Z test

Two group comparison : Do women have a greater risk of developing tennis elbow?

Null hypothesis, $H_0 \rightarrow P_{women} : P_{men}$.

In USA.

$\hat{p}_{women} = 0.11$

$\hat{p}_{men} = 0.11$

USA

	Tennis Elbow (TE)	Non-TE	Total
Women	11	89	100
Men.	10	90	100
Total	21	179	200.

But does this mean, we have enough evidence. No.

AUSTRALIA.

	TE	Non-TE	Total
Women	20	80	100
Men	6	94	100
Total	26	174	200

No enough evidence.

$\hat{p} = 0.20$ $\hat{p} = 0.06$

In Australia sample, is the difference in the SAMPLE above big enough to infer a difference in the population?

→ Use Z test, $n_{women} = 100$ $\hat{p}_{women} = 0.20$

$n_{men} = 100$ $\hat{p}_{men} = 0.06$ $\hat{p} = 0.13$

$$z = \frac{\hat{p}_{women} - \hat{p}_{men}}{\sqrt{\hat{p} \times (1-\hat{p}) \times \left(\frac{1}{n_{women}} + \frac{1}{n_{men}}\right)}}$$

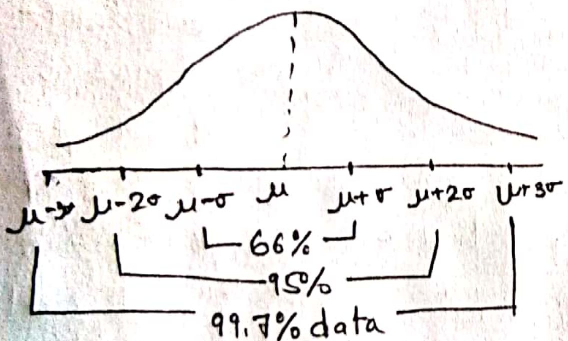
, $z = 2.943$, pvalue $= 0.0032$. So, z score is ~~accepting~~ ~~based on~~ rejecting null hypothesis as P value is less than 0.01.

Another method. Chi-Square test, chi-sq = 8.664, p-value = 0.0032.

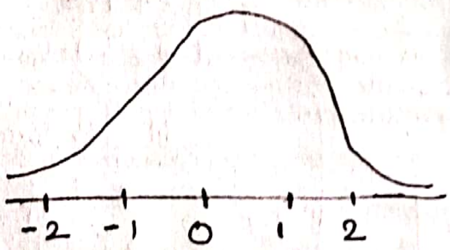
So, by seeing p = 0.0032, yes there is a significant difference between the men & women.

Z Score

Gaussian / Normal Distribution $\xrightarrow[\sigma=1 \text{ Convert}]{\mu=0}$ Standard Normal Distribution



With help of Zscore we will be able to convert this part.

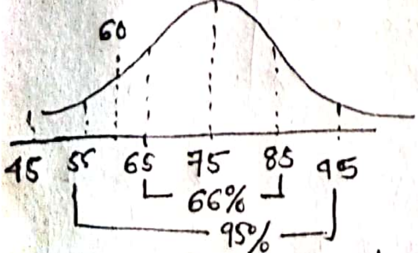


$$Zscore = \frac{x_i - \mu}{\sigma}$$

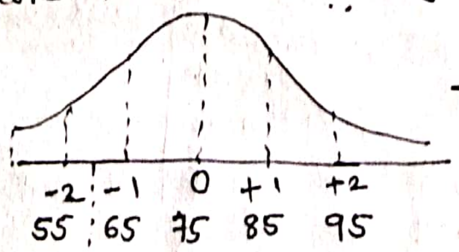
In feature engineering we use standard normilization which is also a Z score.

Suppose in a class, avg student marks is $\mu = 75$ and standard deviation $\sigma = 10$.

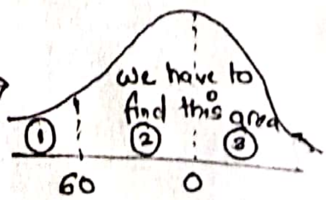
What is the probability student will score more than 60? $P(x > 60)$



convert to standard normal dist



→



But by this we cannot get exactly at 60.

60, which is equal to −1.5

So through z score we will find area < −1.5 which is area ① that is around 6%

We found the Z table / Z chart.

For ③, it is half of chart that is 50% of distribution.

Area = ① + ② + ③

100% = 6% + ② + 50%

② = 44%.

So, 44% is the probability student will score > 60.