Goodness of fit test

Test Statistic.

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - e_i)^2}{e_i} \sim \chi^2(DF = k)$$

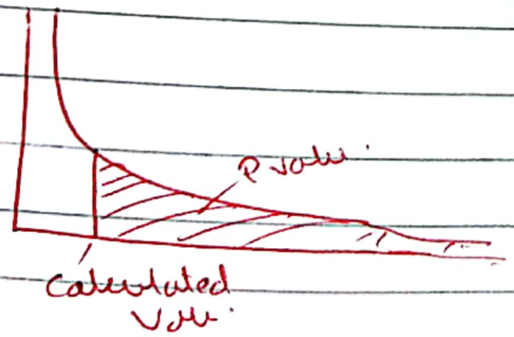
Note: All the expected ~~values~~ frequencies are greater than 5. if not we can't use this formula.

Decision Rule: (when to reject or accept)

Reject if $\chi_i^2 >$ tabulated value at given level of significance.

Do not reject elsewise.

P value approach. (Area in the ~~tail~~ above that particular point)



P value > 0.05 (given level of sig) can't reject.

where is the hidden normal distribution in this formula!

The frequency greater than 5 ensures the CLT that the distribution is normal.

for ex: let you have a data from binomial distribution.

$$z = \frac{x - n\pi}{\sqrt{n\pi(1-\pi)}} \sim N(0,1)$$

(frequency greater than 5 ensure CLT)

$$z^2 = \frac{(x - n\pi)^2}{n\pi(1-\pi)}$$

$$= \frac{(x-n\pi)^2}{n\pi} + \frac{(n-x-n(1-\pi))^2}{n(1-\pi)}$$

$$= \sum_{i=1}^{k} \frac{(O_i - E_k)^2}{E_i}$$ (get χ^2 value.)