1)s_Y^2}{\sigma_Y^2} \sim \chi^2_{n_Y - 1}$$

it follows that

$$\frac{s_X^2/\sigma_X^2}{s_Y^2/\sigma_Y^2} \sim \text{Fisher}(n_X - 1, n_Y - 1)$$

Transformations (cont.)

Note that if $X \sim \text{Fisher}(k_1, k_2)$, then

$$\begin{aligned} \mathbb{E}_X[X] &= \frac{k_2}{k_2 - 2} \quad (\text{if } k_2 > 2) \\ \mathbb{V}\text{ar}_X[X] &= 2\left(\frac{k_2}{k_2 - 2}\right)^2 \frac{(k_1 + k_2 - 2)}{k_1(k_2 - 4)} \quad (\text{if } k_2 > 4) \end{aligned}$$

Also

$$\begin{aligned} X \sim \text{Fisher}(k_1, k_2) &\implies \frac{1}{X} \sim \text{Fisher}(k_2, k_1) \\ X \sim \text{Student}(k_1) &\implies X^2 \sim \text{Fisher}(1, k_1) \\ X \sim \text{Fisher}(k_1, k_2) &\implies \frac{(k_1/k_2)X}{1 + (k_1/k_2)X} \sim \text{Beta}\left(\frac{k_1}{2}, \frac{k_2}{2}\right) \end{aligned}$$

Calculations for the Multivariate Normal

The Multivariate Normal Distribution

Suppose that vector random variable $\mathbf{X} = (X_1, X_2, \ldots, X_k)^\top$ has a multivariate normal distribution with pdf given by

$$f_{\mathbf{X}}(\mathbf{x}) = \left(\frac{1}{2\pi}\right)^{k/2} \frac{1}{|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{x}^\top \Sigma^{-1}\mathbf{x}\right\} \tag{1}$$

where Σ is the $k \times k$ variance-covariance matrix (we can consider here the case where the expected value $\underset{\sim}{\mu}$ is the $k \times 1$ zero vector; results for the general case are easily available by transformation).

Consider partitioning $\mathbf{X}$ into two components $\mathbf{X}_1$ and $\mathbf{X}_2$ of dimensions k_1 and $k_2 = k - k_1$ respectively, that is,

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}.$$

We attempt to deduce

The Multivariate Normal Distribution (cont.)

(a) the **marginal** distribution of $\mathbf{X}_1$, and

(b) the **conditional** distribution of $\mathbf{X}_2$ **given** that $\mathbf{X}_1 = \mathbf{x}_1$.

First, write

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

where Σ_{11} is $k_1 \times k_1$, Σ_{22} is $k_2 \times k_2$, $\Sigma_{21} = \Sigma_{12}^\top$, and

$$\Sigma^{-1} = \mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix}$$

so that $\Sigma\mathbf{V} = \mathbf{I}_k$ ($\mathbf{I}_r$ is the $r \times r$ identity matrix) gives

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{k_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{k_2} \end{bmatrix}$$

The Multivariate Normal Distribution (cont.)

where $\mathbf{0}$ represents the zero matrix of appropriate dimension. More specifically,

$$\begin{aligned}
\Sigma_{11}\mathbf{V}_{11} + \Sigma_{12}\mathbf{V}_{21} &= \mathbf{I}_{k_1} && (2)\\
\Sigma_{11}\mathbf{V}_{12} + \Sigma_{12}\mathbf{V}_{22} &= \mathbf{0} && (3)\\
\Sigma_{21}\mathbf{V}_{11} + \Sigma_{22}\mathbf{V}_{21} &= \mathbf{0} && (4)\\
\Sigma_{21}\mathbf{V}_{12} + \Sigma_{22}\mathbf{V}_{22} &= \mathbf{I}_{k_2}. && (5)
\end{aligned}$$

From the multivariate normal pdf in equation (1), we can re-express the term in the exponent as

$$\mathbf{x}^\top \Sigma^{-1}\mathbf{x} = \mathbf{x}_1^\top \mathbf{V}_{11}\mathbf{x}_1 + \mathbf{x}_1^\top \mathbf{V}_{12}\mathbf{x}_2 + \mathbf{x}_2^\top \mathbf{V}_{21}\mathbf{x}_1 + \mathbf{x}_2^\top \mathbf{V}_{22}\mathbf{x}_2. \quad (6)$$

In order to compute the marginal and conditional distributions, we must complete the square in $\mathbf{x}_2$ in this expression. We can write

$$\mathbf{x}^\top \Sigma^{-1}\mathbf{x} = (\mathbf{x}_2 - \mathbf{m})^\top \mathbf{M}(\mathbf{x}_2 - \mathbf{m}) + \mathbf{c} \quad (7)$$

The Multivariate Normal Distribution (cont.)

and by comparing with equation (6) we can deduce that, for quadratic terms in $\mathbf{x}_2$,

$$\mathbf{x}_2^\top \mathbf{V}_{22}\mathbf{x}_2 = \mathbf{x}_2^\top \mathbf{M}\mathbf{x}_2 \qquad \therefore \qquad \mathbf{M} = \mathbf{V}_{22} \tag{8}$$

for linear terms

$$\mathbf{x}_2^\top \mathbf{V}_{21}\mathbf{x}_1 = -\mathbf{x}_2^\top \mathbf{M}\mathbf{m} \qquad \therefore \qquad \mathbf{m} = -\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_1 \tag{9}$$

and for constant terms

$$\mathbf{x}_1^\top \mathbf{V}_{11}\mathbf{x}_1 = \mathbf{c} + \mathbf{m}^\top \mathbf{M}\mathbf{m} \qquad \therefore \qquad \mathbf{c} = \mathbf{x}_1^\top (\mathbf{V}_{11} - \mathbf{V}_{21}^\top \mathbf{V}_{22}^{-1}\mathbf{V}_{21})\mathbf{x}_1 \tag{10}$$

thus yielding all the terms required for equation (7), that is

$$\mathbf{x}^\top \Sigma^{-1}\mathbf{x} = (\mathbf{x}_2+\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_1)^\top \mathbf{V}_{22}(\mathbf{x}_2+\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_1)+\mathbf{x}_1^\top (\mathbf{V}_{11}-\mathbf{V}_{21}^\top \mathbf{V}_{22}^{-1}\mathbf{V}_{21})\mathbf{x}_1 \tag{11}$$

The Multivariate Normal Distribution (cont.)

which, crucially, is a sum of two terms, where the first can be interpreted as a function of $\mathbf{x}_2$, given $\mathbf{x}_1$, and the second is a function of $\mathbf{x}_1$ only.

Hence we have an immediate factorization of the full joint pdf using the chain rule for random variables;

$$f_{\mathbf{X}}(\mathbf{x}) = f_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2|\mathbf{x}_1) f_{\mathbf{X}_1}(\mathbf{x}_1) \tag{12}$$

where

$$f_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2|\mathbf{x}_1) \propto \exp\left\{-\frac{1}{2}(\mathbf{x}_2 + \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_1)^\top \mathbf{V}_{22}(\mathbf{x}_2 + \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_1)\right\} \tag{13}$$

giving that

$$\mathbf{X}_2|\mathbf{X}_1 = \mathbf{x}_1 \sim \mathcal{N}_{k_2}\left(-\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{x}_1, \mathbf{V}_{22}^{-1}\right) \tag{14}$$

The Multivariate Normal Distribution (cont.)

and

$$f_{\mathbf{x}_1}(\mathbf{x}_1) \propto \exp\left\{-\frac{1}{2}\mathbf{x}_1^\top(\mathbf{V}_{11} - \mathbf{V}_{21}^\top\mathbf{V}_{22}^{-1}\mathbf{V}_{21})\mathbf{x}_1\right\} \tag{15}$$

giving that

$$\mathbf{X}_1 \sim \mathcal{N}_{k_1}\left(0, (\mathbf{V}_{11} - \mathbf{V}_{21}^\top\mathbf{V}_{22}^{-1}\mathbf{V}_{21})^{-1}\right). \tag{16}$$

But, from equation (3), $\Sigma_{12} = -\Sigma_{11}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}$, and then from equation (2), substituting in Σ_{12},

$$\Sigma_{11}\mathbf{V}_{11} - \Sigma_{11}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21} = \mathbf{I}_{k_1}$$

so that

$$\Sigma_{11} = (\mathbf{V}_{11} - \mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21})^{-1} = (\mathbf{V}_{11} - \mathbf{V}_{21}^\top\mathbf{V}_{22}^{-1}\mathbf{V}_{21})^{-1}.$$

The Multivariate Normal Distribution (cont.)

Hence, by inspection of equation (16), we conclude that

$$\boxed{\mathbf{X}_1 \sim \mathcal{N}_{k_1}(0, \Sigma_{11}),} \tag{17}$$

that is, we can extract the Σ_{11} block of Σ to define the marginal sigma matrix of $\mathbf{X}_1$.

Using similar arguments, we can define the conditional distribution from equation (14) more precisely. First, from equation (3), $\mathbf{V}_{12} = -\Sigma_{11}^{-1}\Sigma_{12}\mathbf{V}_{22}$, and then from equation (5), substituting in $\mathbf{V}_{12}$

$$-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\mathbf{V}_{22} + \Sigma_{22}\mathbf{V}_{22} = \mathbf{I}_{k-k_1}$$

so that

$$\mathbf{V}_{22}^{-1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} = \Sigma_{22} - \Sigma_{12}^{\top}\Sigma_{11}^{-1}\Sigma_{12}.$$

The Multivariate Normal Distribution (cont.)

Finally, from equation (3), taking transposes on both sides, we have that $\mathbf{V}_{21}\Sigma_{11} + \mathbf{V}_{22}\Sigma_{21} = \mathbf{0}$. Then pre-multiplying by $\mathbf{V}_{22}^{-1}$, and post-multiplying by Σ_{11}^{-1}, we have

$$\mathbf{V}_{22}^{-1}\mathbf{V}_{21} + \Sigma_{21}\Sigma_{11}^{-1} = \mathbf{0} \qquad \therefore \qquad \mathbf{V}_{22}^{-1}\mathbf{V}_{21} = -\Sigma_{21}\Sigma_{11}^{-1},$$

so we have, substituting into equation (14), that

$$\boxed{\mathbf{X}_2|\mathbf{X}_1 = \mathbf{x}_1 \sim \mathcal{N}_{k_2}\left(\Sigma_{21}\Sigma_{11}^{-1}\mathbf{x}_1, \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right).} \qquad (18)$$

Thus any marginal, and any conditional distribution of a multivariate normal joint distribution is also multivariate normal, as the choices of $\mathbf{X}_1$ and $\mathbf{X}_2$ are arbitrary.

The Sample Mean and Variance Statistics

THEOREM

Suppose that $X_1, \ldots, X_n$ is a random sample from a distribution, with finite expectation μ and variance σ^2. Consider the sample mean and sample variance statistics $\overline{X}$ and s^2 and denote

$$T_1 = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad T_2 = s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$

Then

(a) $\mathbb{E}_{T_1}[T_1] = \mu$

(b) $\mathbb{V}\mathrm{ar}_{T_1}[T_1] = \dfrac{\sigma^2}{n}$

(c) $\mathbb{E}_{T_2}[T_2] = \sigma^2$

The Sample Mean and Variance Statistics (cont.)

Proof: (a) and (b) follow from elementary properties of expectations and variances for independent random variables. For (c), note that

$$\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 = \sum_{i=1}^{n} X_i^2 - n\overline{X}^2.$$

The Sample Mean and Variance Statistics (cont.)

Hence

$$\begin{aligned}
\mathbb{E}_{T_2}[T_2] &= \frac{1}{n-1}\mathbb{E}_{f_X}\left[\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right] \\
&= \frac{1}{n-1}\left[\sum_{i=1}^{n} \mathbb{E}_{X_i}[X_i^2] - n\mathbb{E}_X\left[\overline{X}^2\right]\right] \\
&= \frac{1}{n-1}\left[n(\sigma^2+\mu^2) - n\left(\frac{\sigma^2}{n}+\mu^2\right)\right] \qquad (19) \\
&= \sigma^2
\end{aligned}$$

where line (19) follows from the fact that for any random variable X

$$\sigma^2 = \mathbb{E}_X[X^2] - \mathbb{E}_X[X]^2 = \mathbb{E}_X[X^2] - \mu^2$$

and the result of parts (a) and (b).

Sampling from a Normal Family

Sampling the Normal Family

Recall the fundamental transformation results for Normal random variables:

(i) If $X \sim \mathcal{N}(0,1)$, then

$$X^2 \sim \chi^2_1 \equiv \text{Gamma}\left(\frac{1}{2}, \frac{1}{2}\right)$$

(ii) If $X_1, \ldots, X_r \sim \mathcal{N}(0,1)$ are independent random variables, then

$$Y = \sum_{i=1}^{r} X_i^2 \sim \chi^2_r \equiv \text{Gamma}\left(\frac{r}{2}, \frac{1}{2}\right)$$

(iii) If $Y_1 \sim \chi^2_{r_1}$ and $Y_2 \sim \chi^2_{r_2}$ are independent random variables, then

$$Y = Y_1 + Y_2 \sim \chi^2_{r_1+r_2}$$

Sampling the Normal Family (cont.)

THEOREM

Suppose that $X_1, \ldots, X_n$ is a random sample from a normal distribution, say $X_i \sim \mathcal{N}(\mu, \sigma^2)$. Define the sample mean and sample variance statistics $\overline{X}$ and s^2 as the random variables

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad\qquad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$

Then

(a) $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$

(b) $\overline{X}$ is independent of $\left\{X_i - \overline{X}, i = 1, \ldots, n\right\}$, and $\overline{X}$ and s^2 are independent random variables

(c) The random variable

$$\frac{(n-1)s^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

has a **chi-squared distribution** with $n-1$ degrees of freedom.

Sampling the Normal Family (cont.)

(a) Proof straightforward using mgfs.

(b) The joint pdf $X_1, \ldots, X_n$ is the normal density

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right\}$$

Consider the multivariate transformation to $Y_1, \ldots, Y_n$ where

$$\left.\begin{array}{ll} Y_1 & = \overline{X} \\ Y_i & = X_i - \overline{X},\ i = 2, \ldots, n \end{array}\right\} \Longleftrightarrow \left\{\begin{array}{ll} X_1 & = Y_1 - \sum_{i=2}^{n} Y_i \\ & \\ X_i & = Y_i + Y_1,\quad i = 2, \ldots, n \end{array}\right.$$

Sampling the Normal Family (cont.)

Thus $\mathbf{Y} = A\mathbf{X}$, or equivalently $\mathbf{X} = A^{-1}\mathbf{Y}$, where A is the $n \times n$ matrix with (i,j)th element

$$[A]_{ij} = \begin{cases} \dfrac{1}{n} & i = 1, j = 1, 2, \ldots, n \\ 1 - \dfrac{1}{n} & i = j = 2, 3, \ldots, n \\ -\dfrac{1}{n} & \text{otherwise} \end{cases}$$

that is, we have a linear transformation.

Sampling the Normal Family (cont.)

Note that

$$\begin{aligned}\sum_{i=1}^{n}(x_i-\mu)^2 &= \sum_{i=1}^{n}(x_i-\bar{x}+\bar{x}-\mu)^2 \\ &= \sum_{i=1}^{n}\left[(x_i-\bar{x})^2+2(x_i-\bar{x})(\bar{x}-\mu)+(\bar{x}-\mu)^2\right] \\ &= \sum_{i=1}^{n}(x_i-\bar{x})^2+n(\bar{x}-\mu)^2\end{aligned}$$

where $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ is the observed sample mean.

Sampling the Normal Family (cont.)

Thus the joint pdf of $X_1, \ldots, X_n$ takes the form

$$f_{X_1,..,X_n}(x_1, .., x_n) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - \mu)^2\right]\right\}$$

Now

$$x_1 - \bar{x} = -\sum_{i=2}^{n}(x_i - \bar{x}) = -\sum_{i=2}^{n} y_i$$

and so

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = (x_1 - \bar{x})^2 + \sum_{i=2}^{n}(x_i - \bar{x})^2 = \left(-\sum_{i=2}^{n} y_i\right)^2 + \sum_{i=2}^{n} y_i^2$$

Sampling the Normal Family (cont.)

The Jacobian of the transformation is n, so the joint density of $Y_1, \ldots, Y_n$ is given by the multivariate transformation theorem as

$$
\begin{aligned}
f_{Y_1,..,Y_n}(y_1,..,y_n) &= n\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left\{-\frac{1}{2\sigma^2}\left[\left(-\sum_{i=2}^{n}y_i\right)^2+\sum_{i=2}^{n}y_i^2+n\left(y_1-\mu\right)^2\right]\right\} \\
&= n\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left\{-\frac{1}{2\sigma^2}\left[\left(-\sum_{i=2}^{n}y_i\right)^2+\sum_{i=2}^{n}y_i^2\right]\right\} \\
&\quad \times \exp\left\{-\frac{n}{2\sigma^2}\left(y_1-\mu\right)^2\right\} \\
&= f_{Y_2,..,Y_n}(y_2,..,y_n)f_{Y_1}(y_1)
\end{aligned}
$$

and therefore Y_1 is independent of $Y_2, \ldots, Y_n$. Hence $\overline{X}$ is **independent** of the random variables $\left\{Y_i = X_i - \overline{X}, i = 2, \ldots, n\right\}$.

Sampling the Normal Family (cont.)

Finally, $\overline{X}$ is also independent of $X_1 - \overline{X}$ as

$$X_1 - \overline{X} = -\sum_{i=2}^{n} (X_i - \overline{X})$$

and of s^2, which is a function only of $\{X_i - \overline{X},\ i = 1, \ldots, n\}$. As $\overline{X}$ is independent of these variables, $\overline{X}$ and s^2 are also independent.

(c) The random variables that appear as sums of squares terms in the joint pdf are

$$\frac{\sum_{i=1}^{n} (X_i - \mu)^2}{\sigma^2} = \frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}{\sigma^2} + \frac{n\left(\overline{X} - \mu\right)^2}{\sigma^2}$$

Sampling the Normal Family (cont.)

or $V_1 = V_2 + V_3$, say. Now, $X_i \sim \mathcal{N}(\mu, \sigma^2)$, so therefore

$$\frac{(X_i - \mu)^2}{\sigma^2} \sim \mathcal{N}(0, 1) \quad \Longrightarrow \quad \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi^2_1 \equiv \text{Gamma}\left(\frac{1}{2}, \frac{1}{2}\right)$$

so that

$$V_1 = \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{\sigma^2} \sim \chi^2_n$$

as the X_is are independent, and the sum of n independent Gamma$(1/2, 1/2)$ variables has a Gamma$(n/2, 1/2)$ distribution.

Sampling the Normal Family (cont.)

Similarly, as $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$, $V_3 \sim \chi^2_1$ By part (b), V_2 and V_3 are independent, and so the mgfs of V_1, V_2 and V_3 are related by

$$M_{V_1}(t) = M_{V_2}(t) M_{V_3}(t) \Longrightarrow M_{V_2}(t) = \frac{M_{V_1}(t)}{M_{V_3}(t)}$$

As V_1 and V_3 are Gamma random variables, M_{V_1} and M_{V_3} are given by

$$M_{V_1}(t) = \left(\frac{1/2}{1/2 - t}\right)^{n/2} \quad \text{and} \quad M_{V_3}(t) = \left(\frac{1/2}{1/2 - t}\right)^{1/2}.$$

Sampling the Normal Family (cont.)

So therefore

$$M_{V_2}(t) = \left(\frac{1/2}{1/2 - t}\right)^{(n-1)/2}$$

which is also the mgf of a Gamma random variable, and hence

$$V_2 = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

and the result follows.

Sampling the Normal Family (cont.)

Alternative inductive proof of (c): Let $\overline{X}_k$ and s_k^2, $k = 1, 2, \ldots, n$ denote the sample mean and sample variance random variables derived from the first k variables. Now, for $k \geq 2$, it can be shown after some manipulation that

$$(k-1)s_k^2 = (k-2)s_{k-1}^2 + \left(\frac{k-1}{k}\right)(X_k - \overline{X}_{k-1})^2 \qquad (20)$$

For $k = 2$

$$(2-1)s_2^2 = \frac{1}{2}(X_2 - X_1)^2 = \left(\frac{X_2 - X_1}{\sqrt{2}}\right)^2 = Z^2$$

say, where $Z \sim \mathcal{N}(0, 1)$. Thus $s_2^2 \sim \chi_1^2$. Now for the inductive hypothesis, presume that

$$(k-1)s_k^2 \sim \chi_{k-1}^2$$

Sampling the Normal Family (cont.)

so that, using the identity in (20),

$$ks_{k+1}^2 = (k-1)s_k^2 + \left(\frac{k}{k+1}\right)(X_{k+1} - \overline{X}_k)^2$$

The two terms on the right hand side are independent (using the result in (b)); the first term is χ^2_{k-1} distributed, the second term is χ^2_1 distributed, so ks^2_{k+1} is χ^2_k distributed and the inductive argument is completed.