Helmert (1875) , Karl Pearson. (1900)

Chi- Squared distribution

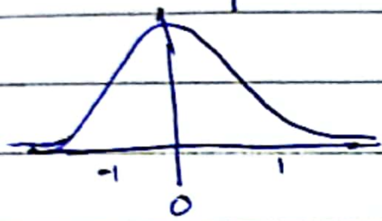
$$\chi^2_k = \sum_{i=1}^{k} z_i^2$$

$E(\chi^2_k) = k$

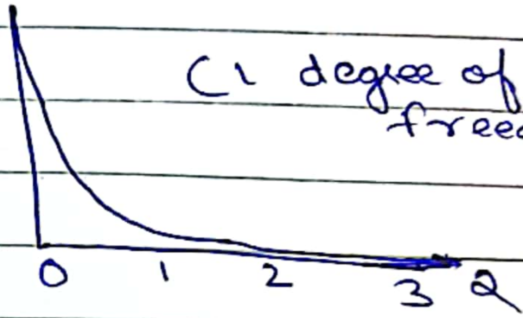
$V(\chi^2_k) = 2k$

z = Std Normal Dist

$z \sim N(0,1)$

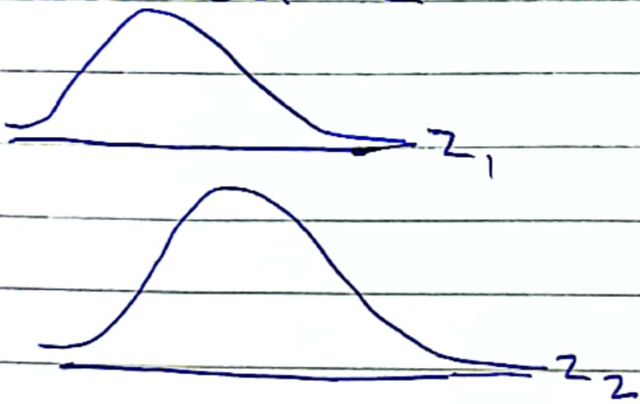


Consider $Q = Z_1^2 \sim \chi_1^2$

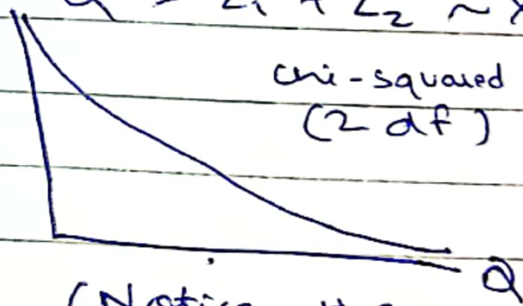


(1 degree of freedom)

Consider. 2



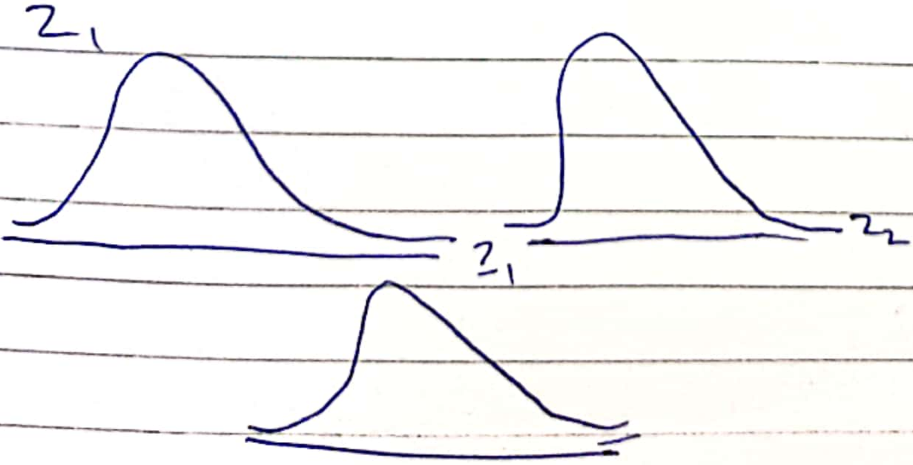
$Q = Z_1^2 + Z_2^2 \sim \chi_2^2$

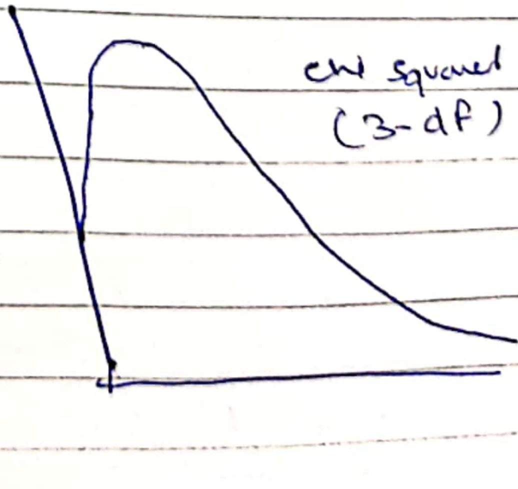


chi - squared (2 df)

(Notice the difference, see how getting value at tails is higher than the first)

Consider 3





chi squared (3-df)

$$\boxed{Q = \sum_{i=1}^{k} z_i^2 \sim \chi_k^2}$$ k degree of freedom

Mean $= k$
Variance $= 2k$

Try to understand the concept:

$V(z_1) = E(z_1^2) - (E(z_1))^2$
$= E(z_1^2) - 0 = E(z_1^2)$

for 1 df:

$E(Q) = E(z_1^2) = 1 =$ Mean.

why Variance 2k?

$V(Q) = E(Q^2) - (E(Q))^2$
$= E(z_1^4) - [E(z_1^2)]^2$ ← kurtosis
$= 3 - [1]^2 = 2$

$V(\chi_k^2) = V\left(\sum_{i=1}^{k} z_i^2\right) = V(z_1^2) + V(z_2^2) + \cdots + V(z_k^2)$
$= 2 + 2 \cdots + 2$
$= 2k$

(becaus $Cov(z_i, z_j) = 0$)

$\boxed{\text{USE}}$: (i) Goodness of fit test
(ii) Test for Independence.

Basicaly these tests check:
Theoratical categorical distribution vs
Observed categorical distribution