01/29/2020

Class #2

We observe "data"

$x = \langle 0, 0, 1, 0, 1, 0 \rangle$ realization from a random process, Assume a parametric process

$\mathcal{F}$ = iid Bernoulli = $\{ \underbrace{p(x;\theta)}_{\text{PMF}} : \theta \in \Theta \} = \{ \theta^x (1-\theta)^{1-x} : \theta \in (0,1) \}$

PMF $p(x) = p(x;\theta)$ constant θ need to compare the probability

i.e. $f(x) = f(x;a)$
$= \sin(ax)$

We'd like to learn about θ (inference) $\mathcal{F}$

$p(x;\theta) = p(\langle 0,0,1,0,1,0 \rangle; \theta) = (\theta^0 (1-\theta)^{1-0}) \cdot (\theta^0 (1-\theta)^{1-0}) \cdots (\theta^0 (1-\theta)^{1-0})$

$= \theta^2 (1-\theta)^4$

what if $\theta = 0.5$ $p(x;\theta) = 0.5^2 (1-0.5)^4 = 0.0156$

what if $\theta = 0.25$ $p(x;\theta) = 0.25^2 (1-0.25)^4 = 0.0198$

"likelihood function" "probability"

$\mathcal{L}(\theta; x) = p(x; \theta)$

θ: input, x: constant; x: input, θ: constant

$\Rightarrow$ $p(x;\theta) \in (0,1)$

$\mathcal{L}(x;\theta) \in (0,1)$

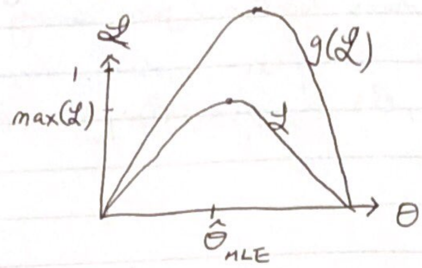
$\underset{\theta \in \Theta}{\operatorname{argmax}} \{ \mathcal{L}(\theta; x) \} := \hat{\Theta}_{MLE}$ "maximum likelihood estimate"

argument that results in the maximum

consider g is a stricly increasing function.

$= \underset{\theta \in \Theta}{\operatorname{argmax}} \{ \underset{\ln}{g}(\mathcal{L}(\theta : x)) \}$

$= \operatorname{argmax} \{ \ell(\theta; x) \}$



$\ell(\theta; x) := \ln(\mathcal{L}(\theta; x)) = \ln\left(\prod_{i=1}^{6} \theta^{x_i}(1-\theta)^{1-x_i}\right)$

"log-likelihood" monotonically increasing function

$= \sum_{i=1}^{6} \ln\left(\theta^{x_i}(1-\theta)^{1-x_i}\right)$

$= \sum_{i=1}^{6} x_i \ln(\theta) + (1-x_i)\ln(1-\theta)$

$= \ln(\theta) \sum_{i=1}^{6} x_i + \ln(1-\theta) \sum_{i=1}^{6} 1 - x_i$

let $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

$\Rightarrow \Sigma x_i = n\bar{x}$ sample avg

$= \ln(\theta)(6\bar{x}) + \ln(1-\theta)(6-6\bar{x}) = \ell(\theta; x)$

$= 6(\ln(\theta)\bar{x} + \ln(1-\theta)(1-\bar{x}))$

① Find $\ell(\theta; x)$

② Find $\frac{d}{d\theta}[\ell(\theta; x)] = \ell'(\theta; x)$

③ set $\ell'(\theta; x) = 0$

④ solve for Θ_{MLE}

$\ell'(\theta; x) = 6\left(\frac{\bar{x}}{\theta} - \frac{1-\bar{x}}{1-\theta}\right) \overset{\text{set}}{=} 0$

$\Rightarrow \frac{\bar{x}}{\theta} = \frac{1-\bar{x}}{1-\theta} \Rightarrow \bar{x}(1-\theta) = (1-\bar{x})\theta \Rightarrow \bar{x} - \bar{x}\theta = \theta - \bar{x}\theta \Rightarrow \hat{\Theta}_{MLE} = \bar{X}$

$x = \langle 0, 0, 1, 0, 1, 0 \rangle \Rightarrow \bar{x} = \frac{1}{3} = \hat{\Theta}_{MLE}$

$\bar{x} = \frac{1}{6}\sum_{i=1}^{6} x_i = \frac{0+0+1+0+1+0}{6}$

$X_1, \ldots, X_n \mid \Theta \overset{iid}{\sim} \text{Bernoulli}(\Theta)$

$\Rightarrow \hat{\Theta}_{MLE} = \bar{X}$

MLE is not the only point estimation estrategy but it is common as it has nice proportion.

I) $\hat{\theta}_{MLE} \xrightarrow{P} \theta$ "consistency" i.e. it converges to the value in n.

standard error $= \sqrt{var}$

II) $\hat{\theta}_{MLE} \approx N(\theta, SE[\hat{\theta}_{MLE}]^2)$ "Asymptotic Normality"

III) Among all consistent estimates, $\hat{\theta}_{MLE}$ has lowest variance. "Efficiency".

$\mathcal{F}$ = iid Geometric $\quad X \sim Geom(\theta) = (1-\theta)^x \theta$

$$\mathcal{L}(\theta; x) = \prod_{i=1}^{n} (1-\theta)^{x_i} \theta$$

more covenient

$$\Rightarrow \ell(\theta; x) = \sum_{i=1}^{n} \ln((1-\theta)^{x_i}\theta) = \sum_{i=1}^{n} x_i \ln(1-\theta) + \ln(\theta) = n(\bar{x}\ln(1-\theta) + \ln(\theta))$$

$$\ell'(\theta; x) = \left(-\frac{\bar{x}}{1-\theta} + \frac{1}{\theta}\right) \overset{set}{=} 0$$

$$\Rightarrow \frac{1}{\theta} = \frac{\bar{x}}{1-\theta} = 1-\theta = \theta\bar{x}$$

$$\Rightarrow 1 = \theta + \theta\bar{x}$$

$$\Rightarrow 1 = \theta(1-\bar{x})$$

$$\Rightarrow \hat{\theta}_{MLE} = \frac{1}{1+\bar{x}}$$

example: $\bar{x} = 99$

$\hat{\theta}_{MLE} = \frac{1}{1+99} = 1\%$

$\bar{x} = 1$

$\hat{\theta}_{MLE} = \frac{1}{1+1} = 50\%$

$\mathcal{F}$ = iid Bernoulli $\quad \hat{\theta}_{MLE} = \bar{X}$

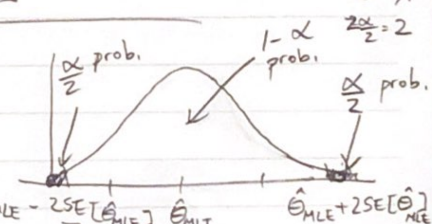
$$\hat{\theta}_{MLE} \overset{d}{\approx} N(\theta, SE[\hat{\theta}_{MLE}]^2) \overset{\text{I}}{\approx} N\left(\hat{\theta}_{MLE}, SE[\hat{\theta}_{MLE}]\Big|_{\theta=\hat{\theta}_{MLE}}\right) \approx N\left(\bar{x}, \sqrt{\frac{\bar{x}(1-\bar{x})}{n}}^2\right)$$

$$SE[\hat{\theta}_{MLE}] = SE[\bar{X}] = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{\theta(1-\theta)}{n}}$$

$\mathcal{F}$ = iid Geometric $\quad \hat{\theta}_{MLE} = \frac{1}{1+\bar{X}}, \quad SE[\hat{\theta}_{MLE}] = SE\left[\frac{1}{1+\bar{X}}\right] = ?$ MATH 369

$\alpha = 5\%$

$\frac{2\alpha}{2} = 2$



2nd Goal of inference "Confidence sets"

i.e. provide a range of plausible values of θ.

$$CI_{\theta, 1-\alpha} := \left[\hat{\theta}_{MLE} \pm Z_{\frac{\alpha}{2}} SE[\hat{\theta}_{MLE}]\Big|_{\theta=\theta_{MLE}}\right]$$

confidence interval for θ of size $1-\alpha$

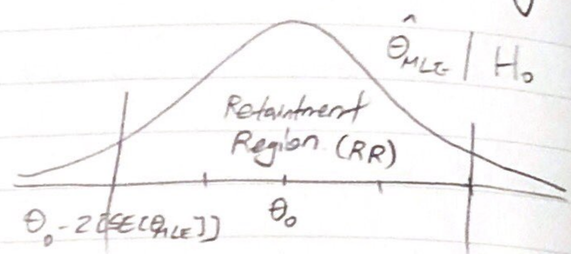
Third Goal of Inference : "Testing" also called "Hypothesis Testing"

"Null hypothesis" $H_0 : \theta = \theta_0$ ← some specific value.

"alternative hypothesis" $H_a : \theta \neq \theta_0$

Assume my theory is true and let the data tell me if I'm rigth or wrong

$$\Rightarrow \hat{\theta}_{MLE} \approx N\left(\theta_0 ,\ SE[\hat{\theta}_{MLE}]\Big|^{2}_{\theta=\hat{\theta}_{MLE}}\right)$$



$$RR_{\theta_0, 1-\alpha} := \left[\theta_0 \pm z_{\frac{\alpha}{2}} SE[\hat{\theta}_{MLE}]\Big|_{\theta=\theta_0}\right]$$

If $\hat{\theta}_{MLE} \in RR \Rightarrow$ Retains H_0
If $\hat{\theta}_{MLE} \notin RR \Rightarrow$ Rejects H_0 / Accept H_a

We have a strategy for all 3 informal goals, We have done "freqrencist inference" which is historically classic