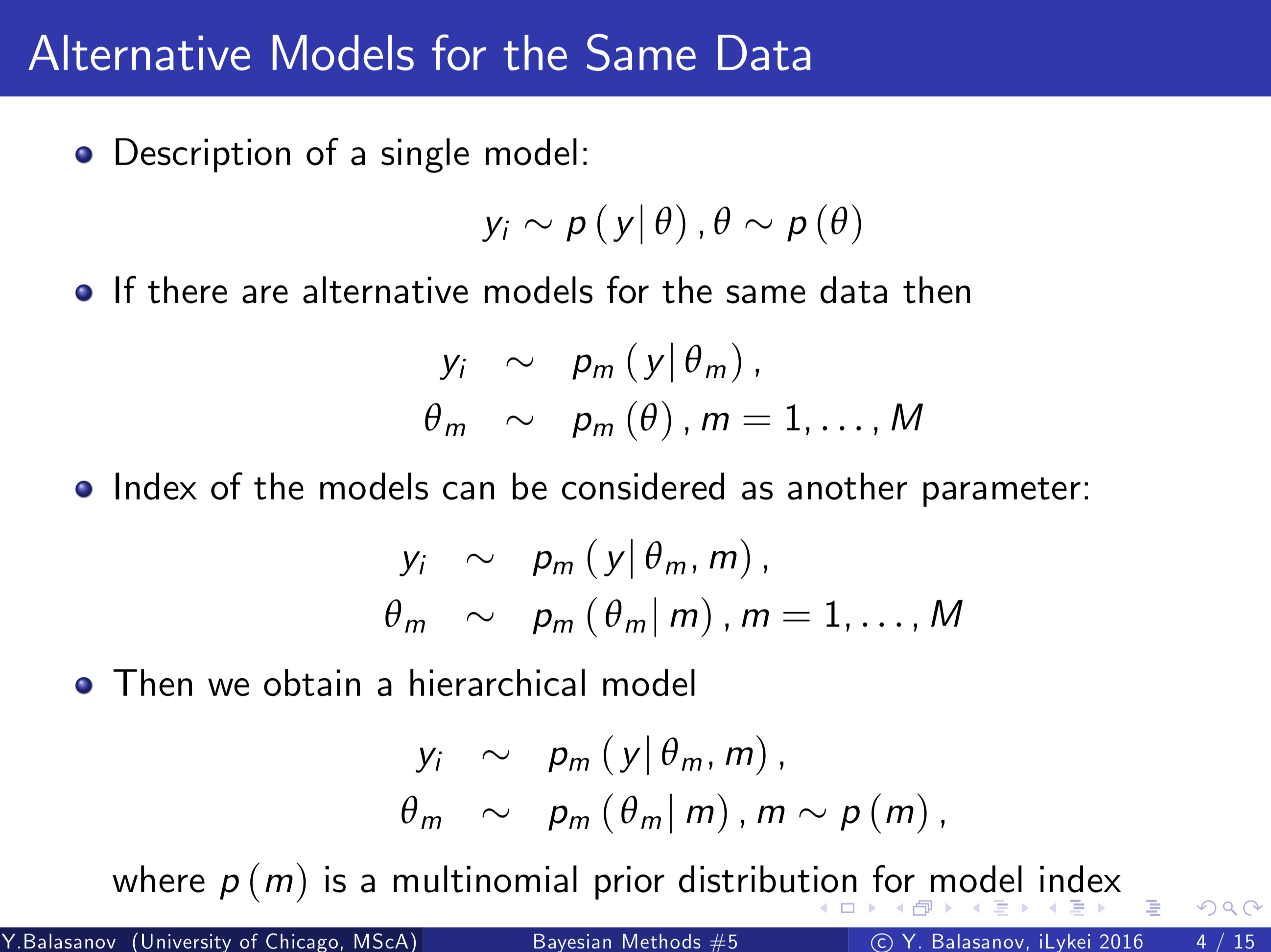
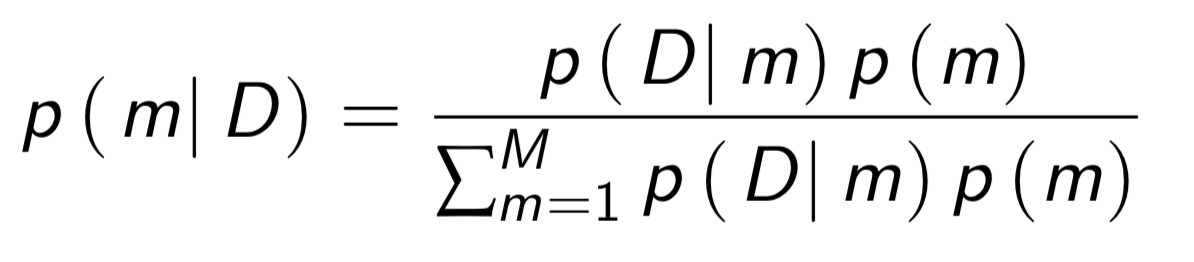
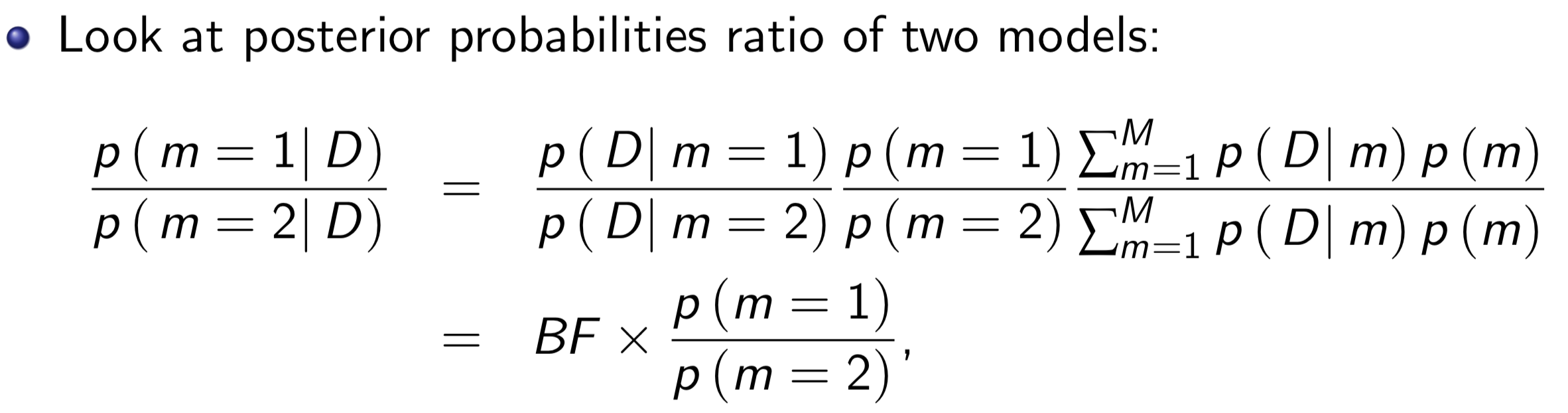
* + Index through models with m
  + 

* + Parameters are included as one of, so we can break them down into product;
    - The poster distribution will be proportional to product
    - If we do not estimate theta simultaneous with m, we use Bayes formula
      * 
      * Same data is used for every model
      * Probability of Data given m is the likelihood function for m
      * Distribution of m depends on prior distribution of Theta
  + **Bayes factor**
    - Look at odds ratio of m (posterior distributions)
    - 
    - Odds ratio of posteriors equals Bayes factor multiplied by odds ratio of priors
      * If Bayes factor is high, we have preference for model 1, otherwise model 2
        + But it depends on the ratio of the two priors
        + If one prior numerator is zero, then there is no way ratio of posterior distributions will be different than zero
    - Rule of thumb: if BF > 3, then we have a preference for m 1; if BF < 1/3, then we favor m2
  + **Example**
    - We set priors for m to .5 if we have no expectation either model is favored
    - Maybe it is best to provide no prior so each model has a fair chance of winning
    - If more than two models exist, use Markov Chain Monte Carlo to see which model state is the most frequent to select the best model
  + **Predicting by Averaging Models**
    - Use posterior probabilities as weights to ensemble the models
  + **Role of Model Complexity**
    - In Bayes, big model vs. little model is taken care of automatically
      * Big model with 5 parameters and small model with 2 parameters
      * Models are nested since parameters are nested
      * When you have distribution over more dimensions, the distribution is much smaller
    - Big model may have more parameters or more potential values for the parameters
    - Prior probabilities for larger model are smaller and are multiplied to get the posterior
    - Dimensionality is the number of parameters in the model
      * If we look at big model versus small models (nested models: sum of parameters are conditioned to zero)