

Bayesian statistics

# So sánh tỷ lệ Frequentist vs. Bayesian

---

Khương Quỳnh Long

Hà Nội, 08/2019

<https://gitlab.com/LongKhuong/adhere-bayesian-statistics>

# Hiệu quả thuốc RU-486

- ▶ **Câu hỏi nghiên cứu:** Thuốc tránh thai khẩn cấp RU-486 có hiệu quả tránh thai cao hơn phương pháp thông thường hay không?
- ▶ **Đối tượng:** 40 phụ nữ đến cơ sở y tế sử dụng thuốc tránh thai khẩn cấp
- ▶ **Thiết kế:** 40 phụ nữ này được phân ngẫu nhiên vào 2 nhóm (mỗi nhóm 20): sử dụng RU-486 và biện pháp thông thường
- ▶ **Dữ liệu:**
  - ✓ 4/20 phụ nữ nhóm RU-486 mang thai
  - ✓ 16/20 phụ nữ nhóm thông thường mang thai

Source: <https://www.coursera.org/learn/bayesian>

# Hiệu quả thuốc RU-486

- ▶ Đơn giản hóa: so sánh một tỉ lệ
- ✓ Tính trên 20 phụ nữ mang thai (4 trong nhóm RU-486 và 16 trong nhóm thông thường)
- ✓ RU-486 có hiệu quả như thế nào, so sánh với biện pháp thông thường?
  - Nếu hiệu quả 2 phương pháp là như nhau  
→  $P(\text{có thai từ nhóm RU-486}) = 0.5$
  - Nếu RU-486 hiệu quả hơn →  $P(\text{có thai từ nhóm RU-486}) < 0.5$
  - Nếu RU-486 kém hiệu quả hơn →  $P(\text{có thai từ nhóm RU-486}) > 0.5$

# Cách tiếp cận theo Frequentist

# Giả thuyết

---

$H_0: (p = 0.5)$  – Không có khác biệt. Có thai xảy ra ở 2 nhóm là như nhau

$H_A: (p < 0.5)$  – RU-486 hiệu quả hơn. Có thai xảy ra ít hơn ở nhóm RU-486

## P-value

- ▶  $K = 4$ ,  $n = 20$  (tổng cộng 20 phụ nữ có thai, 4 trong nhóm RU-486)
- ▶ P-value: xác suất quan sát được kết cuộc và những kết cuộc “hiếm hơn”.
- ▶  $P\text{-value} = P(k \leq 4; n = 20; p = 0.5)$

```
sum(dbinom(0:4, size = 20, p = 0.5))  
[1] 0.005908966
```

Bác bỏ  $H_0 \rightarrow$  tỉ lệ có thai ở nhóm RU-486 thấp hơn so với nhóm thông thường, P 1 tail = 0.0059

# Cách tiếp cận theo Bayesian

# Giả thuyết (#mô hình)

---

- ▶ Bayesian không có giả thuyết  $H_0$
- ▶  $p$  (xác suất có thai xảy ra ở nhóm RU-486) có thể nhận giá trị bất kì từ 0 tới 1  $\rightarrow$  vô số trường hợp (#mô hình)
- ▶ Giả định  $p$  có thể là  
0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9
- ▶ Cân nhắc 9 mô hình, so với 1 như frequentist



# Bayes formula

---

$$\text{Posterior} = \frac{\text{Likelihood} * \text{Prior}}{\text{Marginal likelihood}}$$

# Prior

- ▶ Mức độ tin tưởng (xác suất có thai trong nhóm RU-486 xảy ra) **trước khi thu thập dữ liệu**
- ▶ Giả sử prior cho 9 mô hình như sau

Model (p)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Tổng
Prior	0.06	0.06	0.06	0.06	0.52	0.06	0.06	0.06	0.06	1

- ▶ Hàm ý
- ✓ Hiệu quả thuốc RU-486 là cân bằng: có thể tốt hơn hoặc xấu hơn
- ✓ 52% không có khác biệt giữa 2 phương pháp

# Likelihood

- ▶  $P(\text{data} \mid \text{model})$  cho mỗi mô hình (9 mô hình)
- ▶ Xác suất này được gọi là **likelihood**

$$P(\text{data} \mid \text{model}) = P(k = 4 \mid n = 20, p)$$

Model (p)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Tổng
Prior, $P(\text{model})$	0.06	0.06	0.06	0.06	0.52	0.06	0.06	0.06	0.06	1
Likelihood	0.0898	0.2182	0.1304	0.035	0.0046	0.0003	0.00	0.00	0.00	

**Likelihood:** Nếu model 1 đúng (xác suất có thai từ nhóm RU-486 là 0.1) thì xác suất quan sát được 4 trường hợp có thai từ nhóm RU-486 trong 20 trường hợp là 0.0898

# Posterior

---

Tính xác suất hậu định bằng công thức Bayes  $P(model | data)$

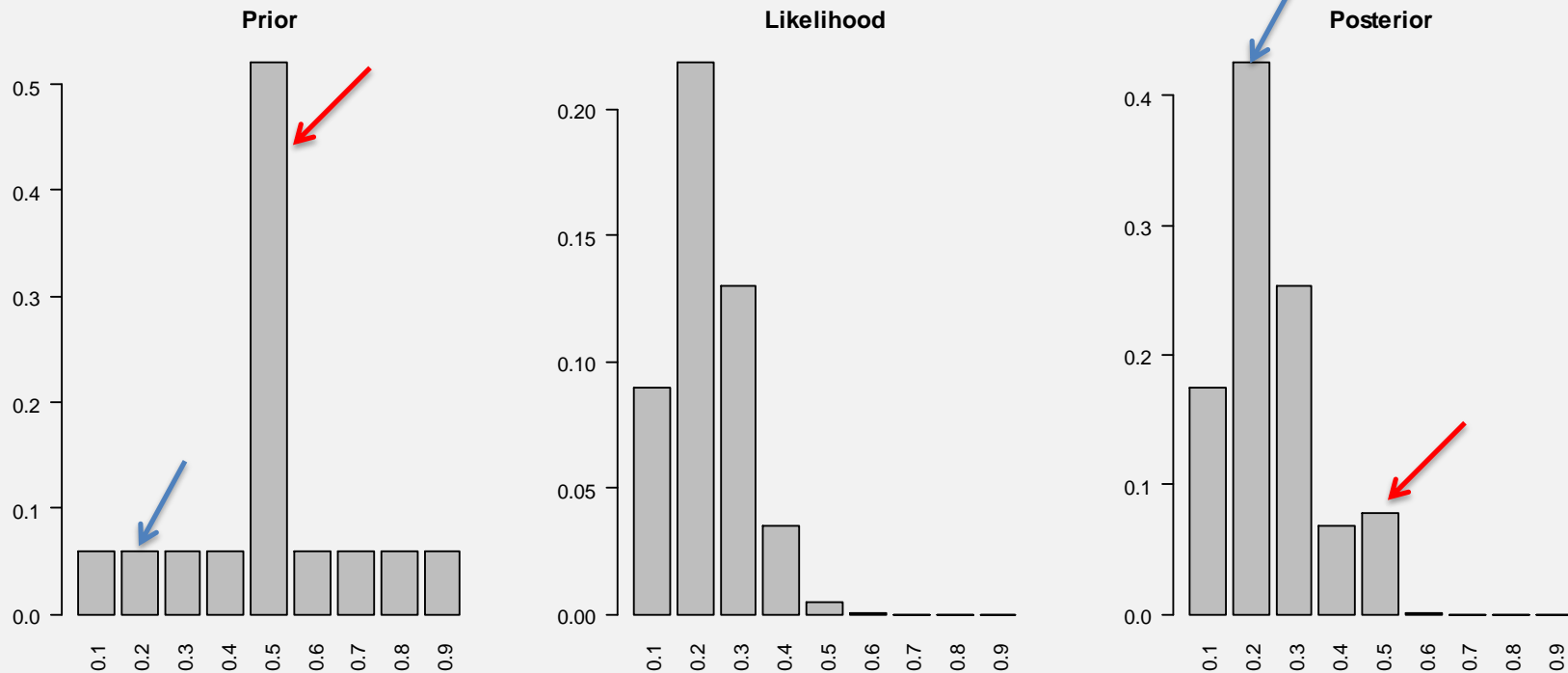
$$P(model | data) = \frac{P(data | model) * P(model)}{P(data)}$$

Tổng  $P(data | model) * P(model)$  của tất cả trường hợp (9 trường hợp)

Model (p)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Tổng
Prior, $P(\text{model})$	0.06	0.06	0.06	0.06	0.52	0.06	0.06	0.06	0.06	1
Likelihood, $P(\text{data} \text{model})$	0.0898	0.2182	0.1304	0.035	0.0046	0.0003	0	0	0	
$P(\text{data} \text{model}) * P(\text{model})$	0.0054	0.0131	0.0078	0.0021	0.0024	0	0	0	0	0.0308 $P(\text{data})$
Posterior, $P(\text{model}   \text{data})$	0.1748	0.4248	0.2539	0.0681	0.0780	0.0005	0	0	0	1

- ▶ Mô hình 0.2 (xác suất có thai từ nhóm RU-486 = 0.2) có xác suất hậu định lớn nhất = 42.48%
- ▶ Mô hình 0.5 (giả định không có khác biệt giữa 2 nhóm) có Prior là 52% giảm xuống 7.8% ở Posterior → update mức độ tin tưởng sau khi thu thập dữ liệu

# Visualization



Model (p)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Tổng
Prior, $P(\text{model})$	0.06	0.06	0.06	0.06	0.52	0.06	0.06	0.06	0.06	1
Likelihood, $P(\text{data} \text{model})$	0.0898	0.2182	0.1304	0.035	0.0046	0.0003	0	0	0	
$P(\text{data} \text{model}) * P(\text{model})$	0.0054	0.0131	0.0078	0.0021	0.0024	0	0	0	0	0.0308 $P(\text{data})$
Posterior, $P(\text{model}   \text{data})$	0.1748	0.4248	0.2539	0.0681	0.0780	0.0005	0	0	0	1

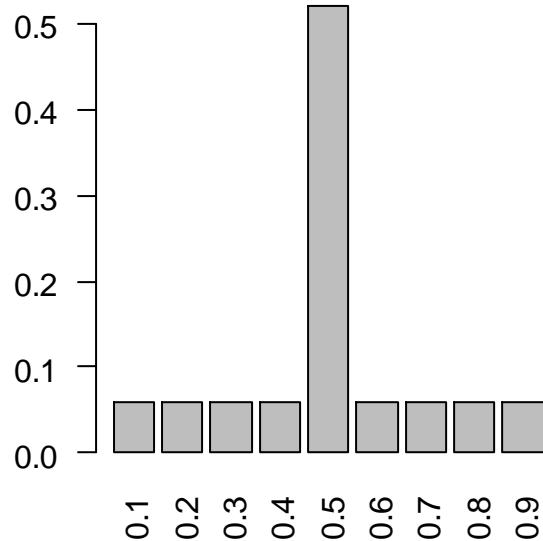
- ▶ Kết luận: RU-486 có hiệu quả tránh thai cao hơn phương pháp thông thường.
- ▶ Xác suất Posterior  $< 0.5 = 92.16\%$
- ▶ Bayesian cho phép suy luận một cách trực tiếp xác suất của một giả thuyết (không cần qua giả thuyết đảo của frequentist)

# Tác động của cỡ mẫu lên Posterior

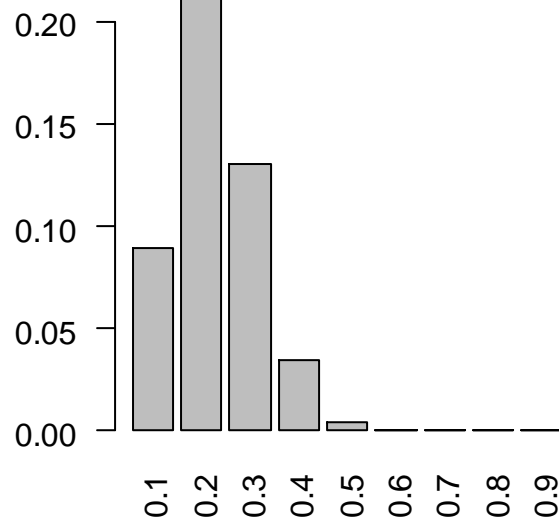


$k = 4, n = 20$

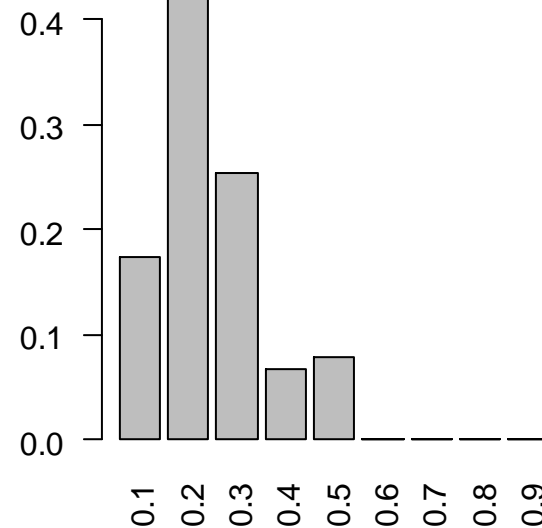
**Prior**



**Likelihood**

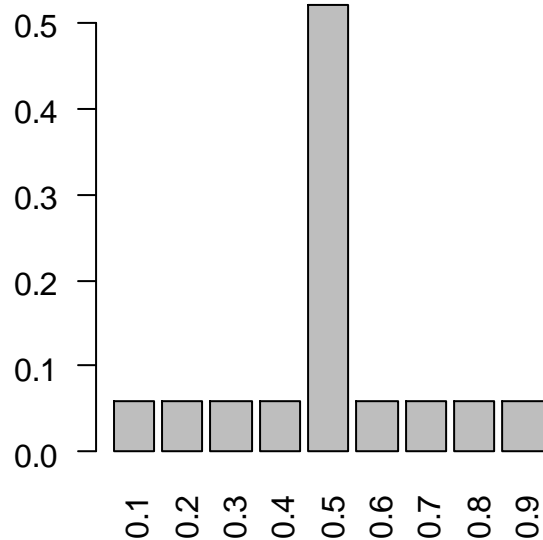


**Posterior**

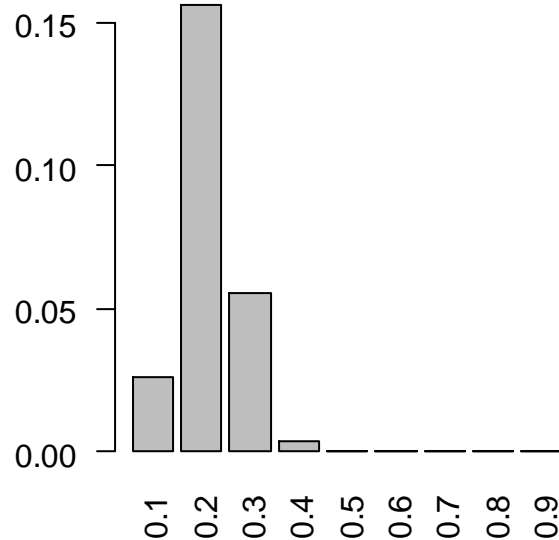


$k = 8, n = 40$

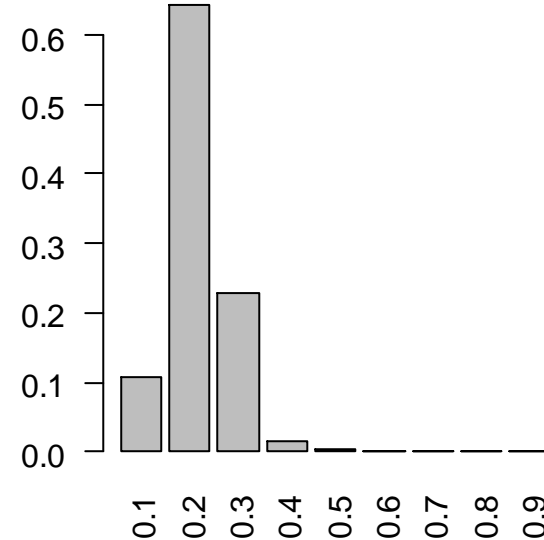
**Prior**



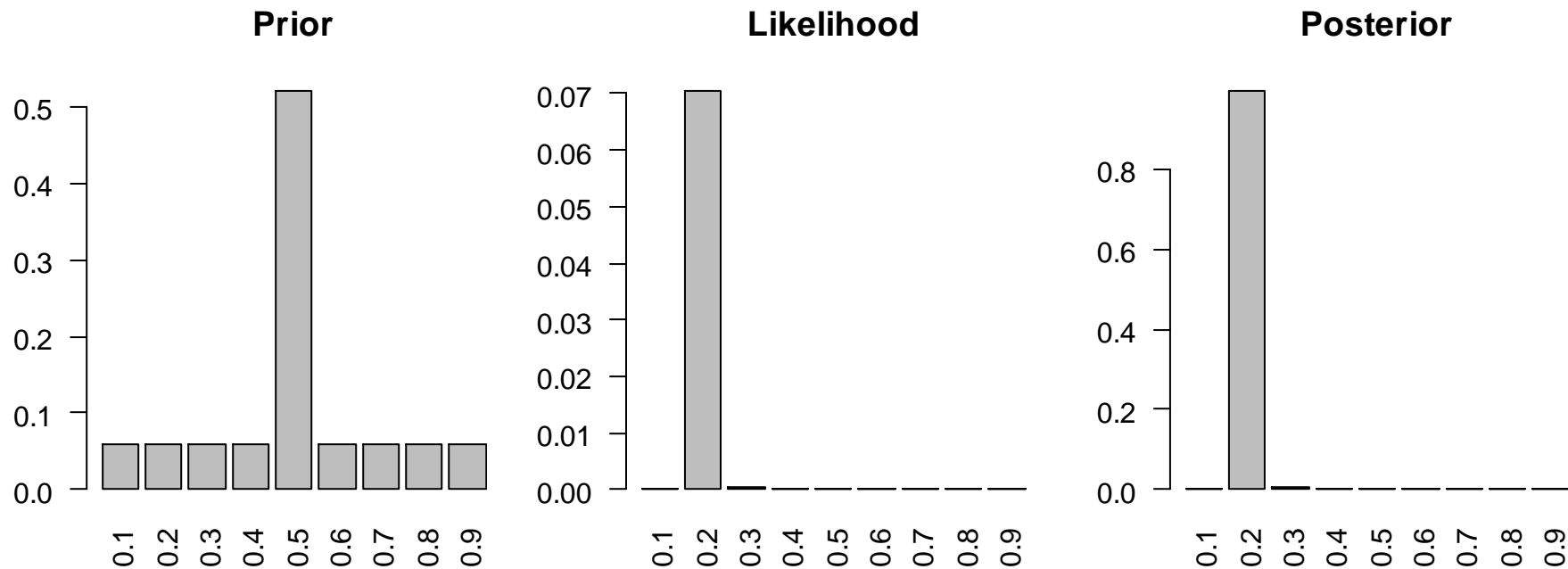
**Likelihood**



**Posterior**



$k = 40, n = 200$

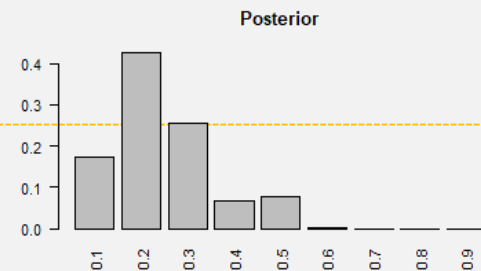
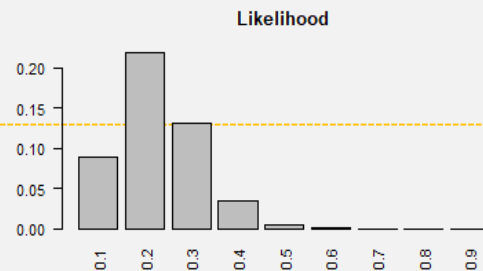
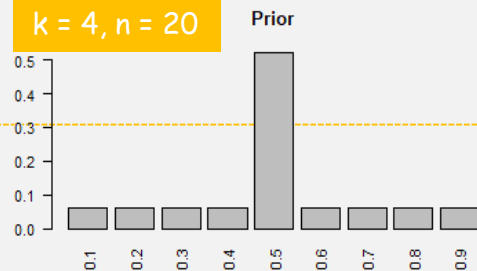


# Khi tăng cỡ mẫu

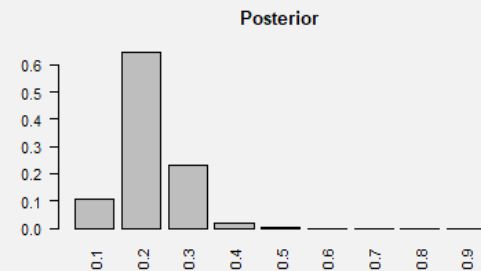
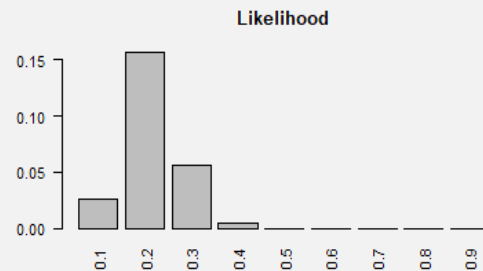
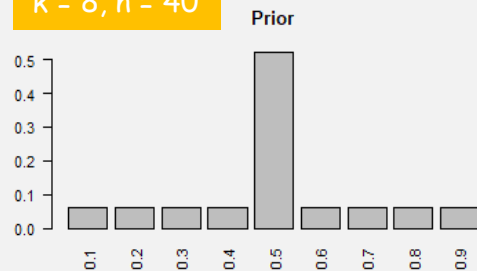
---

- ▶ Mức độ **phân tán** của posterior (“**uncertainty**”) giảm → bằng chứng mạnh hơn
- ▶ Likelihood (data) càng “lấn át” Prior. Ảnh hưởng của Prior lên Posterior giảm khi thu thập thêm dữ liệu.
- ▶ Khi cỡ mẫu đạt  $\infty$  (long-run) → suy luận phụ thuộc vào likelihood của dữ liệu

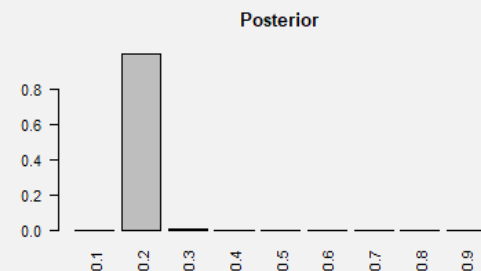
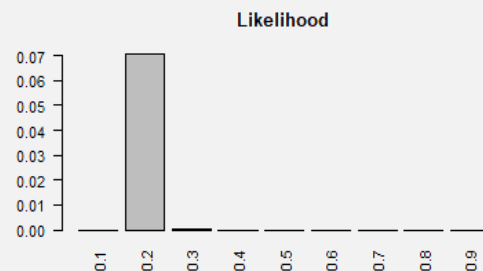
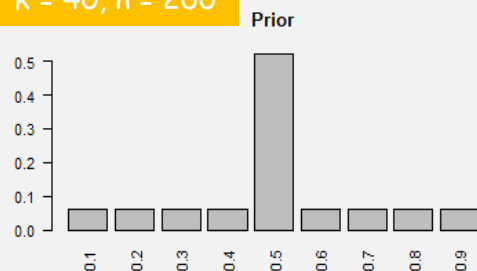
$k = 4, n = 20$



$k = 8, n = 40$



$k = 40, n = 200$



# Tóm tắt

---

- ▶ Sử dụng suy luận Bayes cho so sánh tỉ lệ
- ▶ Bayesian cho phép suy luận một cách **trực tiếp** xác suất của một giả thuyết (không cần qua giả thuyết đảo của frequentist)
- ▶ Ảnh hưởng của **Likelihood** và **Prior** lên **Posterior** khi thay đổi cỡ mẫu

**Thank you !**