

1 Scalar Normal Random Variables

Let $X \sim \text{Normal}(0, \sigma^2)$. Then $X^2/\sigma^2 \sim \chi_1^2$, a “chi-square” random variable with 1 degree of freedom. You might recall this from a probability class: $Z = X/\sigma \sim \text{Normal}(0, 1)$, a standard Normal random variable, and when you square it, it becomes a chi-square random variable. If you’ve taken my 3120 class, this might look familiar.

2 Multivariate Normal Random Vectors

Now let’s think of \mathbf{X} as an n dimensional vector. Assume that $\mathbf{X} \sim \text{Normal}(\mathbf{0}, \mathbf{\Sigma})$. How do we standardize this into a \mathbf{Z} standard Normal? Well we can do something like this

$$\mathbf{\Sigma}^{-1/2}\mathbf{X} \sim \text{Normal}(\mathbf{0}, \mathbf{I}_n).$$

If $\mathbf{\Sigma}$ is analogous to σ^2 in the univariate case, this is kind of like dividing by σ . It turns out that for legitimate covariance matrices $\mathbf{\Sigma}$, there always exists a square root matrix $\mathbf{\Sigma}^{1/2}$ such that $\mathbf{\Sigma} = \mathbf{\Sigma}^{1/2}\mathbf{\Sigma}^{1/2'}$, and such that $\mathbf{\Sigma}^{1/2}$ invertible. We call $[\mathbf{\Sigma}^{1/2}]^{-1} = \mathbf{\Sigma}^{-1/2}$. We won’t discuss the details, but these are guaranteed as long as this covariance matrix is both symmetric and positive definite.

This might be more convincing if you consider the special case where we’re looking at diagonal covariance matrices:

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}.$$

You might guess that

$$\mathbf{\Sigma}^{1/2} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{bmatrix}$$

and that

$$\mathbf{\Sigma}^{-1/2} = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_n} \end{bmatrix}$$

and that

$$\mathbf{\Sigma}^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_n^2} \end{bmatrix}.$$

And you would be correct.

It's also important to assume that all the standard deviations are strictly positive in the last part, otherwise we would be dividing by zero. This is related to the more general positive definite property. Also, this is just for demonstration, because these results are still true for more general covariance matrices.

3 “Squaring” Multivariate Normal Vectors

So $\Sigma^{-1/2}\mathbf{X}$ is a vector, and it's a standard Normal. What happens if we “square” it? Recall that we can square vectors \mathbf{y} by taking their norm, or the inner product of the vector with itself:

$$\mathbf{y}'\mathbf{y} = \sum_i y_i^2$$

So let's do that with our standard Normal random vector:

$$[\Sigma^{-1/2}\mathbf{X}]'[\Sigma^{-1/2}\mathbf{X}] = \mathbf{X}'\Sigma^{-1/2'}\Sigma^{-1/2}\mathbf{X} = \mathbf{X}'\Sigma^{-1}\mathbf{X}.$$

If we started off by assuming that $\mathbf{X} \sim \text{Normal}(\boldsymbol{\mu}, \Sigma)$ (not mean 0), then we could end up with the multivariate “z-score” I was mentioning earlier in class. It turns out that this quadratic form is a chi-square random variable as well.

$$[\mathbf{X} - \boldsymbol{\mu}]'\Sigma^{-1}[\mathbf{X} - \boldsymbol{\mu}] \sim \chi_n^2.$$

The number of degrees of freedom is the same as how long the vector is.

4 A Special Case

Consider the case where all the elements of your \mathbf{X} vector are independent. Then Σ is diagonal again, the same as in the earlier example. And we have

$$\begin{aligned} [\mathbf{X} - \boldsymbol{\mu}]'\Sigma^{-1}[\mathbf{X} - \boldsymbol{\mu}] &= \sum_i \sum_j (X_i - \mu_i)(X_j - \mu_j)\Sigma_{i,j}^{-1} \\ &= \sum_i (X_i - \mu_i)(X_i - \mu_i)\Sigma_{i,i}^{-1} \\ &\quad \text{(only have to worry about diagonals of the matrix)} \\ &= \sum_i (X_i - \mu_i)(X_i - \mu_i)\frac{1}{\sigma_i^2} \\ &= \sum_i \frac{(X_i - \mu_i)^2}{\sigma_i^2} \end{aligned}$$

which might look more familiar to you. It's the sum of n independent, squared standard Normal random variables.

5 What about that result from class?

In class we had a result (that we didn't prove, but it comes from the Central Limit Theorem) that stated

$$\sqrt{n}(\hat{\phi} - \phi) \stackrel{\text{approx.}}{\sim} \text{Normal}(\mathbf{0}, \sigma^2 \mathbf{\Gamma}^{-1}).$$

This means, if we take the “z-score”,

$$\begin{aligned} [\sqrt{n}(\hat{\phi} - \phi)]' [\sigma^2 \mathbf{\Gamma}^{-1}]^{-1} [\sqrt{n}(\hat{\phi} - \phi)] &= \frac{n}{\sigma^2} (\hat{\phi} - \phi)' \mathbf{\Gamma} (\hat{\phi} - \phi) \\ &\stackrel{\text{approx.}}{\sim} \chi_p^2. \end{aligned}$$

Because the size of this vector is p . Also, you might recall that the expectation of a chi-square random variable is the same as its degrees of freedom. Then, using linearity of expectation,

$$\frac{n}{\sigma^2} E \left[(\hat{\phi} - \phi)' \mathbf{\Gamma} (\hat{\phi} - \phi) \right] = p$$

which means

$$E \left[(\hat{\phi} - \phi)' \mathbf{\Gamma} (\hat{\phi} - \phi) \right] = \sigma^2 \frac{p}{n}.$$