Lecture 8 02-20-2020

$H_0: \theta = \theta_0$
$H_a: \theta \neq \theta_0 \Rightarrow$ two sided test

IF $Pval := P(H_0|x) < \alpha \Rightarrow$ Reject H_0/Accept H_a
$= P(\theta = \theta_0 | x) = 0 \Rightarrow$ Problem

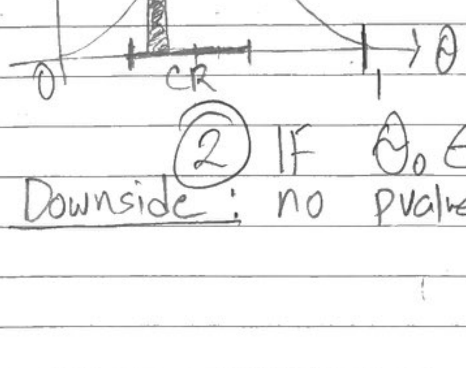
$P(\theta) = U(0,1)$

Two Ideas:
① You declare δ e.g. $\delta = 0.01$, a "margin of equivalance." Then you modify the hypothesis —

$H_0: \theta \in [\theta_0 \pm \delta] \stackrel{e.g.}{=} [0.49, 0.51]$
$H_a: \theta \notin [\theta_0 \pm \delta]$

$Pval = P(H_0|X) = P(\theta \in [\theta_0 \pm \delta] | x)$

e.g. $n=100$, $x=61$
$= qbeta(.51, 62, 40) - qbeta(0.49, 62, 40)$
$= .609 - .607 = .002 < \alpha = 5\%$
$\Rightarrow$ Reject H_0.



② IF $\theta_0 \in CR_{\theta, 1-\alpha} \Rightarrow$ Retain H_0 else Reject.
Downside: no pvalue!

⊗ Modeling — Explanation (inference)

Modeling — Prediction

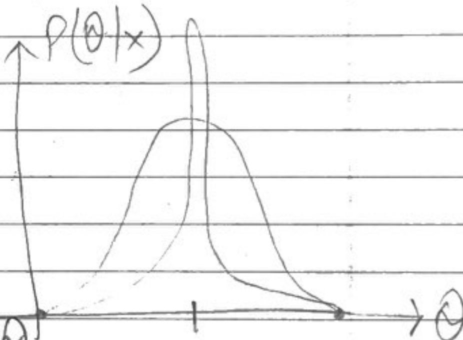
Modeling → try to approximate reality

Seen $x_1, \dots, x_n$ and you want to know how X_* (future data) will be distributed.

$P(X_* | X)$, the posterior prediction distribution.

If θ is known, what is the posterior prediction dist.?

$P(X^* | \theta)$. The best you can do. But it's not possible since θ is unknown.



$$P(X) = \int P(X,y)\,dy$$

continuous posterior

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

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

If θ is unknown, x does not add any information

$$= \int_{\Theta} \underbrace{P(X^* | \theta)}_{\text{(liklihood)}}\ \underbrace{P(\theta | X)}_{\text{posterior}}\, d\theta$$

discrete posterior

$$= \sum_{\theta \in \Theta} P(X^* | \theta)\, P(\theta | X)$$

$\textcircled{x}$ $\Theta_0 = \{0.5, 0.75\}$, $X = \langle 0,1,1 \rangle$

$X^* \sim$? One future observation

$P(X^* | X) = \text{Bern}(?)$

$\text{Supp}[X^* | X] = \{0,1\}$

(A) Consider Rich theory

$X^* | X \sim \text{Bern}\left(\hat{\theta}_{MLE} = \frac{2}{3}\right)$

Problems:

1. $\hat{\theta}_{MLE}$ may not be $\in \Theta_0$
2. $\hat{\theta}_{MLE}$ could be 0 or 1.
3. If multiple future observations n_*

$X_* \sim \text{Bin}(n_*, \hat{\theta}_{MLE})$?

Bad Idea

(B) Assum prior of indifference

Posterior

$P(\theta = 0.75 | X) = 0.53$, $P(\theta = 0.5 | X) = 0.47$

$$P(X^* | X) = P(X^* | \theta = 0.75)\, P(\theta = 0.75 | X) + P(X^* | \theta = 0.5)\, P(\theta = 0.5 | X)$$

$$= P(x^* \mid \theta = 0.75)\cdot(0.53) + P(x^* \mid \theta = 0.5)\cdot(0.47)$$

$$= (0.75)^{x^*}(0.25)^{1-x^*}\cdot(0.53) +$$

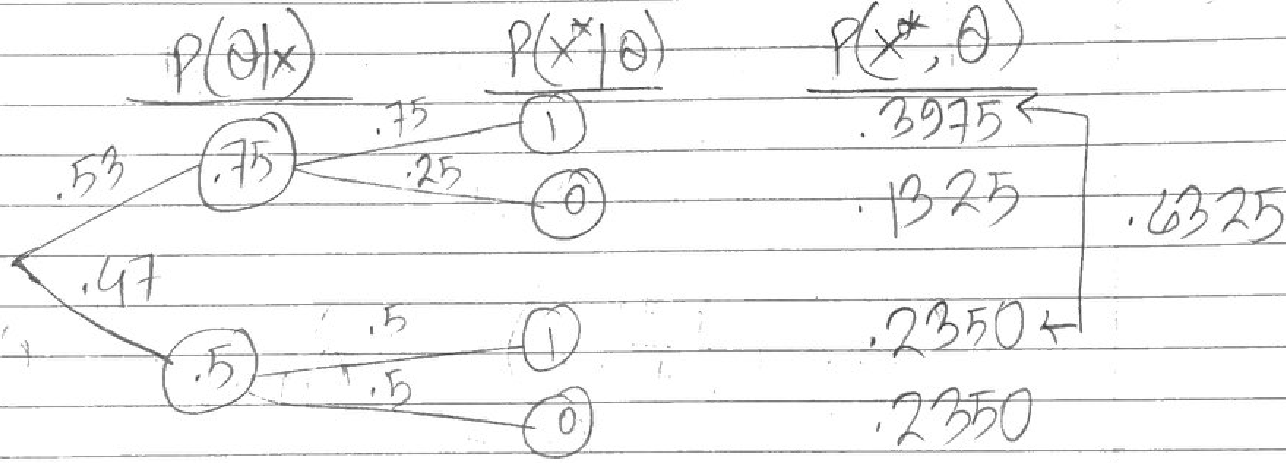
$$(0.5)^{x^*}(0.5)^{1-x^*}\cdot(0.47)$$

$P(x^* \mid x) = \text{Bern}(?) = \text{Bern}(0.6325)$

trick complete $P(x^* = 1 \mid x) = (0.75)^1(0.25)^0(0.53) +$

$(0.5)^1(0.5)^0(0.47)$

$= 0.6325$



: Binomial fixed n.

$P(\theta) = \text{Beta}(\alpha, \beta)$

$\Rightarrow P(\theta \mid x) = \text{Beta}(\alpha + x, \beta + n - x)$

What is the posterior predictive dist. for $n_* = 1$.

$$P(X_* \mid x) = \int_{\Theta} p(X_* \mid \theta)\, p(\theta \mid x)\, d\theta$$

$$= \int_0^1 \left(\theta^{x_*}(1-\theta)^{1-x_*}\right)\left(\frac{1}{B(\alpha+x, \beta+n-x)}\, \theta^{\alpha+x-1}(1-\theta)^{\beta+n-x-1}\right) d\theta$$

$$= \frac{1}{B(\alpha+x, \beta+n-x)} \int_0^1 \theta^{\alpha+x+x_*-1}(1-\theta)^{\beta+n-x-x_*+1-1}\, d\theta$$

$$= \frac{B(\alpha+x+x_*, \beta+n-x-x_*+1)}{B(\alpha+x, \beta+n-x)}$$

$$= \mathrm{Bern}\left(\frac{\alpha+x}{\alpha+\beta+n}\right) \qquad \hat{\theta}_{MMSE} = E[\theta \mid x]$$

Trick is $P(X_*=1 \mid x) = \dfrac{B(\alpha+x+1, \beta+n-x)}{B(\alpha+x, \beta+n-x)}$

$$= \frac{\dfrac{\Gamma(\alpha+x+1)\,\Gamma(\beta+n-x)}{\Gamma(\alpha+\beta+n+1)}}{B(\alpha+x, \beta+n-x)}$$

$$= \frac{\dfrac{(\alpha+x)\,\Gamma(\alpha+x)\,\Gamma(\beta+n-x)}{(\alpha+\beta+n)\,\Gamma(\alpha+\beta+n)}}{B(\alpha+x,\ \beta+n-x)}$$

$$= \frac{\alpha+x}{\alpha+\beta+n}\ \cancel{\frac{B(\alpha+x,\ \beta+n-x)}{B(\alpha+x,\ \beta+n-x)}}$$

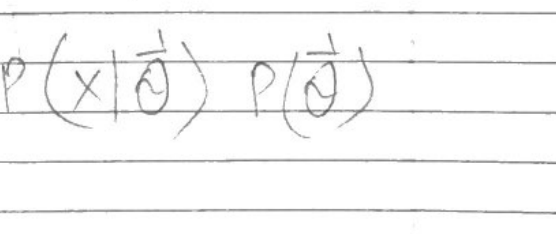
Midterm II material starts here

(x) Mixture Distributions:

component distr's

mixing proportions

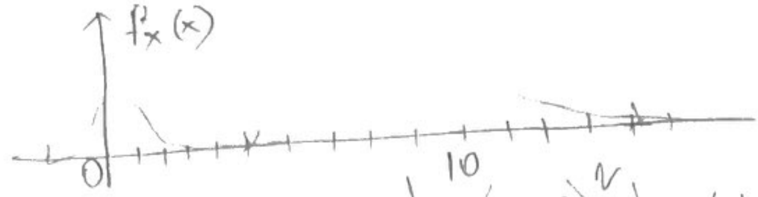
$$X \sim \begin{cases} N(0,1^2) & \text{wp } \frac{1}{2} \\ N(10,2^2) & \text{wp } \frac{1}{2} \end{cases}$$



$f_X(x) = ?$

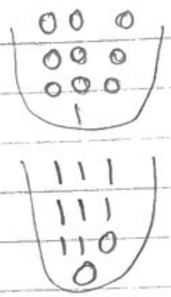
$$P(x) = \sum_{\theta \in \Theta} P(x, \vec{\theta}) = \sum_{\theta \in \Theta} P(x \mid \vec{\theta})\, P(\vec{\theta})$$

⇒

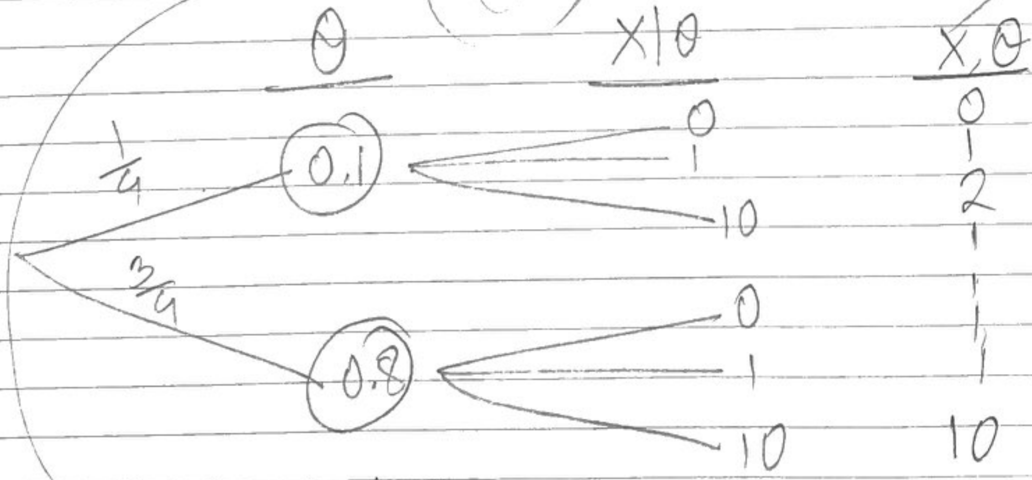


$$= \left(\frac{1}{\sqrt{2\pi(1^2)}}\, e^{-\frac{1}{2\cdot 1^2}(x-0)^2}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{\sqrt{2\pi(2^2)}}\, e^{-\frac{1}{2\cdot 2^2}(x-10)^2}\right)\left(\frac{1}{2}\right)$$

$$X \sim \begin{cases} \text{Bin}(10, 0.1) & \text{wp } \frac{1}{4} \\ \text{Bin}(10, 0.8) & \text{wp } \frac{3}{4} \end{cases}$$



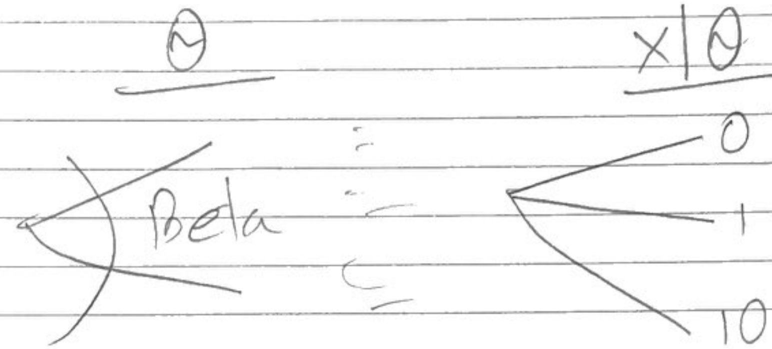
$$P_X(x) = \left(\binom{10}{x} 0.1^x\, 0.9^{10-x}\right)\left(\frac{1}{4}\right) + \left(\binom{10}{x} 0.8^x\, 0.2^{10-x}\right)\left(\frac{3}{4}\right)$$



$\neq$ Binomial

Mixture dist's have a discrete # of component.
Compound dist's do not.

$$X \sim \left\{ \begin{array}{l} Bin(n, \theta) \\ Bin(n, \theta) \end{array} \right\} \quad \theta\text{'s come from a Beta}$$



$$P(\Theta | x) = \frac{P(x|\Theta)\, P(\Theta)}{P(x)}$$

$$= \frac{P(x|\Theta)\, P(\Theta)}{\underbrace{\int_{\Theta} P(x|\Theta)\, P(\Theta)\, d\Theta}_{\text{component dist.}}}$$