

NATIONAL INSTITUTE OF TECHNOLOGY SILCHAR

Cachar, Assam

B.Tech. IVth Sem

Subject Code: CS216

Subject Name: Applied Probability

Submitted By:

Name : Subhojit Ghimire

Sch. Id. : 1912160

Branch : CSE – B

2. Solve the following problems using R programming:

- a. The Stat students of a reputed college have the following test scores: 58, 64, 93, 71, 67, 85, 71, 93, 82, 81, 75, 78, 86, 90 and 87. Determine and interpret the Quartiles of these scores.



The screenshot shows the RStudio interface. In the Console, the following commands and output are visible:

```
> scores <- c(58, 64, 93, 71, 67, 85, 71, 93, 82, 81, 75, 78, 86, 90, 87)
> summary(scores)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  58.00  71.00   81.00   78.73  86.50   93.00
> quantile(scores)
 0% 25% 50% 75% 100%
58.0 71.0 81.0 86.5 93.0
> |
```

The Environment pane on the right shows the variable 'scores' as a numeric vector of length 15, with the first few values displayed: 58, 64, 93, 71, 67, 85, 71, 93.

Interpretation: $Q_1 = 25\%$ implies that 25% of the test scores fall below or equal to 71.0, while the other 75% of it is above 71.0. $Q_2 = 50\%$ is the median, i.e., half the scores are below or equal to 81.0, while the other half scores are more than 81.0. $Q_3 = 75\%$ implies that three-fourth of the data are below or equal to 86.5, while the remaining one-fourth is above 86.5. The minimum and maximum values are 58.0 and 93.0, respectively.

- b. The surveyed weights (in kilogram) of the students in Stat dept. were the following: 50, 65, 72, 62, 77, 84, 57, 74, 66, 68, 75, 58, 52, 69, and 87. Compute and interpret the Deciles of these weights.



```

> weights <- c(50, 65, 72, 62, 77, 84, 57, 74, 66, 68, 75, 58, 52, 69, 87)
> summary(weights)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  50.00  60.00   68.00  67.73  74.50   87.00
> quantile(weights, prob = seq(0, 1, length = 11), type = 5)
 0%  10%  20%  30%  40%  50%  60%  70%  80%  90% 100%
50.0 52.0 57.5 62.0 65.5 68.0 70.5 74.0 76.0 84.0 87.0
>

```

The screenshot shows the RStudio interface. The console on the left displays the R code and its output. The code defines a vector 'weights' with 15 values, computes a summary, and then uses the 'quantile' function with 'prob = seq(0, 1, length = 11)' and 'type = 5' to calculate deciles. The output shows the deciles from 0% to 100% in increments of 10%. The Environment pane on the right shows the 'weights' variable as a numeric vector of length 15.

Interpretation: The first decile $D_1 = 10\%$ implies that one-tenth of the weights fall below or equal to 52.0, and the remaining nine-tenth fall above 52.0. The $D_5 = 50\%$ is the median, thus half of the students' weights weigh below or equal to 68.0, while the other half fall above this. The $D_9 = 90\%$ implies that nine-tenth of the weights fall below or equal to 84.0, while the other one-tenth fall above 84.0. The minimum and maximum values are 50.0 and 87.0, respectively.

To Note: There is a slight difference between the codes of quartiles and deciles computations. The appropriate algorithm for decile is argument type 5, which is set as seen, while the quartile is type 7, which is the default value for quantile function. In addition, the prob argument is used in decile, which is the position to be measured. Since, deciles divide the data points into ten parts, the sequence function seq is used to determine prob's value that is from 0 to 1, the length forming to be 11.

- c. Compute the power of a study to show a difference between group 1 ($n=28$) in which the event probability is 30% and group 2 ($n=28$) in which the event probability is 55%.



```
> power.prop.test (n = 28, p1 = 0.3, p2 = 0.55)

Two-sample comparison of proportions power calculation

      n = 28
     p1 = 0.3
     p2 = 0.55
sig.level = 0.05
  power = 0.4720963
alternative = two.sided

NOTE: n is number in *each* group

> |
```

Interpretation: The power of a study which includes 28 subjects in each of the two experimental groups to see a difference between the event probabilities is 48% under the assumption that the event probabilities are 30% in group 1 and 55% in group 2.

- d. Compute the sample size of a study to show a difference between group 1 in which the event probability is 25% and group 2 in which the event probability is 45% with a power of 65%.



```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Project: (None)
Environment History Connections Tutorial
R Global Environment
List
Environment is empty
Files Plots Packages Help Viewer
Zoom Export
```

```
> power.prop.test(power = 0.65, p1 = 0.25, p2 = 0.45)

Two-sample comparison of proportions power calculation

      n = 62.1105
      p1 = 0.25
      p2 = 0.45
sig.level = 0.05
power = 0.65
alternative = two.sided

NOTE: n is number in *each* group

> |
```

Interpretation: In order to achieve a power of 65% under the assumed probabilities, the study should include at least 63 subjects in each of the experimental groups.