STAT 391 Homework 6

Chongyi Xu
University of Washington
STAT 391 Spring 2018
chongyix@uw.edu

---

1. Problem 1 - Statistical decision making

a Make a neatly labeled table of the outcome space for this problem.
Table1 is the outcome space.

1st Day(Room B)	2nd Day(Room A)
$F, G\|B$	–
$!F, G\|B$	$!F\|B$
$!F, G\|A$	$F, G\|A$
	$F, !G\|A$
	$!F\|A$

Table 1: Outcome Space

b What is the probability that Rob finds the graphics card on the second day?

$$\begin{aligned} P(F \text{ on the second day}) &= P(\bar{F}|B) * P(F|A) * P(A) \\ &= 1 \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6} \end{aligned}$$

c What is the probability that he finds the graphics card?

$$\begin{aligned} P(F) &= P(F|A) + P(F|B) \\ &= \frac{1}{2} + \frac{2}{5} \\ &= \frac{9}{10} \end{aligned}$$

d What is the probability that he finds the card and the card is still good?

$$\begin{aligned} P(F,G) &= P(F|B) \cdot P(G|B) + P(F|A) \cdot P(G|A) \\ &= \frac{2}{5} + \frac{1}{2} \cdot \frac{4}{5} \\ &= \frac{4}{5} \end{aligned}$$

e What is the expected value of Rob's search policy?

1st Day(Room B)	2nd Day(Room A)
$F,G|B = \frac{2}{5} \cdot \frac{2}{3} = \frac{4}{15}$	–
$!F,G|B = \frac{3}{5} \cdot \frac{2}{3} = \frac{6}{15}$	$!F|B = 1$
$!F,G|A = 1$	$F,G|A = \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{4}{5} = \frac{2}{15}$
	$F,!G|A = \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{5} = \frac{1}{30}$
	$!F|A = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

Table 2: Outcome Space with probability

Considering the outcome space with probability Table2. So the expectation value is

$$\begin{aligned} E(\text{policy}) = &\frac{4}{15} \times (+60 - 10) + \frac{6}{15} \cdot 1 \times (-10 - 10 \times 2 - 5) \\ &+ 1 \cdot \frac{2}{15} \times (+60 - 10 \times 2 - 5) + 1 \cdot \frac{1}{30} \times (0 - 10 \times 2 - 5) \\ &+ 1 \cdot \frac{1}{6} \times (-10 - 10 \times 2 - 5) \\ = &-\frac{8}{3} \approx -2.67 \end{aligned}$$

f Help Rob compare the following policies:
Plan 1: Search B on 1st day, search A on 2nd day if necessary.
Plan 2: Search A on 1st ady, search B on 2nd day if necessary.
Plan 3: Search B on 1st day, don't search 2nd day.
Plan 4: Search A on 1st day, don't search 2nd day.
Plan 5: Don't search at all!
Rob would still have to choose his own decision criteria. Justify your answer by showing your reasoning or your calculations in all cases.

From part(e), we have found the $E(\text{Plan } 1) = -\frac{8}{3}$.

$$\begin{aligned}
E(\text{Plan } 2) &= \frac{1}{2\times 3}\times(+60-10) + \frac{1}{2\times 3}\cdot 1\times(-10-10\times 2-5)\\
&\quad + 1\cdot\frac{2\times 2\times 3}{3\times 5\times 5}\times(+60-10\times 2-5)\\
&\quad + 1\cdot\frac{2\times 2\times 2}{3\times 5\times 5}\times(0-10\times 2-5)\\
&\quad + 1\cdot\frac{2\times 3}{3\times 5}\times(-10-10\times 2-5)\\
&\approx -8.57\\
E(\text{Plan } 3) &= \frac{2}{5}\cdot\frac{2}{3}\times(+60-10) + \frac{3}{5}\cdot\frac{2}{3}\times(-10-10) + \frac{1}{3}\times(-10-10)\\
&\approx -1.33\\
E(\text{Plan } 4) &= \frac{1}{2}\cdot\frac{1}{3}\times(+60-10) + \frac{1}{2}\cdot\frac{1}{3}\times(-10-10) + \frac{2}{3}\times(-10-10)\\
&\approx -8.33\\
E(\text{Plan } 5) &= -10
\end{aligned}$$

Therefore, the best optimized policy is Plan 3 by maximizing the expected gain.

2. Problem 3- Bayesian Inference

a Compute the probability that a customer buys all three books under $\bar{P}_{ABC}$.

$$\bar{P}_{ABC}(1,1,1) = 0.6\cdot 0.3\cdot 0.4 = 0.072$$

b Compute the probability that a customer buys all three books under $\tilde{P}$

$$\tilde{P}_{ABC}(1,1,1) = 0.1 \cdot 0.4 = 0.04$$

c Using $\bar{P}$ and $\tilde{P}$ from above, determine if the likelihood that Robin Hood is a man is higher than the likelihood that the (s)he's a woman.

$$\bar{P}_{ABC}(1,1,0) = 0.6 \cdot 0.3 \cdot (1 - 0.4) = 0.108$$
$$\tilde{P}_{ABC}(1,1,0) = 0.1 \cdot (1 - 0.4) = 0.06$$

Therefore, the likelihood tells Robin Hood is more likely to be a woman.

d Determine the posterior probability that Robin Hood is a man.

$$\begin{aligned} P(man|A=1,B=1,C=0) &= \frac{P((1,1,0)|man)P(man)}{P(1,1,0)} \\ &= \frac{0.06 \cdot \frac{2}{3}}{0.06 + 0.108} \\ &= \frac{0.04}{0.168} \\ &\approx 0.238 \end{aligned}$$

e Determine the posterior probability that RObin Hood is a man if Al doesn't recall whether Robin ordered Book C or not.

$$\begin{aligned} P(man|(1,1)) &= P(man|(1,1,1)) + P(man|(1,1,0)) \\ &= \frac{P((1,1,1)|man)P(man)}{P(1,1,1)} + 0.238 \\ &= \frac{0.04 \cdot \frac{2}{3}}{0.112} + 0.238 \\ &= 0.238 + 0.238 = 0.476 \end{aligned}$$

f Compute the value of Likelihood Ratio(LR) for the data $A = 1, B = 1, C = 0$

$$LR(A=1,B=1,C=0) = \frac{\bar{P}_{ABC}(1,1,0)}{\tilde{P}_{ABC}(1,1,0)} = \frac{0.108}{0.06} = 1.8$$

Compute the value of the LR if the data consists of 3 customers $A_1 = 1, B_1 = 0, C_1 = 0, A_2 = 0, B_2 = 1, C_2 = 0, A_3 = 1, B_3 = 0, C_3 = 1$.

$$\begin{aligned}
LR(D) &= \frac{\bar{P}_{ABC}(1,0,0) \cdot \bar{P}_{ABC}(0,1,0) \cdot \bar{P}_{ABC}(1,0,1)}{\tilde{P}_{ABC}(1,0,0) \cdot \tilde{P}_{ABC}(0,1,0) \cdot \tilde{P}_{ABC}(1,0,1)} \\
&= \frac{0.6 * (1-0.3) * (1-0.4) * (1-0.6) * 0.3 * (1-0.4) * 0.6 * (1-0.3) * 0.4}{0.5 * (1-0.4) * 0.2 * 0.6 * 0.5 * 0.4} \\
&= 0.423
\end{aligned}$$

g Give an example of a data set where $LR > 1$.

$$D : A_1 = 1, B_1 = 1, C_1 = 0, A_2 = 1, B_2 = 1, C_2 = 1$$

3. Problem 5 - Dirichlet/Beta Distribution

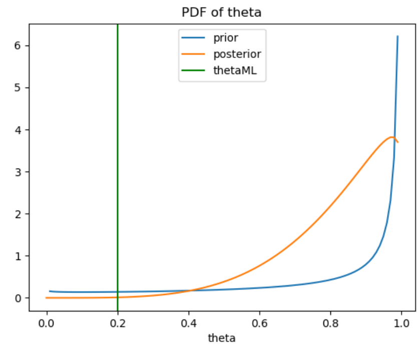
a Change the variables θ_j to $\xi_j = ln\theta_j$ and express L as a function of ξ.

$$L = P(\theta_1, \theta_2) = \theta_1^{n_1}\theta_2^{n_2} = n_1\xi_1 n_2\xi_2$$

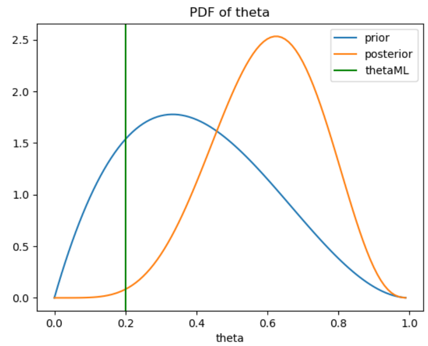
b Let $m = 2, S = \{1, 2\}$,and $D = \{1, 1, 2, 1, 1\}$. For the following 3 dirichlet priors, give numerical values of the fictitious sample size α, and the posterior parameters $\alpha'_{1,2}$

- $Diri(\theta_1, \theta_2; 0.9, 0.1)$
  For this case, $\alpha = 0.9 + 0.1 = 1$, $\alpha'_1 = 0.9 + 4 = 4.9$, $\alpha'_2 = 0.1 + 1 = 1.1$
- $Diri(\theta_1, \theta_2; 2, 3)$
  For this case, $\alpha = 2 + 3 = 5$, $\alpha'_1 = 2 + 4 = 6$, $\alpha'_2 = 3 + 1 = 4$
- $Diri(\theta_1, \theta_2; 20, 20$
  For this case, $\alpha = 20 + 20 = 40$, $\alpha'_1 = 20 + 4 = 24$, $\alpha'_2 = 20 + 1 = 21$

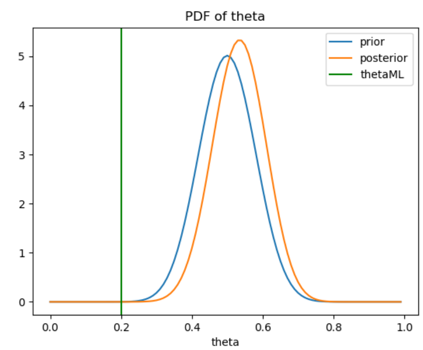
c Let $\theta = \theta_2$. For each of the 3 cases above, make a plot showing the prior, posterior as functions of θ, as well as the location of ML estimate θ_2^{ML} on the θ axis.



(a) PDF of *theta* for case $Diri(\theta_1, \theta_2; 0.9, 0.1)$



(b) PDF of *theta* for case $Diri(\theta_1, \theta_2; 2, 3)$



(c) PDF of *theta* for case $Diri(\theta_1, \theta_2; 20, 20)$

Figure 1: Estimation of logistic density

d Assume now the prior is uniform, that is $Diri(\theta_1, \theta_2; 1, 1)$. Show that the posterior of (θ_1, θ_2) is a Beta distribution and calculate its parameters for the data in **b**.
Consider we have a data set with n_1 observations of outcome 1 and n_2 observations of outcome 2. Then the posterior of (θ_1, θ_2) is

$$\begin{aligned} Diri(\theta_1, \theta_2; 1+n1, 1+n2) &= \frac{\Gamma(1+n1)}{\Gamma(1+n1)\Gamma(1+n2)} \theta_1^{n1} \theta_2^{n2} \\ &= Beta(\theta_1, \theta_2; 1+n1, 1+n2) \end{aligned}$$

So the posterior parameters for data in **b** is $\alpha_1^{'} = 1+4 = 5, \alpha_2^{'} =$

$1 + 1 = 2$.

e Same as **c**, plot prior, posterior and location of ML estimate for θ_2^{ML} for the uniform prior.

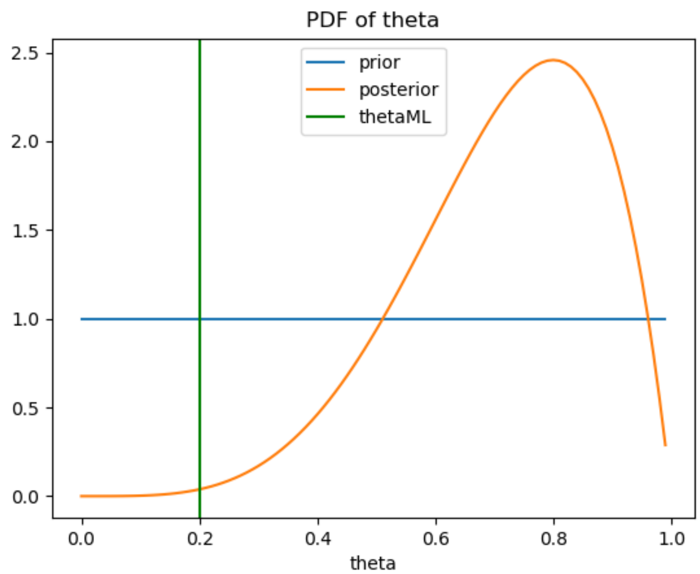


Figure 2: PDF of θ for $Diri(\theta_1, \theta_2; 5, 2)$

Figure2 shows the plot.