210 = even
210 =

2/17 maximum a posteriori – max after the fact
↳ most probable θ after you see the data.

$$\hat{\theta}_{MAP} := \underset{\theta \in \Theta_0}{\arg\max}\{P(\theta \mid x)\}$$

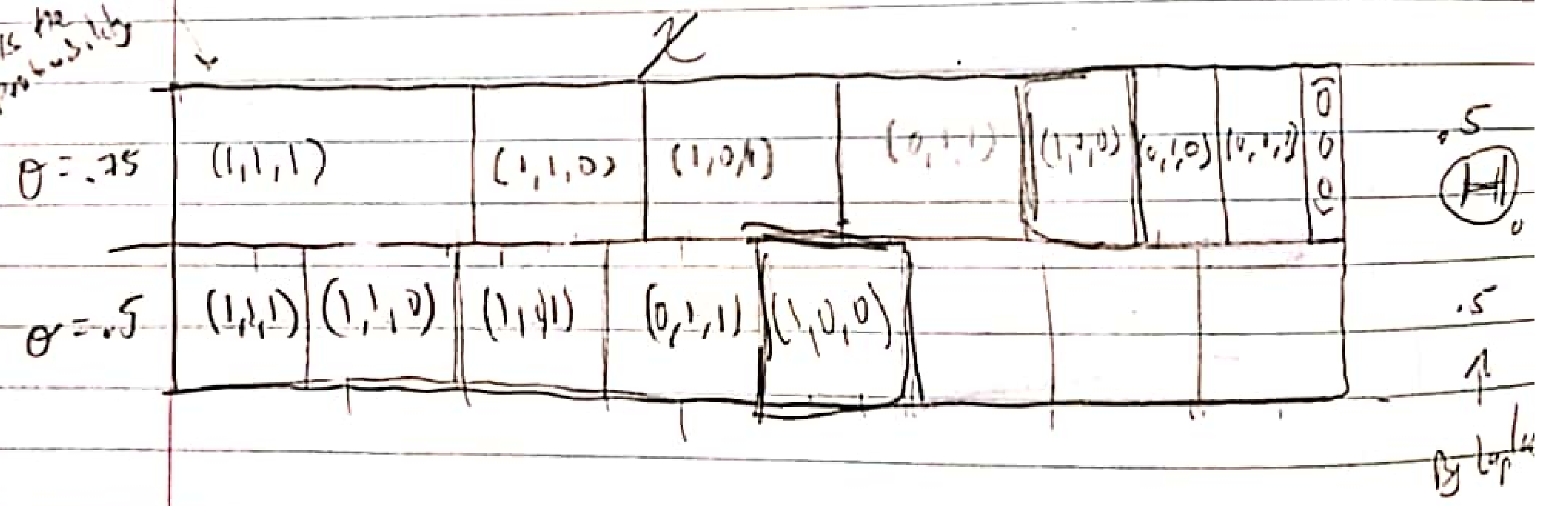
↑ After draw

This is one of 3 Bayesian Point Estimates.

$\mathcal{F}$ = iid Bern(θ), $n=3$, $x=(0,1,1)$ $\Theta_0 = \{.5, .75\}$

$x \in \mathcal{X} = \{0,1\} \times \{0,1\} \times \{0,1\}$ $2^3 = 8$ data sets

The area of the boxes is the probability



$P(X=(1,1,1) \mid \theta = .75) = (.75)^3 = .422$ 42%
$P(X=(1,1,0) \mid \theta = .75) = (.75)^2(.25) = .141$ 14%

Since $\mathcal{F}$ is iid Bern(θ) $P(X=(1,1,0) \mid \theta = .75) = P(X=(1,0,1) \mid \theta = .75)$

$P(X=(1,0,0) \mid \theta = .75) = .047$ 4%

$P(X=(0,0,0) \mid \theta = .75) = .25^3 = .016$ 1.5%

2/17

$P(X=(1,1,1) \mid \theta=.5) = .125$

: All will be the same.

hence the bottom half should be 8 even boxes

$P(X=(1,0,0)) = P(X=(1,0,0), \theta=.75) + P(X=(1,0,0), \theta=.5)$

$.047 + .125$

$$P(\theta=.5 \mid X=(1,0,0)) = \frac{P(\theta=.5 \cap X=(1,0,0))}{P(1,0,0)}$$

$$= \frac{.125}{.047+.125} \qquad \frac{\square}{\square + \square}$$

For discrete parameter spaces

$$\sum_{\theta \in \Theta} P(\theta) = 1 \qquad \sum_{\theta \in \Theta} P(\theta \mid x) = 1$$

$\sum_{\theta \in \Theta} P(x \mid \theta)$ = could be anything

(If $P(\theta)$ is constant Laplaces idea)

$$P(\theta \mid x) = \frac{P(x, \theta)}{P(x)} \propto P(x \mid \theta) \propto P(x \mid \theta) P(\theta) \propto P(x \mid \theta)$$

By principle of indifference

$$\hat{\theta}_{MAP} = \underset{\theta \in \Theta_0}{\arg\max} \{P(\theta \mid x)\} = \arg\max \{P(x \mid \theta)\, P(\theta)\} = \arg\max \{P(x \mid \theta)\} \overset{??}{=} \hat{\theta}_{MLE}$$

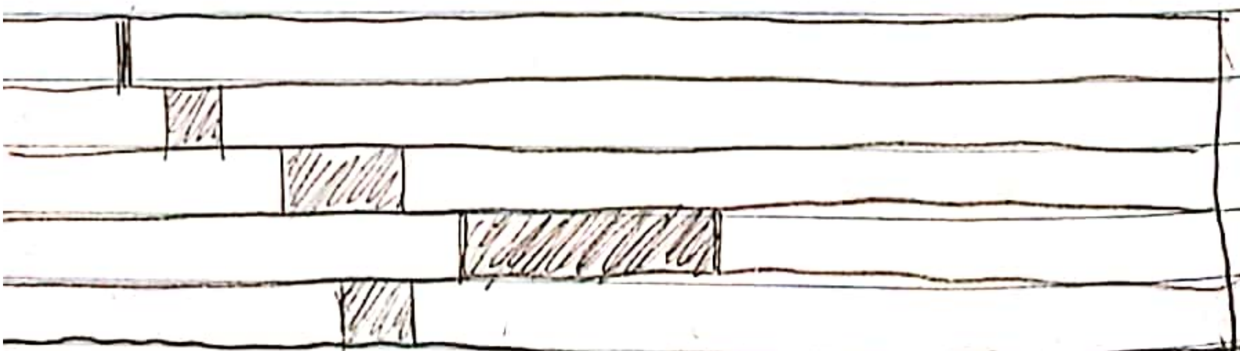
only equal if the MLE is in the parameter set you specify. ↳ Θ_0

Let $\Theta_0 = \{.1, .25, .5, .75, .9\}$, (Laplace Prior)

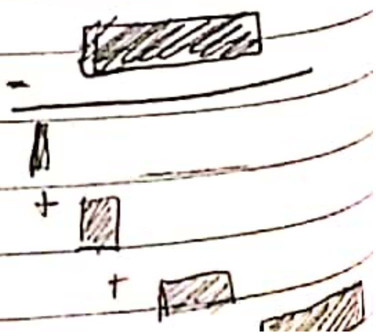
let $x = (1, 1, 0)$

$P(x \mid \theta = .1) = .009$
$P(x \mid \theta = .25) = .047$
$P(x \mid \theta = .5) = .125$
$P(x \mid \theta = .75) = .141$
$P(x \mid \theta = .9) = .081$

$\frac{1}{5}$



$P(\theta = .75 \mid x = (1,1,0)) =$



$\hat{\theta}_{MAP} = .75$

(lar)gest area

7/17

MLE is not in this set Θ_0. So what do we do?
Lets examine laplace's prior under many different parameter spaces (approaching the full space.

Trying to give every θ in $(0,1)$ equal prob by laplace prior results in $U(0,1)$ for the full parameter set for Bern

$$\Theta_{0,3} = \left\{ \frac{1}{4}, \frac{1}{2}, \frac{3}{4} \right\} \Rightarrow P(\theta) = U(\Theta_{0,3}) = \left\{ \frac{1}{3} \ \forall \theta \right.$$

$$\Theta_{0,9} = \left\{ \frac{1}{10}, \frac{2}{10}, \ldots, \frac{9}{10} \right\} \Rightarrow P(\theta) = U(\Theta_{0,9}) = \left\{ \frac{1}{9} \ \forall \theta \right.$$

$$\Theta_{0,n} = \left\{ \frac{1}{n+1}, \frac{2}{n+1}, \ldots, \frac{n}{n+1} \right\}$$

$$= P(\theta) = U(\Theta) = \left\{ \frac{1}{n} \ \forall \theta \right.$$

$$(0,1) = \lim_{n\to\infty} \Theta_{0,n} \Rightarrow P(\theta) = 0 \ \forall \theta \ \underline{\text{not a PMF}}!$$

$$\lim_{n\to\infty} F_n(\theta) = F(\theta) = \theta \Rightarrow P(\theta) = F'(\theta) = 1$$

$$\Rightarrow P(\theta) = U(0,1)$$

<u>ie cts</u>

$F \stackrel{iid}{=} \text{Bern}(\theta) \quad X = (1,1,0) \quad P(\theta) = U(0,1)$

$$P(\theta|X) = \frac{P(X|\theta)\overbrace{P(\theta)}^{=1}}{P(X)} = \frac{P(X|\theta)}{P(X)} = \frac{P(X|\theta)}{\int_{\Theta} P(X,\theta)\,d\theta}$$

$$= \frac{P(X|\theta)}{\int_{\Theta} P(X|\theta)\overbrace{P(\theta)}^{=1}\,d\theta} = \frac{\theta^2(1-\theta)}{\int_0^1 \theta^2(1-\theta)\,d\theta} = \frac{\theta^2(1-\theta)}{\left[\frac{\theta^3}{3} - \frac{\theta^4}{4}\right]_0^1} = \frac{\theta^2(1-\theta)}{\frac{1}{12}}$$

$\rightarrow$

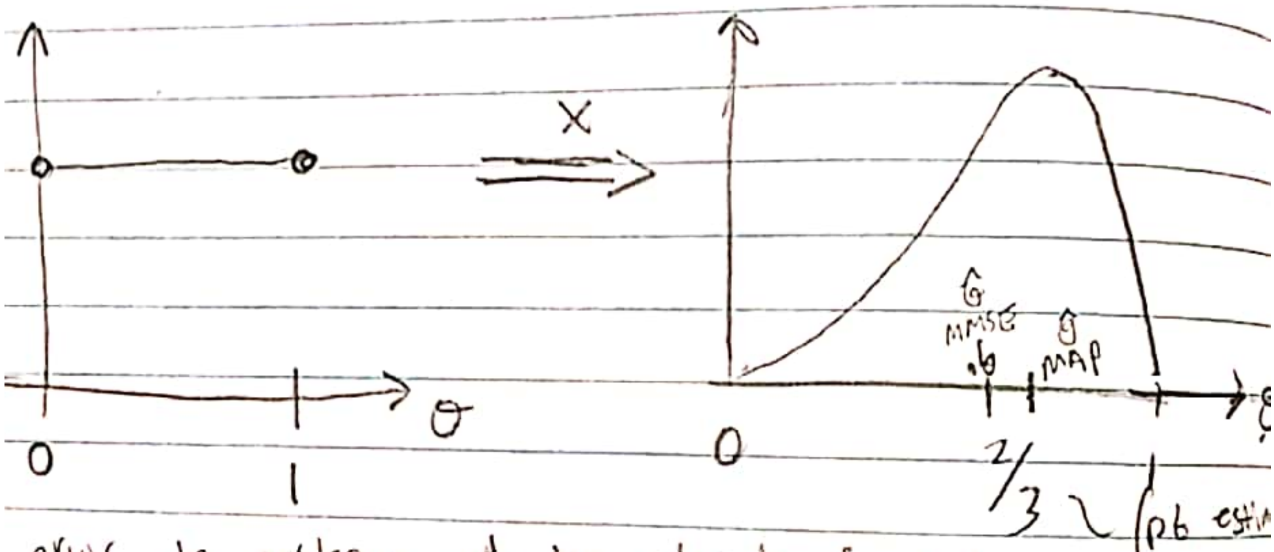
mean, median, mode

$= 12\theta^2(1-\theta) = P(\theta|x)$

$\hat{\theta}_{MAP} = \arg\max\{12\theta^2(1-\theta)\} = \arg\max\{\theta^2(1-\theta)\}$

(take derivative and set to 0) $= \arg\max\{\mathcal{L}(\theta;x)\} = \frac{2}{3}$

$P(\theta) \xRightarrow{x} P(\theta|x)$ $\quad p(\theta|x)$



prior to posterior with the advent of data

$P(\theta > .5|x) = \int_{.5}^{1} 12\theta^2(1-\theta)\,d\theta = .688$

the coin is tilted heads

We talked about the MAP Bayesian point estimate. Are there other measures of "best guess theta" if you have posterior distribution $P(\theta|x)$?

2/17

$$\hat{\theta}_{MMSE} := E[\theta | X] = \arg\min_{\theta \in \Theta} \{ E[(\theta - \hat{\theta})^2 | X] \}$$

The minimum mean squared error Bayesian point estimate is the posterior mean (expectation). In our case:

$$\hat{\theta}_{MMSE} = \int_{\Theta} \theta \, P(\theta | X) \, d\theta = \int_0^1 \theta \, 12 \theta^2 (1-\theta) \, d\theta = \frac{12}{20} = \_$$

(This is the default estimator)