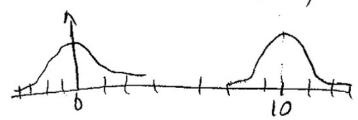
LECTURE-7
02-24-16

(1)

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



$$= \frac{1}{2} N(0,1) + \frac{1}{2} N(10,1^2)$$

$$= \sum_{m=1} f_m(x)$$

Compound dist:

$$p(x) = \sum_{\theta \in \Theta} p(x;\theta)\, p(\theta)$$

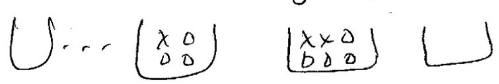
Mix Models

$$X \sim \begin{cases} Bin(n, 0.25) & \text{np } 0.1 \\ Bin(n, 0.4) & \text{np } 0.9 \end{cases}$$

$$p(x) = \sum_{\theta \in \{.25, .4\}} \binom{n}{x} \theta^x (1-\theta)^{n-x} p(\theta)$$

$$\binom{n}{x}\left(.25^x \cdot .75^{n-x} \cdot .1 + .4^x \cdot .6^{n-x} \cdot .9\right)$$

Compound, Marginal



(2)

$$\theta \sim Beta(\alpha, \beta) = U(0,1) \quad \alpha=1, \beta=1$$

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

$$p(x) = \int_{\Theta} \underbrace{p(x|\theta)p(\theta)}_{p(x,\theta)} d\theta = \int_0^1 \binom{n}{x} \theta^x (1-\theta)^{n-x} (1)\, d\theta$$

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

$$\sum_{x=0}^{n} \binom{n}{x} B(x+1, n-x+1) = 1$$

$$\{p.m.f\}$$

$$\int_0^1 \binom{n}{x} \theta^x (1-\theta)^{n-x} \underbrace{\frac{1}{B(\alpha,\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}}_{p.d.f \text{ of Beta}} d\theta$$

$$= \underbrace{\frac{\binom{n}{x}}{B(\alpha,\beta)} B(x+\alpha, n-x+\beta)}_{= X \sim Beta\,Bin(n,\alpha,\beta)}$$

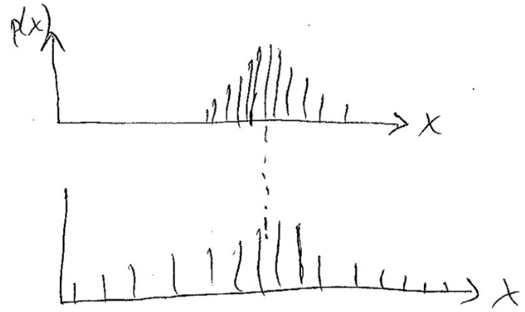
$$E[X] = n\theta$$

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

Beta $\quad E[X] = n \frac{\alpha}{\alpha+\beta}$

$$Var[X] =$$

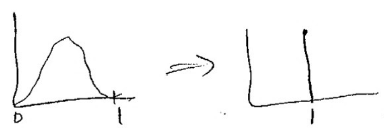
③



$\alpha = \beta$

$$\lim_{\alpha \to \infty} \text{Beta Bin}(n, \alpha, \alpha) = \text{Bin}(n, 1/2)$$

$\boxed{n \cdot \frac{1}{2} \cdot \frac{1}{2}}$



Laplace (Men are born more than woman)

6,115 families with 13 children

Dataset is good of first 12 children.

# male child	0	1	2	3	4	5	6	7	8	9	10	11	12	
freq	3	24	104	286	670	1033	1343	1112	829	478	181	45	7	Total / 6115·12 = .519
$\hat{X} \sim$ Beta Bin (12, 34, 32)	2	23	5	311	656	1036	1258	1182	854	462	178	44	5	
$\hat{X} \sim$ Bin (12, .519)	1	12	72	259	628	1085	1367	1266	854	410	132	26	2	

Beta (34, 32)

④

$X_1, \ldots, X_n \sim \text{Bin}(n, \theta)$

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

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

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

Given n data points;

Predict the next m data points i.e.

$$P(\underset{m}{\underline{x^*}} \mid \underset{n}{\underline{x}}) = \int_{\theta \in \Theta} p(x^* \mid \theta)\, p(\theta \mid x)\, d\theta$$

← over n' over average

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

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

$= \text{Beta Bin}(m, x+\alpha, n-x+\beta)$

If θ known, posterior Beta

$x^* \mid \cancel{X}_{\theta} \sim \text{Bin}(m, \theta)$

This wrap up the uncertainty.

(5)

$$P(\theta|x) = \frac{P(x|\theta)P(\theta)}{P(x)}$$ } function of all θ.

$\propto$ (proportional to p.d.f)

$P(x|\theta)P(\theta)$

{posterior}

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

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

$$\propto \theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}$$

(function of 'θ')

$$\frac{P(\theta=\theta_1|x)}{P(\theta=\theta_2|x)} = \frac{\theta_1^{x+\alpha-1}(1-\theta_1)^{n-x+\beta-1}}{\theta_2^{x+\alpha-1}(1-\theta_2)^{n-x+\beta-1}}$$

$$\left(\frac{\theta_1}{\theta_2}\right)\left(\frac{1-\theta_1}{1-\theta_2}\right)$$

$$\int_{x \in supp[x]} \underbrace{\theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}}_{g(x;\theta)} dx \neq 1$$

$= C < \infty$

$$\int \underbrace{\frac{g(x;\theta)}{c}}_{f(x;\theta)} dx = 1$$

Take p.d.f and

$$f(x;\theta) = \underbrace{h(\theta)}_{\text{normalizing constant}} \underbrace{g(x;\theta)}_{\text{kernel}}$$

(6)

$$P(x;\theta,n) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$$

$$= \frac{n!}{x!(n-x)!}\theta^x(1-\theta)^{n}(1-\theta)^{-x}$$

$$\propto \frac{1}{x!(n-x)!}\left(\frac{\theta}{1-\theta}\right)^x$$

normalizing constant $\left(\frac{n!}{(1-\theta)^n}\right)$

$$\theta \sim \text{Beta}(\alpha,\beta) = \frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}$$

$$\propto \underbrace{\boxed{\theta^a(1-\theta)^b}}_{\text{kernel}}$$

$a = \alpha - 1$

$b = \beta - 1$

$\theta | X \sim \text{Beta}(\alpha, \beta) = U(0,1)$

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

$$\hat{\theta}_{MMSE} = \frac{x+1}{n+2}$$ } Wilson Estimate, Law of succession

$\hat{\theta}_{MMSE} =$

(7)

$$\hat{\theta}_{MMSE} = \frac{\alpha + X}{\alpha+\beta+n} = \frac{\alpha}{\alpha+\beta+n} + \frac{X}{\alpha+\beta+n}$$

$$= \frac{\alpha}{\alpha+\beta+n}\left(\frac{\alpha+\beta}{\alpha+\beta}\right) + \frac{X}{\alpha+\beta+n}\left(\frac{n}{n}\right)$$

$$= \frac{\alpha+\beta}{\alpha+\beta+n}\left(\frac{\alpha}{\alpha+\beta}\right) + \left(\frac{n}{\alpha+\beta+n}\right)\underbrace{\left(\frac{X}{n}\right)}_{\hat{\theta}_{MLE}}$$

$$= \frac{\alpha+\beta}{\alpha+\beta+n}\underbrace{E[\theta]}_{\text{prior}} + \left(\frac{n}{\alpha+\beta+n}\right)\underbrace{\hat{\theta}_{MLE}}_{\text{data}}$$

$$= \rho E(\theta) + (1-\rho)\hat{\theta}_{MLE}$$

$$\rho := \frac{\alpha+\beta}{\alpha+\beta+n} = f(\alpha, \beta, n) = \frac{2}{2+n}$$

Shrinkage Estimator:

Shrink to your $E(\theta)$

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

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

proportional to $\boxed{\alpha}$ Pseudo counts

$$\theta^{(\alpha-1)+(X)}(1-\theta)^{(\beta-1)+(n-X)}$$

data success

#data fails

H.W 3

3i

$\alpha, \beta \to 0$ e.g. 1E-6

as close as to zero

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

$\hookrightarrow \alpha, \beta \in (0, \infty)$ 3i

if u include '0', not legal p.d.f.