Stat 499: Expectations and Sampling methods

Kai Gong

In this class we mainly focus on the sampling and its methods as following:

Sampling: the selection of a subset (a statistical sample) of individuals from within a statistical population to estimate characteristics of the whole population.

Sampling method:

1. Importance Sampling
    - Approximate E[f] by drawing samples from a“proposal distribution” q, and correcting appropriately by a weighting ratio.
    - Suppose dealing with p(z) is harder, i.e., we can’t even evaluate p(z) but can only do so up to proportionality constant, and only $\tilde{p}(z)$ can be evaluated. We can still apply importance sampling by applying the importance weight.
2. Rejection Sampling
    - Need to set up a proposal function q(z) and M, so that $Mq(z) \geq \tilde{p}(z)$, for all z.
    - Simulate U~Unif(0,1) and candidate X~g from the candidate density.
    - Use $U<\tilde{p}(z)/ Mq(z)$ to test if reject candidate X or not.

Bayesian inference:

- A method of statistical inference in which Bayes' theorem is used to update the probability for a hypothesis as more evidence or information becomes available.
- P(θ|D)=(P(D|θ) X P(θ))/P(D)
- Here, P(θ) is the prior, P(D|θ) is the likelihood of observing our result given our distribution for θ. P(D) is the evidence. P(θ|D) is the posterior belief of our parameters after observing the evidence i.e the number of heads .
- Use P(θ|D) to estimate the probability of θ given the data.

Maximum likelihood estimation:

- A method of estimating the parameters of a probability distribution by maximizing the likelihood function, so that under the assumed statistical model the observed data is most probable.
- In practice, it is often convenient to work with log likelihood.