Exercise 2:

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad \text{and} \quad \sigma^2_{ad} = \frac{1}{n-p-1} \hat{\varepsilon}^T \hat{\varepsilon}.$$

1.
$$\hat{\beta} = (X^T X)^{-1} X^T y$$
$$= (X^T X)^{-1} X^T (X\beta + \varepsilon) \quad [\because y_i = \beta_0 + x_i \beta_1 + \varepsilon_i]$$
$$= (X^T X)^{-1} (X^T X)\beta + (X^T X)^{-1} X^T \varepsilon$$
$$= \beta + (X^T X)^{-1} X^T \varepsilon$$

Now $\mathbb{E}[\hat{\beta}] = \beta + (X^T X)^{-1} X^T \mathbb{E}[\varepsilon]$
$$= \beta \quad [\because \mathbb{E}[\varepsilon] = 0]$$

We know,
$$\text{cov}(\hat{\beta}) = \mathbb{E}[(\hat{\beta} - \mathbb{E}[\hat{\beta}])(\hat{\beta} - \mathbb{E}[\hat{\beta}])^T]$$
$$= \mathbb{E}[(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T]$$

Now, $\hat{\beta} - \beta = (X^T X)^{-1} X^T y - \beta$
$$= (X^T X)^{-1} X^T (X\beta + \varepsilon) - \beta$$
$$= \beta + (X^T X)^{-1} X^T \varepsilon - \beta$$
$$= (X^T X)^{-1} X^T \varepsilon.$$

So, $\text{cov}(\hat{\beta}) = \mathbb{E}[((X^T X)^{-1} X^T \varepsilon)((X^T X)^{-1} X^T \varepsilon)^T]$
$$= \mathbb{E}[(X^T X)^{-1} (X^T X) \varepsilon \varepsilon^T (X^T X)^{-1}]$$
$$= \mathbb{E}[\varepsilon \varepsilon^T (X^T X)^{-1}]$$

$$= (X^T X)^{-1} \underbrace{\mathbb{E}[\varepsilon \varepsilon^T]}$$

$$= (X^T X)^{-1} \cdot \boxed{\sigma^2 I_n}$$

$$= \sigma^2 (X^T X)^{-1}.$$

So, $\mathrm{cov}(\hat{\beta}) = \sigma^2 (X^T X)^{-1}.$

2.

$$\hat{\varepsilon} = y - X\hat{\beta}$$

so,

$$\mathbb{E}[\hat{\varepsilon}^T \hat{\varepsilon}] = \mathbb{E}[(y - X\hat{\beta})^T (y - X\hat{\beta})]$$

$$= \mathbb{E}[(y - X(X^T X)^{-1} X^T y)^T (y - X(X^T X)^{-1} X^T y)]$$

$$= \mathbb{E}[(y - Hy)^T (y - Hy)]$$

[Take $X(X^T X)^{-1} X^T = H$]

$$= \mathbb{E}[y^T (I_n - H)^T (I_n - H) y]$$

$$= \mathbb{E}[y^T (I_n - H) y]$$

$$= \mathrm{tr}((I_n - H)\sigma^2 I_n) + \beta^T X^T (I_n - H) X\beta$$

[$\mathbb{E}[x^T A x] = \mathrm{tr}(A\Sigma) + \mu^T A \mu$]

$$= \sigma^2 (n - p - 1) + \beta^T X^T (I_n - X(X^T X)^{-1} X^T) X\beta$$

[The matrix $I_n - H$ is also symmetric and Idempotent with rank $(I_n - H) = n - p - 1$.]

$$= \sigma^2(n-p-1) + \beta^T X^T X \beta - \beta^T X^T X (X^T X)^{-1} X^T X \beta$$

$$= \sigma^2(n-p-1) + \beta^T X^T X \beta - \beta^T X^T X \beta$$

$$= \sigma^2(n-p-1)$$

Now, $\mathbb{E}[\sigma^2_{ad}] = \dfrac{1}{n-p-1} \mathbb{E}[\hat{\varepsilon}^T \hat{\varepsilon}]$

$$\Rightarrow \mathbb{E}[\sigma^2_{ad}] = \frac{1}{n-p-1} \cdot \sigma^2 (n-p-1)$$

$$\Rightarrow \mathbb{E}[\sigma^2_{ad}] = \sigma^2.$$