A Brief Introduction to R Shiny

Dane Alabran

University of Pittsburgh

October 14, 2017

Overview

What Exactly is R Shiny?

The technical definition straight from R:

Definition

Shiny is an R package that makes it easy to build interactive web apps straight from R. You can host standalone apps on a webpage or embed them in R Markdown documents or build dashboards. You can also extend your Shiny apps with CSS themes, htmlwidgets, and JavaScript actions.

What Exactly is R Shiny?

Simply put:

1. Shiny makes many kinds of interactive apps
2. Apps can be customized and published
3. Multiple functions can be added
4. If you can imagine it, you can make it.

But what makes Shiny work?

What Drives R Shiny?

All Shiny creations must come from a computer running R (Shiny is an R package after all). Because of this, all Shiny apps can be housed in R or R Studio like it was just another code. Every application must have 4 components in order for Shiny to work:

1. The library(shiny)
2. The user interface definition ui
3. The server definition
4. The shinyApp declaration

The above, when implemented, create the most basic of apps, which we will see in the Hello World example.

Hello World Example

Our first app will be a friendly greeting called "Hello World", which is typically the example used for the most simple programs in many programming languages.

The User Interface

The user interface defines what our application page will look like. Though Shiny has many kinds of page styles, we will focus on the most versatile one called fluidPage(). We will see several options for ui in the following examples.

What can this "fluid page" do for you?

1. Set up panels or "place holders" for titles, sidebar options, and plots
2. Allow for user input of any kind
   - Numeric input can come from sliders
   - Text input can be entered directly or selected in a drop down menu
   - Date inputs can be selected from a calendar
   - etc.
3. Allow users to define "Action Buttons" to force specific application events, e.g. "Exit" closes the app.

The Server

We must define servers using a function command in R, and tell R what to do with the input and the output. So all servers will have the form

$$\text{server} = \text{function(input, output)}\{\ldots\}$$

Server functions:

- Houses all the program logic
- Uses Reactive() to update the results based on text input
- Uses RenderPlot() to update all plot output based off the input
- Defines what should happen when buttons are pressed

Scatter Plots of Anderson's Iris Data

Suppose we wanted to visualize the iris dataset by producing multiple scatter plots of all pairs of quantitative variables. Instead of changing the code over and over again, we can create a Shiny App to change the input for us and react by updating the scatter plots.

Extending Functionality: Looking up Stocks

We will create a Shiny App that will look up whatever stock index we desire, granted we know what the symbol for that stock is, and plot the stock's prices for a user defined date range.

This app uses library(quantmod) which lets us pull up-to-date stock prices directly from google or yahoo finance. This app also uses a special type of plot called chart series to make the output look more professional.

But this is statistics, right?

What if wanted to do something a little more complicated like see how the choice of a conjugate prior effects the posterior distribution? First we need to understand what I just said and give a brief introduction to Bayesian Statistics.

Theorem 1

Theorem (Normal Mean Conjugate Prior)

Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution with unknown mean θ and known variance $\sigma^2 > 0$. Suppose the prior distribution of θ is selected to be another normal distribution with mean μ_0 and variance ν_0^2. Then the posterior distribution of θ given $X_i = x_i$, for $i = 1, \ldots, n$, is also a normal distribution with mean μ_1 and variance ν_1^2 where

$$\mu_1 = \frac{\sigma^2 \mu_0 + n\nu_0^2 \overline{X}_n}{\sigma^2 + n\nu_0^2} \text{ and } \nu_1^2 = \frac{\sigma^2 \nu_0^2}{\sigma^2 + n\nu_0^2}$$

Likelihood

Since $X_1, X_2, \ldots, X_n$ is a random sample from a normal distribution with mean θ and variance σ^2; the likelihood function of the data, $f(x|\theta)$, will have the form

$$
\begin{aligned}
f(x|\theta) &= \prod_{i=1}^{n} \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left(x_i - \theta\right)^2\right\} \\
&= \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i - \theta\right)^2\right\}
\end{aligned}
$$

Removing all proportionality constants, i.e. all the terms that are not dependent on θ, this relation simplifies to

$$
f(x|\theta) \propto \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i - \theta\right)^2\right\}
$$

Complete the Square

The likelihood can be further simplified by completing the squares. Applying this principle to $\sum_{i=1}^{n}(x_i-\theta)^2$ with $a_i=1$, $b_i=x_i$, $y=\theta$ and $c=0$ then,

$$\begin{aligned}
\sum_{i=1}^{n}(x_i-\theta)^2 &= \left(\sum_{i=1}^{n}1\right)\left[\theta-\frac{\sum_{i=1}^{n}x_i-0}{\sum_{i=1}^{n}1}\right]^2+\sum_{i=1}^{n}1\left(x_i-\frac{\sum_{i=1}^{n}x_i}{\sum_{i=1}^{n}1}\right)^2 \\
&= n\left(\theta-\frac{1}{n}\sum_{i=1}^{n}x_i\right)^2+\sum_{i=1}^{n}\left(x_i-\frac{1}{n}\sum_{i=1}^{n}x_i\right)^2 \\
&= n\left(\theta-\overline{x}_n\right)^2+\sum_{i=1}^{n}\left(x_i-\overline{x}_n\right)^2
\end{aligned}$$

Likelihood and Prior

Finally using the equation from the last slide, the likelihood function can be expressed as

$$f(x|\theta) \propto \exp\left\{-\frac{1}{2\sigma^2}n(\theta - \overline{x}_n)^2\right\}.$$

The selected prior distribution was a normal distribution with mean μ_0 and variance ν_0^2, which has the form

$$\xi(\theta) \propto \exp\left\{-\frac{1}{2\nu_0^2}(\theta - \mu_0)^2\right\}.$$

Posterior

The posterior, $\xi(\theta|x)$, must satisfy the relation

$$\xi(\theta|x) \propto f(x|\theta)\xi(\theta) \propto \exp\left\{-\frac{1}{2}\left[\frac{n}{\sigma^2}(\theta-\overline{x}_n)^2+\frac{1}{\nu_0^2}(\theta-\mu_0)^2\right]\right\}.$$

Expanding the equation over a common denominator yields

$$\xi(\theta|x) \propto \exp\left\{-\frac{1}{2}\left[\frac{n\nu_0^2(\theta^2-2\theta\overline{x}_n+\overline{x}_n^2)+\sigma^2(\theta^2-2\theta\mu_0+\mu_0^2)}{\sigma^2\nu_0^2}\right]\right\}$$

$$\xi(\theta|x) \propto \exp\left\{-\frac{1}{2}\left[\frac{\theta^2(\sigma^2+n\nu_0^2)-2\theta(\sigma^2\mu_0+n\nu_0^2\overline{x}_n)}{\sigma^2\nu_0^2}\right]\right\}$$

Creating a Recognizable Form

To get the previous expression of the posterior into a recognizable form, the following steps are taken:

$$
\begin{aligned}
\xi(\theta|x) &\propto \exp\left\{-\frac{1}{2}\left[\frac{1/(\sigma^2+n\nu_0^2)}{1/(\sigma^2+n\nu_0^2)}\left(\frac{\theta^2(\sigma^2+n\nu_0^2)-2\theta(\sigma^2\mu_0+n\nu_0^2\overline{x}_n)}{\sigma^2\nu_0^2}\right)\right]\right\} \\
&= \exp\left\{-\frac{1}{2}\left[\frac{\theta^2-2\theta\frac{\sigma^2\mu_0+n\nu_0^2\overline{x}_n}{\sigma^2+n\nu_0^2}}{\frac{\sigma^2\nu_0^2}{\sigma^2+n\nu_0^2}}\right]\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left[\frac{\left(\theta-\frac{\sigma^2\mu_0+n\nu_0^2\overline{x}_n}{\sigma^2+n\nu_0^2}\right)^2}{\frac{\sigma^2\nu_0^2}{\sigma^2+n\nu_0^2}}\right]\right\}
\end{aligned}
$$

Remark

Why is

$$\exp\left\{-\frac{1}{2}\left[\frac{\theta^2 - 2\theta\frac{\sigma^2\mu_0 + n\nu_0^2\overline{x}_n}{\sigma^2+n\nu_0^2}}{\frac{\sigma^2\nu_0^2}{\sigma^2+n\nu_0^2}}\right]\right\} \propto \exp\left\{-\frac{1}{2}\left[\frac{\left(\theta - \frac{\sigma^2\mu_0 + n\nu_0^2\overline{x}_n}{\sigma^2+n\nu_0^2}\right)^2}{\frac{\sigma^2\nu_0^2}{\sigma^2+n\nu_0^2}}\right]\right\}?$$

Consider the case where we are interested in a parameter ψ and there is some constant c that does not depend on ψ. Then,

$$(\psi - c)^2 = \psi^2 - 2\psi c + c^2 \propto \psi^2 - 2\psi c$$

So let $\psi = \theta$ and let $c = \frac{\sigma^2\mu_0 + n\nu_0^2\overline{x}_n}{\sigma^2+n\nu_0^2}$.

Reparameterize and Identify the Distribution

Let $\mu_1 = \frac{\sigma^2\mu_0 + n\nu_0^2\overline{X}_n}{\sigma^2 + n\nu_0^2}$ and $\nu_1^2 = \frac{\sigma^2\nu_0^2}{\sigma^2 + n\nu_0^2}$. But then,

$$\xi(\theta|x) \propto \exp\left\{-\frac{1}{2\nu_1^2}(\theta - \mu_1)^2\right\}$$

which is recognizable as a normal distribution with mean μ_1 and variance ν_1^2, thus proving the theorem.

Definitions

Definition (Precision)

The precision ϕ of a normal distribution with mean μ and variance σ^2 is defined as the reciprocal of the variance; that is, $\phi = 1/\sigma^2$.

Definition (Normal p.d.f using precision)

If a random variable has a normal distribution with mean μ and precision ϕ, then its p.d.f is given by

$$f(x|\mu,\phi) = \left(\frac{\phi}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{\phi}{2}(x-\mu)^2\right\} \text{ for } -\infty < x < \infty$$

Theorem 2

Theorem (Gamma Precision Conjugate Prior)

Let $X_1, \ldots, X_n$ be a random sample from a normal distribution with known mean μ and unknown precision ϕ. Suppose the prior distribution on ϕ is selected to be Gamma(a, b). Then the posterior distribution of ϕ given $X_i = x_i$, for $i = 1, \ldots, n$, is also a Gamma distribution with parameters α and β for

$$\alpha = a + \frac{n}{2} \text{ and } \beta = b + \frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2$$

Likelihood and Prior

The likelihood function of the data, $f(x|\phi)$, will be

$$
\begin{aligned}
f(x|\phi) &= \prod_{i=1}^{n} \left(\frac{\phi}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{\phi}{2}(x_i-\mu)^2\right\} \\
&= \left(\frac{\phi}{2\pi}\right)^{\frac{n}{2}} \exp\left\{-\frac{\phi}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right\} \\
&\propto \phi^{n/2} \exp\left\{-\phi\left(\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\right\}.
\end{aligned}
$$

The selected prior distribution was a Gamma(a, b), which has the form

$$
\xi(\phi) = \frac{b^a}{\Gamma(a)}\phi^{a-1}\exp\{-\phi b\} \propto \phi^{a-1}\exp\{-\phi b\}.
$$

Posterior

Thus, it follows that the posterior $\xi(\phi|x)$ must satisfy the relation

$$\begin{aligned}
\xi(\phi|x) &\propto f(x|\phi)\xi(\phi) \\
&\propto \phi^{n/2} \exp\left\{-\phi\left(\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\right\}\phi^{a-1}\exp\{-\phi b\} \\
&= \phi^{(a+n/2)-1}\exp\left\{-\phi\left(b+\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\right\}
\end{aligned}$$

Play Name that Distribution

Letting $\alpha = a + n/2$ and $\beta = b + \frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2$, makes the posterior distribution have the form

$$\xi(\phi|x) \propto \phi^{\alpha-1}\exp\{-\phi\beta\}$$

which is recognizable as a Gamma distribution with parameters α and β, hence proving the theorem.

Simulation

A R Shiny App was created to show the effect on how the choice of sample size, prior mean, and prior variance effect the posterior distribution for the normal mean normal conjugate prior.

Questions?

The End