Assignment 2

2022-10-21

Question 1

a)

The test statistic for the hypothesis is

$$T = \frac{\hat{\theta}_3 - 2}{\sqrt{\hat{\Sigma}_{3,3}}} = \frac{4-2}{\sqrt{1.1}} = 1.9069252.$$

Since $1.9069252 > t_{97,0.95} = 1.6607146$ so we reject H_0.

b)

Let $\alpha = 0.05$, the 95% confidence interval is

$$\hat{\theta}_3 \pm t_{97,0.975}\sqrt{\hat{\Sigma}_{3,3}} = 4 \pm 1.98 \times 1.04 = [1.94, 6.05].$$

c)

We have the statistic

$$T = \frac{f(0,\hat{\theta})}{\sqrt{\hat{v}_x^T \hat{\Sigma} \hat{v}}} = \frac{0.54}{0.85} = 0.63$$

which is lower than $t_{97,0.975}$ so we can not reject H_0.

d)

The 95% confidence interval is

$$f(0,\hat{\theta}) \pm t_{97,0.975} \times \sqrt{\hat{v}_x^T \hat{\Sigma} \hat{v}} = [-1.14, 2.22].$$

e)

Since the dataset with the funtion $f(x, \theta)$, we can compute the matrix $\hat{V}$ where

$$\hat{V}_{ij} = \partial f(x_i, \hat{\theta}) / \partial \theta_j.$$

Thus, we can compute $(\hat{V}^T \hat{V})^{-1}$ and

$$\hat{\sigma}^2 = \hat{\Sigma}(\hat{V}^T \hat{V})$$

where $\hat{\Sigma}$ is the given estimated covariance matrix.

Question 2

a)

The estimates of θ are $\hat{\theta} = (0.81, -0.44, 1.98, 1.27)$ and $\hat{\sigma}^2 = \frac{S(\hat{\theta})}{n-p}$ where $S(\hat{\theta}) = RSE^2(n-p)$ where RSE is the residual standard error. Thus, we have $\hat{\sigma}^2 = RSE^2 = 0.275$. The residual sum of squares can be recovered from the residual standard error and the degree of freedoms because

$$RSE = \sqrt{\frac{RSS}{n-p}}$$

where p is the number of parameters θ, which in this case, is 4.

b)

We first obtain the estimated covariance matrix or just the values on the diagonal of $\hat{\Sigma}$ using $x_i = \frac{3(i-1)}{n-1}$ with $n = 100$ and the partial derivatives $\partial f(x_i, \hat{\theta})/\partial\theta_j$. Note that the partial derivatives calculations are quite straightforward so they will be omitted. We have $\hat{\Sigma}_{1,1} = 0.114$, $\hat{\Sigma}_{2,2} = 0.028$, $\hat{\Sigma}_{3,3} = 0.0076$, and $\hat{\Sigma}_{4,4} = 0.1104$. Like before, we have the test statistic

$$T = \frac{0.8}{\sqrt{0.114}} = 2.37 > t_{96,1-0.05/2} = 1.984$$

so we reject H_0. The 95% confidence interval is

$$\hat{\theta_1} \pm t_{96,1-0.05/2}\sqrt{\hat{\Sigma}_{1,1}} = 0.81 \pm 1.984 \times 0.337 = [0.141, 1.478].$$

c)

Likewise, we have the statistic

$$T = \frac{\hat{\theta_4} - 1}{\sqrt{\hat{\Sigma}_{4,4}}} = 0.81 < t_{96,1-0.05/2}$$

so we can not reject H_0. Let $\alpha = 0.02$ then the 98% confidence interval is

$$\hat{\theta_2} \pm t_{96,1-0.02/2}\sqrt{\hat{\Sigma}_{2,2}} = -0.44 \pm 2.36 \times 0.028 = [-0.5, -0.37].$$

d)

e)

Let the global model Ω be the full model, i.e, $S(\hat{\theta_q}) = \min_{\theta_q} ||Y - f_\Omega(x, \theta_q)||^2$, where $\theta_q \in \mathbb{R}^q$, and the linear submodel to be $\omega : S(\hat{\theta_p}) = \min_{\theta_p} ||Y - f_\omega(x, \theta_p)||^2$ where $\theta_p \in \mathbb{R}^p$ with $p < q$ such that $f_\omega(x, \theta_p)$ is linear. We can test for H_0 that the linear submodel fits well at significance level α by using the test statistic

$$V = \frac{[S(\hat{\theta_p}) - S(\hat{\theta_q})]/(q-p)}{S(\hat{\theta_q})/(n-q)}.$$

If $V > F_{q-p,n-q,1-\alpha}$ then we reject H_0 that the linear submodel fits well.