Assignment-01

AUTHOR
Yifan Meng

PUBLISHED
February 16, 2025

1 Question 1 [12pts]

A performance test of manual dexterity was administered to a group of subjects. The frequency distribution is reported as follows.

Score	50	60	70	80	90
Frequency	1	2	3	3	1

1.1 Part 1 [3pts]

1.1.1 Question

Compute the mean, variance, and standard deviation of the score distribution.

Hint: You can use mean(), var(), and sd() functions in R.

1.1.2 Answer

```
# Create the data
scores <- rep(c(50, 60, 70, 80, 90), times=c(1, 2, 3, 3, 1))

# Calculate statistics
mean(scores)
```

```
[1] 71
```

```
var(scores)
```

```
[1] 143.3333
```

```
sd(scores)
```

```
[1] 11.97219
```

```
sd(scores)/sqrt(length(scores))
```

```
[1] 3.785939
```

```
## Values of the answer below is calculated by function round(), e.g.: round(sd(scores),2)
```

> **Report**
>
> - The mean of this score distribution is 71

- The variance of the score distribution is 143.33
- The standard deviation of the score distribution is 11.97

1.2 Part 2 [9pts]

1.2.1 Question

The following table contains raw scores, z-scores, and T-scores for Joan and Peter from the group. Complete the table by finding out the unknown values (a), (b), and (c).

Name	Raw score	z-score	T-score
Joan	80.00	0.75	57.50
Peter	49.45	-1.80	32.00

1.2.2 Answer

Solution Process

According to **Part 1**, μ estimated by the distribution is 71 and σ is 11.97.

1.2.2.1 (a) Joan's z-score:

$$z = \frac{X - \mu}{\sigma}$$
$$z = \frac{80 - 71}{11.97} = \frac{9}{11.97} \approx 0.75$$

1.2.2.2 (b): Peter's raw score

$$X = z \cdot \sigma + \mu$$

Substituting the values for Peter's z-score:

$$X = (-1.80) \cdot 11.97 + 71$$
$$X = -21.53 + 71 = 49.47$$

1.2.2.3 (c): Peter's T-score:

The formula of T score is :

$$T = 10 \cdot z + 50$$

Calculate the Joan's T score by z-score

$$T = 10 \cdot (-1.80) + 50 = 32$$

2 Question 2 [11pts]

James is interested in finding the relationship between Introversion and Social Dominance. He collected the following data and calculated the descriptive statistics of the data.

Examinee	1	2	3	4	5	6	7	8	Mean	SD
Introversion	0	1	1	2	2	3	4	5	2.25	1.67
Social Dominance	4	5	4	2	3	3	2	1	3.00	1.31

2.1 Part 1 [3pts]

2.1.1 Question

Obtain a scatter plot of the two variables by using plot() function in R. Please put Introversion on the x-axis and Social Dominance on the y-axis.

Hint: You can construct the dataset in .csv or .txt format and import the data OR you can also create data objects by running the codes below:

2.1.2 Answer

```
Introv <- c(0, 1, 1, 2, 2, 3, 4, 5)
Socialdom <- c(4, 5, 4, 2, 3, 3, 2, 1)
#df <- data.frame(Introv, Socialdom)

plot(Introv,Socialdom,main = "Scatter Plot of Introversion and Social Dominance",
     xlab = "Introversion", ylab = "Social Dominance")
```

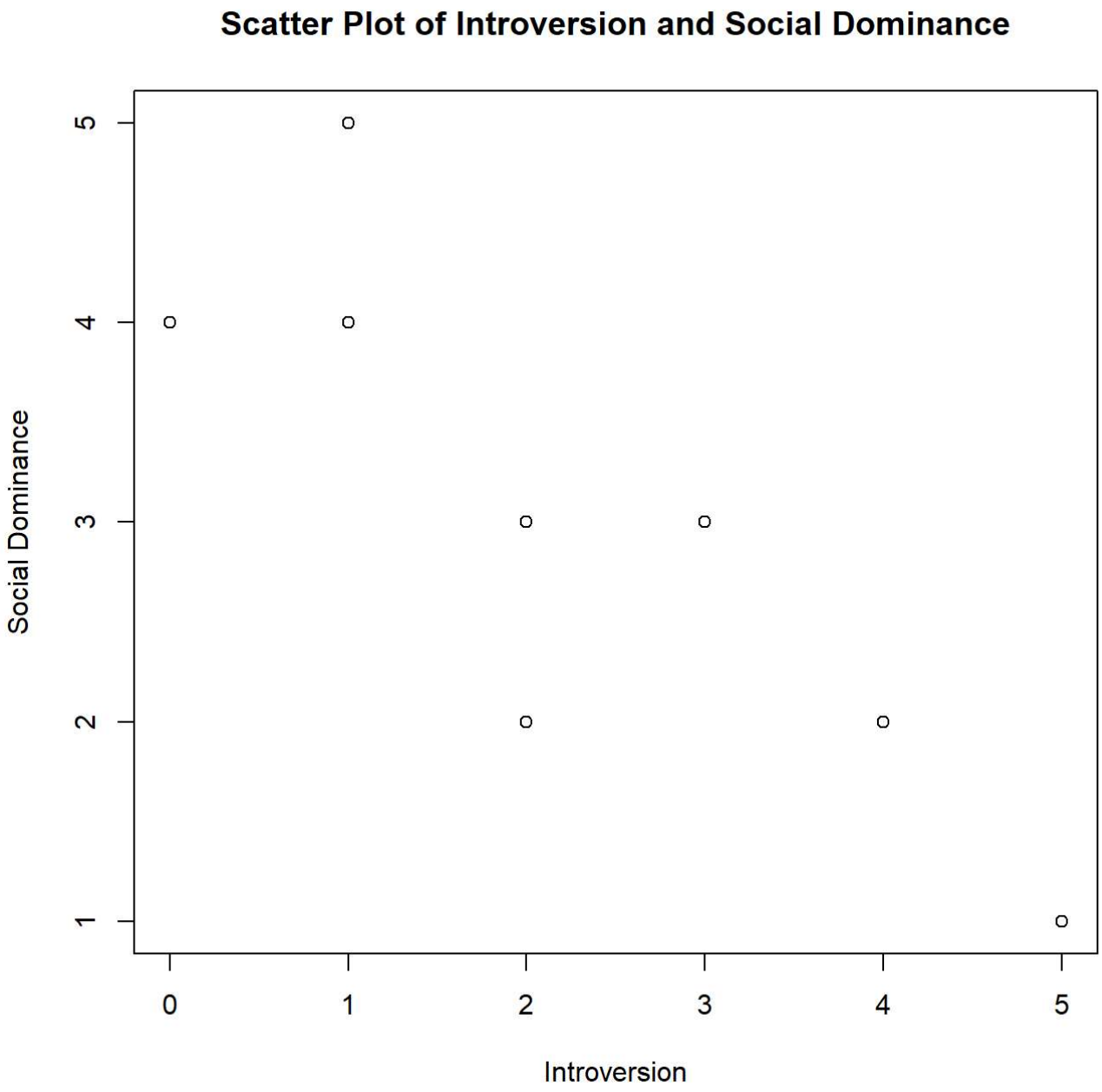

Figure 1

2.2 Part 2 [6pts]

2.2.1 Question

Based on the scatter plot you obtained in 2-1), how would you describe the: i) trend ii) strength iii) direction of the association between the two variables? Select your response from the following options for each element:

a. Positive b) Negative
b. Linear d) Non-linear
c. Strong f) Weak

2.2.2 Answer

iii. direction: b) Negative

iv. strength: e) Strong
v. trend: c) Linear

> **Details**
>
> - **Direction (Negative)**: As Introversion increases, Social Dominance tends to decrease.
> - **Strength (Strong)**: The points are closely clustered around a downward trend, indicating a strong relationship (also can be supported by the correlation coefficient below).
> - **Trend (Linear)**: The pattern follows a straight line rather than a curve.

2.3 Part 3 [2pts]

2.3.1 Question

Compute the Pearson correlation coefficient between Introversion and Social Dominance using cor() function in R.

2.3.2 Answer

```
cor(Introv,Socialdom)
```

```
[1] -0.8498366
```

The Pearson correlation between Introversion and Social Dominance is -0.85

3 Question 3 [8pts]

Determine the scale of measurement for each of the following variables.

1. [2pts] Ranking of baseball teams (determined by their won-lost records at the end of the season): **Ordinal Scale**
2. [2pts] Social Security numbers: **Nominal Scale**
3. [2pts] Air distance between New York City and other cities in the United States: **Interval Scale**
4. [2pts] pH value: **Ratio Scale**

4 Question 4 [9pts]

The scores of a reference population on the Wechsler Intelligence Scale for Children (WISC) are normally distributed with $\mu = 100$ and $\sigma = 15$.

4.0.1 Question 1 [3pts]

Approximately what percentage of the population would score below 100 on the test? **50%**

4.0.2 Answer 1

Solution

The WISC scores follow a normal distribution:

$$X \sim \mathcal{N}(100, 15^2)$$

Here we want to estimate

$$P(X < 100)$$

Because μ=100, so z-score of 100 is 0.

From the standard normal table, the probability of a z-score being less than 0 is 50%

Or directly calculate using R

```
pnorm(100, mean = 100, sd = 15)
```

```
[1] 0.5
```

4.0.3 Question 2 [3pts]

Approximately what percentage of the population would score above 130 on the test?

4.0.4 Answer 2

The z score of 130 in this test is 2, so there will be approximately 2.5% of the population would score above 130 on the test.

Solution

Here we want to estimate:

$$P(X > 130)$$

Then, calculate the z-score for 130:

$$z = \frac{130 - 100}{15}$$

$$z = \frac{30}{15} = 2$$

Based on the standard norm table:

$$P(Z > 2) = 1 - P(Z < 2) \approx 1 - 0.975 \approx 0.025$$

For z = 2:

- Area above z = 2 ≈ 2.5%
- Therefore, approximately **2.5%** of the population would score above 130

Or directly calculate using R

```r
1-pnorm(130,mean=100,sd=15)
```

```
[1] 0.02275013
```

```r
#or
pnorm(130,mean=100,sd=15,lower.tail = FALSE)
```

```
[1] 0.02275013
```

4.0.5 Question 3 [3pts]

Approximately what percentage of the population would score between 100 and 130 on the test? **47.5%**

4.0.6 Answer 3

Solution

Here we want to calculate:

$$P(100 < X < 130)$$

From 2 subquestions above,

$$P(X < 130) \approx 0.975$$
$$P(X > 100) = 0.5$$

So:

$$P(100 < X < 130) = P(X < 130) - P(X > 100) = 0.975 - 0.5 = 47.5$$

In R:

```r
pnorm(130, mean = 100, sd = 15) - pnorm(100, mean = 100, sd = 15)
```

```
[1] 0.4772499
```

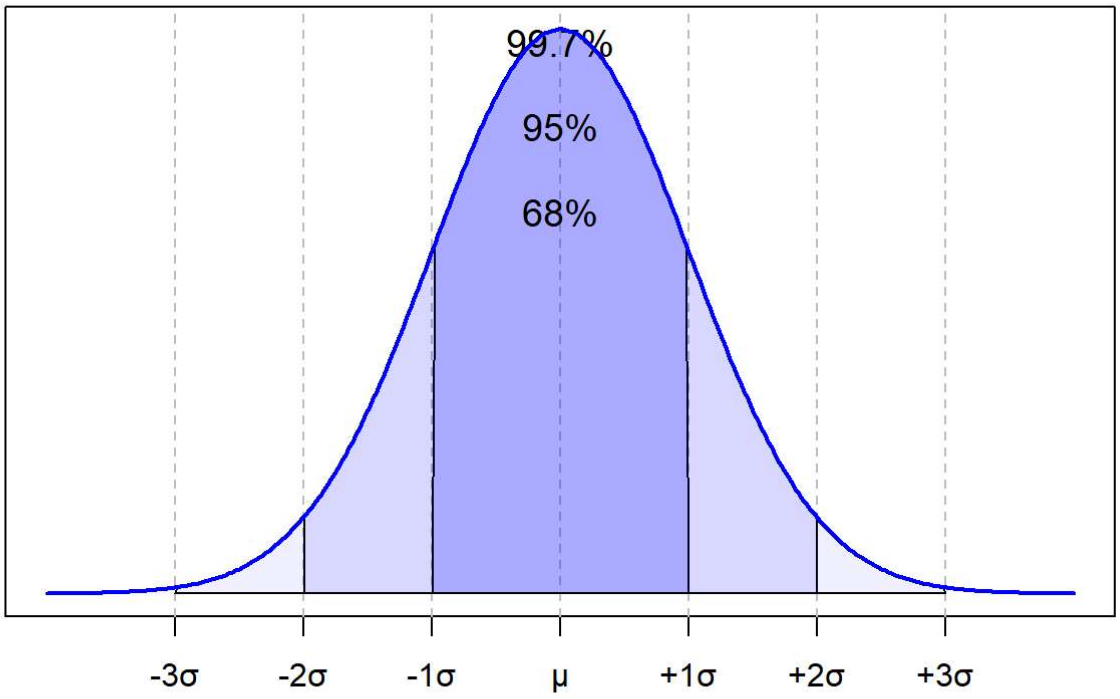

Figure 2