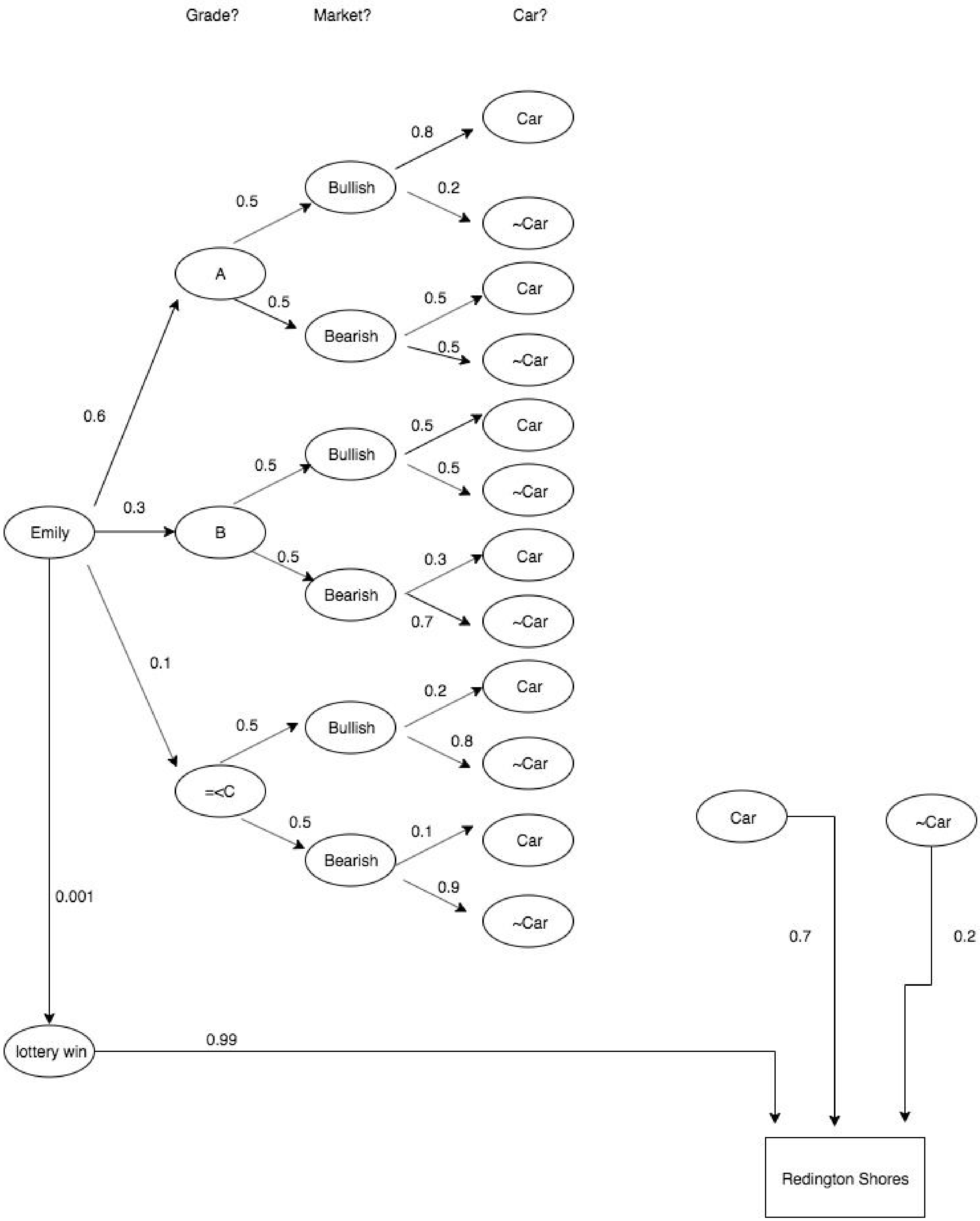
1.


Grade?
Market?
Car?
Emily
A
B
=<C
Bullish
Bearish
Car
~Car
lottery win
Redington Shores
0.6
0.3
0.1
0.001
0.5
0.8
0.2
0.3
0.7
0.9
0.99

Let R be the event that Emily goes to Redington Shores

Let C be the event that Emily gets a Car.

Let S be the event that Emily wins Sweepstakes lottery.

(a) P (C|R) = P(R|C) * P (C) / P(R)

Here,

P (R|C) = 0.7

P (R|~C) = 0.2

P (S) = 0.001

P (R|S) = 0.99

P(C) can be obtained by summing below from the above diagram:

Emily gets P(Grade=A) * P(Market=Bullish |A) * P(Car| Bullish, A) = 0.6 * 0.5 * 0.8

Emily gets P(Grade=A) * P(Market=Bearish |A) * P(Car| Bearish, A) = 0.6 * 0.5 * 0.5

Emily gets P(Grade=B) * P(Market=Bullish |B) * P(Car| Bullish, B) = 0.3 * 0.5 * 0.5

Emily gets P(Grade=B) * P(Market=Bearish |B) * P(Car| Bearish, B) = 0.3 * 0.5 * 0.3

Emily gets P(Grade<=C) * P(Market=Bullish |<=C) * P(Car| Bullish, <=C) = 0.1 * 0.5 * 0.2

Emily gets P(Grade<=C) * P(Market=Bearish |<=C) * P(Car| Bearish, <=C) = 0.1 * 0.5 * 0.1

=> P(C) = 0.525

P(~C) = 1 - P(C) = 0.475

P(R) can be given by:

P(R|C)*P(C) + P(R|~C) * P(~C) + P (R|S) * P (S)

= 0.7 * 0.525 + 0.2 * 0.475 + 0.99 * 0.001

= 0.46349

Therefore,

P (C|R) = P(R|C) * P (C) / P(R)

= 0.7 * 0.525/0.46349

= 0.79289736563

(b) P (S|R) = P(R|S) * P(S)/P(R)

(P(R) is calculated above and P(R|S) = 0.99, P(S) = 0.001 are given in the question)

= 0.99 * 0.001/0.46349

= 0.00213596841

(c) P (Grade = B | R) = P (R|Grade = B) * P(Grade = B)/P(R)

P(R) is calculated above and has value 0.46349
P(Grade = B) = 0.3

P (R|Grade = B) can be calculated by summing below:
Path via Car in the diagram
P(Market=Bullish |B) * P(Car| Bullish, B) * P(R|Car, Bullish, B) = 0.5 * 0.5 * 0.7
P(Market=Bearish |B) * P(Car| Bearish, B) * P(R|Car, Bearish, B) = 0.5 * 0.3 * 0.7

Path via ~Car in the diagram
P(Market=Bullish |B) * P(~Car| Bullish, B) * P(R|~Car, Bullish, B) = 0.5 * 0.5 * 0.2
P(Market=Bearish |B) * P(~Car| Bearish, B) * P(R|~Car, Bearish, B) = 0.5 * 0.7 * 0.2

=> P (R|Grade = B) = 0.4

Therefore,
P (Grade = B | R) = P (R|Grade = B) * P(Grade = B)/P(R)
= 0.4*0.3/0.46349
= 0.25890526224

(d)P(Market=Bearish|R) = P (R|Market=Bearish) * P (Market = Bearish)/P (R)
P(R) is calculated above and has value 0.46349
P (Market = Bearish) = 0.5

P (R|Market=Bearish) can be calculated by summing the following:
P(R|Car) * P(Car|Bearish, Grade = A)*P(Grade = A) = 0.7 * 0.5 * 0.6
P(R|Car) * P(Car|Bearish, Grade = B)*P(Grade = B) = 0.7 * 0.3 * 0.3
P(R|Car) * P(Car|Bearish, Grade <= C)*P(Grade <= C) = 0.7 * 0.1 * 0.1
P(R|~Car) * P(~Car|Bearish, Grade = A)*P(Grade = A) = 0.2 * 0.5 * 0.6
P(R|~Car) * P(~Car|Bearish, Grade = B)*P(Grade = B) = 0.2 * 0.7 * 0.3
P(R|~Car) * P(~Car|Bearish, Grade <= C)*P(Grade <= C) = 0.2 * 0.9 * 0.1

=> P (R|Market=Bearish) = 0.4
Therefore,
P(Market=Bearish|R) = P (R|Market=Bearish) * P (Market = Bearish)/P (R)
= 0.4 * 0.5/ 0.46349
= 0.43150877041

2. This is an example of Negative binomial distribution.

Let Y be the number of failures before the r^{th} success.

$P(Y = y | p) = {}^{r+y-1}C_y p^r (1-p)^y$

$\Rightarrow f(y|p) = {}^{r+y-1}C_y p^r (1-p)^y$

From the question, we know that there were in total 11 independent experiments with below failure values till the 4^{th} success:

$y_i = [5, 2, 2, 0, 1, 4, 3, 5, 0, 7, 1]$

r = 4

n =11

=> Likelihood = $\prod_{i=1}^{n} f(y_i|p) = \prod_{i=1}^{n} {}^{r+y_i-1}C_{y_i} p^r (1-p)^{y_i}$

=> Likelihood $\propto p^{\sum_{i=1}^{n} r} * (1-p)^{\sum_{i=1}^{n} y_i}$

Prior = $\pi(p)$ = Beta (a, b) $\propto p^{a-1}(1-p)^{b-1}$

Therefore,

Prior mean can be given by:

a/(a+b)

Posterior: $\pi(p \mid y_1, y_2, \ldots, y_n) \propto Likelihood * Prior$

$\propto p^{a-1+\sum_{i=1}^{n} r}(1-p)^{b-1+\sum_{i=1}^{n} y_i}$

From above it can be clearly seen that Posterior has a kernel of Beta (a + nr, b + $\sum_{i=1}^{n} y_i$)

Therefore, posterior mean can be given by:

(a + nr)/(a+nr + b + $\sum_{i=1}^{n} y_i$)

(a) If a = b = 1
    => Prior has distribution of Beta (a, b) = Beta (1, 1)
    => prior mean = 1/(1+1) = 0.5
    We know from question that:
    n = 11
    r = 4
    $\sum_{i=1}^{n} y_i = 5 + 2 + 2 + 0 + 1 + 4 + 3 + 5 + 0 + 7 + 1 = 30$
    => Posterior has the distribution: Beta (a + nr, b + $\sum_{i=1}^{n} y_i$)
    = Beta (1+44, 1 + 30) = Beta (45, 31)

    (i) Therefore, the bayes estimator of p can be given by:
    45/(45+31) = 0.59210526315

(ii) To find 95% (equal-tailed) credible set for p, we run the following two commands in R:

qbeta(0.025, 45, 31, 0)

```
[1] 0.4803705
```

qbeta(0.975, 45, 31, 0)

```
[1] 0.6992464
```

Therefore, our 95% (equal-tailed) credible set for p is (0.4803705, 0.6992464)

(iii) The posterior probability of hypothesis H: p >= 0.8 can be given by:

$${}_{0.8}\int^{\infty} Beta(45,\ 31)dp$$

$$=\ 1 - {}_{-\infty}\int^{0.8} Beta(45,\ 31)dp$$

Now,

The probability of ${}_{-\infty}\int^{0.8} Beta(45,\ 31)dp$ can be calculated from MATLAB using below code:

betacdf(0.8,45,31)

```
ans =   0.99998
```

Therefore, posterior probability of hypothesis H: p >= 0.8 is:

1 - 0.99998 = 0.00002

(b)

If a = b = 1/2

=> Prior has distribution of Beta (a, b) = Beta (1/2, 1/2)

=> prior mean = (½)/((½)+(1/2)) = 0.5

We know from question that:

n = 11

r = 4

$\sum_{i=1}^{n} y_i = 5 + 2 + 2 + 0 + 1 + 4 + 3 + 5 + 0 + 7 + 1 = 30$

=> Posterior has the distribution: Beta (a + nr, b + $\sum_{i=1}^{n} y_i$)

= Beta (1/2+44, 1/2 + 30) = Beta (44.5, 30.5)

(i) Therefore, the bayes estimator of p can be given by:

44.5/(44.5 + 30.5) = 0.59333333333

(ii) To find 95% (equal-tailed) credible set for p, we run the following two commands in R:

qbeta(0.025, 44.5, 30.5, 0)

```
[1] 0.4808762
```

qbeta(0.975, 44.5, 30.5, 0)

```
[1] 0.7010734
```

Therefore, our 95% (equal-tailed) credible set for p is (0.4808762, 0.7010734)

(iii) The posterior probability of hypothesis H: p >= 0.8 can be given by:

$$\int_{0.8}^{\infty} Beta(44.5,\ 30.5)dp$$

$$= \ 1 - \int_{-\infty}^{0.8} Beta(44.5,\ 30.5)dp$$

Now,

The probability of $\int_{-\infty}^{0.8} Beta(44.5,\ 30.5)dp$ can be calculated from MATLAB using below code:

```
betacdf(0.8,44.5,30.5)
ans =   0.99998
```

Therefore, posterior probability of hypothesis H: p >= 0.8 is:

1 - 0.99998 = 0.00002

(c) If a = 9, b = 1

=> Prior has distribution of Beta (a, b) = Beta (9, 1)

=> prior mean = 9/(9+1) = 0.9

We know from question that:

n = 11

r = 4

$\sum_{i=1}^{n} y_i = 5 + 2 + 2 + 0 + 1 + 4 + 3 + 5 + 0 + 7 + 1 = 30$

=> Posterior has the distribution: Beta (a + nr, b + $\sum_{i=1}^{n} y_i$)

= Beta (9 + 44, 1 + 30) = Beta (53, 31)

(i) Therefore, the bayes estimator of p can be given by:

53/(53+31) = 0.63095238095

(ii) To find 95% (equal-tailed) credible set for p, we run the following two commands in R:

```
qbeta(0.025, 53, 31, 0)
[1] 0.5257093
qbeta(0.975, 53, 31, 0)
[1] 0.7302874
```

Therefore, our 95% (equal-tailed) credible set for p is (0.5257093, 0.7302874)

(iii) The posterior probability of hypothesis H: p >= 0.8 can be given by:

$$\int_{0.8}^{\infty} Beta(53,\ 31)dp$$

$$= \ 1 - \int_{-\infty}^{0.8} Beta(53,\ 31)dp$$

Now,

The probability of $\int_{-\infty}^{0.8} Beta(53,\ 31)dp$ can be calculated from MATLAB using below code:

betacdf(0.8,53,31)

```
ans =  0.99981
```

Therefore, posterior probability of hypothesis H: p >= 0.8 is:
1 - 0.99981 = 0.00019

3.
The joint distribution is given by:

$f(y\,|u,\tau\,) = \prod^{n}_{i=1} f(y_i|u,\tau\,) * \pi(u) * \pi(\tau\,)$

$\propto \tau^{\,n/2} exp(\tau\,/2 * \sum_{i=1}{}^{n}(y_i - u)^2) * exp(-(u - u_0)^2/2) * \tau^{\,a-1} exp(-b\,\tau\,)$

To find the full conditional probability of u, we select terms from $f(y\,,u,\tau\,)$ that contains u and normalize.

$\pi\,(\,u\,|\,\tau\,,y\,\,) = \pi\,(\,u\,,\tau\,|y\,\,)/\,\pi\,(\,\tau\,|y\,\,) = \pi\,(\,u\,,\tau\,,y\,\,)/\pi\,(\,\tau\,,y\,\,) \propto \pi(u,\tau\,,y\,)$

Thus,

$\pi\,(\,u\,|\,\tau\,,y\,\,) \propto exp\{-\tau\,/2 * \sum_{i=1}{}^{n}(y_i - u)^2\} * exp\{\tau_0/2 * (u\, - u_0)^2\}$

$\propto exp\{-1/2\, * (\tau_0 + n\,\tau\,\,) * (u - (\tau \sum_{i=1}{}^{n} y_i + u_0\tau_0)/(\tau_0 + n\,\tau\,))^2\}$

From the above expression, we can see that this is a kernel of :

$N(\{\tau \sum_{i=1}{}^{n} y_i + u_0\tau_0\}/\{\tau_0 + n\,\tau\,\},\; 1/(\tau_0 + n\,\tau\,))$

y_i = [41, 44, 43, 47, 43, 46, 45, 42, 45, 45, 43, 45, 47, 40]

$u_0 = 45$

$\tau_0 = 1/4$

$n\, = 14$

(I am not putting the exact values of these constants in the kernel, as this is not asked in the question. My Octave code will calculate the exact values.)

Similarly,

$\pi\,(\,\tau\,\,|\,u,y\,) \;\propto \tau^{\,n/2} * exp\{-\tau\,/2 * \sum_{i=1}{}^{n}(y_i - u)^2\}\tau^{\,a-1} * exp(-b\,\tau\,)$

$= \tau^{\,n/2+a-1} * exp\{-\tau\,[b\, + 1/2 * \sum_{i=1}{}^{n}(y_i - u)^2]\}$

Clearly the above expression is the kernel of:

$Ga(a\, + n\,/2,\; b\, + 1/2 * \sum_{i=1}{}^{n}(y_i - u)^2)$

Where:

y_i = [41, 44, 43, 47, 43, 46, 45, 42, 45, 45, 43, 45, 47, 40]

$a\, = 4$

$b\, = 2$

$n\, = 14$

(I am not putting the exact values of these constants in the kernel, as this is not asked in the question. My Octave code will calculate the exact values.)

For sampling, I will use the original algorithm mentioned in the lectures which is given below:

$$
\begin{array}{l}
\text{Draw } \theta_1^{n+1} \text{ from } \pi(\theta_1|\theta_2^n, \theta_3^n, \ldots, \theta_p^n) \\
\text{Draw } \theta_2^{n+1} \text{ from } \pi(\theta_2|\theta_1^{n+1}, \theta_3^n, \ldots, \theta_p^n) \\
\text{Draw } \theta_3^{n+1} \text{ from } \pi(\theta_3|\theta_1^{n+1}, \theta_2^{n+1}, \theta_4^n, \ldots, \theta_p^n) \\
\ldots \\
\text{Draw } \theta_{p-1}^{n+1} \text{ from } \pi(\theta_{p-1}|\theta_1^{n+1}, \theta_2^{n+1}, \ldots, \theta_{p-2}^{n+1}, \theta_p^n) \\
\text{Draw } \theta_p^{n+1} \text{ from } \pi(\theta_p|\theta_1^{n+1}, \theta_2^{n+1}, \ldots, \theta_{p-1}^{n+1})
\end{array}
$$

For my case, I'll sample u and τ independently from one another and fixing one of them at a time.

Below is my exact code which I used for sampling using Octave (also attached with filename q3.m if you would like to run it). Note that I have set my initial values of $u = 45$ and $\tau = 1/4$ (The initial values won't matter much because we are going to burn 1000 samples anyway):

```
 close all
 clear all
  randn("state",1)
  randg("state",1)
  y = [41.00 44.00 43.00 47.00 43.00 46.00 45.00 42.00 45.00 45.00 43.00 45.00 47.00 40.00];
  n = length(y)
%------------------------------------------
NN = 11000;
mus = []; taus = [];
sumdata = sum(y);
%hyperparameters
mu0=45; tau0 = 1/4;
a= 4; b= 2;
% start, initial values
mu = 45;   tau = 1/4; %
for i = 1 : NN
  newmu  = sqrt(1/(tau0+n*tau)) * randn + (tau * sumdata+tau0*mu0)/(tau0+n*tau);
  %par   = b+1/2 * sum ( (y - mu).^2);
  par   = b+1/2 * sum ( (y - newmu).^2);
  newtau = gamrnd(a + n/2, 1/par); %par is rate
  mus = [mus newmu];
  taus = [taus newtau];
  mu=newmu;
  tau=newtau;
end
```

```
burn =1000;
mus = mus(burn+1:end);
taus=taus(burn+1:end);

mean(mus)
mean(taus)
 prctile(mus,2.5)
 prctile(mus,97.5)
length(mus(mus<45))/length(mus)
 prctile(taus,2.5)
 prctile(taus,97.5)
```

Through the above sampler code, I got, the following values as output:

```
ans =  44.049
ans =  0.33795
ans =  43.112
ans =  45.009
ans =  0.97430
ans =  0.16312
ans =  0.57161
```

The outputs are in the order of: mean of posterior u, mean of posterior τ, 2.5 percentile of posterior u, 97.5 percentile of posterior u, proportion of posterior u < 45 to all posterior u sampled, 2.5 percentile of posterior τ, 97.5 percentile of posterior τ.

Therefore,
Bayes estimate of u = 44.049
Bayes estimate of τ = 0.33795

(a) For calculating posterior probability of hypothesis that the researcher was interested in
$H_0 : u < 45$,
I have counted all the sampled posterior u < 45 and have divided by count of total sampled posterior u.

From this the ratio came out to be the below value:

```
ans =  0.97430
```

Therefore, the probability of u < 45 is roughly equal to 0.97430.

(b)
From the Octave code, I got the following values for 2.5 percentile and 97.5 percentile for my equitailed credible set for τ:

```
ans =  0.16312
ans =  0.57161
```

Therefore, 95% equitailed credible set for τ = [0.16312, 0.57161]