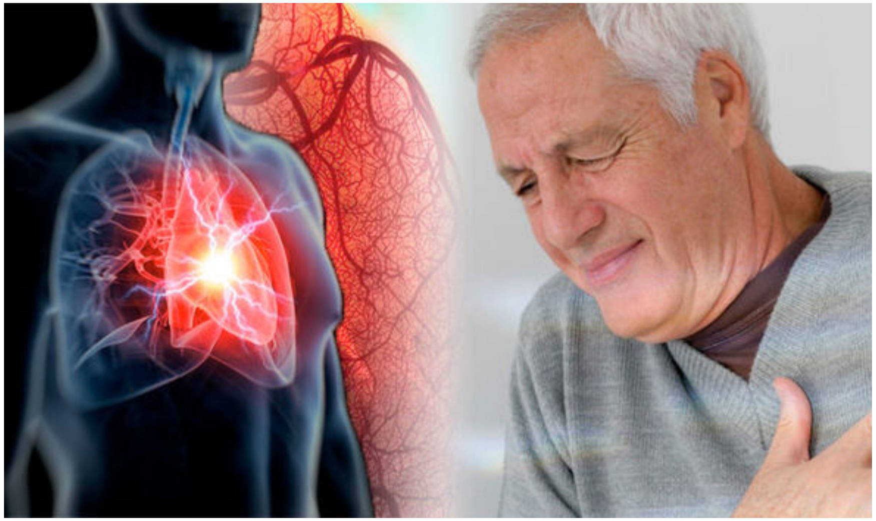
What Can Probabilistic Modeling Do For Me?

Arya Pourzanjani

Data is random, and randomness is described by probability distributions



- Example: In a study of 39 people who had heart attacks, 3 (8%) died within a year

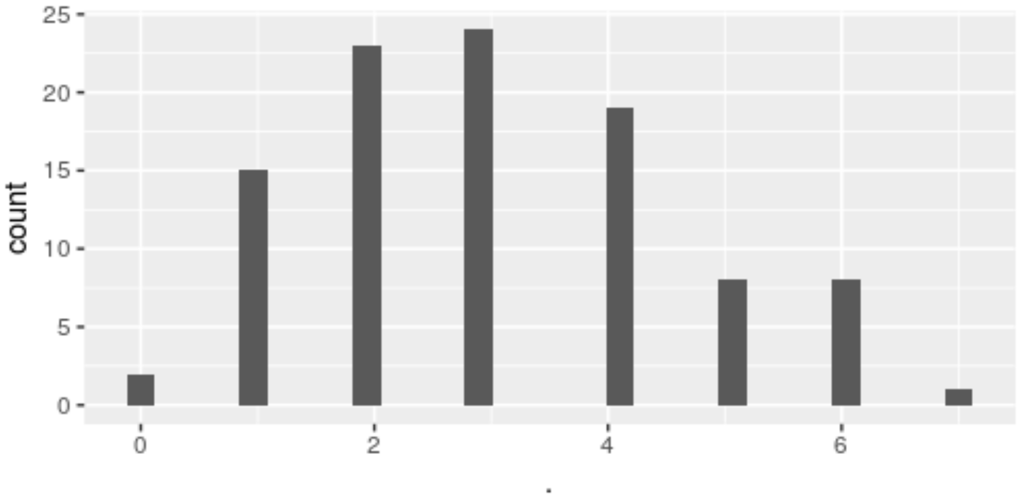
Data is random, and randomness is described by probability distributions

- Random variables are described by distributions, functions that tell you the probability the random variable will take on some value

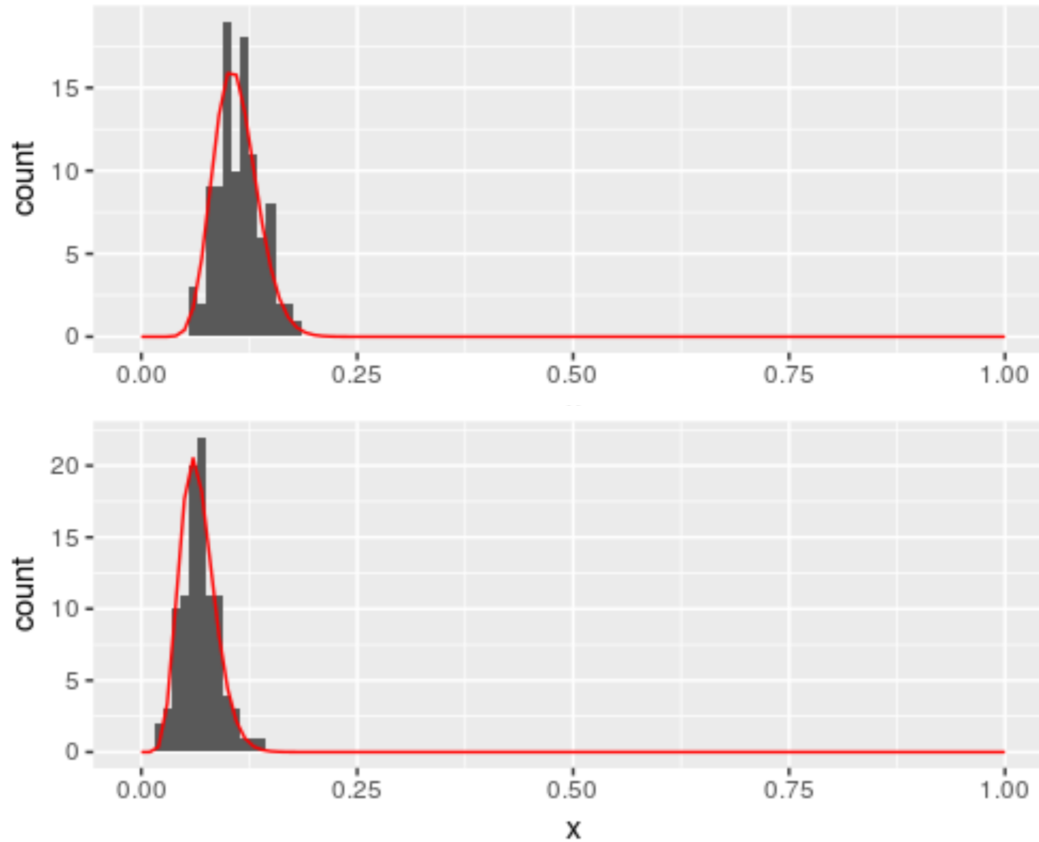
$$N \sim \mathrm{Binomial}(39, \theta)$$

$$p(N = n \,|\, \theta) = \binom{39}{n} \theta^n (1 - \theta)^{39-n}$$

Data is random, and randomness is described by probability distributions



If we can accurately describe our data we can answer questions with our data



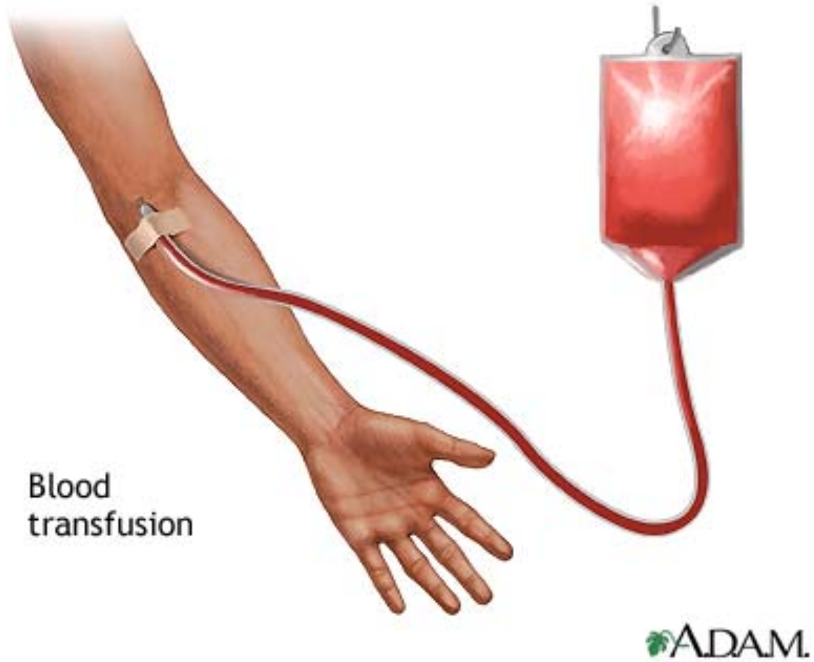
There are three steps to probabilistic modeling

1. Choosing a model for your data

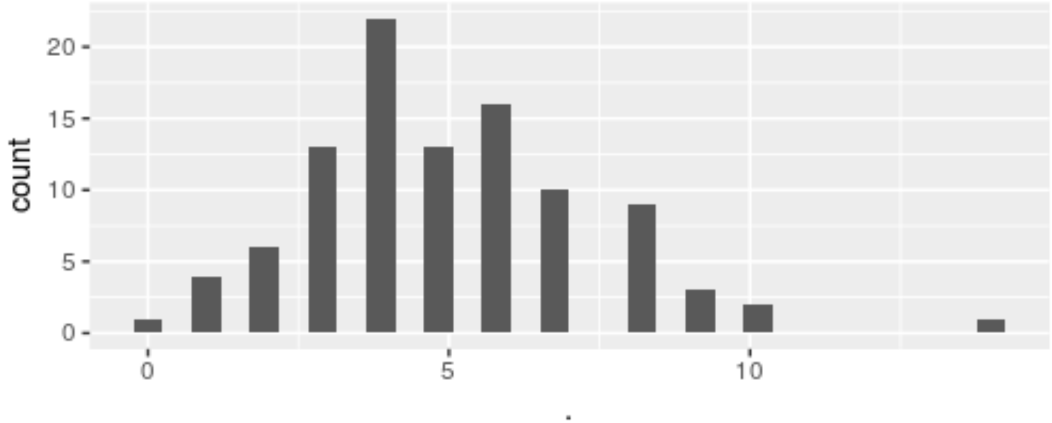
2. Doing Bayesian inference

3 .Checking your model

Choosing a model or distribution for your data depends on the kind of data you have



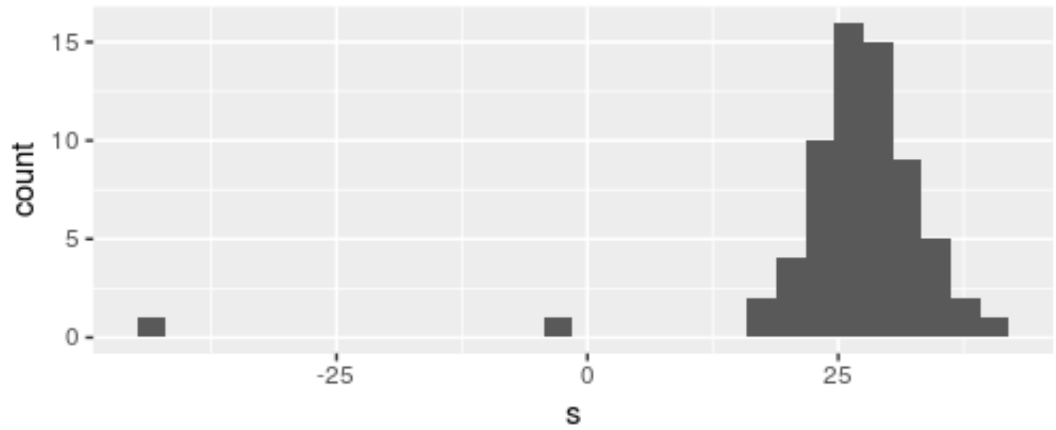
Choosing a model or distribution for your data depends on the kind of data you have



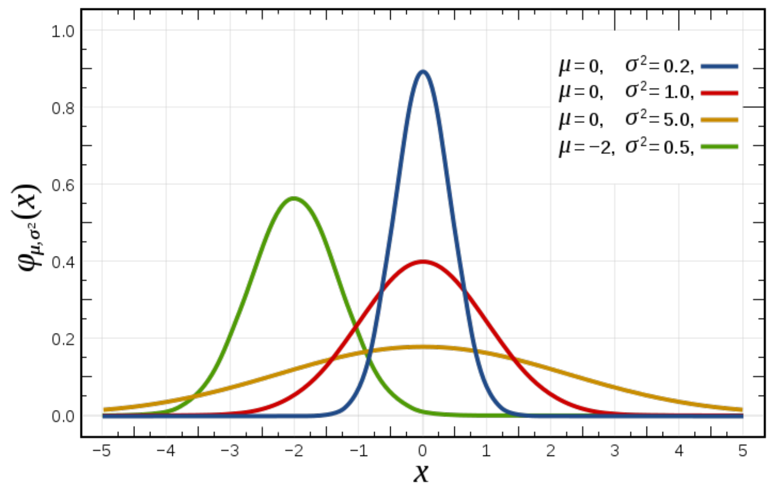
$$B \sim \mathrm{Poisson}(\lambda)$$

$$p(B = b \,|\, \lambda) = \frac{\lambda^b e^{-\lambda}}{b!}$$

Choosing a model or distribution for your data depends on the kind of data you have



Choosing a model or distribution for your data depends on the kind of data you have



$$Y \sim \mathcal{N}(\mu, \sigma^2)$$

$$f(y \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

There are three steps to probabilistic modeling

1. Choosing a model for your data

2. Doing Bayesian inference

   -The old-fashioned way

   -The computational way

3 .Checking your model

Bayesian inference allows us to estimate parameters of a model from data

$$N \sim \text{Binomial}(39, \theta)$$

$$B \sim \text{Poisson}(\lambda)$$

$$Y \sim \mathcal{N}(\mu, \sigma^2)$$

Bayesian inference allows us to estimate parameters of a model from data

$$p(B \mid A) \cdot p(A) = p(A, B) = p(A \mid B) \cdot p(B)$$

$$p(B \mid A) = \frac{p(A \mid B) \cdot p(B)}{p(A)}$$

$$\overbrace{p(\theta \mid N = n)}^{\text{Posterior}} = \frac{\overbrace{p(N = n \mid \theta)}^{\text{Model}} \; \overbrace{p(\theta)}^{\text{Prior}}}{\underbrace{p(N = n)}_{\text{Normalization Term}}}$$

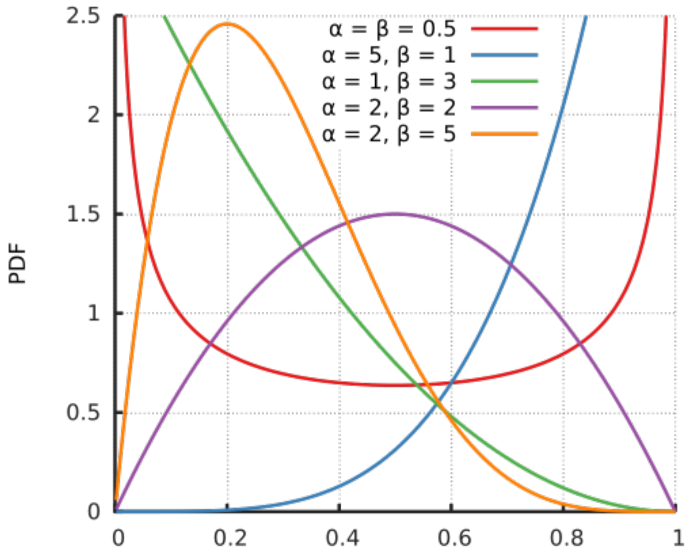
Bayesian inference allows us to estimate parameters of a model from data

$$\overbrace{p(\theta \,|\, N = n)}^{\text{Posterior}} = \frac{\overbrace{p(N = n \,|\, \theta)}^{\text{Model}} \; \overbrace{p(\theta)}^{\text{Prior}}}{\underbrace{p(N = n)}_{\text{Normalization Term}}}$$

$$N \sim \text{Binomial}(39, \theta)$$

$$p(N = n \,|\, \theta) = \binom{39}{n} \theta^n (1 - \theta)^{39-n}$$

Bayesian inference allows us to estimate parameters of a model from data



$$p(\theta \mid N = n) = \frac{\overbrace{p(N = n \mid \theta)}^{\text{Model}} \; \overbrace{p(\theta)}^{\text{Prior}}}{\underbrace{p(N = n)}_{\text{Normalization Term}}}$$

$$p(\theta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\mathrm{B}(\alpha, \beta)}$$

Bayesian inference allows us to estimate parameters of a model from data

$$p(\theta \mid N = n) = \frac{\overbrace{p(N = n \mid \theta)}^{\text{Model}} \; \overbrace{p(\theta)}^{\text{Prior}}}{\underbrace{p(N = n)}_{\text{Normalization Term}}}$$

$$p(N = n) = \int_0^1 p(N = n \mid \theta) \cdot p(\theta) \, d\theta$$

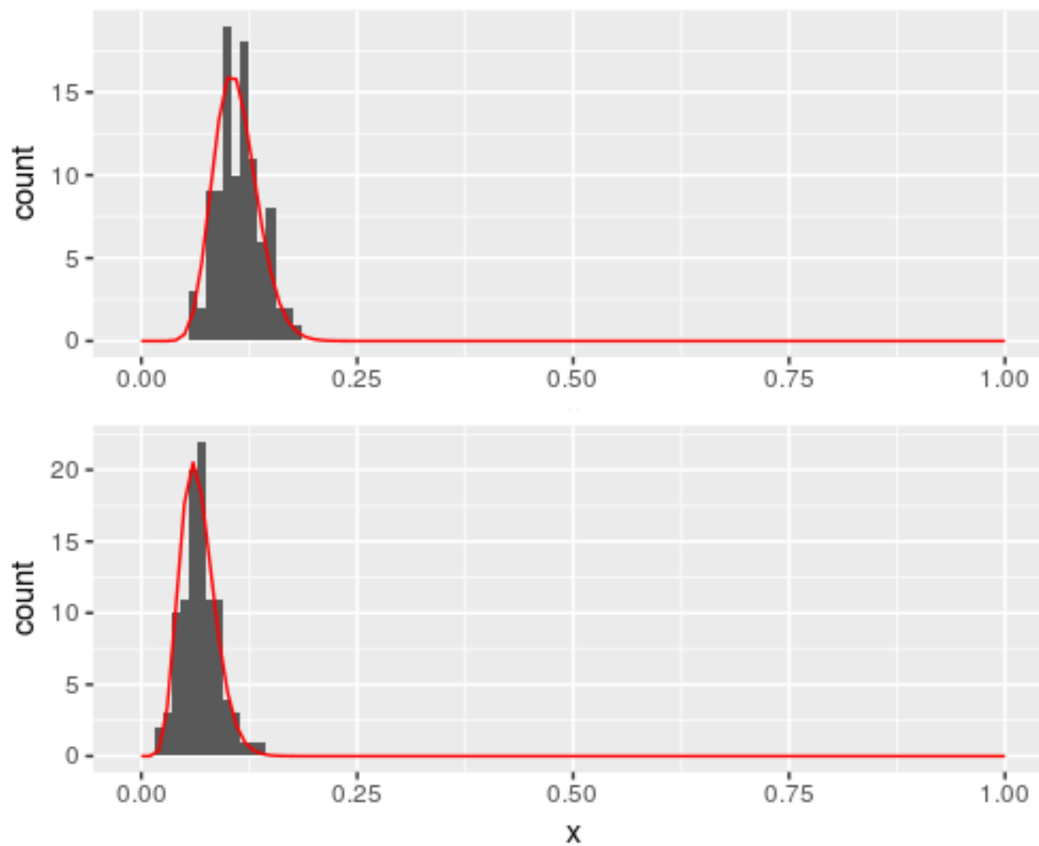
Bayesian inference allows us to estimate parameters of a model from data

$$p(N = n \,|\, \theta) = \binom{39}{n} \theta^n (1-\theta)^{39-n}$$

$$p(\theta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\mathrm{B}(\alpha, \beta)}$$

$$p(\theta \,|\, N = n) \propto \theta^{n+\alpha-1}(1-\theta)^{39-n+\beta-1}$$

Bayesian inference allows us to estimate parameters of a model from data



Optimization is an approximation of full inference

$$p(\theta \mid N = n) = \frac{\overbrace{p(N = n \mid \theta)}^{\text{Model}} \; \overbrace{p(\theta)}^{\text{Prior}}}{\underbrace{p(N = n)}_{\text{Normalization Term}}}$$

$$p(N = n) = \int_0^1 p(N = n \mid \theta) \cdot p(\theta) \, d\theta$$

There are three steps to probabilistic modeling

1. Choosing a model for your data

2. Doing Bayesian inference

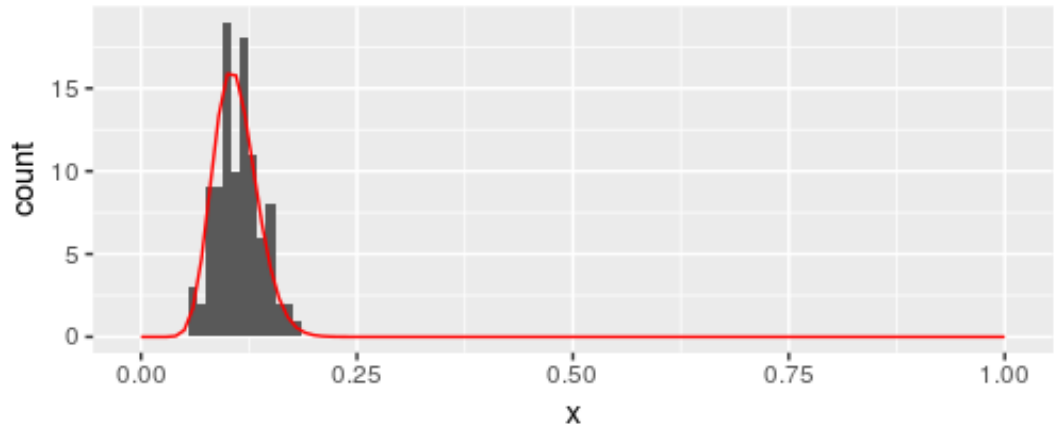
   -The old-fashioned way

   -The computational way

3 .Checking your model

Samples from the posterior distribution are just as good as the posterior distribution

$$p(\theta \mid N = n) = \frac{\overbrace{p(N = n \mid \theta)}^{\text{Model}} \; \overbrace{p(\theta)}^{\text{Prior}}}{\underbrace{p(N = n)}_{\text{Normalization Term}}}$$



Samples from the posterior distribution are just as good as the posterior distribution

$$p(\theta \mid N = n) = \frac{\overbrace{p(N = n \mid \theta)}^{\text{Model}}\ \overbrace{p(\theta)}^{\text{Prior}}}{\underbrace{p(N = n)}_{\text{Normalization Term}}}$$

$$p(\theta_1, \theta_2 \mid N = n)$$

Stan allows us to specify models and
it gets the posterior distribution samples for us

- http://mc-stan.org/

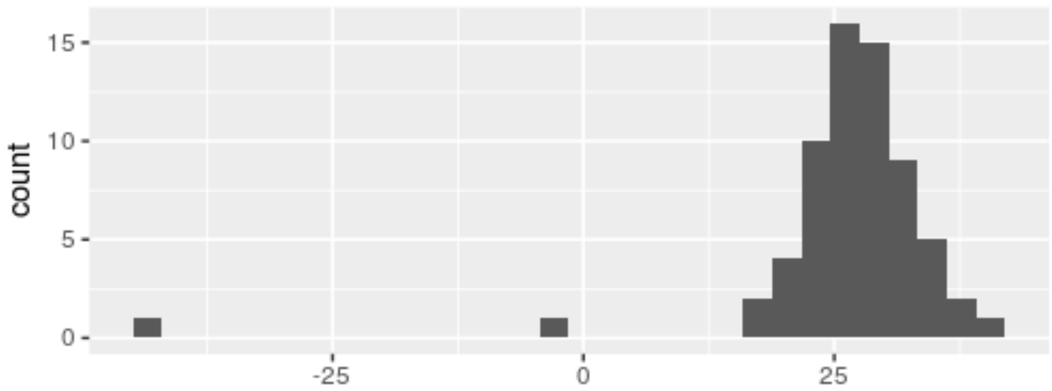
There are three steps to probabilistic modeling

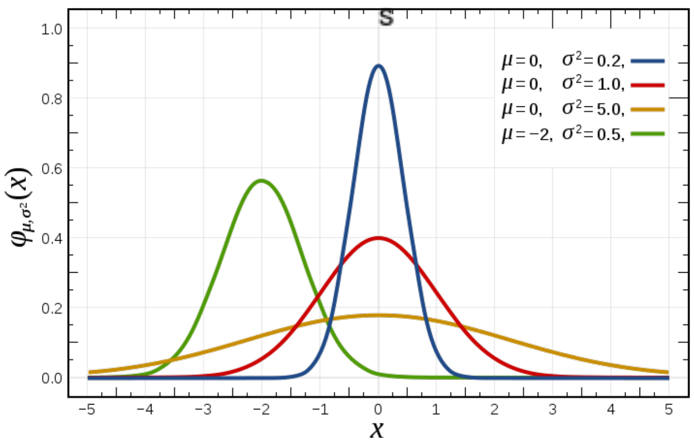
1. Choosing a model for your data

2. Doing Bayesian inference

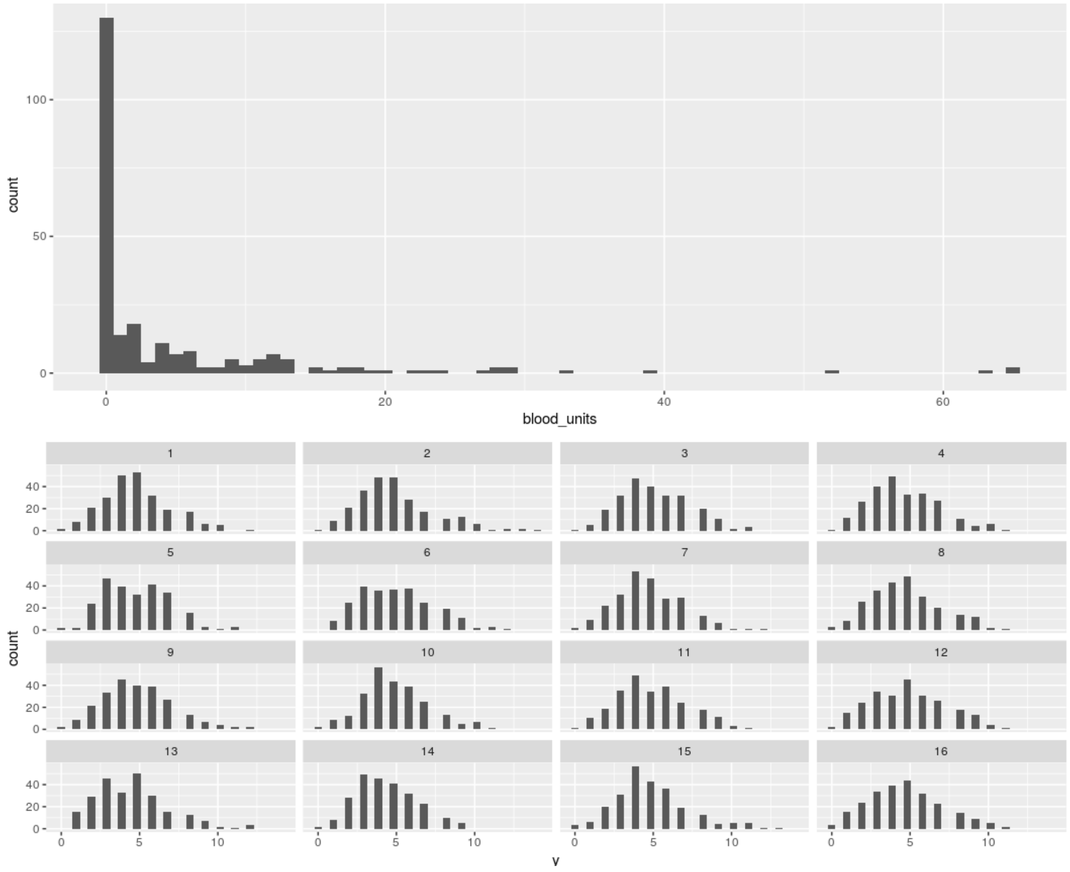
3 .Checking your model

We can generate fake data from fit models to check the plausibility of our models





We can generate fake data from fit models to check the plausibility of our models



We can generate fake data from fit models to check the plausibility of our models

$$p(y \mid \theta, \lambda) = \begin{cases} \theta + (1-\theta) \cdot \mathrm{Poisson}(0 \mid \lambda) & \text{if } y = 0 \\ (1-\theta) \cdot \mathrm{Poisson}(y \mid \lambda) & \text{if } y > 0 \end{cases}$$

There are three steps to probabilistic modeling

1. Choosing a model for your data

2. Doing Bayesian inference

3 .Checking your model

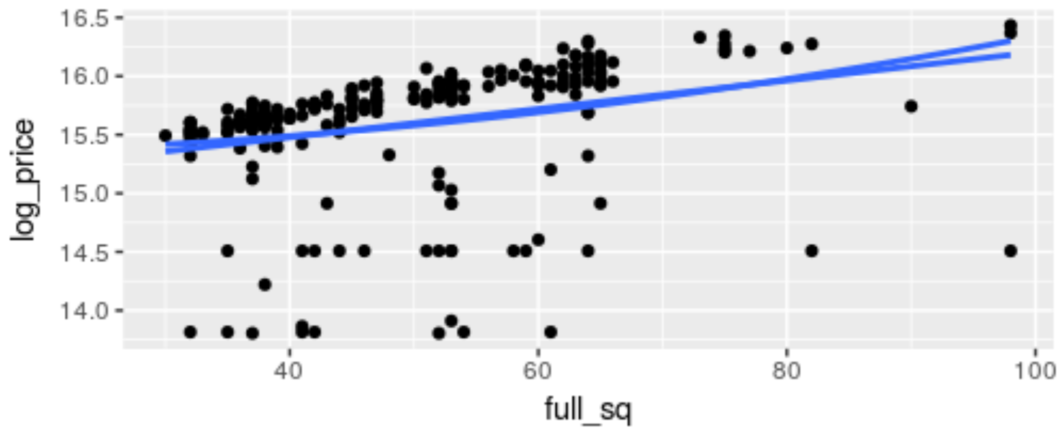
- More complicated models

Hierarchical models are useful when there are different types of observational units

$$N_i \sim \mathrm{Binomial}(39, \theta_i)$$

$$\theta_i \sim \mathrm{Beta}(\alpha, \beta)$$

Regression is a powerful tool for learning functions



$$y_i \sim \mathcal{N}(x_i\beta_1 + \beta_0, \sigma^2)$$