SUMMARY

1.Properties of an estimator:

- Let $X_1, X_2, ..., X_n \sim f_\theta$ such that θ is a parameter.
- Let $\theta = g(X_1, X_2, ..., X_n)$ is an estimator of θ.
- We have some properties θ as following:
    1. θ is an unbiased estimator of θ if $E(\theta) = \theta$
    2. θ is an efficient estimator of θ if $V(\theta) \xrightarrow{n \to \infty} 0$
    3. Cramer-Rao Lower Bound (CRLB): gives the lower estimate for the variance of an unbiased estimator

       We use only one density function

       $$V(\theta) \geq \frac{1}{nI(\theta)} \quad \text{with} \quad I(\theta) = E\left[\left[(\ln(f_\theta))'\right]^2\right] = -E\left[(\ln(f_\theta))''\right]$$

       We use likelihood function

       Let $l(\theta) = \ln\left(\prod_{i=1}^{n} f_\theta(x_i)\right)$, then

       $$V(\theta) \geq \frac{1}{I(\theta)} \quad \text{with} \quad I(\theta) = E\left[(l'(\theta))^2\right] = -E\left[l''(\theta)\right]$$

    4. θ is a consistent estimator of θ if $\theta \xrightarrow[n \to \infty]{P} \theta$ $\lim_{n \to \infty} \hat{\theta} = \theta$
    5. Central Limit Theorem

       $$\theta \xrightarrow[n \to \infty]{L} N\left(E(\theta), V(\theta)\right)$$

II. Maximum Likelihood Estimator(MLE)

- Let $X_1, X_2, ..., X_n \sim f_\theta$ such that θ is a parameter.
- Let $l(\theta) = \ln\left(\prod_{i=1}^{n} f_\theta(x_i)\right)$
- MLE, $\theta_{MLE} = \theta = \underset{\theta \in \Theta \subset IR}{Arg \max}\, l(\theta)$

Example1.

1. Let $X_1, X_2, ..., X_n \sim Ber(p)$, then $p = \overline{X}$
2. Let $X_1, X_2, ..., X_n \sim Poi(\lambda)$, then $\lambda = \overline{X}$

Unbiased.

$E(\hat{p}) = p$

efficiency

$V(\hat{p}) = \frac{p(1-p)}{n} \xrightarrow{n \to \infty} 0$

$E(\hat{\mu}) = \mu$, $V(\hat{\mu}) = \frac{\sigma^2}{n}$. (efficiency)

unbiased

3. Let $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$, then $\hat{\mu} = \overline{X}$, $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$

Example2. Calculate the Expected value and Variance of the estimator is Example1.

- Properties of MLE
  1. MLE(θ) is an unbiased, efficient and consistent estimator of θ
  2. MLE(θ) reaches(more or less) the CRLB
  3. CLT
     a. Let θ be the p (in Bernoulli), λ (in Poisson) or μ (in Normal), then

$$\hat{\theta} \xrightarrow[n\to\infty]{L} N\left(E(\hat{\theta}), V(\hat{\theta})\right)$$

$\hat{p} \rightarrow N(\underbrace{E(\hat{p})}_{p}, \underbrace{V(\hat{p})}_{\frac{p(1-p)}{n}})$

$\hat{p} \rightarrow N\left(p, \frac{p(1-p)}{n}\right)$

     b. CLT for σ^2

The MLE of σ^2 is $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ which is biased estimator.

$\overline{X} = \frac{1}{n}\sum x_i$

$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$

Let $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$, then S^2 is an unbiased estimator of σ^2.

We have

$$(n-1)\frac{S^2}{\sigma^2} \sim \chi^2(n-1)$$

$\hat{\mu} \rightarrow N(\underbrace{E(\hat{\mu})}_{\mu}, \underbrace{V(\hat{\mu})}_{\frac{\sigma^2}{n}})$

Remark: We use CTL to construct the Confidence Interval.

III. Confidence Interval

$Z = \frac{\hat{\mu} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \rightarrow N(0,1)$

$\boxed{\hat{\mu} \rightarrow N\left(\mu, \frac{\sigma^2}{n}\right)}$

- Suppose that α is known, e.g $\alpha = 5\%$. It is called Confidence Level.
- Definition: given $a, b \in IR$ with $a < b$, then $[a, b]$ is called (1-α) confidence interval for parameter θ if $P(a \leq \theta \leq b) = 1 - \alpha$.

$Z \sim N(0,1)$

$X \sim N(\mu, \sigma^2)$

1. Confidence Interval for means with known σ^2

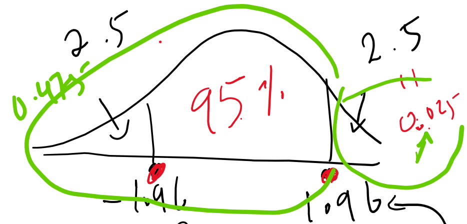
- Suppose that $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$ such that σ^2 is known. Our aim: to estimate μ.
- Point estimator for μ is $\hat{\mu} = \overline{X}$.
- We want now to construct the Confidence Interval (Interval Estimation) for μ.
- We know that $E\left(\overline{X}\right) = \mu, V\left(\overline{X}\right) = \frac{\sigma^2}{n}$
- CLT, $\hat{\mu} \xrightarrow[n\to\infty]{L} N\left(\mu, \frac{\sigma^2}{n}\right)$

$Z = \frac{X - \mu}{\sqrt{\sigma^2}}$

$-\sqrt{1296}$

- 95% confidence interval for parameter μ is $\left[\overline{X} \pm 1.96\frac{\sigma}{\sqrt{n}}\right]$

- General: (1-α) confidence interval for parameter μ is $\left[\overline{X} \pm Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$



- Example:

  a. Let X equal the length of life of a 60-watt light bulb marketed by a certain manufacturer. Assume that the distribution of X is $N(\mu, 1296)$. If a random sample of $n = 27$ bulbs is tested until they burn out, yielding a sample mean of $\bar{x} = 1478$ hours.

  $P(Z > \lambda) = 0.025$

  $\lambda = 1.96$

  $\bar{X}$

  a.1 What is the point estimator for μ.

  a a.2 Construct a95% confidence interval for μ $\rightarrow 1478 \pm 1.96 \frac{\sqrt{1296}}{\sqrt{27}}$

  .3 Construct a90% confidence interval for μ $\rightarrow 1478 \pm 1.645 \frac{\sqrt{1296}}{\sqrt{27}}$

  a.4 Interpret the result in a.3

  b. Lake Macatawa, an inlet lake on the east side of Lake Michigan, is divided into an east basin and a west basin. To measure the effect on the lake of salting city streets in the winter, students took 32 samples of water from the west basin and measured the amount of sodium in parts per million in order to make a statistical inference about the unknown mean μ. They obtained the following data:

13.0	18.5	16.4	14.8	19.4	17.3	23.2	24.9
20.8	19.3	18.8	23.1	15.2	19.9	19.1	18.1
25.1	16.8	20.4	17.4	25.2	23.1	15.3	19.4
16.0	21.7	15.2	21.3	21.5	16.8	15.6	17.6

  b.1 Construct a95% confidence interval for μ if $\sigma^2 = 9$ $\boxed{\bar{X} \pm 1.96 \frac{\sigma}{\sqrt{n}}}$

  b.2 Construct a95% confidence interval for μ if σ^2 is unknown $\boxed{\bar{X} \pm 1.96 \frac{S}{\sqrt{n}}}$ $n > 30$

2. Confidence Interval for means with unknown σ^2

   We use Student T distribution

3. Confidence Interval for σ^2

   We use Chi-Square distribution

σ^2 unknow but $n > 30 \rightarrow Z(0,1)$

σ^2 ______ and $n < 30 \rightarrow$ Student distribution

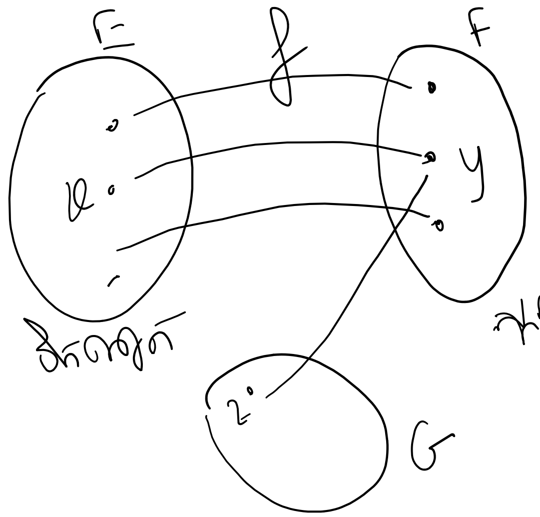
I. One-variable function $y = x^3 - 3x$ (critical value)

only one x

- step1: $y' = 3x^2 - 3$, $x_1 = -1$, $x_2 = +1$
- step2: $y'' = 6x$

$y''(-1) = -6 < 0 \Rightarrow$ max

$y''(+1) = +6 > 0 \Rightarrow$ min.



$y = G(x, z)$

II. 2-variable function: $z = f(x,y) = x^2 + y^2$

step1: z'_x & z'_y $\rightarrow \begin{cases} z'_x = 0 \\ z'_y = 0 \end{cases} \rightarrow (x_0, y_0)$

$z = x^2 + y^2$ $\quad z'_x = 2x$ $\quad z'_y = 2y$ $\quad \begin{cases} 2x = 0 \\ 2y = 0 \end{cases} \rightarrow (x,y) = (0,0)$

Step2

$z'_x \rightarrow z''_{xx}, \; z''_{xy}$

$z'_y \rightarrow z''_{yx}, \; z''_{yy}$ ← 4: Second order derivative

let $\Delta = \left(z''_{xx} z''_{yy}\right) - \left(z''_{xy} z''_{yx}\right)$

$\Delta > 0$ & $z''_{xx} > 0 \Rightarrow$ min.
$\Delta > 0$ & $z''_{xx} < 0 \Rightarrow$ max

$$\boxed{l(\mu, \sigma^2) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln \sigma^2 - \frac{1}{2\sigma^2}\sum (x_i - \mu)^2}$$

$$l'_\mu = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu), \quad l'_{\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum (x_i - \mu)^2$$

$$\begin{cases} l'_\mu = 0 \\ l'_{\sigma^2} = 0 \end{cases} \Rightarrow \boxed{\hat{\mu} = \bar{x}, \quad \hat{\sigma}^2 = \frac{1}{n}\sum (x_i - \bar{x})^2}$$

$$l'_\mu = \frac{1}{\sigma^2}\sum (x_i - \mu) \quad \begin{cases} \boxed{l''_{\mu\mu} = -\frac{n}{\sigma^2}} \\ l''_{\mu\sigma^2} = -\frac{1}{\sigma^4}\sum_{i=1}^{n}(x_i - \mu). \end{cases}$$

$$l'_{\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\Sigma(x_i-\mu)^2$$

$$l''_{\sigma^2\mu} = -\frac{1}{\sigma^4}\Sigma(x_i-\mu)$$

$$l''_{\sigma^2\sigma^2} = \boxed{+\frac{n}{2\sigma^4} + \frac{(-2)}{2\sigma^6}\Sigma(x_i-\mu)^2}$$

Bernoulli $\hat{p} = \bar{x}$, $E(\bar{x}) = ?$ $V(\bar{x}) = ?$

Normal $\hat{\mu} = \bar{x}$, $E(\bar{x})$? $V(\bar{x}) = ?$

$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$, $E(\hat{\sigma}^2)$, $V(\hat{\sigma}^2)$?

Bernoulli: $X_1, \dots, X_n \sim Ber(p)$ $E(X) = p$

$V(X) = p(1-p)$

$E(ax) = aE(x)$

$E[X+Y] = E(X) + E(Y)$

$$E(\hat{p}) = E(\bar{x}) = E\left(\frac{1}{n}(x_1+x_2+\dots+x_n)\right)$$
$$= \frac{1}{n}E[x_1+x_2+\dots+x_n]$$
$$= \frac{1}{n}\left(E(x_1)+E(x_2)+\dots+E(x_n)\right)$$
$$= \frac{1}{n}(p+p+\dots+p) = \frac{np}{n} = p$$

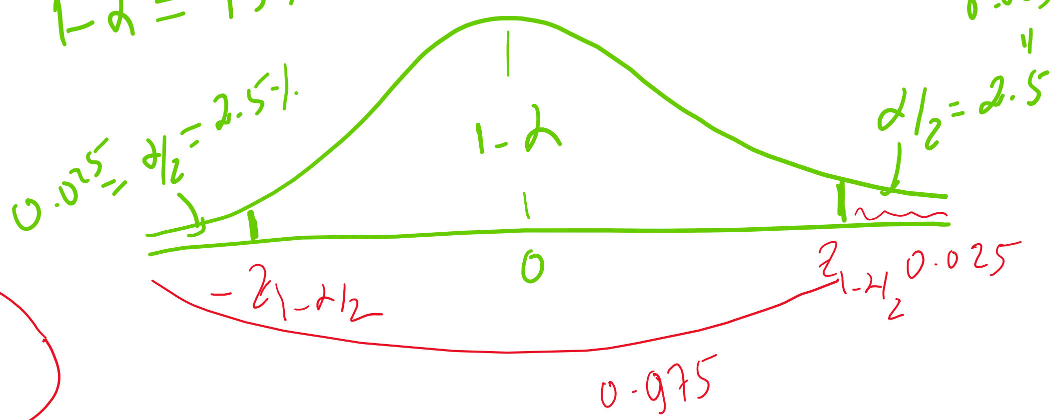
$E(\hat{p}) = p \Rightarrow \hat{p}$ is an unbiased estimator.

$V(\hat{p}) =$

CLT for $\hat{\mu}$: $\hat{\mu} \longrightarrow N\left(\mu, \frac{\sigma^2}{n}\right)$

$$z = \frac{\hat{\mu}-\mu}{\sqrt{\frac{\sigma^2}{n}}} = \sqrt{n}\left(\frac{\bar{X}-\mu}{\sigma}\right) \longrightarrow N(0,1)$$

$\alpha = 5\%$ $\quad 1-\alpha = 95\%$



$P(z < \lambda) = 0.025$

$P(z < \lambda) = 0.975 \Rightarrow \lambda = 1.96$

$$z_{1-\alpha/2} = 1.96, \quad -z_{1-\alpha/2} = 1.96.$$

$$P(-1.96 < z < 1.96) = 0.95$$

$$P\left(-1.96 < \sqrt{n}\frac{(\bar{x}-\mu)}{\sigma} < 1.96\right) = 0.95$$

$$P(-1.96\sigma < \sqrt{n}(\bar{x}-\mu) < 1.96\sigma) = 0.95$$

$$P\left(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

$$P(a < \mu < b) = 0.95$$

$\Rightarrow (1-\alpha)$ CI for μ is $\left[\bar{x} \pm 1.96\frac{\sigma}{\sqrt{n}}\right]$