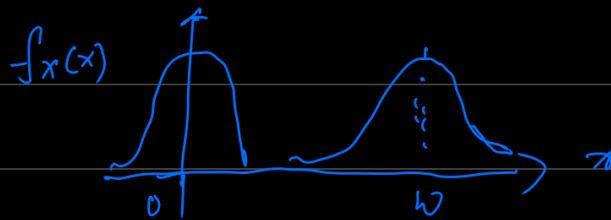
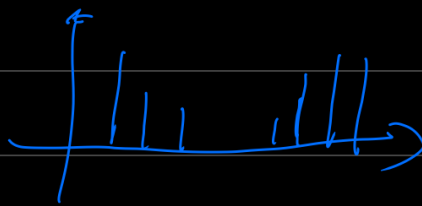


$$X \sim \begin{cases} N(0, 1^2) & \text{wp } \frac{1}{2} \\ N(0, 2^2) & \text{wp } \frac{1}{2} \end{cases}$$



$$X \sim \begin{cases} \text{Bin}(0, 0.1) & \text{wp } \frac{1}{4} \\ \text{Bin}(0, 0.6) & \text{wp } \frac{3}{4} \end{cases}$$



$$p(x) = \int_{\Theta} p(x|\theta) p(\theta) d\theta = \sum_{\theta \in \Theta} P(x|\theta) p(\theta) d\theta$$

$$P(\theta|x) = \frac{P(x|\theta) P(\theta)}{p(x)} \rightarrow \int_{\Theta} P(x|\theta) P(\theta) d\theta$$

likelihood prior

$$\text{Binomial} \quad p(\theta) = \text{Beta}(d, \ell) \Rightarrow P(\theta|x) = \text{Beta}(x+d, n-x+\ell)$$

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

$$= \text{Beta Binomial}(n, d, \ell)$$

$$\text{Supp } X = \{0, 1, \dots, n\}$$

$$E(X) = \sum_{x=0}^n x p(x) = \frac{n d}{d+\ell}$$

$$\text{Var}(X) = n \frac{d\ell(d+\ell+n)}{(d+\ell)^2(d+\ell-1)} \quad n \in \mathbb{N} \quad d, \ell > 0$$

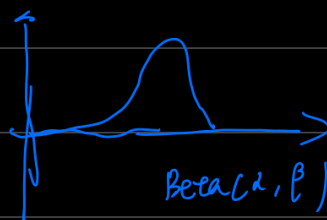
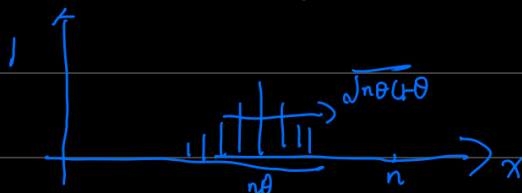
$$\alpha = \frac{d\theta}{(1-\theta)}$$

$$\frac{d}{d\theta} \frac{\ell}{d\theta}$$

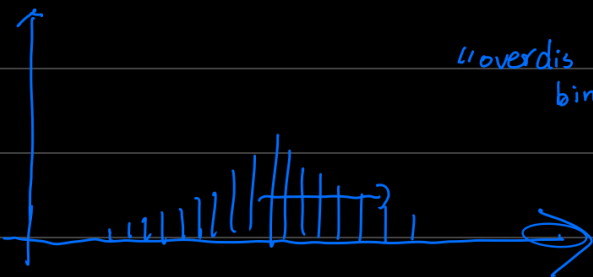
A difference parametrization let: $\theta = \frac{d}{d+\ell} \Rightarrow \ell = \frac{d(1-\theta)}{\theta} \quad X \sim \text{BetaBin}(n, d, \ell) = \text{BetaBin}(n, \theta, d)$

$$E(X) = n\theta \quad \text{Var}(X) = \dots = n\theta(1-\theta) \frac{\frac{d}{\theta} + n}{\frac{d}{\theta} + 1} \quad \lim_{d \rightarrow \infty} \text{Var}(X) = n\theta(1-\theta) \lim_{d \rightarrow \infty} \frac{\frac{d}{\theta} + n}{\frac{d}{\theta} + 1} = n\theta(1-\theta)$$

$$\theta = 0.7 = \frac{d}{d+\ell}, \text{ if } d=3.5 \Rightarrow \ell=1.5 \quad X \sim \text{Bin}(n, \theta)$$



"overdispersion"



$$\text{BetaBin}(n, \alpha, \beta)$$

A model: Birth Gender Data example $p(\text{Male}) = 0.51$ 6,115 ^{women} will more than ≥ 13 children.

# Boy	0	1	2	3	4	5	6	7	8	9	10	11	12	
X	3	24	104	286	670	1033	1340	1112	829	478	181	45	7	6115
Mod 1 pr $\hat{\theta}$	1	12	72	237	628	1085	1367	1266	854	410	152	26	2	6115
Mod 2 $\hat{\theta}_2$	2	23	105	311	656	1036	1258	1182	854	462	178	94	5	

$X|\theta$ 服从 Beta Bin (12, 34, 32) 分布.

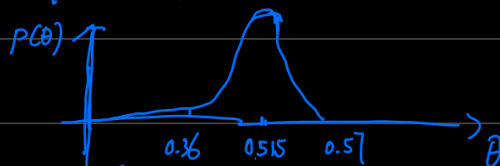
Model 1: $\text{Bin}(12, 0.51)$

Model 2: $\text{BetaBin}(12, 34, 32)$

$d_{MLE} = 34$ $\hat{\theta}_{MLE} = 32$ $E(\theta) = 0.515$

$q_{\text{beta}}(0.005, 34, 32) = 0.36$

$q_{\text{beta}}(0.995, 34, 32) = 0.57$



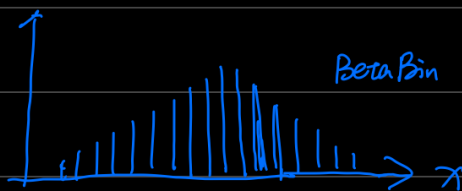
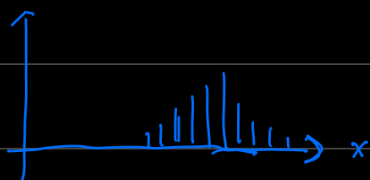
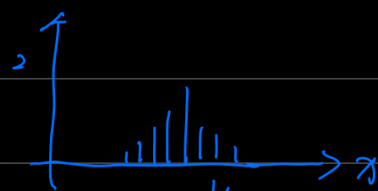
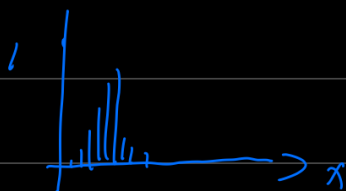
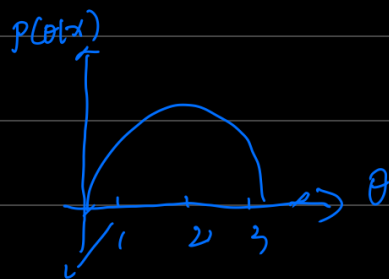
$\hat{\theta}$ Binomial fixed n $p(\theta) = \text{Beta}(d, \beta) \Rightarrow p(\theta|x) = \text{Beta}(x+d, n-x+\beta) \Rightarrow \hat{\theta}_{\text{marg}} = \frac{x+d}{n+d+\beta}$

Imagine n finite observations $n \rightarrow \infty$, X_n be # of success if θ was known.

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

In real life θ is unknown let's Bayes $\Rightarrow p(\theta|x)$

$$P(X_n | x) = \int_{\theta} P(X_n | \theta) p(\theta|x) d\theta = \text{Beta Bin}(n, d+x, \beta+n-x) \text{ posterior prediction dist.}$$



(Laplace)
For $\hat{\theta}$ Binomial indifferent prior, an example of an uninformative prior

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

Halldare Prior (1932) $p(\theta) = \text{Beta}(0, 0)$ not a legal dist ("improper") don't care.

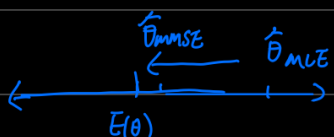
$\Rightarrow p(\theta|x) = \text{Beta}(x, n-x)$ will be proper if $x \neq 0$ and $x \neq n$

$$\hookrightarrow \hat{\theta}_{\text{marg}} = \frac{x}{n} = \hat{\theta}_{MLE} \text{ (no shrinkage)}$$

Objectivist: the data must speak for itself. (Halldare)

$$p(\theta) = U(0,1) = \overset{E(\theta)=0.5}{\text{Beta}(1,1)} \Rightarrow n_0=2 \Rightarrow \rho = \frac{\alpha+\beta}{\alpha+\beta+n} > 0$$

Informative prior: $\theta \sim \text{Beta}(\alpha, \beta)$ α, β related to n s.t. "large"



$E(\theta) = \frac{\alpha}{\alpha+\beta}$ $\rho = \frac{\alpha+\beta}{n+\alpha+\beta}$ is large