# STA 206 - Homework 1 (Solution)

1. (a) $\mathbf{E}(A\mathbf{Z}) = A\mathbf{E}(\mathbf{Z})$.

*Proof.* LHS(left hand side) and RHS(right hand side) are both vectors, so it suffices to show they are equal element-wisely.

$$
\begin{aligned}
(A\mathbf{Z})_j &= \sum_k a_{jk}\mathbf{Z}_k, \quad j = 1, \cdots, s, \\
(E(A\mathbf{Z}))_j &= E((A\mathbf{Z})_j) = E(\sum_k a_{jk}\mathbf{Z}_k) \\
&= \sum_k a_{jk}E(\mathbf{Z}_k) = (AE(\mathbf{Z}))_j, \quad j = 1, \cdots, s.
\end{aligned}
$$

□

(b) $\mathbf{Cov}(A\mathbf{Z}, B\mathbf{Z}) = A\Sigma B^T$. So in particular, $\mathbf{Var}(A\mathbf{Z}) = A\Sigma A^T$.

*Proof.* Method 1: Define

$$
W = A\mathbf{Z},\ U = B\mathbf{Z},\ C = \mathrm{Cov}(W, U),\ D = A\Sigma B^T.
$$

Then

$$
\begin{aligned}
C_{ij} &= \mathrm{Cov}(W_i, U_j) = \mathrm{Cov}(\sum_k a_{ik}\mathbf{Z}_k, \sum_k b_{jk}\mathbf{Z}_k) \\
&= \sum_k \sum_l a_{ik}b_{jl}\mathrm{Cov}(\mathbf{Z}_k, \mathbf{Z}_l) \\
&= \sum_k \sum_l a_{ik}b_{jl}\Sigma_{kl} = D_{ij}, \quad i = 1, \cdots, s, \quad j = 1, \cdots, t.
\end{aligned}
$$

Method 2:

$$
\begin{aligned}
\mathbf{Cov}(A\mathbf{Z}, B\mathbf{Z}) &= \mathbb{E}\left[(A\mathbf{Z} - \mathbb{E}(A\mathbf{Z}))(B\mathbf{Z} - \mathbb{E}(B\mathbf{Z}))^T\right] \\
&= \mathbb{E}\left[(A\mathbf{Z} - \mathbb{E}(A\mathbf{Z}))(\mathbf{Z}^T B^T - (\mathbb{E}(B\mathbf{Z}))^T)\right] \\
&= \mathbb{E}\left[A\mathbf{Z}\mathbf{Z}^T B^T - A\mathbf{Z}(\mathbb{E}(B\mathbf{Z}))^T - \mathbb{E}(A\mathbf{Z})\mathbf{Z}^T B^T + \mathbb{E}(A\mathbf{Z})(\mathbb{E}(B\mathbf{Z}))^T\right] \\
&= A\mathbb{E}[\mathbf{Z}\mathbf{Z}^T]B^T - A\mathbb{E}[\mathbf{Z}](\mathbb{E}(B\mathbf{Z}))^T - A\mathbb{E}(\mathbf{Z})\mathbb{E}[\mathbf{Z}^T]B^T + A\mathbb{E}(\mathbf{Z})(\mathbb{E}(B\mathbf{Z}))^T \\
&= A\mathbb{E}[\mathbf{Z}\mathbf{Z}^T]B^T - A\mathbb{E}(\mathbf{Z})(\mathbb{E}[\mathbf{Z}^T])B^T \\
&= A\left[\mathbb{E}[\mathbf{Z}\mathbf{Z}^T] - \mathbb{E}(\mathbf{Z})(\mathbb{E}[\mathbf{Z}])^T\right]B^T \\
&= A\mathrm{Var}(\mathbf{Z})B^T \\
&= A\Sigma B^T
\end{aligned}
$$

In particular,

$$
\mathrm{Var}(A\mathbf{Z}) = \mathrm{Cov}(A\mathbf{Z}, A\mathbf{Z}) = A\Sigma A^T.
$$

□

2. Derive the following.

(a) $\sum_{i=1}^n (X_i - \overline{X}) = 0$

*Proof.*

$$\sum_{i=1}^n (X_i - \overline{X}) = \sum_{i=1}^n X_i - \sum_{i=1}^n \overline{X} = n\overline{X} - n\overline{X} = 0.$$

□

(b) $\sum_{i=1}^n (X_i - \overline{X})^2 = \sum_{i=1}^n (X_i - \overline{X})X_i = \sum_{i=1}^n X_i^2 - n(\overline{X})^2$.

*Proof.*

$$\begin{aligned}
\sum_{i=1}^n (X_i - \overline{X})^2 &= \sum_{i=1}^n (X_i - \overline{X})(X_i - \overline{X}) = \sum_{i=1}^n (X_i - \overline{X})X_i - \sum_{i=1}^n (X_i - \overline{X})\overline{X} \\
&= \sum_{i=1}^n (X_i - \overline{X})X_i - \overline{X}\sum_{i=1}^n (X_i - \overline{X}) \\
&= \sum_{i=1}^n (X_i - \overline{X})X_i && \text{(from part (a))} \\
&= \sum_{i=1}^n X_i^2 - \overline{X}\sum_{i=1}^n X_i = \sum_{i=1}^n X_i^2 - n(\overline{X})^2
\end{aligned}$$

□

(c) $\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y}) = \sum_{i=1}^n (X_i - \overline{X})Y_i = \sum_{i=1}^n X_i(Y_i - \overline{Y}) = \sum_{i=1}^n X_iY_i - n\overline{X}\,\overline{Y}$.

*Proof.* We can write $(X_i - \overline{X})(Y_i - \overline{Y})$ as $(X_i - \overline{X})Y_i - (X_i - \overline{X})\overline{Y}$ and also as $X_i(Y_i - \overline{Y}) - \overline{X}(Y_i - \overline{Y})$, then apply the same technique used in part (b) to conclude that

$$\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y}) = \sum_{i=1}^n (X_i - \overline{X})Y_i = \sum_{i=1}^n X_i(Y_i - \overline{Y}).$$

Moreover,

$$\begin{aligned}
\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y}) &= \sum_{i=1}^n (X_i - \overline{X})Y_i \\
&= \sum_{i=1}^n X_iY_i - \overline{X}\sum_{i=1}^n Y_i \\
&= \sum_{i=1}^n X_iY_i - n\overline{X}\,\overline{Y}.
\end{aligned}$$

□

**Remark.** *Notice here $X_i$'s and $Y_i$'s can be either fixed values or random variables.*

3. Least-squares principle.

(a) State the least-squares principle.

*Solution.* For a given line: $y = b_0 + b_1 x$, the *sum of squared vertical deviations* of the observations $\{(X_i, Y_i)\}_{i=1}^n$ from the corresponding points on the line is:

$$Q(b_0, b_1) = \sum_{i=1}^{n} (Y_i - (b_0 + b_1 X_i))^2 .$$

The *least squares (LS) principle* is to fit the observed data by minimizing the sum of squared vertical deviations. □

(b) Derive the LS estimators for simple linear regression model.

*Solution.* From the lecture notes, the LS estimators can be derived by finding $b_0$ and $b_1$ which satisfy the normal equations:

$$nb_0 + b_1 \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} Y_i, \tag{1}$$

$$b_0 \sum_{i=1}^{n} X_i + b_1 \sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} X_i Y_i. \tag{2}$$

From equation (1),

$$b_0 = \overline{Y} - b_1 \overline{X}.$$

Using this in equation (2) we have

$$\begin{aligned} b_0 n\overline{X} + b_1 \sum_{i=1}^{n} X_i^2 &= \sum_{i=1}^{n} X_i Y_i \\ \Rightarrow n\overline{XY} + b_1 [\sum_{i=1}^{n} X_i^2 - n\overline{X}^2] &= \sum_{i=1}^{n} X_i Y_i \\ \Rightarrow b_1 &= \frac{\sum_{i=1}^{n} X_i Y_i - n\overline{XY}}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2} \end{aligned}$$

Now from 2(a) and (b),

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} X_i Y_i - n\overline{XY}}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2} = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n} (X_i - \overline{X})^2}.$$

From equation (1),

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}.$$

□

(c) Assume the observations follow:

$$Y_i = \exp(a + bX_i) + \epsilon_i, \quad i = 1, \cdots, n,$$

where $a, b \in \mathbb{R}$ are unknown parameters and $\epsilon_i$s are uncorrelated random variables with $E(\epsilon_i) = 0, Var(\epsilon_i) = \sigma^2$. Describe how to estimate the regression function (equivalently, $a, b$) by least-squares principle. (Notes: You only need to provide a description. This is an example of a nonlinear regression model. )

*Solution.* Plot implies linear regression model is not accurate for non-linear dataset. The red linear is a linear fit which is inappropriate. Instead, we use non-linear regression and define the following sum of square error (SSE)

$$Q(a,b) = \sum_{i=1}^{n} (Y_i - \exp(a + bX_i))^2 .$$

The *least squares (LS) principle* is to fit the observed data by minimizing the sum of squared vertical deviations:

$$(\hat{a}, \hat{b}) = \underset{a,b}{\operatorname{argmin}} \sum_{i=1}^{n} (Y_i - \exp(a + bX_i))^2 .$$

□

4. Tell true or false (with a brief explanation) of the following statements with regard to simple linear regression.

   (a) The least squares line always passes the center of the data $(\overline{X}, \overline{Y})$.

**ANS.** True. since $y = \overline{Y} + \hat{\beta}_1(x - \overline{X})$. This can also be derived from equation (1).

   (b) If $\overline{X} = 0,\ \ \overline{Y} = 0$, then $\hat{\beta}_0 = 0$ no matter what is $\hat{\beta}_1$.

**ANS.** True since $\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}$

   (c) Given the sample size, the larger the range of $X_i$s, the smaller the standard errors of $\hat{\beta}_0, \hat{\beta}_1$ tend to be.

**ANS.** True. since $s\{\hat{\beta}_0\}$ and $s\{\hat{\beta}_1\}$ has $\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}$ in the denominator.

5. Properties of the residuals under simple linear regression model. Recall that

$$\begin{aligned} e_i &= Y_i - \widehat{Y}_i = Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i), \quad i = 1, \cdots n. \\ &= (Y_i - \overline{Y}) - \hat{\beta}_1(X_i - \overline{X}). \end{aligned}$$

Show that

   (a) $\sum_{i=1}^{n} e_i = 0$.

   *Proof.*

$$\begin{aligned} \sum_{i=1}^{n} e_i &= \sum_{i=1}^{n}(Y_i - \overline{Y}) - \sum_{i=1}^{n} \hat{\beta}_1(X_i - \overline{X}) \\ &= (\sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} \overline{Y}) - \hat{\beta}_1(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \overline{X}) \\ &= (n\overline{Y} - n\overline{Y}) - \hat{\beta}_1(n\overline{X} - n\overline{X}) \\ &= 0. \end{aligned}$$

□

   (b) $\sum_{i=1}^{n} X_i e_i = 0$.

*Proof.* Using results of 2(b) and 2(c),

$$\begin{aligned}
\sum_{i=1}^{n} X_i e_i &= \sum_{i=1}^{n} X_i((Y_i - \overline{Y}) - \hat{\beta}_1(X_i - \overline{X})) \\
&= (\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \overline{Y}) - \hat{\beta}_1(\sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} X_i \overline{X}) \\
&= (\sum_{i=1}^{n} X_i Y_i - n\overline{X}\ \overline{Y}) - \hat{\beta}_1(\sum_{i=1}^{n} X_i^2 - n(\overline{X})^2) \\
&= \sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}) - \hat{\beta}_1 \sum_{i=1}^{n}(X_i - \overline{X})^2 \\
&= \sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}) - \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \sum_{i=1}^{n}(X_i - \overline{X})^2 \\
&= 0.
\end{aligned}$$

□

(c) $\sum_{i=1}^{n} \widehat{Y}_i e_i = 0$.

*Proof.* By parts (a) and (b), and $\widehat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$.

$$\sum_{i=1}^{n} \hat{Y}_i e_i = \hat{\beta}_0 \sum_{i=1}^{n} e_i + \hat{\beta}_1 \sum_{i=1}^{n} X_i e_i = 0 + 0 = 0.$$

□

6. Under the simple linear regression model:

(a) Show that the LS estimator $\hat{\beta}_0$ is an unbiased estimator of $\hat{\beta}_0$.

*Proof.*

$$\begin{aligned}
E(\hat{\beta}_0) &= E(\bar{Y}) - E(\hat{\beta}_1 \bar{X}) \\
&= \frac{1}{n} \sum(\beta_0 + \beta_1 X_i) - \bar{X} E(\hat{\beta}_1) \\
&= \beta_0 + \beta_1 \bar{X} - \bar{X}\beta_1 && \text{(since } E(\hat{\beta}_1) = \beta_1\text{)} \\
&= \beta_0.
\end{aligned}$$

Then to show $E(\hat{\beta}_1) = \beta_1$, using results of 2(a), 2(b) and noticing $E(\epsilon_i) = 0$, we have

$$\begin{aligned}
E(\hat{\beta}_1) &= E\left(\frac{\sum_{i=1}^{n}(X_i - \overline{X})Y_i}{\sum_{i=1}^{n}(X_i - \overline{X})^2}\right) \\
&= E\left(\frac{\sum_{i=1}^{n}(X_i - \overline{X})(\beta_1 X_i + \beta_0 + \epsilon_i)}{\sum_{i=1}^{n}(X_i - \overline{X})^2}\right) \\
&= \frac{\sum_{i=1}^{n}(X_i - \overline{X})X_i}{\sum_{i=1}^{n}(X_i - \overline{X})^2}\beta_1 = \beta_1
\end{aligned}$$

□

(b) Derive the variance formula for $\hat{\beta}_0$.

*Proof.*

$$
\begin{aligned}
\mathrm{Var}(\hat{\beta}_0) &= \mathrm{Var}(\bar{Y} - \hat{\beta}_1 \bar{X}) \\
&= \mathrm{Var}(\bar{Y}) - 2\mathrm{Cov}(\bar{Y}, \hat{\beta}_1 \overline{X}) + \mathrm{Var}(\hat{\beta}_1 \overline{X}) \\
&= \frac{1}{n^2}\mathrm{Var}\left(\sum Y_i\right) + (\overline{X})^2 \mathrm{Var}(\hat{\beta}_1) && (\text{since } \mathrm{Cov}(\bar{Y}, \hat{\beta}_1 \overline{X}) = 0) \\
&= \frac{n\sigma^2}{n^2} + (\overline{X})^2 \left[\frac{\sigma^2}{\sum_{i=1}^n (X_i - \overline{X})^2}\right] \\
&= \sigma^2 \left[\frac{1}{n} + \frac{\overline{X}^2}{\sum_{i=1}^n (X_i - \overline{X})^2}\right]
\end{aligned}
$$

To show that $\mathrm{Cov}(\bar{Y}, \hat{\beta}_1 \overline{X}) = 0$, we need to note that

$$
\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \overline{X}) Y_i}{\sum_{i=1}^n (X_i - \overline{X})^2} = \sum_{i=1}^n K_i Y_i \quad \text{where} \quad K_i = \frac{(X_i - \overline{X})}{\sum_{i=1}^n (X_i - \overline{X})^2}.
$$

Then we can see that

$$
\begin{aligned}
\mathrm{Cov}(\bar{Y}, \hat{\beta}_1 \overline{X}) = \overline{X}\mathrm{Cov}(\bar{Y}, \hat{\beta}_1) &= \mathrm{Cov}\left(\frac{\sum_{i=1}^n Y_i}{n}, \sum_{i=1}^n K_i Y_i\right) \\
&= \frac{\sum_{i=1}^n K_i \mathrm{Var}(Y_i)}{n} \\
&= \frac{\sigma^2}{n}\sum_{i=1}^n K_i = 0. && (\text{because } \textstyle\sum K_i = 0)
\end{aligned}
$$

□

7. R in appendix

8. Random design simple linear regression.

(a) Show that the LS estimators are unbiased.

*Solution.* We know from lecture and homework that, conditional on $\mathbf{X} = (X_1, \cdots, X_n)$, $\mathbb{E}(\hat{\beta}_1 \mid \mathbf{X}) = \beta_1$, $\mathbb{E}(\hat{\beta}_0 \mid \mathbf{X}) = \beta_0$. Thus using the formula:

$$
E(\hat{\beta}_0) = E(E(\hat{\beta}_0 | \mathbf{X})) = E(\beta_0) = \beta_0, \quad E(\hat{\beta}_1) = E(E(\hat{\beta}_1 | \mathbf{X})) = E(\beta_1) = \beta_1.
$$

Therefore both estimators are unbiased. □

(b) Derive their variances.

*Solution.* From lecture and homework, we know that conditional on $\mathbf{X}$, we have

$$
\mathrm{Var}(\hat{\beta}_1 \mid \mathbf{X}) = \frac{\sigma^2}{S_{xx}}, \qquad \mathrm{Var}(\hat{\beta}_0 \mid \mathbf{X}) = \sigma^2\left(\frac{1}{n} + \frac{\bar{X}^2}{S_{xx}}\right).
$$

□

Therefore, using the formula:

$$\mathrm{Var}(\hat{\beta}_1) = \mathbb{E}[\mathrm{Var}(\hat{\beta}_1 \mid \mathbf{X})] + \mathrm{Var}(\mathbb{E}[\hat{\beta}_1 \mid \mathbf{X}]) = \mathbb{E}\left[\frac{\sigma^2}{S_{xx}}\right] + Var(\beta_1) = \sigma^2\,\mathbb{E}\left[\frac{1}{S_{xx}}\right].$$

Similarly,

$$\mathrm{Var}(\hat{\beta}_0) = \mathbb{E}[\mathrm{Var}(\hat{\beta}_0 \mid \mathbf{X})] + \mathrm{Var}(\mathbb{E}[\hat{\beta}_0 \mid \mathbf{X}]) = \mathbb{E}\left[\frac{\sigma^2}{n} + \frac{\bar{X}^2\sigma^2}{S_{xx}}\right] + Var(\beta_0) = \sigma^2\left(\frac{1}{n} + \mathbb{E}\left[\frac{\bar{X}^2}{S_{xx}}\right]\right).$$

(c) Large-sample behavior.

*Solution.* Let $\mu_X = \mathbb{E}[X_i]$ and $\sigma_X^2 = \mathrm{Var}(X_i)$ (assuming they exist), then by law of large numbers $S_{xx}/n \to \sigma_X^2$, $\bar{X} \to \mu_X$. Hence

$$\mathrm{Var}(\hat{\beta}_1) = \frac{\sigma^2}{n\sigma_X^2} + o\left(\frac{1}{n}\right), \qquad \mathrm{Var}(\hat{\beta}_0) = \frac{\sigma^2}{n}\left(1 + \frac{\mu_X^2}{\sigma_X^2}\right) + o\left(\frac{1}{n}\right).$$

Thus as $n \to \infty$, we have $\mathrm{Var}(\hat{\beta}_1) \to 0$ and $\mathrm{Var}(\hat{\beta}_0) \to 0$.
Note that, by CLT, we can show the asymptotic distributions are:

$$\sqrt{n}(\hat{\beta}_1 - \beta_1) \Rightarrow \mathcal{N}\left(0, \frac{\sigma^2}{\sigma_X^2}\right), \quad \sqrt{n}(\hat{\beta}_0 - \beta_0) \Rightarrow \mathcal{N}\left(0, \sigma^2\left(1 + \frac{\mu_X^2}{\sigma_X^2}\right)\right).$$

$\square$

# HW1: Question 7

Jingzhi SUn

2025/10/10

## Simulation by R

You need to submit your codes alongside with the answers, plots, outputs, etc. For this homework, you are encouraged (though not required) to use R Markdown: Please submit a .rmd file and its corresponding .html file. Later in the quarter, you may be required to use R Markdown for some homework or quiz problems. *(Hint: Use the `help` function if needed)*

### (a)

Create a sequence of consecutive integers ranging from 1 to 100. Record these in a vector x. *(Hint: use the `seq` function)*

```
x<-seq(1,100)
#x <- 1:100 #equivalent
```

### (b)

Create a new vector w by the formula: $w = 2 + 0.5 * x$.

```
w<-2+0.5*x
```

### (c)

Randomly sample 100 numbers from a Normal distribution with mean zero and standard deviation 5. Calculate the sample mean and sample variance and draw a histogram. What do you observe? *(Hint: use the `rnorm` function)*

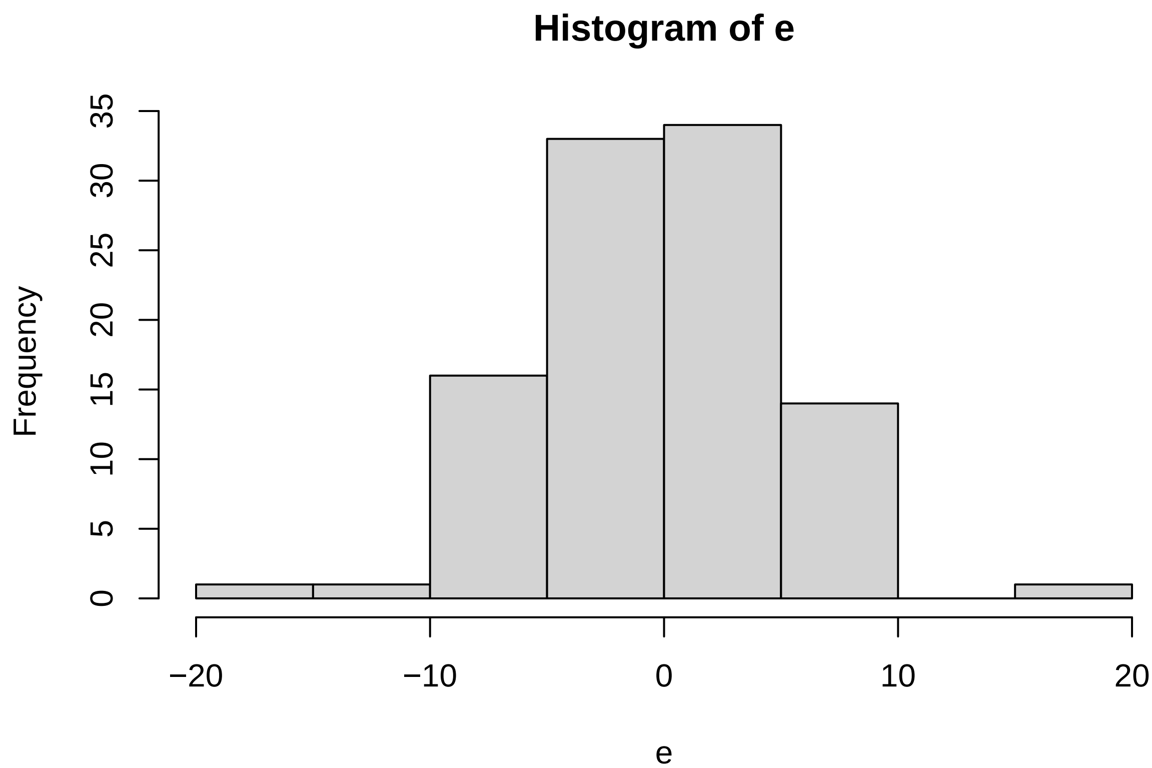
```
e <- rnorm(n = 100 , mean = 0 , sd = 5)
mean(e)
```

```
## [1] -0.1553721
```

```
var(e)
```

```
## [1] 26.20812
```

```
hist(e)
```




We observe that the histogram is approximate to a normal distribution, and the sample mean and sample variance are not too far from their true values.

### (d)

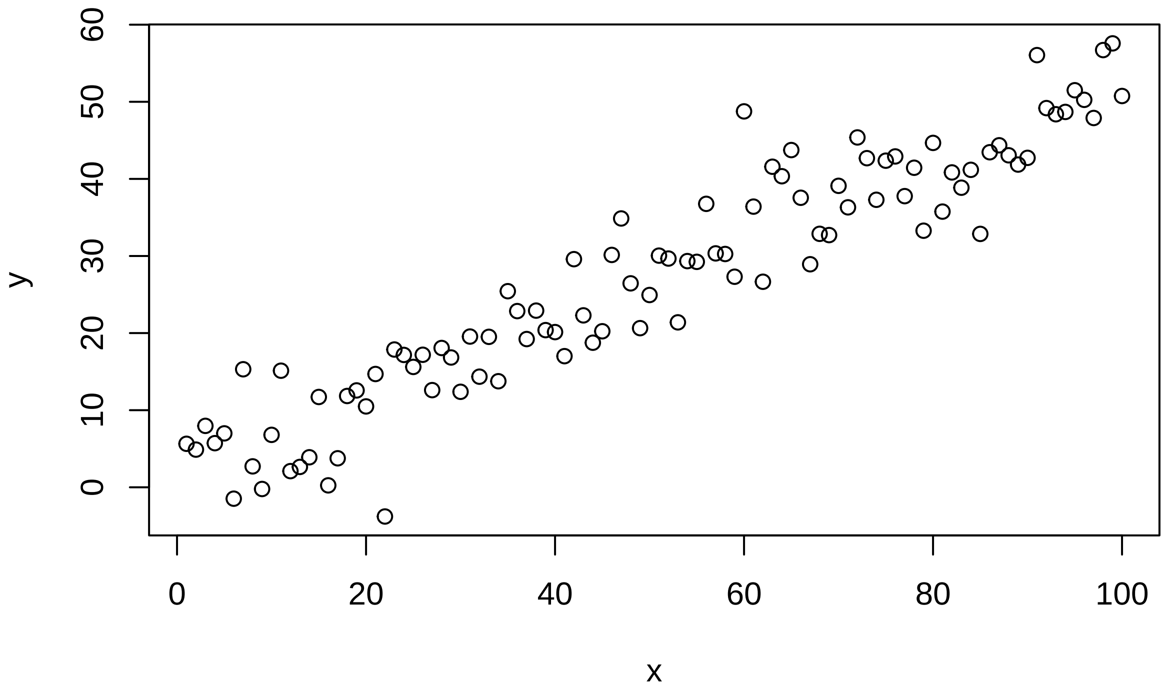
Add (element-wise) the numbers created in part (c) to the vector w. Record the new vector as y.

```
y<-w+e
```

### (e)

Draw the scatter plot of y versus x.

```
plot(x,y)
```



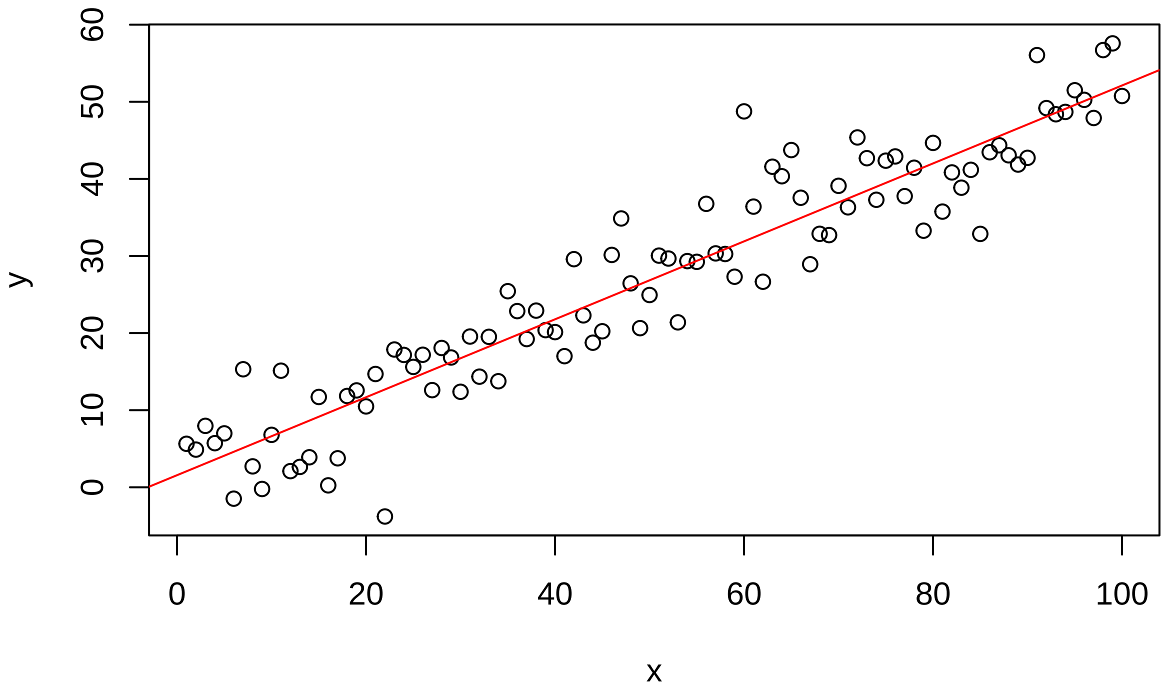

### (f)

Estimate the regression coefficients of y on x. Add the fitted regression line to the scatter plot in part (e). What do you observe?

```r
fit<- lm( y ~ x )
beta0 <- fit$coef[ 1 ]
beta1 <- fit$coef[2]
c(beta0,beta1)
```

```
## (Intercept)           x
##   1.5673452   0.5054907
```

```r
plot (x, y)
abline ( a = beta0 , b = beta1 , col = "red")
```

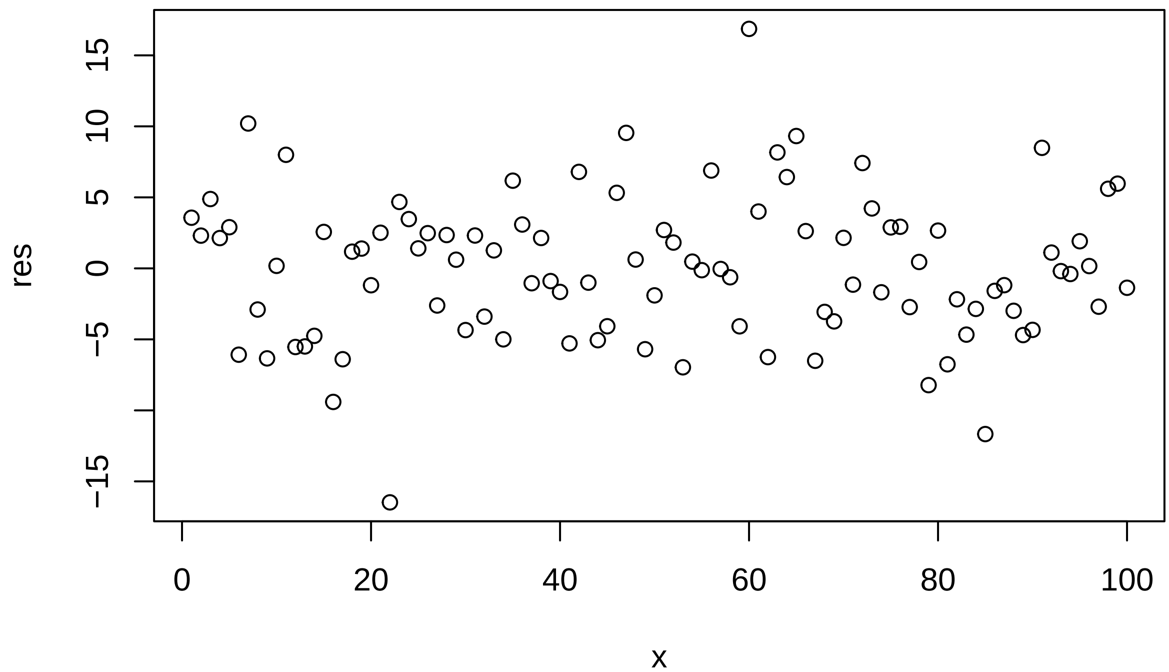



We observe the fitted regression line passing through the data cloud.

**(g)**

Calculate the residuals and draw a scatter plot of residuals versus x. What do you observe? Derive MSE.

```
res <- residuals (fit)
plot (x , res )
```




```
sse <- sum( res ^2)
mse <- sse /(100 - 2)
mse
```

```
## [1] 26.44991
```

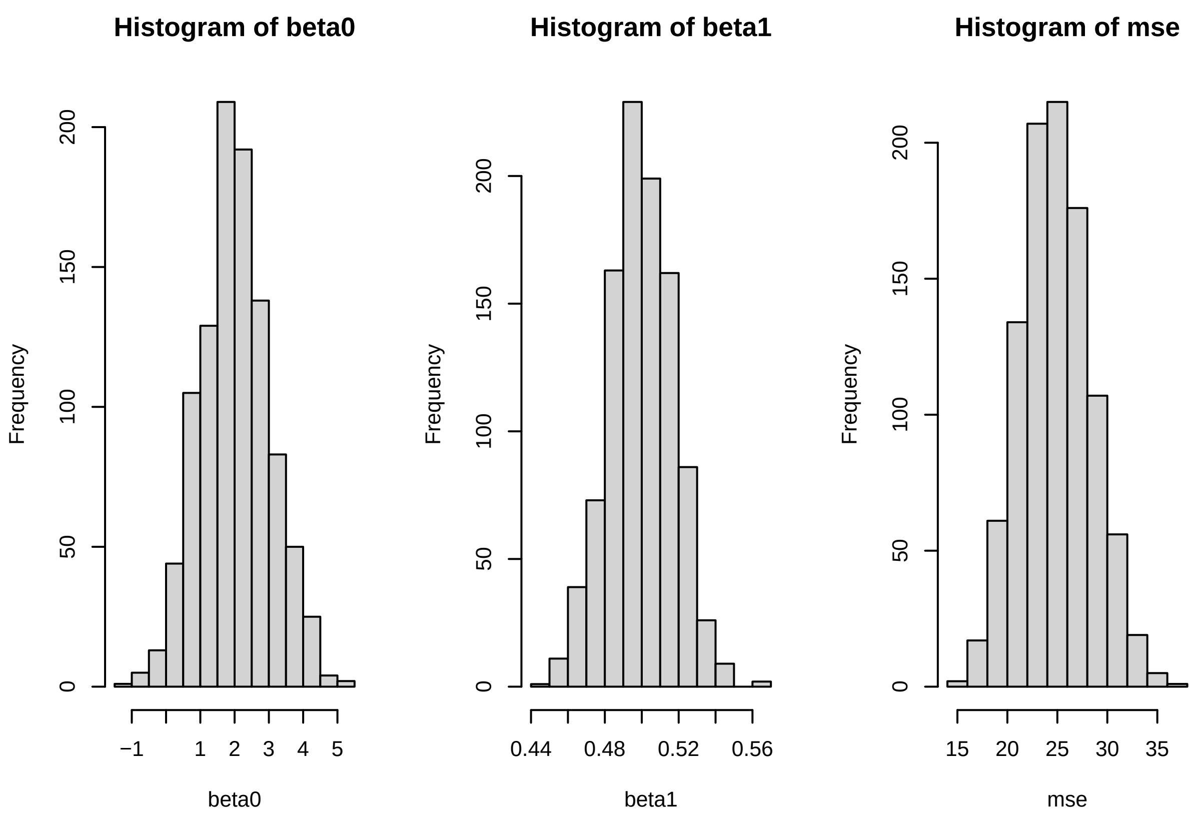
We observe that residuals are randomly distributed, and the mean squared error is not too far from the true variance.

### (h)

Repeat parts (c) – (d) 1000 times. Each time, derive the fitted regression coefficients and MSE and record them. Draw histogram and calculate sample mean and sample variance for each of the three estimators. Summarize your observations. *(Hint: use the `for` loop)*

```
beta0 <- c ( )
beta1 <- c ( )
mse <- c ( )
for ( k in 1 : 1000 ) {
  e <- rnorm(n = 100 , mean = 0 , sd = 5)
  y <- w + e
  fit <- lm( y ~ x )
  beta0 [k] <- fit$coef[ 1 ]
  beta1 [k] <- fit$coef[ 2 ]
  res <- residuals (fit)
```

```r
  mse[k]<-sum(res^2)/(100-2)
}
par(mfrow=c(1,3))
hist(beta0)
hist(beta1)
hist(mse)
```




```r
mean(beta0)
```

```
## [1] 2.016542
```

```r
mean(beta1)
```

```
## [1] 0.4997548
```

```r
mean(mse)
```

```
## [1] 24.85165
```

```r
var(beta0)
```

```
## [1] 1.056001
```

```r
var(beta1)
```

```
## [1] 0.0003106439
```

```r
var(mse)
```

```
## [1] 12.26074
```

The distribution of each of the three estimators is bell shaped, with sample mean close to true parameter.