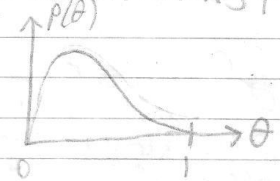
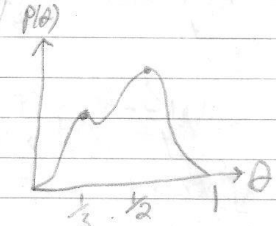
5/2/17

$X_1, \dots, X_n | \theta, \sigma^2 \overset{iid}{\sim} N(\theta, \sigma^2)$
$\theta \sim N(\mu_0, \tau^2)$
$\sigma^2 \sim InvGamma(\frac{n_0}{2}, \frac{n_0 \sigma_0^2}{2})$
$\Rightarrow P(\theta, \sigma^2 | X) = N(\theta_p, \sigma_p^2) K(\sigma^2 | X)$ θ_p is fn of σ^2
$\hat{\theta}_{MMSE} =$ derive from above

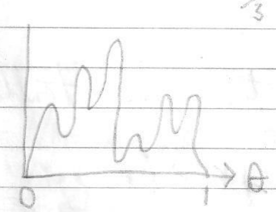
Recall
$X|\theta \sim Bin(n, \theta)$
$\theta \sim Beta(\alpha, \beta)$
$\theta | X \sim Beta(\alpha + x, \beta + n - x)$



What if you want:



Cannot be a Beta!



If you knew the function $P(\theta)$ then you can compare:

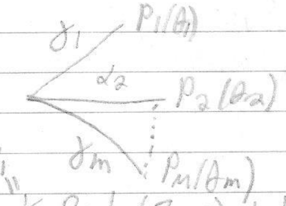
$P(\theta|X) \propto P(X|\theta)P(\theta) = k(\theta|X)$
and use a grid sample

$G = \langle \underset{0}{\theta_{min}}, \theta_{min} + \Delta\theta, \theta_{min} + 2\Delta\theta, \dots, \underset{1}{\theta_{max}} \rangle$

Do $\sum k \frac{1}{c}$, where c is sum of those ks
Can we still conjugacy?
Imagine $P(\theta)$ is a mixture (compound distr. of a

discrete # of beta components

$$P(\theta) = \sum_{m=1}^{M} \gamma_m \underset{\substack{\| \\ Beta(\alpha_m, \beta_m)}}{P_m(\theta)} \quad s.t. \sum \gamma_m = 1$$



$$\underset{\substack{\| \\ }}{\gamma_1} \qquad P(\theta) = \tfrac{1}{2} Beta(\underset{\substack{\| \\ \alpha_1}}{3}, \underset{\substack{\| \\ \beta_1}}{3}) + \tfrac{1}{2} Beta(\underset{\substack{\| \\ \alpha_2}}{2}, \underset{\substack{\| \\ \beta_2}}{4})$$

Recall

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

$$P(\theta) = \sum_{m=1}^{M} \gamma_m P_m(\theta)$$

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)} = \frac{P(X|\theta) \sum \gamma_m P_m(\theta)}{P(X)}$$

$$= \sum_{m=1}^{M} \gamma_m \frac{\overbrace{P(X|\theta)}^{\binom{n}{x}\theta^x(1-\theta)^{n-x}} P_m(\theta)}{P(X)} = \sum \underbrace{\frac{\gamma_m P_m(x)}{P(X)}}_{\gamma_m'} \cdot \underbrace{\frac{P(X|\theta) P_m(\theta)}{P_m(X)}}_{P_m(\theta|X)}$$

$$\frac{P_m(x)}{P_m(x)}$$

$$= \sum_{m=1}^{M} \gamma_m' \; \underset{\substack{\| \\ Beta(\alpha_m', \beta_m')}}{P_m(\theta|X)} \qquad \alpha_m' = \alpha_m + x \qquad \beta_m' = \beta_m + n - x$$

Recall

$$P(X) = \int_{\Theta} P(X|\theta) P(\theta) d\theta = \int_{\Theta} P(X|\theta) \sum \gamma_m P_m(\theta) d\theta$$

$$= \sum_{m=1}^{M} \gamma_m \underbrace{\int_{\Theta} P(X|\theta) P_m(\theta) d\theta}_{P_m(x) = BetaBinom(n, \alpha_m, \beta_m)}$$

cf $\gamma_m = \frac{1}{M} \; \forall m$

$$\gamma_m' = \frac{\gamma_m p_m(X)}{p(X)} = \frac{\gamma_m p_m(X)}{\sum \gamma_m p_m(X)} = \frac{p_m(X)}{\sum p_m(X)}$$

if $\gamma_m = \frac{1}{M} \ \forall m$

equal components allow implication of formulas

$\gamma_1 = \gamma_2 = \frac{1}{2}$; $\alpha_1 = 3$, $\beta_1 = 3$; $\alpha_2 = 2$, $\beta_2 = 4$, $n = 10$, $x = 5$

$$p(\theta|X) = \sum \underset{\substack{\shortparallel \\ \frac{p_m(x)}{\sum p_m(X)}}}{\gamma_m'} p_m(\theta|X)$$

$$= \frac{1}{\sum p_m(X)} \sum p_m(X) \underbrace{p_m(\theta|X)}_{\text{Beta distribution}}$$

$$= \frac{1}{p_1(5)+p_2(5)} \left(p_1(5)\, p_1(\theta|X=5) + p_2(5)\, p_2(\theta|X=5)\right)$$

$p_1(5) =$

$p_1(X) = \text{Betabinom}(\underset{10}{n}, \underset{3}{\alpha_1}, \underset{3}{\beta_1})$ pmf

evaluate pmf at $x = 5$

$p_1(5) = \text{dbetabinom}(5, 10, 3, 3) = .147$ "dbb"

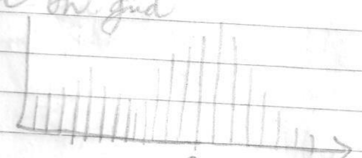
$p_2(5) = \text{dbb}(5, 10, 2, 4) = .112$

prior predictive distribution

$$= \frac{1}{\text{dbb}(5,10,3,3) + \text{dbb}(5,10,2,4)} \left(\text{dbb}(5,10,3,3)\,\text{dbeta}(\theta, 8, 8) + \text{dbb}(5,10,2,4)\,\text{dbeta}(\theta, 7, 9) \right)$$

$= .57\,\text{dbeta}($
$+ .43\,\text{dbeta}($

Plot on grid



Sample from $p(\theta|X)$

① Sample $\theta_{s,1}$ from Beta(8, 8), use rbeta(8, 8) rbeta samples from the distribution

② Sample $\theta_{0,2}$ from Beta(7,9) use rbeta(7,9)
③ Return $\theta_0 = \gamma_1' \theta_{0,1} + \gamma_2' \theta_{0,2}$
④ Repeat 1-3 many times

<u>Point Est.</u>

$$\hat{\theta}_{MMSE} = E[\theta|X] = \int_{\Theta} \theta \sum \gamma_m' P_m(\theta|X)\, d\theta$$

$$= \sum \gamma_m' \int_{\Theta} \theta P_m(\theta|X)\, d\theta$$

$$= \sum \gamma_m' E_m[\theta|X] = \sum_{m=1}^{M} \gamma_m' \frac{\alpha_m'}{\alpha_m' + \beta_m'}$$

$$= .57 \cdot \frac{8}{16} + .43 \cdot \frac{7}{16}$$

Not covered but
need CDF of beta mixtures - none available

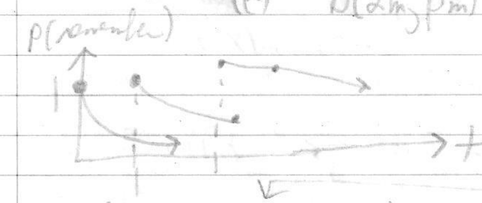
(not on exam) $\hat{\theta}_{MAE} = \cdots$ sample median

$P(\theta|X) = \sum \gamma_m' \int_{\Theta} \theta P_m(\theta|X)\, d\theta$, take middle

$\hat{\theta}_{MAP} = \text{argmax}\{P(\theta|X)\} = \text{argmax}\{k(\theta|X)\}$

$$P(\theta|X) \propto \sum \gamma_m P_m(X) P_m(\theta|X) = k(\theta|X)$$

On final
$$= \sum \gamma_m \left(\binom{n}{x} \frac{B(x+\alpha_m, n-x+\beta_m)}{B(\alpha_m, \beta_m)} \right) \left(\frac{1}{B(x+\alpha_m, n-x+\beta_m)} \theta^{x+\alpha_m-1} (1-\theta)^{n-x+\beta_m-1} \right)$$



Spacing effect

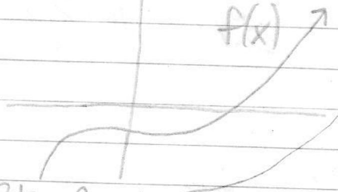
$$0 \overset{set}{=} \frac{d}{d\theta}\left[\quad \right] = \binom{n}{x} \sum \frac{d}{d\theta} \left[\frac{\overbrace{\theta^{x+\alpha_m-1}}^{u} \overbrace{(1-\theta)^{n-x+\beta_m-1}}^{v}}{B(\alpha_m, \beta_m)} \right]$$

$$= \sum \frac{1}{B(\alpha_m, \beta_m)} \Big((1-\theta)^{n-x+\beta_m-1} (x+\alpha_m-1)\, \theta^{x+\alpha_m-2}$$

$$-\theta^{x+\alpha m-1}(n-x+\beta m-1)(1-\theta)^{n-x+\beta m-2}\Big) = 0$$

can't be solved

Assume $f(x)$ is continuous and diff. and has one zero on $\mathcal{X}$. Want x^* s.t. $f(x^*) = 0$.



Step 1: Guess $y_0 = x^*$

Step 2: Draw tangent line

Step 3: Set $x_1 =$ x-intercept of tangent line

Step 4: Repeat by setting $x_0 = x_m$ until $|x_{t+1} - x_t| < \varepsilon$

ε is your accuracy/tolerance level

Step 2 <

$y - b = m(x - a)$

$y - f(x_0) = f'(x_0)(x - x_0)$

Step 3

Solve for x-intercept (x_1):

$$-f(x_0) = f'(x_0)(x_1 - x_0)$$

$$\Rightarrow x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

$$x_{t+1} = x_t - \frac{f(x_t)}{f'(x_t)}$$

Prior is a known mixture.

What if likelihood model is a mixture?

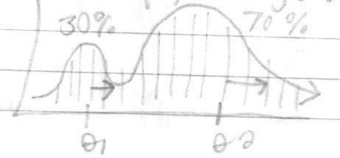
$$X_1, \ldots, X_n \mid \theta \underset{\gamma_1, \ldots, \gamma_m}{\overset{iid}{\sim}} \sum_{m=1}^{M} \gamma_m p_m(X \mid \theta)$$

E.g. $X_1, \ldots, X_n \mid \theta_1, \sigma_1^2, \theta_m, \sigma_1^2, \rho$

$$\overset{iid}{\sim} \underset{\gamma_1}{\rho} N(\theta_1, \sigma_1^2) + \underset{\gamma_2}{(1-\rho)} N(\theta_2, \sigma_2^2)$$

$\rho \; N(\theta_1, \sigma_1^2)$

$1-\rho \; N(\theta_2, \sigma_2^2)$



Goal: get posterior or functions of posterior:

$$P(\theta_1, \sigma_1^2, \theta_2, \sigma_2^2, \rho \mid X) \propto \left(\prod_{i=1}^{n} P(X_i \mid \theta_1, \sigma_1^2, \theta_2, \sigma_2^2, \rho)\right) P(\theta_1, \sigma_1^2, \theta_2, \sigma^2, \rho)$$

5/4/17

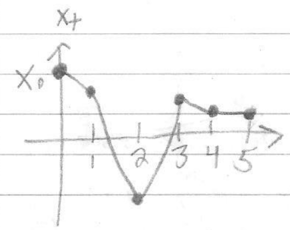
N-R Method

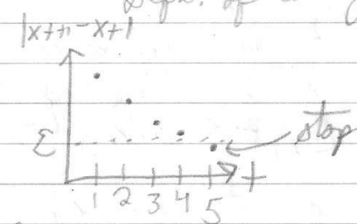
$f(x) = 0$, solve for x. Given ε.

① Guess solution in x_0.

② Calculate $x_1 = x_0 - \dfrac{f(x_0)}{f'(x_0)}$ } iterative steps in an iterative algorithm

③ Repeat step 2 until $\underbrace{|x_{t+1} - x_t| < \varepsilon}_{\text{Defn. of convergence}}$





(Likelihood is a mixture) Consider the model

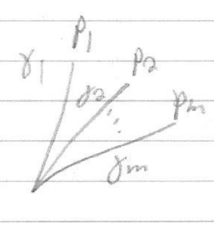
$$X_1, \ldots, X_n \mid \vec{\theta}_1, \ldots, \vec{\theta}_M, \gamma_1, \ldots, \gamma_M \overset{iid}{\sim} \sum_{m=1}^{M} \gamma_m P_m(\vec{\theta}_m)$$

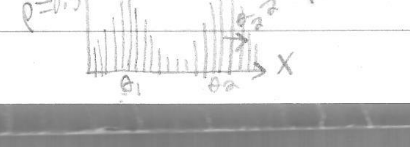
s.t. $\gamma_1 + \gamma_2 + \ldots + \gamma_m = 1$

e.g.

$$X_1, \ldots, X_n \mid \underbrace{\theta_1, \sigma_1^2}_{\vec{\theta}_1}, \underbrace{\theta_2, \sigma_2^2}_{\vec{\theta}_2}, \underbrace{\rho}_{\gamma_1, \gamma_2}$$

$$\overset{iid}{\sim} \underset{\gamma_1}{\rho}\, \overbrace{N(\theta_1, \sigma_1^2)}^{P_1(\vec{\theta}_1)} + \underset{\gamma_2}{(1-\rho)}\, \underbrace{N(\theta_2, \sigma_2^2)}_{P_2(\vec{\theta}_2)}$$





Bayesian Inference: Here, infering to get parameter values