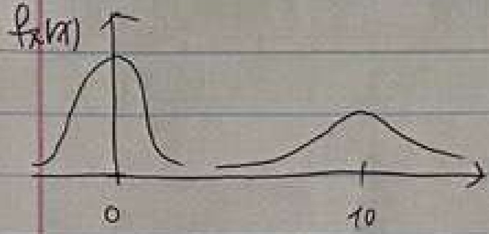
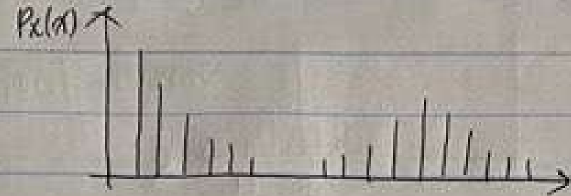
Lecture 9

3/3 上. [Math 341]

$X \sim \begin{cases} N(0, 1^2) & w.p.\ \frac{1}{2} \\ N(10, 2^2) & w.p.\ \frac{1}{2} \end{cases}$ $\qquad$ $X \sim \begin{cases} Bin(10, 0.1) & w.p.\ \frac{1}{4} \\ Bin(10, 0.6) & w.p.\ \frac{3}{4} \end{cases}$





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

$P(\theta|x) = \dfrac{\overbrace{P(x|\theta)}^{\text{assumed with } F} P(\theta)}{P(x)}$

$= \int_\Theta \underbrace{P(x|\theta)}_{\text{likelihood model}} \underbrace{P(\theta)}_{\text{mixing proportion (prior)}} d\theta$

F: Binomial. $P(\theta) = Beta(\alpha, \beta)$.

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

$P(x) = ?$

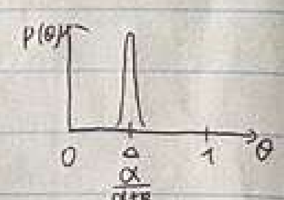
$P(x) = \int_0^1 \left(\binom{n}{x} \theta^x (1-\theta)^{n-x}\right) \left(\frac{1}{B(\alpha,\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}\right) d\theta.$

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

$= \binom{n}{x} \dfrac{B(x+\alpha, n-x+\beta)}{B(\alpha,\beta)}$ $\quad X \sim$ Beta Binomial(n, α, β)

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

parameter space $\begin{cases} n \in \mathbb{N} \\ \alpha > 0 \\ \alpha > 0 \end{cases}$) from Beta.



A different parametrization

Let $\theta := \dfrac{\alpha}{\alpha+\beta} \Rightarrow \beta = \alpha \dfrac{1-\theta}{\theta}$

$X \sim BetaBinomial(n, \alpha, \beta) = Beta\,Bin(n, \theta, \alpha)$

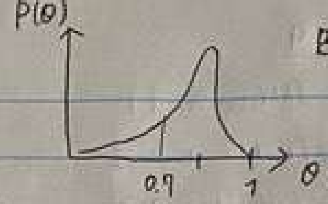
$E[X] = n \cdot \theta$

$Var[X] = \dots = n\theta(1-\theta) \dfrac{\frac{\alpha}{\theta} + n}{\frac{\alpha}{\theta} + 1}$

$\lim_{\alpha \to \infty} Var[X] = n \cdot \theta(1-\theta) \cdot \lim_{\alpha \to \infty} \dfrac{\frac{\alpha}{\theta} + n}{\frac{\alpha}{\theta} + 1}$

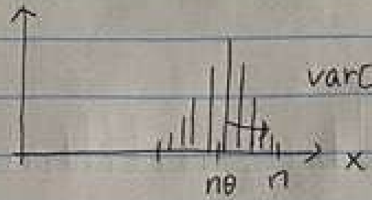
$= n \cdot \theta(1-\theta)$

$Beta(3.5, 1.5)$



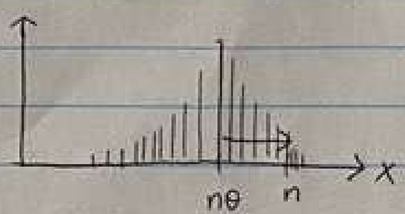
$\theta = 0.7 = \frac{\alpha}{\alpha+\beta}$, $\alpha = \overset{3.5}{\cancel{}} \Rightarrow \beta = \overset{1.5}{\cancel{}}$

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



$var[X] = \sqrt{n\theta(1-\theta)}$

$X \sim Betabin(n, \alpha, \beta)$ "Overdisposal Binomial"



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

$\alpha_{MLE} = 34$
$\beta_{MLE} = 32$

Birth Gender Data Example

$P(MALE) = 0.51$

6,115 women with 213 children

- Model1 : $Bin(\overset{12}{\cancel{}}, 0.51)$
- Model2 : $Betabin(12, 34, 32)$

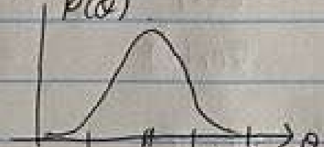
# of males	0	1	2	3	4	5	6	7	8	9	10	11	12	~~13~~	
X	3	24	104	286	670	1033	1343	1112	829	478	181	45	7		6115
F_1 Model1 prediction	1	12	72	259	628	1085	1367	1266	854	410	152	26	2		6115
F_2 Model2 "	2	23	105	311	656	1036	1253	1192	854	462	178	44	5		

- Model2 : $BetaBin(12, 34, 32)$ ← fits better

$\alpha_{MLE} = 34$

$\beta_{MLE} = 32$

- $E[\theta] = 0.515$
- qbeta(0.005, 34, 32) ≈ .36
- qbeta(0.995, 34, 32) ≈ .37



$Bin(12, 0.51) \Rightarrow$ 모든 여자, 0.51 확률로 남자아이 낳음

$BetaBin(12, 34, 32) \Rightarrow$ 각 여자들이 남자아이 낳을 확률 다름

ex) 누구는 1/3, 누구는 2/3 으로 etc

F Binomial fixed n.

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

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

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

- If θ was known,

  $X_* \sim Bin(n_*, \theta)$

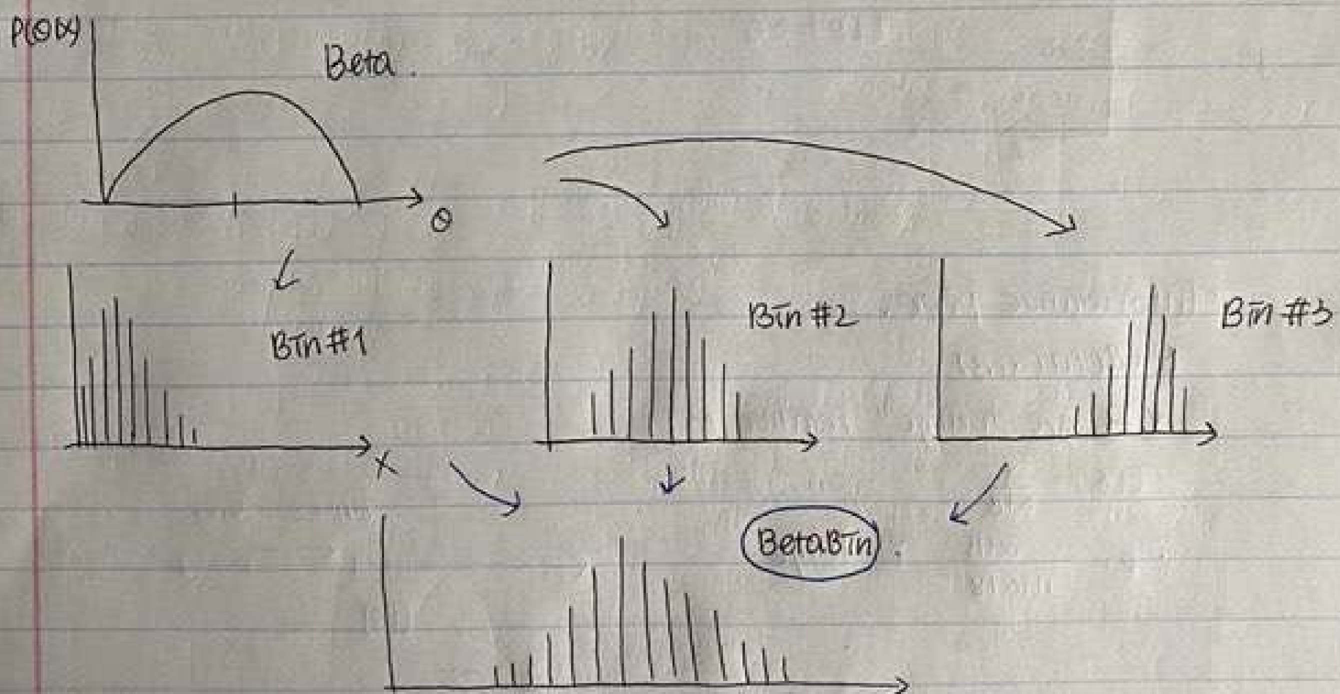
- In real life θ is unknown. Let's use Bayesian Inference.

  We obtain $P(\theta|x)$

· posterior predictive distribution

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

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



$\mathcal{F}$: Binomial fixed n

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

$P(\theta) = Beta(1,1) = U(0,1)$ Indifference prior, an example of an informative prior (Laplace).

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

• Hald Prior (1932)

$P(\theta) = Beta(0,0)$ but not a legal distribution ("improper")

Ha : "I don't care!"

$\Rightarrow P(\theta|X) = Beta(x, n-x)$ will be "proper" if $x \neq 0$ and $x \neq n$. $\Rightarrow \hat{\theta}_{MMSE} = \frac{x}{n} = \hat{\theta}_{MLE}$ (no shrinkage)

Objectivist: the data must speak for indifference.

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

$E[\theta] = 0.5$.

• Informative prior.

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

α, β are "Large" relative to n.

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

· $\rho = \frac{\alpha+\beta}{n+\alpha+\beta}$ is large