| QUARTILES, DECILES, AND PERCENTILES |
| R Programming |

Al-Ahmadgaid B. Asaad
alstated@gmail.com

The measures of position such as quartiles, deciles, and percentiles are available in the `quantile` function. This function has a usage,

```
quantile(x, probs = seq(0, 1, 0.25), na.rm = FALSE,
         names = TRUE, type = 7, ...)
```

where:

- `x` - the data points;

- `prob` - the location to measure;

- `na.rm` - if `FALSE`, `NA` (Not Available) data points are not ignored;

- `names` - for attributes, FALSE means no attributes, hence speeds-up the computation;

- `type` - type of the quantile algorithms; and,

- `...` - further arguments.

Example 1. The junior BS Stat students of MSU-IIT have the following SASE scores: 88, 84, 83, 80, 94, 90, 81, 79, 79, 81, 85, 87, 86, 89, and 92. Determine and interpret the quartiles of these scores.

```
> scores <- c(88, 84, 83, 80, 94, 90, 81,
+             79, 79, 81, 85, 87, 86, 89, 92)
> quantile(scores)
  0%   25%   50%   75%  100%
79.0  81.0  85.0  88.5  94.0
```

Interpretation: Therefore, $Q_1 = 25\%$ implies that, 25% of the SASE scores fall below or equal to 81.0, while the other 75% of it is above 81.0. $Q_2 = 50\%$ is the median, and thus half of the scores are below or equal to 85.0, while the other half, are above 85.0. $Q_3 = 75\%$, implies that three-fourth of the data are below or equal to 88.5, while the remaining one-fourth are above 88.5. And the minimum and maximum values are 79.0 and 94.0, respectively.

Example 2. The surveyed weights (in kilograms) of the students in Stat 131 were the following: 69, 70, 75, 66, 83, 88, 66, 63, 61, 68, 73, 57, 52, 58, and 77. Compute and interpret the deciles of these weights.

```
> weights <- c(69, 70, 75, 66, 83, 88, 66,
+               63, 61, 68, 73, 57, 52, 58, 77)
> quantile(weights, prob = seq(0, 1, length = 11), type = 5)
  0%  10%  20%  30%  40%  50%  60%  70%  80%  90% 100%
52.0 57.0 59.5 63.0 66.0 68.0 69.5 73.0 76.0 83.0 88.0
```

Notice the difference between the codes of quartiles and deciles? This time the function `quantile` has an argument `type` which is set to 5. With this, the quantile algorithm between the quartiles and deciles differ. Hence, the appropriate algorithm for decile is type 5, while the quartile is type 7, which is the default one. For further reading about the quantile algorithm run `?quantile`. In addition, the `prob` argument above is the position to be measured, and since deciles divide the data points into ten parts, then the sequence function, `seq`, is used for `prob`'s value that is from 0 to 1 of length 11 (`length = 11`, 11 because zero is included, which is the minimum of the data points).

Interpretation: The first decile is $D_1 = 10\%$, implies that one-tenth of the weights fall below or equal to 57.0, and the remaining nine-tenth fall above 57.0. 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 it. And so on.

Example 3. Compute the 15th, 25th, and 35th percentiles of weights in **Example 2**.

```
> quantile(datad, prob = c(0.15, 0.25, 0.35))
 15%  25%  35%
58.3 62.0 65.7
```

Interpretation: The fifteenth percentile $P_{15} = 15\%$ is interpreted as 15% of the samples fall below or equal to 58.3 while 85% fall above 58.3. The thirty-fifth percentile $P_{35} = 35\%$, implies that 35% of the weights fall below or equal to 65.7, and that 65% of it fall above 65.7.

References

Yau, Chi. R Tutorial: Percentile.

Labels

Descriptive Statistics, R, Tutorial