Solution to Exercise 9.20

★9.20 Let the components z_t of the $n-$vector $\boldsymbol{z}$ be IID drawings from the $\mathrm{N}(\mu, \sigma^2)$ distribution, and let s^2 be the OLS estimate of the error variance from the regression of $\boldsymbol{z}$ on the constant vector $\boldsymbol{\iota}$. Show that the variance of s^2 is $2\sigma^4/(n-1)$.

Would this result still hold if the normality assumption were dropped? Without this assumption, what would you need to know about the distribution of the z_t in order to find the variance of s^2?

The OLS estimator of σ^2 is

$$s^2 = \frac{\boldsymbol{z}^\top \boldsymbol{M_\iota z}}{n-1},$$

where $\boldsymbol{M_\iota}$ is the matrix that projects orthogonally off the space spanned by the constant vector $\boldsymbol{\iota}$. The variance of s^2 is

$$\mathrm{E}\bigg(\Big(\frac{\boldsymbol{z}^\top \boldsymbol{M_\iota z}}{n-1}\Big)^2\bigg) - \bigg(\mathrm{E}\Big(\frac{\boldsymbol{z}^\top \boldsymbol{M_\iota z}}{n-1}\Big)\bigg)^2,$$

and, since s^2 is unbiased, this is just

$$\mathrm{E}\Big(\frac{\boldsymbol{z}^\top \boldsymbol{M_\iota z}}{n-1}\Big)^2 - \sigma^4. \tag{S9.43}$$

The problem, then, is to calculate the first term in (S9.43), which is $\mathrm{E}(s^4)$.

The expectation of s^4 can be written as

$$\frac{1}{(n-1)^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\sum_{r=1}^{n}\sum_{l=1}^{n}\mathrm{E}(z_t z_s z_r z_l) M_{ts} M_{rl},$$

where M_{ts} is the (t, s) element of $\boldsymbol{M_\iota}$. Now $\mathrm{E}(z_t z_s z_r z_l)$ is zero unless $t = s$ and $r = l$ with $t \neq r$, or else $t = r$ and $s = l$ with $t \neq s$, or else $t = l$ and $s = r$ with $t \neq s$, or else $t = s = l = r$. These cases are mutually exclusive. For each of the first three cases, there are exactly $n(n-1)$ terms in the quadruple sum that satisfy the conditions, and there are exactly n terms that satisfy the conditions of the fourth case.

For the first three cases, the expectation is $\sigma^2 \times \sigma^2 = \sigma^4$. For the fourth, because of the normality assumption, it is $\mathrm{E}(z_t^4) = 3\sigma^4$. Thus the expectation of s^4 is $(n-1)^{-2}\sigma^4$ times

$$\sum_{t=1}^{n}\sum_{r=1, r\neq t}^{n} M_{tt}M_{rr} + \sum_{t=1}^{n}\sum_{s=1, s\neq t}^{n} M_{ts}M_{ts} + \sum_{t=1}^{n}\sum_{s=1, s\neq t}^{n} M_{ts}M_{st} + 3\sum_{t=1}^{n} M_{tt}M_{tt}.$$

It is easy to see that the last term fills in the terms missing from the first three sums (here the factor of 3 serves to provide exactly enough terms), and so the expectation can be written more simply as

$$\frac{\sigma^4}{(n-1)^2}\Big(\sum_{t=1}^{n}\sum_{r=1}^{n} M_{tt}M_{rr} + \sum_{t=1}^{n}\sum_{s=1}^{n} M_{ts}M_{ts} + \sum_{t=1}^{n}\sum_{s=1}^{n} M_{ts}M_{st}\Big). \tag{S9.44}$$

The first term inside the large parentheses in (S9.44) is just the product of two independent sums, $\sum M_{tt}$ and $\sum M_{rr}$. Each of these sums is equal to $\mathrm{Tr}(\boldsymbol{M_\iota}) = n-1$. Thus the first term is $(n-1)^2$. The next two terms are equal, since $\boldsymbol{M_\iota}$ is symmetric. Given that

$$\sum_{s=1}^{n} M_{ts}M_{st} = (\boldsymbol{M_\iota M_\iota})_{tt} = M_{tt},$$

each of these two terms is equal to $\mathrm{Tr}(\boldsymbol{M_\iota}) = n-1$. Thus

$$\begin{aligned} \mathrm{E}(s^4) &= \frac{\sigma^4}{(n-1)^2}\big((n-1)^2 + 2(n-1)\big) = \frac{\sigma^4(n^2-1)}{(n-1)^2} \\ &= \frac{\sigma^4(n-1)(n+1)}{(n-1)^2} = \frac{\sigma^4(n+1)}{n-1}. \end{aligned} \tag{S9.45}$$

From (S9.43), the variance of s^2 is the above expression minus σ^4, that is,

$$\mathrm{Var}(s^4) = \sigma^4\Big(\frac{n+1}{n-1} - 1\Big) = \frac{2\sigma^4}{n-1}. \tag{S9.46}$$

This is the result we were asked to prove.

An easier approach would be to use the formula which tells us that, if $\boldsymbol{w}$ is standard normal and $\boldsymbol{A}$ and $\boldsymbol{B}$ are symmetric, then

$$\mathrm{E}(\boldsymbol{w}^\top\boldsymbol{A}\boldsymbol{w}\boldsymbol{w}^\top\boldsymbol{B}\boldsymbol{w}) = 2\,\mathrm{Tr}(\boldsymbol{AB}) + \mathrm{Tr}(\boldsymbol{A})\mathrm{Tr}(\boldsymbol{B});$$

this result, which was not discussed in the text, can be proved in the same way that we proved (S9.45).

The result (S9.46) would evidently not hold without the normality assumption. We used the fact that the fourth central moment of z_t is $3\sigma^4$ in order to get (S9.44). If the z_t followed some other distribution, we would need to know the fourth central moment of that distribution.