Ketaki Mahajan
A – 2 / 16014022050
Tutorial 6: Sampling using R (15 / 03 / 2024)

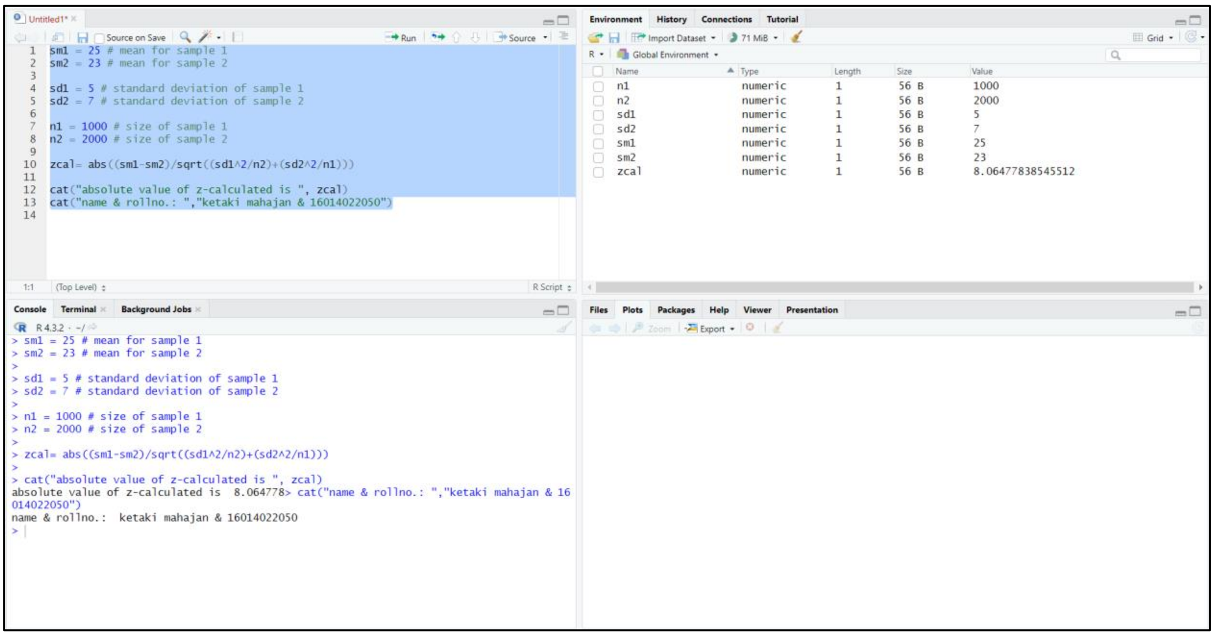
1. **Test the significance of the difference between the means of two normal population with the same standard deviation from the following data –**

	Size	Mean	Standard Deviation
Sample 1	1000	25	5
Sample 2	2000	23	7

Code –

```
sm1 = 25 # mean for sample 1
sm2 = 23 # mean for sample 2

sd1 = 5 # standard deviation of sample 1
sd2 = 7 # standard deviation of sample 2

n1 = 1000 # size of sample 1
n2 = 2000 # size of sample 2

zcal= abs((sm1-sm2)/sqrt((sd1^2/n2)+(sd2^2/n1)))

cat("absolute value of z-calculated is ", zcal)
cat("name & rollno.: ","ketaki mahajan & 16014022050")
```

R-studio Output –



Steps of Hypothesis Testing –

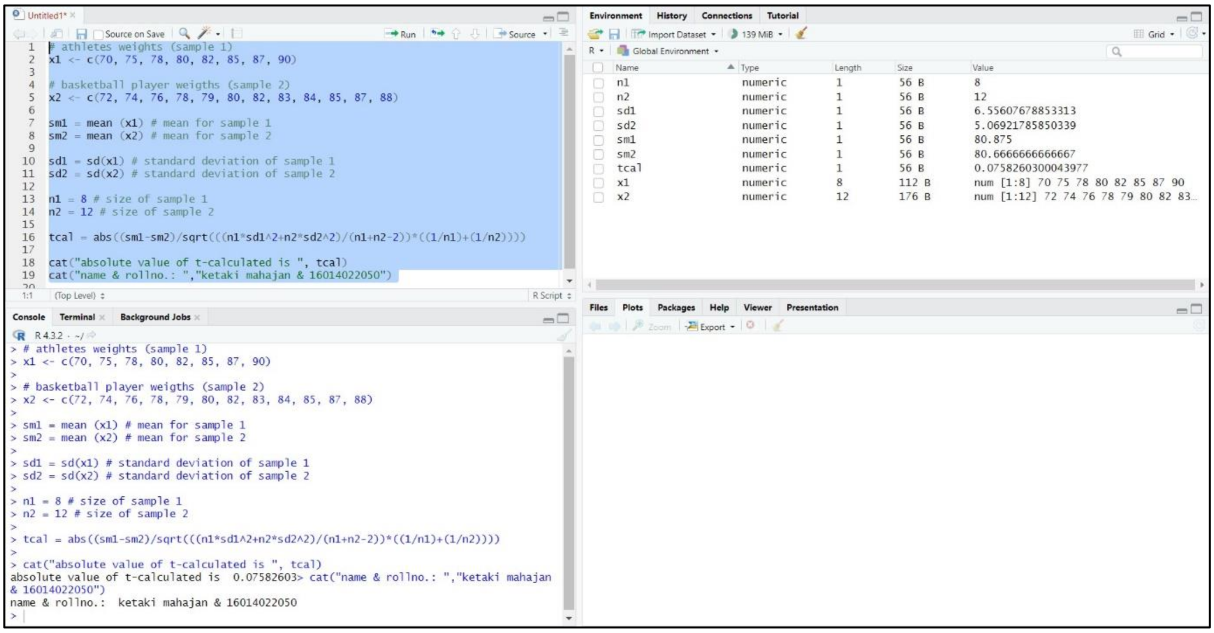
- **Ho :** $\mu_1 = \mu_1$
- **Ha :** $\mu_1 \neq \mu_1$ (Nature of the test is two tailed)
- LOS is 5%
- Table value of Z_α is: 1.96
- Calculated value of Z: $Z_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_2} + \frac{s_2^2}{n_1}}} = 8.064778$
- Since, $Z_{cal} > Z_\alpha$, we **reject** the null hypothesis.
- Hence, to conclude, we can say that the **difference between the population means is significant**.

2. **The weights of eight randomly selected athletes are recorded in kilograms:**
   70, 75, 78, 80, 82, 85, 87, 90.
   The weights of twelve randomly selected basketball players are recorded in kilograms:
   72, 74, 76, 78, 79, 80, 82, 83, 84, 85, 87, 88.
   Can it be concluded that basketball players, on average, weigh more than athletes?

   Code –

```
# athletes weights (sample 1)
x1 <- c(70, 75, 78, 80, 82, 85, 87, 90)

# basketball player weigths (sample 2)
x2 <- c(72, 74, 76, 78, 79, 80, 82, 83, 84, 85, 87, 88)

sm1 = mean (x1) # mean for sample 1
sm2 = mean (x2) # mean for sample 2

sd1 = sd(x1) # standard deviation of sample 1
sd2 = sd(x2) # standard deviation of sample 2

n1 = 8 # size of sample 1
n2 = 12 # size of sample 2

tcal = abs((sm1-sm2)/sqrt(((n1*sd1^2+n2*sd2^2)/(n1+n2-
2))*((1/n1)+(1/n2))))

cat("absolute value of t-calculated is ", tcal)
cat("name & rollno.: ","ketaki mahajan & 16014022050")
```

R-Studio Output –



Steps of Hypothesis Testing –

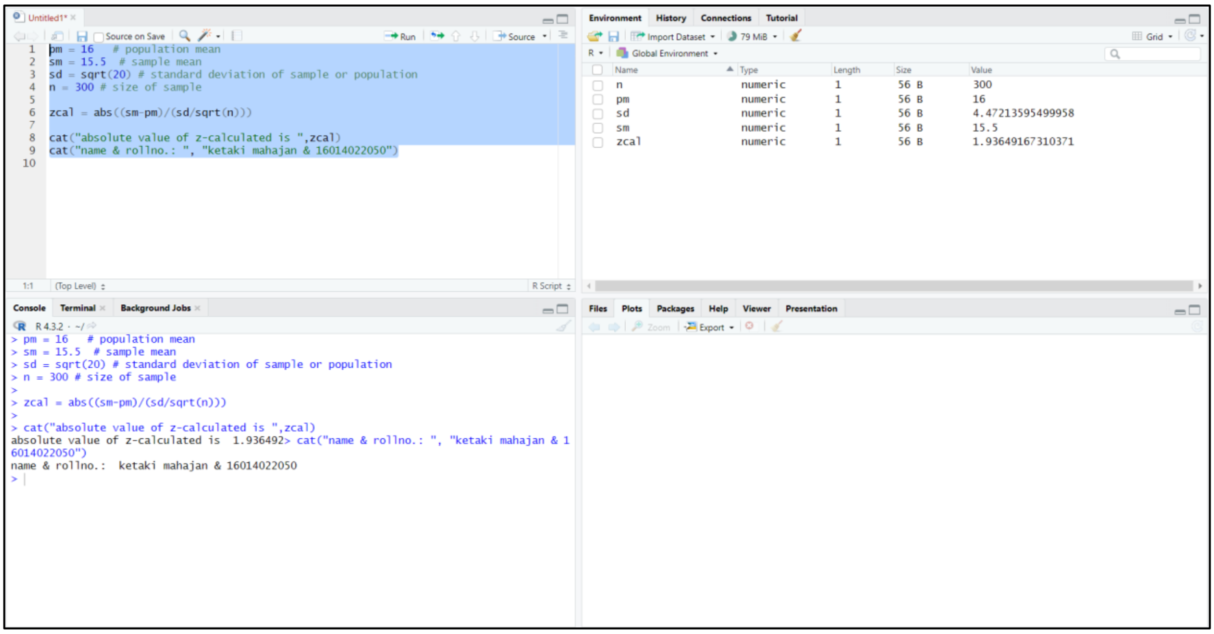
- **H₀ :** $\mu_1 = \mu_2$
- **Hₐ :** $\mu_1 < \mu_2$ (Nature of the test is one tailed)
- LOS is assumed as 5%
- DOF = 8 + 12 – 2 = 18
- Table value of t_α is: 1.7341
- Calculated value of t: $t_{cal} = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}} \times \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ = 0.07582603
- Since, $t_{cal} < t_\alpha$, we **accept** the null hypothesis.
- Hence, to conclude, there is **no evidence** that says basketball players, on average, weigh more than athletes.

3. A random sample of 300 observations has a mean of 15.5 kg. Can it be a random sample from a population whose mean is 16 kg and variance are 20 kg?

<u>Code</u> –

```
pm = 16   # population mean
sm = 15.5  # sample mean
sd = sqrt(20) # standard deviation of sample or population
n = 300 # size of sample

zcal = abs((sm-pm)/(sd/sqrt(n)))

cat("absolute value of z-calculated is ",zcal)
cat("name & rollno.: ", "ketaki mahajan & 16014022050")
```

<u>R-Studio Output</u> –



<u>Steps of Hypothesis Testing</u> –

- **H₀ :** $\mu = 16$
- **Hₐ :** $\mu \neq 16$ (Nature of the test is two tailed)
- LOS is 5%
- Table value of Z_α is: 1.96
- Calculated value of Z: $Z_{cal} = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = 1.936492$
- Since $Z_{cal} < Z_\alpha$, we **<u>accept</u>** the null hypothesis.
- Hence, to conclude, we can say that the **<u>sample is drawn from a population</u>** with mean 16.