STAT 840 Chapter 1 Exercises

Vsevolod Ladtchenko 20895137

TODO

[1] 1 [1] 2 [1] 3 [1] 4 [1] 5 [1] 6 [1] 7 [1] 8 [1] 9 [1] 10 [1] 11 [1] 12 [1] 13 [1] 14

TRICKS TO REMEMBER

- mode = highest point of density. get it by taking derivative
- gamma/beta mean = rewrite as Gam(a+1,b) using Gam(n) = (n-1)!
- chap 1 variance decomposition: use only decomposition for xbar

PROBABLY CORRECT BUT I SHOULD ASK ANYWAYS

18

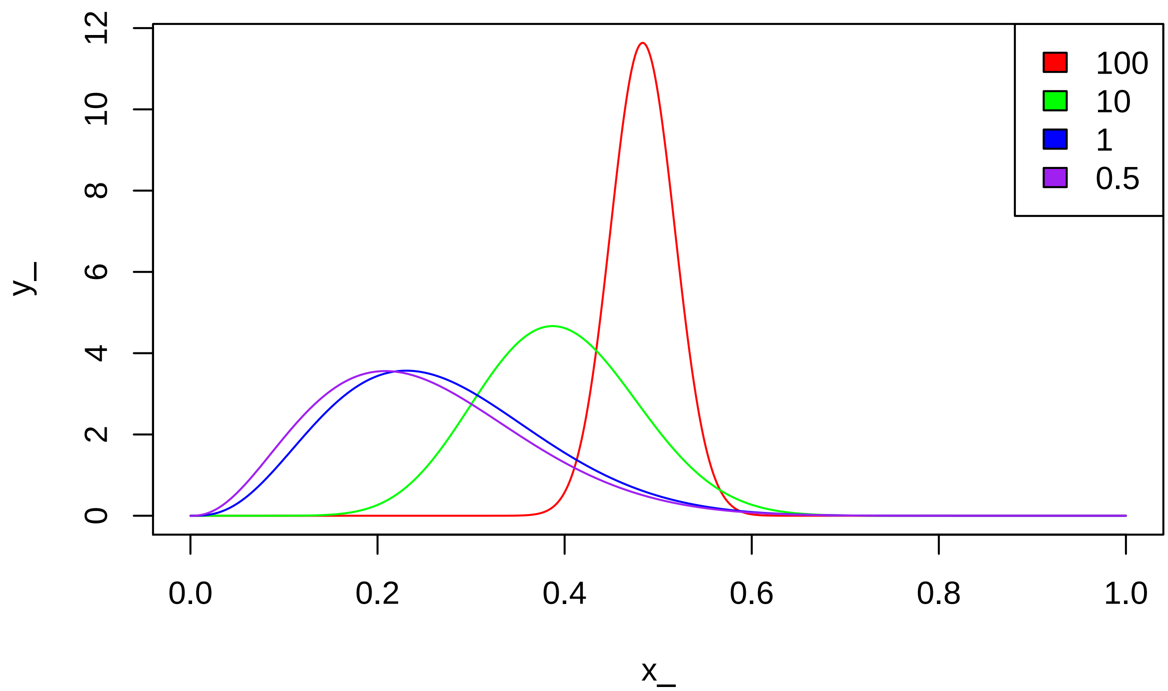
Problem 1.14

Problem 1.15

$$
\begin{aligned}
f(x \mid \theta) &\propto \theta^x (1-\theta)^{1-x} \\
f(X \mid \theta) &\propto \prod_{i=1}^{n} \theta^{x_i} (1-\theta)^{1-x_i} \\
&\propto \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i} \\
f(\theta) &\propto \theta^{\alpha-1} (1-\theta)^{\beta-1} \\
f(\theta \mid X) &= f(X \mid \theta) f(\theta) \\
&\propto \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i} \theta^{\alpha-1} (1-\theta)^{\beta-1} \\
&\propto \theta^{\alpha-1+\sum_{i=1}^n x_i} (1-\theta)^{\beta-1+n-\sum_{i=1}^n x_i} \\
&\sim \text{Beta}(\alpha_{prior} + \sum_{i=1}^{n} x_i, \beta_{prior} + n - \sum_{i=1}^{n} x_i)
\end{aligned}
$$

```
problem_1_15 = function()
{
  x = c(0, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0)

  bay_est = function (a,b) a/(a+b)

  map_est = function (a,b) (a-1)/(a+b-2)
```

```r
  params = c(100, 10, 1, 0.5)
  colors = c('red','green','blue','purple')

  for (i in 1:4)
  {
    a_prior = params[i]
    b_prior = params[i]

    a_post = a_prior + sum(x)
    b_post = b_prior + length(x) - sum(x)

    x_ = seq(0,1,1/1000)
    y_ = dbeta(x_, a_post, b_post)

    if (i==1)
    {
      plot(x=x_,y=y_,type='l', col=colors[i])
      legend(x = "topright", legend=params, fill = colors)
    }
    else lines(x=x_,y=y_, col=colors[i])

    print(paste("for prior (", a_prior, ",",b_prior, ")"))
    print(round(c(Bay=bay_est(a_post, b_post), MAP=map_est(a_post, b_post)),4))
  }
}

problem_1_15()
```




```
## [1] "for prior ( 100 , 100 )"
##    Bay    MAP
## 0.4836 0.4834
## [1] "for prior ( 10 , 10 )"
##    Bay    MAP
## 0.3939 0.3871
## [1] "for prior ( 1 , 1 )"
##    Bay    MAP
## 0.2667 0.2308
## [1] "for prior ( 0.5 , 0.5 )"
##    Bay    MAP
## 0.2500 0.2083
```

Problem 1.16

$$\begin{aligned}
X_i &\sim Pois(\theta) \\
f(x_i \mid \theta) &= \frac{\theta^{x_i} e^{-\theta}}{x_i!} \\
f(X \mid \theta) &= \prod_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i!} \\
f(\theta) &= \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} \\
f(\theta \mid X) &= f(X \mid \theta) f(\theta) \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} \prod_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i!} \\
&\propto \theta^{\alpha-1} e^{-\beta\theta} \prod_{i=1}^{n} \theta^{x_i} e^{-\theta} \\
&\propto \theta^{\alpha-1} e^{-\beta\theta - n\theta} \theta^{\sum_{i=1}^{n} x_i} \\
&\propto \theta^{\alpha-1+\sum_{i=1}^{n} x_i} e^{-\beta\theta - n\theta} \\
&\propto \theta^{(\alpha+\sum_{i=1}^{n} x_i)-1} e^{-\theta(\beta+n)} \\
&\propto \theta^{(\alpha+\sum_{i=1}^{n} x_i)-1} e^{-(\beta+n)\theta} \\
&\sim \text{Gamma}(\alpha_{post} = \alpha_{prior} + \sum_{i=1}^{n} x_i, \beta_{post} = \beta_{prior} + n)
\end{aligned}$$

Posterior mode:

$$\begin{aligned}
f(\theta) &= \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} \\
f(\theta) &\propto \theta^{\alpha-1} e^{-\beta\theta} \\
f'(\theta) &\propto \theta^{\alpha-1}(-\beta) e^{-\beta\theta} + (\alpha-1)\theta^{\alpha-2} e^{-\beta\theta} \\
0 &= \theta^{\alpha-1}(-\beta) + (\alpha-1)\theta^{\alpha-2} \\
0 &= (-\beta\theta + \alpha - 1)\theta^{\alpha-2} \\
&\quad \text{assume } \theta \neq 0 \text{ and later check value at } 0 \\
0 &= -\beta\theta + \alpha - 1 \\
\beta\theta &= \alpha - 1 \\
\theta &= \frac{\alpha-1}{\beta} \\
&\quad \text{we know support is positive} \\
&\quad \text{this is positive for } \alpha > 1 \\
&\quad \alpha < 1 \text{ implies max at } 0
\end{aligned}$$

Posterior mean:

$$
\begin{aligned}
f(\theta) &= \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta} \\
\mathbb{E}[\theta] &= \int_0^\infty \theta\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}d\theta \\
&= \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{(\alpha+1)-1}e^{-\beta\theta}d\theta \\
&= \int_0^\infty \frac{1}{\beta}\frac{\beta^{\alpha+1}}{\Gamma(\alpha)}\theta^{(\alpha+1)-1}e^{-\beta\theta}d\theta \\
&\Gamma(\alpha) = (\alpha-1)! \\
&\Gamma(\alpha+1) = \alpha! \\
&\Gamma(\alpha+1) = \alpha\Gamma(\alpha) \\
&= \int_0^\infty \frac{\alpha}{\beta}\frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)}\theta^{(\alpha+1)-1}e^{-\beta\theta}d\theta \\
&= \frac{\alpha}{\beta}\int_0^\infty \frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)}\theta^{(\alpha+1)-1}e^{-\beta\theta}d\theta \\
&= \frac{\alpha}{\beta}
\end{aligned}
$$

Thus the posterior mean and mode of θ are:

$$
\begin{aligned}
\theta_{mode} &= \frac{\alpha_{prior} + \sum_{i=1}^n x_i - 1}{\beta_{prior} + n} \\
\theta_{mean} &= \frac{\alpha_{prior} + \sum_{i=1}^n x_i}{\beta_{prior} + n}
\end{aligned}
$$

Problem 1.17

$$
\begin{aligned}
\theta &\sim \text{Beta}(\alpha = 5, \beta = 10) \\
f(\theta) &\propto \theta^{\alpha-1}(1-\theta)^{\beta-1} \\
&\propto \theta^4(1-\theta)^9 \\
f(x \mid \theta) &= \binom{n}{k}\theta^k(1-\theta)^{n-k} \\
&\propto \theta^1(1-\theta)^{19} \\
f(\theta \mid x) &= f(x \mid \theta)f(\theta) \\
&\propto \theta^1(1-\theta)^{19}\theta^4(1-\theta)^9 \\
&\propto \theta^5(1-\theta)^{28} \\
&\sim \text{Beta}(6, 29)
\end{aligned}
$$

MAP estimate (mode):

$$
\begin{aligned}
f(\theta \mid x) &\propto \theta^{\alpha-1}(1-\theta)^{\beta-1} \\
f'(\theta \mid x) &\propto -(\beta-1)\theta^{\alpha-1}(1-\theta)^{\beta-2} + (\alpha-1)\theta^{\alpha-2}(1-\theta)^{\beta-1} \\
0 &= \Bigg(-(\beta-1)\theta + (\alpha-1)(1-\theta)\Bigg)\theta^{\alpha-2}(1-\theta)^{\beta-2} \\
&\quad \text{assume } \theta, (1-\theta) \neq 0 \text{ later check value at } 0 \\
0 &= -(\beta-1)\theta + (\alpha-1)(1-\theta) \\
0 &= (1-\beta)\theta + \alpha(1-\theta) - (1-\theta) \\
0 &= \theta - \beta\theta + \alpha - \alpha\theta - 1 + \theta \\
1-\alpha &= 2\theta - \alpha\theta - \beta\theta \\
1-\alpha &= (2-\alpha-\beta)\theta \\
\theta &= \frac{1-\alpha}{2-\alpha-\beta}\frac{-1}{-1} \\
\theta &= \frac{\alpha-1}{\alpha+\beta-2}
\end{aligned}
$$

Bayes estimate (mean):

$$
\begin{aligned}
f(\theta \mid x) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1} \\
\mathbb{E}[\theta] &= \int_0^1 \theta\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}d\theta \\
&= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{(\alpha+1)-1}(1-\theta)^{\beta-1}d\theta \\
&\quad \Gamma(\alpha+1) = \alpha\Gamma(\alpha) \\
&\quad \Gamma(\alpha+\beta+1) = (\alpha+\beta)\Gamma(\alpha+\beta) \\
&= \int_0^1 \frac{\alpha}{\alpha+\beta}\frac{\Gamma(\alpha+1+\beta)}{\Gamma(\alpha+1)\Gamma(\beta)}\theta^{(\alpha+1)-1}(1-\theta)^{\beta-1}d\theta \\
&= \frac{\alpha}{\alpha+\beta}\int_0^1 \frac{\Gamma(\alpha+1+\beta)}{\Gamma(\alpha+1)\Gamma(\beta)}\theta^{(\alpha+1)-1}(1-\theta)^{\beta-1}d\theta \\
&= \frac{\alpha}{\alpha+\beta}
\end{aligned}
$$

So in our problem:

$$
\begin{aligned}
\theta_{mode} &= \frac{6-1}{6+29-2} \\
\theta_{mean} &= \frac{6}{6+29}
\end{aligned}
$$

Problem 1.18

$$
\begin{aligned}
P(\text{positive}) &= P(\text{positive} \cap \text{sensitive}) + P(\text{positive} \cap \text{not-sensitive}) \\
&\quad \text{law of total probability} \\
&= P(\text{positive} \mid \text{sensitive})P(\text{sensitive}) + P(\text{positive} \mid \text{not-sensitive})P(\text{not-sensitive}) \\
&\quad \text{definition of conditional probability} \\
&= P(\text{positive} \mid \text{sensitive})0.5 + P(\text{positive} \mid \text{not-sensitive})0.5 \\
&= p0.5 + 0.25 \\
0.5p &= P(\text{positive}) - 0.25 \\
p &= 2P(\text{positive}) - 0.5
\end{aligned}
$$

Now we need to get the MLE of the probability of a positive response. Since this involves counting a group of binary outcomes, it is a binomial distribution.

$$
\begin{aligned}
f(k) &= \binom{n}{k}\theta^k(1-\theta)^{n-k} \\
\frac{d}{d\theta}f(k) &= -(n-k)\binom{n}{k}\theta^k(1-\theta)^{n-k-1} + k\binom{n}{k}\theta^{k-1}(1-\theta)^{n-k} \\
0 &= \left(-(n-k)\theta + k(1-\theta)\right)\theta^{k-1}(1-\theta)^{n-k-1} \\
&\quad \text{assume } \theta, (1-\theta) \neq 0 \text{ later check value at } 0 \\
0 &= -(n-k)\theta + k(1-\theta) \\
0 &= (k-n)\theta + k - k\theta \\
0 &= k\theta - n\theta + k - k\theta \\
0 &= -n\theta + k \\
n\theta &= k \\
\theta &= k/n
\end{aligned}
$$

Because of the invariance property of the MLE, we know that any function of the MLE is the MLE of that function. So in our case:

$$
\begin{aligned}
P(\text{positive})_{MLE} &= X/n \\
p_{MLE} &= 2P(\text{positive})_{MLE} - 0.5 \\
p_{MLE} &= 2X/n - 0.5
\end{aligned}
$$

Problem 1.19

i)

$$\begin{aligned}
\sum_{i=1}^{n}(x_i-\bar{x}_n)^2 &= \sum_{i=1}^{n} x_i^2 - 2\bar{x}_n\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\bar{x}_n^2 \\
&= \sum_{i=1}^{n} x_i^2 - 2\bar{x}_n n\bar{x}_n + n\bar{x}_n^2 \\
&\qquad \text{since } n\bar{x}_n = \sum_{i=1}^{n} x_i \\
&= \sum_{i=1}^{n} x_i^2 - n\bar{x}_n^2
\end{aligned}$$

ii)

$$\begin{aligned}
\bar{x}_{n+1} &= \frac{1}{n+1}\sum_{i=1}^{n+1} x_i \\
&= \frac{1}{n+1}x_{n+1} + \frac{1}{n+1}\sum_{i=1}^{n} x_i \\
&= \frac{1}{n+1}x_{n+1} + \frac{n}{n+1}\bar{x}_n \\
&= \bar{x}_n + \frac{1}{n+1}(x_{n+1}-\bar{x}_n)
\end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^{n+1}(x_i-\bar{x}_{n+1})^2 &= \sum_{i=1}^{n+1}(x_i-\bar{x}_n-\frac{1}{n+1}(x_{n+1}-\bar{x}_n))^2 \\
&\quad \text{from previous part, plug in } \bar{x}_{n+1} \\
&= \sum_{i=1}^{n+1}(x_i-\bar{x}_n)^2 - 2\frac{1}{n+1}(x_{n+1}-\bar{x}_n)\sum_{i=1}^{n+1}(x_i-\bar{x}_n) + \sum_{i=1}^{n+1}\frac{1}{(n+1)^2}(x_{n+1}-\bar{x}_n)^2 \\
&\quad \text{note: } \sum_{i=1}^{n+1}(x_i-\bar{x}_n) = (n+1)(\bar{x}_{n+1}-\bar{x}_n) \\
&\quad \text{from previous part, } \bar{x}_{n+1}-\bar{x}_n = \frac{1}{n+1}(x_{n+1}-\bar{x}_n) \\
&= \sum_{i=1}^{n+1}(x_i-\bar{x}_n)^2 - \frac{2}{n+1}(x_{n+1}-\bar{x}_n)^2 + \frac{1}{n+1}(x_{n+1}-\bar{x}_n)^2 \\
&= \sum_{i=1}^{n}(x_i-\bar{x}_n)^2 + \frac{n}{n+1}(x_{n+1}-\bar{x}_n)^2
\end{aligned}$$