Lecture 8

02/20/20

$H_0: \theta = \theta_0$

$H_A: \theta \neq \theta_0$ Two-Sided test

$Pval = P(H_0 | X) < \alpha \Rightarrow$ Reject H_0 / Accept H_A

$= P(\theta = \theta_0 | X) = 0 \Rightarrow$ Problem

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

Two Ideas

1) Declare δ (delta) e.g. $\delta = 0.01$ the "margin of equivalence"

Then you modify the hypothesis:

$H_0: \theta \in [\theta_0 \pm \delta] = [0.49, 0.51]$

$H_A: \theta \notin [\theta_0 \pm \delta]$

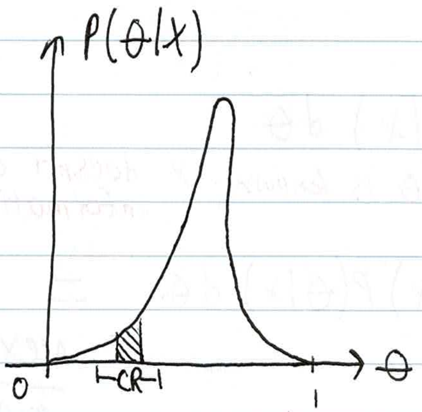
Back to Coin Toss Example...

$n = 100, \ X = 61$

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

$= qbeta(.51, 62, 40) - qbeta(0.49, 62, 40) = .609 - .602 =$

$= .002 < \alpha = 5\% \Rightarrow$ Reject H_0



2) If $\theta_0 \in CR_{\theta, 1-\alpha} \Rightarrow$ Retain H_0 else Reject H_0

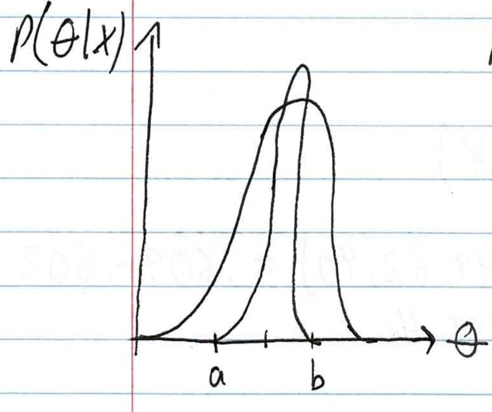
Downside: No pvalue

Try to approximate reality ← Modeling
- Explanation (Inference)
- Prediction
  Seen $x_1, \dots, x_n$ and you must know how X_* (future data) will be distributed

$P(X_* | x)$, the Posterior Predictive Distribution

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

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



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

continuous Posterior

$P(x) = \int p(x, y)\, dy$

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

If θ is known, X doesn't add any information

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

Next page

$$= \int_{\Theta} \underbrace{p(X^*|\theta)}_{\text{likelihood}} \underbrace{p(\theta|X)}_{\text{Posterior}} d\theta$$

Posterior : Everything known about θ after data is seen

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

↓ Discrete posterior

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

$X^* \sim ?$ one future observation

$p(X^*|X) =$ Bernoulli

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

Consider: $X^*|X \sim Bern(\underbrace{\hat{\theta}}_{\text{Frequentist}}{}_{MLE} = \frac{2}{3})$

<u>Problems with this:</u>

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_*

$$X^* \sim Bin(n_*, \hat{\theta}_{MLE})$$

Assume Prior of Indifference

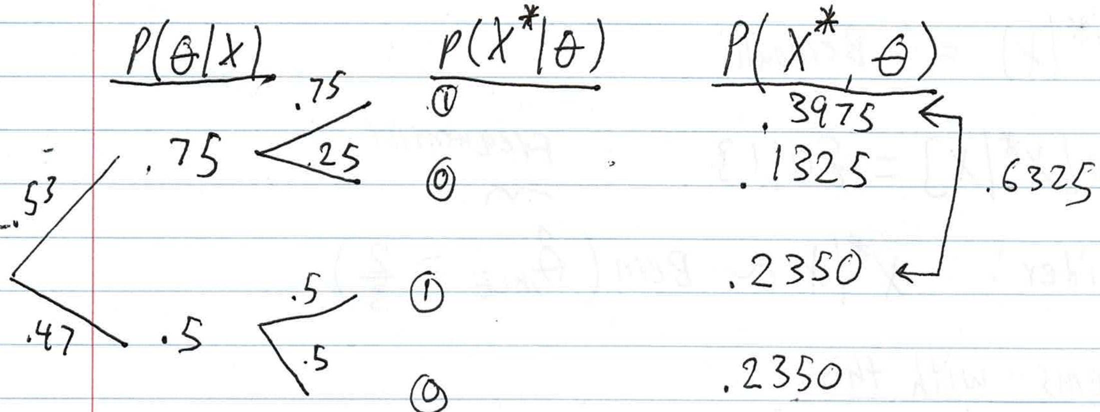
Posterior:

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

$$P(X^*|x)=P(X^*|\theta=0.75)P(\theta=0.75|x)+P(X^*|\theta=0.5)P(\theta=0.5|x)$$
$$=P(X^*|\theta=0.75)\cdot 0.53+P(X^*|\theta=0.5)\cdot .47$$
$$=(0.75)^{X^*}(0.25)^{1-X^*}(.53)+(.5)^{X^*}(.5)^{1-X^*}(.47)$$

Compute $P(X^*=1|x) = .75^1(.25^0)(.53)+(.5^0)(.5^1)(.47) = .6325$

$$P(X^*|x) = Bern(.6325)$$



$\tilde{F}$ binomial fixed n
$P(\theta)=Beta(\alpha,\beta)$

$$\Rightarrow P(\theta|X)=Beta(\alpha+X,\ \beta+n-x)$$

What is the posterior distribution for $n_0=1$?

$$P(X^*|x) = \int_{\Theta_0} P(X^*|\theta)\,P(\theta|x)\,d\theta$$

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

$$=\frac{1}{B(\alpha+X,\beta+n-x)}\int_0^1 \theta^{\alpha+X^*+X-1}(1-\theta)^{n-X+\beta-X^*+1-1}\,d\theta = \xrightarrow{\text{next}} \text{page}$$

as seen on previous page

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

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

$$= Beta(?) = Beta\left(\frac{\alpha+x}{n+\alpha+\beta}\right) \rightarrow \hat{\theta}_{MMSE} = E[\theta|x]$$

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

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

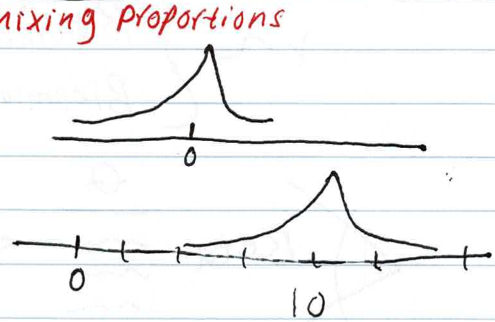
$$= \frac{\alpha+x}{\alpha+\beta+n} \cdot \frac{B(\alpha+x,\ \cancel{\beta+n-x})}{B(\alpha+x,\ \cancel{\beta+n-x})}$$

End of Midterm 1 Material ↑

Mixture distributions

component Distribution ↙ Mixing Proportions ↙

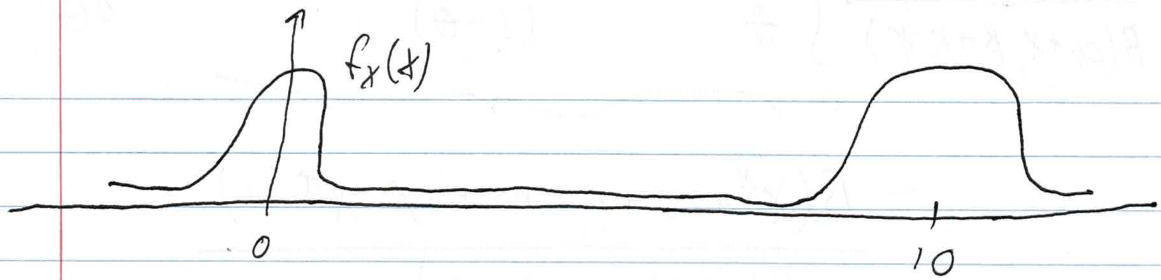
$$X \sim \begin{cases} N(0,1^2) & w.p.\ \frac{1}{2} \\ N(10,2^2) & w.p.\ \frac{1}{2} \end{cases}$$



$f_X(x) = ?$

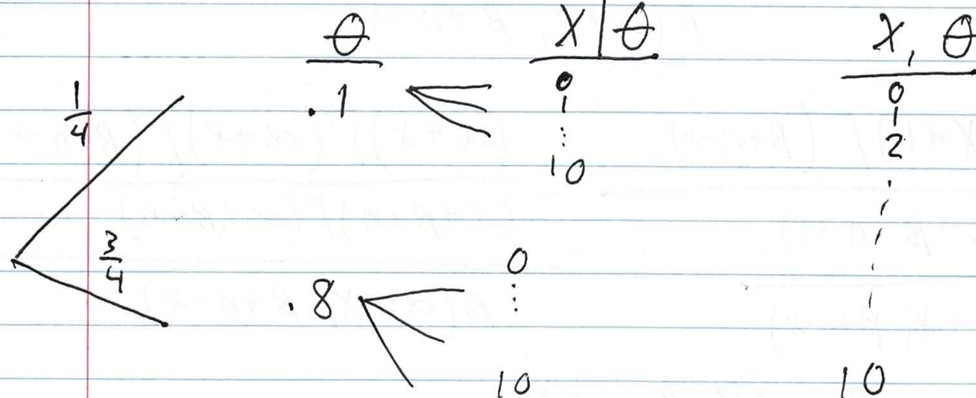
$$P(x) = \sum_{\theta\in\Theta} P(x,\vec{\theta}) = \sum_{\theta\in\Theta} P(x|\vec{\theta})P(\vec{\theta}) =$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\left(\frac{1}{2}\right) + \frac{1}{\sqrt{2\pi(2^2)}} e^{-\frac{1}{4}(x-10)^2}\left(\frac{1}{2}\right)$$



$$X \sim \begin{cases} \text{Binomial}(10, 0.1) & \text{w.p. } \frac{1}{4} \\ \text{Binomial}(10, 0.8) & \text{w.p. } \frac{3}{4} \end{cases}$$

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

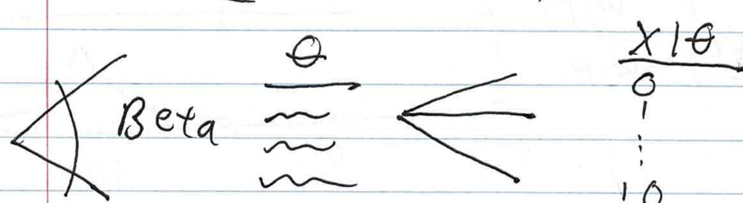


$P_X(x) \neq$ Binomial

Mixture distributions have a discrete # of components

Compound Distributions do NOT

$$X \sim \begin{cases} \text{Binomial}(n, \theta) \\ \text{Binomial}(n, \theta) \end{cases} \quad \theta\text{'s come from Beta}$$



$$P(\theta|x) = \frac{P(x|\theta)P(\theta)}{P(x)} = \frac{P(x|\theta)P(\theta)}{\underbrace{\int_{\Theta} P(x|\theta)\, P(\theta)\, d\theta}_{\text{Compound Distribution}}}$$