Median and Interquartile Range -- Nonparametric Univariate Statistics for Quantitative Variables

Application: To obtain a summary of the distribution of scores (center and spread) for a variable when the data are not normally distributed, are not measured on an interval scale, and/or there is only a small sample. Thus, they are used in the same data situation as the mean and standard deviation, except that they are used when the data are either importantly non-normally distributed, the measurement scale of the dependent variable is ordinal (not interval or ratio), or from a too-small sample.

The **median** is the value that is in the "middle" of the distribution, with 50% of the scores having a value larger than the median, and 50% of the scores having a value smaller than the median. It is important to keep in mind the difference between this definition of "middle" and that used when describing the mean.

The **interquartile range** (**IQR**) is the range of values within which reside the middle 50% of the scores. The lower bound of the interquartile range is called the first quartile (Q1) -- 25% of the scores have a value lower than Q1 and 75% of the scores have a value larger than Q1. The upper bound of the interquartile range is called the third quartile (Q3) -- 75% of the scores have a value lower than Q3 and 25% of the scores have a value larger than Q3. (By the way -- Q2 is the median.)

The data: The researcher is interested in describing the data about the number of fish displayed in the Pet stores. Specifically, the researcher wants to tell what is the typical number of fish displayed. The data from the 12 stores are displayed below.

```
32    41    31    38    21    13    17    22    24    11    17    20
```

Step 1 Arrange the data for analysis, putting the scores in order from smallest to largest values.

```
11    13    17    17    20    21    22    24    31    32    38    41
```

Step 2 Assign the depth of each score -- the rank of that score from the nearest extreme value.

```
depth =    1    2    3    4    5    6    6    5    4    3    2    1

value =    11   13   17   17   20   21   22   24   31   32   38   41
```

Computing the Median

Step 3 Determine the sample size.

```
N = 12
```

Step 4 Compute the depth of the median using the formula.

```
Depth of the median =   (N + 1) / 2    =    (12+1) / 2   =   13 / 2    =   6.5
```

Step 5 Determine the value of the median.

- IF the depth of the median is a whole number, then it will identify that score which has the median value.

- IF the depth of the median is a fraction (it can only be ?.5), then the value of the median is the average of the values that have that depth.

```
For these data, the depth of the median is 6.5, so we would compute the median as
the average of the two scores with depths of 6 (21 and 22).
```

```
Median =   (21 + 22) / 2   =   21.5
```

Computing the IQR

Step 6 Compute the depth of Q1 and Q3 using the formula.

```
depth of Q1 & Q3   =   N / 4  = 12 / 4   =   3
```

Step 7 Determine the value of Q1 and Q3

-   IF the depth of the quartiles is a whole number, then it will identify the scores which have the values of Q1 and Q3.

    ```
    For these data with a quartile depth of 3, Q1 = 17 and Q3 = 32
    ```

-   IF the quartile depth is a fraction then you must interpolate to find the values of Q1.

    ```
    For Example, if the quartile depth had been 3.25, we would need the value
    that is 1/4 of the way between the values of the scores at depths 3 and 4.
    ```

    ```
    For Q1 -- This is simple for this example, because the value of the scores at
    the lower depths 3 and 4 are both 17, so Q1 would be 17.
    ```

    ```
    For Q3 -- The computation would be, 31 + [.25 * (32-31)] = 31.25.
    ```

Step 8 Compute the IQR as the difference between Q3 and Q1

```
           IQR = Q3 - Q1  =  32  -  17  =  15
```

Step 9 Reporting the results.

You should report the univariate statistics for the involved variables before reporting the results of any significance testing. Usually one reports the value of the median and the IQR. One limitation of reporting the IQR as a value is that the IQR might be either symmetrical or asymmetrical around the median. Consider the data in the example. Q1 (17) is much closer to the median (21.5) than is Q3 (32), however this is not conveyed by reporting that IQR = 15. For this reason, it is more useful to report the IQR as a range (reporting Q1 and Q3), rather than as a value.

```
    The 12 stores in this sample displayed a Mdn of 21.5 fish (IQR = 17-32).
```