10/1 Matthew Maslow

Hw 2

1) Hoff's book 3.3

$y_A = (12, 9, 12, 14, 13, 13, 15, 8, 15, 6)$
$y_B = (11, 11, 10, 9, 9, 8, 7, 10, 6, 8, 8, 9, 7)$

a) Find Posterior dist, means, variances, and 95% quantile-bas CI for θ_A and θ_B...

We know that...

$y_A \sim Poisson(\theta_A)$
$y_B \sim Poisson(\theta_B)$

* Independent...
$P(\theta_A \cdot \theta_B) = P(\theta_A) \cdot P(\theta_B)$

$\frac{\alpha}{\lambda} = 12$ *
↳ $\frac{\alpha}{10} = 12 \Rightarrow \alpha = 120$

given we have priors...

$\theta_A \sim Gamma(120, 10)$
$\theta_B \sim Gamma(12, 1)$

Now, establishing Likelihood func for poisson...

$$P(y|\theta) \propto \theta^{\Sigma y} \cdot e^{-n \cdot \theta}$$, where, n=

Therefore,

posterior for θ_A... using $n_A = 10$ and $\Sigma y_A = 117$

now, combining prior w/ post...

$120+117$ ↗ $10+10$ ↗

$$\boxed{\theta_A \sim Gamma(237, 20)}$$

Mean θ_A: $E(\theta_A|Y_A) = 237/20 \Rightarrow \boxed{E(\theta_A|Y_A) = 11.85}$

Variance θ_A: $Var(\theta_A|Y_A) = 237/20^2 \Rightarrow \boxed{Var(\theta_A|Y_A) = 0.5925}$

95% CI for θ_A:
↳ $\boxed{(10.39, 13.41)}$

L_{θ_A} = qgamma(0.025, 237, 20) = 10.38924
U_{θ_A} = qgamma(0.975, 237, 20) = 13.40545

1) 3.3) a cont'd)

Now, for θ_B ....

Posterior θ_B... using $n_0 = 13$, $\Sigma y_B = 118$

$$\boxed{\theta_B \sim \text{Gamma}(\underbrace{130}_{12+118}, \underbrace{14}_{1+13})}$$

Mean θ_B...

$$\frac{130}{14} \Rightarrow \boxed{E(\theta_B \mid Y_B) = 9.29}$$

Variance θ_B... $\frac{130}{14^2} \Rightarrow \boxed{Var(\theta_B \mid Y_B) = 0.663}$

95% CI for θ_B:

L_{θ_B} = qgamma(0.025, 130, 14) = 7.75819

U_{θ_B} = qgamma(0.975, 130, 14) = 10.94047

$$\boxed{(7.76, 10.95)}$$

---

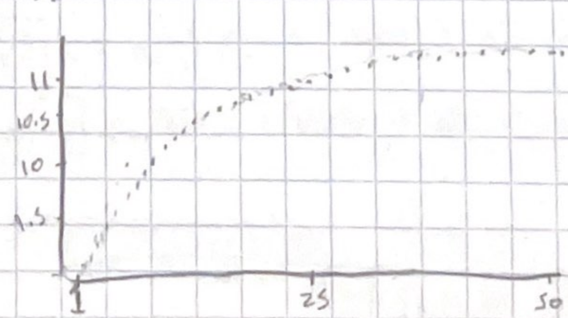
b) Compute/plot posterior expectation of θ_B under the prior dist... $\theta_B \sim \text{Gamma}(12 \times n_0, n_0)$

for $n_0 \in \{1, 2, ..., 50\}$

therefore we would believe that when values of "n_0" increases 1 to 50

↳ smaller n_0's have less influence while bigger ones have greater influence

Plot...



Larger n_0's have higher θ's but 11.5 ~~orange~~ by showing not too much influence

1) 3.3)

c.) Does it make sense to have $P(\theta_A, \theta_B) = P(\theta_A) \times P(\theta_B)$

↳ Since Both population B and A are related in a way it does make sense combining them in some way, as both share similar pdfs (prior + post where they use poisson and gamma.

3.4) mix of Beta priors...

prob θ of teen recidivism w/

$n = 43$ released

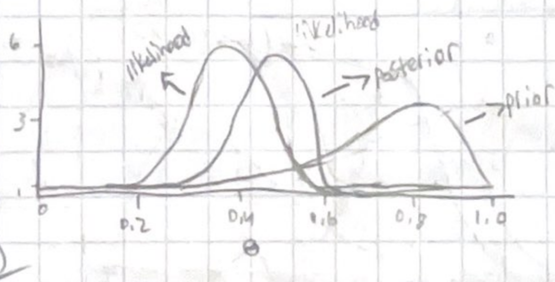
$y = 15$ re-offenders] → w/n 36 months

a) using beta(2, 8) prior for θ, plot $p(\theta)$, $P(y|\theta)$, and $p(\theta|y)$ as functions of θ. Then find posterior mean, mode, and SD of θ. And 95% CI for θ

w/ a $p(\theta) = beta(2,8)$ then

$p(y|\theta) = binomial(n, \theta) \sim y$

$p(\theta|y) = Beta(17, 36)$



beta(2,8)

Mean of θ...

$$E(\theta) = \frac{17}{17+36} \Rightarrow \boxed{E(\theta) = 0.321}$$

Mode of θ...

$$Mode = \frac{17-1}{17+36-2} \Rightarrow \boxed{Mode = 0.314}$$

SD of θ...

$$\sqrt{\frac{17 \cdot 36}{(17+36)^2 \cdot (17+36+1)}} \Rightarrow \boxed{SD = 0.064}$$

95% CI...

$L_b = qbeta(0.025, 17, 36) = 0.2032978$

$U_b = qbeta(0.975, 17, 36) = 0.4510240$] → $\boxed{(0.203, 0.451)}$

2) 3.4)

b) Beta(8,2) using $\theta | y \sim Beta(\overset{15+8}{23}, \overset{43-15+2}{30})$

Mean of θ

$$E[\theta | y] = \alpha / \alpha+\beta$$
$$= 23/23+30 \Rightarrow \boxed{E(\theta|y) = 0.434}$$

Mode of θ:

$$mode = \frac{\alpha - 1}{\alpha + \beta - 2} = \frac{23-1}{23+30-2} = \frac{22}{51} \Rightarrow \boxed{mode = 0.431}$$

Standard Dev of θ:

$$SD = \sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}} = \sqrt{\frac{23 \times 30}{(53)^2(54)}} \Rightarrow \boxed{SD_\theta = 0.0674}$$

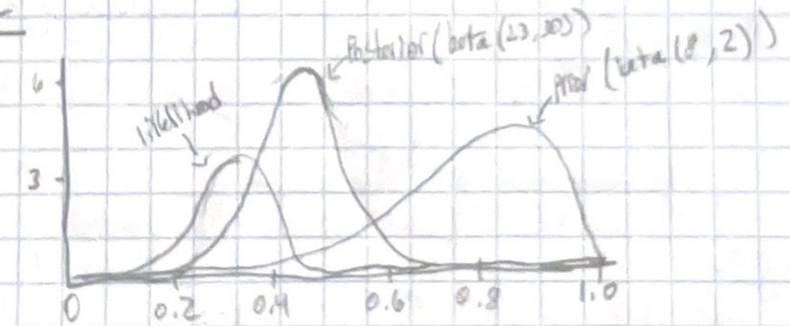
95% CI of θ:

L_θ = qbeta(0.025, 23, 30) = 0.3046956

U_θ = qbeta(0.975, 23, 30) = 0.5679528

$\boxed{(0.305, 0.568)}$

Plot



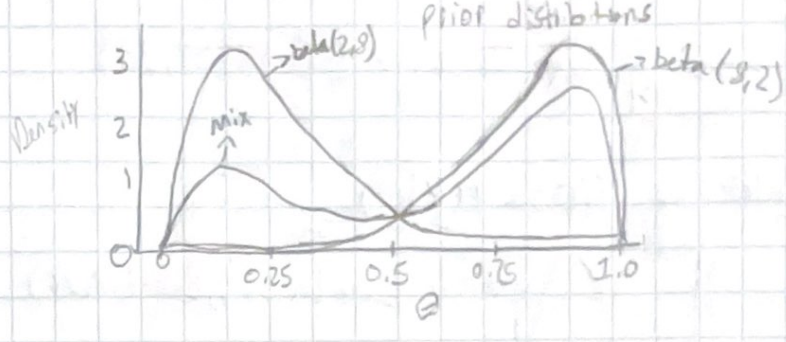
Part C →

2) 3.4) c.)

Consider...

$$p(\theta) = \frac{1}{4}\frac{\Gamma(10)}{\Gamma(2)\Gamma(8)}\left[3\theta(1-\theta)^7 + \theta^7(1-\theta)\right]$$

↳ mix of beta (2,8) and beta (8,2) prior distributions



Looking at the plot of the 3 priors, we can conclude that w/ the 75% weight on beta(8,2) and 25% on beta(2,8) it represents the prior for teen recidivism is more likely to skew w/ higher values, along w/ some consideration for lower values. Overall is emphasizing a strong belief in high recidivism

d) For prior in c)...

i) write out $p(\theta) \times p(y|\theta)$ and simplify

$$p(\theta) = \frac{1}{4}\text{Beta}(2,8) + \frac{3}{4}\text{Beta}(8,2)$$

using $y=15$ and $n=43$

$$p(y|\theta) \propto \theta^y(1-\theta)^{n-y} = \theta^{15}(1-\theta)^{28}$$

now...

$$p(\theta) * p(y|\theta) \propto \left(\frac{1}{4}\theta(1-\theta)^7 + \frac{3}{4}\theta^7(1-\theta)^1\right) * \theta^{15}(1-\theta)^{28}$$

↳ simplifies to...

$$\boxed{p(\theta|y) \propto \frac{1}{4}\theta^{16}(1-\theta)^{35} + \frac{3}{4}\theta^{22}(1-\theta)^{29}}$$

d)

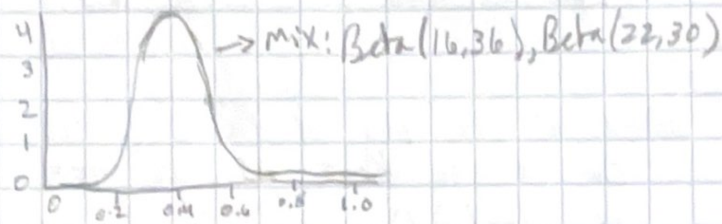
ii) ID two Posterior dist

1) Beta (16, 36) → beta(2,8) w/ likelihood

2) Beta (22, 30 → beta(8,2) w/ likelihood

↳ $\frac{1}{4}$ Beta(16,36) + $\frac{3}{4}$ Beta(22,30)

iii) plot of $P(\theta) \times P(y|\theta)$



$$\text{Mode}_{B(16,36)} = 0.30$$

$$\text{Mode}_{B(22,30)} = 0.41$$

The mix ~~prior~~ posterior Mode compares more w/ beta(8,2) ~~prior~~ posterior Mode as Mix prior is weighted more heavily w/ beta(8,2) prior.

e)

i) General weights formula for mix prior / Posterior

$$P(\theta|y) \propto P(y|\theta) \times P(\theta)$$

where $i = 1, 2$ $\quad W_i = \dfrac{P(y|\text{Prior}_i) \times w_i}{P(y|\text{Prior}_1) \times w_1 + P(y|\text{Prior}_2) \times w_2}$

↳ where w_1 and w_2 represent weights of priors, and

w_1 and prior_1 → goes w/ Beta(2,8)

w_2 and prior_2 → goes w/ Beta(8,2)

3) Jeffery's Prior

a) If $\theta \sim Beta(\alpha, B)$, show that $\lambda = \log(\theta/(1-\theta)) = \rightarrow$ logit(θ)

$$p(\theta) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1}$$

w/ $\lambda = \log\left(\frac{\theta}{1-\theta}\right)$

$$\theta = \frac{e^\lambda}{1+e^\lambda} \rightarrow \frac{d\theta}{d\lambda} = \frac{e^\lambda}{(1+e^\lambda)^2}$$

combine...

$$p(\lambda) \propto e^{(\alpha-1)\lambda}(1+e^\lambda)^{-\alpha-\beta+2} * \frac{e^\lambda}{(1+e^\lambda)^2}$$

$$\boxed{\propto \frac{e^{\alpha\lambda}}{(1+e^\lambda)^{\alpha+\beta}}}$$

b) Jeffery's prior w/ λ: is Beta-logit$(\frac{1}{2}, \frac{1}{2})$

Log-likelihood of $x|\lambda \sim Binom(n, logit^{-1}(\lambda))$ is...

$$\log P(X|\lambda) = n * \lambda * X - n\log(1+e^\lambda)$$

Jefferys prior should be proportionally square root to fourer criteria...

$$I(\lambda) \propto \frac{e^\lambda}{(1+e^\lambda)^2}$$

So,

$$p(\lambda) \propto \sqrt{I(\lambda)} \propto \frac{e^{\lambda/2}}{1+e^\lambda}$$

which is exact form of beta-logit w/ parameters $\alpha = \frac{1}{2}$, $B = \frac{1}{2}$

$\therefore \theta \sim Beta(\frac{1}{2}, \frac{1}{2})$

since $\lambda = \log\left(\frac{\theta}{1-\theta}\right)$, transformed $\theta = logit^{-1}(\lambda)$, meaning if λ has a beta-logit dist w/ $\alpha = \frac{1}{2}$ and $B = \frac{1}{2}$ then dist will be $Beta(\frac{1}{2}, \frac{1}{2})$