Multivariate Delta Method
(Greene, p.914, appendix D.)

If some vector of estimates, $\hat{\theta}_1, \ldots, \hat{\theta}_k$, has the following asymptotic distribution

$$\sqrt{n}\left(\begin{bmatrix} \hat{\theta}_1 \\ : \\ \hat{\theta}_k \end{bmatrix} - \begin{bmatrix} \theta_1 \\ : \\ \theta_k \end{bmatrix}\right) \overset{d}{\to} N\left(\begin{bmatrix} 0 \\ : \\ 0 \end{bmatrix}, \Sigma\right)$$

where $\theta_1, \ldots, \theta_k$ is the true parameter and Σ is the kxk asymptotic covariance matrix,

then some J continuos functions of the parameters, $g_1(\theta_1, \ldots, \theta_k), \ldots, g_J(\theta_1, \ldots, \theta_k)$, have the following asymptotic distribution

$$\sqrt{n}\left(\begin{bmatrix} g_1(\hat{\theta}_1, \ldots, \hat{\theta}_k) \\ : \\ g_J(\hat{\theta}_1, \ldots, \hat{\theta}_k) \end{bmatrix} - \begin{bmatrix} g_1(\theta_1, \ldots, \theta_k) \\ : \\ g_J(\theta_1, \ldots, \theta_k) \end{bmatrix}\right) \overset{d}{\to} N\left(\begin{bmatrix} 0 \\ : \\ 0 \end{bmatrix}, \Omega\right),$$

where Ω is JxJ asymptotic covariance matrix.
$\Omega = \Psi\Sigma\Psi'$ and Ψ is the Jxk matrix of derivatives.

$$\Psi = \begin{bmatrix} \frac{\partial g_1(\hat{\theta}_1, \ldots, \hat{\theta}_k)}{\partial \theta_1} & \ldots & \frac{\partial g_1(\hat{\theta}_1, \ldots, \hat{\theta}_k)}{\partial \theta_k} \\ : & & : \\ \frac{\partial g_J(\hat{\theta}_1, \ldots, \hat{\theta}_k)}{\partial \theta_1} & \ldots & \frac{\partial g_J(\hat{\theta}_1, \ldots, \hat{\theta}_k)}{\partial \theta_k} \end{bmatrix}$$

An example using the normal distribution: $z \sim N(\mu, \sigma^2)$

In section, we derived the MLE estimator of μ and σ^2 for a random sample of n observations.
$\hat{\mu}_{ml} = 1/n \sum x_i$ and $\hat{\sigma}^2_{ml} = 1/n \sum (x_i - \hat{\mu}_{ml})^2$

We also derived the joint asymptotic distribution:

$$\sqrt{n}\left(\begin{bmatrix} \widehat{\mu}_{ml} \\ \widehat{\sigma}^2_{ml} \end{bmatrix} - \begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}\right) \xrightarrow{d} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & 0 \\ 0 & 2(\sigma^2)^2 \end{bmatrix}\right)$$

Let's assume we're actually interested in some other parameters, η_1 and η_2. We know that these parameters are a function of the normal distribution parameters.

$$\eta_1 = g_1(\mu, \sigma^2) = 3\mu$$

$$\eta_2 = g_2(\mu, \sigma^2) = 5\left(\frac{\mu}{\sigma^2}\right)^2$$

By the Invariance Property of MLE, the MLE for η_1 and η_2 are

$$\widehat{\eta}_{1,ml} = g_1(\widehat{\mu}_{ml}, \widehat{\sigma}^2_{ml}) = 3\widehat{\mu}_{ml} = 3/n \sum x_i$$

$$\widehat{\eta}_{2,ml} = g_2(\widehat{\mu}_{ml}, \widehat{\sigma}^2_{ml}) = 5\left(\frac{\widehat{\mu}_{ml}}{\widehat{\sigma}^2_{ml}}\right)^2 = 5\left(\frac{\sum x_i}{\sum (x_i - \widehat{\mu}_{ml})^2}\right)^2$$

By the Delta Method, the joint asymptotic distribution of these two estimators is

$$\sqrt{n}\left(\begin{bmatrix} \widehat{\eta}_{1,ml} \\ \widehat{\eta}_{2,ml} \end{bmatrix} - \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}\right) \xrightarrow{d} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Omega\right)$$

where in this case Ω is 2x2 (JxJ)

$$\Omega = \begin{bmatrix} \frac{\partial g_1(\mu,\sigma^2)}{\partial \mu} & \frac{\partial g_1(\mu,\sigma^2)}{\partial \sigma^2} \\ \frac{\partial g_2(\mu,\sigma^2)}{\partial \mu} & \frac{\partial g_2(\mu,\sigma^2)}{\partial \sigma^2} \end{bmatrix} \begin{bmatrix} \sigma^2 & 0 \\ 0 & 2(\sigma^2)^2 \end{bmatrix} \begin{bmatrix} \frac{\partial g_1(\mu,\sigma^2)}{\partial \mu} & \frac{\partial g_2(\mu,\sigma^2)}{\partial \mu} \\ \frac{\partial g_1(\mu,\sigma^2)}{\partial \sigma^2} & \frac{\partial g_2(\mu,\sigma^2)}{\partial \sigma^2} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial g_1(\mu,\sigma^2)}{\partial \mu}\sigma^2 & \frac{\partial g_1(\mu,\sigma^2)}{\partial \sigma^2}2(\sigma^2)^2 \\ \frac{\partial g_2(\mu,\sigma^2)}{\partial \mu}\sigma^2 & \frac{\partial g_2(\mu,\sigma^2)}{\partial \sigma^2}2(\sigma^2)^2 \end{bmatrix} \begin{bmatrix} \frac{\partial g_1(\mu,\sigma^2)}{\partial \mu} & \frac{\partial g_2(\mu,\sigma^2)}{\partial \mu} \\ \frac{\partial g_1(\mu,\sigma^2)}{\partial \sigma^2} & \frac{\partial g_2(\mu,\sigma^2)}{\partial \sigma^2} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial g_1(\mu,\sigma^2)}{\partial \mu}\sigma^2 \frac{\partial g_1(\mu,\sigma^2)}{\partial \mu} + \frac{\partial g_1(\mu,\sigma^2)}{\partial \sigma^2} 2(\sigma^2)^2 \frac{\partial g_1(\mu,\sigma^2)}{\partial \sigma^2} & con't \\ \frac{\partial g_2(\mu,\sigma^2)}{\partial \mu}\sigma^2 \frac{\partial g_1(\mu,\sigma^2)}{\partial \mu} + \frac{\partial g_2(\mu,\sigma^2)}{\partial \sigma^2} 2(\sigma^2)^2 \frac{\partial g_1(\mu,\sigma^2)}{\partial \sigma^2} & con't \end{bmatrix}$$

$$\begin{bmatrix} con't & \frac{\partial g_1(\mu,\sigma^2)}{\partial \mu}\sigma^2 \frac{\partial g_2(\mu,\sigma^2)}{\partial \mu} + \frac{\partial g_1(\mu,\sigma^2)}{\partial \sigma^2} 2(\sigma^2)^2 \frac{\partial g_2(\mu,\sigma^2)}{\partial \sigma^2} \\ con't & \frac{\partial g_2(\mu,\sigma^2)}{\partial \mu}\sigma^2 \frac{\partial g_2(\mu,\sigma^2)}{\partial \mu} + \frac{\partial g_2(\mu,\sigma^2)}{\partial \sigma^2} 2(\sigma^2)^2 \frac{\partial g_2(\mu,\sigma^2)}{\partial \sigma^2} \end{bmatrix}$$

Computing the derivatives:

$\frac{\partial g_1(\mu,\sigma^2)}{\partial \mu} = 3$ and $\frac{\partial g_1(\mu,\sigma^2)}{\partial \sigma^2} = 0$

$\frac{\partial g_2(\mu,\sigma^2)}{\partial \mu} = \frac{10\mu}{\sigma^4}$ and $\frac{\partial g_2(\mu,\sigma^2)}{\partial \sigma^2} = -10\frac{\mu^2}{\sigma^6}$

$$\Omega = \begin{bmatrix} 9\sigma^2 & 3\sigma^2 \frac{10\mu}{\sigma^4} \\ 3\sigma^2 \frac{10\mu}{\sigma^4} & \frac{100\mu^2}{\sigma^8}\sigma^2 + 100\frac{\mu^4}{\sigma^{12}} 2(\sigma^2)^2 \end{bmatrix}$$

Note: Although there is no correlation between the MLE estimates of μ and σ^2 (cov($\widehat{\mu}_{ml}, \widehat{\sigma}^2_{ml}) = 0$), there is correlation between the MLE estimates of η_1 and η_2.

This is because the MLE estimates of these parameters are functions of the same random variable, in this case $\widehat{\mu}_{ml}$.