Family of random variables
(distributions)

pmf or pdf

→ class of functions
each function in the class indexed by
a parameter

① Bernoulli: $X \to [0,1]$ $p \in [0,1]$

$p(0) = (1-p)$

$p(1) = p$

② Binomial random with parameters n (positive integer)
$p \in [0,1]$

sum of n independent
Bernoulli random variables

$X \to \{0, 1, \ldots, n\}$

$$P(i) = \binom{n}{i} p^i (1-p)^{n-i}$$

$$\binom{n}{i} = \frac{n!}{i!(n-i)!}$$

(3) Poisson random $X \to 0, 1, 2, \ldots$

$\lambda > 0$

$$P(i) = \frac{e^{-\lambda} \lambda^i}{i!}$$

(4) Uniform distribution $[a, b]$

$$f_X(x) = \begin{cases} 1/_{b-a}, & x \in [a, b] \\ 0, & \text{otherwise} \end{cases}$$

(5) Normal distribution

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$\mu \in \mathbb{R}$

$\sigma \in \mathbb{R}^+$

(6) Exponential distribution $\lambda > 0$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

(7) Chi-squared distribution

n degrees of freedom

$X_1, \ldots, X_n \overset{iid}{\sim} N(0, 1)$

$$\sum_{i=1}^{n} X_i^2 \sim \chi_n^2$$

(8) Student's-T distribution.

$X \sim N(0,1)$

$S \sim \chi_n^2$, X & S are independent

$$T = \frac{X}{\sqrt{S/n}} \sim T_n$$

Central Limit Theorem

$X_1, \dots, X_n \overset{iid}{\sim} \mu, \sigma^2$

$$\bar{X}_n = \frac{1}{n} \sum X_i$$

$$Z_n = \sqrt{n} \cdot \frac{\bar{X}_n - \mu}{\sigma^2} \longrightarrow N(0,1)$$

n large (30?)

$$\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Def (point estimator)

Suppose we have a sample $X_1, \ldots, X_n$

where X_i are iid $f(x \mid \theta_1, \ldots, \theta_k)$

$p(x \mid \theta_1, \ldots, \theta_k)$

a point estimator of $g(\theta_1, \ldots, \theta_n)$

$\phi_n(X_1, \ldots, X_n)$ is a statistic (function of data)

that attempts to estimate $g(\theta_1, \ldots, \theta_n)$.

Ex Suppose $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

$\theta_1 = \mu \quad \theta_2 = \sigma$

suppose $g(\mu, \sigma^2) = \mu$

What would be a reasonable estimator of μ?

$$\phi_n = \frac{1}{n} \sum X_i$$

$$\phi_n = \max\{x_1, \ldots, x_n\} - \min\{x_1, \ldots, x_n\}$$

$$\phi_n = \text{median } X_i$$

$\phi_n = X$

$\varphi_n = \lambda_n$

$g(\mu, \sigma) = \sigma^2$

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$$

$g(\mu, \sigma) = \sigma$

$$S_n = \sqrt{S_n^2}$$

$g(\mu, \sigma) = \mu/\sigma$

$$\frac{\bar{X}_n}{S_n}$$

Ex $X_1, \ldots, X_n \overset{iid}{\sim} \exp(\lambda)$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & o/w \end{cases}$$

$E[X] = 1/\lambda$

$Var[X] = 1/\lambda^2$

What would be a good estimator of λ?

$n / \sum X_i$

$1/\bar{X}_n \qquad \frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}$

What would be good estimator of $1/\lambda$?

$$\frac{1}{n}\sum_{i=1}^{n} X_i$$

<u>Def</u> a point estimator with data $X_1, \dots, X_n \sim f(x \mid \theta_1, \dots, \theta_k)$

$\phi_n(X_1, \dots, X_n)$

is said to be unbiased for $g(\theta_1, \dots, \theta_k)$

$$E[\phi_n(X_1, \dots, X_n)] = g(\theta_1, \dots, \theta_k)$$

An estimator is biased if not.

<u>Ex</u> $X_1, \dots, X_n \sim N(\mu, \sigma^2)$.

$$E\left[\frac{1}{n}\sum X_i\right] = \mu$$

<u>Ex</u> $X_1, \dots, X_n \sim \exp(\lambda)$

estimator for λ is

$$\frac{n}{}$$

$$\hat{p}_n = \frac{}{\sum X_i}$$

$$E[\hat{p}_n] = \frac{n}{n-1} \cdot \lambda$$

$\Longrightarrow$ biased estimator,

can unbias by doing what?

$$E\left[\frac{n-1}{n} \hat{p}_n\right] = \lambda.$$

<u>Def</u> The bias of estimator is defined

as the difference between $E[\hat{p}_n]$

$- g(\theta_1, \ldots, \theta_n)$

$$\frac{n}{n-1}\lambda - \lambda = \lambda\left(\frac{1}{n-1}\right)$$

---

all else being equal, (two estimators have the same bias) we prefer

the estimator that has less variance.

<u>Ex</u> $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$

$\bar{X}_n$ unbiased

but is X_1, which has less variance

$Var(\bar{X}_n) = \frac{\sigma^2}{n}$

$Var(X_1) = \sigma^2$

Def An estimator is asymptotically unbiased if $\lim_{n \to \infty} bias(\phi_n) = 0$

Def An estimator ϕ_n is consistent for $g(\theta_1, \dots, \theta_k)$ if for $\varepsilon > 0$,

$$P(|\phi_n - g(\theta_1, \dots, \theta_k)| > \varepsilon) \to 0$$

as $n \to \infty$

$E[X] = \mu$

$g(\theta_1, \dots, \theta_k) = \mu$

then sample mean is consistent for μ

Maximum Likelihood estimator

$X_1, \dots, X_n \sim f(x \mid \theta_1, \dots, \theta_k)$

I have observed $X_1, \dots, X_n$. I want to find the vals $\theta_1, \dots, \theta_k$ that produce the distribution closest to my data (from this particular parametric family).

$$L(\theta_1, \dots, \theta_k \mid X_1, \dots, X_n)$$

$$= \underline{\prod_{i=1}^{n} f(x_i \mid \theta_1, \dots, \theta_n)}$$

$\Longrightarrow$ find the parameters $\hat{\theta}_1^{MLE}, \dots, \hat{\theta}_k^{MLE}$ that maximize this function

$\Longrightarrow$ $$\ell(\theta_1, \dots, \theta_n \mid X_1, \dots, X_n) = \sum_{i=1}^{n} \log f(x_i \mid \theta_1, \dots, \theta_n)$$

<u>Ex</u> $X_1, \dots, X_n \sim N(\mu, \sigma^2)$

$$L(\mu, \sigma \mid X_1, \dots, X_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{(x_i - \mu)^2}{2\sigma^2}\right)$$

$$l(\mu, \sigma \mid x_1, \ldots, x_n) = -\frac{n}{2} \cancel{\log(2\pi)}$$

$$- n \cancel{\log(\sigma)} - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

$$\frac{\partial l}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) = 0$$

$$\hat{\mu}^{MLE} = \frac{1}{n} \sum x_i$$

MLE for the normal dist'n μ is sample

$$\frac{\partial l}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum (x_i - \mu)^2$$

$$\hat{\sigma}_{MLE} = \sqrt{\frac{1}{n} \sum (x_i - \hat{\mu}_{MLE})^2}$$

<u>Ex</u> Suppose that $X_1, \ldots, X_n \sim \exp(\lambda)$

$$L(\lambda \mid X_1, \ldots, X_n) = \lambda^n \prod_{i=1}^{n} \exp(-\lambda x_i)$$

$$l(\lambda \mid X_1, \ldots, X_n) = n \log(\lambda) - \lambda \sum_{i=1}^{n} x_i$$

$$\frac{\partial l}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i$$

$$\hat{\lambda}_{MLE} = \frac{1}{n}\sum x_i$$

$\underline{Ex}$ Suppose $X_1, \ldots, X_n \sim \text{Unif}[0, \theta]$

↳ what is the MLE of θ?

$$\mathbb{1}_{\{A\}}(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

$$f_X(x|\theta) = \frac{1}{\theta}\mathbb{1}_{\{0<x<\theta\}}(x)$$

$$L(\theta | X_1, \ldots, X_n) = \prod_{i=1}^{n} \frac{1}{\theta} \cdot \mathbb{1}_{\{x \le \theta\}}(x_i)$$

$$= \frac{1}{\theta^n} \prod_{i=1}^{n} \underbrace{\mathbb{1}_{\{0 \le x \le \theta\}}(x_i)}_{L}$$

$$\hat{\theta}_{MLE} = \max_i x_i$$