Lecture - 08 02/20/2020

$H_0 : \theta = \theta_0$
$H_a : \theta \neq \theta_0 = 0.5$ two-sided test

if Pval := $P(H_0 \mid X) < \alpha \Rightarrow$ Reject H_0 / Accept H_a

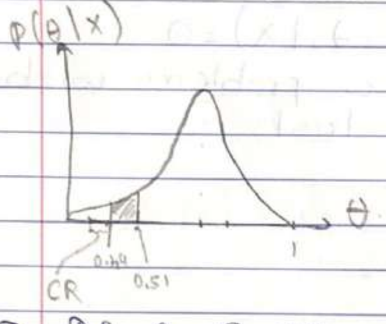
$= P(\theta = \theta_0 \mid x) \overset{P(\theta) = U(0,1)}{=} 0$

Two Ideas

I Declair δ e.g $\delta = 0.01$; a "margin of equavalence" then you modify the hypotheses.

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

$Pval = P(H_0 \mid X) = P(\theta \in [\theta_0 \pm \delta] \mid X)$
$= qbeta(0.51, 62, 40) - qbeta(0.49, 62, 40)$
$=$ (small) $0.609 - 0.607$
$= 0.002 < \alpha = 5\% \Rightarrow$ Reject H_0.



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

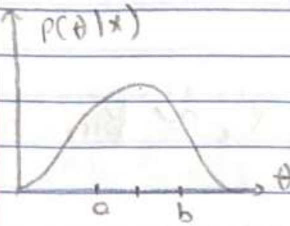
modeling → Explanation (inferance)
modeling → Prediction $\Rightarrow$ seen $x_1, \ldots, x_n$ and you want to know how future X_* (future data) will be distributed.

↓
try to approximate reality

$P(X_* | X)$; the posterior predictive distribution.

If θ was known, what is the posterior predictive distribution?

$P(X^* | \theta)$ the best you can do, but it is not possible since θ is unknown.



$$P(x^* | \theta) = P(x^* | \theta = a) P(\theta = a) + P(x^* | \theta = b) P(\theta = b)$$

weighted average — continuous posterior

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

$$= \int_{\Theta} p(x^* | \theta, x) \, p(\theta | x) \, d\theta.$$

If θ is unknown; x doesn't any more information.

$$= \int_{\Theta} \underbrace{p(x^* | \theta)}_{\text{likelihood}} \, \underbrace{p(\theta | x)}_{\text{posterior}} \, d\theta.$$

discrete posterior $= \sum_{\theta \in \Theta} p(x^* | \theta) \, p(\theta | x)$

$\Theta_0 = [0.5, 0.75]$; $x = \langle 0, 1, 1 \rangle$

$X^* \sim$? One future observation

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

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

Consider Rich's theory

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

Problems

① $\hat{\theta}_{MLE}$ may not be $\in \Theta$.

② $\hat{\theta}_{MLE}$ could be 0 or 1

③ If multiple future observation $X_* \sim \text{Bin}(n_*, \theta_{MLE})$ // BAD IDEA.

Assume prior of indifference posterior.

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

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

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

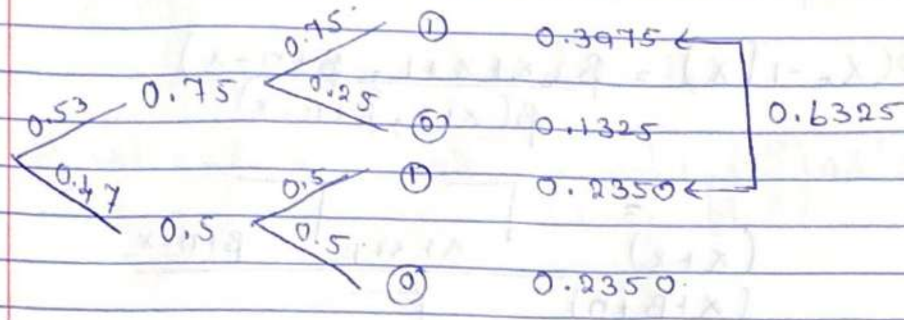
$$= (0.75)^{x^*} (0.25)^{1-x^*} (0.53) + (0.5)^{x^*} (0.5)^{1-x^*} (0.47)$$

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

trick; compute $P(X^* = 1 | X) = (0.75)^1 \cdot (0.25)^0 (0.53)$
$= (0.5)^0 (0.5)^1 (0.47)$
$= 0.6325$

Posterior after see the data.

$P(\theta \mid x)$ $\quad$ $P(x^* \mid \theta)$

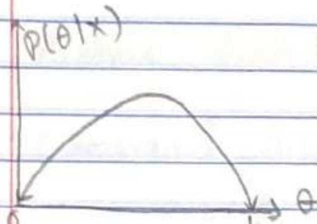


F: Binomial Fixed n

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

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

What is the posterior predictive distribution for $n_* = n$?



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

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

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

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

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

Trick compute.

$$P(X_* = 1 \mid x) = \frac{\beta(\alpha + x + 1,\ \beta + n - x)}{\beta(\alpha + x,\ \beta + n - x)}$$

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

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

$$= \frac{(\alpha + x)}{(\alpha + \beta + n)}$$

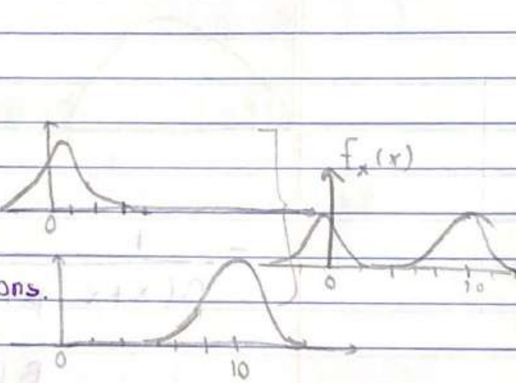
Mixture Distribution

component distribution

$X \sim$ $N(0, 1^2)$ up $\frac{1}{2}$

mixing proportions.

$N(10, 2^2)$ up $\frac{1}{2}$



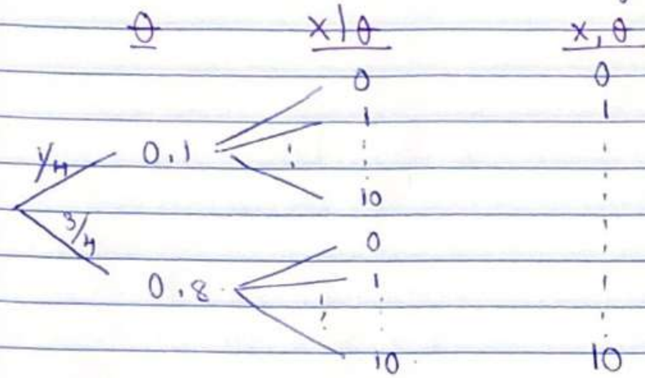
$F_x(x) = ?$ $\quad p(x) = \sum_{\theta \in \Theta} P(x, \vec{\theta})$

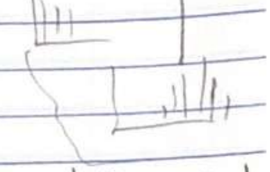
$$= \sum_{\theta \in \Theta} P(x \mid \vec{\theta})\, P(\vec{\theta})$$

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

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

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



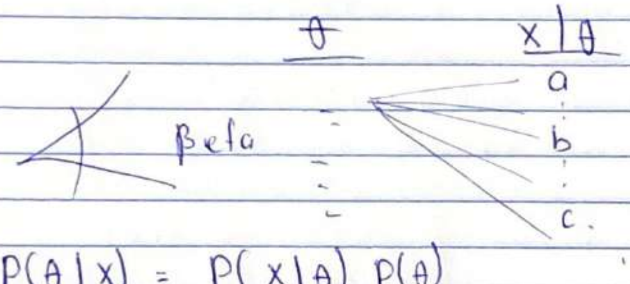


≠ Binomial
can't have two bump

Mixture distribution's has a discrete # of components.

Compound distribution's do not.

$$X \sim \left.\begin{cases} Bin(n, \theta) \\ Bin(n, \theta) \end{cases}\right\} \theta\text{'s come from Beta}$$



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

$$= \frac{P(x \mid \theta)\, P(\theta)}{\underbrace{\int_{\Theta} P(x \mid \theta)\, P(\theta)\, d\theta}_{\text{compound distribution.}}}$$