Functions of Random Variables

Given a random variable X with some known distribution, we are interested in analyzing the random variable $f(X)$ by looking at its expected value $E[f(X)]$ and its variance $\text{var}[f(X)]$. For example consider X having a uniform distribution on $(-\pi/2, \pi/2)$, and let $f(x) = \cos(x)$. What can we say about the expected value and variance of the random variable $f(X) = \cos(X)$?

- We can get an idea by generating data in R. Generate `reps` realizations of $\cos(X)$. Plot a histogram of the data, as well as a plot of the curve $f(x) = \cos(x)$ for $-\pi/2 \le x \le \pi/2$:

```
## Simulate Data
reps = 1e5
y.vec = cos(runif(n=reps, min=-pi/2, max=pi/2))
hist(y.vec)

## plot f(x) = cos(x)
x=seq(from=-pi/2, to=pi/2, by=0.01)
fx = cos(x)
plot(x, fx, type="l")
```

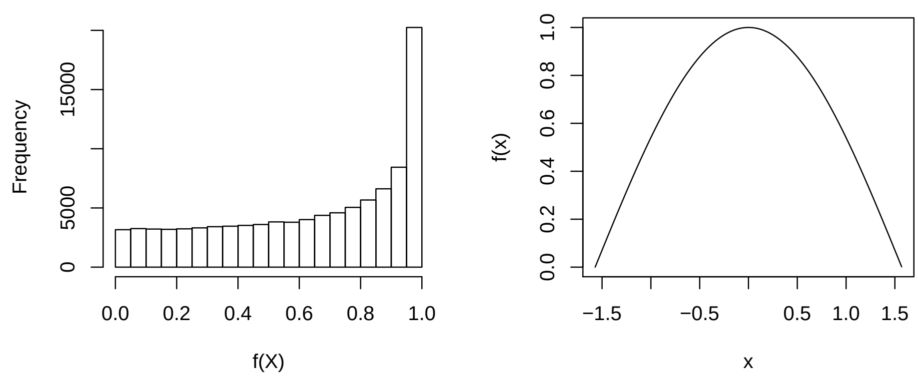

Figure 1: (left) Histogram of `reps` realizations of $f(X)$; (right) A plot of $f(x) = \cos(x)$ on $[-\pi/2, \pi/2]$.

- We can simulate an estimate of of the exact values of $E[\cos(X)]$ and $\text{var}[\cos(X)]$in R:

```
> est.mean = mean(y.vec)
> est.mean
[1] 0.6364099
> est.var = var(y.vec)
> est.var
[1] 0.09481122
```

- How close are these values to the actual value of $E[\cos(X)]$ and $\mathrm{var}[\cos(X)]$? We can check the math:

$$\begin{aligned}
E[\cos(X)] &= \int_{-\pi/2}^{\pi/2} \cos(x)\frac{1}{\pi}dx = \frac{1}{\pi}\sin(x)|_{x=-\pi/2}^{x=\pi/2} \\
&= \frac{1}{\pi}(1-(-1)) = \frac{2}{\pi} \approx 0.6366198 \\
E[\cos(X)^2] &= \int_{-\pi/2}^{\pi/2} \cos(x)^2\frac{1}{\pi}dx = \frac{1}{\pi}\left(\frac{1}{2}\cos(x)\sin(x) + x/2|_{x=-\pi/2}^{x=\pi/2}\right) \\
&= \frac{1}{\pi}(\pi/4 - (-\pi/4)) = \frac{1}{2} \\
\mathrm{var}[\cos(X)] &= E[\cos(X)^2] - E[\cos(X)]^2 \\
&= \frac{1}{2} - \left(\frac{2}{\pi}\right)^2 = 1/2 - \frac{4}{\pi^2} \approx 0.094715
\end{aligned}$$

Linear and Quadratic Approximations:

- Suppose we wished to approximate $f(x) = \cos(x)$ with a quadratic function g around the point $E[X] = 0$. Applying a second order Taylor series expansion we have:

$$g(x) = f(E[X]) + (x - E[X])f'(E[X]) + \frac{1}{2}(x - E[X])^2 f''(E[X])$$

  Plugging in: $f'(E[X]) = -\sin(E[X])$ and $f''(E[X]) = -\cos(E[X])$, and $E[X] = 0$ we have a quadratic function:

$$\begin{aligned}
g(x) &= \cos(0) + (x-0)(-\sin(0)) + \frac{1}{2}(x-0)^2(-\cos(0)) \\
g(x) &= 1 - \frac{1}{2}x^2
\end{aligned}$$

- Suppose we wished to approximate $f(x) = \cos(x)$ with a linear function h around the point $E[X] = 0$. Applying a first order Taylor series expansion we have:

$$h(x) = f(E[X]) + (x - E[X])f'(E[X])$$

  Plugging in: $f'(E[X]) = -\sin(E[X])$ and $E[X] = 0$ we have a linear function:

$$\begin{aligned}
h(x) &= \cos(0) + (x-0)(-\sin(0)) \\
h(x) &= 1
\end{aligned}$$

- How good are these quadratic and linear functions $g(x), h(x)$ at approximating $f(x) = \cos(x)$ on the interval $[-\pi/2, \pi/2]$? We can plot all three curves to see how well our approximations do:

```
## plot f(x) = cos(x) and g(x) = 1-0.5 x^2
x=seq(from=-pi/2, to=pi/2, by=0.01)
fx = cos(x)
gx = 1-0.5*x^2
hx = array(1, length(x) )
plot(x, fx, type="l", lty = 1)
lines(x, gx, lty=2)
lines(x, hx, lty=3)
legend(x=1, y=1, legend=c("f(x)", "g(x)", "h(x)"), lty=c(1,2,3))
```

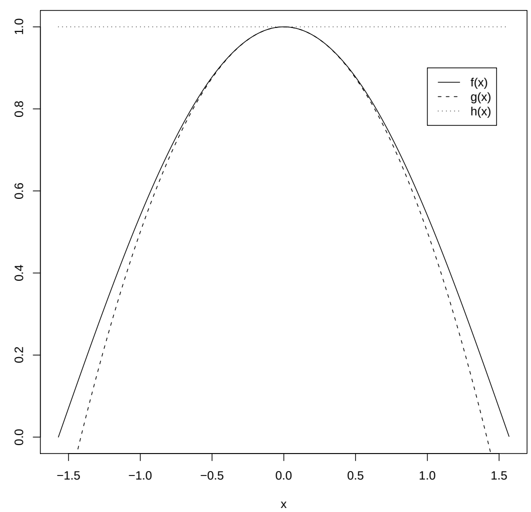

Figure 2: A plot of $f(x) = \cos(x)$, $g(x) = 1 - 0.5x^2$, and $h(x) = 1$ on $[-\pi/2, \pi/2]$.

We can see that $g(x)$, the quadratic approximation to $f(x) = cos(x)$ is pretty good, while $h(x)$, the linear approximation is not as good.

Exercises

For all exercises let X have a uniform distribution on $(-\pi/2, \pi/2)$, and let $f(x) = \cos(x)$

1. Consider the approximation:

$$\text{var} f(X) \approx \text{var}(X) f'(EX)^2$$

Calculate the approximate result, and an unbiased estimate of the exact value using simulation. Also calculate the "relative error", which is $|(A - E)/E|$, where A is the approximate result and E is the exact result from the simulation.

Solution:
The approximation of $\mathrm{var} f(X)$ is found by plugging in $EX = 0$, $\mathrm{var}(X) = \pi^2/12$, and $f'(EX) = -\sin(EX)$:

$$A = (\pi^2/12)(-\sin(0))^2 = 0$$

```
reps=1e5
A=0
E = var(cos(runif(n=reps, min=-pi/2, max=pi/2)))
rel.err = abs((A-E)/E)
```

2. Consider the crude approximation:

$$E[f(X)] \approx f(EX)$$

Estimate the exact value of $Ef(X)$ using simulation and compare it to the crude approximation.

Solution:
The approximation of $E[f(X)]$ is found by plugging in $EX = 0$:

$$A = \cos(0) = 1$$

```
reps=1e5
A=cos(0)
E = mean(cos(runif(n=reps, min=-pi/2, max=pi/2)))
c(A,E)
```

3. Consider a better approximation:

$$E[f(X)] \approx f(E[X]) + \frac{1}{2}\mathrm{var}[X]f''(E[X])$$

Estimate the exact value of $Ef(X)$ using simulation and compare it to the better approximation.

Solution:
The approximate result is found by plugging in $EX = 0$, $\mathrm{var}(X) = \pi^2/12$, and $f''(EX) = -\cos(EX)$:

$$A = \cos(0) + 0.5(\pi^2/12)(-\cos(0)) = 1 - \pi^2/24$$

```
reps=1e5
A=1-pi^2/24
E = mean(cos(runif(n=reps, min=-pi/2, max=pi/2)))
c(A,E)
```

Relating these approximations to the Taylor expansions

These three approximations are related to the first and second order Taylor expansions. Both the crude approximation of $Ef(X)$ and the approximation of $\mathrm{var} f(X)$ come from the linear approximation $h(X)$ of $f(X)$:

$$E[f(X)] \approx E[h(X)] = E\left[f(E[X]) + (X - E[X])f'(E[X])\right] = f(E[X])$$

and

$$\begin{aligned}\mathrm{var} f(X) \approx \mathrm{var}[h(X)] &= \mathrm{var}\left[f(E[X]) + (X - E[X])f'(E[X])\right] \\ &= \mathrm{var}[f(E[X])] + \mathrm{var}\left[(X - E[X])f'(E[X])\right] = \mathrm{var}(X)f'(E[X])^2\end{aligned}$$

These approximations were not very good for the example we gave, as the function f was far from linear on the interval $[-\pi/2, \pi/2]$.

The better approximation of $Ef(X)$ came from the quadratic approximation $g(X)$ of $f(X)$:

$$\begin{aligned}E[f(X)] \approx E[g(X)] &= E\left[f(E[X]) + (X - E[X])f'(E[X]) + \frac{1}{2}(X - E[X])^2 f''(E[X])\right] \\ &= E[f(E[X])] + E[(X - E[X])f'(E[X])] + \frac{1}{2}E[(X - E[X])^2 f''(E[X])] \\ &= f(E[X]) + \frac{1}{2}\mathrm{var}[X]f''(E[X])\end{aligned}$$

This approximation was reasonable for the example we gave, as the function f was close to the approximate function g.

Another Example

Let X have a uniform distribution on $(5, 6)$, and let $f(x) = \log(x)$. Simulate data to find the exact value of $E \log(X)$ and $\text{var}(\log(X))$, then compare it to the approximations. Note that $f'(x) = 1/x$ and $f''(x) = -\frac{1}{x^2}$, $E[X] = 5.5, \text{var}(X) = 1/12$.

$$\text{var} f(X) \approx \text{var}(X) f'(EX)^2$$

```
## Variance Approximation/Exact with relative error
reps=1e5
A= (1/12)*(1/5.5)^2
E = var(log(runif(n=reps, min=5, max=6)))
rel.err = abs((A-E)/E)
```

$$E[f(X)] \approx f(EX)$$

```
## Crude Expected Value Approximation
reps=1e5
A=log(5.5)
E = mean(log(runif(n=reps, min=5, max=6)))
c(A,E)
```

$$E[f(X)] \approx f(E[X]) + \frac{1}{2}\text{var}[X] f''(E[X])$$

```
## Second order Expected Value Approximation
reps=1e5
A=log(5.5) + 0.5*(1/12)*(-1/(5.5)^2)
E = mean(log(runif(n=reps, min=5, max=6)))
c(A,E)
```

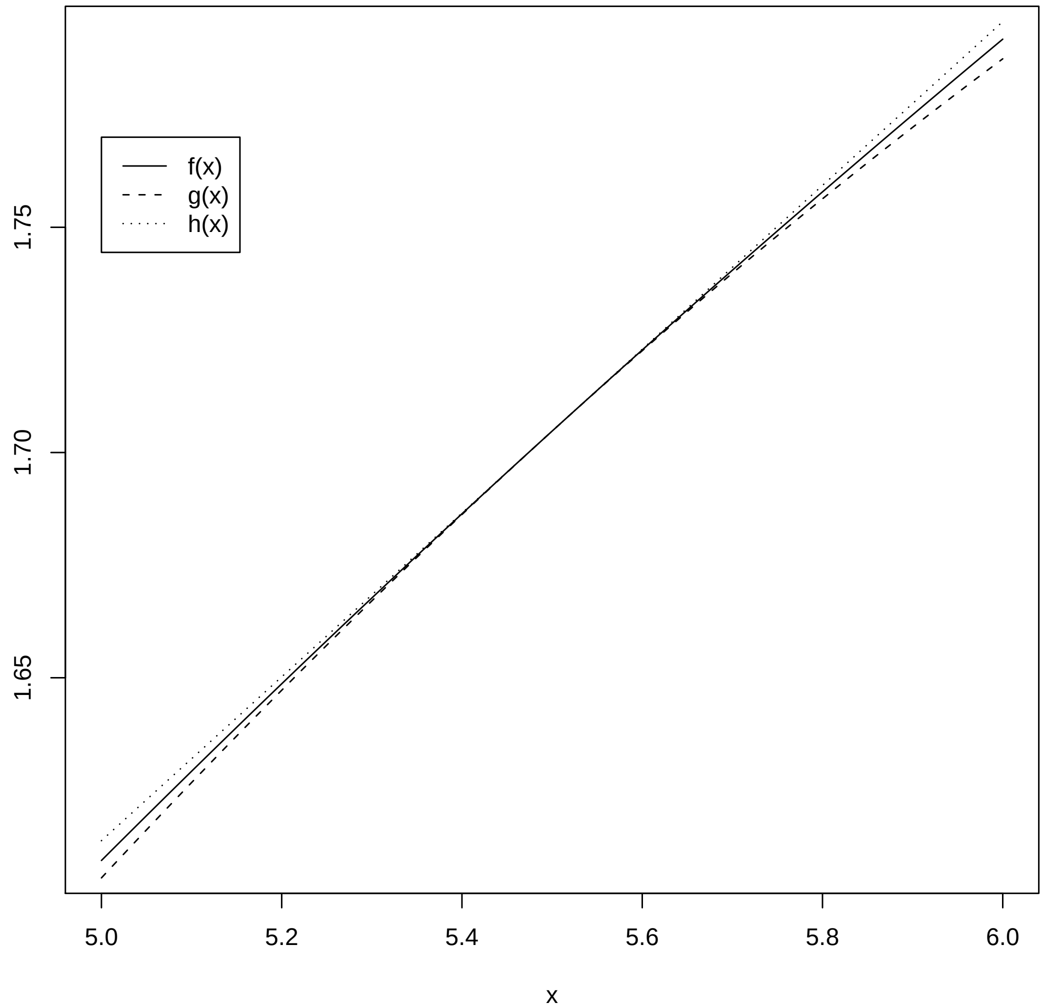

Figure 3: A plot of $f(x) = \log(x)$, $g(x) = \log(5.5) + (x - 5.5)/5.5 - 0.5(x - 5.5)^2/5.5^2$, and $h(x) = \log(5.5) + (x - 5.5)/5.5$ on $[5, 6]$.