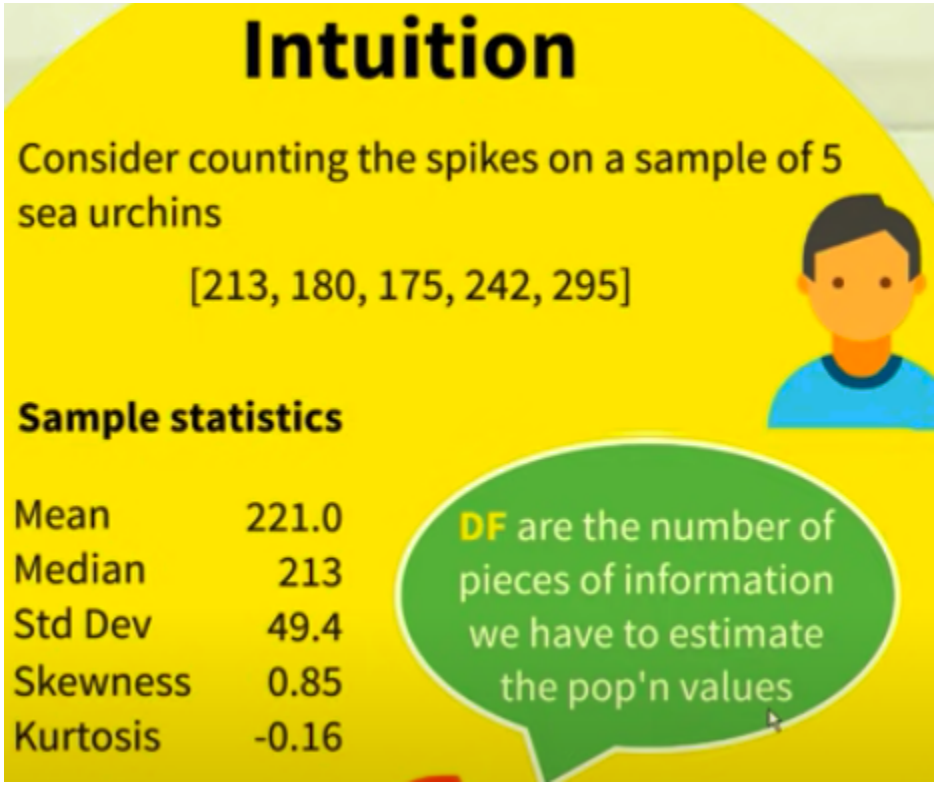
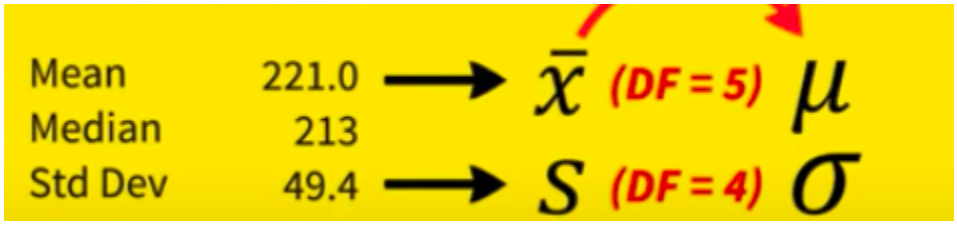
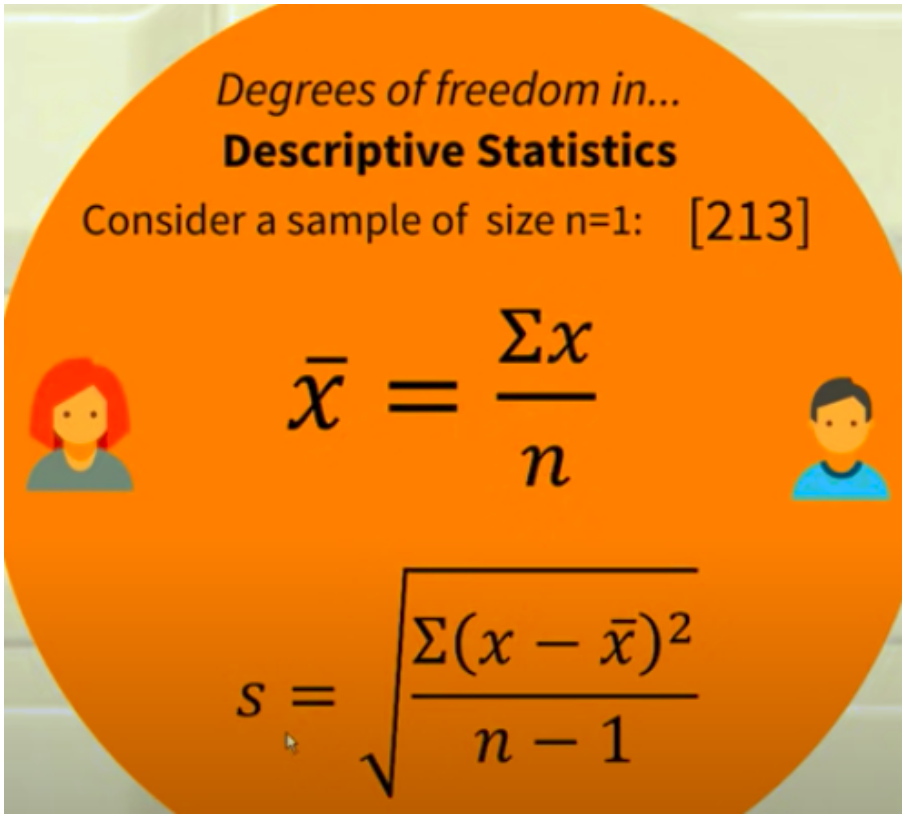
DEGREES OF FREEDOM



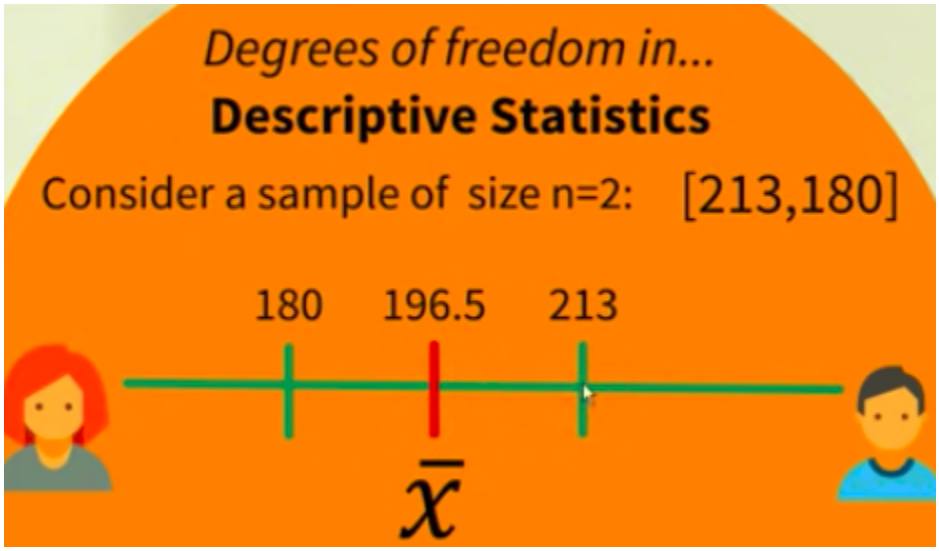
- Degrees of freedom - the number of independent pieces of information that we have.



- We only have 4 degrees of freedom for standard deviation (sigma).
- We have 5 degrees of freedom for mean
- In statistics we are taking a sample of something, and we are using the sample to estimate a population value. So the number of pieces we have to estimate that population value is the degrees of freedom.

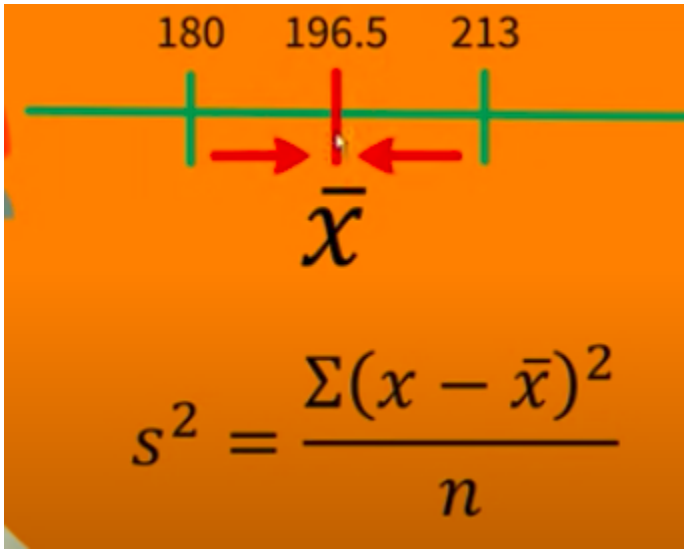


- The standard deviation is undefined, because there is only one sample. We do not know what population it came from, and have zero degrees of freedom.
- When we are looking for the mean, we have 1 degree of freedom as we only need that number.



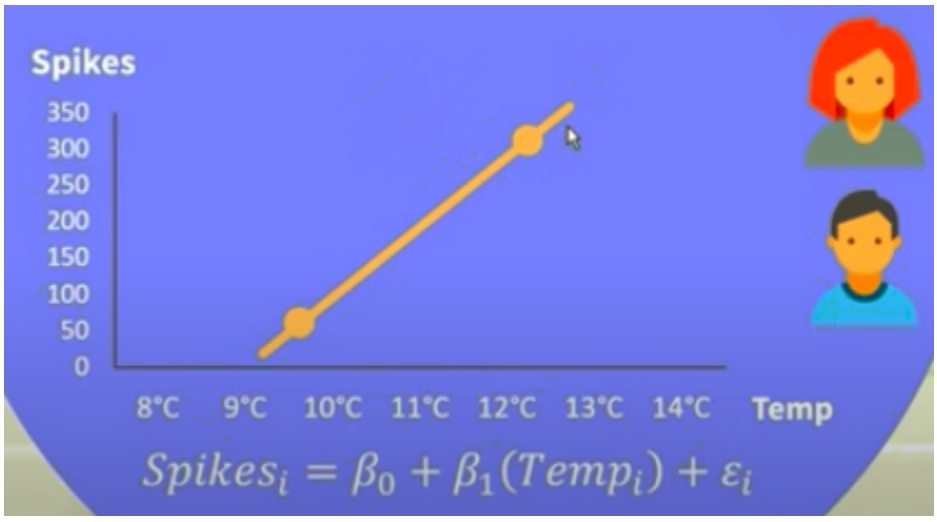
- In this case, if we look to find the average deviation from the mean it will not work. The mean was calculated between these two numbers, rendering this technique useless. Instead we use variance.

- We find the average squared deviation from the mean. Negative ones become positive, and we find some robust measure of the spread.



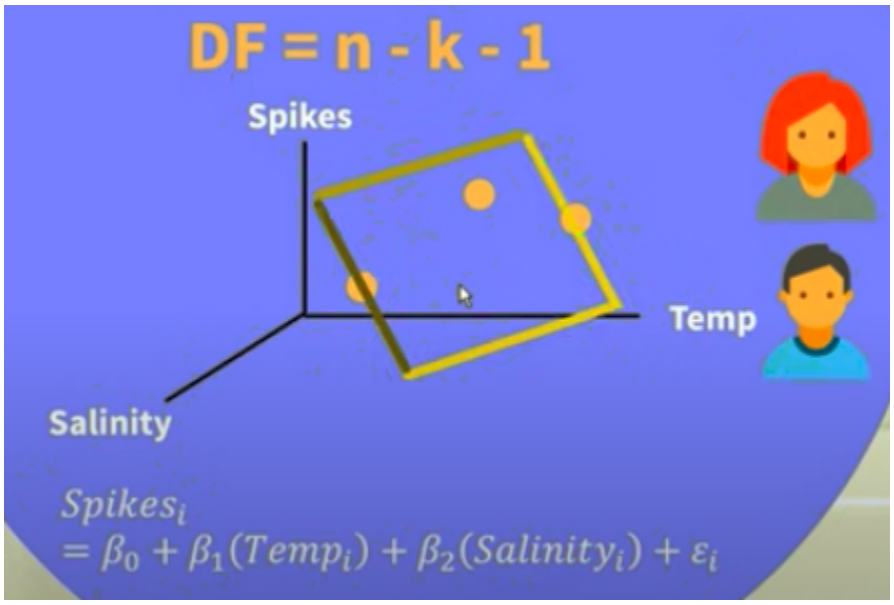
- There is a problem with this formula - X bar is the best case scenario for this formula. Therefore we must change the denominator, and artificially increase the calculation. The degrees of freedom is n-1. It is only with the second observation that we are able to get a calculation for the variance, leading to a standard deviation.
- That second number is required in order to make this work, therefore we have 1 degree of freedom.

REGRESSION



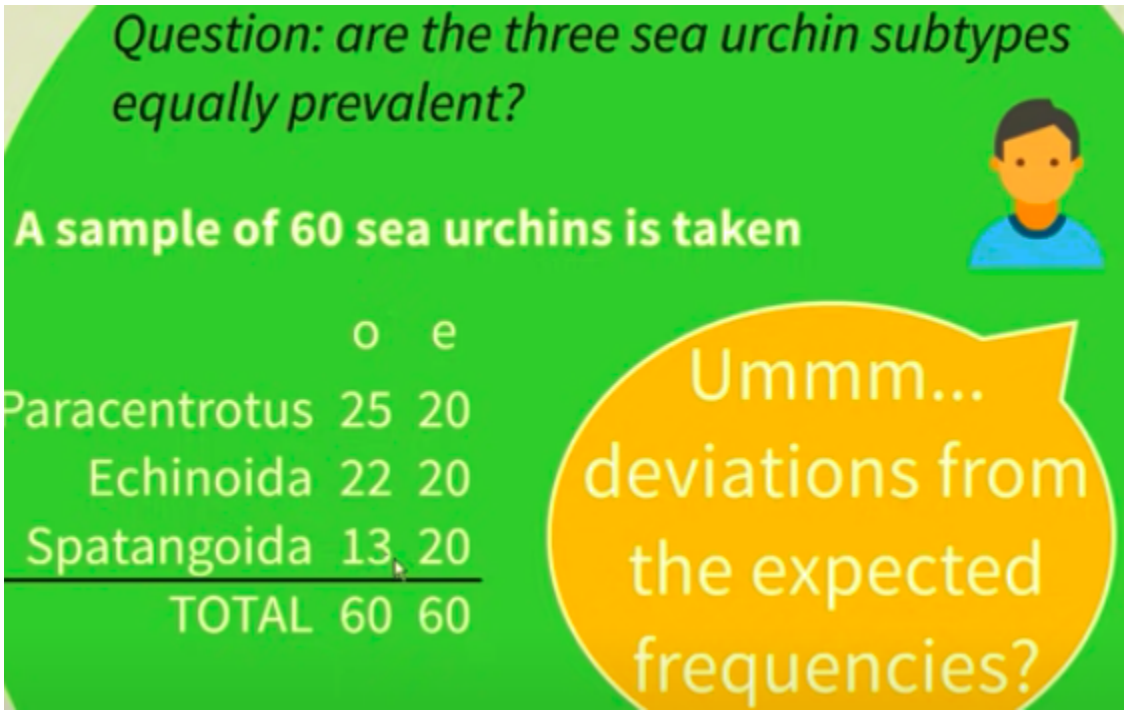
- Regression estimates the uncertainty we have with our given values. With only 2 points we cannot estimate this as the line fits perfectly with whatever 2 points supplied. Therefore we need a third.
- Only with the third observation do we have a degree of freedom.
- Degrees of Freedom = # of observations - # of x variables (k) - 1
- Df = 3 - 1 - 1 There are 3 points, and 1 degree of freedom.
  Df = 2 -1 -1 There are only 2 points, and no freedom.

- When we add another variable say salinity, and temperature take the following diagram

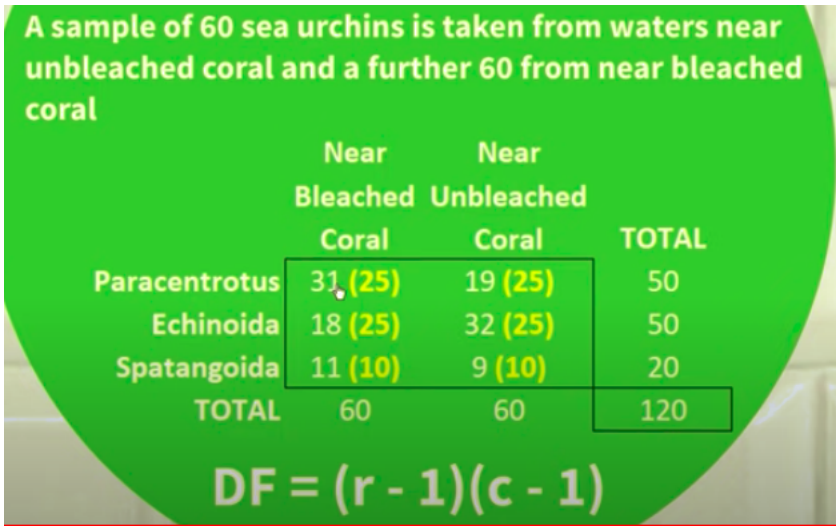


- Question - what is the minimum number of observations we need to run a regression?
- We are drawing a plane of best fit as we have a 3 dimensional space. It is only with a 4th observation are we able to find uncertainty, and a degree of freedom.
- DF = 4 (observations) - 2 (x variables) - 1
  1 degree of freedom.

CHI SQUARED



- We are estimating deviations from the expected frequencies. Given we have 60 total observations, equal frequency would be 20 x 3.
- We are asking how many degrees of freedom do we have in the necessity of creating the Chi Squared statistic. Because we have the total we can acquire the third through simple algebra.
  DF = k (number of obs) - 1
  DF = 3 - 1
  Two degrees of freedom.



DF = (3 -1) (2-1) = 2 degrees of freedom