

Bayesian Data Analysis

Bayesian Inference with applications to SPTs



Sabarno Saha

2025-03-30

IISERK



1. The Mighty Bayes theorem	2
2. Toy Example 1 : Coin Toss	9
2.1 Uniform Prior	14
2.2 Uniform Prior for 20 coin tosses.	15
2.3 Beta Prior	16
2.4 Beta Prior for 20 coin tosses.	17
2.5 Priors, LL, and Posterior for Beta prior	18
2.6 Comparison Between Uniform and Beta Priors	19
2.7 Bayesian to Frequentist thinking	20
3. Toy Example 2: Finding Evidence (Z)	23
4. Information Theory	24
4.1 Nested Sampling	24
4.2 Likelihood Determination	24

1. The Mighty Bayes theorem

1. The Mighty Bayes theorem



The mighty Bayes Theorem

Definition 1.1 (Bayes Theorem)

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Let us slightly modify the above equation to suit our needs.

1. The Mighty Bayes theorem



Problem Description :

Suppose we get some data D and we want to infer the model parameters θ and the model M_i that generated the data.

1. The Mighty Bayes theorem



Problem Description :

Suppose we get some data D and we want to infer the model parameters θ and the model M_i that generated the data.

Aloha! Bayes

We will try to use Bayes Theorem to infer the model parameters.

1. The Mighty Bayes theorem



Let M be the model and D be the data we obtained. Then Bayes Theorem can be written as

$$\begin{array}{c} \text{Posterior Distribution} \\ \overbrace{P(M|D)} \end{array} = \frac{\overbrace{P(D|M)}^{\text{Likelihood}} \overbrace{P(M)}^{\text{Prior}}}{\underbrace{P(D)}_{\text{Evidence}}}$$

1. The Mighty Bayes theorem



- **Posterior Distribution** [$P(M|D)$] : The probability of the model given the data.
- **Likelihood** [$P(D|M)$] : The probability of the data given the model.
- **Prior** [$P(M)$] : The probability of the model.
- **Evidence** [$P(D)$] : The probability of the data.

1. The Mighty Bayes theorem



Let $M_i \in \mathbb{M}$ where \mathbb{M} is the set of models considered for the analysis. Let θ be a vector consisting of the parameters of the model. The *practical* Bayes Theorem as

$$P(\theta|\mathbf{D}, M_i) = \frac{P(\mathbf{D}|\theta, M_i)P(\theta|M_i)}{P(\mathbf{D}|M_i)}$$

1. The Mighty Bayes theorem



- **Posterior Distribution** [$P(\theta|D, M_i)$] : The probability of the parameters given the data and the model.
- **Likelihood** [$P(D|\theta, M_i)$] : The probability of the data given the parameters and the model.
- **Prior** [$P(\theta|M_i)$] : The probability of the parameters given the model.
- **Evidence** [$P(D|M_i)$] : The probability of the data given the model.

2. Toy Example 1 : Coin Toss

2. Toy Example 1 : Coin Toss



Problem:

Suppose we have a coin and we want to infer the probability of getting a head, say p . Let us take n , say $n = 20$ tosses out of which r are heads. We want to infer the value of p given the data

$$\theta = \{p\}$$

. Here r is our data and p is the parameter we want to infer.

2. Toy Example 1 : Coin Toss



Now we know that the random variable representing one coin toss $X_c \sim \text{Bernoulli}(p)$. Then after getting r heads in n tosses, the likelihood function is a Binomial distribution. The likelihood function is given as

$$P(r|\boldsymbol{\theta}, n) = P(r|p, n) = \binom{n}{r} p^r (1 - p)^{n-r}$$

The Beta Prior is a conjugate prior for the Binomial likelihood function. The posterior distribution for the Uniform Prior is given as

2. Toy Example 1 : Coin Toss



We will use two priors here :

- **Uniform Prior :**

$$P(p) = 1$$

for $0 \leq p \leq 1$

- **Beta Prior :**

$$P(p) = \beta(p|a, b) = \left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \right) p^{a-1} (1-p)^{b-1}$$

for $0 \leq p \leq 1$

2. Toy Example 1 : Coin Toss



The Uniform Prior is used when we have no prior information about the parameter.

The Beta Prior is used when we have some prior information about the parameter. We assume that the coin is somewhat fair and choose $a = b = \frac{n}{2} = 10$.

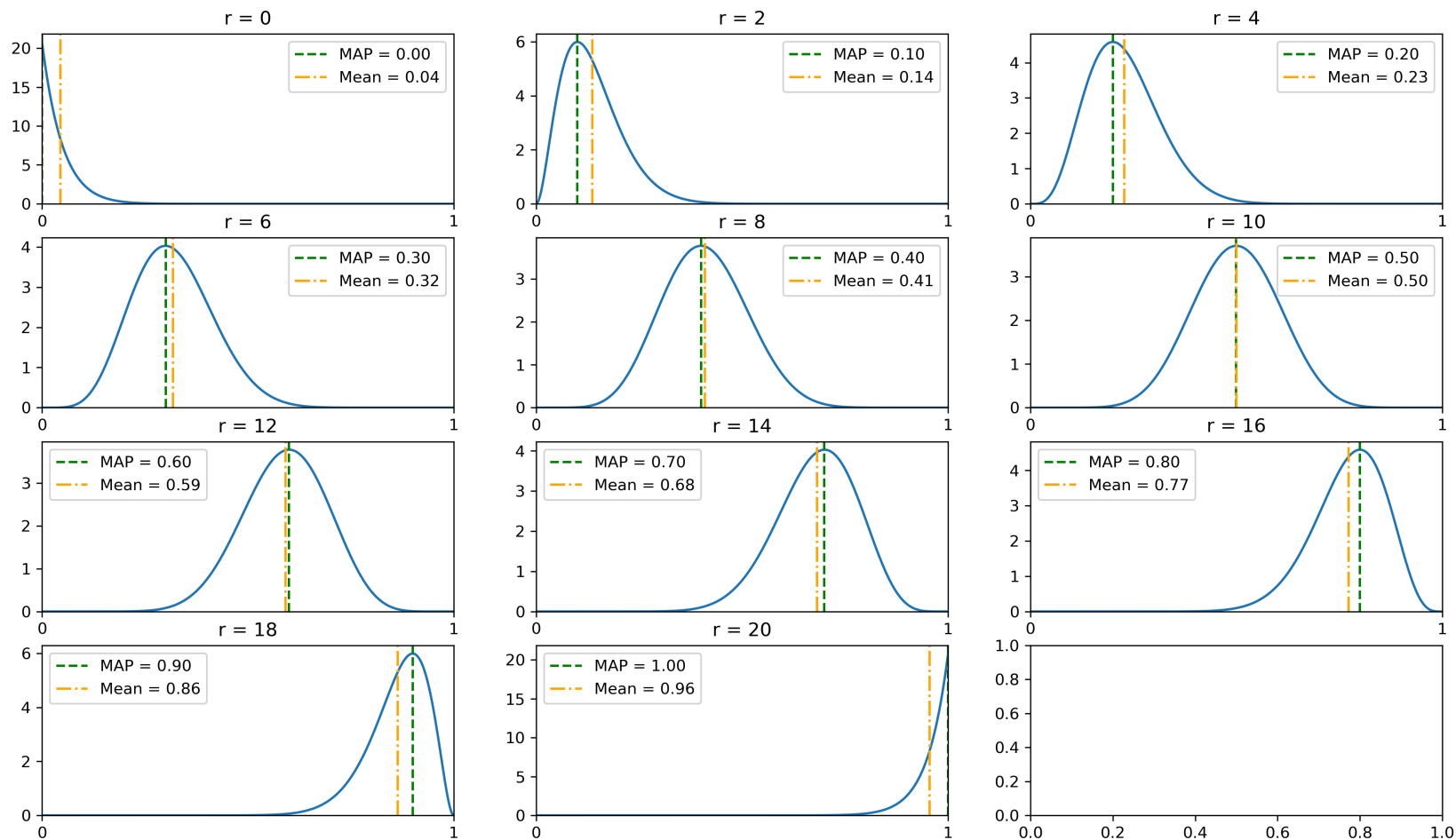


Using bayes theorem, we can calculate the posterior distribution for the Uniform Prior to be just proportional to the likelihood function. The posterior distribution is given as

$$P(p|r, n) = \frac{1}{Z} \binom{n}{r} p^r (1 - p)^{n-r}$$

where Z is the evidence or the normalizing constant.

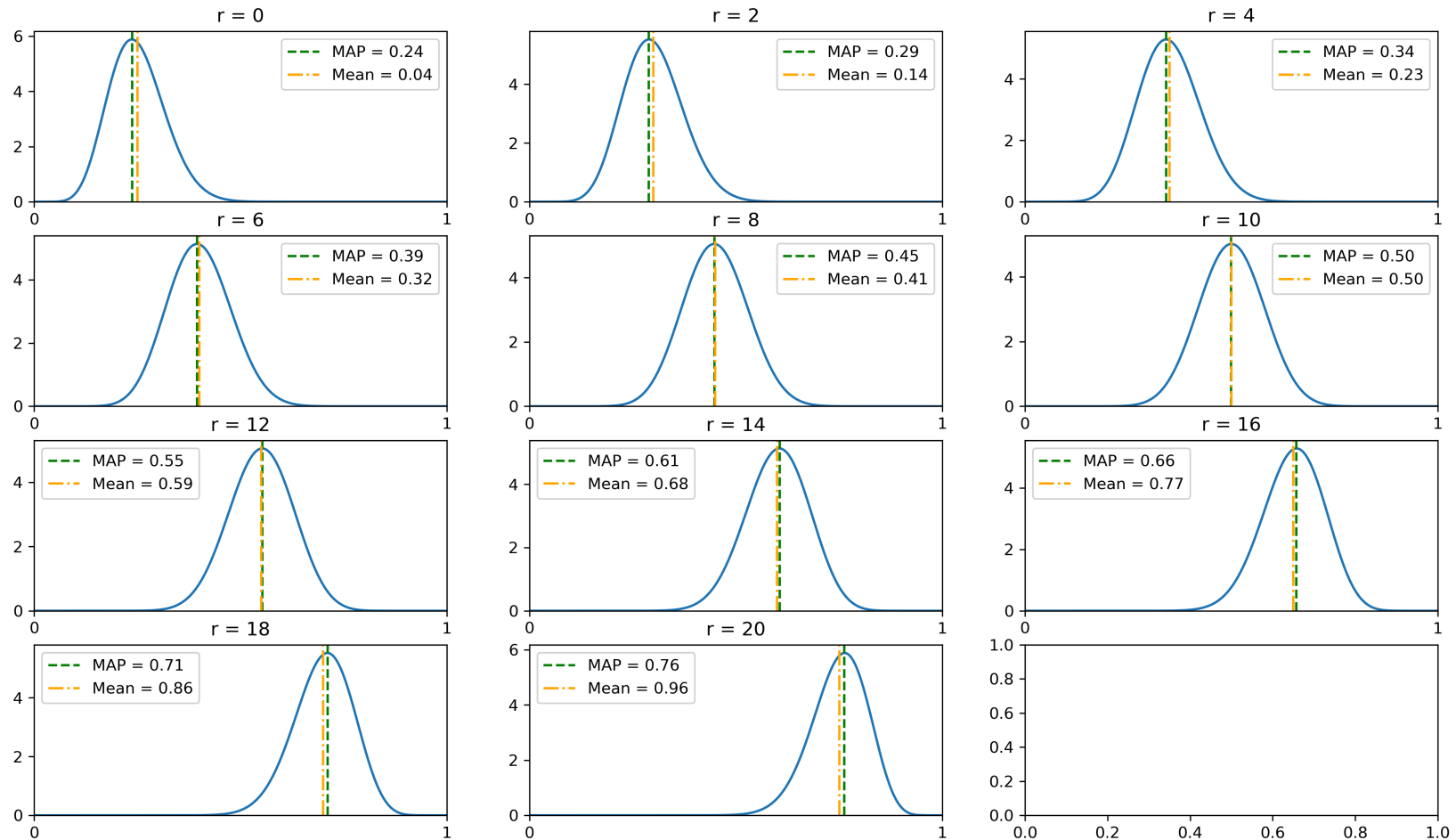
2.2 Uniform Prior for 20 coin tosses.



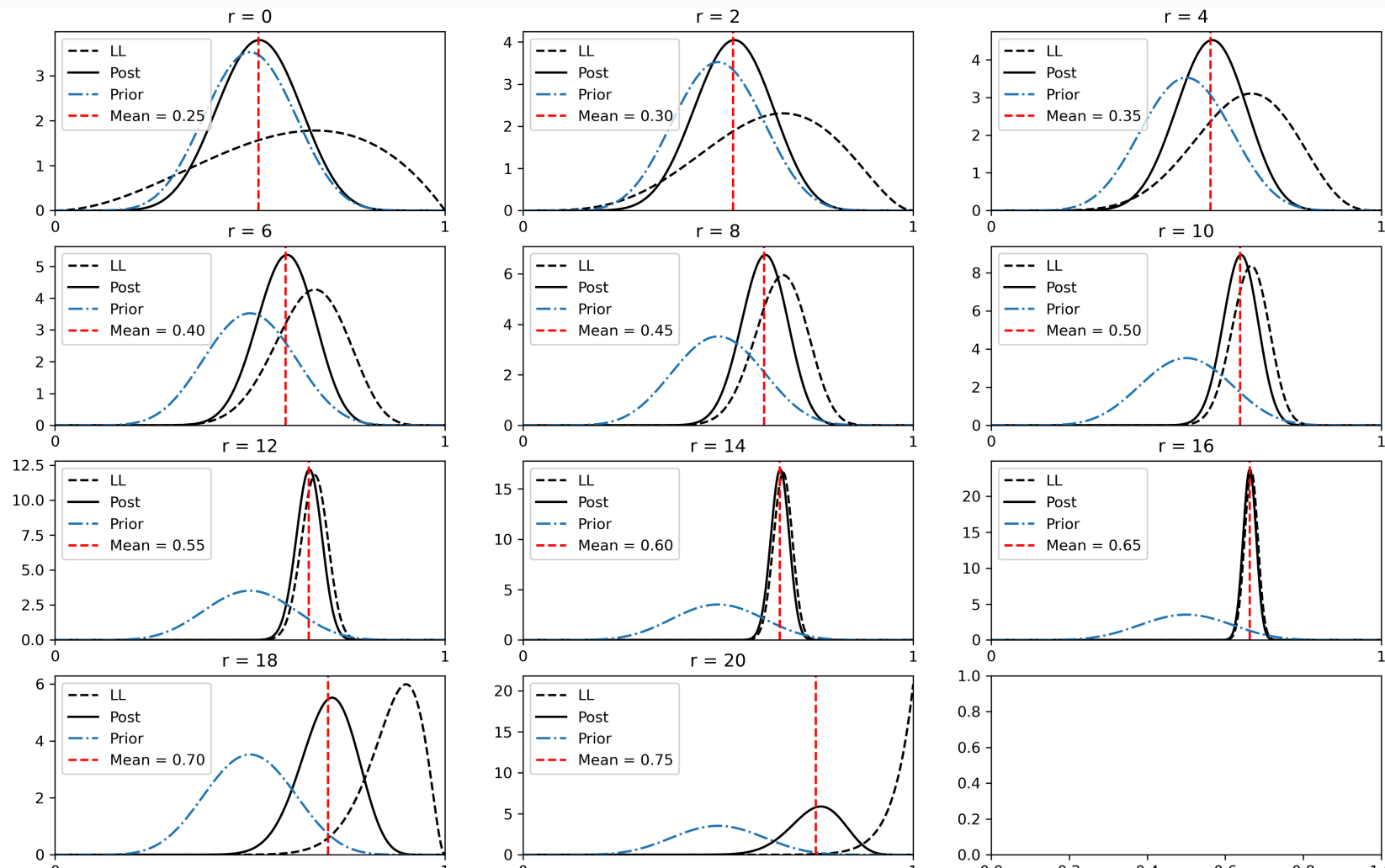


1

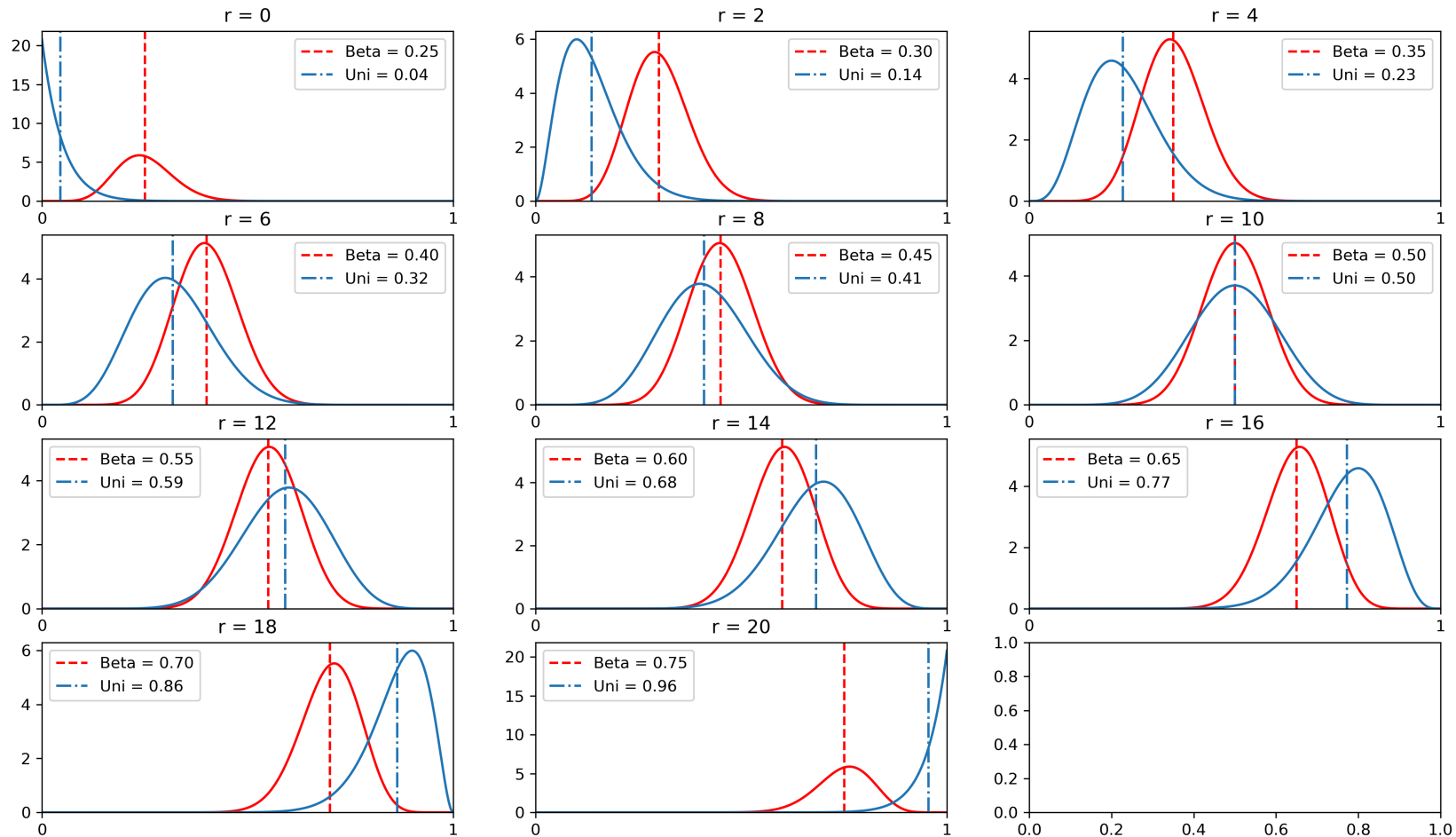
2.4 Beta Prior for 20 coin tosses.



2.5 Priors, LL, and Posterior for Beta prior



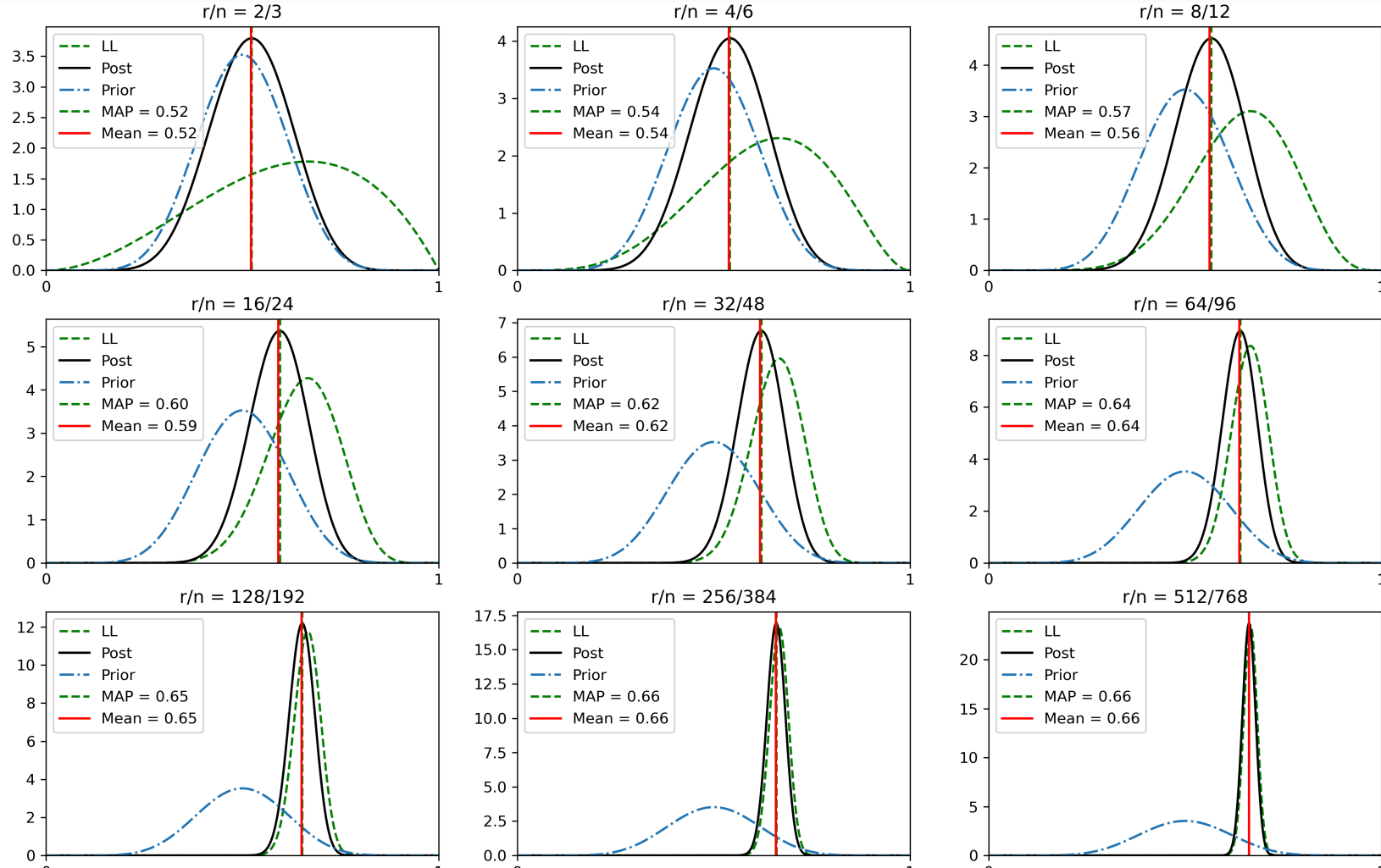
2.6 Comparison Between Uniform and Beta Priors





- **Bayesian** : We have a prior distribution for the parameter and we update it using the data to get the posterior distribution.
- **Frequentist** : We , theoretically, have an infinite number of data points and we use the data to estimate the parameter.

2.7 Bayesian to Frequentist thinking



2.7 Bayesian to Frequentist thinking



—

3. Toy Example 2: Finding Evidence (Z)

4. Information Theory
