Bayesian Statistics

Qianqian Yu

Basic idea:

1. Bayesian statistics is a mathematical procedure that applies probabilities to statistical problems. It provides people the tools to update their beliefs in the evidence of new data

2. Bayesian philosophy provides a formal framework for such knowledge creation. This framework depends upon prior information, data, and the balance between them.

Bayesian Model:

1. Prior: the information we know at first
2. Likelihood: It indicates how likely a particular population is to produce an observed sample.
3. posterior probability: It is the revised or updated probability of an event occurring after taking into consideration new information

Formula:

1. $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(B)L(B|A)}{P(A)}$
2. $Posterior = \frac{Prior \cdot Likelihood}{Normalizing\ Constant}$

Beta-Binomial Model:

1. The Beta Model: : π~Beta(α,β)
   π be random variable between [0,1] , α, β are shape hyperparameters and both > 0
   Special case: $\pi \sim Beta(1,1)$, equivalent to $\pi \sim Unif(1,1)$.
2. Y$|\pi \sim Bin(n, \pi)$
   will give $\pi|(Y = y) \sim Beta(\alpha + y, \beta + n - y)$
   $\pi \sim Beta(\alpha, \beta)$

Latent Dirichlet Allocation (LDA) Model:

1. LDA is a Bayesian model used in natural language processing (NLP) to extract topics from a corpus of text documents
2. The term latent means something that exists but is not yet developed.
3. Documents are a mixture of topics
4. Topics are a mixture of words
5. Probability of document:

$$P(W, Z, \theta, \varphi, \alpha, \beta) = \prod_{j=1}^{M} P(\theta_j; \alpha) \prod_{i=1}^{K} P(\varphi_i; \beta) \prod_{t=1}^{N} P(Z_{j,t}|\theta_j) P(W_{j,t}|\varphi_{Z_{j,t}})$$

W -- words in document Z ---Latent topic θ -- topic distribution

φ -- topic-word distribution α -- per-document topic distribution β -- per topic word distribution