So that's the posterior:

$$\Pr(p \mid k \text{ heads}) = (n+1)\binom{n}{k} p^k (1-p)^{n-k}$$

$\boxed{\Rightarrow \text{Let's plot this} \Rightarrow \text{slides!}}$

Bayes Factor (Model Selection)

$$P(M \mid D) = \frac{P(D \mid M) P(M)}{P(D)} \qquad M\text{-model, } D\text{-data}$$

Suppose we have two models, M_1 and M_2. We can compare them relative to the data w/ the posterior odds ratio:

$$\text{Posterior odds ratio} = \frac{P(M_1 \mid D)}{P(M_2 \mid D)} = \frac{P(M_1)}{P(M_2)} \underbrace{\frac{P(D \mid M_1)}{P(D \mid M_2)}}_{\substack{\text{"Bayes Factor"} \\ = K}}$$

$$K = \frac{P(D \mid M_1)}{P(D \mid M_2)} = \frac{\int P(D \mid \beta_1, M_1) P(\beta_1 \mid M_1) d\beta_1}{\int P(D \mid \beta_2, M_2) P(\beta_2 \mid M_2) d\beta_2}$$

(more generally assuming model M_i is parameterized by a parameter β_i)

In the context of our coin problem:

Compare unbiased coin ($= M_1$) vs. biased coin ($" M_2$):

$n = 100$, $k = 60$. as before

models: $P(r \mid M_1) = \delta_{r,\frac{1}{2}} = \begin{cases} 1 & \text{if } r = 1/2, \\ 0 & \text{if } r \neq 1/2. \end{cases}$ "delta"

$P(D \mid \beta, M) \Rightarrow P(D \mid r, M)$

$P(r \mid M_2) = 1$ "uniform"

$$K = \frac{\int P(D \mid r, \text{unbiased}) P(r \mid \text{unbiased}) dr}{\int P(D \mid r, \text{biased}) P(r \mid \text{biased}) dr} = \frac{\int P(D \mid r, \text{unbiased}) \delta_{r,1/2} dr}{\int P(D \mid r, \text{biased}) \cdot 1 \, dr}$$

$$= \frac{P(D \mid 1/2)}{\int_0^1 P(D \mid r) dr} = \frac{\binom{n}{k}\left(\frac{1}{2}\right)^n}{\int_0^1 \binom{n}{k} r^k (1-r)^{n-k} dr} = \frac{0.01084}{\frac{1}{101}} = 1.095 = K$$

If instead $n=100$, $k=70 \Rightarrow K = 0.00234$. or $\frac{1}{K} = \underline{427.3}$

Interpreting K:

$K > 1$ model 1 better supported by data than model 2.

$1 < K < 3$	barely worth mentioning
$3 < K < 10$	substantial
$10 < K < 30$	strong
$30 < K < 100$	very strong
$K > 100$	decisive

(Jeffreys, 1961)