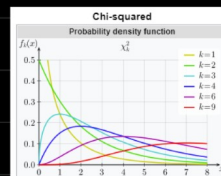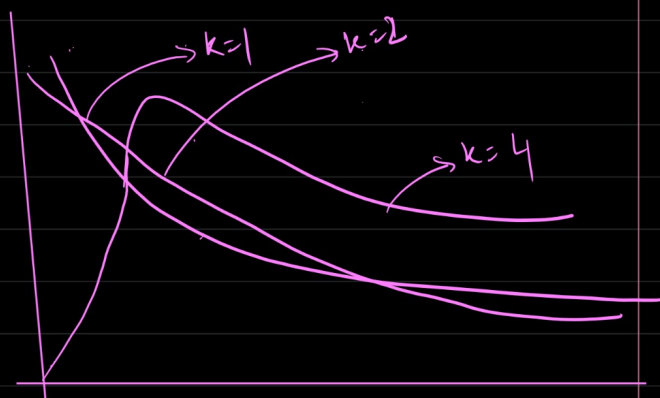# Chi-Square distribution and Chi-square test

* The chi-square distribution is a probability distribution that describes the distribution of a sum of squares of K random variables.

Chi-squared
Probability density function

→ degree of freedom $(k) = n-1$

→ Chi-square distribution shape is determined by `k`

→ Non-negative distribution

→ right skewed distribution

* Chi-square test $(X^2 \text{ test})$

→ Chi-square distribution.

→ Goodness of fit test — Used to compare observed and expected Categorical data.

→ Test of Independence — To determine the relationship b/w two categorical variable.

→ It test the claims about population proportions.
↑ Categorical

→ Non-parametric test — No assumption about the population.

Types of car	(Expected) theory	(observed) Sample
Sports Car	$\frac{1}{3}$	22
SUV	$\frac{1}{3}$	17
Sedan	$\frac{1}{3}$	59

↑
theoretical categorical distribution.

→ observed categorical distribution
⇓
Using this observed Sample distⁿ, you have to verify it

the theoretical distribution
is true   or not
→ goodness of fit

ex: In a class of 75 students, 11 are left handed.

theory → 12% of the people   are   left handed.

                                    O   E

                    left handed   11   9   12% of 75

                    right handed   64   66 $\frac{12}{100} \times 75 = 9$

$\chi^2_{statistic} = \sum \dfrac{(Observed - Expected)^2}{Expected}$

$$\underbrace{\dfrac{(11-9)^2}{9}}_{left} + \underbrace{\dfrac{(64-66)^2}{66}}_{\searrow right hand}$$