Introduction to Bayesian Modeling part 2

Jingchen (Monika) Hu

Vassar College

Statistical Data Privacy

Outline

1. Recap
2. Posterior predictive and synthetic data
3. Case study: bike sharing rental counts
4. Summary and References

Outline

From part 1 lecture

- The foundation of Bayesian inference
  - Prior
  - Likelihood
  - Bayes' rule (discrete and continuous) and posterior
- Markov chain Monte Carlo (MCMC)
  - Estimation
  - Diagnostics

Outline

Posterior predictive

- One important type of statistical inference is making predictions
- Bayesian methods: through the **posterior predictive distribution**

Let Y^* be the random variable for the predicted value, and y be the observed value. Let θ represent the model parameter(s). The posterior predictive distribution is

$$\pi(Y^* = y^* \mid y) = \int f(y^* \mid \theta)\pi(\theta \mid y)d\theta. \tag{1}$$

Beta-binomial example

- In the beta-binomial conjugate model, the posterior predictive distribution is analytically available

$$\begin{aligned} \pi(Y^* = y^* \mid y) &= \int_0^1 f(y^* \mid p)\pi(p \mid y)dp \\ &= \binom{n^*}{y^*} \frac{B(y^* + a + y, b + n - y + n^* - y^*)}{B(a + y, b + n - y)}, \quad (2) \end{aligned}$$

where n^* is the fixed number of trials in the prediction (usually $n^* = n$), and a and b are the parameters for the beta prior for p

Beta-binomial example

- In the beta-binomial conjugate model, the posterior predictive distribution is analytically available

$$\begin{aligned} \pi(Y^* = y^* \mid y) &= \int_0^1 f(y^* \mid p)\pi(p \mid y)dp \\ &= \binom{n^*}{y^*}\frac{B(y^* + a + y, b + n - y + n^* - y^*)}{B(a + y, b + n - y)}, \quad (2) \end{aligned}$$

where n^* is the fixed number of trials in the prediction (usually $n^* = n$), and a and b are the parameters for the beta prior for p

- This is a known distribution, the **beta-binomial distribution** with parameters n^*, a and b
- The `rbetabinom.ab()` function in the `VGAM` R package can simulate from a beta-binomial distribution (Yee (2021))

When ppd is not readily available

The general steps for making posterior predictions if ppd is not readily available:

1. Extract the posterior draws of all parameters from MCMC estimation at one iteration (after convergence)
2. Sample one or more predicted data points from the data model using the extracted posterior parameter draws

Normal example

To make a prediction of Y^* at MCMC iteration s, i.e., generate one predicted value:

1. Extract $\{\mu^s, \sigma^s\}$ from estimated MCMC chain after diagnostic checks
2. Sample $Y^* \sim \mathrm{Normal}(\mu^s, \sigma^s)$

Synthetic data

- Recall Overview of synthetic data lecture. . .
- Generating synthetic data from Bayesian models for privacy protection is in essence equivalent to generating posterior predictions from estimated Bayesian models
  - chosen Bayesian models are fitted and estimated on the confidential data,
  - synthetic values of confidential data are generated from the posterior predictive distributions

Bayesian models and utility

- If the selected Bayesian models are suitable for the confidential data at hand, the posterior predictions (i.e., simulated synthetic data) will resemble the confidential data, indicating high utility

- If not, then we should consider building more sophisticated models for the confidential data to achieve better posterior predictions.

Outline

The dataset

- A sample dataset about bike sharing rentals in Washington D.C. from Fanaee-T and Gama (2014)
- Bike rentals information for 250 work days in 2012, including total count and other information

```
bikesharing_data <- readr::read_csv(file = "bikeshare.csv")
bikesharing_data[1:3, ]
```

```
## # A tibble: 3 x 6
##   season weather temperature humidity windspeed count
##   <chr>    <dbl>       <dbl>    <dbl>     <dbl> <dbl>
## 1 spring       1       0.126    0.441     0.366    22
## 2 spring       2       0.119    0.415     0.185    24
## 3 spring       1       0.278    0.524     0.130    33
```

The data dictionary

Variable name	Variable information
season	Categorical; 4 levels: spring, summer, and winter.
weather	Categorical; 1 = clear, few clouds, partly cloudy, 2 = mist + cloudy, mist + broken clouds, mist + few clouds, mist, 3 = light snow, light rain + thunderstorm + scattered clouds, light rain + scattered clouds, 4 = heavy rain + ice pallets + thunderstorm + mist, snow + fog.
temperature	Continuous; Normalized feeling temperature in Celsius. The values are divided to 50 (max).
humidity	Continuous; Normalized humidity. The values are divided to 100 (max).
windspeed	Continuous; Normalized wind speed. The values are divided to 67 (max).
count	Count; Count (in 100s) of total bike rentals including both casual and registered.

Some inferential questions

- Suppose we are interested in modeling and making inference about the count variable in this bike sharing rentals dataset
- Discussion question: What kind of inferential questions you might have given this dataset and the variables it contains?

What we attempt to answer in this lecture

- First, we introduce the basic setup of using a Poisson model for count data, where we only know about the daily counts but none of the additional information such as temperature

What we attempt to answer in this lecture

- First, we introduce the basic setup of using a Poisson model for count data, where we only know about the daily counts but none of the additional information such as temperature
- Next, we deal with the situation where some additional information becomes available, such as temperature, and how a Poisson regression model is built up for taking into account that bike rentals might vary with the daily temperature
- In each stage
  - model setup and prior choices
  - computing techniques for posterior estimation (MCMC)
  - making posterior predictions

A gamma-Poisson conjugate model: Poisson data model

- The Poisson distribution is a **one-parameter distribution**, commonly used to model count data
- It is a discrete distribution, where any sampled draw is a non-negative integer, i.e., a count

The probability mass function of $\mathrm{Poisson}(\lambda)$ is:

$$Pr(Y = y \mid \lambda) = \frac{\lambda^y \exp(-\lambda)}{y!}, y \in \{0, 1, 2, \cdots\}, \lambda > 0, \tag{3}$$

where y is an observed count and λ is the rate parameter. The mean and variance of $Y \sim \mathrm{Poisson}(\lambda)$ are both λ.

A gamma-Poisson conjugate model: Poisson data model

- If we know the value of λ, say $\lambda = 5$, we can get the probability of obtaining a count of, say 8, by calculating
$$Pr(Y = 8 \mid \lambda = 5) = \frac{\lambda^y \exp(-\lambda)}{y!} = \frac{5^8 \exp(-5)}{8!} = 0.065. \tag{4}$$

A gamma-Poisson conjugate model: Poisson data model

- If we know the value of λ, say $\lambda = 5$, we can get the probability of obtaining a count of, say 8, by calculating

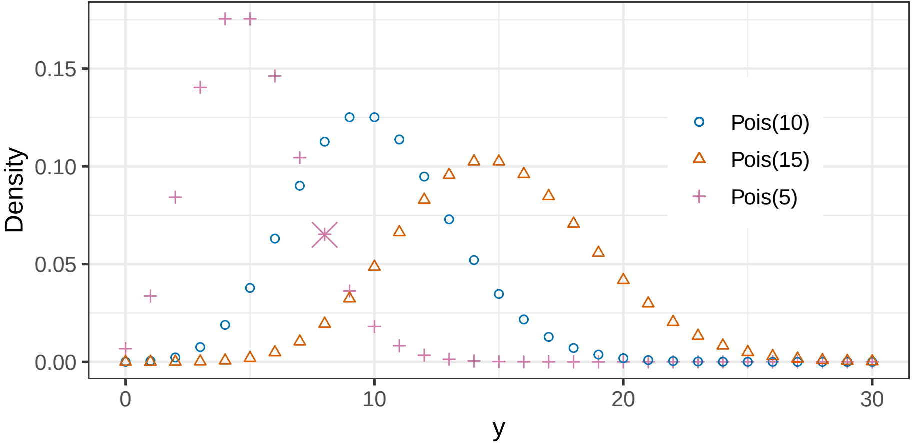
$$Pr(Y = 8 \mid \lambda = 5) = \frac{\lambda^{y} \exp(-\lambda)}{y!} = \frac{5^{8} \exp(-5)}{8!} = 0.065. \tag{4}$$

- In R, we can get it using the `dpois()` function, which returns the density (probability mass function) value

```
stats::dpois(x = 8, lambda = 5)
```

```
## [1] 0.06527804
```

A gamma-Poisson conjugate model: Poisson data model



- The peak value increases as λ increases (about mean)
- The spread increases as λ increases (about variance)

A gamma-Poisson conjugate model: the likelihood function

- We propose to work with a Poisson model for each count, that is:
$$Y_i \mid \lambda \overset{i.i.d.}{\sim} \mathrm{Poisson}(\lambda), \tag{5}$$
where $i = 1, \cdots, n$ and n is the total number of observations ($n = 250$ in our sample)
- Our data model assumes that each day's bike sharing rental count is **identically and independently distributed (i.i.d.)** according to $\mathrm{Poisson}(\lambda)$

A gamma-Poisson conjugate model: the likelihood function

- We can write out the **joint probability mass function** of n observations as:

$$Pr(Y_1 = y_1, \cdots, Y_n = y_n \mid \lambda) = \prod_{i=1}^{n} \frac{\lambda^{y_i} \exp(-\lambda)}{y_i!} = \frac{\lambda^{\sum_{i=1}^{n} y_i} \exp(-n\lambda)}{\prod_{i=1}^{n} y_i!}. \tag{6}$$

A gamma-Poisson conjugate model: the likelihood function

- We can write out the **joint probability mass function** of n observations as:

$$Pr(Y_1 = y_1, \cdots, Y_n = y_n \mid \lambda) = \prod_{i=1}^{n} \frac{\lambda^{y_i} \exp(-\lambda)}{y_i!} = \frac{\lambda^{\sum_{i=1}^{n} y_i} \exp(-n\lambda)}{\prod_{i=1}^{n} y_i!}. \tag{6}$$

- From this, the likelihood function of λ can be then expressed as:

$$L(\lambda) = \frac{\lambda^{\sum_{i=1}^{n} y_i} \exp(-n\lambda)}{\prod_{i=1}^{n} y_i!} \propto \lambda^{\sum_{i=1}^{n} y_i} \exp(-n\lambda). \tag{7}$$

- As mentioned last time, the likelihood function is a function of the unknown parameter(s)
- Since $\{y_1, \cdots, y_n\}$ are observed, they can be simplified into the proportional sign

A gamma-Poisson conjugate model: the likelihood function

- We can write out the **joint probability mass function** of n observations as:

$$Pr(Y_1 = y_1, \cdots, Y_n = y_n \mid \lambda) = \prod_{i=1}^{n} \frac{\lambda^{y_i} \exp(-\lambda)}{y_i!} = \frac{\lambda^{\sum_{i=1}^{n} y_i} \exp(-n\lambda)}{\prod_{i=1}^{n} y_i!}. \tag{6}$$

- From this, the likelihood function of λ can be then expressed as:

$$L(\lambda) = \frac{\lambda^{\sum_{i=1}^{n} y_i} \exp(-n\lambda)}{\prod_{i=1}^{n} y_i!} \propto \lambda^{\sum_{i=1}^{n} y_i} \exp(-n\lambda). \tag{7}$$

- As mentioned last time, the likelihood function is a function of the unknown parameter(s)
- Since $\{y_1, \cdots, y_n\}$ are observed, they can be simplified into the proportional sign
- Bayesian inference: we assume λ is unknown and we will give a prior distribution for it

A gamma-Poisson conjugate model: a gamma prior

- The gamma distribution is a two-parameter continuous distribution
- It takes positive values, and therefore is commonly used for modeling positive quantities
  - the rate parameter in a Poisson data model
  - the precision parameter (i.e., the reciprocal of the variance parameter) in a normal data model

A gamma-Poisson conjugate model: a gamma prior

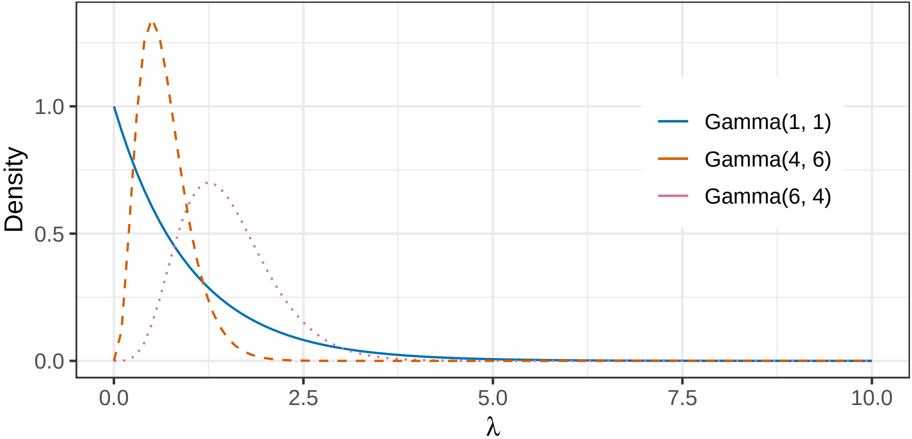
- The gamma distribution is a two-parameter continuous distribution
- It takes positive values, and therefore is commonly used for modeling positive quantities
    - the rate parameter in a Poisson data model
    - the precision parameter (i.e., the reciprocal of the variance parameter) in a normal data model

The probability density function of $\mathrm{Gamma}(a, b)$ is:

$$f(\lambda) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp(-b\lambda), \lambda > 0, a > 0, b > 0, \tag{8}$$

where a is the shape parameter and b is the rate parameter. The mean and variance of $\lambda \sim \mathrm{Gamma}(a, b)$ are a/b and a/b^2 respectively

A gamma-Poisson conjugate model: a gamma prior



- $\text{Gamma}(1, 1)$ is a common prior choice when we do not have much prior information (**weakly informative priors**)

A gamma-Poisson conjugate model: posterior derivation

- Continuous Bayes' rule
- The joint likelihood function of λ:

$$L(\lambda) = \frac{\lambda^{\sum_{i=1}^{n} y_i} \exp(-n\lambda)}{\prod_{i=1}^{n} y_i!} \propto \lambda^{\sum_{i=1}^{n} y_i} \exp(-n\lambda).$$

- The prior density of λ using $\mathrm{Gamma}(a, b)$:

$$\pi(\lambda) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp(-b\lambda) \propto \lambda^{a-1} \exp(-b\lambda),$$

where we put known quantities $1/\prod_{i=1}^{n} y_i!$, b^a, and $\Gamma(a)$ into the proportional sign

- Discussion question: Can you derive the posterior to be $\mathrm{Gamma}(a + \sum_{i=1}^{n} y_i, b + n)$?

A gamma-Poisson conjugate model: conjugacy theorem

For a Poisson data model where $Y_1, \cdots, Y_n \overset{iid}{\sim} \mathrm{Poisson}(\lambda)$, a $\mathrm{Gamma}(a, b)$ prior for λ gives a $\mathrm{Gamma}(a + \sum_{i=1}^{n} y_i, b + n)$ posterior.

A gamma-Poisson conjugate model: posterior mean

- We can express the **posterior mean**, that is the mean of $\mathrm{Gamma}(a + \sum_{i=1}^{n} y_i, b + n)$ as:

$$\frac{a + \sum_{i=1}^{n} y_i}{b + n} = \frac{b}{b + n} \times \frac{a}{b} + \frac{n}{b + n} \times \frac{\sum_{i=1}^{n} y_i}{n} = \frac{b}{b + n} \times \frac{a}{b} + \frac{n}{b + n} \times \bar{y}, \tag{9}$$

  where a/b and $\bar{y}$ are the **prior mean** and the **sample mean**, respectively

A gamma-Poisson conjugate model: posterior mean

- We can express the **posterior mean**, that is the mean of $\mathrm{Gamma}(a + \sum_{i=1}^{n} y_i, b + n)$ as:
$$\frac{a + \sum_{i=1}^{n} y_i}{b + n} = \frac{b}{b+n} \times \frac{a}{b} + \frac{n}{b+n} \times \frac{\sum_{i=1}^{n} y_i}{n} = \frac{b}{b+n} \times \frac{a}{b} + \frac{n}{b+n} \times \bar{y}, \tag{9}$$
where a/b and $\bar{y}$ are the **prior mean** and the **sample mean**, respectively

- The posterior mean as a **weighted average** of the prior mean and the sample mean, with weight $b/(b+n)$ for the prior and $n/(b+n)$ for the data sample

A gamma-Poisson conjugate model: posterior mean

- We can express the **posterior mean**, that is the mean of $\mathrm{Gamma}(a + \sum_{i=1}^{n} y_i, b + n)$ as:

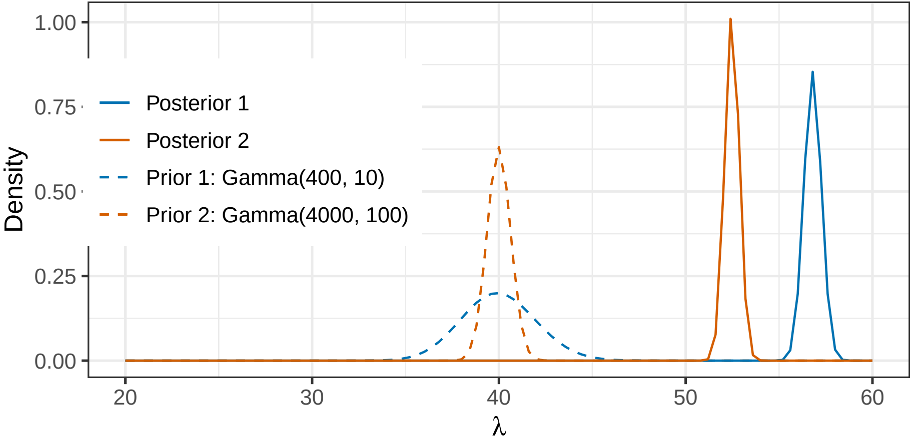
$$\frac{a + \sum_{i=1}^{n} y_i}{b+n} = \frac{b}{b+n} \times \frac{a}{b} + \frac{n}{b+n} \times \frac{\sum_{i=1}^{n} y_i}{n} = \frac{b}{b+n} \times \frac{a}{b} + \frac{n}{b+n} \times \bar{y}, \tag{9}$$

  where a/b and $\bar{y}$ are the **prior mean** and the **sample mean**, respectively

- The posterior mean as a **weighted average** of the prior mean and the sample mean, with weight $b/(b+n)$ for the prior and $n/(b+n)$ for the data sample

- Given fixed b, the larger the value of n, the stronger the influence of the data on the posterior; similarly, given fixed n, the larger the value of b, the stronger the influence of the prior on the posterior

A gamma-Poisson conjugate model: prior influence



- Discussion question: Which prior has stronger influence? What about the change of spread from prior to posterior?

A gamma-Poisson conjugate model: inference

- Interested in posterior mean and variance

1. Analytically

```
a <- 1
b <- 1
(a + sum_y) / (b + n)
```

```
## [1] 57.24701
```

```
(a + sum_y) / (b + n)^2
```

```
## [1] 0.2280757
```

A gamma-Poisson conjugate model: posterior inference

2. Simulation-based

```
S <- 1000
lambda_draws <- stats::rgamma(n = S, shape = (a + sum_y), rate = (b + n))

mean(lambda_draws)
```

```
## [1] 57.25217
```

```
var(lambda_draws)
```

```
## [1] 0.2163508
```

- The larger the Monte Carlo sample is, i.e., the larger S is, the better the Monte Carlo approximation

A gamma-Poisson conjugate model: posterior prediction

- Can we predict the next year's bike share rental counts, i.e., make a predicted sample of 250 rental counts?
- For $i = 1, \cdots, n$ where $n = 250$, we first simulate a posterior draw of λ, denoted as λ^* and next simulate a value for each prediction Y_i^*

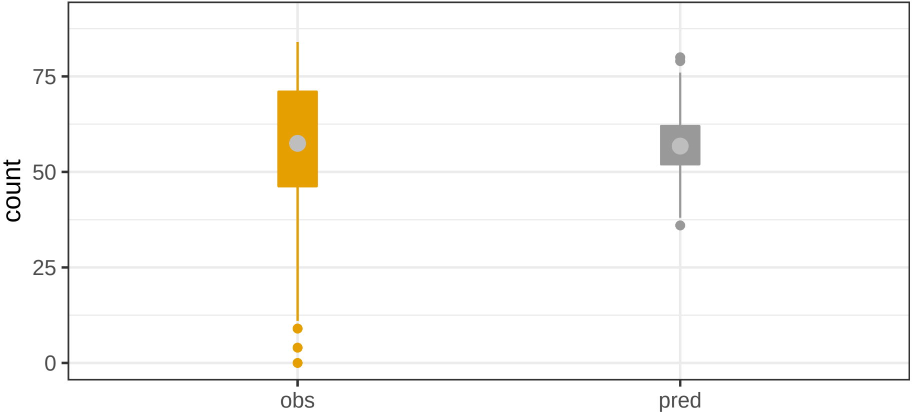
$$
\begin{aligned}
& & \rightarrow \quad & \text{sample } Y_1^* \sim \mathrm{Poisson}(\lambda^*) \\
& & & \vdots \\
\text{sample } \lambda^* \sim \mathrm{Gamma}\left(a + \sum_{i=1}^{n} y_i, b + n\right) & & \rightarrow \quad & \text{sample } Y_i^* \sim \mathrm{Poisson}(\lambda^*) \\
& & & \vdots \\
& & \rightarrow \quad & \text{sample } Y_n^* \sim \mathrm{Poisson}(\lambda^*)
\end{aligned}
$$

A gamma-Poisson conjugate model: posterior prediction

- In R
  - `rgamma()` function to simulate a posterior draw of λ^*
  - `rpois()` function to simulate predicted $Y_1^*, \cdots, Y_n^*$
  - `set.seed()` with a particular seed value to obtain reproducible results

```
set.seed(248)
lambda_draw <- stats::rgamma(n = 1, shape = (a + sum_y), rate = (b + n))
y_pred <- stats::rpois(n = n, lambda = lambda_draw)
```

A gamma-Poisson conjugate model: posterior prediction



- Mean seems okay, spread (variance) not so much
- Issues with Poisson data model (mean and variance are both λ)

A gamma-Poisson conjugate model: prior choices

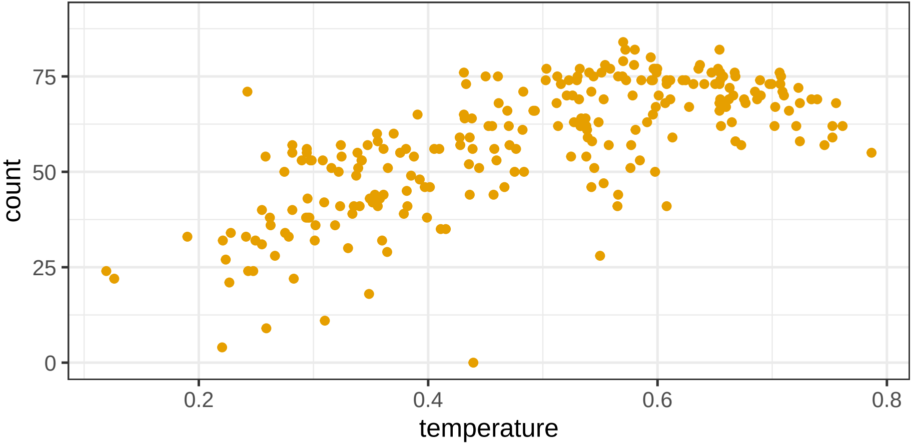
- Prior distributions are used to encode our subjective belief about the unknown parameter
- Do not choose convenient priors, such as gamma for Poisson due to conjugacy
- If there are other prior distributions that reflect our belief better, we should use those prior choices
- MCMC estimation software can help with posterior estimation

A gamma-Poisson conjugate model: prior choices

- Prior distributions are used to encode our subjective belief about the unknown parameter
- Do not choose convenient priors, such as gamma for Poisson due to conjugacy
- If there are other prior distributions that reflect our belief better, we should use those prior choices
- MCMC estimation software can help with posterior estimation

- Next, we try using additional information to better model the counts as well as better prediction results

A Poisson regression model: motivation



- A positive relationship between counts and temperature: as temperature increases, the counts of bike rentals increase as well
- Discussion question: How can we utilize such useful information in modeling and predicting counts?

A Poisson regression model: model specification

- An observation-specific Poisson data model can be expressed as:
$$Y_i \overset{ind}{\sim} \mathrm{Poisson}(\lambda_i), \tag{10}$$
where $\lambda_i > 0$ is the rate parameter of the Poisson model for observation i ($i = 1, \cdots, n$). The *ind* symbol indicates that observations are independent from each other

A Poisson regression model: model specification

- An observation-specific Poisson data model can be expressed as:
$$Y_i \overset{ind}{\sim} \mathrm{Poisson}(\lambda_i), \tag{10}$$
where $\lambda_i > 0$ is the rate parameter of the Poisson model for observation i ($i = 1, \cdots, n$). The *ind* symbol indicates that observations are independent from each other

- Given their relationship, we would want to express λ_i in terms of the temperature of day i
- One way to do so is to express $\log(\lambda_i)$ as a linear function of temperature i, denoted by X_i:
$$\log(\lambda_i) = \beta_0 + \beta_1 X_i, \tag{11}$$
where $\{\beta_0, \beta_1\}$ are the parameters of our model (λ_i's are not parameters)
- Discussion question: Why $\log(\lambda_i)$ instead of λ_i?

A Poisson regression model: MCMC estimation (brms)

```
PoissonReg_fit <- brms::brm(data = bikesharing_data,
                            family = poisson(link = "log"),
                            count ~ 1 + temperature,
                            prior = c(prior(normal(0, 10),
                                            class = Intercept),
                                      prior(normal(0, 10), class = b)),
                            iter = 5000,
                            warmup = 3000,
                            thin = 1,
                            chains = 1,
                            seed = 123)
```

A Poisson regression model: MCMC estimation (brms)

- `brm()` function is the main function we use from the `brms` package
- `data = bikesharing_data`: dataset
- `family = poisson(link = "log")`: data model
- `count ~ 1 + temperature`: model expression
- Prior statements (if left blank, default priors will be used)
- `iter, warmup, thin, chains, seed`: MCMC configurations

A Poisson regression model: MCMC estimation (brms)

```
post_PoissonReg <- brms::posterior_samples(x = PoissonReg_fit)
post_PoissonReg[1:3, ]
```

```
##   b_Intercept b_temperature      lp__
## 1    3.318428      1.463756 -1120.863
## 2    3.381624      1.316238 -1120.003
## 3    3.329988      1.437574 -1120.039
```

- `posterior_samples()` to extract posterior parameter draws
- Discussion question: how many rows in `post_PoissonReg`?

A Poisson regression model: MCMC estimation (brms)

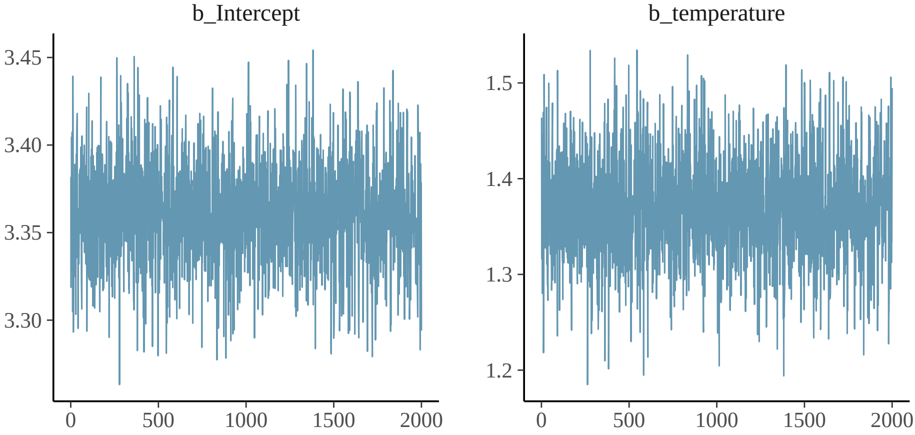
```
post_PoissonReg <- brms::posterior_samples(x = PoissonReg_fit)
post_PoissonReg[1:3, ]
```

```
##   b_Intercept b_temperature      lp__
## 1    3.318428      1.463756 -1120.863
## 2    3.381624      1.316238 -1120.003
## 3    3.329988      1.437574 -1120.039
```

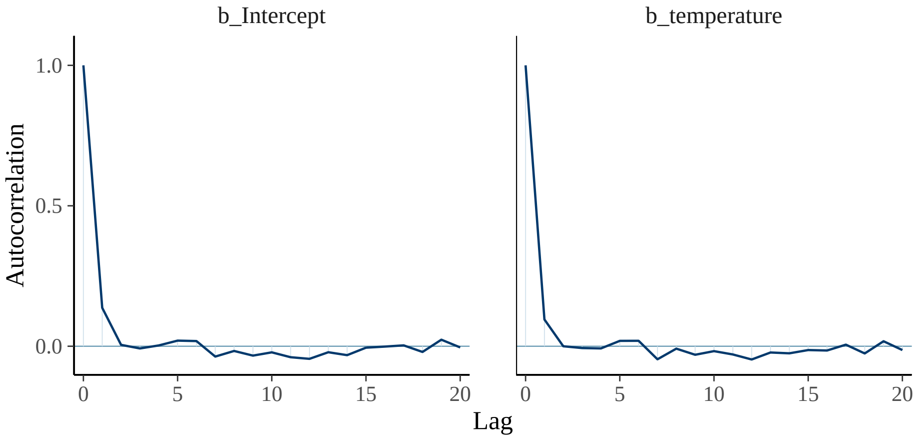
- `posterior_samples()` to extract posterior parameter draws
- Discussion question: how many rows in `post_PoissonReg`?

- Later in prediction, we will need to exponeniate $\log(\lambda_i) = \beta_0 + \beta_1 X_i$ to obtain λ_i to be used in `rpois()` to predict a new value Y_i^*

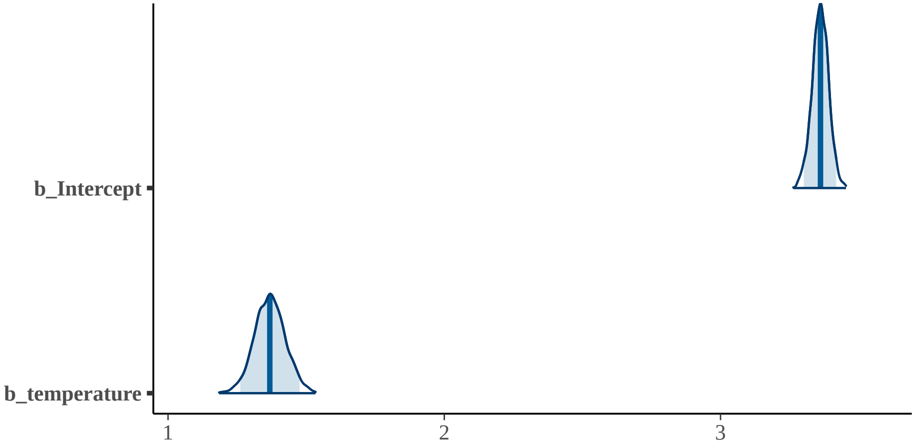
A Poisson regression model: MCMC diagnostics



A Poisson regression model: MCMC diagnostics



A Poisson regression model: posterior inference



- The thickened middle bar in each plot is the mode, while the shaded area gives the central 95% credible interval: there is a 95% posterior probability that the parameter falls into this interval

A Poisson regression model: posterior prediction

$$
\begin{array}{ccc}
\text{compute } \log(\lambda_1^*) = \beta_0^* + \beta_1^* X_1 & \rightarrow & \text{sample } Y_1^* \sim \mathrm{Poisson}(\lambda_1^*) \\
 & \vdots & \\
\text{compute } \log(\lambda_i^*) = \beta_0^* + \beta_1^* X_i & \rightarrow & \text{sample } Y_i^* \sim \mathrm{Poisson}(\lambda_i^*) \\
 & \vdots & \\
\text{compute } \log(\lambda_n^*) = \beta_0^* + \beta_1^* X_n & \rightarrow & \text{sample } Y_n^* \sim \mathrm{Poisson}(\lambda_n^*)
\end{array}
$$

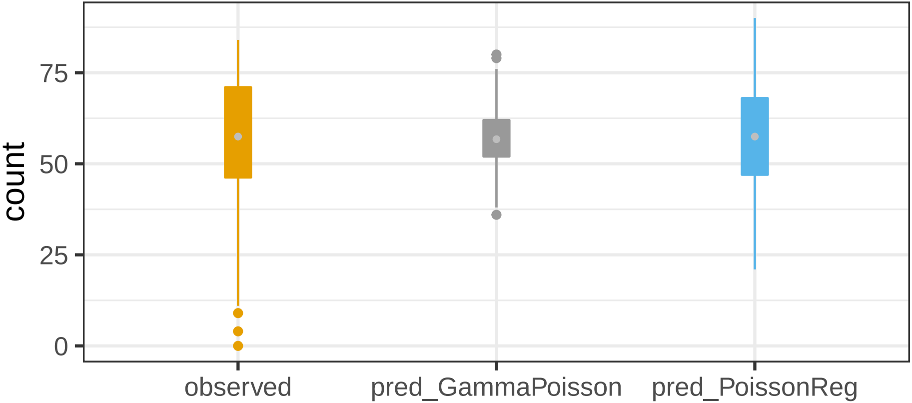
- $\{\beta_0^*, \beta_1^*\}$ represent one set of posterior draws of the two parameters

A Poisson regression model: posterior prediction

Below shows how we can generate one predicted dataset using the first pair of $\{\beta_0^*, \beta_1^*\}$ posterior draws

```
y_pred_PoissonReg <- rep(NA, n)
for (i in 1:n){
  y_pred_PoissonReg[i] <- rpois(n = 1,
                                lambda =
                                  exp(post_PoissonReg[1, "b_Intercept"] +
                                        post_PoissonReg[1, "b_temperature"] *
                                        bikesharing_data$temperature[i]))
}
```

A Poisson regression model: posterior prediction



- Discussion question: What are your findings between the gamma-Poisson conjugate model and the Poisson regression model for this bike rental sharing dataset?

Outline

Summary

- Posterior predictive
- Synthetic data
- Case study
  - Gamma-Poisson conjugate model
  - Poisson regression model
  - Posterior prediction, prior choices, posterior inference, MCMC, MCMC diagnostics

Summary

- Posterior predictive
- Synthetic data
- Case study
  - Gamma-Poisson conjugate model
  - Poisson regression model
  - Posterior prediction, prior choices, posterior inference, MCMC, MCMC diagnostics

- Homework 3: a few derivation and R programming exercises
  - Submission on Moodle and prepare to discuss next time
- Lecture 4: Bayesian synthesis models part 1: continuous and binary
  - Section 3 of Kinney et al. (2011)
  - Section 12.4 of Albert and Hu (2019): https://bayesball.github.io/BOOK/proportion.html (a different MCMC software is used)

References I

Albert, J., and J. Hu. 2019. Probability and Bayesian Modeling. Texts in Statistical Science, Chapman Hall CRC Press.

Fanaee-T, Hadi, and Joao Gama. 2014. "Event Labeling Combining Ensemble Detectors and Background Knowledge." Progress in Artificial Intelligence, 113–27. https://doi.org/10.1007/s13748-013-0040-3.

Kinney, S. K., J. P. Reiter, A. P. Reznek, J. Miranda, R. S. Jarmin, and J. M. Abowd. 2011. "Towards Unrestricted Public Use Business Microdata: The Synthetic Longitudinal Business Database." International Statistical Review 79 (3): 362–84.

Yee, Thomas W. 2021. VGAM: Vector Generalized Linear and Additive Models. https://CRAN.R-project.org/package=VGAM.