2/28/17

$$B := \frac{P_{Ha}(x)}{P_{H_0}(x)} = \frac{\int_{\Theta_{Ha}} P_{Ha}(x\mid\theta)\, P_{Ha}(\theta)\, d\theta}{\int_{\Theta_{H_0}} P_{H_0}(x\mid\theta)\, P_{H_0}(\theta)\, d\theta}$$

<u>example</u>: $H_0: \theta = 0.5$ } $\theta \sim Deg(0.5)$

$H_a: \theta \neq 0.5$

$F$ = Binomial

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

$n = 100$

$x = 61$

$$= \frac{\int_0^1 \cancel{\binom{n}{x}}\, \theta^x (1-\theta)^{n-x}\, \overbrace{(1)}^{\text{b/c uniform}}\, d\theta}{\int_{\{0.5\}} \cancel{\binom{n}{x}}\, 0.5^x (1-.5)^{n-x}\, \underbrace{(1)}_{\text{Deg}}\, d\theta}$$

$$= \frac{B(x-1,\ n-x+1)}{0.5^n}$$

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

if $B < 1 \Rightarrow$ 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

$H_0: \theta = 0.5$

$H_a: \theta \neq 0.5$

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

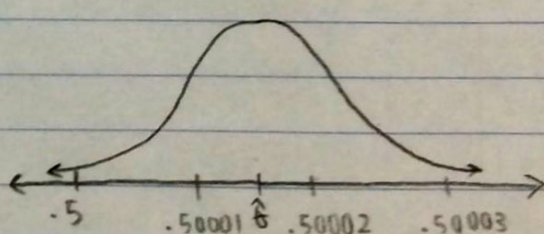
$x = 52{,}263{,}920$

$\hat{\theta} = .50001768$

pval $= .0003 < 5\% \Rightarrow$ reject $H_0$

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

$\theta \mid x \sim Beta(52{,}263{,}921,\ 52{,}226{,}080)$




$$B = \frac{B(52{,}263{,}921,\ 52{,}226{,}080)}{.5^{104{,}490{,}000}} = \frac{1}{12} \Rightarrow B < 1 \Rightarrow \text{no evidence}$$

why different? when $n\uparrow$, $H_0$ will be rejected

Midterm 1 ⸺↑⸺

## Mixture Distributions:

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

$\frac{1}{2}$ — $N(0,1)^2$

$\frac{1}{2}$ — $N(10,1^2)$

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

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

$\frac{1}{2}$ — $Bin(10, 0.1) \to$ 9 0's & 1 1's

$\frac{1}{2}$ — $Bin(10, 0.9) \to$ 9 1's & 1 0's

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

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

$$P(X) = \sum_{\theta \in \Theta} P(x|\theta)P(\theta) = P(x|\theta=0.1)P(\theta=0.1) + P(x|\theta=0.9)P(\theta=0.9)$$

$$= \binom{10}{x} 0.1^x \, 0.9^{10-x}\left(\tfrac{1}{2}\right) + \binom{10}{x} 0.9^x \, 0.1^{10-x}\left(\tfrac{1}{2}\right)$$

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

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

$Beta(\alpha, \beta)$ — $Bin(n, \theta)$

$\Rightarrow$ different $\theta$ for each "bag"

$\theta$ is continuous so:

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

ie) average of what happened *

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

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

$supp[X] = \{0, 1, \ldots, n\}$

parameter space: $n \in \mathbb{N}$, $\alpha, \beta > 0$

$E[X] = n\frac{\alpha}{\alpha+\beta}$ $\in (1, n)$

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

variance-inflated binomial.

let $\theta = \overset{\nearrow \text{pinning}}{\dfrac{\alpha}{\alpha+\beta}} \Rightarrow \beta = \dfrac{\alpha}{\theta} - \alpha$

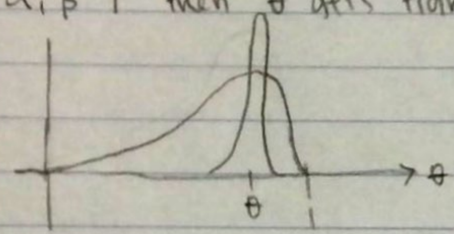
$\Rightarrow E[X] = n\theta$

$\lim\limits_{\alpha \to \infty} E[X] = n\theta \qquad \lim\limits_{\alpha \to \infty} Var[X] = \lim\limits_{\alpha \to \infty} n \overbrace{\dfrac{\alpha}{\alpha+\beta}}^{\theta} \cdot \overbrace{\dfrac{\beta}{\alpha+\beta}}^{1-\theta} = \dfrac{\alpha+\beta+n}{\alpha+\beta+1}$

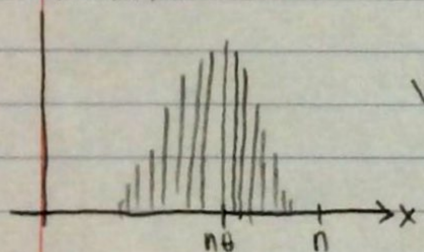
$= \overbrace{n\theta(1-\theta)}^{\text{variance of binom}} \lim\limits_{\alpha \to \infty} \dfrac{\cancel{\alpha} + \frac{\alpha}{\theta} - \cancel{\alpha} + n}{\cancel{\alpha} + \frac{\alpha}{\theta} - \cancel{\alpha} + 1}$

$\lim\limits_{\alpha \to \infty} \dfrac{\alpha + n\theta}{\alpha + \theta} = 1$

$\alpha, \beta \uparrow$ then $\theta$ gets tighter & becomes degenerate (more binomial-like)

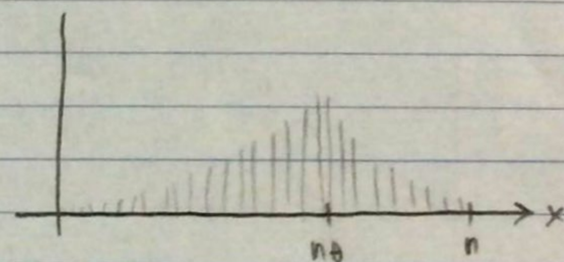



Bin(n, θ)




$E[X] = n\theta$

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




<u>Gender Birth Data</u>

6,115 females with > 12 children $\qquad P(\text{male}) = 0.511$

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 |
| Beta Bin(12, 34, 32) | 2 | 23 | 105 | 311 | 656 | 1036 | 1258 | 1132 | 854 | 462 | 178 | 47 | 5 | |

$\hookrightarrow E[X] = \dfrac{34}{34+32} = 0.515$

* tails are not close enough, use Beta Binomial to inflate the variance.

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

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

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

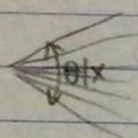
$\nearrow$ next 1 observation

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

↳ $\hat{\theta}_{MMSE}$ = posterior expectation

what if $x^*$ is next $n^*$ observations?

$$P(x^*|x) = \int_{\Theta} \underbrace{P(x^*|\theta)}_{\text{Binom}} \underbrace{P(\theta|x)}_{\text{Beta}} d\theta = \int_0^1 \binom{n^*}{x^*} \theta^{x^*} (1-\theta)^{n^*-x^*} \frac{1}{B(\alpha+x, \beta+n-x)} \theta^{x+\alpha-1} (1-\theta)^{n-x+\beta-1} d\theta$$




$X^*|\theta$

$$= BetaBinom(n^*, \alpha + x, \beta + n - x)$$

Project ideas of $\theta$ into future $x^*$