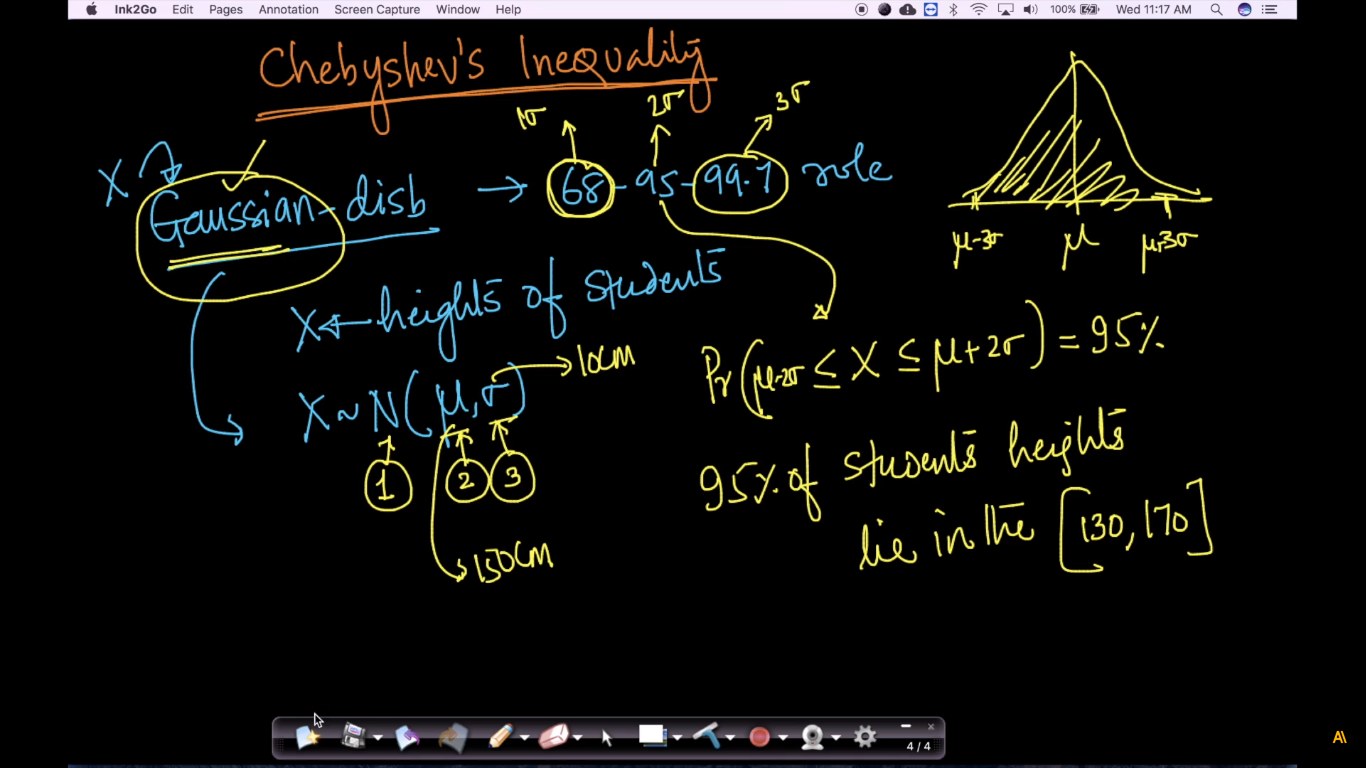
If give random variable is gaussian distributed then with help of 68-95-99.7 rule, we can easily say how much population lies within mean+ k \* var and mean – k\*var, where k is 1, 2, 3 respectively.

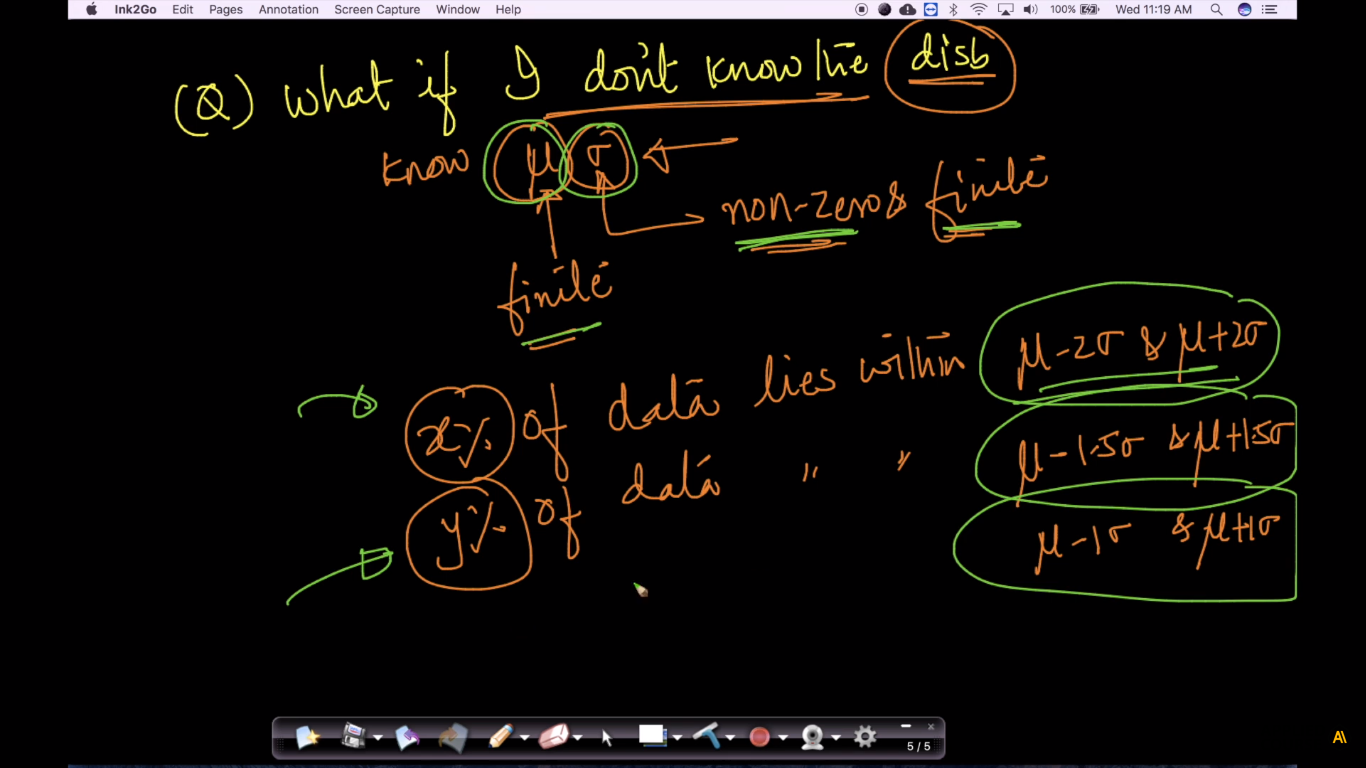
In below fig there is example of height where we can see that since height are normally distributed and therefore we can easily say that 95% of height are between 130 and 170.



Now what if we don’t know the distribution, the only thing we know is that:

* Mean is finite
* SD is non-zero and finite.

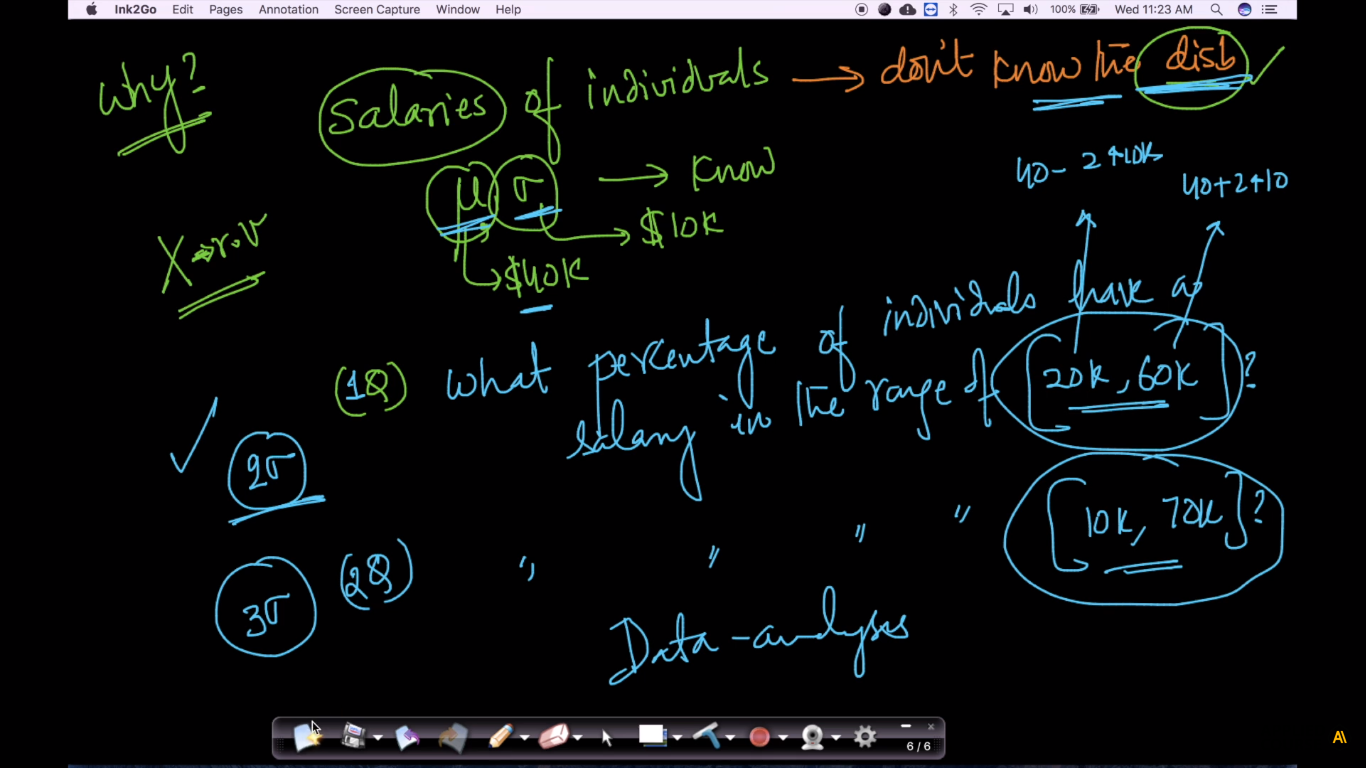
So just by knowing this without knowing distribution we can’t tell how much %age of observations lies within particular range.



Below fig shows that we have salaries but we don’t know distribution, only thing we know is

* Mean is $40k
* SD is $10k

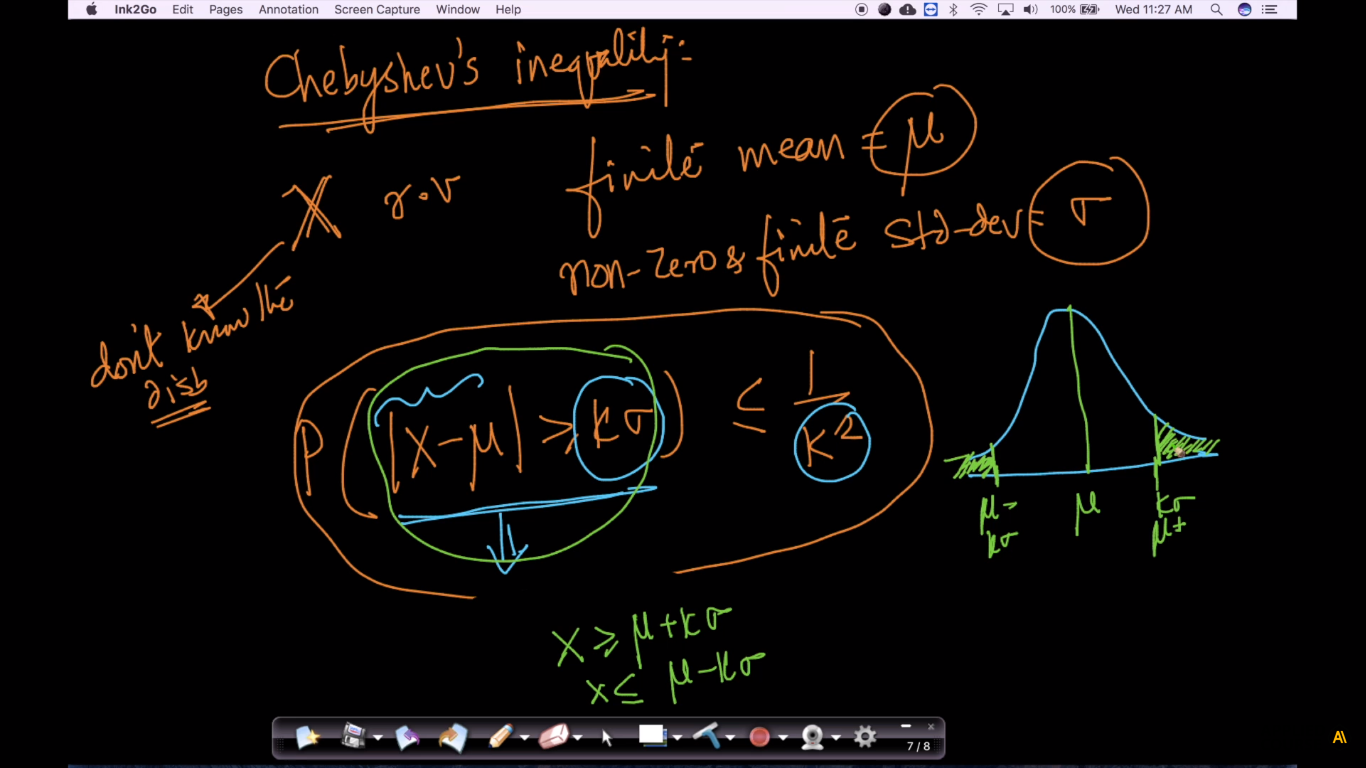
And we want the percentage of individuals having salary in range 20k – 60k, and 10k – 70k.



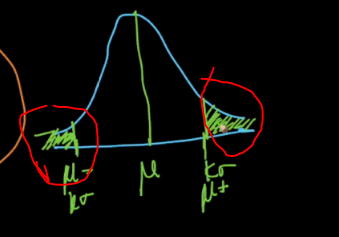
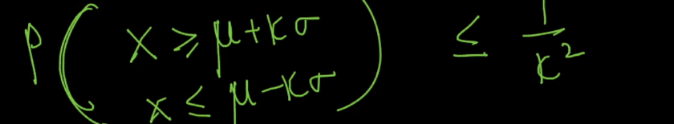
So, to overcome this limitation **chebyshev’s inequality** is used.

**chebyshev’s inequality**

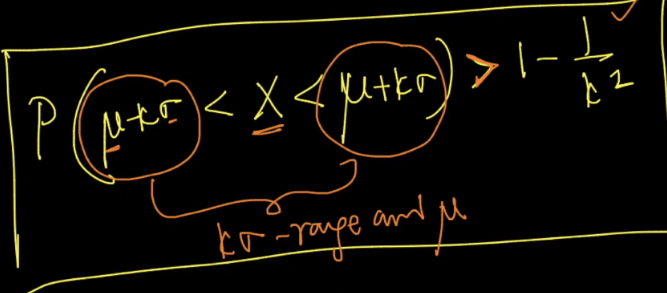
Given finite mean and non-zero & finite SD, the Chebyshev’s inequality is given as equation mentioned in below fig.

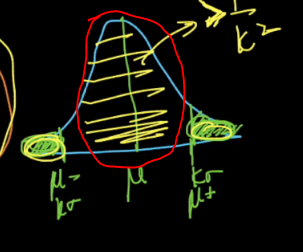


This equation states that Probability : P( X < mean – k \* SD ) or P ( X > mean + K \* SD) is at least 1/k2

Since we now the probability of outer is 1/k2, probability for inner area is at least 1-1/k2





**Note:** Remember that Chebyshev’s enequality gives the at least probability, the probability can be greater the prob we’ve find using Chebyshev’s enequality

Hence now using Chebyshev’s enequality we can answer the salaries range question as

Salary range between 20-60, that means k is 2 here for mean=40 and SD=10, is 0.75 or 75%

