베이즈 통계 스터디 1주차

발제자:이현정

1. Statistical Modeling

- objectives:
  1. Quantify uncertainty
  2. Inference
  3. Measure support for hypotheses
  4. Prediction
- process:
  1. Understand the problem
  2. Plan ad collect data
  3. Explore data
  4. Postulate model
  5. Fit model
  6. Check model
  7. Iterate
  8. Use model

2. Bayesian Modeling

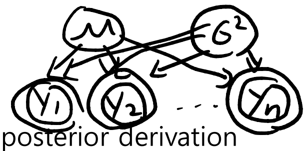
- components of Bayesian models
  1. Likelihood : $P(y|\theta)$
  2. Prior : $P(\theta)$
  3. Posterior : $P(\theta|y)$
- model specification

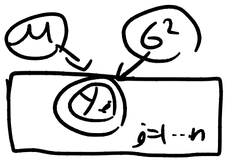
$Y_i | \mu, \sigma^2 \overset{iid}{\sim} N(\mu, \sigma^2)$ $i=1, \dots, n$: likelihood

$p(\mu, \sigma^2) = p(\mu) \cdot p(\sigma^2)$: prior

$\mu \sim N(\mu_0, \sigma_0^2)$

$\sigma^2 \sim IG(\nu_0, \beta_0)$





: graphical representation.

- posterior derivation

- non-conjugate models

3. **Monte Carlo Estimation**

What Is a Monte Carlo Simulation?

Monte Carlo simulations are used to model the probability of different outcomes in a process that cannot easily be predicted due to the intervention of random variables. It is a technique used to understand the impact of risk and uncertainty in prediction and forecasting models.

A Monte Carlo simulation can be used to tackle a range of problems in virtually every field such as finance, engineering, supply chain, and science. It is also referred to as a multiple probability simulation.

The basis of a Monte Carlo simulation is that the probability of varying outcomes cannot be determined because of random variable interference. Therefore, a Monte Carlo simulation focuses on constantly repeating random samples to achieve certain results.

A Monte Carlo simulation takes the variable that has uncertainty and assigns it a random value. The model is then run and a result is provided. This process is repeated again and again while assigning the variable in question with many different values. Once the simulation is complete, the results are averaged together to provide an estimate.

- monte carlo integration

$$E(\theta) = \int_{-\infty}^{\infty} \theta\, p(\theta)\, d\theta \approx \frac{1}{m}\sum_{i=1}^{m} \theta_i^*$$

$$E[h(\theta)] = \int_{-\infty}^{\infty} h(\theta)\, p(\theta)\, d\theta \approx \frac{1}{m}\sum_{i=1}^{m} \theta_i^*$$

- monte carlo error and marginalization

$$\bar{\theta}^* = \frac{1}{m}\sum_{i=1}^{m} \theta_i^* \qquad \bar{\theta}^* \sim N\left(E[\theta], \frac{Var(\theta)}{m}\right)$$

$$\widehat{Var}(\theta) = \frac{1}{m}\sum_{i=1}^{m}(\theta_i^* - \bar{\theta}^*)^2, \quad SE = \sqrt{\frac{\widehat{Var}(\theta)}{m}}$$

- computing examples & computing monte carlo error

Lesson3 test

5 & 6

```
> m = 100000

> phi = rbeta(n=m, shape1=5.0, shape2=3.0)

> mean(phi)

[1] 0.6250099

> h = phi/(1-phi)

> mean(h)

[1] 2.499682

> prob = h>1

> mean(prob)

[1] 0.77298
```

7.

```
> qnorm(p=0.3,0,1)

[1] -0.5244005
```

8.

> sqrt(5.2/5000)

[1] 0.03224903

4. Markov Chains

- *Markov assumption:* given the entire past history, the probability distribution for the random variable at the next time step *only* depends on the current variable

$p(X_{t+1}|X_t,X_{t-1},\ldots,X_2,X_1)=p(X_{t+1}|X_t)$

joint probability function:

$p(X_1,X_2,\ldots X_n)=p(X_1)\cdot p(X_2|X_1)\cdot p(X_3|X_2,X_1)\cdot\ldots\cdot p(X_n|X_{n-1},X_{n-2},\ldots,X_2,X_1)$.

➔ $p(X_1,X_2,\ldots X_n)=p(X_1)\cdot p(X_2|X_1)\cdot p(X_3|X_2)\cdot p(X_4|X_3)\cdot\ldots\cdot p(X_n|X_{n-1})$

one more assumption: that the transition probabilities do not change with time

$p(X_{t+2}=3 \mid X_t=1) = p(X_{t+3} \mid X_{t+1}=1)$

Examples

- **Discrete Markov chain**

<Transition Matrix>

$$Q = \begin{array}{c|ccccc} & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 0 & .5 & 0 & 0 & .5 \\ 2 & .5 & 0 & .5 & 0 & 0 \\ 3 & 0 & .5 & 0 & .5 & 0 \\ 4 & 0 & 0 & .5 & 0 & .5 \\ 5 & .5 & 0 & 0 & .5 & 0 \end{array}$$

$p(X_{t+2}=3|X_t=1)$

$= \sum_{k=1}^{5} p(X_{t+2}=3|X_{t+1}=k)\cdot p(X_{t+1}=k|X_t=1)$ $= p(X_{t+2}=3 \mid X_{t+1}=2)$

$\times p(X_{t+1}=2 \mid X_t=1)$

<Stationary distribution>

$p(X_{t+h}|X_t)$ => Q^h matrix $\boxed{Q^h}$

If we let h get really large, and take it to the limit, all the rows of the long-range transition matrix will become equal to $(.2,.2,.2,.2,.2)$. That is, if you run the Markov chain for a very long time, the probability that you will end up in any particular state is $1/5=.2$ for each of the five states. These long-range probabilities are equal to what is called the **stationary** distribution of the Markov chain.

가치

As we have just seen, if you simulate a Markov chain for many iterations, the samples can be used as a Monte Carlo sample from the stationary distribution. This is exactly how we are going to use Markov chains for Bayesian inference. In order to simulate from a complicated posterior distribution, we will set up and run a Markov chain whose stationary distribution is the posterior distribution

- **Random walk**