Applied Statistics, assignment 3

KOUAGOU N'Dah Jean

October 13, 2018

Question 1

Suppose $\hat{\theta}$ is an estimator of θ.
The bias is given by:$b(\hat{\theta}) = E(\hat{\theta}) - \theta$.
Let's prove this theorem: $MSE(\hat{\theta}) = \text{var}(\hat{\theta}) + [b(\hat{\theta})]^2$.

We have:

$$\begin{aligned}
MSE(\hat{\theta}) &= E[(\hat{\theta} - \theta)^2] && (1)\\
&= E(\hat{\theta}^2 - 2\hat{\theta}\theta + \theta^2) && (2)\\
&= E(\hat{\theta}^2) - 2\theta E(\hat{\theta}) + \theta^2 && (3)\\
&= E(\hat{\theta}^2) - [E(\hat{\theta})]^2 - 2\theta E(\hat{\theta}) + [E(\hat{\theta})]^2 + \theta^2 && (4)\\
&= \text{var}(\hat{\theta}) + [b(\hat{\theta})]^2 && (5)\\
\text{Then, } MSE(\hat{\theta}) &= \text{var}(\hat{\theta}) + [b(\hat{\theta})]^2. && (6)
\end{aligned}$$

Question 2

Proof of some equalities:

1. Show that $SST = \sum_{i=1}^{a} \sum_{j=1}^{n} y_{ij}^2 - \frac{y_{\bullet\bullet}^2}{N}$ where $N = a \times n$

We have:

$$\begin{aligned}
SST &= \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{\bullet\bullet})^2 && (7)\\
&= \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}^2 - 2y_{ij}\bar{y}_{\bullet\bullet} + \bar{y}_{\bullet\bullet}^2) && (8)\\
&= \sum_{i=1}^{a}\sum_{j=1}^{n}y_{ij}^2 - 2\bar{y}_{\bullet\bullet}\sum_{i=1}^{a}\sum_{j=1}^{n}y_{ij} + \sum_{i=1}^{a}\sum_{j=1}^{n}\bar{y}_{\bullet\bullet}^2 && (9)\\
&= \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}^2) - \frac{2}{N}y_{\bullet\bullet}\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}) + N\bar{y}_{\bullet\bullet}^2 && (10)\\
&= \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}^2) - 2\frac{y_{\bullet\bullet}^2}{N} + N\frac{y_{\bullet\bullet}^2}{N^2} && (11)\\
&= \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}^2) - \frac{y_{\bullet\bullet}^2}{N} && (12)\\
\text{Therefore, } SST &= \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}^2) - \frac{y_{\bullet\bullet}^2}{N} && (13)
\end{aligned}$$

2. Show that $SS_{Treatment} = \sum_{i=1}^{a}(\frac{y_{i\bullet}^2}{n}) - \frac{y_{\bullet\bullet}^2}{N}$

   We have:

$$\begin{aligned}
SS_{Treatment} &= \sum_{i=1}^{a} n(\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet})^2 && (14)\\
&= \sum_{i=1}^{a} n(\bar{y}_{i\bullet}^2 - 2\bar{y}_{\bullet\bullet}\bar{y}_{i\bullet} + \bar{y}_{\bullet\bullet}^2) && (15)\\
&= \sum_{i=1}^{a} n(\frac{\bar{y}_{i\bullet}^2}{n^2} - 2\frac{y_{i\bullet}}{n}\frac{y_{\bullet\bullet}}{N} + \frac{y_{\bullet\bullet}^2}{N^2}) && (16)\\
&= \sum_{i=1}^{a}(\frac{\bar{y}_{i\bullet}^2}{n} - 2y_{i\bullet}\frac{y_{\bullet\bullet}}{N} + n\frac{y_{\bullet\bullet}^2}{N^2}) && (17)\\
&= \sum_{i=1}^{a}\frac{\bar{y}_{i\bullet}^2}{n} - 2\frac{y_{\bullet\bullet}}{N}\sum_{i=1}^{a}y_{i\bullet} + n\frac{y_{\bullet\bullet}^2}{N^2}\sum_{i=1}^{a}1 && (18)\\
&= \sum_{i=1}^{a}(\frac{\bar{y}_{i\bullet}^2}{n}) - 2\frac{y_{\bullet\bullet}^2}{N} + n \times a \times \frac{y_{\bullet\bullet}^2}{N^2} && (19)\\
&= \sum_{i=1}^{a}(\frac{\bar{y}_{i\bullet}^2}{n}) - 2\frac{y_{\bullet\bullet}^2}{N} + \frac{y_{\bullet\bullet}^2}{N} && (20)\\
\text{As a result, } SS_{Treatment} &= \sum_{i=1}^{a}(\frac{y_{i\bullet}^2}{n}) - \frac{y_{\bullet\bullet}^2}{N} && (21)
\end{aligned}$$

3. Show that MSE is an unbiased estimator of σ^2, i.e $MSE = \sigma^2$

We have:

$$
\begin{aligned}
\mathrm{E}(MSE) &= \mathrm{E}(\frac{SSE}{N-a}) \text{ and } \mathrm{E}(SSE) = \mathrm{E}(SST - SS_{Treatment}) && (22)\\
\text{Therefore:} & && (23)\\
\mathrm{E}(SSE) &= \mathrm{E}[(\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}^2) - \frac{y_{\bullet\bullet}^2}{N}) - (\sum_{i=1}^{a}(\frac{y_{i\bullet}^2}{n}) - \frac{y_{\bullet\bullet}^2}{N})] && (24)\\
&= \mathrm{E}[\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}^2) - \sum_{i=1}^{a}(\frac{y_{i\bullet}^2}{n})] && (25)\\
&= \mathrm{E}[\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i\bullet})^2 + 2\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}\bar{y}_{i\bullet}) && (26)\\
&- \sum_{i=1}^{a}\sum_{j=1}^{n}(\bar{y}_{i\bullet})^2 - \sum_{i=1}^{a}(\frac{y_{i\bullet}^2}{n})] && (27)\\
&= \mathrm{E}[\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i\bullet})^2 + 2\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}\bar{y}_{i\bullet}) && (28)\\
&- \sum_{i=1}^{a}\sum_{j=1}^{n}(\bar{y}_{i\bullet})^2 - \sum_{i=1}^{a}(\frac{y_{i\bullet}^2}{n})] && (29)\\
&= \mathrm{E}[\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i\bullet})^2 + 2\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}\frac{y_{i\bullet}}{n}) && (30)\\
&- \sum_{i=1}^{a}\sum_{j=1}^{n}(\frac{y_{i\bullet}}{n})^2 - \sum_{i=1}^{a}(\frac{y_{i\bullet}^2}{n})] && (31)\\
&= \mathrm{E}[\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i\bullet})^2 + 2\sum_{i=1}^{a}(\frac{y_{i\bullet}}{n})\sum_{j=1}^{n}(y_{ij}) && (32)\\
&- \sum_{i=1}^{a}(\sum_{j=1}^{n}(\frac{y_{i\bullet}}{n})^2) - \sum_{i=1}^{a}(\frac{y_{i\bullet}^2}{n})] && (33)\\
&= \mathrm{E}[\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i\bullet})^2 + 2\sum_{i=1}^{a}(\frac{y_{i\bullet}^2}{n}) && (34)\\
&- \sum_{i=1}^{a}(\frac{y_{i\bullet}^2}{n}) - \sum_{i=1}^{a}(\frac{y_{i\bullet}^2}{n})] && (35)\\
&= \mathrm{E}[\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i\bullet})^2] && (36)\\
&= \sum_{i=1}^{a}[(n-1)\sigma^2] && (37)\\
\text{cause } \ \sigma^2 = \mathrm{E}[\frac{\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i\bullet})^2}{n-1}] & && (38)
\end{aligned}
$$

$$\text{Then, } \mathrm{E}(SSE) = a \times (n-1)\sigma^2 \tag{39}$$
$$= (N-a)\sigma^2 \tag{40}$$
$$\text{As a result, } \mathrm{E}(MSE) = \sigma^2 \tag{41}$$

We conclude that MSE is an unbiased estimator of σ^2.

4. Show that $MS_{treatment}$ is an unbiased estimator of σ^2

   Under the null hypothesis, $\theta_1 = \theta_2 = \ldots \theta_a = \theta$, each $\bar{y}_{i\bullet}$ is normally distributed with the mean θ and the standard deviation $\frac{\sigma^2}{n}$.
   It also means that each $\sqrt{n}\bar{y}_{i\bullet}$ is normally distributed with the mean $\sqrt{n}\theta$ and the standard deviation σ^2.
   Since $\bar{y}_{\bullet\bullet}$ is an estimation of θ, we have:

$$\begin{aligned}
\mathrm{E}(MS_{treatment}) &= \mathrm{E}[\frac{SS_{treatment}}{a-1}] && (42)\\
&= \frac{1}{a-1}\mathrm{E}[\sum_{i=1}^{a}(\sqrt{n}\bar{y}_{i\bullet} - \sqrt{n}\bar{y}_{\bullet\bullet})^2] && (43)\\
&= \frac{1}{a-1}\mathrm{E}[\sum_{i=1}^{a} n(\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet})^2] && (44)\\
&= \sigma^2 && (45)
\end{aligned}$$

Question 3

Proof of some equalities:

1. Show that $SST = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{y_{\bullet\bullet}^2}{N}$ where $N = a \times b$

We have:

$$\begin{aligned}
SST &= \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij}-\bar{y}_{\bullet\bullet})^2 && (46)\\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij}^2-2y_{ij}\bar{y}_{\bullet\bullet}+\bar{y}_{\bullet\bullet}^2) && (47)\\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}y_{ij}^2-2\bar{y}_{\bullet\bullet}\sum_{i=1}^{a}\sum_{j=1}^{b}y_{ij}+\sum_{i=1}^{a}\sum_{j=1}^{b}\bar{y}_{\bullet\bullet}^2 && (48)\\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij}^2)-\frac{2}{N}y_{\bullet\bullet}\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij})+N\bar{y}_{\bullet\bullet}^2 && (49)\\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij}^2)-2\frac{y_{\bullet\bullet}^2}{N}+N\frac{y_{\bullet\bullet}^2}{N^2} && (50)\\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij}^2)-\frac{y_{\bullet\bullet}^2}{N} && (51)\\
\text{Therefore, } SST &= \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij}^2)-\frac{y_{\bullet\bullet}^2}{N} && (52)
\end{aligned}$$

2. Show that $SS_{Treatment}=\sum_{i=1}^{a}(\frac{y_{i\bullet}^2}{b})-\frac{y_{\bullet\bullet}^2}{N}$

We have:

$$\begin{aligned}
SS_{Treatment} &= \sum_{i=1}^{a}b(\bar{y}_{i\bullet}-\bar{y}_{\bullet\bullet})^2 && (53)\\
&= \sum_{i=1}^{a}b(\bar{y}_{i\bullet}^2-2\bar{y}_{\bullet\bullet}\bar{y}_{i\bullet}+\bar{y}_{\bullet\bullet}^2) && (54)\\
&= \sum_{i=1}^{a}b(\frac{\bar{y}_{i\bullet}^2}{b^2}-2\frac{y_{i\bullet}}{b}\frac{y_{\bullet\bullet}}{N}+\frac{y_{\bullet\bullet}^2}{N^2}) && (55)\\
&= \sum_{i=1}^{a}(\frac{\bar{y}_{i\bullet}^2}{b}-2y_{i\bullet}\frac{y_{\bullet\bullet}}{N}+b\frac{y_{\bullet\bullet}^2}{N^2}) && (56)\\
&= \sum_{i=1}^{a}\frac{\bar{y}_{i\bullet}^2}{b}-2\frac{y_{\bullet\bullet}}{N}\sum_{i=1}^{a}y_{i\bullet}+b\frac{y_{\bullet\bullet}^2}{N^2}\sum_{i=1}^{a}(1) && (57)\\
&= \sum_{i=1}^{a}(\frac{\bar{y}_{i\bullet}^2}{b})-2\frac{y_{\bullet\bullet}^2}{N}+b\times a\times\frac{y_{\bullet\bullet}^2}{N^2} && (58)\\
&= \sum_{i=1}^{a}(\frac{\bar{y}_{i\bullet}^2}{b})-2\frac{y_{\bullet\bullet}^2}{N}+\frac{y_{\bullet\bullet}^2}{N} && (59)\\
\text{As a result, } SS_{Treatment} &= \sum_{i=1}^{a}(\frac{y_{i\bullet}^2}{b})-\frac{y_{\bullet\bullet}^2}{N} && (60)
\end{aligned}$$

3. Show that $SS_{Block} = \sum_{j=1}^{b}(\frac{y_{\bullet j}^2}{b}) - \frac{y_{\bullet\bullet}^2}{N}$

   We have:

$$
\begin{aligned}
SS_{Block} &= \sum_{j=1}^{b} a(\bar{y}_{\bullet j} - \bar{y}_{\bullet\bullet})^2 && (61)\\
&= \sum_{j=1}^{b} a(\bar{y}_{\bullet j}^2 - 2\bar{y}_{\bullet\bullet}\bar{y}_{\bullet j} + \bar{y}_{\bullet\bullet}^2) && (62)\\
&= \sum_{j=1}^{b} a(\frac{\bar{y}_{\bullet j}^2}{a^2} - 2\frac{y_{\bullet j}}{a}\frac{y_{\bullet\bullet}}{N} + \frac{y_{\bullet\bullet}^2}{N^2}) && (63)\\
&= \sum_{j=1}^{b}(\frac{\bar{y}_{\bullet j}^2}{a} - 2y_{\bullet j}\frac{y_{\bullet\bullet}}{N} + a\frac{y_{\bullet\bullet}^2}{N^2}) && (64)\\
&= \sum_{j=1}^{b}\frac{\bar{y}_{\bullet j}^2}{a} - 2\frac{y_{\bullet\bullet}}{N}\sum_{j=1}^{b} y_{\bullet j} + a\frac{y_{\bullet\bullet}^2}{N^2}\sum_{j=1}^{b}(1) && (65)\\
&= \sum_{j=1}^{b}(\frac{\bar{y}_{\bullet j}^2}{a}) - 2\frac{y_{\bullet\bullet}^2}{N} + a \times b \times \frac{y_{\bullet\bullet}^2}{N^2} && (66)\\
&= \sum_{j=1}^{b}(\frac{\bar{y}_{\bullet j}^2}{a}) - 2\frac{y_{\bullet\bullet}^2}{N} + \frac{y_{\bullet\bullet}^2}{N} && (67)\\
\text{As a result, } SS_{Block} &= \sum_{j=1}^{b}(\frac{y_{\bullet j}^2}{a}) - \frac{y_{\bullet\bullet}^2}{N} && (68)
\end{aligned}
$$