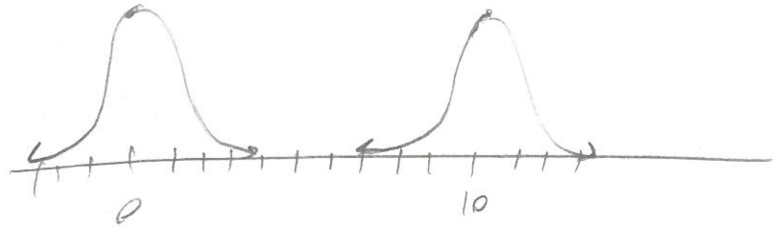
Lec 10

Math 341 3/12/18

Mixture Distr's

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



What is the PDF of X?

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

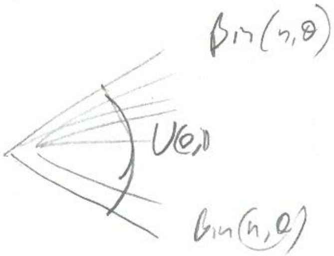
$$= \frac{1}{2} \frac{1}{\sqrt{2\pi}} \left(e^{-\frac{1}{2}(x)^2} + e^{-\frac{1}{2}(x-10)^2} \right)$$

this is called a "mixture distr"

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

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

Imagine...



Hierarchical Model

Bin(n,θ) but...

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

Now let's go back to $X|\Theta \sim \text{Binom}(n,\theta)$, $\theta \sim \text{Beta}(1,1) = U(0,1)$ ...
and look at the beginning again:

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

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

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

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

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

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

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

"Overdispersed" Binomial

New distr! Beta-Binomial

named after the hierarchy / mixture

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

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

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

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 $\alpha \to \infty$

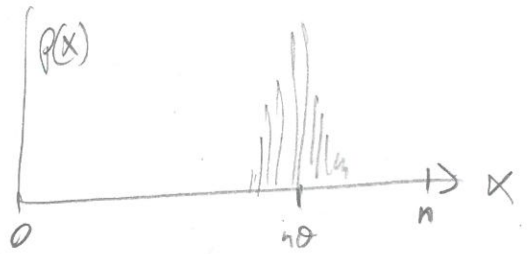
$$\lim_{\alpha\to\infty} \mathbb{E}(X) = \lim_{\alpha\to\infty} n \underbrace{\left(\frac{\alpha}{\alpha+\beta}\right)}_{(\theta)} \underbrace{\left(\frac{\beta}{\alpha+\beta}\right)}_{(1-\theta)} \frac{\alpha+\overbrace{\beta}^{\frac{\alpha}{\theta}-\alpha}+n}{\alpha+\underbrace{\beta}_{\frac{\alpha}{\theta}-\alpha}+1}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

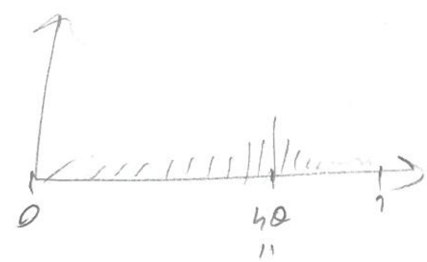
In the limit of $\alpha \to \infty$, this is a binomial ... why should this be?

Beta

$\mathrm{Bin}(n, \theta)$

$\mathrm{Beta\,Bin}(n, \alpha, \beta)$, $\theta := \frac{\alpha}{\alpha+\beta}$





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

$E(X) = n\theta$

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

if n fixed, Var fixed

$E(X) = n\theta$

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

if α, β small $Var(X) \uparrow$

if α, β big $Var(X) \to Var[\text{Binomial}]$

Overdispersion or (flexibility)

you can control the spread

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

$= \mathrm{Deg}(\theta)$

i.e. one prob

i.e. a binomial

$P(\text{Male}) = 0.511 \neq 0.5$

why? No one knows

Is this useful?

Gender birth 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 eqs

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

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

4

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

we have shown that

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

But what if X^* is n^* test obs? $X^*|X \sim$?

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

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

$$P(x^*|x) = \int_{\Theta} P(x^*|\theta)\,P(\theta|x)\,d\theta \quad \text{mixture distr !!}$$

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

Same thing we saw before!!!

$$= \text{BetaBin}(n^*,\ \underbrace{\alpha+x}_{\alpha'},\ \underbrace{\beta+n-x}_{\beta'}) = \binom{n^*}{x^*}\frac{B(\alpha+x+x^*,\ \beta+n-x+n^*-x^*)}{B(x+\alpha,\ \beta+n-x)}$$

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

$$X|\theta \sim \text{Bin}(n,\theta)$$
$$\theta \sim \text{Beta}(\alpha,\beta)$$
$$\Rightarrow P(\theta|x) = \frac{P(x|\theta)\,P(\theta)}{P(x)}$$

prior predictive distr.

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

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