Homework 1

Jeremy Lin

1/15/2021

Problem 1

Compute the posterior distribution $p(\theta|x = 3)$.

θ	0.3	0.4	.5	.6	.7
$p(\theta)$	0.022	0.038	0.831	0.057	0.051

Compute the posterior mean, $E(\theta|x = 3)$, which can be assumed as a possible estimator of the population proportion, based on the observed data.

$$E(\theta|x = 3) = \sum \theta * p(\theta|x = 3) \tag{1}$$
$$E(\theta|x = 3) = 0.507778 \tag{2}$$

Suppose I conduct a more thorough study, and obtain data from 10 more individuals. Among them, eight clicked on the post. Based on the currently available information, how would you update your inference on the parameter θ?

The updated posterior distribution $p(\theta|x = 8)$

θ	0.3	0.4	.5	.6	.7
$p(\theta)$	0.001	0.007	0.653	0.124	0.214

Based on the results from (3), would you conclude that the campaign marketing post is successful, or you'd look for a better post? [This is a thought-provoking question at this point; try your best about to answer it]

To conclude that campaign marketing post is successful, I will use a cut off of 50% chance or greater that a person will click on the post ($\theta > 0.6$. The probability of theres is more than 60% chance a person will click on the post is equal to 0.463, Therefore I think this marketing is succesfull.

Problem 2

Based on the interim analysis, what is the posterior distribution of θ?

From interim analysis 20(y = 20) out of 30(n = 30) patients had a positive outcome. Assume uniform prior on θ.

Posterior distribution :

$$\theta|y \sim Beta(21, 11)$$

What is a possible point estimate for θ?

Possible point estimate :

$$E(\theta|y) = \frac{a+y}{a+b+n} = \frac{21}{32} \approx 0.656 \tag{3}$$

Find the 5% and 95% quantiles of the posterior distribution.

The 5% quantile is 0.5145839 and the 95% quantile is 0.7866395

In the continuation of the trial, 50 out of the remaining patients have had a positive outcome. What is the updated posterior of θ? What is a possible estimate for θ? Find the 5% and 95% quantiles of the distribution.

The updated posterior θ is :

$$\theta|y \sim Beta(71, 31)$$

Possible point estimate :

$$E(\theta|y) = \frac{a+y}{a+b+n} = \frac{71}{102} \approx 0.696 \tag{4}$$

The 5% quantile is 0.619 and the 95% quantile is 0.768

Based on the results from the analysis, does the treatment appear promising?

Since we are 90% sure that the percentage of patient having a positive outcome because of the treatment is between between 61.935% and 76.841%. therefore I believe that the treatment appear promising.

Problem 3

Describe your model for studying the clutch success probability of each subject, including the likelihood and prior.

We want to study the clutch success probability, i.e the proportion of number of clutch success over the number of clutch attempts. For this type of data, we assume Binomial likelihood, with a uniform prior (Beta(1,1)) distribution. Therefore our likelihood function is :

$$L(y|\theta) \propto \theta^y (1-\theta)^{n-y}$$

with prior distribution

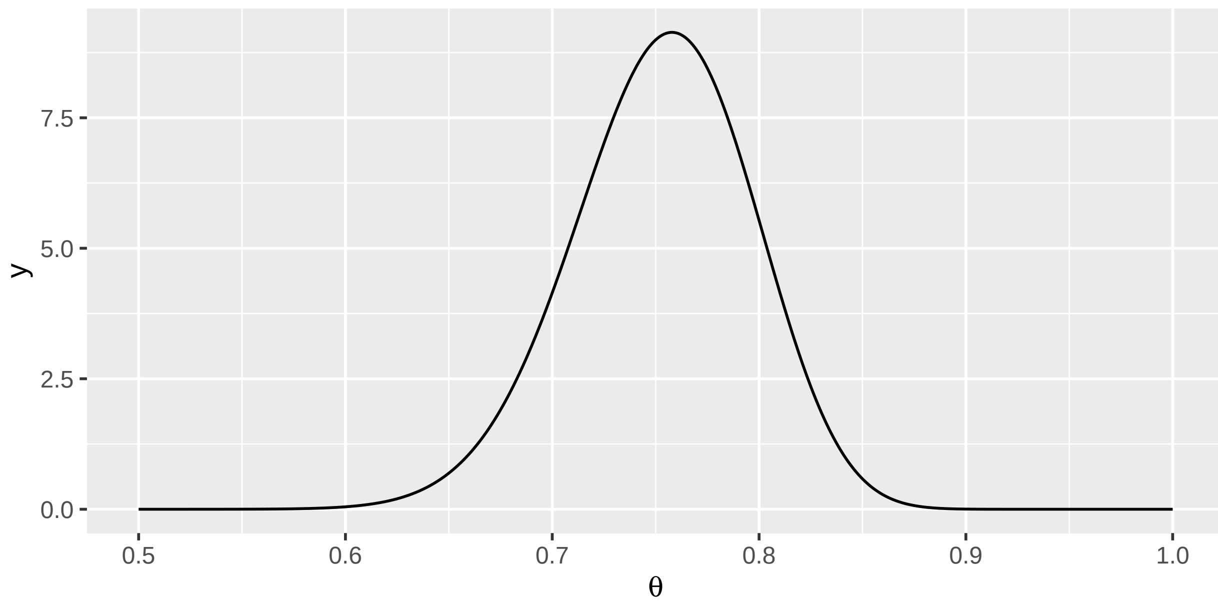
$$\theta_{prior} \sim Beta(1, 1)$$

Plot the posterior distribution of the clutch success probabilities for James Harden and Le Bron James.

Posterior Distribution of James Harden :

$$p(\theta|y) \sim Beta(73, 24)$$

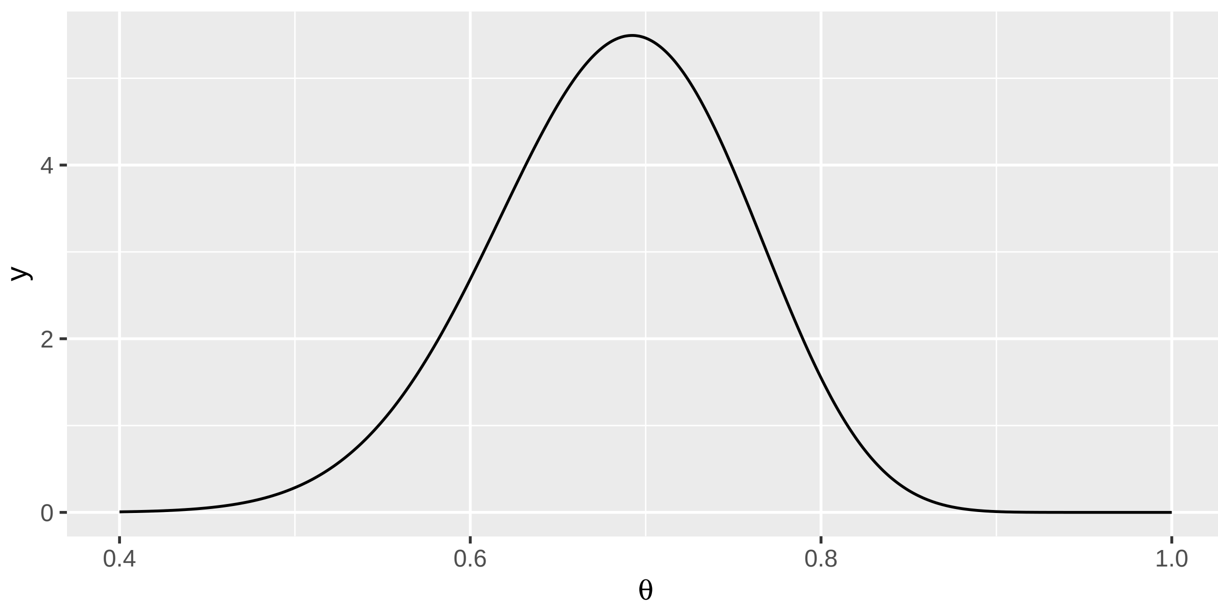
with the distribution plot:



Posterior Distribution of Lebron James

$$p(\theta|y) \sim Beta(28, 13)$$

with the distribution plot:



Summarize the posterior distribution *for each player* in a table. The table should include (at least) the posterior mean and the (5%, 25%, 50%, 75%, 95%) posterior quantiles.

Table 3 is the summary of the posterior distribution for each player.

Table 3: Player Summary

Player	post_mean	post_5	post_25	post_50	post_75	post_95
Russel Westbrook	0.8441558	0.7718157	0.8180802	0.8471422	0.8734561	0.9062892
James Harden	0.7525773	0.6779739	0.7240035	0.7543189	0.7830364	0.8212314
Kawhi Leonard	0.8615385	0.7857443	0.8349456	0.8652532	0.8921340	0.9246310
LeBron James	0.6829268	0.5597203	0.6354949	0.6859271	0.7335766	0.7958659
Isaiah Thomas	0.8941176	0.8347164	0.8735612	0.8972081	0.9180112	0.9429563
Stephen Curry	0.8928571	0.7847000	0.8599061	0.9021872	0.9356802	0.9690216
Giannis Antetokoumpo	0.6744186	0.5535649	0.6276477	0.6771455	0.7241164	0.7859418
John Wall	0.7976190	0.7220548	0.7694267	0.7999885	0.8283730	0.8650869
Anthony Davis	0.7321429	0.6310221	0.6937741	0.7349227	0.7735056	0.8237581
Kevin Durant	0.7777778	0.6043586	0.7179257	0.7882150	0.8485597	0.9153549

Do you find evidence that any of the players have a different clutch percentage than overall percentage? [*Hint: test the hypothesis that the probability of "clutch" for a player is greater than their corresponding overall proportion*]

Let the overall proportion be π , Table 4 is the probability of clutch proportion is greater than their corresponding overall proportion ($p(\theta > \pi)$). James Harden, Giannis Antetokoumpo, Anthony Davis, and Kevin Durant has a pretty low probably that could indicate that the player have a different clutch percentage than their corresponding overall percentage. This could due because these players tend to avoid taking clutch shots (low number of clutch attempt), with an exception of James Harden Who have the most attempt through all the player.

Table 4 also provided the 90% credible interval of clutch percentage and James Harden overall proportion is the only one inside the 90% CI, therefore we have a strong evidence that James Harden have a different clutch percentage than his shot percentage.

Table 4: Player Summary

Player	probability	post_5	post_95	overal_proportion
Russel Westbrook	0.5207121	0.7718157	0.9062892	0.845
James Harden	0.0090208	0.6779739	0.8212314	0.847
Kawhi Leonard	0.3597559	0.7857443	0.9246310	0.880
LeBron James	0.5645547	0.5597203	0.7958659	0.674
Isaiah Thomas	0.3553753	0.8347164	0.9429563	0.909
Stephen Curry	0.5293229	0.7847000	0.9690216	0.898
Giannis Antetokoumpo	0.0833433	0.5535649	0.7859418	0.770
John Wall	0.4907634	0.7220548	0.8650869	0.801
Anthony Davis	0.1133361	0.6310221	0.8237581	0.802
Kevin Durant	0.1542541	0.6043586	0.9153549	0.875