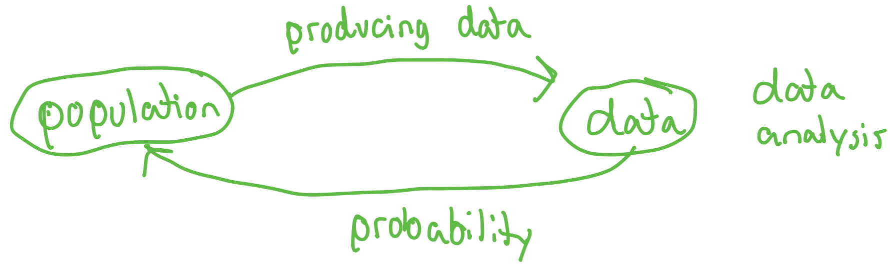
4-20-18 Lesson 32 J.K.

Fitting Models to data

Comparing Estimators

- Methods for fitting Models to
  - Methods of Moments
  - Method of maximum likelihood
  - Method of least squares



- infer of average height of all undergrads at PSU
- collect data from subset of students sample mean $\bar{X}$ to estimate μ

- F (CDF), P (PMF), f (pdf), X (underlying r.v.)
  - used to denote population

- Parameters - are attributes attached a population

e.g. $\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx$

$\sigma^2 = V(x) = \int_{-\infty}^{\infty} [x - E(x)]^2 f x\,dx$

- Sample $\{x_1, \dots, x_n\}$ already seen and fixed

  n: sample size $\{X_1, X_2, \dots X_n\}$ ← they're random unseen

- identically distributed, but they're not necessarily going to give the same number

- Sample statistic → population parameter

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

- height ex. - infer about average height
  $N(\mu, \sigma^2)$ μ_0 : true average height
  - choose most suitable μ from μ_1, μ_2, μ_k
- $X_1, \ldots X_n$ $N(\mu, \sigma^2)$
  $\hat{\mu} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ to estimate μ
  - under normality it is the best in one unite rg
- $\hat{\Theta} = \hat{\theta}(X_1 \ldots X_n)$ estimator ← concepted
  $\hat{\theta} = \hat{\theta}(x_1 \ldots x_n)$ estimate ← observation
- unbiased
  $E_\theta(\hat{\theta}) = \theta$ for all $\theta \in \Theta$
  $\{F_\theta \ \theta \in \Theta\}$
  $\text{bias}_\theta(\hat{\theta}) = E_\theta(\hat{\theta}) - \theta$
- $X_1, \ldots X_n \sim Ber(p)$ $0 < p < 1$
  $\hat{p} := \frac{X_1 + \ldots X_n}{n}$ sample proportion
  unbiased estimator of p
- $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ $E[S]^2 = \sigma^2$
- standard error of estimator
  (standard deviation)
- $\sigma_\theta x^2 = Var_\theta(\hat{\theta})$
  $SE\,\hat{\theta} = \sigma_{\hat{\theta}} = \sqrt{\sigma_{\hat{\theta}}^2}$
  You will most likely have θ in SE_θ replace θ with $\hat{\theta}$
- $X_1, \ldots X_n$ iid $\sim$ Bern(p)
  $\hat{p} = \frac{X_1 + \ldots X_n}{n}$
  $Var(\hat{p}) = \frac{Var(X_1 + \ldots X_n)}{n} = \frac{n Var(X)}{n^2} = \frac{p(1-p)}{n}$

- $SE_{\hat{p}} = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$ $\qquad X_1 \dots X_n$ iid $\sim (\mu, \sigma^2)$

$SE_{\bar{x}}$, $n = 36$, $S = 1.3$

$$Var(\bar{x}) = \frac{n}{n^2} Var(x_i) = \frac{Var(x_i)}{n} = \frac{1.3^2}{36}$$

$$SE_{\bar{x}} = \sqrt{\frac{1.3^2}{36}} = \frac{1.3}{6}$$

- Mean Square Error (MSE) of $\hat{\theta}$ of θ is defined

$$MSE_\theta(\hat{\theta}) = E_\theta[(\hat{\theta} - \theta)^2]$$

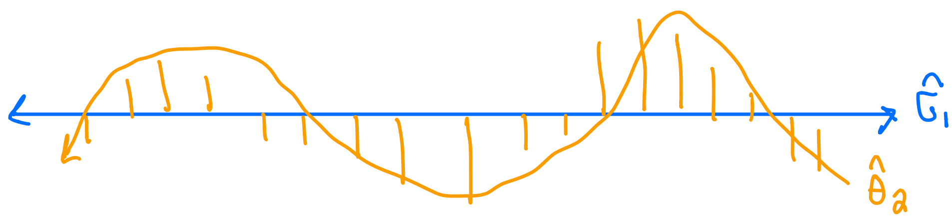
$$= E_\theta[(\hat{\theta} - E[\hat{\theta}] + E[\hat{\theta}] - \theta)^2]$$

$$= E_\theta[(\hat{\theta} - E[\hat{\theta}])^2 + (E[\hat{\theta}] - \theta)^2 + 2(E[\hat{\theta}] - \theta)(\hat{\theta} - E[\hat{\theta}])]$$

$$= E_\theta[(\hat{\theta} - E(\hat{\theta}))^2] + (E[\hat{\theta}] - \theta)^2 + 0$$

$$= Var(\hat{\theta}) + (bias(\hat{\theta}))^2$$

- bias variance trade off



- Ex) $X_1 \dots X_{10}$ iid $\sim (\mu, \sigma^2)$

$Y_1, \dots, Y_{10}$ iid $\sim (\mu, 4\sigma^2)$

$\{X_i\}_{i=1}^{10} \perp\!\!\!\perp \{Y_i\}_{i=1}^{10}$

$\check{\mu} = \delta \bar{X} + (1-\delta)\bar{Y} \qquad 0 \le \delta \le 1$

$$MSE(\hat{\mu}) = E[(\delta \bar{X} + (1+\delta)\bar{Y} - \mu)^2)$$

$$= E[(\delta \bar{x} + (1-\delta)\bar{y} - \delta\mu - (1-\delta)\mu)^2]$$

$$= E[(\delta(\bar{x} - \mu) + (1-\delta)(\bar{Y} - \mu)^2]$$

$$= E[\delta^2(\bar{x} - \mu)^2 + (1-\delta)^2(\bar{y} - \mu)^2 + 2\delta(1-\delta)(\bar{x} - \mu)(\bar{y} - \mu)]$$

$$= \delta^2 E(\bar{x} - \mu) + (1-\delta)^2 E(\bar{y} - \mu) + 2\delta(1-\delta)$$

$$= \delta^2 Var(\bar{x}) + (1-\delta)^2 Var(\bar{Y})$$

$= \delta^2 \frac{\sigma^2}{10} + (1-\delta)^2 \frac{4\sigma^2}{10}$

what δ minimizes this MSE

- Moments

$\mu_k(\theta) = E_\theta(X^k) \qquad k=1,2\ldots$

Sample

$\hat{\mu}_k = \frac{1}{n} \sum_{i=1}^{n} x_i^k \qquad k=1,2\ldots \qquad \theta=(\mu,\sigma^2)$

- Method of Moments, estimate θ from $\{X_1, \ldots X_n\}$

$\mu_k(\theta) = \hat{\mu}_k \qquad k=1,2$

- θ is the unkown solution for this we call the moment estimator $\hat{\theta}_{mom}$ of θ

Variant of MOM

$\bar{X} \qquad \mu(\theta)$

$s^2 \qquad \sigma^2(\theta)$

$\mu(\sigma) = E_\theta(X) = \bar{X}$

$\sigma^2(\theta) = Var_\theta(X) = s^2$