**Latex Command**

\documentclass[a4paper]{article}

\begin{document}

\end{document}

\title{}

\date{}

\author{}

\maketitle

\newpage

\pagenumbering{gobble} – no number

\pagenumbering{arabic} – Arabic number

\pagenumbering{roman} – roman number

\section

\subsection

\paragraph

\subparagraph

\begin{equation}

\usepackage{amsmath}, \begin{equation\*}

\begin{align\*}

1 + 2 &= 3\\

1 &= 3 - 2

\end{align\*}

\int^a\_b – integral a to b

frac{1}{x} – 1 / x

sqrt{x}

\left[

\begin{matrix}

1 & 0\\

0 & 1

\end{matrix}

\right]

Embed formulas in text by surrounding them with dollar signs $

The align environment will align formulas at the & symbol

Split formulas using \\

\lambda

insert graph

\begin{figure}[h!,t!,b!,p!]

\usepackage{graphicx}

\graphicspath{{}}

\includegraphics{filename}

\caption{} – text shown below image

\label{} – invisible, for reference purpose

\begin{table}

\begin{appendix}, \listoffigures, \listoftables

\cite{}

\bibliography{bibfilename}

\bibliographystyle{ieeetr}

\autocite{} – generate footnote

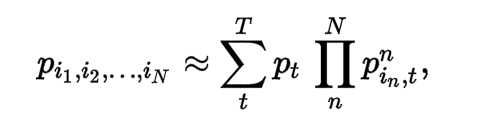
\footnote{\label{}}

\centering

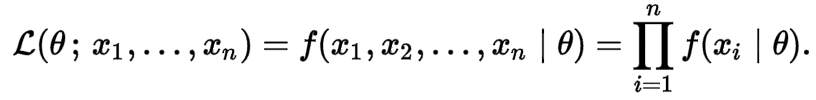
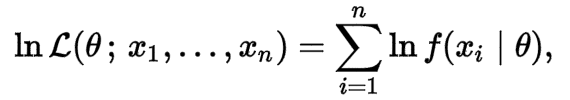
\begin{tabular}{}

\hline

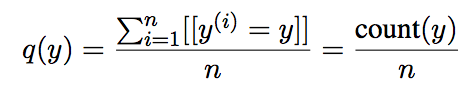
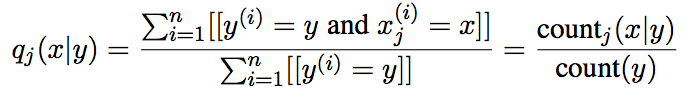
**Latent Class Model**

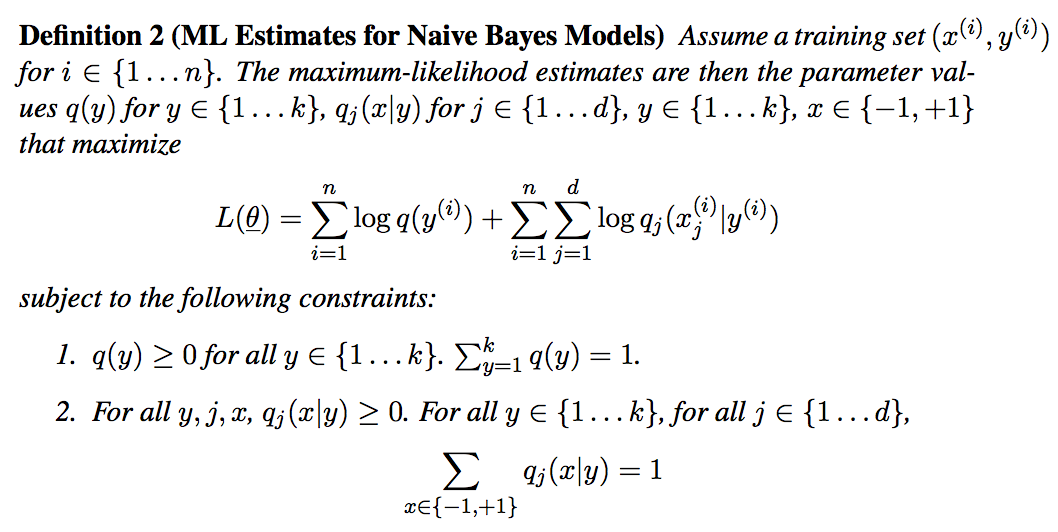
1. Within each latent class, the observed variables are statistically independent.
2. Usually the observed variables are statistically dependent. By introducing the latent variable, independence is restored in the sense that within classes variables are independent (local independence)
3. We then say that the association between observed variables is explained by the classes of the latent variables
4. 

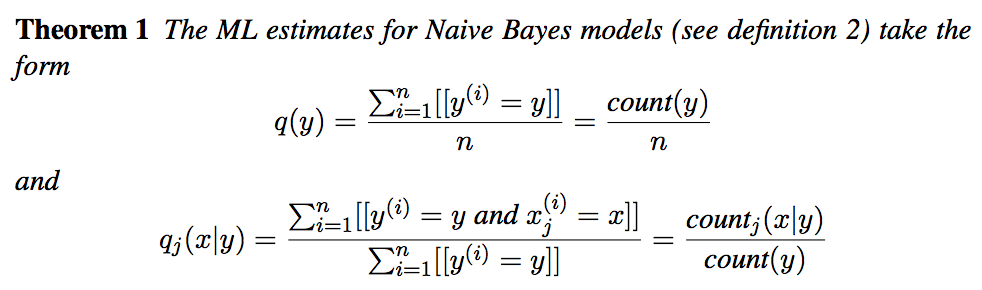
**Maximum Likelihood Estimation**

1. Method to estimate the parameters of a statistical model given observations, by finding the parameter values that maximize the likelihood of making the observations given the parameters.
2. Special case of the **maximum a posteriori estimation**
3. Assume a uniform prior distribution of the parameters
4. For a fixed set of data and underlying model, it selects the set of values of model parameters that maximizes the likelihood function.
5. Example:
   1. Suppose there is a sample from a distribution with unknown probability density function F
   2. F belongs to a certain family of distributions (Theta is a vector of parameters for this family)
   3. We need to find the Theta such that it is close to the true Theta
   4. 
   5. **likelihood function**:
   6. **log-likelihood function**: 

**Naïve Bayes**

1. It has two types of parameters
   1. 
   2. 
2. Maximum likelihood estimate
   1. 
   2. 
   3. Derivation





Then use lagrange multiplier to prove it

# Annealed EM

1. EM will only converge to a local maximum
2. Annealed EM can help you find the global maximum
3. Deterministic Annealing is to “smooth” the posterior “landscape” by raising it a to a temperature, and then gradually cooling it, all the while slowly tracking the global maximum

**Latent Dirichlet Allocation**

1. A generative statistical model that allows sets of observations to be explained by unobserved groups that explain why some parts of the data are similar
2. Each document may be viewed as a mixture of various topic where each document is considered to have a set of topics that are assigned to it via LDA.
3. Each topic is a mixture of words
4. The topic distribution is assumed to have a sparse Dirchlet prior
5. Sparse Dirchlet prior: documents cover only a small set of topics and that topics use only a small set of words frequently
6. Dirichlet Distribution:
   1. continuous multivariate probability distributions parameterized by a vector of positive reals.
   2. Very often used as prior distributions in Bayesian statistics
   3. 
   4. 1,2,3,4,…k is the category. X1, X2…Xk is the probability. a1,a2,…ak is the count of appearance for each category

**Gamma Function**

1. 

**Beta Function**

1. 

**Probabilistic Latent Semantic Analysis**

1. Consider the observations in the form of co-occurences (w,d) w is word , d is document
2. 
3. c is the words topic
4. Number of topics must be chosen in advance and is not estimated from the data
5. It is not a generative model of new documents