

# Statistics (ABC260S) Test 2 Memo 2025

■ represents one mark; ▲ represents half a mark

## Questions

---

1. [10 marks] Answer the following multiple choice questions (two marks each).

a. [2 marks] What is the capital city of Viet Nam?

- i. Hanoi Correct ■ ■
- ii. Lome
- iii. Kingstown
- iv. Panama

b. [2 marks] What is the deepest lake in the world?

- i. Matano
- ii. Kivu
- iii. Sarez
- iv. Baikal Correct ■ ■

c. [2 marks] What is  $4 + 10$ ?

- i. 16
- ii. 30
- iii. 11
- iv. 14 Correct ■ ■

d. [2 marks] Which of the below capital cities is furthest south?

- i. Minsk
- ii. Bucuresti
- iii. Gaborone Correct ■ ■
- iv. Chisinau

2. [11 marks] Below are given the final marks of a class of fifteen students.

40 57 31 84 37 58 43 54 71 40 83 61 70 72 29

a. [3 marks] Find the mean of the final marks.

$$\begin{aligned}
 \bar{x} &= \frac{1}{n} \sum_{i=1}^n x_i \blacksquare \\
 &= \frac{1}{15} [40 + 57 + 31 + 84 + 37 + 58 + 43 + 54 + 71 + 40 + 83 + 61 + 70 + 72 + 29] \blacksquare \\
 &= \frac{830}{15} = 55.3333 \blacksquare
 \end{aligned}$$

b. [3 marks] Find the median of the final marks.

First we put the data in increasing order:  $\blacksquare$

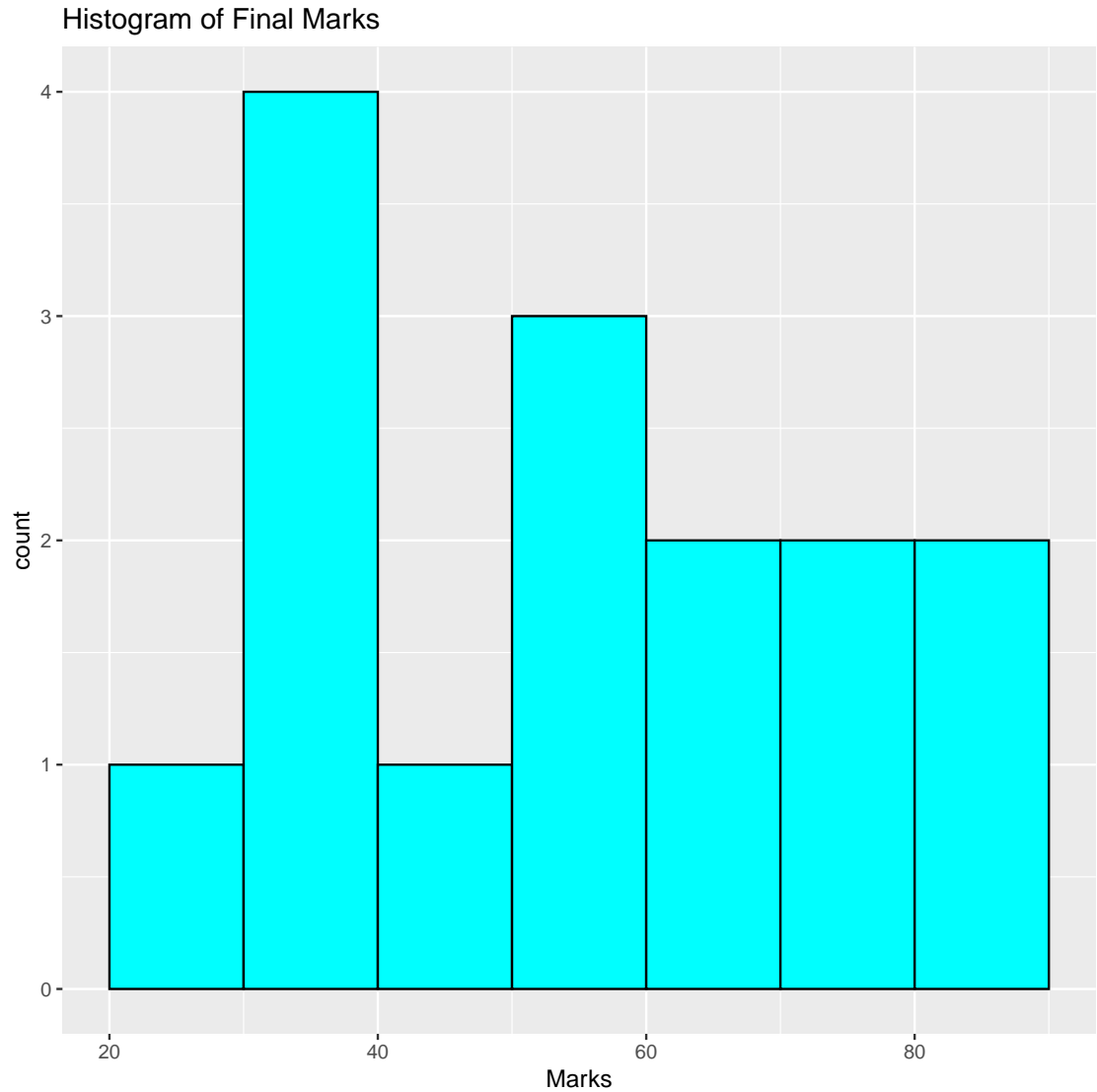
29 31 37 40 40 43 54 57 58 61 70 71 72 83 84

Then we take the 8th value.  $\blacksquare$ . Hence, the median is 57  $\blacksquare$

c. [3 marks] Find the standard deviation of the final marks.

$$\begin{aligned}
 s^2 &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \blacksquare \\
 &= \frac{1}{14} \left[ \left(40 - \frac{830}{15}\right)^2 + \left(57 - \frac{830}{15}\right)^2 + \cdots + \left(29 - \frac{830}{15}\right)^2 \right] \blacksquare \\
 s &= \sqrt{329.5238095} \\
 &= 18.1528 \blacksquare
 \end{aligned}$$

d. [2 marks] Draw a histogram of the final marks. Use a bin width of 10.



■ for correctly labelled axes; ■ for correct drawn histogram

3. [9 marks] Consider the random variable  $X$  with probability density function as follows:

$$f_X(x) = \begin{cases} k(6x + 6x^2) & \text{for } 0 \leq x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

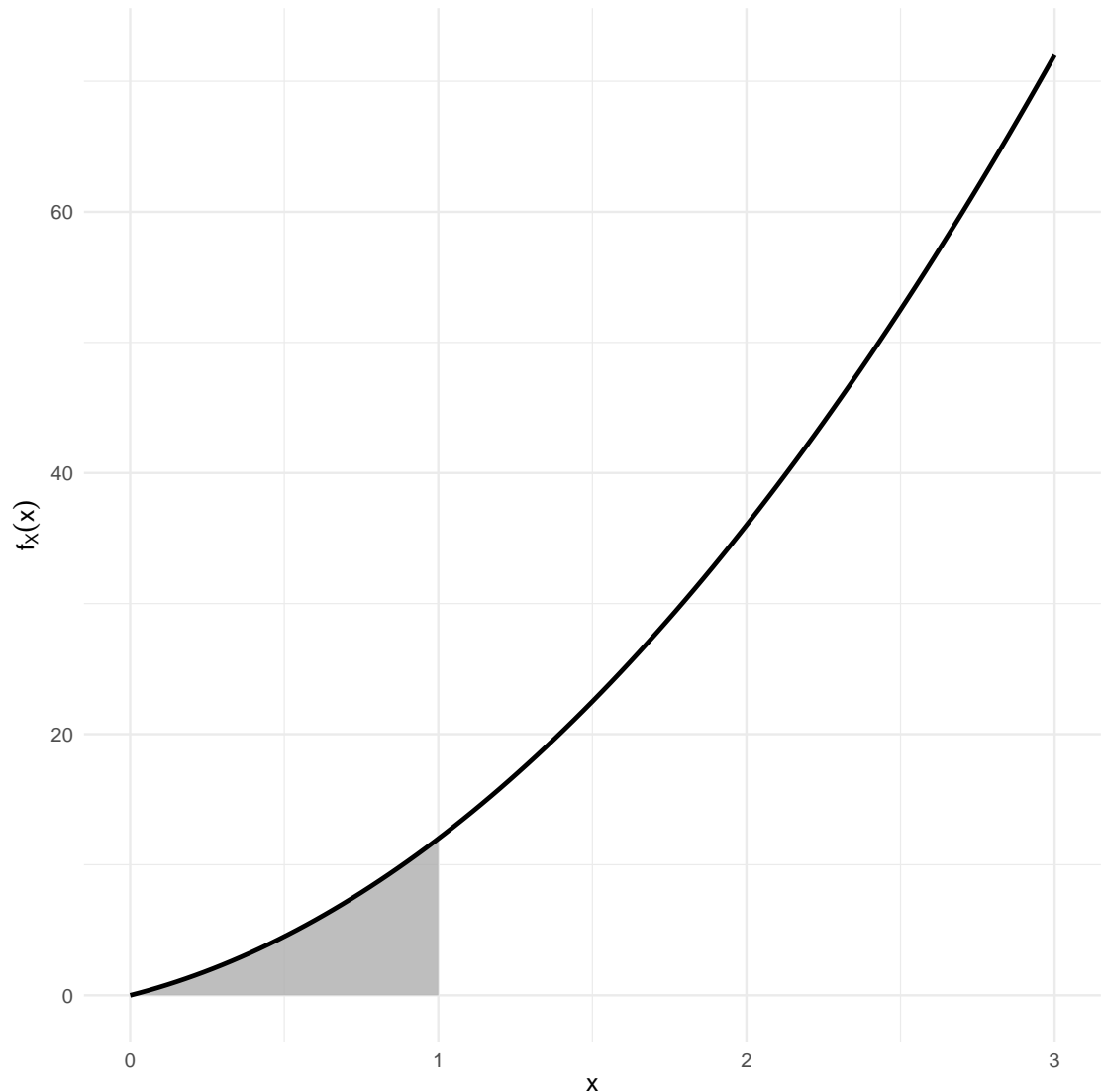
a. [3 marks] Find the value of  $k$ .

$$\begin{aligned}
 k \int_0^3 f_X(x) dx &= 1 \quad \blacksquare \\
 k \int_0^3 6x + 6x^2 dx &= 1 \\
 k [3x^2 + 2x^3]_0^3 &= 1 \quad \blacksquare \\
 k (81) &= 1 \\
 k &= \frac{1}{81} = 0.012346 \quad \blacksquare
 \end{aligned}$$

b. [3 marks] Find  $\Pr(X < 1)$ .

$$\begin{aligned}
 \Pr(X < 1) &= \int_0^1 6x + 6x^2 dx \quad \blacksquare \\
 &= [3x^2 + 2x^3]_0^1 \quad \blacksquare \\
 &= \left(\frac{5}{81}\right) - (0) \\
 &= \frac{5}{81} = 0.06173 \quad \blacksquare
 \end{aligned}$$

c. [3 marks] Show the probability just calculated on a graph of  $f_X(x)$ .



■ for correctly labeled axes, ■ for correctly plotted  $f_X(x)$ , ■ for correct region shaded

4. [8 marks] Consider the dataset provided in the Excel file `surf4cars_Toyota_Corolla.xlsx`. A description of the columns in the dataset is as follows. Answer the following questions using R code and output only.

a. [4 marks] Fit a linear regression model to the data where Asking Price is the response variable and EngineSize and YearModel are explanatory variables. Display a summary of the model output.

```
toyota <- read_excel("surf4cars_Toyota_Corolla.xlsx")
mylm <- lm(AskingPrice ~ EngineSize + YearModel, data = toyota)
summary(mylm)
```

```
##
## Call:
## lm(formula = theformula, data = toyota)
##
## Residuals:
```

```
##      Min      1Q Median      3Q      Max
## -52586 -14936      64    237 155569
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -42525782   3276015 -12.981  <2e-16 ***
## EngineSize    -8951     30812  -0.291    0.772
## YearModel     21172     1639   12.917  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 28820 on 191 degrees of freedom
## (2 observations deleted due to missingness)
## Multiple R-squared:  0.6287, Adjusted R-squared:  0.6248
## F-statistic: 161.7 on 2 and 191 DF,  p-value: < 2.2e-16
```

■ for importing data; ■ for calling `lm`; ■ for formula and data arguments; ■ for summary

- b. [4 marks] Predict the asking price of a Toyota Corolla with a Engine Size (l) value of 1.8 and a Year of Manufacture value of 2021. Include 95% prediction limits.

```
newdata <- data.frame(EngineSize = 1.8, YearModel = 2021)
predict(my1m, newdata = newdata, interval = "prediction", level = 0.95)
```

```
##      fit      lwr      upr
## 1 247585.9 190402.2 304769.7
```

■ for creating new data frame; ■ for calling `predict`; ■ for `interval = "prediction"`; ■ for correct level