APM1111 Statistical Theory

Summative Assessment 1

Submitted by:

Lopez, Ian Darylle

Pineda, Erianne

Sindayen, Lorenzo Danilo

Submitted to:

Sir Frederick Gella

Submitted on:

October 2025

Part I.

1. Provide descriptive statistical summaries of the entire data using JASP, Python or R.

R CODE:

```
male <- c(
        12, 4, 11, 13, 11, 7, 9, 10, 10, 7,
        7, 12, 6, 9, 15, 10, 11, 12, 7, 8,
        8, 9, 11, 10, 9, 10, 9, 9, 7, 9,
        11, 7, 10, 10, 11, 9, 12, 12, 8, 13,
        9, 10, 8, 11, 10, 13, 13, 9, 10, 13
)

female <- c(
        11, 9, 7, 10, 9, 10, 10, 7, 9, 10,
        11, 8, 9, 6, 11, 10, 7, 9, 12, 14,
        11, 12, 12, 8, 12, 12, 9, 10, 11, 7,
        12, 7, 9, 8, 11, 10, 8, 13, 8, 10,
        9, 9, 9, 11, 9, 9, 8, 9, 12, 11
)

both <- c(male, female)
cat("1. Descriptive statistical summaries of the entire data:", "\n")
Length <- round(length(both), 2)
Mean <- round(mean(both), 2)
Median <- round(median(both), 2)
Define_Mode <- table(both)
Mode <- names(sort(Define_Mode, decreasing = TRUE))[1]
```

```
Minimum <- round(min(both), 2)
Maximum <- round(max(both), 2)
Range <- round(Maximum - Minimum, 2) Variance <- round(var(both), 2)
Standard_Deviation <- round(sd(both), 2)

Summary_Both <- data.frame(
Statistics = c("Number of Observations:", "Mean:", "Median:", "Mode:", "Minimum:",
        "Maximum:", "Range:", "Variance:", "Standard Deviation:"),
Values = c(Length, Mean, Median, Mode, Minimum, Maximum, Range, Variance,
        Standard_Deviation)
)
cat("\n==========================================\n")
cat("Descriptive Statistics: Male and Female Students\n")
cat("==========================================\n")
print(Summary_Both, row.names = FALSE)
```

OUTPUT:

```
1. Descriptive statistical summaries of the entire data:

================================================
Descriptive Statistics: Male and Female Students
================================================
              Statistics Values
 Number of Observations:    100
                   Mean:   9.76
                 Median:     10
                   Mode:      9
                Minimum:      4
                Maximum:     15
                  Range:     11
               Variance:   3.86
     Standard Deviation:   1.96
```

---

2. Provide descriptive statistical summaries of the data for each gender category using JASP, Python or R.

R CODE:

```
cat("\n\n", "2. Descriptive statistical summaries of the data for each gender category:", "\n")


  Summary_MF <- function(x) {
        Length_MF <- round(length(x),2)
        Mean_MF <- round(mean(x), 2)
        Median_MF <- round(median(x), 2)
        Define_Mode_MF <- table(x)
        Mode_MF <- names(sort(Define_Mode_MF, decreasing = TRUE))[1]
```

```
    Minimum_MF <- round(min(x), 2)
    Maximum_MF <- round(max(x), 2)
    Range_MF <- round(Maximum_MF - Minimum_MF, 2)
    Variance_MF <- round(var(x), 2)
    Standard_Deviation_MF <- round(sd(x), 2)

  data.frame(
  Statistics = c("Number of Observations:", "Mean:", "Median:", "Mode:", "Minimum:",
      "Maximum:", "Range:", "Variance:", "Standard Deviation:"),
  Values = c(Length_MF, Mean_MF, Median_MF, Mode_MF, Minimum_MF, Maximum_MF,
      Range_MF, Variance_MF, Standard_Deviation_MF)
       )
}

male_summary <- Summary_MF(male)
female_summary <- Summary_MF(female)

cat("\n========================================\n")
cat("Descriptive Statistics: Male Students\n")
cat("========================================\n")
print(male_summary, row.names = FALSE)

cat("\n========================================\n")
cat("Descriptive Statistics: Female Students\n")
cat("========================================\n")
print(female_summary, row.names = FALSE)
```

OUTPUT:

```
 2. Descriptive statistical summaries of the data for each gender category:

======================================
Descriptive Statistics: Male Students
======================================
              Statistics Values
 Number of Observations:     50
                   Mean:   9.82
                 Median:     10
                   Mode:      9
                Minimum:      4
                Maximum:     15
                  Range:     11
               Variance:   4.64
     Standard Deviation:   2.15

======================================
Descriptive Statistics: Female Students
======================================
              Statistics Values
 Number of Observations:     50
                   Mean:    9.7
                 Median:    9.5
                   Mode:      9
                Minimum:      6
                Maximum:     14
                  Range:      8
               Variance:   3.15
     Standard Deviation:   1.78
```

3. Make a report based on the statistical summaries, including the results for both the combined data (not split by gender) and the gender-specific categories.

R CODE:

```
cat("\n\n", "3. Make a report based on the statistical summaries, including the results for both the combined data (not split by gender) and the gender-specific categories:", "\n\n")


cat("- Inferring from the statistical summaries, the data shows varying patterns between male and female students in the prospect of using their phones. On average, the students have a total usage rate of 9.76 hours per week, meaning that it takes up most of their time in their daily lives. Even considering the aspect of gender, the results don't vary much (9.82 hours for males and 9.7 hours for females). The variation for both typically ranges from 4 hours to 16 hours with the cluster mostly being on 9. From the data, we can see that both groups of students are engaged deeply in mobile device usage, and neither groups are well-regulated with their discipline of phone usage.\n")
```

OUTPUT:

```
 3. Make a report based on the statistical summaries, including the results for
    both the combined data (not split by gender) and the gender-specific categories:

- Inferring from the statistical summaries, the data shows varying patterns between
male and female students in the prospect of using their phones. On average, the students
have a total usage rate of 9.76 hours per week, meaning that it takes up most of their
time in their daily lives. Even considering the aspect of gender, the results don't vary
much (9.82 hours for males and 9.7 hours for females). The hours spent for both genders
typically ranges from 4 hours to 15 hours with the cluster mostly being on 9. From the
data, we can see that both groups of students are engaged deeply in mobile device usage,
and neither groups are well-regulated with their discipline of phone usage.
```

Part II.

1. Find the (a) first, (b) second, (c) third, and (d) fourth moments of the set 2, 3, 7, 8, 10.

R CODE:

```
x <- c(2, 3, 7, 8, 10)
n <- length(x)

m1_prime <- mean(x)
m2_prime <- mean(x^2)
m3_prime <- mean(x^3)
m4_prime <- mean(x^4)

cat("1.\n")
cat("(a) m1' =", m1_prime, "\n")
cat("(b) m2' =", m2_prime, "\n")
cat("(c) m3' =", m3_prime, "\n")
cat("(d) m4' =", m4_prime, "\n\n")
```

OUTPUT:

```
1.
(a) m1' = 6
(b) m2' = 45.2
(c) m3' = 378
(d) m4' = 3318.8
```

---

2. Find the (a) first, (b) second, (c) third, and (d) fourth moments about the mean of the set 2, 3, 7, 8, 10.

R CODE:

```
mean_x <- m1_prime
y <- x - mean_x

m1 <- mean(y)
m2 <- mean(y^2)
m3 <- mean(y^3)
m4 <- mean(y^4)

cat("2.\n")
cat("(a) m1 =", m1, "\n")
cat("(b) m2 =", m2, "\n")
cat("(c) m3 =", m3, "\n")
cat("(d) m4 =", m4, "\n\n")
```

OUTPUT:

```
2.
(a) m1 = 0
(b) m2 = 9.2
(c) m3 = -3.6
(d) m4 = 122
```

---

3. Verify: $m = m'_4 - 4m'_1 m'_3 + 6m'^2_1 m'_2 - 3m'^4_1$ for the set 2, 3, 7, 8, 10.

R CODE:

```
formula <- m4_prime - 4*m1_prime*m3_prime + 6*(m1_prime^2)*m2_prime -3*(m1_prime^4)
cat("3.\n")
cat("Theoretical: m4 =", formula, "\n")
cat("Calculated: m4 =", m4, "\n")
```

OUTPUT:

```
3.
Theoretical: m4 = 122
Calculated: m4 = 122
```

Part III.

Prove that $m_4' = m_4 + 4hm_3 + 6h^2m_2 + h^4, \quad$ where $h = m_1'$.

Proof:
The raw moment about the origin is defined as:

$$m_r' = \frac{1}{n}\sum x_i^r,$$

Also, the central moment about the mean is defined as:

$$m_r = \frac{1}{n}\sum (x_i - \bar{x})^r.$$

Expand $(x_i - \bar{x})^4$ using the binomial theorem:

$$(x_i - \bar{x})^4 = x_i^4 - 4x_i^3\bar{x} + 6x_i^2\bar{x}^2 - 4x_i\bar{x}^3 + \bar{x}^4$$

Taking the mean of both sides:

$$m_4 = \frac{1}{n}\sum (x_i - \bar{x})^4 = m_4' - 4\bar{x}m_3' + 6\bar{x}^2m_2' - 4\bar{x}^3m_1' + \bar{x}^4$$

Since $\bar{x} = h = m_1'$, substitute this into the equation:

$$m_4 = m_4' - 4hm_3' + 6h^2m_2' - 4h^3m_1' + h^4$$

Substitute the relationships between raw and central moments for lower orders:

$$\begin{aligned} m_3' &= m_3 + 3hm_2 + h^3, \\ m_2' &= m_2 + h^2, \\ m_1' &= h. \end{aligned}$$

Substitute these into the expression for m_4':

$$\begin{aligned} m_4' &= m_4 + 4h(m_3 + 3hm_2 + h^3) - 6h^2(m_2 + h^2) + 4h^3(h) - h^4 \\ &= m_4 + 4hm_3 + 12h^2m_2 + 4h^4 - 6h^2m_2 - 6h^4 + 4h^4 - h^4 \end{aligned}$$

Combine like terms:

$$\begin{aligned} m_4' &= m_4 + 4hm_3 + (12h^2m_2 - 6h^2m_2) + (4h^4 - 6h^4 + 4h^4 - h^4) \\ &= m_4 + 4hm_3 + 6h^2m_2 + h^4 \end{aligned}$$

Therefore,

$$\boxed{m_4' = m_4 + 4hm_3 + 6h^2m_2 + h^4}$$