$$= \frac{\int_0^1 \theta^{61}(1-\theta)^{39}\,d\theta}{\underbrace{.5^{61}(1-.5)^{39}}_{.5^{100}}} = \frac{B(62,40)}{.5^{100}} = 1.39$$

2/28/17

Bayesian Hypothesis Testing

$B := \frac{P_{H_a}(x)}{P_{H_0}(x)}$ Bayes Factor

$$= \frac{\int_{\Theta_{H_a}} P_{H_a}(x|\theta)\,P_{H_a}(\theta)\,d\theta}{\int_{\Theta_{H_0}} P_{H_0}(x|\theta)\,P_{H_0}(\theta)\,d\theta}$$

$H_0: \theta = 0.5 \quad \} \ \theta \sim Deg(0.5)$

$H_a: \theta = \theta \neq 0.5 \quad \} \ \theta \sim U(0,1)$

$f = \text{Binomial}$

$n = 100 \quad x = 61$

$$= \frac{\int_0^1 \binom{n}{x}\theta^{x}(1-\theta)^{n-x}(1)\,d\theta}{\int_{\{0.5\}} \binom{n}{x} 0.5^{x}(1-0.5)^{n-x}(1)\,d\theta} = \frac{B(x+1,\ n-x+1)}{0.5^{n}}$$

$$= \frac{B(62,40)}{0.5^{100}} = 1.39$$

$B < 1:1 \Rightarrow$ no evidence

$B \in [1:1,\ 3:1] \Rightarrow$ 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

$B > 100:1 \Rightarrow$ decisive

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

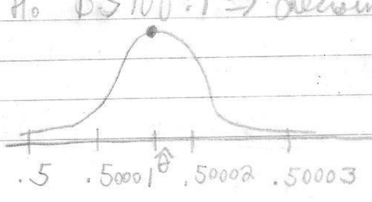
$x = 52{,}263{,}920$

$\hat{\theta} = .50001768$

$P_{val} = .0003 < 5\% \Rightarrow$ Reject H_0

$\theta \sim Beta(1,1)$

$\theta | x \sim Beta(\quad , \quad)$



n large

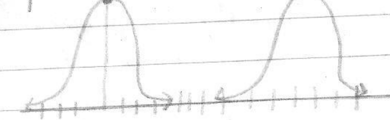
$$B = \frac{(52, 263, 92, 52, 226, \dots)}{.5^{104490000}} = \frac{1}{12} \Rightarrow B < 1$$

$\Rightarrow$ no evidence whatsoever

MIDTERM ONE MATERIAL CONCLUDED

Mixture Distributions

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



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

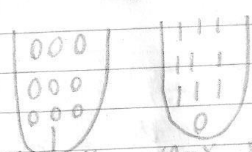
$\frac{1}{2}$ $N(10, 1^2)$

$$P(X) = \sum_{\theta \in \Theta} P(X|\theta)P(\theta) = P(X|\theta=0)P(\theta=0) + P(X|\theta=10)P(\theta=10)$$

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

$\frac{1}{2}$ $\text{Bin}(10, 0.1)$

$\frac{1}{2}$ $\text{Bin}(10, 0.9)$



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

$$= P(X|\theta=0.1)P(\theta=0.1) + P(X|\theta=0.9)P(\theta=0.9)$$

$$= \binom{10}{x} 0.1^x \, 0.9^{10-x} \cdot \frac{1}{2} + \binom{10}{x} 0.9^x \, 0.1^{10-x} \cdot \frac{1}{2}$$

Hierarchical Model

$$\theta \sim \begin{cases} 0.1 & \text{wp } \frac{1}{2} \\ 0.9 & \text{wp } \frac{1}{2} \end{cases}$$

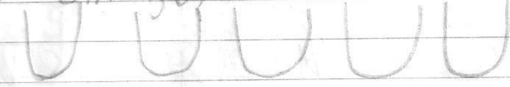
$$X|\theta \sim \text{Bin}(n, \theta)$$

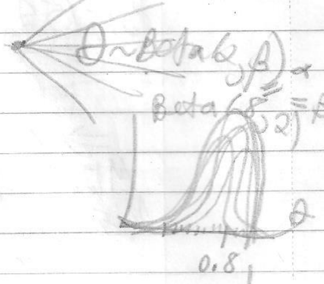
over

Mixture Distribution

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

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





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

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

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

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

Beta-Binomial

$Supp(X) = \{0, 1, 2, \dots, n\}$

Parameter Space: $n \in \mathbb{N}$

$\alpha > 0$

$\beta > 0$

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

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

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

$$\Rightarrow E(X) = n\theta$$

$$\lim_{\alpha \to \infty} E(X) = n\theta \qquad \lim_{\alpha \to \infty} Var(X) = \lim_{\alpha \to \infty} n \underbrace{\frac{\alpha}{\alpha+\beta}}_{\theta} \cdot \underbrace{\frac{\beta}{\alpha+\beta}}_{1-\theta} \cdot \frac{\alpha+\beta+n}{\alpha+\beta+1}$$

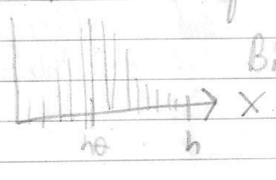
$$= \underbrace{n\theta(1-\theta)}_{\text{variance of binomial}} \lim_{\alpha \to \infty} \frac{\cancel{\alpha} + \frac{\alpha}{\theta} - \cancel{\alpha} + n}{\cancel{\alpha} + \frac{\alpha}{\theta} - \cancel{\alpha} + 1} = \lim_{\alpha \to \infty} \frac{\alpha + n\theta}{\alpha + \theta} = 1$$

Irreducible noise

$$E(X) = n\,\frac{\alpha}{\alpha+\beta} \qquad Var(X) = \frac{n\alpha\beta}{(\alpha+\beta)^2}\left[\frac{(\alpha+\beta+n)}{(\alpha+\beta+1)}\right]$$

$\in (1, n)$

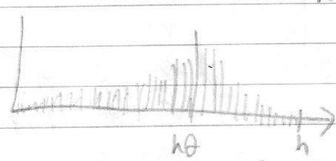
Variance-inflated binomial



$Bin(n, \theta)$

$E(X) = n\theta$

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



Gender Birth Data

6,115 females with >12 children

count # of boys in 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	823	478	181	45	7	6115
Bin(12, 0.511)	1	12	72	259	628	1085	1367	1266	854	410	152	26	2	6115
Betabin(12, 34, 32)	2	32	105	311	656	1036	1258	1182	854	462	178	47	5	

$E(X) = \frac{34}{34+32} = 0.515$

$P(Male) = 0.511$

Inflate variance using BetaBin model

ON TEST

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

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

$X^*|X \sim Beta\left(\frac{x+\alpha}{n+\alpha+\beta}\right) \longrightarrow \hat{\theta}_{MMSE}$ post. exp.

X^* is the next one observation

What if X^* was the next n^* observations?

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

Beta

All Files

notepad

. .tex

$X^*|\theta$

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

$= \text{BetaBin}(n^*, \alpha+x, \beta+n-x)$

$P(\theta|x)$: if loose, inflate variance
if tight, will know θ better

3/2/17

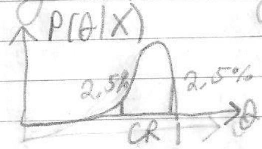
$\theta \sim \text{Beta}(\alpha, \beta)$

$E(\theta) = \frac{\alpha}{\alpha+\beta} = \frac{1}{2}$ if $\alpha = \beta$

$x|\theta \sim \text{Binom}(n, \theta)$

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

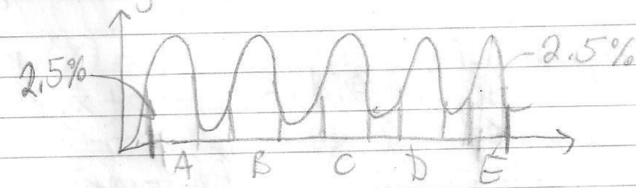
Credible region is a Bayesian confidence interval



Two-sided credible region

$P(\theta \in CI_{\theta, 1-\alpha}) = 1-\alpha$

$P(\theta \in CR_{\theta, 1-\alpha}) = 1-\alpha$



HDR method

$CR = A \cup B \cup C \cup D \cup E$