Today

Statistical Inference.

$f_X(x;\theta)$ — prob. model.

θ → parameters of the model

Estimated from observations

Classical methods. (frequentialist)

Treat θ as unknown vector/scalar. (constant)

- Point Estimation
- Interval Estimation
- MLE (Maxm-likelihood)

Bayesian methods

* θ is random.
* make assumption on the distribution of θ

- MMSE - estimator.
- L-MMSE - "
- MAP - estimator.

---

① Point Estimation

$X_1, X_2, \ldots, X_n$ → iid outcomes.

Goal: Estimate the mean

$E[X_i] = \mu$

(We do not know μ)

$$\hat{\mu}_{(n)} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

$M_{(n)}$ — Sample mean. → this is an estimate of the unknown mean μ.

Sample mean is a Random Variable.

(Desired) Properties of a Point Estimate

$\hat{\gamma}$ is an estimate of γ

(1) Unbiased Estimator

$$E[\hat{\gamma}] = \gamma$$

(Bias $= 0$). Bias $= (E[\hat{\gamma}] - \gamma)$

is $\hat{\mu}_n$ / $M_{(n)}$ an unbiased estimate of μ ?

is $E[\hat{\mu}_n] \overset{?}{=} \mu$

$$E\left[\frac{X_1 + \dots + X_n}{n}\right] = \frac{n\mu}{n} = \mu. \checkmark$$

Yes

$\hat{\mu}^{(my)} = X_1$ Yes this is also unbiased.

$E[\hat{\mu}^{(my)}] = E[X_1] = \mu$ $\checkmark$

(2) $\boxed{\text{Consistency of an Estimator}}$

$\gamma \rightsquigarrow \hat{R}_1, \hat{R}_2, \dots, \hat{R}_n$ (Sequence) of estimator)

The seq. is consistent if for any $\varepsilon > 0$

$$\lim_{n \to \infty} P(|\hat{R}_n - \gamma| \geq \varepsilon) = 0$$

$$\hat{R}_n \xrightarrow[\text{as } n \to \infty]{\text{(in Prob)}} \gamma$$

③ MSE of estimate $\hat{R}$

$$\text{error} = E[(\hat{R}-r)^2].$$

if a sequence of estimators $\hat{R}_1, \hat{R}_2, \ldots \hat{R}_n$,

$$e_n = E[(\hat{R}_n - r)^2] = Var(\hat{R}_n).$$

if $\hat{R}_n$ is also an unbiased estimate of r.

$$E[\hat{R}_n] = r.$$

if $\boxed{\lim_{n\to\infty} e_n = 0}$;

then the sequence of estimators $\hat{R}_n$ is consistent.

$$\hat{R}_n \xrightarrow[n\to\infty]{\text{in M.S.}} r$$

$$\begin{tikzcd}
\boxed{A.S.} \arrow[dr] & & \\
& \boxed{P} \arrow[r] & \boxed{D} \\
\boxed{M.S.} \arrow[ur] & &
\end{tikzcd}$$

Sample mean. $\overline{X}_{(n)} = \dfrac{X_1 + X_2 \cdots + X_n}{n}$.

① Unbiased ✓.

② Consistent ?

$$\underbrace{E\left[\left(\overline{X}_{(n)} - \mu\right)^2\right]}_{} \stackrel{?}{=} Var\left(\overline{X}_{(n)}\right)$$

$$= \frac{\sigma^2}{n} \xrightarrow[]{\text{as } n \to \infty.} 0$$

$$\Rightarrow \overline{X}_{(n)} \xrightarrow[\text{as } n \to \infty.]{\text{in M.S.}} \mu$$

$$\Rightarrow \overline{X}_{(n)} \xrightarrow[\text{as } n \to \infty.]{\text{in Prob.}} \mu$$

$\Rightarrow$ $\overline{X}_{(n)}$ is a <u>consistent estimate</u> of μ.

$$\hat{\mu}^{(my)}_{(n)} = X_1$$

MSE of this estimator $\stackrel{?}{=} E\left[(X_1 - \mu)^2\right]$

$$= \underbrace{\sigma^2}_{\text{const.}} \cancel{\xrightarrow[n \to \infty.]{\text{as}} 0}$$

Estimating Variance

$X_1, X_2, \ldots, X_n$ iid.

(1) Known mean, unknown variance.

(2) Unknown mean, Unknown Variance.

We know μ.

Goal: estimate σ^2.

$$\sigma^2 = E\left[(X-\mu)^2\right]$$

R.V.

$$\hat{\sigma}^2_{(n)} = \frac{(X_1-\mu)^2 + (X_2-\mu)^2 + \ldots + (X_n-\mu)^2}{n}.$$

$$\boxed{\hat{\sigma}^2_{(n)} = \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{n}.}$$

① Unbiased (?) $\quad E\left[\hat{\sigma}^{(2)}_{(n)}\right] \stackrel{?}{=} \sigma^2$

$$E\left[\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{n}\right] \stackrel{?}{=} \sigma^2.$$

$$\frac{\sum_{i=1}^{n} E[(X_i-\mu)^2]}{n} = \frac{\cancel{n}\,\sigma^2}{\cancel{n}} = \sigma^2$$

② Consistent (?).

$$\underbrace{P\left(\left| \hat{\sigma}^{(2)}_{(n)} - \sigma^2 \right| \geq \varepsilon \right)} \xrightarrow[\text{as } n \to \infty]{} 0$$

↓

Chebychev's inequality

$\boxed{\text{Homework}}$ → DIY

Unknown mean, Unknown Variance.

↓

Estimate of mean $\overset{?}{=}$ $\bar{X} = \dfrac{x_1 + \cdots + x_n}{n}$

(sample mean)

X_i's are iid

$$\underbrace{\hat{\sigma}^{(2)}_{(n)}} = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \bar{X} \right)^2$$

① Is the above estimator unbiased ?

$$E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\right] \overset{?}{=} \sigma^2.$$

$$LHS = E\left[\frac{1}{n}\sum_{i=1}^{n}\left(X_i^2+\bar{X}^2-2X_i\bar{X}\right)\right].$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left[\underbrace{E[X_i^2]}+\underbrace{E[\bar{X}^2]}-2E[X_i\bar{X}]\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left[(\sigma^2+\mu^2)+\frac{\sigma^2}{n}+\mu^2-2E[X_i\bar{X}].\right]$$

$$Var(A) = E[A^2]-(E[A])^2.$$

$$Var(\bar{X}) = E[\bar{X}^2] - (E[\bar{X}])^2.$$

$$E[\bar{X}^2] = \underbrace{Var(\bar{X})} + \underbrace{(E(\bar{X}))^2}.$$

$$= \sigma^2/n. + \mu^2$$

$$\left(\bar{X} = \frac{X_1+\cdots+X_n}{n}\right)$$

$$\boxed{E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\right]=\left(\frac{n-1}{n}\right)\sigma^2}$$

$$\Downarrow$$

$$E\left[\hat{\sigma}^{(2)}_{(n)}\right].$$

$\boxed{\text{Biased Estimate of } \sigma^2}$

As $n \to \infty$.

Bias $= E\left[\hat{\sigma}^{(2)}_{(n)}\right] - \sigma^2$

$$= \frac{(n-1)}{n}\sigma^2 - \sigma^2$$

$$= -\frac{\sigma^2}{n} \xrightarrow[n\to\infty]{\text{as}} 0$$

<u>Asymptotically</u> unbiased estimate

$$E\left[\frac{1}{\cancel{n}}\sum_{i=1}^{n}(X_i-\bar{X})^2\right]=\left(\frac{n-1}{\cancel{n}}\right)\sigma^2.$$

divide by $(n-1)$

$$E\left[\frac{1}{(n-1)}\sum_{i=1}^{n}(X_i-\bar{X})^2\right] = \sigma^2.$$

unbiased

$$\hat{\sigma}^{2,(new)}_{(n)} = \frac{1}{(n-1)}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

Sample Variance

② Check if this is a consistent estimator