

L2 Histograms Practice #3

For the three histograms here:

- discuss whether the distribution is Normal or not;
- estimate mean and median, and discuss any difference;
- estimate standard deviation.

Chart 7

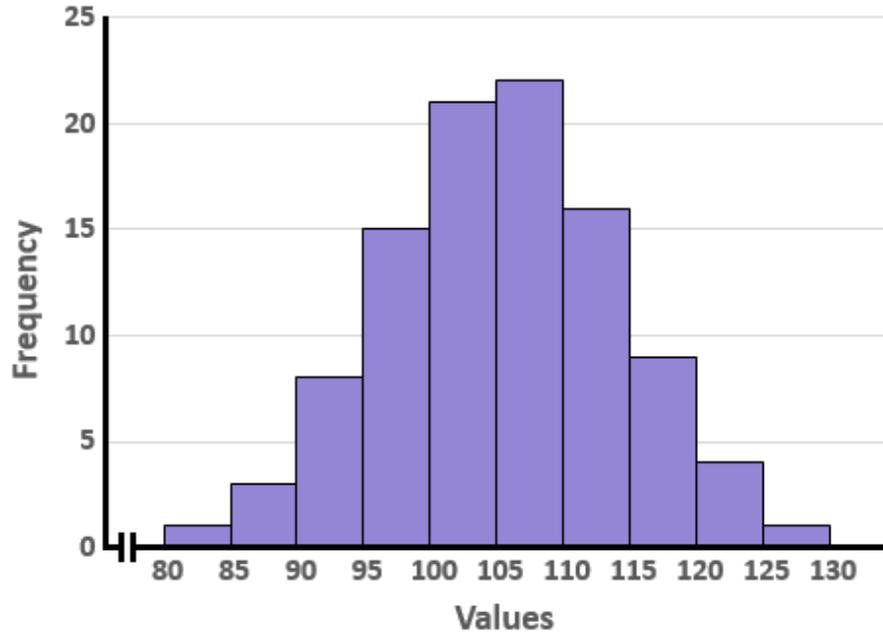


Chart 8

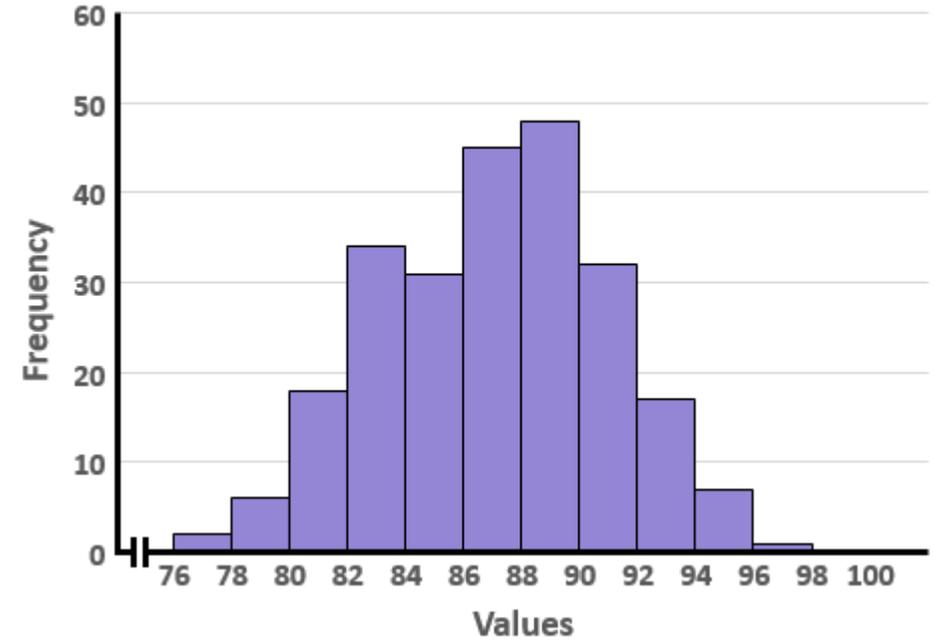
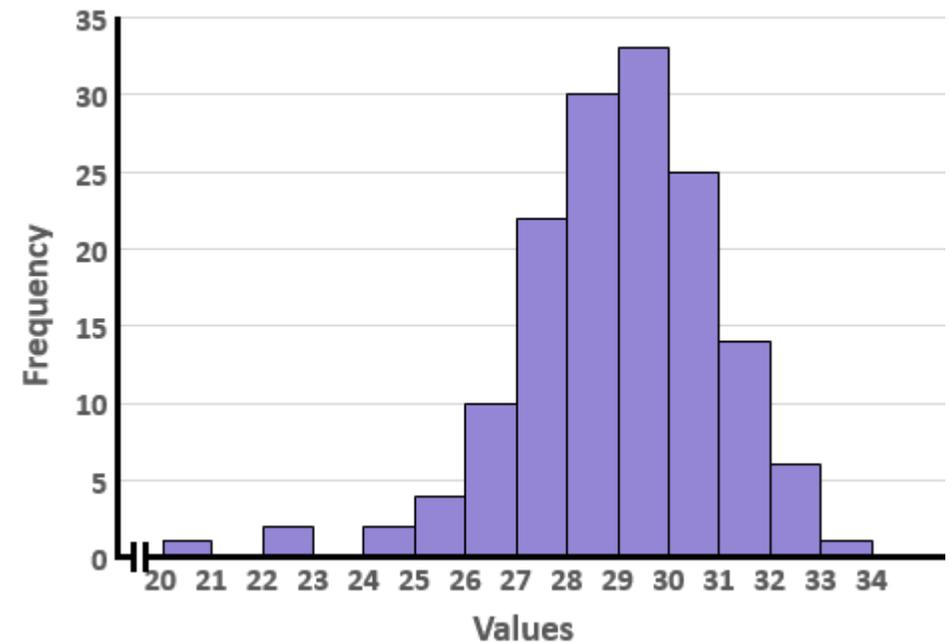


Chart 9



Answers – L2 Histograms Practice #3

Marks will depend as much on discussion of evidence as values given. Key terminology should be used rather than general terms.

The key features for a distribution being Normal should be discussed in each answer:

- “bell curve” shape fits the middle of the columns;
- symmetrical, allowing for some random variation in practice;
- drops off to effectively zero at 3 standard deviations.

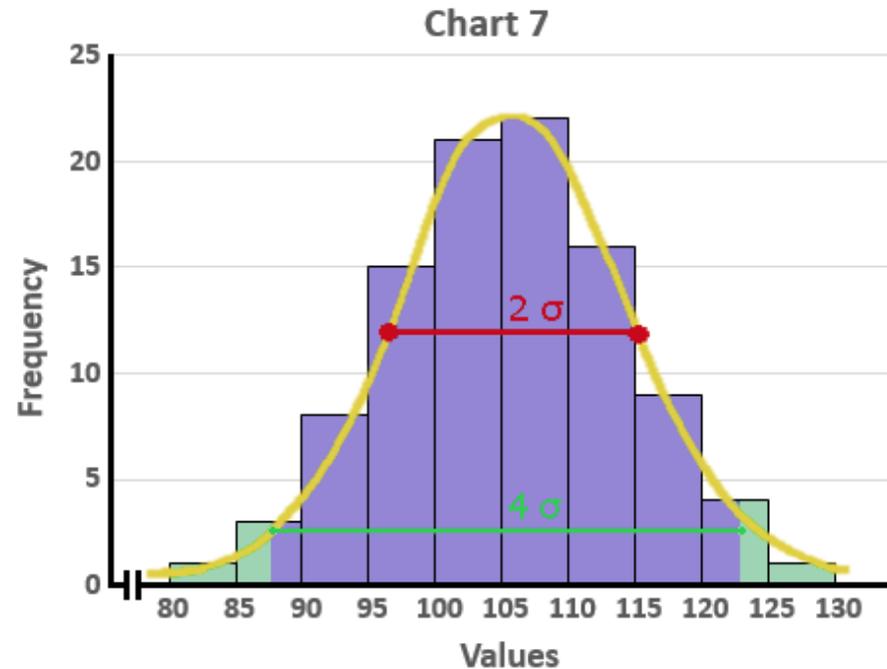
The median can be found by adding up the total frequency and counting in from either end.

The mean will be the median, moved towards any skew away from symmetrical.

The standard deviation can be estimated by two methods:

- shown in red, the distance between the inflection points (where the curve stops getting steeper) is 2σ ;
- shown in green, the middle 95% of values covers 4σ , so the outermost 5% of values can be counted off and the distance found $\div 4$.

Note, while all distributions have a standard deviation that we cannot make probability calculations for non-Normal ones. If a distribution is wildly non-Normal then σ cannot be calculated as if it is Normal.

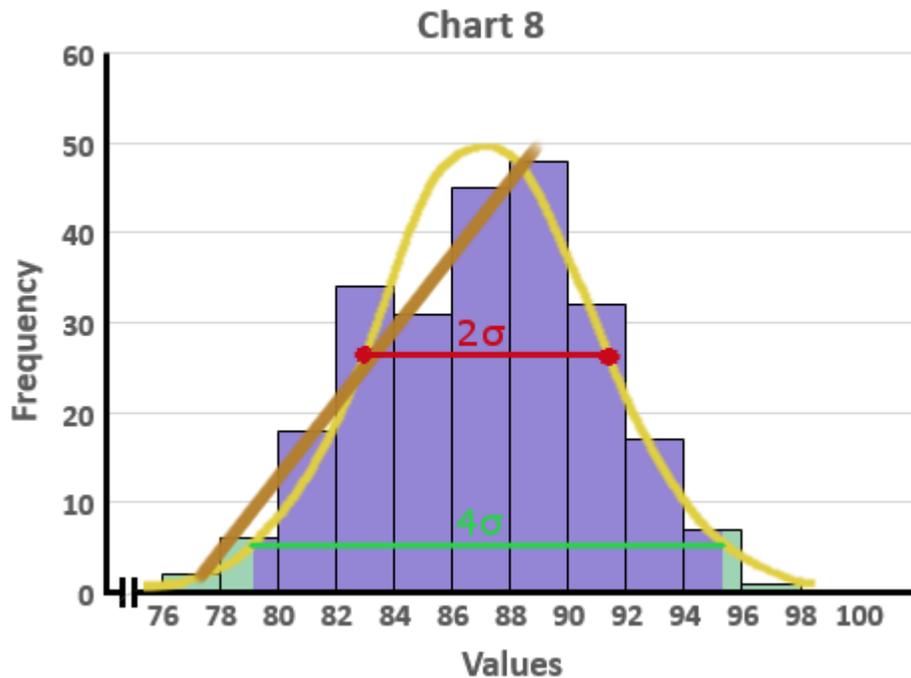


This is pretty much a perfect Normal distribution.

A curve through the centres of each column head is a bell shape, completely symmetrical. The ends tail off to zero outside 4σ .

The mean and median are probably about 106, but certainly somewhere in the region of 105 to 107. This is based on the peak of the curve and in the median's case also because there are 100 values and counting in from the left there are 48 up to 105, so it must be just into the 105–110 column.

Standard deviation is about 9 or so, definitely in the range 8–10. This can be seen from the red 2σ distance between the two points of inflection which are 18 apart, indicating $\sigma = 9$. Alternatively the green 4σ distance containing the middle 95% of values is 35, indicating $\sigma = 8.75$.



This appears to not be Normal.

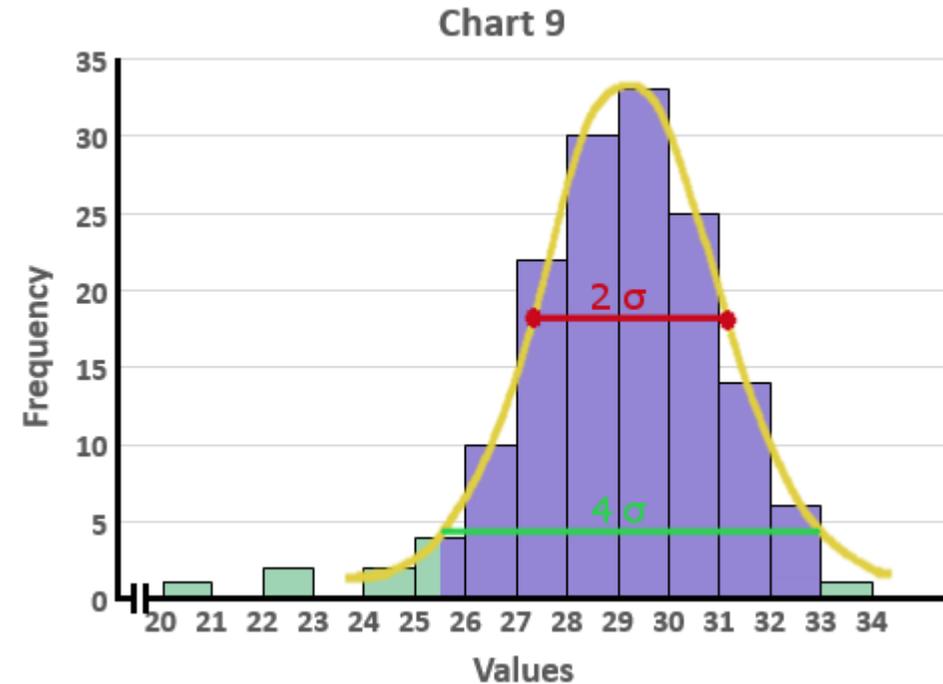
- The data is asymmetric, with far more values below the peak than above.
- While a bell curve approximates the right hand side, the left hand side fits a line much better than a curve.

Note however that there are only 108 data values, so it may just be a particularly unlikely sample of a Normal distribution. We can't tell for sure.

Counting 54 data points in, the median is around 111-112.

The mean will be lower than the median, as the tail values to the left are more distant (although not a lot further away as the skew is not great) so perhaps 109—110.

Standard deviation is in the range 11–12, although if non-Normal the s.d. is not useful for calculations of probability.



This is not Normal, but only because of the tail to the left.

The bulk of the curve fits a bell curve very well, symmetrical and tailing off to the right correctly. But the chance of 3 values out of 150 at more than 3.5σ is a billion to one, so we can strongly assume it is not Normal. (It may be a Normal distribution into which three values that are false in some way have entered – perhaps measured wrong or from another population altogether.)

The median is around 29.2. There are 150 values, so the median is the 75th/76th, on the left hand side of the 29–30 column. The mean will be a bit lower, as the low tail values will drag it away from the median – but there are only 3 of them out of 150, so it won't be a huge effect. (It ends up being 29.0.)

Standard deviation is a touch under 2. This can be seen from the red 2σ distance between the two points of inflection which are just under 4 apart, indicating $\sigma = 1.95$ or so. Alternatively the green 4σ distance containing the middle 95% of values is 7.5, indicating $\sigma = 1.9$. However this σ is based only on the main 147 values and has to ignore the three very low values.