

Answer on Question #52567 – Math – Statistics and Probability

A quality audit of a bottling company showed the number of defective batches outside of the specification limits for a particular oil product. The grouped data shown in the table below indicated the number of b=defective items per unit of measurement:

Volume (ml) f

20 - < 25 1

25 - < 30 2

30 - < 35 8

35 - < 40 5

40 - < 45 3

45 - < 50 1

3.1 Complete the table under the given columns: (4)

CLASS LIMITS f CUM < f CLASS MARK (X) X f X - X̄ (X - X̄)² (X - X̄)² f

3.2 Calculate the mean value. (2)

3.3 Calculate the standard deviation. (3)

3.4 Draw the less-than ogive graph.

Use your graph to determine the median value. (5)

3.5 Calculate the 80th percentile value. (3)

3.6 Use the values from 3.2 and 3.4 to comment on the nature if the distribution.

Give a reason for your answer.

Solution

3.1. We start with the frequency histogram. First, we find the midpoint of each interval.

Midpoint of interval = $\frac{1}{2}$ (Lower class limit + Upper class limit).

Then we get the sum of all the frequencies (f) and the sum of all the fx.

We obtained the following Table

Volume (ml)	f	Relative Frequency	Midpoints (x)	f*x	(x - \bar{x})	(x - \bar{x}) ²	(x - \bar{x}) ² · f
20 - < 25	1	0.05	22.5	22.5	-12.5	156.25	156.25
25 - < 30	2	0.1	27.5	55	-7.5	56.25	112.5
30 - < 35	8	0.4	32.5	260	-2.5	6.25	50
35 - < 40	5	0.25	37.5	187.5	2.5	6.25	31.25
40 - < 45	3	0.15	42.5	127.5	7.5	56.25	168.75
45 - < 50	1	0.05	47.5	47.5	12.5	156.25	156.25
Sum	20	1	210	700	0	437.5	675

3.2. Next, we determine the mean by dividing the sum of fx by sum of f to get the mean.

$$\text{Mean} = \frac{\text{Sum of } (fx)}{\text{Sum of } f} = \frac{700}{20} = 35$$

3.3. To calculate the standard deviation

$$S = \sqrt{\frac{\sum(x - \bar{x})^2 f}{n}}$$

we substitute the given values into the formula noted above:

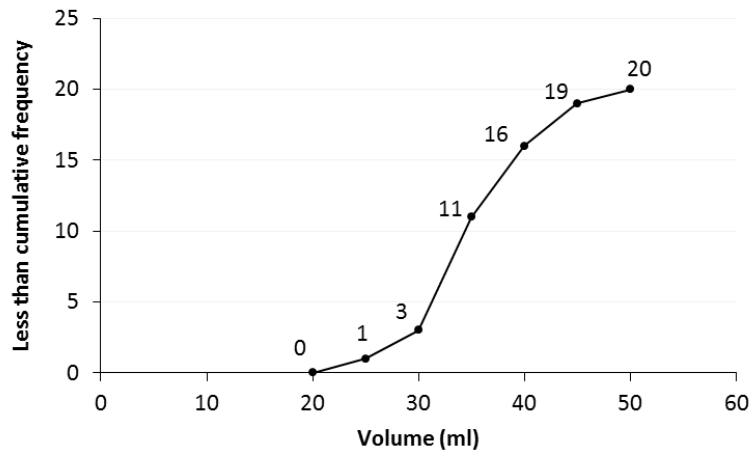
$$S = \sqrt{\frac{675}{20}} = \sqrt{33.75} = 5.809$$

During calculations, when a variable is grouped by class intervals, the midpoint of the interval is used in place of every other value in the interval. Thus, the spread of observations within each interval is ignored. This makes the standard deviation always less than the true value. It should, therefore, be regarded as an approximation.

3.4. We create the following Table

Volume (ml)	f	Less than cumulative frequency
20 - 25	1	1
25 - 30	2	3
30 - 35	8	11
35 - 40	5	16
40 - 45	3	19
45 - 50	1	20

Then, we can plot the points having abscissa as upper limits and ordinates as the cumulative frequencies (20, 0), (25, 1), (30, 3), (35, 11), (40, 16), (45, 19), (50, 20) and join the points by a smooth curve. The information is provided in Figure below.



3.5. Now we have to find the 80th percentile. Percentiles and percentile ranks provide information about how a person or thing relates to a larger group. Relative measures of this type are often extremely valuable to researchers employing statistical techniques. Percentiles are used to represent the percentage of scores "lower than" a given score. To calculate the given value we need to apply the following formula:

$$P_{80} = \frac{k}{100} \times n$$

We can note the given values to calculate the percentile. We have $n = 20$, $k = 80$. So, we can substitute into the formula and calculate the value.

$$P_{80} = \frac{80}{100} \times 20 = 16$$

We obtained the value of p , which is equal to 16. Thus we can write P_{80} is in the 16th location. From the table noted above we can note that $P_{80} = 37.5$ (interval from 35 to 40).

3.6. The median is the middle term in the arranged data set. The median is the mean of the middle two numbers (the 10th and 11th values). But if we need to estimate a single Median value we can use this formula:

$$\text{Estimated Median} = L + \frac{\frac{n}{2} + cf}{f} w = 30 + \frac{\frac{20}{2} - 11}{8} \cdot 5 = 29.375$$

Based on the value of mean, which is equal to 35 and the obtained graph the less-than ogive graph we can assume that the distributions is normal, because the normal distributions are symmetric around their mean.