01/30/2020.

We observe "data"; $x = <0,0,1,0,1,0>$

Realization from a random process; assume a parametric model

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

$p(x) = p(x;\theta)$; constant θ needed to complete probability.
$f(x) = f(x;a) = \sin(ax)$

We'd like to learn about θ (inference)

$p(x;\theta) = p(<0,0,1,0,1,0>;\theta) \overset{\mathcal{F}}{=} (\theta^0(1-\theta)^{1-0}).(\theta^0(1-\theta)^{1-0})\dots\dots(\theta^0(1-\theta)^{1-0}) = \theta^2(1-\theta)^4.$

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

input constant input constant

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

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

$\hat{\theta}_{MLE} = \underset{\theta \in \Theta}{\arg\max}[\mathcal{L}(\theta;x)] = \arg\max[g(\mathcal{L}(\theta;x))]$.

$= \arg\max[\ell(\theta;x)]$

argument that results; if that maximum.

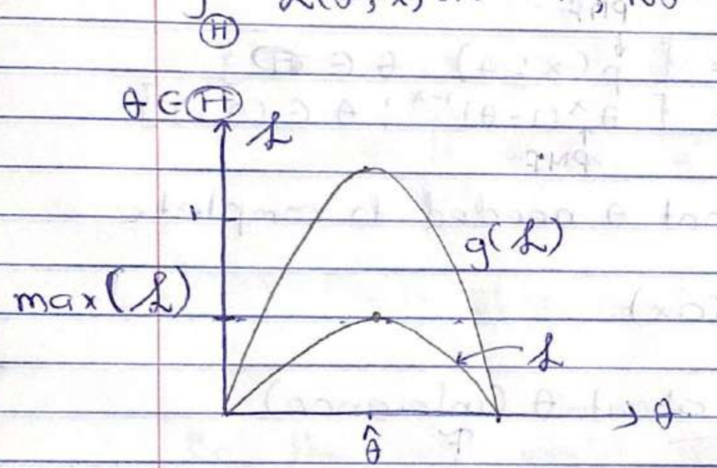
maximum likelihood estimate

consider g is a strickly increasing function

$\int_{\Theta} \mathcal{L}(\theta;x)\,d\theta \overset{?}{=} 1$; No

$\left.\begin{aligned}\sum_x p(x;\theta) \\ \int_x p(x;\theta)\,dx\end{aligned}\right] = 1$

$\theta \in \Theta$



$\ell(\theta;x) := \ln(\mathcal{L}(\theta;x)) =$

log-likelihood

monotonically increasing function

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

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

let $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \Rightarrow \sum x_i = n\bar{x}$ sample avg.

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

$= \ell(\theta;x)$

I Find $\ell(\theta; x)$
II Find $\frac{d}{d\theta}[\ell(\theta; x)] = \ell'(\theta; x)$
III $\ell'(\theta; x) \overset{set}{=} 0$
IV Solve for $\hat{\theta}_{MLE}$

Exercise

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

$$\Longrightarrow \bar{x}/\theta = \frac{(1-\bar{x})}{(1-\theta)} \Longrightarrow \bar{x}(1-\theta) = (1-\bar{x})\theta$$

$$\Longrightarrow \bar{x} - \bar{x}\theta = \theta - \bar{x}\theta \Longrightarrow \boxed{\hat{\theta}_{MLE} = \bar{x}}$$

$x = \langle 0, 0, 1, 0, 1, 0 \rangle$

$$\Longrightarrow \bar{x} = 1/3 = \hat{\theta}_{MLE}$$

$$\left(\bar{x} = 1/6 \sum_{i=1}^{6} x_i = \frac{0+0+1+0+1+0}{6} = 2/6 = 1/3\right)$$

Most commonly
$$\boxed{\begin{gathered} X_1, \ldots, X_n \mid \theta \overset{iid}{\sim} \text{Bern}(\theta) \\ \Longrightarrow \hat{\theta}_{MLE} = \bar{x} \end{gathered}}$$

MLE is not the only point estimates strategy, but it is common as it has nice properties.

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

standard error = $\sqrt{\text{variance}}$

(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"]

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

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

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

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

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

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

$\Rightarrow 1 = \theta(1+\bar{x})$; $\hat{\theta}_{MLE} = 1/(1+\bar{x})$

99 failure: want to wait long for success (1%)

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

$\bar{x} = 1 \Rightarrow \hat{\theta}_{MLE} = \frac{1}{1+1} = 1/2 = 50\%$

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

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

$\hat{\theta}_{MLE} \approx N(\theta, SE[\hat{\theta}_{MLE}]^2) \overset{\text{(I)}}{\approx} N(\hat{\theta}_{MLE}, SE[\hat{\theta}_{MLE}]\Big|^2_{\theta=\hat{\theta}_{MLE}})$

" $\approx N(\bar{x}, \sqrt{\frac{\bar{x}(1-\bar{x})}{n}}^2$ " $)$

T = iid Geometric $\hat{\theta}_{MLE} = \frac{1}{1+\bar{x}}$; $SE[\hat{\theta}_{MLE}]$

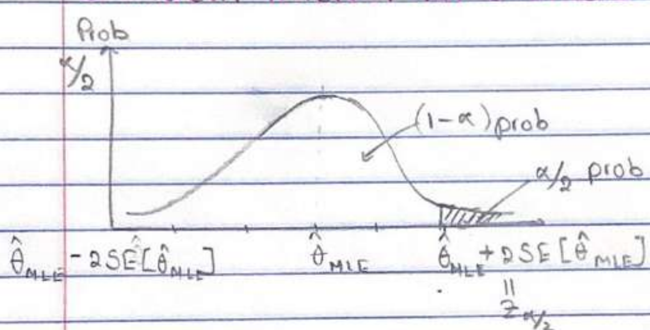
$= SE\left[\frac{1}{1+\bar{x}}\right] = ??$ [will study in higher math courses]

2^{nd} goal of inference ; " Confidence set "

- Provide a range of plaussible value of θ.

$$CI_{\theta, 1-\alpha} = \left[\hat{\theta}_{MLE} \pm z_{\alpha/2} \, SE[\hat{\theta}_{MLE}] \Big|_{\theta = \hat{\theta}_{MLE}} \right]$$

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



$\alpha = 5\%$

$\Rightarrow z_{\alpha/2} \approx 2$

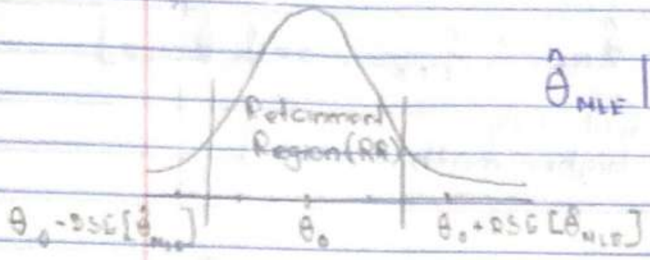
3^{rd} goal of inference ; Testing also called hypothesis testing.

$H_0 : \theta = \theta_0 \leftarrow$ Some specific value $\quad H_a : \theta \neq \theta_0$

idea " Null hypothesis " $\qquad$ Alternative hypothesis.

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

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



$\hat{\theta}_{MLE} \mid H_0 \quad RR_{\theta_0, 1-\alpha} := \left[\theta_0 \pm z_{\alpha/2} SE[\hat{\theta}_{MLE}] \Big|_{\theta = \theta_0} \right]$

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

We've a strategy for all 3 inferential goals; we've done "frequentist inference" which is historically classic.