Bayesian Inference for a Proportion

Jingchen (Monika) Hu

Vassar College

MATH 347 Bayesian Statistics

Outline

Outline

Tokyo Express customers' dining preference

- Tokyo Express is a popular Japanese restaurant on Lagrange Ave (10-15 minutes walk from Vassar).

Tokyo Express customers' dining preference

- Tokyo Express is a popular Japanese restaurant on Lagrange Ave (10-15 minutes walk from Vassar).
- Suppose the restaurant owner wants to improve the business even more, especially for dinner.

Tokyo Express customers' dining preference

- Tokyo Express is a popular Japanese restaurant on Lagrange Ave (10-15 minutes walk from Vassar).
- Suppose the restaurant owner wants to improve the business even more, especially for dinner.
- The owner plans to conduct a survey by asking their customers: "what is your favorite day to eat out for dinner?"
- The owner wants to find out how popular is choice of Friday.

Outline

Bayesian inference with discrete priors

- Three general steps of Bayesian inference:
  - Step 1: express an opinion about the location of the proportion p before sampling (prior).
  - Step 2: take the sample and record the observed proportion of preferring Friday (data/likelihood).
  - Step 3: use Bayes' rule to sharpen and update the previous opinion about p given the information from the sample (posterior).

Bayesian inference with discrete priors

- Three general steps of Bayesian inference:
  - Step 1: express an opinion about the location of the proportion p before sampling (prior).
  - Step 2: take the sample and record the observed proportion of preferring Friday (data/likelihood).
  - Step 3: use Bayes' rule to sharpen and update the previous opinion about p given the information from the sample (posterior).
- Bayesian inference with discrete priors:
  - Step 1 on priors: list the finite number of possible values for the proportion p, and assign probability to each value.

Bayesian inference with discrete priors

- Three general steps of Bayesian inference:
  - Step 1: express an opinion about the location of the proportion p before sampling (prior).
  - Step 2: take the sample and record the observed proportion of preferring Friday (data/likelihood).
  - Step 3: use Bayes' rule to sharpen and update the previous opinion about p given the information from the sample (posterior).
- Bayesian inference with discrete priors:
  - Step 1 on priors: list the finite number of possible values for the proportion p, and assign probability to each value.
  - Step 2 on data/likelihood: Binomial distribution.

Bayesian inference with discrete priors

- Three general steps of Bayesian inference:
  - Step 1: express an opinion about the location of the proportion p before sampling (prior).
  - Step 2: take the sample and record the observed proportion of preferring Friday (data/likelihood).
  - Step 3: use Bayes' rule to sharpen and update the previous opinion about p given the information from the sample (posterior).
- Bayesian inference with discrete priors:
  - Step 1 on priors: list the finite number of possible values for the proportion p, and assign probability to each value.
  - Step 2 on data/likelihood: Binomial distribution.
  - Step 3 on posterior: use the discrete version of the Bayes' rule (summation Σ) for to sharpen and update the probability of each specified possible values of p.

Step 1: Prior distribution

- Consider the percentage of customers' choice is Friday, p.
- Before giving out the survey, let's consider:
  - the possible value(s) of proportion p;
  - the probability associate with each value of p.

Step 1: Prior distribution

- Consider the percentage of customers' choice is Friday, p.
- Before giving out the survey, let's consider:
  - the possible value(s) of proportion p;
  - the probability associate with each value of p.
- Suppose p can take 6 possible values

$$p = \{0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}. \tag{1}$$

Step 1: Prior distribution

- Consider the percentage of customers' choice is Friday, p.
- Before giving out the survey, let's consider:
  - the possible value(s) of proportion p;
  - the probability associate with each value of p.
- Suppose p can take 6 possible values

$$p = \{0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}. \tag{1}$$

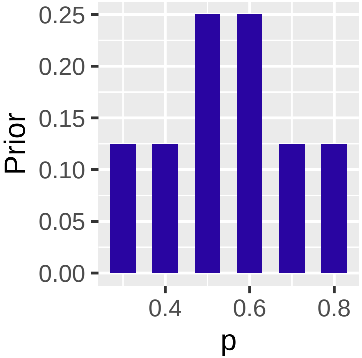
- Further suppose Tokyo Express owner believes that some values are more likely than the others, specifically, a prior distribution:

$$\pi_{owner}(p) = (0.125, 0.125, 0.250, 0.250, 0.125, 0.125). \tag{2}$$

- Exercise: Is the prior distribution $\pi_{owner}(p)$ reasonable? (Hint: 3 axioms of probability)

Using R/RStudio to express and plot the prior $\pi_{owner}(p)$

```
bayes_table <- data.frame(p = seq(.3, .8, by=.1),
                          Prior = c(0.125, 0.125, 0.250,
                                    0.250, 0.125, 0.125))
ggplot(data=bayes_table, aes(x=p, y=Prior)) +
  geom_bar(stat="identity", fill=crcblue, width = 0.06)
```



Step 2: Data/likelihood of proportion p

- Now the Tokyo Express owner gives a survey to 20 customers.
- Out of the 20 responses, 12 say that their favorite day for eating out for dinner is Friday.

Step 2: Data/likelihood of proportion p

- Now the Tokyo Express owner gives a survey to 20 customers.
- Out of the 20 responses, 12 say that their favorite day for eating out for dinner is Friday.
- Quantity of interest: p, the proportion of customers prefer eating out for dinner on Friday.

Step 2: Data/likelihood of proportion p

- Now the Tokyo Express owner gives a survey to 20 customers.
- Out of the 20 responses, 12 say that their favorite day for eating out for dinner is Friday.
- Quantity of interest: p, the proportion of customers prefer eating out for dinner on Friday.
- The data/likelihood is a function of the quantity of interest.
- What would be the function of 12 out of 20 preferring Friday, in terms of the proportion p?

The Binomial distribution

- A Binomial experiment:
  1. One is repeating the same basic task or trial many times – let the number of trials be denoted by n.
  2. On each trial, there are two possible outcomes that are called `success" or` failure".
  3. The probability of a success, denoted by p, is the same for each trial.
  4. The results of outcomes from different trials are independent.
- Do you think the survey is a Binomial experiment?

The Binomial distribution

- A Binomial experiment:
  1. One is repeating the same basic task or trial many times – let the number of trials be denoted by n.
  2. On each trial, there are two possible outcomes that are called `success"` `orfailure"`.
  3. The probability of a success, denoted by p, is the same for each trial.
  4. The results of outcomes from different trials are independent.
- Do you think the survey is a Binomial experiment?
- The probability of y successes in a Binomial experiment is given by

$$P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}, y = 0, \cdots, n, \tag{3}$$

  where n is the number of trials and p is the success probability.

The likelihood function

- The probability of y successes in a binomial experiment is given by
$$P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}, y = 0, \cdots, n, \tag{4}$$
where n is the number of trials and p is the success probability.

The likelihood function

- The probability of y successes in a binomial experiment is given by
$$P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}, y = 0, \cdots, n, \tag{4}$$
where n is the number of trials and p is the success probability.
- The likelihood is the chance of 12 successes in 20 trials viewed as a function of the probability of success is p:
$$\text{Likelihood} = L(p) = \binom{20}{12} p^{12}(1-p)^8. \tag{5}$$
    - L is a function of p.
    - n is fixed and known.
    - Y is the random variable.
    - p is the quantity of interest, also the unknown parameter in the Binomial distribution.

Use R/RStudio to compute the likelihood function

$$\text{Likelihood} = L(p) = \binom{20}{12} p^{12}(1-p)^8 \tag{6}$$

- Need: sample size n (20), number of successes k (12), and possible values of proportion p ($\{0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}$).
- Do not need: the assigned probabilities (0.125, 0.125, 0.250, 0.250, 0.125, 0.125) in the prior distribution $\pi_{owner}(p)$.

Use R/RStudio to compute the likelihood function

```
bayes_table$Likelihood <- dbinom(12, size=20,
                                 prob=bayes_table$p)
bayes_table
```

```
##     p Prior  Likelihood
## 1 0.3 0.125 0.003859282
## 2 0.4 0.125 0.035497440
## 3 0.5 0.250 0.120134354
## 4 0.6 0.250 0.179705788
## 5 0.7 0.125 0.114396740
## 6 0.8 0.125 0.022160877
```

Step 3: Posterior distribution

- Notations:
  - $\pi(p)$ the prior distribution of p.
  - $L(p)$ is the likelihood function.
  - $\pi(p \mid y)$ the posterior distribution of p after observing the number of successes y.

Step 3: Posterior distribution

- Notations:
  - $\pi(p)$ the prior distribution of p.
  - $L(p)$ is the likelihood function.
  - $\pi(p \mid y)$ the posterior distribution of p after observing the number of successes y.
- The Bayes' rule for a discrete parameter has the form

$$\pi(p_i \mid y) = \frac{\pi(p_i) \times L(p_i)}{\sum_j \pi(p_j) \times L(p_j)}, \tag{7}$$

  - $\pi(p_i)$ the prior probability of $p = p_i$.
  - $L(p_i)$ the likelihood function evaluated at $p = p_i$.
  - $\pi(p_i \mid y)$ the posterior probability of $p = p_i$ given the number of successes y.
  - the denominator gives the marginal distribution of the observation y (by the Law of Total Probability).

Use R/RStudio to compute and plot the posterior

```
bayesian_crank(bayes_table) -> bayes_table
bayes_table
```

```
##     p Prior  Likelihood      Product   Posterior
## 1 0.3 0.125 0.003859282 0.0004824102 0.004975901
## 2 0.4 0.125 0.035497440 0.0044371799 0.045768032
## 3 0.5 0.250 0.120134354 0.0300335884 0.309786454
## 4 0.6 0.250 0.179705788 0.0449264469 0.463401326
## 5 0.7 0.125 0.114396740 0.0142995925 0.147495530
## 6 0.8 0.125 0.022160877 0.0027701096 0.028572757
```

- Inference question: What is the posterior probability that over half of the customers prefer eating out on Friday for dinner?

Use R/RStudio to compute and plot the posterior

```
bayesian_crank(bayes_table) -> bayes_table
bayes_table
```

```
##     p Prior  Likelihood      Product   Posterior
## 1 0.3 0.125 0.003859282 0.0004824102 0.004975901
## 2 0.4 0.125 0.035497440 0.0044371799 0.045768032
## 3 0.5 0.250 0.120134354 0.0300335884 0.309786454
## 4 0.6 0.250 0.179705788 0.0449264469 0.463401326
## 5 0.7 0.125 0.114396740 0.0142995925 0.147495530
## 6 0.8 0.125 0.022160877 0.0027701096 0.028572757
```

- Inference question: What is the posterior probability that over half of the customers prefer eating out on Friday for dinner?

$$Prob(p > 0.5) = 0.463 + 0.147 + 0.029 = 0.639. \quad (8)$$

Use R/RStudio to compute and plot the posterior

```
bayesian_crank(bayes_table) -> bayes_table
sum(bayes_table$Posterior[bayes_table$p > 0.5])
```

```
## [1] 0.6394696
```

Use R/RStudio to compute and plot the posterior

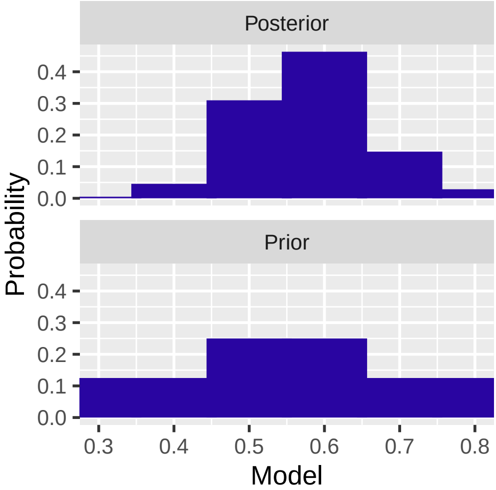
```
bayesian_crank(bayes_table) -> bayes_table
sum(bayes_table$Posterior[bayes_table$p > 0.5])
```

```
## [1] 0.6394696
```

- Exercise: What is the posterior probability that less than 40% of the customers prefer eating out on Friday for dinner?

Using R/RStudio to compute and plot the posterior

```
prior_post_plot(bayes_table, Color = crcblue) +
  theme(text=element_text(size=10))
```



Outline

Step 1: Prior distribution

- Before giving out the survey, we need to specify a prior distribution for unknown parameter p.
- Previously, p can take 6 possible values
$$p = \{0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}. \tag{9}$$
- And a discrete prior distribution:
$$\pi_{owner}(p) = (0.125, 0.125, 0.250, 0.250, 0.125, 0.125). \tag{10}$$

Step 1: Prior distribution

```
bayes_table
```

```
##     p Prior  Likelihood      Product   Posterior
## 1 0.3 0.125 0.003859282 0.0004824102 0.004975901
## 2 0.4 0.125 0.035497440 0.0044371799 0.045768032
## 3 0.5 0.250 0.120134354 0.0300335884 0.309786454
## 4 0.6 0.250 0.179705788 0.0449264469 0.463401326
## 5 0.7 0.125 0.114396740 0.0142995925 0.147495530
## 6 0.8 0.125 0.022160877 0.0027701096 0.028572757
```

- Anything unsatisfactory?

Move to continuous priors

- A limitation of specifying a discrete prior for p
  - If a plausible value is not specified in the prior distribution (e.g. $p = 0.2$), it will be assigned a 0 probability in the posterior distribution.

Move to continuous priors

- A limitation of specifying a discrete prior for p
  - If a plausible value is not specified in the prior distribution (e.g. $p = 0.2$), it will be assigned a 0 probability in the posterior distribution.
- Ideally, we want a distribution that allows p to be any value in [0, 1].
- The continuous Uniform distribution:
  - Any value of p is equally likely.
  - The probability density function of the continuous Uniform on the interval $[a, b]$ is

$$\pi(p) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq p \leq b, \\ 0 & \text{for } p < a \text{ or } p > b. \end{cases} \tag{11}$$

  - $p \sim \mathrm{Uniform}(0, 1)$: a very special case of p.
- The Beta distribution!

The Beta distribution

- Notation: $\mathrm{Beta}(a, b)$.
- For a random variable falling between 0 and 1, suitable for proportion p.
- Beta distribution has two shape parameters a and b.
- Probability density function (pdf) is:
$$\pi(p) = \frac{1}{B(a, b)} p^{a-1}(1-p)^{b-1}, \ 0 \leq p \leq 1. \tag{12}$$
  - $B(a, b)$ is the Beta function $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.
  - Γ is the Gamma function.

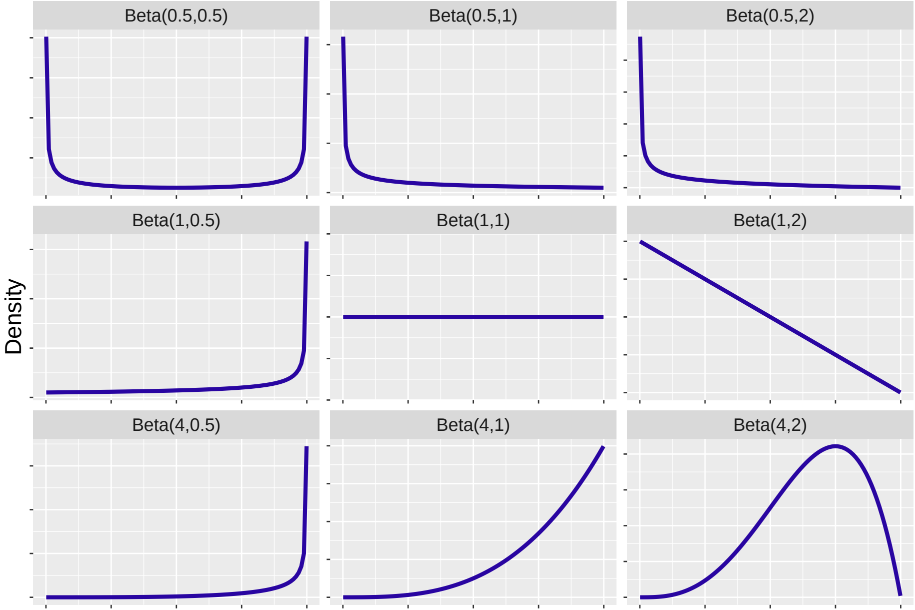
The Beta distribution

- Notation: $\mathrm{Beta}(a, b)$.
- For a random variable falling between 0 and 1, suitable for proportion p.
- Beta distribution has two shape parameters a and b.
- Probability density function (pdf) is:

$$\pi(p) = \frac{1}{B(a, b)} p^{a-1}(1-p)^{b-1},\ 0 \leq p \leq 1. \tag{12}$$

  - $B(a, b)$ is the Beta function $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.
  - Γ is the Gamma function.
  - Continuous Uniform on [0, 1] is a special case of Beta with $a = b = 1$: $\mathrm{Uniform}(0, 1) = \mathrm{Beta}(1, 1)$

Examples of Beta curves



Choose a Beta curve to represent prior opinion

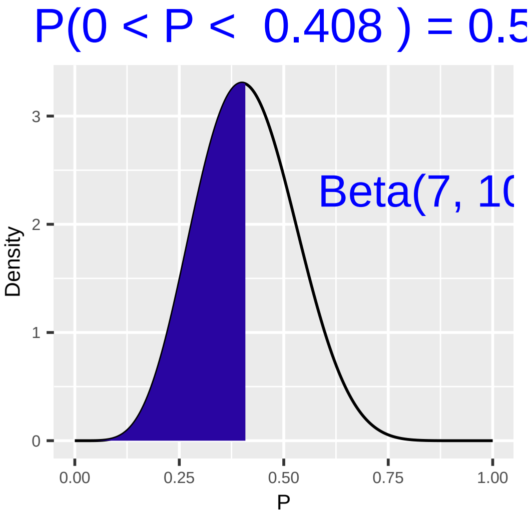
- Prior opinion: values of p and associated probabilities.
- Difficult to guess values of a and b in $\mathrm{Beta}(a, b)$.
- Solution: specifying a Beta prior by specification of quantiles of the distribution.

Choose a Beta curve to represent prior opinion

- Prior opinion: values of p and associated probabilities.
- Difficult to guess values of a and b in $\mathrm{Beta}(a, b)$.
- Solution: specifying a Beta prior by specification of quantiles of the distribution.
    - Quantiles are about rank order of values.
    - e.g. Middle quantile/50-th percentile: median.
    - `beta_quantile()` function in R: inputs a probability measure p and outputs the value of x such that $Prob(X <= x) = p$. (e.g. $x = 0.408$ when $p = 0.5$ for $\mathrm{Beta}(7, 10)$)

Choose a Beta curve to represent prior opinion

```
beta_quantile(0.5, c(7, 10), Color = crcblue) +
  theme(text=element_text(size=8))
```



Use `beta.select()` to choose a Beta curve

- The `beta.select()` function needs us to
  - first think about specifying two quantiles
  - then it finds the Beta curve that matches with these quantiles

Use beta.select() to choose a Beta curve

- The beta.select() function needs us to
  - first think about specifying two quantiles
  - then it finds the Beta curve that matches with these quantiles
- Example:
  - Suppose we believe that $p_{50} = 0.55$ (50-th quantile)
  - Suppose we also believe that $p_{90} = 0.8$ (90-th quantile)
  - Input these two sets of values into beta.select()

```
beta.select(list(x = 0.55, p = 0.5),
            list(x = 0.80, p = 0.9))
```

```
## [1] 3.06 2.56
```

Use beta.select() to choose a Beta curve

- The `beta.select()` function needs us to
  - first think about specifying two quantiles
  - then it finds the Beta curve that matches with these quantiles
- Example:
  - Suppose we believe that $p_{50} = 0.55$ (50-th quantile)
  - Suppose we also believe that $p_{90} = 0.8$ (90-th quantile)
  - Input these two sets of values into `beta.select()`

```
beta.select(list(x = 0.55, p = 0.5),
            list(x = 0.80, p = 0.9))
```

```
## [1] 3.06 2.56
```

- Exercise 1: Verify that $\mathrm{Beta}(3.06, 2.56)$ has $p_{50} = 0.55$ and $p_{90} = 0.8$. (Hint: use the `beta_quantile()` function)
- Exercise 2: Come up with your own Beta prior distribution, and share it with your neighbors.

Outline

Step 2: Data/likelihood of proportion p

- Recall that
  - The Tokyo Express owner gives a survey to 20 customers, and 12 respond that their favorite day is Friday.
  - The data/likelihood is a function of the quantity of interest, p.
  - It is a Binomial experiment, and

$$P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}, y = 0, \cdots, n, \tag{13}$$

where n is the number of trials and p is the success probability.

Step 2: Data/likelihood of proportion p

- Recall that
  - The Tokyo Express owner gives a survey to 20 customers, and 12 respond that their favorite day is Friday.
  - The data/likelihood is a function of the quantity of interest, p.
  - It is a Binomial experiment, and

$$P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}, y = 0, \cdots, n, \tag{13}$$

    where n is the number of trials and p is the success probability.
- Exercise: Write out the likelihood function for $n = 20$ and $y = 12$.

Step 2: Data/likelihood of proportion p

- Recall that
  - The Tokyo Express owner gives a survey to 20 customers, and 12 respond that their favorite day is Friday.
  - The data/likelihood is a function of the quantity of interest, p.
  - It is a Binomial experiment, and

$$P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}, y = 0, \cdots, n, \tag{13}$$

    where n is the number of trials and p is the success probability.
- Exercise: Write out the likelihood function for $n = 20$ and $y = 12$.
- Solution:

$$\text{Likelihood} = L(p) = \binom{20}{12} p^{12} (1-p)^8. \tag{14}$$

Bayes' rule for continuous priors

- The likelihood function is the same, regardless of the prior distribution.
- Recall: the Bayes' rule for a discrete parameter has the form
$$\pi(p_i \mid y) = \frac{\pi(p_i) \times L(p_i)}{\sum_j \pi(p_j) \times L(p_j)} \tag{15}$$
- What about for continuous p? Unfortunately we can not list each value of p anymore.

Bayes' rule for continuous priors

- The likelihood function is the same, regardless of the prior distribution.
- Recall: the Bayes' rule for a discrete parameter has the form

$$\pi(p_i \mid y) = \frac{\pi(p_i) \times L(p_i)}{\sum_j \pi(p_j) \times L(p_j)} \tag{15}$$

- What about for continuous p? Unfortunately we can not list each value of p anymore.
- Solution: With continuous p, the denominator changes from summation Σ to integration $\int$.
- Since $\int \pi(p) \times L(p)dp = f(y)$ is fixed, we can write the Bayes' rule for continuous p in proportional sign:

$$\pi(p \mid y) \propto \pi(p) \times L(p). \tag{16}$$

Step 3: Derive the posterior

$$\pi(p \mid y) \propto \pi(p) \times L(p) \tag{17}$$

- For prior $\pi(p)$, we have $p \sim \mathrm{Beta}(3.06, 2.56)$.
- For data/likelihood $L(p)$, we have $Y \sim \mathrm{Binomial}(20, p)$.
- We need to derive $\pi(p \mid y)$.

Step 3: Derive the posterior

$$\pi(p \mid y) \propto \pi(p) \times L(p) \tag{17}$$

- For prior $\pi(p)$, we have $p \sim \mathrm{Beta}(3.06, 2.56)$.
- For data/likelihood $L(p)$, we have $Y \sim \mathrm{Binomial}(20, p)$.
- We need to derive $\pi(p \mid y)$.
- Exercise: Derive $\pi(p \mid y)$ with the setup below.
  - The prior distribution:

$$\pi(p) = \frac{1}{B(3.06, 2.56)} p^{3.06-1}(1-p)^{2.56-1}.$$

  - The likelihood:

$$f(Y = 12 \mid p) = L(p) = \binom{20}{12} p^{12}(1-p)^{8}.$$

  - The posterior?

Step 3: Derive the posterior cont'd

- The posterior:

$$\pi(p \mid Y = 12) \quad \propto \quad \pi(p) \times f(Y = 12 \mid p)$$

Step 3: Derive the posterior cont'd

- The posterior:

$$
\begin{aligned}
\pi(p \mid Y = 12) &\propto \pi(p) \times f(Y = 12 \mid p) \\
&= \frac{1}{B(3.06, 2.56)} p^{3.06-1}(1-p)^{2.56-1} \times \\
&\quad \binom{20}{12} p^{12}(1-p)^{8}
\end{aligned}
$$

Step 3: Derive the posterior cont'd

- The posterior:

$$
\begin{aligned}
\pi(p \mid Y = 12) &\propto \pi(p) \times f(Y = 12 \mid p) \\
&= \frac{1}{B(3.06, 2.56)} p^{3.06-1}(1-p)^{2.56-1} \times \\
&\quad \binom{20}{12} p^{12}(1-p)^{8} \\
\texttt{[drop the constants]} &\propto p^{12}(1-p)^{8} p^{3.06-1}(1-p)^{2.56-1} \\
\texttt{[combine the powers]} &= p^{15.06-1}(1-p)^{10.56-1}. \qquad (18)
\end{aligned}
$$

- That is,

$$\pi(p \mid Y = 12) \propto p^{15.06-1}(1-p)^{10.56-1},$$

which means

Step 3: Derive the posterior cont'd

- The posterior:

$$\begin{aligned}
\pi(p \mid Y=12) &\propto \pi(p) \times f(Y=12 \mid p) \\
&= \frac{1}{B(3.06, 2.56)} p^{3.06-1}(1-p)^{2.56-1} \times \\
&\quad \binom{20}{12} p^{12}(1-p)^{8} \\
\texttt{[drop the constants]} &\propto p^{12}(1-p)^{8} p^{3.06-1}(1-p)^{2.56-1} \\
\texttt{[combine the powers]} &= p^{15.06-1}(1-p)^{10.56-1}. \qquad (18)
\end{aligned}$$

- That is,

$$\pi(p \mid Y=12) \propto p^{15.06-1}(1-p)^{10.56-1},$$

which means

$$p \mid Y=12 \sim \mathrm{Beta}(15.06, 10.56).$$

From Beta prior to Beta posterior: conjugate priors

- The prior distribution:
$$p \sim \mathrm{Beta}(a, b)$$
- The sampling density:
$$Y \sim \mathrm{Binomial}(n, p)$$
- The posterior distribution:
$$p \mid Y = y \sim \mathrm{Beta}(a + y, b + n - y)$$
- Conjugate priors: from Beta prior to Beta posterior.

From Beta prior to Beta posterior: conjugate priors

- The prior distribution:
$$p \sim \mathrm{Beta}(a, b)$$
- The sampling density:
$$Y \sim \mathrm{Binomial}(n, p)$$
- The posterior distribution:
$$p \mid Y = y \sim \mathrm{Beta}(a + y, b + n - y)$$
- Conjugate priors: from Beta prior to Beta posterior.

Source	Successes	Failures
Prior	a	b
Data/Likelihood	y	$n - y$
Posterior	$a + y$	$b + n - y$

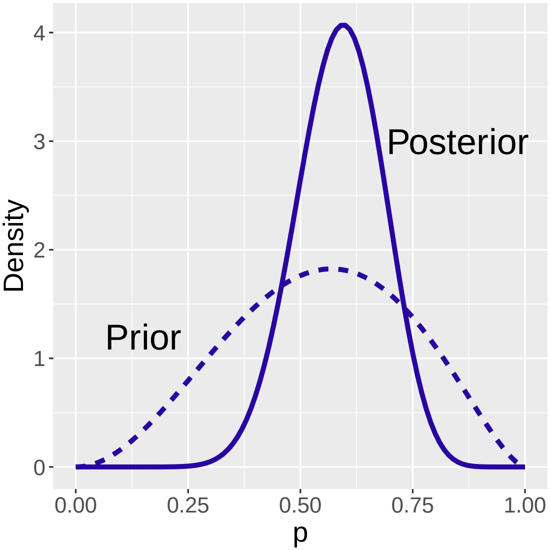
Use R/RStudio to compute and plot the posterior

$$p \mid Y = y \sim \mathrm{Beta}(a + y, b + n - y)$$

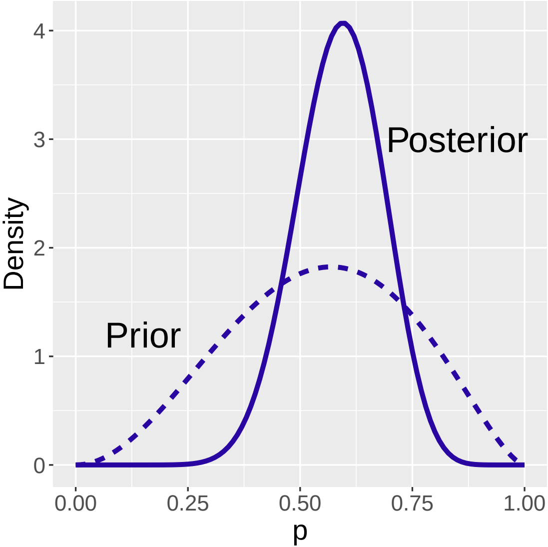
```
ab <- c(3.06, 2.56)
yny <- c(12, 8)
(ab_new <- ab + yny)
```

```
## [1] 15.06 10.56
```

Use R/RStudio to compute and plot the posterior cont'd



Use R/RStudio to compute and plot the posterior cont'd



- Exercise 1: Compare prior mean 0.544 and posterior mean 0.588. (Recall that sample mean is 0.6.)
- Exercise 2: Compare the spreads of the two curves.

Outline

Bayesian hypothesis testing

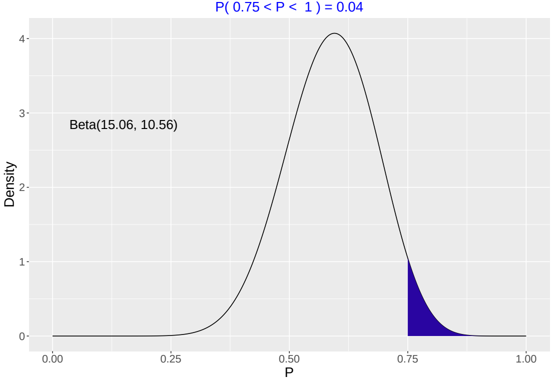
- Suppose one of the Tokyo Express's workers claims that at least 75% of the customers prefer Friday. Is this a reasonable claim?

Bayesian hypothesis testing

- Suppose one of the Tokyo Express's workers claims that at least 75% of the customers prefer Friday. Is this a reasonable claim?
- From a Bayesian viewpoint,
  - Find the posterior probability that $p >= 0.75$.
  - Make a decision based on the value of the posterior probability.
  - If the probability is small, we can reject this claim.

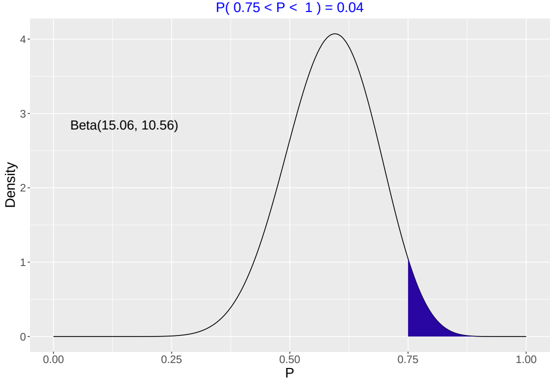
Bayesian hypothesis testing cont'd

```
beta_area(lo = 0.75, hi = 1.0, shape_par = c(15.06, 10.56),
          Color = crcblue) +
  theme(text=element_text(size=18))
```



Bayesian hypothesis testing cont'd

```
beta_area(lo = 0.75, hi = 1.0, shape_par = c(15.06, 10.56),
          Color = crcblue) +
  theme(text=element_text(size=18))
```



- Posterior probability is only 4%, reject the claim.
- Exercise: What about a claim "at most 30% of the customers prefer Friday"?

Bayesian hypothesis testing cont'd

When the posterior distribution is known...

Bayesian hypothesis testing cont'd

When the posterior distribution is known...

- Exact solution: use the `beta_area()` function in the `ProbBayes` package and/or the `pbeta()` R function

```
beta_area(lo = 0.75, hi = 1.0, shape_par = c(15.06, 10.56))
```

```
pbeta(1, 15.06, 10.56) - pbeta(0.75, 15.06, 10.56)
```

```
## [1] 0.03973022
```

Bayesian hypothesis testing cont'd

When the posterior distribution is known...

- Exact solution: use the `beta_area()` function in the `ProbBayes` package and/or the `pbeta()` R function

```
beta_area(lo = 0.75, hi = 1.0, shape_par = c(15.06, 10.56))
```

```
pbeta(1, 15.06, 10.56) - pbeta(0.75, 15.06, 10.56)
```

```
## [1] 0.03973022
```

- Approximation through Monte Carlo simulation using the `rbeta()` R function

```
S <- 1000
BetaSamples <- rbeta(S, 15.06, 10.56)
sum(BetaSamples >= 0.75)/S
```

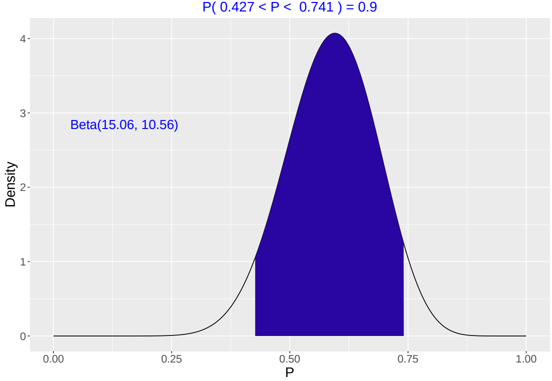
```
## [1] 0.037
```

Bayesian credible intervals

- Consider an interval that we are confident contains p.
- Such an interval provides an uncertainty estimate for our parameter p.
- A 90% Bayesian credible interval is an interval contains 90% of the posterior probability.

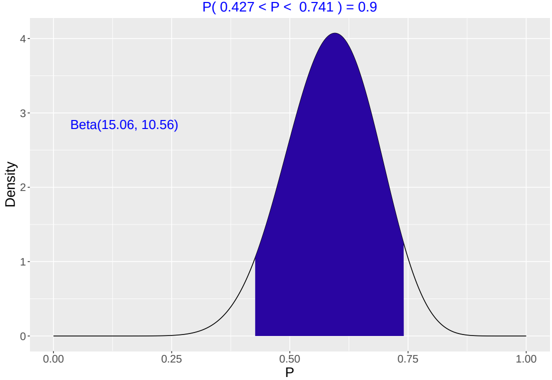
Bayesian credible intervals cont'd

```
beta_interval(0.9, c(15.06, 10.56), Color = crcblue) +
  theme(text=element_text(size=18))
```



Bayesian credible intervals cont'd

```
beta_interval(0.9, c(15.06, 10.56), Color = crcblue) +
  theme(text=element_text(size=18))
```



- The middle 90%.
- Interpretation: different from a confidence interval.
- Exercise: Construct a middle 98% credible interval for p.

Bayesian credible intervals cont'd

When the posterior distribution is known...

Bayesian credible intervals cont'd

When the posterior distribution is known…

- Exact solution: use the `beta_interval()` function in the `ProbBayes` package (only the middle 90%) and or the `qbeta()` R function (not necessarily the middle 90%)

```
beta_interval(0.9, c(15.06, 10.56), Color = crcblue)
```

```
c(qbeta(0.05, 15.06, 10.56), qbeta(0.95, 15.06, 10.56))
```

```
## [1] 0.4266788 0.7410141
```

Bayesian credible intervals cont'd

When the posterior distribution is known...

- Exact solution: use the `beta_interval()` function in the `ProbBayes` package (only the middle 90%) and or the `qbeta()` R function (not necessarily the middle 90%)

```
beta_interval(0.9, c(15.06, 10.56), Color = crcblue)
```

```
c(qbeta(0.05, 15.06, 10.56), qbeta(0.95, 15.06, 10.56))
```

```
## [1] 0.4266788 0.7410141
```

- Approximation through Monte Carlo simulation using the `rbeta()` R function (not necessarily the middle 90%)

```
S <- 1000; BetaSamples <- rbeta(S, 15.06, 10.56)
quantile(BetaSamples, c(0.05, 0.95))
```

```
##        5%       95%
## 0.4024251 0.7425349
```

Bayesian prediction

- Suppose the Tokyo Express owner gives out another survey to m customers, how many would prefer Friday?
- The predictive distribution: $\tilde{Y} \mid Y = y$.

Bayesian prediction

- Suppose the Tokyo Express owner gives out another survey to m customers, how many would prefer Friday?
- The predictive distribution: $\tilde{Y} \mid Y = y$.
  - The exact prediction:

$$\tilde{Y} \mid Y = y \sim \text{Beta-Binomial}(m, a + y, b + n - y).$$

  - Prediction through simulation (our focus):

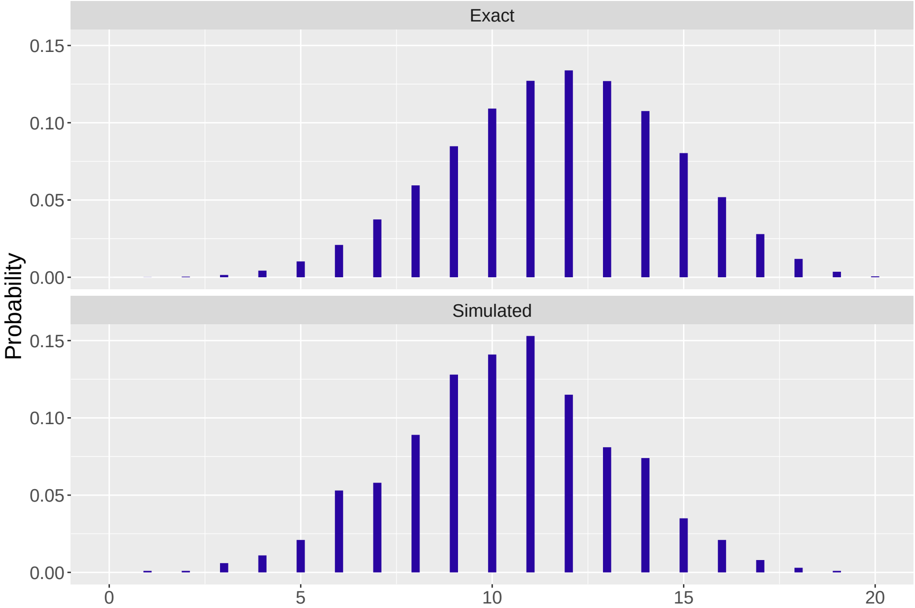
$$\text{sample } p \sim \text{Beta}(a + y, b + n - y) \quad \rightarrow \quad \text{sample } \tilde{y} \sim \text{Binomial}(m, p)$$

Use R/RStudio to make Bayesian predictions

```r
S <- 1000
a <- 3.06; b <- 2.56
n <- 20; y <- 12
m <- 20
pred_p_sim <- rbeta(S, a + y, b + n - y)
pred_y_sim <- rbinom(S, m, pred_p_sim)
sum(pred_y_sim >=5 & pred_y_sim <= 15)/S
```

```
## [1] 0.902
```

Use R/RStudio to make Bayesian predictions



Bayesian predictive checking

- Bayesian prediction is posterior predictive

$$\text{sample } p \sim \mathrm{Beta}(a+y, b+n-y) \quad \rightarrow \quad \text{sample } \tilde{y} \sim \mathrm{Binomial}(m, p)$$

- We can perform posterior predictive checking through simulation

$$\begin{aligned}
\text{sample } p^{(1)} \sim \mathrm{Beta}(a+y, b+n-y) \quad &\rightarrow \quad \text{sample } \tilde{y}^{(1)} \sim \mathrm{Binomial}(n, p^{(1)}) \\
\text{sample } p^{(2)} \sim \mathrm{Beta}(a+y, b+n-y) \quad &\rightarrow \quad \text{sample } \tilde{y}^{(2)} \sim \mathrm{Binomial}(n, p^{(2)}) \\
&\vdots \\
\text{sample } p^{(S)} \sim \mathrm{Beta}(a+y, b+n-y) \quad &\rightarrow \quad \text{sample } \tilde{y}^{(S)} \sim \mathrm{Binomial}(n, p^{(S)})
\end{aligned}$$

The sample $\{\tilde{y}^{(1)}, ..., \tilde{y}^{(S)}\}$ is an approximation to the posterior predictive distribution that can be used for model checking.

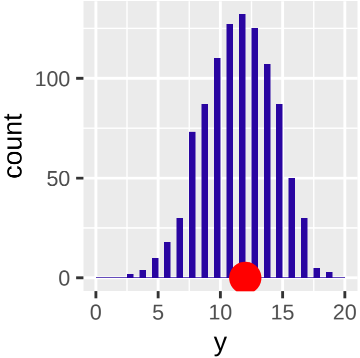
Use R/Rstudio to perform posterior predictive checking

```r
S <- 1000
a <- 3.06; b <- 2.56
n <- 20; y <- 12
newy = as.data.frame(rep(NA, S))
names(newy) = c("y")

set.seed(123)
for (s in 1:S){
  pred_p_sim <- rbeta(1, a + y, b + n - y)
  pred_y_sim <- rbinom(1, n, pred_p_sim)
  newy[s,] = pred_y_sim
}
```

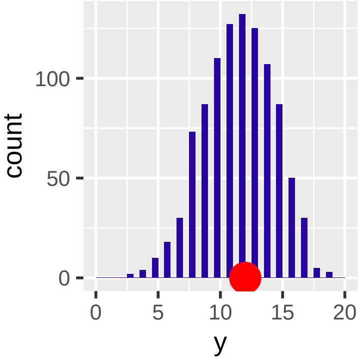
Use R/Rstudio to perform posterior predictive checking cont'd

```
ggplot(data=newy, aes(newy$y)) +
  geom_histogram(breaks=seq(0, 20, by=0.5), fill = crcblue) +
  annotate("point", x = 12, y = 0, colour = "red", size = 5) +
  xlab("y") + theme(text=element_text(size=10))
```



Use R/Rstudio to perform posterior predictive checking cont'd

```
ggplot(data=newy, aes(newy$y)) +
  geom_histogram(breaks=seq(0, 20, by=0.5), fill = crcblue) +
  annotate("point", x = 12, y = 0, colour = "red", size = 5) +
  xlab("y") + theme(text=element_text(size=10))
```



- The observed $y = 12$ is plotted as a red dot. The observed value of y is consistent with simulations of replicated data from this predictive distribution.

Use R/Rstudio to perform posterior predictive checking cont'd

- More formally, one can calculate the following probabilities:

$$Prob(y > \tilde{y} \mid y), \text{ or } 1 - Prob(y > \tilde{y} \mid y). \tag{19}$$

- If either probability is small, it suggests the model does not describe y very well.

Use R/Rstudio to perform posterior predictive checking cont'd

- More formally, one can calculate the following probabilities:

$$Prob(y > \tilde{y} \mid y), \text{ or } 1 - Prob(y > \tilde{y} \mid y). \tag{19}$$

- If either probability is small, it suggests the model does not describe y very well.

```
sum(newy > y)/S
```

```
## [1] 0.407
```

```
1 - sum(newy > y)/S
```

```
## [1] 0.593
```

- Since 0.407 and 0.593 are not small, it suggests the model describe y well. The inference passes the posterior predictive checking.

Outline

Recap

- Bayesian inference procedure:
  - Step 1: express an opinion about the location of the proportion p before sampling (prior).
  - Step 2: take the sample and record the observed proportion (data/likelihood).
  - Step 3: use Bayes' rule to sharpen and update the previous opinion about p given the information from the sample (posterior).

Recap

- Bayesian inference procedure:
  - Step 1: express an opinion about the location of the proportion p before sampling (prior).
  - Step 2: take the sample and record the observed proportion (data/likelihood).
  - Step 3: use Bayes' rule to sharpen and update the previous opinion about p given the information from the sample (posterior).
- For Binomial data/likelihood, the Beta distributions are conjugate priors.
  - The prior distribution: $p \sim \mathrm{Beta}(a, b)$.
  - The sampling density: $Y \sim \mathrm{Binomial}(n, p)$.
  - The posterior distribution: $p \mid Y = y \sim \mathrm{Beta}(a + y, b + n - y)$.

Recap

- Bayesian inference procedure:
  - Step 1: express an opinion about the location of the proportion p before sampling (prior).
  - Step 2: take the sample and record the observed proportion (data/likelihood).
  - Step 3: use Bayes' rule to sharpen and update the previous opinion about p given the information from the sample (posterior).
- For Binomial data/likelihood, the Beta distributions are conjugate priors.
  - The prior distribution: $p \sim \mathrm{Beta}(a, b)$.
  - The sampling density: $Y \sim \mathrm{Binomial}(n, p)$.
  - The posterior distribution: $p \mid Y = y \sim \mathrm{Beta}(a + y, b + n - y)$.
- Bayesian inferences (exact vs simulated)
  - Bayesian hypothesis testing & Bayesian credible intervals
  - Bayesian predictions
  - Bayesian posterior predictive checking