Math 341 Lec 8 2/28/17

1

Define $B := \dfrac{P_{Ha}(X)}{P_{H0}(X)}$

if B big $\Rightarrow$ H_a is a better model for the data, X.

Bayes Factor

denominator in Bayes Rule!! prob of data!

$$= \frac{\displaystyle\int_{\Theta_{Ha}} P(X|\theta)\, P_{Ha}(\theta)\, d\theta}{\displaystyle\int_{\Theta_{H0}} P(X|\theta)\, P_{H0}(\theta)\, d\theta}$$

oops...

$$= \frac{\displaystyle\int_{0}^{1} \cancel{\binom{100}{61}}\, \theta^{61}(1-\theta)^{100-61}\, (1)\, d\theta}{\displaystyle\int_{\{.5\}} \cancel{\binom{100}{61}}\, .5^{61}(1-.5)^{100-61}\, (1)\, d\theta}$$

for a difference

$$= \frac{B(62, 40)}{.5^{100}} = 1.39$$

H_a better model... but is it decisive?

Jeffreys 1961 scale of Bayes Factors interpretations for validity of H_a

$B < 1 \Rightarrow$ no evidence whatsoever

$B \in [1{:}1, 3{:}1] \Rightarrow$ barely worth mentioning

$B \in [3{:}1, 10{:}1] \Rightarrow$ Substantial

$B \in [10{:}1, 30{:}1] \Rightarrow$ strong

$B \in [30{:}1, 100{:}1] \Rightarrow$ very strong

$B \geq 100{:}1 \Rightarrow$ absolutely decisive

our results here .. barely worth mentioning

Famous data

testing psychokinesis (ESP)

$H_0: \theta = 0.5$

$H_a: \theta \neq 0.5$

$\alpha = 5\%$

$n = 104{,}490{,}000$

$\hat{\theta} = .50001768$

$X = 52{,}263{,}470$

[illegible]

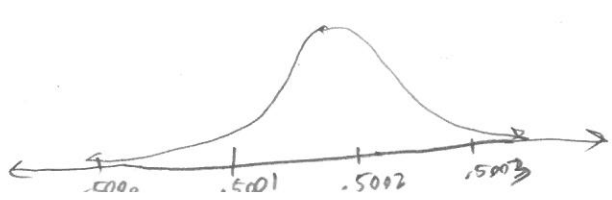
$p_{val} = .0003 < 5\% \Rightarrow$ this guy has the psychokinesis ability!!

But ... Bayesian..

$\theta | X \sim \text{Beta}(52263471, 52226531)$

$H_0: \theta = 0.5$

$H_a: \theta \sim U(0,1)$



C3

$$B = \frac{P_{H_a}(X)}{P_{H_0}(X)} = \frac{B(52{,}263{,}471,\ 52{,}226{,}53)}{.5 \ 104490000}$$

$$\cong \frac{1}{12}$$ ⇒ no evidence whatsoever!

in fact .. evidence for H_0 stronger... no ESP

paradox indeed!!

Clash of frequentist and Bayesian

⇒ Testing is very subtle

Bayes Factor also used for Model selection

Mod 1: $X|\theta \sim N(\mu, \sigma^2)$ vs Mod 2: $X|\theta \sim \text{Logistic}(\ )$

sim. distr.

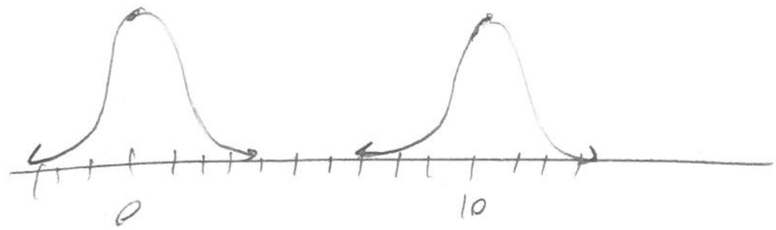
$$B = \frac{P_{M_1}(X)}{P_{M_2}(X)} \ - - -$$

Midterm I ↑

Midterm II ↓

Mixture Distr's

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



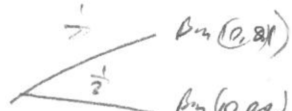
What is the PDF of X?

$$p(x) = \sum_{\theta \in \Theta} p(x|\theta)\,p(\theta) = \frac{1}{2} N(0,1^2) + \frac{1}{2} N(10,1^2)$$

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

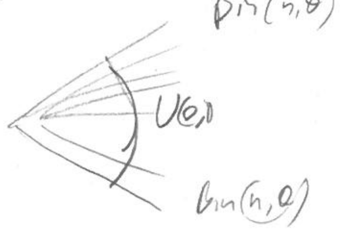
this is called a "mixture distr"

Imagine $X_1 \sim Bin(10, 0.1)$, $X_2 \sim Bin(10, 0.9)$ where each wp $\frac{1}{2}$



$$p(x) = \sum_{\theta \in \Theta} p(x|\theta)\,p(\theta) = \frac{1}{2}\binom{n}{x}\left(0.1^x\, 0.9^{10-x} + 0.9^x\, 0.1^{10-x}\right)$$

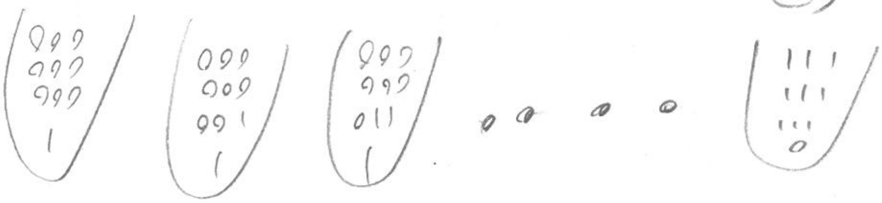
Imagine...



Hierarchical Model

$Bin(n, \theta)$ but...

$\theta \sim U(0,1)$



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

$$= \int_0^1 \left(\binom{n}{x} \theta^x (1-\theta)^{n-x}\right)(1)\, d\theta$$

$$= \binom{n}{x} B(x+1, n-x+1)$$

$X \sim \text{Bin}(n, \theta)$, $\theta \sim \text{Beta}(\alpha, \beta)$

$$p(x) = \int_{\Theta} p(x|\theta)\, p(\theta)\, d\theta = \int_0^1 \binom{n}{x} \theta^x (1-\theta)^{n-x} \frac{1}{B(\alpha,\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} d\theta$$

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

$$X \sim \text{BetaBin}(n, \alpha, \beta) := \binom{n}{x} \frac{B(x+\alpha, n-x+\beta)}{B(\alpha,\beta)}$$

new distr! Beta-Binomial

$Supp(X) = \{0, \ldots, n\}$ why?

Param space $n \in \mathbb{N}$, $\alpha, \beta > 0$ why?

named after the hierarchy / mixture

$$\mathbb{E}(X) = n\frac{\alpha}{\alpha+\beta}, \quad Var(X) = \frac{n\alpha\beta(\alpha+\beta+n)}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

if $\frac{\alpha}{\alpha+\beta} = \theta \Rightarrow \mathbb{E}(X) = n\theta$

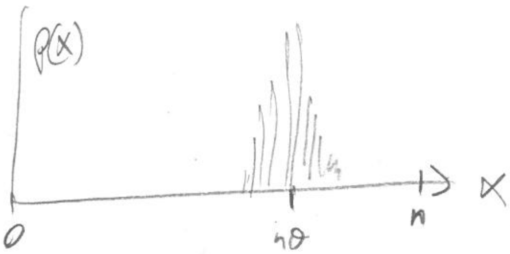
$\Rightarrow \frac{\alpha+\beta}{\alpha} = \frac{1}{\theta} \Rightarrow 1 + \frac{\beta}{\alpha} = \frac{1}{\theta} \Rightarrow \frac{\beta}{\alpha} = \frac{1}{\theta} - 1 \Rightarrow \beta = \frac{\alpha}{\theta} - \alpha$

Now let $\alpha \to \infty$

$$\lim_{\alpha\to\infty} \mathbb{E}(X) = \lim_{\alpha\to\infty} n \underbrace{\left(\frac{\alpha}{\alpha+\beta}\right)}_{\theta} \underbrace{\left(\frac{\beta}{\alpha+\beta}\right)}_{1-\theta} \quad \frac{\alpha+\overset{\frac{\alpha}{\theta}-\alpha}{\beta}+n}{\alpha+\underset{\frac{\alpha}{\theta}-\alpha}{\beta}+1} = \frac{\frac{\alpha}{\theta}+n}{\frac{\alpha}{\theta}+1} = \frac{\alpha+n\theta}{\alpha+\theta} \overset{\alpha\to\infty}{=} 1$$

In the limit of $\alpha \to \infty$, this is a binomial ... why should this be?

$Bin(n, \theta)$

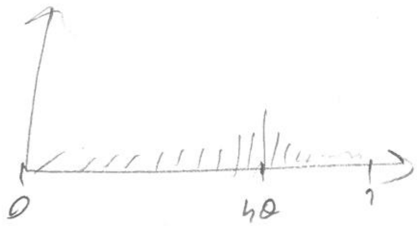


$E(X) = n\theta$

$Var(X) = n\theta(1-\theta)$

if n fixed, var fixed

$BetaBin(n, \alpha, \beta)$, $\theta := \frac{\alpha}{\alpha+\beta}$



$n\theta = n\frac{\alpha}{\alpha+\beta}$

$E(X) = n\theta$

$Var(X) = n\theta(1-\theta)\frac{\alpha+\beta+n}{\alpha+\beta+1}$

if α, β small $Var(X) \uparrow$ $n \times Var$ binomial

if $\alpha \beta$ big $Var(X) \to n\theta(1-\theta)(1-\theta)$

you can control the spread

if α, β large $Beta(\alpha, \beta) \to Deg\left(\frac{\alpha}{\alpha+\beta}\right)$

$= Deg(\theta)$

i.e. one prob

i.e. a binomial

$P(Male) = 0.511 \neq 0.5$

why? No one knows

Is this useful?

Gender births data, 6115 families with ≥ 13 children, genders of first 12 children...

# Boys	0	1	2	3	4	5	6	7	8	9	10	11	12	
X	3	24	104	286	670	1033	1343	1112	829	478	181	45	7	6115
$Bin(12, 0.511)$	1	12	72	259	628	1085	1367	1266	854	410	132	26	2	6118
$BetaBin(12, 34, 32)$	2	23	105	311	656	1036	1258	1182	854	462	178	~~44~~	5	6115

$\frac{34}{32+34} = .515$

via MLE

→ doesn't fit well in tails (extremes)

Meaning?

Women have different propensities to have males?

Return to Bayesian setup

$X|\theta \sim Bin(n,\theta)$ $\qquad P(\theta|X) = \dfrac{P(X|\theta)\,P(\theta)}{P(X)}$

$\theta \sim Beta(\alpha,\beta)$

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

Now, what if we want $X^*|X$ where X^* is next obs.

we have shown that

$X^*|X \sim \cancel{Bern}\left(\dfrac{x+\alpha}{n+\alpha+\beta}\right)$

but what if X^* is n^* next obs? $X^*|X \sim$?

if θ known... $X^*|\theta \sim Bin(n^*, \theta)$ but θ unknown..

but we have idea on θ represented by $\theta|X$, the posterior...

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

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

Same thing we saw before!!!

$= BetaBin\left(n^*,\ \alpha+x,\ \beta+n-x+1\right)$

$X|\theta \sim Bin(n,\theta)$

$\theta \sim Beta(\alpha,\beta)$

$\Rightarrow P(\theta|X) = \dfrac{P(X|\theta)\,P(\theta)}{P(X)}$ $\longrightarrow$ prior predictive distr.

$X \sim$? $\qquad P(X) = \displaystyle\int_{\Theta} P(X|\theta)P(\theta)\,d\theta$ same thing!!!

$= \displaystyle\int_0^1 \binom{n}{x}\theta^x(1-\theta)^{n-x}\frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}\,d\theta = BetaBin(n,\alpha,\beta)$