STAT 536 : Class 4

Last class : Review

→ Objectives : (i) Integrating
(ii) Optimization.

→ why do we care about (i) and (ii).
(ii) estimation (Maximum Lik. estimates).

Usefulness of (i) : Bayes methods :

Three main ingredients :

(i) Data : $\underline{X \sim f(x|\theta)}$.

(ii) Prior distribution :

$$\theta \sim \pi(\theta)$$

iii) Posterior distribution :

$$\pi(\theta|X) = \frac{f(x|\theta)\,\pi(\theta)}{\int_\theta f(x|\theta)\,\pi(\theta)\,d\theta}$$

Then the Bayes' estimate of θ is defined as the <u>Expected Value</u> of the posterior distribution, i.e.,

$$\hat{\theta} = \cancel{E X} \underbrace{\int_{\theta} \theta \cdot \Pi(\theta | X)\, d\theta}_{\text{integral}}.$$

Recall,

Beta distribution : $X \sim B(a, b)$.

$$f(x) = \begin{cases} \dfrac{1}{B(a,b)} x^{a-1} (1-x)^{b-1}, & 0 < x < 1 \end{cases}$$

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1}\, dx.$$

$\rightarrow$ Beta integral.

$$E(X) = \frac{a}{a+b}.$$

Ex: Suppose $X = (\underline{X_1, \ldots X_n}) \sim Ber(p)$.
$p \in (0,1)$.

We know,

$$f(x|p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$

$$= \underline{p^{\sum x_i}(1-p)^{n-\sum x_i}}$$

$\rightarrow$ density of observed data.

assume : $p \sim \underline{Beta(a,b)}$.
(prior dist).

posterior :-

$\pi(p|X) \propto f(x|p)\, \pi(p)$ Proportionality symbol.

$$\propto p^{\sum x_i} (1-p)^{n-\sum x_i} p^{a-1} (1-p)^{b-1}$$

$$\propto \underbrace{p^{\sum x_i + a - 1} (1-p)^{n-\sum x_i + b - 1}}$$

$$\sim \text{Beta}\left(\sum x_i + a,\ n - \sum x_i + b\right).$$

Then Bayes' estimate of p is

$$\hat{p} = E[p|X] = \frac{\sum x_i + a}{\sum x_i + a + n - \sum x_i + b}$$

$$= \frac{\sum x_i + a}{n + a + b}.$$

Such closed form expressions may not always exist.

In general,

$\cancel{\hat{\phi}}$ $\cancel{E}$

$$\hat{\theta} = E(\theta|x) := \int_{\theta} \theta \, \pi(\theta|x) \, d\theta .$$

integral that needs to be computed numerically.

(1) In the following we shall develop stochastic methods to achieve these objectives of integration and optimization

(2) Such methods will rely on our ability to simulate an endless flow of random variables arising from almost any distribution.

Chapter 2

Random Variable generation:

The ability to simulate r.v.'s from any distribution relies on our ability to simulate from a uniform distribution.

→ we assume there exists a method.

to simulate from uniform distributions.

<u>Note</u>: These are typically deterministic algorithms which reproduce the behavior of uniform r.v.'s.

(Uniform pseudo-random number generators).

(chaotic functions, Mandelbrot sets).

Methods for generating other distributions:

1) Inverse method:

ex: Suppose we have $U \sim U(0,1)$.

Can we transform U into an $Exp(1)$ dist?

Define: $X = -\log(1-U)$. then,

cdf of X:

$$F(x) = P[X \leq x].$$

$$= P[-\log(1-U) \leq x].$$

$$= P[(1-U) \geq e^{-x}].$$

$$= P[U \leq (1-e^{-x})].$$

$$= 1-e^{-x}$$

Recall $X \sim Exp(1)$.

$$f(x) = e^{-x}, \quad x > 0.$$

$$F(x) = \int_0^x e^{-x}\,dx.$$

$$= 1 - e^{-x}$$

where did $-\log(1-U)$ come from?

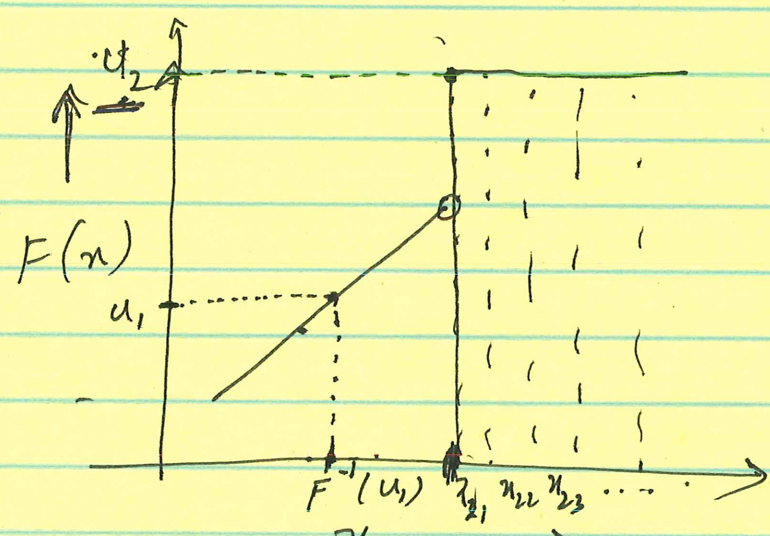
Def: (generalized inverse):

For a non-decreasing function

$$F : \mathbb{R} \to [0,1].$$

the generalized inv. of F is a function F^{-1} defined as,

$$\underline{F^{-1}(u) = \inf\{x : F(x) \geq u\}}.$$



$F(x_{21}) = u_2$

$F(x_{22}) = u_2$

$F(x_{23}) = u_2$

$\vdots$

$F^{-1}(u_2) = \underline{x_{21}}$

Theorem (Probability integral transform):
Let $F(x)$ denote a distribution function (continuous or not). Let $F^{-1}(y)$, $y \in [0,1]$ denote the generalized inverse of F.
Then if $U \sim U(0,1)$, then

$$X = F^{-1}(U) \text{ is distributed as } F,$$

i.e, $P[X \le x] = F(x)$.

In the previous ex:

$$X \sim Exp(\lambda).$$

$$F(x) = 1 - e^{-\lambda x}$$

$$y = 1 - e^{-\lambda x}.$$

$$\Rightarrow \quad x = -\frac{1}{\lambda}\log(1-y).$$