Lecture - 10

Recap:

Poisson random variable,
Geometric random variable,
Negative binomial random variable

---

e.g. Estimate the number of lions in Gir. Assume that the no. of lions is N. You randomly catch & mark m lions. Then you release them. Randomly catch n lions. You count how many of these n lions are marked: let this be a random variable X.

X ranges from 0 to n

$X \in \{0, 1, 2, \cdots, n\}$

$$P(X = i) = \frac{\binom{m}{i}\binom{N-m}{n-i}}{{}^{N}C_{n}}$$

Hypergeometric random variable

N

m are marked

$N - m$ are not marked

$n - i$

i

Estimate N.

Maximum Likelihood Estimation (MLE)

Maximize $$\frac{\binom{m}{i}\binom{N-m}{n-i}}{{}^{N}C_{n}}$$

over N

$$f(N) = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}$$

③

$$f(N-1) = \frac{\binom{m}{i}\binom{N-1-m}{n-i}}{\binom{N-1}{n}}$$

$$f(N) \geq f(N-1)$$

$$\boxed{N \leq \frac{mn}{i}}$$

e.g. $m = 10$ lions

$n = 50$, $i = 5$ are marked

$$N = \frac{mn}{i} = \frac{10 \cdot 50}{5} = 100$$

<u>e.g.</u> Toss 3 dice simultaneously.

$X =$ total on 3 dice

What is $E[X]$?

$X \in \{3, 4, 5, 12, \dots, 18\}$

$P(X=12) =$

<u>Theorem</u>: if $X = X_1 + X_2 + \dots + X_n$, then $E[X] = E[X_1] + \dots + E[X_n]$

$X_i =$ result of i^{th} dice

$E[X_i] = 3.5 = \frac{7}{2}$

$E[X] = 3 \cdot E[X_1] = \frac{21}{2}$.

Cumulative distribution function.

$$F(a) = \sum_{x \le a} p(x)$$

= Probability that x is less than or equal to a

1. It is non-decreasing
   if $a < b$, then
   $$F(a) \le F(b)$$

2. $$\lim_{b \to \infty} F(b) = 1$$

3. $$\lim_{b \to -\infty} F(b) = 0$$

Continuous Random Variables

Discrete
- finite
- countable

Continuous
uncountable

<u>Defn</u>: X is a continuous random variable if there exists a non-negative function f, defined over $\mathbb{R}$, s.t.

set of real numbers

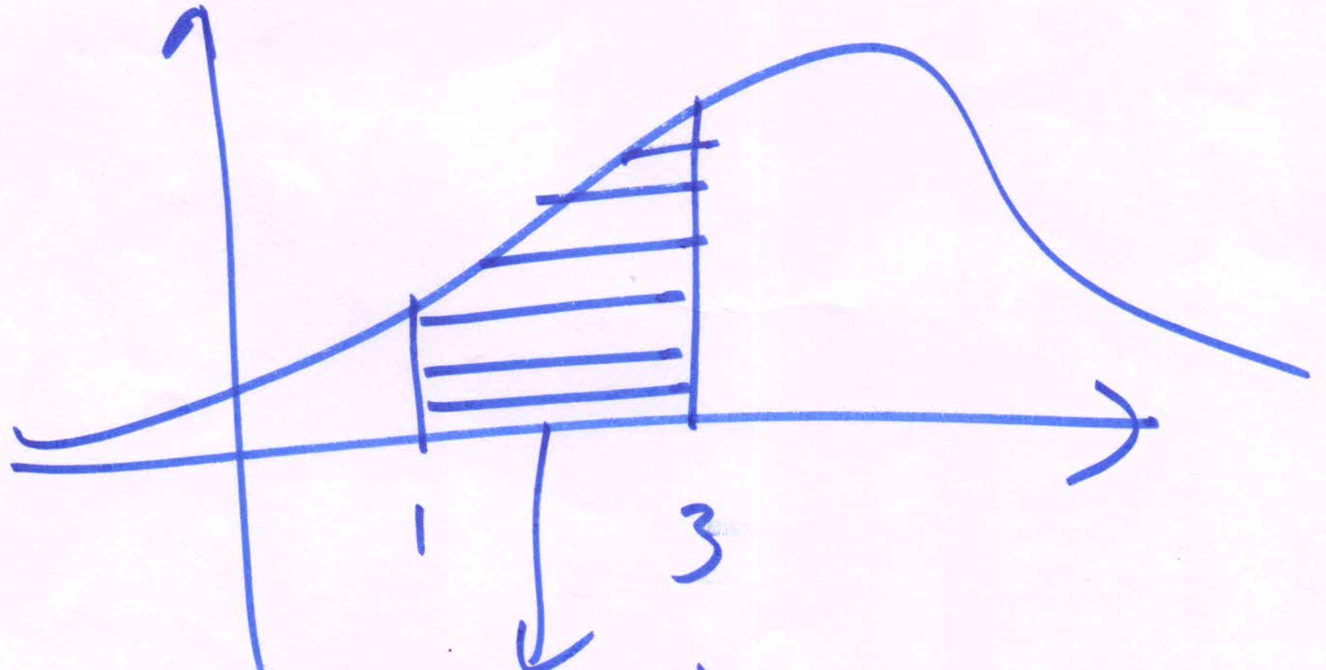
$$P(X \in B) = \int_B f(x)\,dx,$$

$B \subseteq \mathbb{R}$, f is called probability density function e.g.

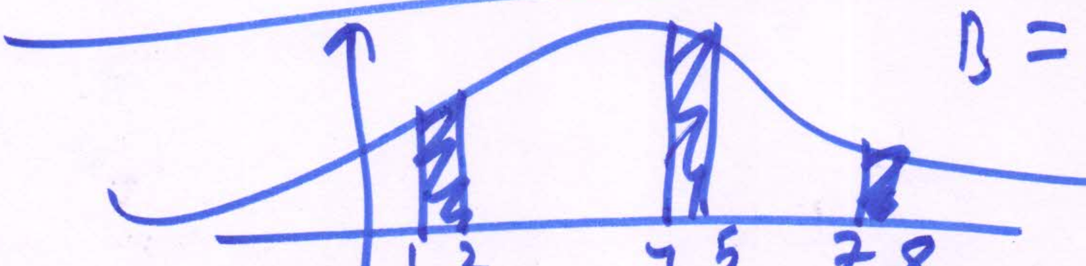
$$B = [a, b]$$

$$B = [1, 3]$$

$$P(X \in B) = \int_1^3 f(x)\,dx$$



$f(X \in B)$



$$B = [1,2] \cup [4,5] \cup [7,8]$$

e.g.: $P(\underset{\substack{\parallel \\ 0}}{X = a}) \neq f(a)$

$$P(X \in B) = \int_B f(x)\,dx$$

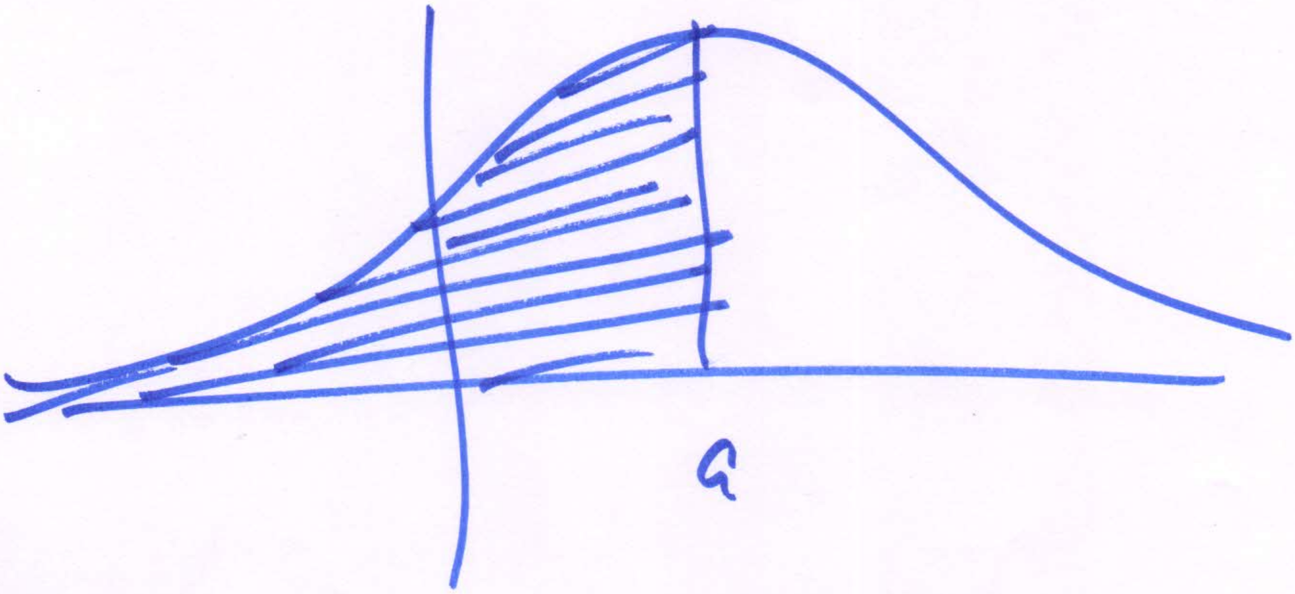
$$B = \{a\}$$

$$= \int_a^a f(x)\,dx = 0$$

$F(a)$ = cumulative probability distribution

$$\begin{gathered} \parallel \\ P(X \leq a) = \int_{-\infty}^{a} f(x)\,d(x) \\ \parallel \\ P(X < a) + \underset{\substack{\parallel \\ 0}}{P(X = a)} \end{gathered}$$

$F(a) =$



e.g. $f(x) = \begin{cases} C(4x - 2x^2) & \text{if } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$

Compute $P(X > 1)$

$= \int_1^\infty f(x)dx = \int_1^2 f(x)dx + \underbrace{\int_2^\infty f(x)dx}_{=0}$

$= \int_1^2 C.(4x - 2x^2)dx =$

Step 1: compute C.

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\underbrace{\int_{-\infty}^{0}}_{=0} + \int_{0}^{2} + \underbrace{\int_{2}^{\infty}}_{=0}$$

$$\int_{0}^{2} C(4x - 2x^2)\,dx = 1$$

$C = ?$

e.g.:

X = no. of hours a computer works before breaking down

$$f(x) = \begin{cases} \lambda e^{-x/100}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

What is the probability that the computer will work between 50 & 150 hours?

Step 1: compute λ.

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$ H.W.