Lecture 4 - Minimal, Complete, Ancillary, Estimation and Method of Moments

Defn: Statistic

If $X_n \overset{iid}{\sim} f_\theta$ where $\theta \in \Theta$ is unknown.

A statistic T is a function of $X_1, \dots, X_N$ whose formula doesn't depend on θ.

Ex. $X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ where $\mu \in \mathbb{R}$

$T = \bar{X}$ is a statistic

but $\rightarrow T = \mu$ isn't a statistic

neither is $T = \bar{X} - \mu$.

Defn: Ancillary Quantity

Ancillary quantity is basically the opposite of a good statistic.

An ancillary quantity is a function of $\underset{\sim}{X}$

$$Q = Q(\underset{\sim}{X})$$

whose distribution doesn't depend on the unknown parameter.

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

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

then $\bar{X} \sim N(\mu, \sigma^2/N)$

so

$$Q = \frac{\bar{X} - \mu}{\sigma/\sqrt{N}}$$ is ancillary to μ and σ^2.

$Q \sim N(0,1)$.

not a statistic b/c it depends on μ, σ^2

Defn: Ancillary Statistic

T is ancillary statistic if it is an ancillary quantity and also a statistic.

T doesn't involve θ in its formula nor in its distribution.

depends on model and what is unknown

Ex. Choice:

$X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ w/ μ unknown but σ^2 known

" or μ known, σ^2 unknown

Ex. $X_n \overset{iid}{\sim} N(\theta, 1)$

and let

$$R = X_{(N)} - X_{(1)}.$$

this is a statistic

Q: is R ancillary to θ?

Notice: $X_n = Z_n + \theta$ where $Z_n \overset{iid}{\sim} N(0,1)$

Similarly

$$X_{(N)} = Z_{(N)} + \theta$$

$$X_{(1)} = Z_{(1)} + \theta$$

hence

$$R = X_{(N)} - X_{(1)} = Z_{(N)} + \theta - (Z_{(1)} + \theta)$$

$$= Z_{(N)} - Z_{(1)}$$

$Z_n \overset{iid}{\sim} N(0,1)$ so they don't depend on θ

Hence R is ancilly to θ.

Concept: Minimal Sufficient Statistic.

Ex. $X_n \overset{iid}{\sim} N(\mu, 1)$ we showed $T = \overline{X}$ sufficient for μ.

What about $T = (\overline{X}, \overline{X}^3)$?

Yes. I can always add an information.

Seems wasteful. Formalize: minimal sufficient statistic.

sufficient but

no extra info.

Concept: Complete Statistic

T is complete if I can't make a ancillary statistic from T.

Deep: connected uniqueness/identifiability.

Basically: the Stat T is rich enough to figure out θ / separate different values of θ.

Theorem: (Don't need to worry too much)

If $X_n \overset{iid}{\sim} f_\theta$ (one parameter $\theta \in \mathbb{R}$)

$$f_\theta(\underline{x}) = h(\underline{x})\, c(\theta) \exp\left(T(\underline{x})\, w(\theta)\right)$$

then T is

1. sufficient for θ
2. minimal sufficient for θ
3. complete.

Theorem: Basu's Theorem

Alluded:
→ Sufficient contains all info about θ
→ ancillary contains no info about θ

→ ancillary contains no info about θ
opposites!

If T is a sufficient statistic and S is an ancillary statistic then $T \perp\!\!\!\perp S$.

Theorem: $X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ then $\bar{X} \perp\!\!\!\perp S^2$.

pf. $\bar{X}$ is sufficient for μ,
and $\frac{N-1}{\sigma^2} S^2 \sim \chi^2(N-1) \iff S^2 \sim \frac{\sigma^2}{N-1}\chi^2(N-1)$

doesn't depend on μ

hence S^2 is ancillary to μ.
Hence $\bar{X} \perp\!\!\!\perp S^2$.

Point Estimation

Setup: $X_n \overset{iid}{\sim} f_\theta$ where $\theta \in \Theta$.

Defn: A point estimator of θ is just a statistic

$$\hat{\theta} = \hat{\theta}(\underline{X})$$

pedantic terminology

$\hat{\theta} : \mathbb{R}^N \to \mathbb{R}^d$

typically $a=1$

and $\hat{\theta}(\underset{\sim}{X})$ is called an estimator (random)

$\hat{\theta}(\underset{\sim}{x})$ is called an estimate (not random).

Q: (1) how do I construct estimates

(2) how do I know if (1) is good?

Our first method for constructing estimators is called the Method of Moments.

Defn: the r^{th} moment of a RV X is

$$\mu_r \overset{\text{def}}{=} \mathbb{E}[X^r]$$

the r^{th} sample moment of a RS $X_1, \dots, X_N$ is

$$m_r \overset{\text{def}}{=} \frac{1}{N}\sum_{n=1}^{N} X_n^r.$$

Ex. $\mu_1 = \mathbb{E}[X]$ and $m_1 = \overline{X}$

$\mu_2 = \mathbb{E}[X^2]$ and $m_2 = \frac{1}{N}\sum X^2$

$$\mu_2 = \mathbb{E}[X^2] \quad \text{and} \quad m_2 = \frac{1}{N}\sum X_n^2$$

$$\vdots \qquad\qquad \vdots$$

Notice:

$$\mathbb{E}[m_r] = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N} X_n^r\right]$$

$$= \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[X_n^r]$$

$$= \frac{1}{N}\sum_{n=1}^{N} \mu_r$$

$$= \mu_r$$

on avg. $m_r \approx \mu_r$

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

I want estimates for μ and σ^2.

MoM:

1) Calculate first two moments of $\cancel{N}(\mu, \sigma^2)$ $X \sim$

$$\mu_1 = \mathbb{E}[X] = \mu$$

$$\mu_2 = \mathbb{E}[X^2] = Var(X) + \mathbb{E}[X]^2 = \sigma^2 + \mu^2$$

$$\left(Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2\right)$$

2) Say that approx.

(2) Say that approx.

$\mu_1 \approx m_1$ and $\mu_2 = m_2$

$$\boxed{\hat{\mu} = \overline{X}}$$

$$\mu^2 + \sigma^2 = \frac{1}{N}\sum_{n=1}^{N} X_n^2$$

$$\sigma^2 = \frac{1}{N}\sum_{n=1}^{N} X_n^2 - \mu^2$$

$$= \frac{1}{N}\sum_{n=1}^{N} X_n^2 - \overline{X}^2$$

$$\boxed{\hat{\sigma}^2 = \frac{1}{N}\sum_{n=1}^{N} (X_n - \overline{X})^2}$$

looks a lot like S_{N-1}^2

Method of Moments:

$X_n \overset{iid}{\sim} f_\theta$ where $\theta = (\theta_1, \theta_2, \ldots, \theta_K) \in \Theta \subset \mathbb{R}^K$

the population moments of f_θ are

$\mu_1, \mu_2, \ldots, \mu_K$

and the corresp. sample moments are

$m_1, \ldots, m_K$

and assuming $m_1 \approx \mu_1$, $m_2 \approx \mu_2$, $\ldots$, $m_K \approx \mu_K$ then and assume $m_1 \approx \mu_1, m_2 \approx \mu_2, \ldots, m_K \approx \mu_K$ then we form the set of K equations in K unknowns

ms are functions of Xs

$$\begin{cases} m_1 = \mu_1 = f_1(\theta_1, \theta_2, \ldots, \theta_K) \\ m_2 = \mu_2 = f_2(\theta_1, \theta_2, \ldots, \theta_K) \\ \vdots \\ m_K = \mu_K = f_K(\theta_1, \ldots, \theta_K) \end{cases}$$

μs are functions of the θs

we solve this set of eqns. for the θs in terms of Xs to get MoM estimates $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_K$.

Ex. $X_n \overset{iid}{\sim} Bin(k, p)$

Lets find MoM estimators for k and p.

(1) Population moments $X \sim Bin(k,p)$

$$\mu_1 = E[X] = kp$$

$$\mu_2 = E[X^2] = Var(X) + E[X]^2$$
$$= kp(1-p) + k^2p^2$$

(2) Form equations

(2) form equations

(1) $\overline{X} = m_1 = \mu_1 = kp$

(2) $\overline{X^2} = \frac{1}{N}\sum_{n=1}^{N} X_n^2 = \mu_2 = \underbrace{kp(1-p) + k^2p^2}$

(3) Solve for k and p.

$\rightarrow \overline{X}^2 = (kp)^2 = k^2p^2$

So $\overline{X^2} - \overline{X}^2 = kp(1-p)$

hence $\dfrac{\overline{X^2} - \overline{X}^2}{\overline{X}} = \dfrac{kp(1-p)}{kp} = 1-p$

hence $\boxed{\hat{p} = 1 - \dfrac{\overline{X^2} - \overline{X}^2}{\overline{X}}}$

and since $\overline{X} = kp$ then $\boxed{\hat{k} = \dfrac{\overline{X}}{\hat{p}}}$

Note: these estimates are not great.

Notice if $m_2 - m_1^2 > m_1$ then $\hat{p} < 0$.

Ex. $X_n \overset{iid}{\sim} \text{Poisson}(\lambda)$

MoM estmate for λ?

① $\mu_1 = E[\text{Poisson}(\lambda)] = \lambda$

$m_1 = \overline{X}$

② Form equation $\overline{X} = \lambda$

③ Solve for λ . . . $\boxed{\hat{\lambda} = \overline{X}}$

Variation of this...

① $Var(\text{Pois}(\lambda)) = \lambda$

set this equal to S^2

hence a MoM estimator is $\boxed{\hat{\lambda} = S^2}$

Ex. $X_n \overset{iid}{\sim} U(0,\theta)$

Way 1

① $\mu_1 = E[X_n] = \theta/2$

② $m_1 = \overline{X} = \theta/2 = \mu_1$

③ $\boxed{\hat{\theta} = 2\overline{X}}$

① $Var(X) = \dfrac{\theta^2}{12}$

② $S^2 \approx Var(X) = \dfrac{\theta^2}{12}$

③ $\hat{\theta} = \sqrt{12 S^2}$

Ex. $X_n \overset{iid}{\sim} \text{Beta}(\alpha, 1)$ — $\mathbb{E}[\text{Beta}] = \dfrac{\alpha}{\alpha+\beta}$

$\beta = 1.$

(1) $\mu_1 = \dfrac{\alpha}{\alpha+1}$

(2) $m_1 = \bar{X} = \mu_1 = \dfrac{\alpha}{\alpha+1}$

(3) Solve for α ...

we find $\boxed{\hat{\alpha} = \dfrac{\bar{X}}{1+\bar{X}}}$