Lecture 28

Part 6 Introduction to Bayesian Statistics

Bayesian Statistics

Estimation

- Question: How should I estimate θ?
- Answer to the question is another question: What is your loss function?
- First, what is the decision space?
- $\mathcal{D} = (0, 1)$, same as the parameter space.
- $d \in \mathcal{D}$ is a guess about the value of θ.
- The loss function is up to you, but surely the more you are wrong, the more you lose.
- How about squared error loss?
- $L(d, \theta) = k(d - \theta)^2$
- We can omit the proportionality constant k.

Minimize Expected Loss

$L(d, \theta) = (d - \theta)^2$

Denote $E(\theta|X = x)$ by μ. Then

$$\begin{aligned} E(L(d,\theta)|X = x) &= E((d - \theta)^2|X = x) \\ &= E((d - \mu + \mu - \theta)^2|X = x) \\ &= \cdots \\ &= E((d - \mu)^2|X = x) + E((\theta - \mu)^2|X = x) \\ &= (d - \mu)^2 + Var(\theta|X = x) \end{aligned}$$

- Minimal when $d = \mu = E(\theta|X = x)$, the posterior mean.
- This was general.
- The Bayes estimate under squared error loss is the posterior mean.

Back to the example

Give the Bayes estimate of θ under squared error loss.

Posterior distribution of θ is Beta, with $\alpha' = \alpha + \sum_{i=1}^{n} x_i = 61$ and $\beta' = \beta + n - \sum_{i=1}^{n} x_i = 41$.

```
> 61/(61+41)
[1] 0.5980392
```

Hypothesis Testing

$\theta > \frac{1}{2}$ means consumers tend to prefer the new blend of coffee.

Test $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$.

- What is the loss function?
- When you are wrong, you lose.
- Try zero-one loss.

	Loss $L(d_j, \theta)$	
Decision	When $\theta \leq \theta_0$	When $\theta > \theta_0$
$d_0 : \theta \leq \theta_0$	0	1
$d_1 : \theta > \theta_0$	1	0

Compare expected loss for d_0 and d_1

	Loss $L(d_j, \theta)$	
Decision	When $\theta \leq \theta_0$	When $\theta > \theta_0$
$d_0 : \theta \leq \theta_0$	0	1
$d_1 : \theta > \theta_0$	1	0

Note $L(d_0, \theta) = I(\theta > \theta_0)$ and $L(d_1, \theta) = I(\theta \leq \theta_0)$.

$$E(I(\theta > \theta_0)|X = x) = P(\theta > \theta_0|X = x)$$
$$E(I(\theta \leq \theta_0)|X = x) = P(\theta \leq \theta_0|X = x)$$

- Choose the smaller posterior probability of being wrong.
- Equivalently, reject H_0 if $P(H_0|X = x) < \frac{1}{2}$.

Back to the example

Decide between $H_0 : \theta \leq 1/2$ and $H_1 : \theta > 1/2$ under zero-one loss.

Posterior distribution of θ is Beta, with $\alpha' = \alpha + \sum_{i=1}^{n} x_i = 61$ and $\beta' = \beta + n - \sum_{i=1}^{n} x_i = 41$.
Want $P(\theta > \frac{1}{2}|X = x)$

```
> 1 - pbeta(1/2,61,41) # P(theta > theta0|X=x)
[1] 0.976978
```

How much worse is a Type I error?

	Loss $L(d_j, \theta)$	
Decision	When $\theta \leq \theta_0$	When $\theta > \theta_0$
$d_0 : \theta \leq \theta_0$	0	1
$d_1 : \theta > \theta_0$	k	0

To conclude H_1, posterior probability must be at least k times as big as posterior probability of H_0.
$k = 19$ is attractive.
A realistic loss function for the taste test would be more complicated.

Computation

- Inference will be based on the posterior.
- Must be able to calculate $E(g(\theta)|X = x)$
- For example, $E(L(d, \theta)|X = x)$
- Or at least
$$\int L(d, \theta) f(x|\theta)\pi(\theta)\, d\theta.$$
- If θ is of low dimension, numerical integration usually works.
- For high dimension, it can be tough.