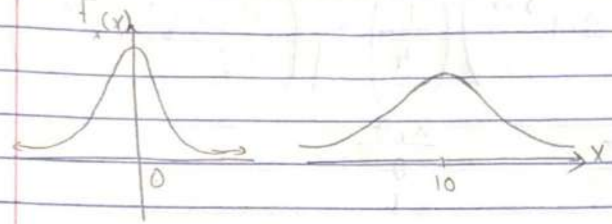
Lecture - 09 03/03/2020

$$X \sim \begin{cases} N(0, 1^2) & \text{up } 1/2 \\ N(10, 2^2) & \text{up } 1/2 \end{cases} \qquad X \sim \begin{cases} Bin(10, 0.1) & \text{up } 1/4 \\ Bin(10, 0.6) & \text{up } 3/4 \end{cases}$$



$$p(x) = \begin{cases} \int_{\Theta} p(x|\theta)\, p(\theta)\, d\theta. \\ \sum_{\theta \in \Theta} p(x|\theta)\, p(\theta) \end{cases}$$

$$p(\theta|x) = \frac{\overbrace{p(x|\theta)\ p(\theta)}^{\text{assumed with F}}}{\underbrace{p(x)}_{\Downarrow}}$$

$$\int_{\Theta} \underbrace{p(x|\theta)}_{\substack{\text{likelihood} \\ \text{model}}} \ \underbrace{p(\theta)}_{\substack{\text{mixing proportions} \\ \text{(prior)}}} d\theta$$

F : Binomial

$p(\theta) = Beta(\alpha, \beta)$

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

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

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

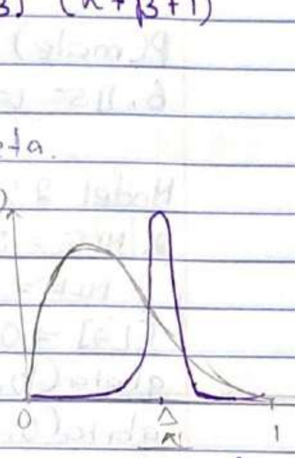
$$= \binom{n}{x} \frac{\beta(x+\alpha,\ n-x+\beta)}{\beta(\alpha+\beta)}$$

$$= x \sim \text{Beta binomial}(n, \alpha, \beta)$$

- $E[x] = \sum_{x=0}^{n} x\, p(x) = \dots = n \cdot \left(\frac{\alpha}{\alpha+\beta}\right)$
- $Supp[x] = \{0, 1, \dots, n\}$
- $Var[x] = \dots = n \cdot \frac{\alpha\beta(\alpha+\beta+n)}{(\alpha+\beta)^2(\alpha+\beta+1)}$

Parameter space $n \in \mathbb{N}$

$\left.\begin{array}{l}\alpha > 0 \\ \beta > 0\end{array}\right)$ from Beta.



A different parameterization

Let $\theta := \frac{\alpha}{\alpha+\beta} \Longrightarrow \beta = \alpha \cdot \frac{(1-\theta)}{\theta}$

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

expectation of x

$\Longrightarrow E[x] = n\theta$

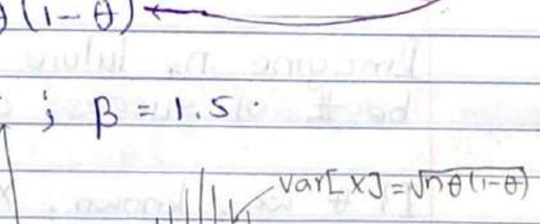
$\Longrightarrow Var[x] = \dots = n\theta(1-\theta) \cdot \frac{\alpha/\theta + n}{\alpha/\theta + 1}$

$\Longrightarrow \lim_{x \to \alpha} Var[x] = n\theta(1-\theta) \lim_{x \to \alpha} \frac{\alpha/\theta + n}{\alpha/\theta + 1}$

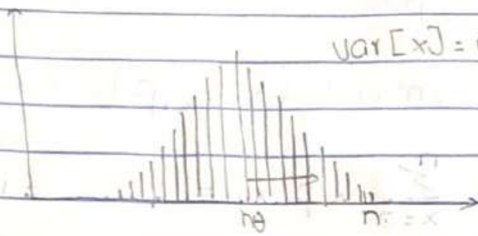
$= n\theta(1-\theta)$

$\theta = 0.7 = \alpha/\alpha+\beta$; $\alpha = 3.5$; $\beta = 1.5$.

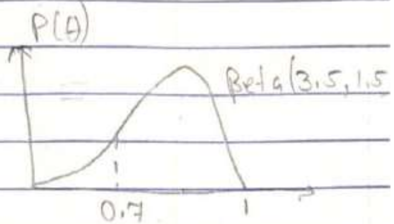
$x \sim \text{Binomial}(n, \theta)$



$x \sim \beta etabin(n, \alpha, \beta)$ "Over disposal binomial"



$var[x] = n\theta(1-\theta) \cdot \frac{\alpha/\theta + n}{\alpha/\theta + 1}$

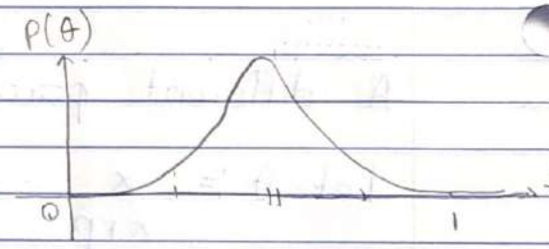


Birth Gender Data Example

Model 1: $Bin(12, 0.51)$
Model 2: $\beta etabin(12, 34, 32)$
$\alpha MLE = 34$
$\beta MLE = 32$

$P(male) = 0.51$
6,115 women with ≥ 12 children.

Model 2: $\beta etabin(12, 34, 32)$
α MLE = 34
β MLE = 32.
$E[\theta] = 0.515$
qbeta(0.005, 34, 32) = 0.36
qbeta(0.995, 34, 32) = 0.67



	# of males	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	6,115
f_1 mod 1:	prediction	1	12	72	259	628	1085	1367	1266	854	410	152	26	2	6,115
f_0 mod 2:	prediction	2	23	105	311	656	1036	1258	1182	854	462	178	44	5	6,115

F: Binomial fixed n
$P(\theta) = \beta eta(\alpha, \beta) \implies P(\theta | x) = \beta eta(x + \alpha, n - \cancel{\alpha} + \beta)$

Imagine n_* future observation where $n_* \geq 1$. Let X_* be # of success of n_* future observations.

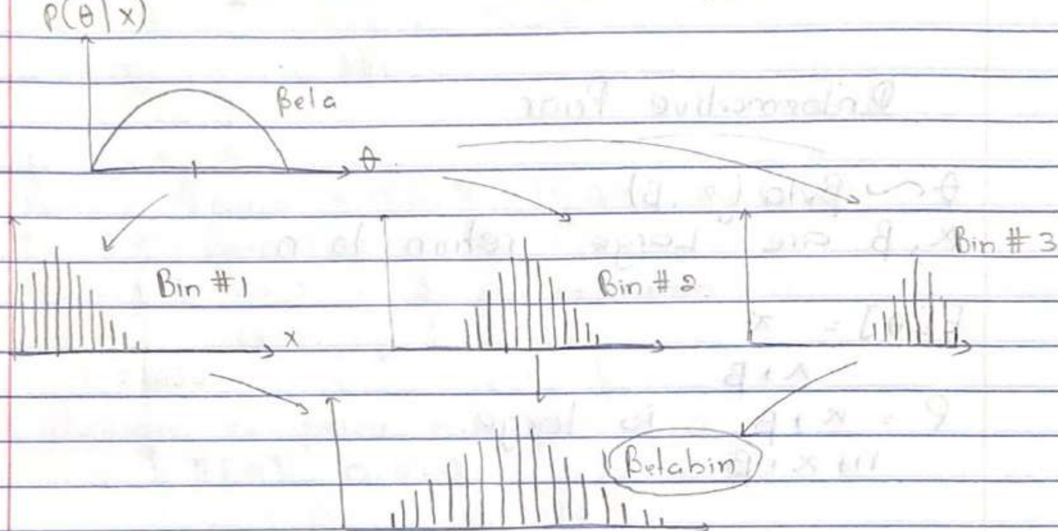
If θ was known, $X_* \sim Bin(n_*, \theta)$

In real life θ is unknown. Let's use Bayesian Inference. We obtain $P(\theta | x)$

$$P(X_* | x) = \int_{\Theta} \underbrace{P(X_* | \theta)}_{= \, Bin(n_*, \theta)} \underbrace{P(\theta | x)}_{= \, Beta(\alpha + x, \beta + n - x)} d\theta.$$

posterior predictive distribution

$$= Betabin(n_*, \alpha + x, \beta + n - x)$$



F: Binomial fixed n

$\Rightarrow P(\theta) = Beta(\alpha, \beta)$

$\Rightarrow P(\theta | x) = Beta(x + \alpha, n - x + \beta) \Rightarrow \hat{\theta}_{MMSE} = \frac{x + \alpha}{n + \alpha + \beta}$

$P(\theta) = Beta(1, 1) = U(0, 1)$ indifferent prior, an example of an uninformative prior (Laplace)

$P(\theta) = Beta(\frac{1}{2}, \frac{1}{2})$

Haldare Prior (1932)

$P(\theta) = Beta(0, 0)$ but not a legal distribution. ("improper")
Haldare retorts, "I don't care!"

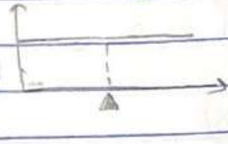
$\Longrightarrow p(\theta|x) = \text{Beta}(x, n-x)$ will be "proper"
If $x \neq 0$ and $x = n$ $\Rrightarrow \hat{\theta}_{MMSE} = x/n = \hat{\theta}_{MLE}$ (no shrinkage)

(Haldane) Objectivist : The data must speak for itself.

$P(\theta) = U(0,1) = \text{Beta}(1,1) \Longrightarrow n_0 = 2 \Longrightarrow \rho = \dfrac{\alpha+\beta}{\alpha+\beta+n} > 0.$

$\Downarrow$

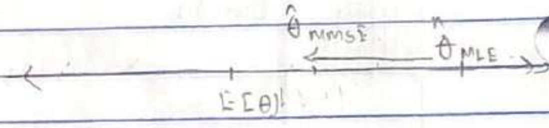
$E[\theta] = 0.5$



Informative Prior

$\theta \sim \text{Beta}(\alpha, \beta)$
α, β are "Large" return to n

$E[\theta] = \dfrac{\alpha}{\alpha+\beta}$



$\rho = \dfrac{\alpha+\beta}{n+\alpha+\beta}$ is large.