

- Such a model is called a *generative model*
 - because it specifies the hypothetical random process that generates the data.

$$p(c|x) = \frac{p(x|c)p(c)}{p(x)}$$

- $P(c|x)$ is the posterior probability of class (c, target) given predictor (x, attributes).
- $P(c)$ is the prior probability of class.
- $P(x|c)$ is the likelihood which is the probability of predictor given class.
- $P(x)$ is the prior probability of predictor.



Step 1: Convert the data set into a frequency table

Step 2: Create Likelihood table by finding the probabilities like:

- $p(\text{Overcast}) = 0.29$, $p(\text{rainy}) = 0.36$, $p(\text{sunny}) = 0.36$
- $p(\text{playing}) = 0.64$, $p(\text{rest}) = 0.36$

Step 3: Now, use Naive Bayesian equation to calculate the posterior probability for each class. The class with the highest posterior probability is the outcome of prediction.

Problem: Players will play if weather is sunny. Is this statement is correct?

We can solve it using above discussed method of posterior probability.

$$P(\text{Yes}|\text{Sunny}) = \frac{P(\text{Sunny}|\text{Yes}) * P(\text{Yes})}{P(\text{Sunny})}$$

Here we have $P(\text{Sunny} | \text{Yes}) = 3/9 = 0.33$, $P(\text{Sunny}) = 5/14 = 0.36$, $P(\text{Yes}) = 9/14 = 0.64$