# STATS 2107
# Statistical Modelling and Inference II
# Tutorial 4
# Solutions

Sharon Lee, Matt Ryan

Semester 2 2021

1. (a) Consider regression data

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

and let

$$S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2, S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}), S_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2.$$

Prove that

$$S_{xy} = \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}.$$

---

**Solutions:**

$$\begin{aligned}
S_{xy} &= \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \\
&= \sum_{i=1}^{n}(x_i y_i - x_i\bar{y} - \bar{x}y_i + \bar{x}\bar{y}) \\
&= \sum_{i=1}^{n} x_i y_i - \bar{y}\sum_{i=1}^{n} x_i - \bar{x}\sum_{i=1}^{n} y_i + n\bar{x}\bar{y} \\
&= \sum_{i=1}^{n} x_i y_i - n\bar{y}\bar{x} - n\bar{x}\bar{y} + n\bar{x}\bar{y} \\
&= \sum_{i=1}^{n} x_i y_i - n\bar{y}\bar{x}.
\end{aligned}$$

---

(b) Consider independent random variables $Y_1, Y_2, \ldots, Y_n$ with

$$E(Y_i) = \beta_0 + \beta_1 x_i \quad \text{and} \quad \mathrm{Var}(Y_i) = \sigma^2.$$

Let

$$\hat{E}_i = Y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i).$$

Prove that

$$E\left[\sum_{i=1}^{n}\{Y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)\}^2\right] = (n-2)\sigma^2.$$

Hence deduce that $S_e^2$ is an unbiased estimator for $\sigma^2$.

---

**Solutions:**
Recall from lectures the following properties of residuals:

$$E[\hat{E}_i] = 0 \quad \text{and} \quad \text{Var}(\hat{E}_i) = \sigma^2 \left(1 - \frac{1}{n} - \frac{(x_i - \bar{x})^2}{S_{xx}}\right).$$

We have

$$\begin{aligned}
E\left[\sum_{i=1}^{n}\{Y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)\}^2\right] &= E\left(\sum_{i=1}^{n}\hat{E}_i^2\right) = \sum_{i=1}^{n} E(\hat{E}_i^2) = \sum_{i=1}^{n}\left\{\text{Var}(\hat{E}_i) + E(\hat{E}_i)^2\right\} \\
&= \sum_{i=1}^{n}\sigma^2\left\{1 - 1/n - (x_i - \bar{x})^2/S_{xx}\right\} + 0^2 \\
&= \sigma^2\left\{\sum_{i=1}^{n} 1 - \sum_{i=1}^{n} 1/n - \sum_{i=1}^{n}(x_i - \bar{x})^2/S_{xx}\right\} \\
&= \sigma^2(n - n/n - S_{xx}/S_{xx}) \\
&= (n-2)\sigma^2.
\end{aligned}$$

Hence,

$$E(S_e^2) = \frac{1}{(n-2)} E\left[\sum_{i=1}^{n}\{Y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)\}^2\right] = \frac{(n-2)\sigma^2}{(n-2)} = \sigma^2.$$

---

2. Suppose $X$ is an $n \times p$ matrix with linearly independent columns and let

$$\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top.$$

In lectures, we have stated that the matrices $\boldsymbol{H}$ and $(\boldsymbol{I} - \boldsymbol{H})$ are symmetric and idempotent. We will prove these properties here.

(a) Show that $\boldsymbol{H}$ is symmetric, *i.e.*, $\boldsymbol{H} = \boldsymbol{H}^\top$.

---

**Solutions:**

$$\begin{aligned}
\boldsymbol{H}^T &= (\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top)^\top \\
&= (\boldsymbol{X}^\top)^T((\boldsymbol{X}^T \boldsymbol{X})^{-1})^\top \boldsymbol{X}^\top && \text{since } (\boldsymbol{ABC})^\top = \boldsymbol{C}^\top \boldsymbol{B}^\top \boldsymbol{A}^\top \\
&= \boldsymbol{X}((\boldsymbol{X}^\top \boldsymbol{X})^\top)^{-1}\boldsymbol{X}^\top && \text{since } (\boldsymbol{A}^\top)^\top = \boldsymbol{A} \\
&= \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top && \text{since } (\boldsymbol{A}^{-1})^\top = (\boldsymbol{A}^\top)^{-1} \\
&= \boldsymbol{H}
\end{aligned}$$

So $\boldsymbol{H}$ is symmetric.

---

(b) Show that $\boldsymbol{H}$ is idempotent, that is, $\boldsymbol{H}^2 = \boldsymbol{H}$.

---

**Solutions:**

$$\begin{aligned}\boldsymbol{H}^2 &= (\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top)(\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top)\\ &= \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\\ &= \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\\ &= \boldsymbol{H}\end{aligned}$$

So $\boldsymbol{H}$ is idempotent.

---

(c) Show that $(\boldsymbol{I}-\boldsymbol{H})$ is symmetric and idempotent, where $\boldsymbol{I}$ is the $n\times n$ identity matrix.

---

**Solutions:**
First symmetry:

$$\begin{aligned}(\boldsymbol{I}-\boldsymbol{H})^T &= \boldsymbol{I}^T-\boldsymbol{H}^T\\ &= \boldsymbol{I}-\boldsymbol{H}\,,\end{aligned}$$

by Part (a) and since the identity matrix is trivially symmetric.

Now to show it is idempotent:

$$\begin{aligned}(\boldsymbol{I}-\boldsymbol{H})^2 &= (\boldsymbol{I}-\boldsymbol{H})(\boldsymbol{I}-\boldsymbol{H})\\ &= \boldsymbol{I}^2-\boldsymbol{H}-\boldsymbol{H}+\boldsymbol{H}^2\\ &= \boldsymbol{I}-2\boldsymbol{H}+\boldsymbol{H} \quad \text{by Part (b)}\\ &= \boldsymbol{I}-\boldsymbol{H}\,.\end{aligned}$$

So $(\boldsymbol{I}-\boldsymbol{H})$ is idempotent.

---

3. Consider the multiple regression model

$$M:\ \boldsymbol{Y}=X\boldsymbol{\beta}+\boldsymbol{\epsilon},$$

where $Y_i\sim N(\eta_i,\sigma^2)$ independently for $i=1,2,\dots,n$ and $E[\boldsymbol{Y}]=\boldsymbol{\eta}=X\boldsymbol{\beta}$. The vector of residuals is defined by $\hat{\boldsymbol{E}}=\boldsymbol{Y}-X\hat{\boldsymbol{\beta}}$. In lectures, we have stated some properties of $\hat{\boldsymbol{E}}$. We will look at these here.

(a) Prove that $\hat{\boldsymbol{E}}=(I-H)\boldsymbol{Y}$, where $H=X(X^TX)^{-1}X^T$.

---

**Solutions:**

$$\begin{aligned}\hat{\boldsymbol{E}} &= \boldsymbol{Y}-X\hat{\beta}\\ &= \boldsymbol{Y}-X(X^TX)^{-1}X^T\boldsymbol{Y}\\ &= (I-X(X^TX)^{-1}X^T)\boldsymbol{Y}\\ &= (I-H)\boldsymbol{Y}.\end{aligned}$$

---

(b) Prove that $E(\hat{\boldsymbol{E}})=\mathbf{0}$ (and hence $E(\hat{E}_i)=0$ for $i=1,2,\dots,n$).

**Solutions:**

$$
\begin{aligned}
E[\hat{\boldsymbol{E}}] &= E[(I-H)\boldsymbol{Y}] \\
&= (I-H)E[\boldsymbol{Y}] \\
&= (I-H)\boldsymbol{\eta} \\
&= (I-H)X\boldsymbol{\beta} \\
&= X\boldsymbol{\beta} - X(X^TX)^{-1}X^TX\boldsymbol{\beta} \\
&= X\boldsymbol{\beta} - X\boldsymbol{\beta} \\
&= \mathbf{0}.
\end{aligned}
$$

So $E(E_i) = 0, i = 1, 2, \ldots, n$.

---

(c) Prove that: $\mathrm{Var}(\hat{E}_i) = \sigma^2(1-h_{ii})$, where $h_{ii}$ is the $(i,i)$th element of $H$. **Hint: Let $(I-H)$ have rows $\boldsymbol{a}_1^T, \boldsymbol{a}_2^T, \ldots, \boldsymbol{a}_n^T$.**

---

**Solutions:**
Let

$$
(I-H)\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \\ \vdots \\ \boldsymbol{a}_n^T \end{pmatrix} \boldsymbol{Y}
$$

So $\hat{E}_i = \boldsymbol{a}_i^T \boldsymbol{Y}$.

From lectures (Lemma 7), we have

$$
\mathrm{Var}(\hat{E}_i) = \sigma^2 \boldsymbol{a}_i^T \boldsymbol{a}_i
$$

Note that the $(i,i)^{th}$ entry of the matrix $(I-H)(I-H)^T$ is $\boldsymbol{a}_i^T\boldsymbol{a}_i$, but from Q2 above we know that

$$
(I-H)(I-H)^T = (I-H)
$$

so we get

$$
\mathrm{Var}(\hat{E}_i) = \sigma^2 \boldsymbol{a}_i^T \boldsymbol{a}_i = \sigma^2(1-h_{ii}),
$$

where $h_{ii}$ is the $i$th diagonal element of $H$.

---

4. Linearise the following equations:
   (a) $Y = \alpha\beta^x$

---

**Solutions:**

$$
\log(Y) = \log(\alpha) + \log(\beta)x\,.
$$

Taking $Y^* = \log(Y)$, $x^* = x$, $\beta_0 = \log(\alpha)$ and $\beta_1 = \log(\beta)$, we get

$$
Y^* = \beta_0 + \beta_1 x^*\,.
$$

(b) $Y = \alpha e^{\frac{\beta}{x}}$

**Solutions:**

$$\log(Y) = \log(\alpha) + \beta \frac{1}{x}\,.$$

Taking $Y^* = \log(Y)$, $x^* = \frac{1}{x}$, $\beta_0 = \log(\alpha)$ and $\beta_1 = \beta$, we get

$$Y^* = \beta_0 + \beta_1 x^*\,.$$

(c) $Y = \alpha + \frac{\beta}{x}$

**Solutions:**

$$Y = \alpha + \beta \frac{1}{x}\,.$$

Taking $Y^* = Y$, $x^* = \frac{1}{x}$, $\beta_0 = \alpha$ and $\beta_1 = \beta$, we get

$$Y^* = \beta_0 + \beta_1 x^*\,.$$

(d) $Y = \frac{\alpha}{\beta + x}$

**Solutions:**

$$\frac{1}{Y} = \frac{\beta}{\alpha} + \frac{1}{\alpha} x\,.$$

Taking $Y^* = \frac{1}{Y}$, $x^* = x$, $\beta_0 = \frac{\beta}{\alpha}$ and $\beta_1 = \frac{1}{\alpha}$, we get

$$Y^* = \beta_0 + \beta_1 x^*\,.$$

(e) $Y = \alpha + \beta x^n$

**Solutions:**

$$Y = \alpha + \beta(x^n)\,.$$

Taking $Y^* = Y$, $x^* = x^n$, $\beta_0 = \alpha$ and $\beta_1 = \beta$, we get

$$Y^* = \beta_0 + \beta_1 x^*\,.$$

(f) $Y = \frac{1}{\alpha + \beta e^{-x}}$

**Solutions:**

$$\frac{1}{Y} = \alpha + \beta e^{-x}\,.$$

Taking $Y^* = \frac{1}{Y}$, $x^* = e^{-x}$, $\beta_0 = \alpha$ and $\beta_1 = \beta$, we get

$$Y^* = \beta_0 + \beta_1 x^*\,.$$

---

(g) $Y = e^{-\alpha x_1 e^{-\frac{\beta}{x_2}}}$

---

**Solutions:**

$$\log(Y) = -\alpha x_1 e^{-\frac{\beta}{x_2}}$$

$$\log(\log(Y)) = \log(-\alpha) + \log(x_1) - \beta\frac{1}{x_2}$$

Taking $Y^* = \log(\log(Y))$, $x_1^* = \log(x_1)$, $x_2^* = \frac{1}{x_2}$, $\beta_0 = \log(-\alpha)$ and $\beta_1 = 1$, and $\beta_2 = -\beta$, we get

$$Y^* = \beta_0 + \beta_1 x_1^* + x_2^*\,.$$

---

***The following question is optional:***

5. In the proof of Theorem 11, we have used the result that if $\boldsymbol{X}$ is a random variable with $E(\boldsymbol{X}) = \boldsymbol{\mu}$ and $\mathrm{Var}(\boldsymbol{X}) = \boldsymbol{\Sigma}$, then
$$E(\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X}) = tr(\boldsymbol{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}^\top \boldsymbol{A}\boldsymbol{\mu}.$$
Prove this result.

---

**Solutions:**
Let $X_i$ and $X_j$ be two elements of $\boldsymbol{X}$. Observe that

$$E(X_iX_j) = \mathrm{Cov}(X_iX_j) + E(X_i)E(X_j) = \sigma_{ij} + \mu_i\mu_j,$$

where $\sigma_{ij}$ is the $ij$th element of $\boldsymbol{\Sigma}$.

Suppose $\boldsymbol{X}$ is of dimension $n$, then $\boldsymbol{A}$ is an $n \times n$ matrix. Note that

$$\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X} = \begin{bmatrix} X_1 & X_2 & \dots X_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}X_iX_j,$$

where $a_{ij}$ is the $ij$th element of $\boldsymbol{A}$.

Recall that the trace is defined as the sum of the diagonal elements of a matrix, that is, $tr(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}$. Observe that

$$\boldsymbol{A}\boldsymbol{\Sigma} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} a_{1i}\sigma_{i1} & \sum_{i=1}^{n} a_{1i}\sigma_{i2} & \dots & \sum_{i=1}^{n} a_{1i}\sigma_{in} \\ \sum_{i=1}^{n} a_{2i}\sigma_{i1} & \sum_{i=1}^{n} a_{2i}\sigma_{i2} & \dots & \sum_{i=1}^{n} a_{2i}\sigma_{in} \\ \vdots & \vdots & & \vdots \\ \sum_{i=1}^{n} a_{ni}\sigma_{i1} & \sum_{i=1}^{n} a_{ni}\sigma_{i2} & \dots & \sum_{i=1}^{n} a_{ni}\sigma_{in} \end{bmatrix}.$$

Hence, $tr(\boldsymbol{A\Sigma}) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ji}\sigma_{ij}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\sigma_{ij}$, as $\boldsymbol{\Sigma}$ is symmetric and hence $\sigma_{ji} = \sigma_{ij}$.

It follows that

$$\begin{aligned}
E(\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X}) &= E\left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}X_iX_j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}E\left[X_iX_j\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\left(\sigma_{ij} + \mu_i\mu_j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\sigma_{ij} + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\mu_i\mu_j \\
&= tr\left(\boldsymbol{A\Sigma}\right) + \boldsymbol{\mu}^\top \boldsymbol{A}\boldsymbol{\mu}.
\end{aligned}$$

---