Ans.1

a) $$gini(t) = 1 - \sum_{j=0}^{c-1} [p(i/t)^2]$$

$$= 1 - \left[\left(\frac{10}{20}\right)^2 + \left(\frac{10}{20}\right)^2\right]$$

$$= 1 - [1/4 + 1/4]$$

$$= 1/2$$

b) Female :

$$gini(t) = 1 - \left[\left(\frac{6}{2}\right)^2 + (1/20)^2\right]$$

$$= 1 - [0.36 + 0.16]$$

$$= 0.48$$

Male :

$$gini(t) = 1 - [(6/20)^2 + (4/20)^2]$$

$$= 1 - 0.52 = 0.48$$

Weighted avg $= [10/26 \times 0.48] + [10/26 \times 0.48] = 0.48$

c) Family :

$$gini(t) = 1 - [(1/4)^2 + (3/4)^2]$$

$$= 1 - 0.625$$

$$= 0.375$$

Luxury :

$$= 1 - [0.0156 + 0.7656]$$

$$= 0.21875$$

Sports :

$$= 1 - [(8/8)^2 + (8/8)^2]$$

$$= 0$$

Average :

$$[4/20 * 0.375] + [8/20 * 0.21875] + [8/20 * 0]$$

$$= 0.163$$

Ans. 2

$$\text{Entropy}(t) = -\sum_{i=0}^{c-1} p(i/t) \log_2 p(i/t)$$

$$= -\left[\frac{-4}{9} \log_2(4/9) + 5/9 \log_2(5/9)\right]$$

$$= -\left[(-0.57997) + (-0.47107)\right]$$

$$= 0.99107$$