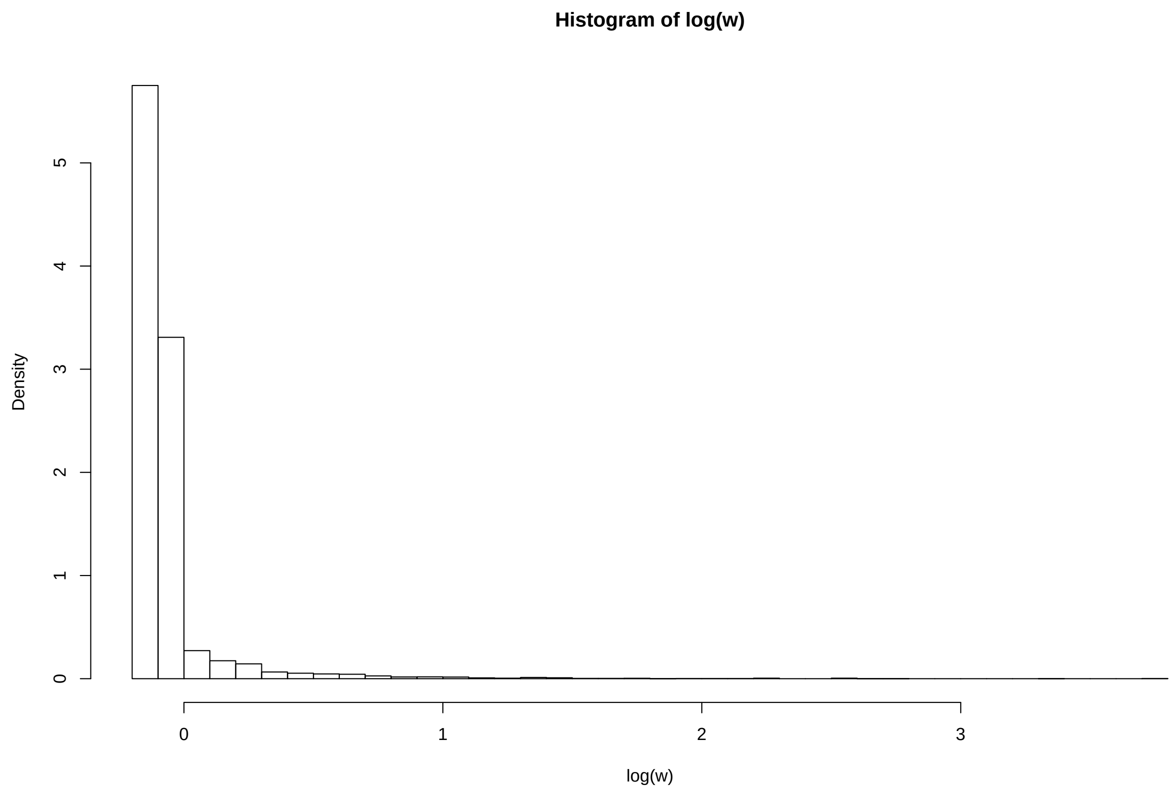
STAT 8700 Homework 7

Brian Detweiler

Friday, October 14, 2016

1. Suppose our target distribution $p(\theta|y)$ is a standard Normal, and we choose $g(\theta)$ to be a t-distribution with 3 degrees of freedom.

```
#2. Repeat Question 1, swapping the
#distributions used for p(??|y) and g(??).

#Suppose our target distribution p(??|y)
#is a now a t-distribution with 3 degrees
#of freedom and we choose g(??) to be a
#standard normal.

#(a) Draw a sample of size S = 100 from g(??)
#and compute the importance ratios.
#Plot a histogram of the log importance ratios.

#(b) Estimate E[??|y] and Var(??|y) using
#importance sampling. Compare to the true
#values.


#(c) Repeat the first two parts of
#the question using S = 10000
#g(theta) = standard normal

#g(theta)=standard normal, S=10000
theta=rnorm(10000,0,1)
#target dist = t-dist, dof=3
w = dt(theta,3)/dnorm(theta,0,1)
hist(log(w), br=30,freq=F)
```




```
#mean of w same as expec.valu
sum(theta*w)/sum(w)
```

```
## [1] 0.01717703
```

```
#var of w
expec.valu = sum(theta*w)/sum(w)
sum(((theta-expec.valu)^2)*w)/sum(w)
```

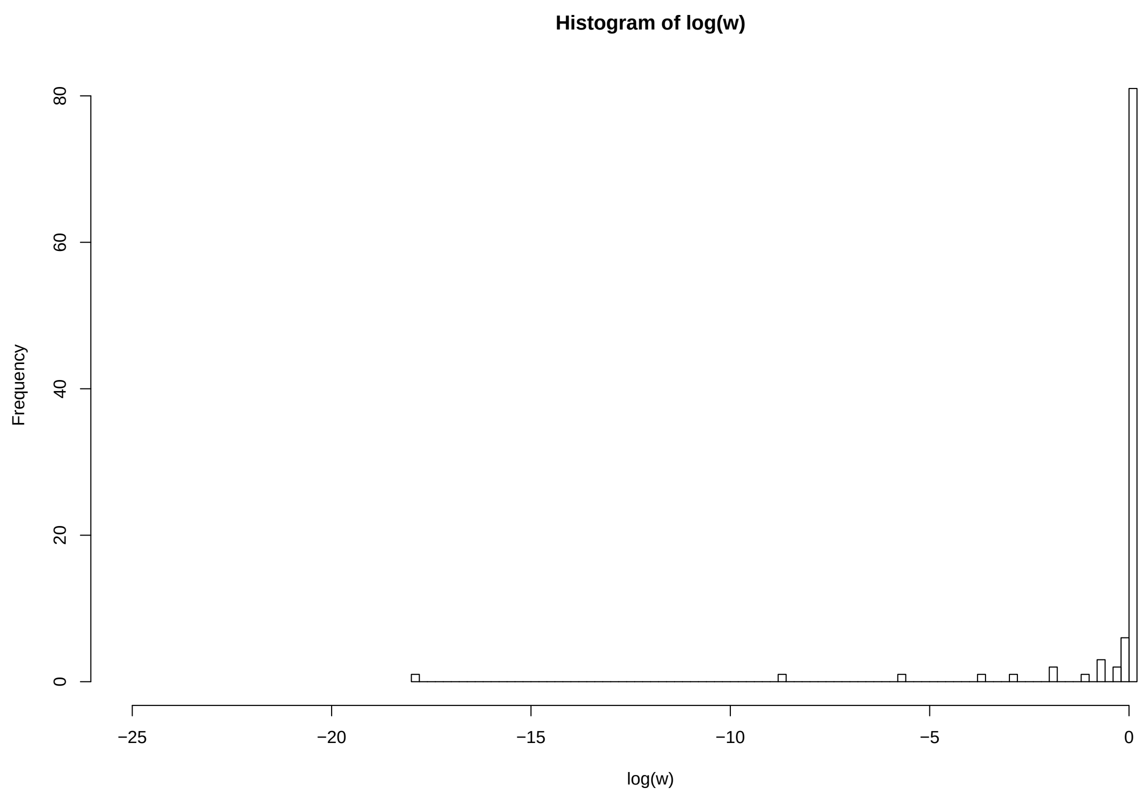
```
## [1] 1.615004
```

```
#(d) Using the sample obtained in the
#previous part, compute an estimate of the
#effective sample size.

normalized_w = w/sum(w)
s_eff=1/sum(normalized_w^2)
s_eff
```

```
## [1] 5816.902
```

(a) Draw a sample of size $S = 100$ from $g(\theta)$ and compute the importance ratios. Plot a histogram of the log importance ratios.

```r
S <- 100
df <- 3
theta <- rt(S, df)

w <- dnorm(theta, 0, 1) / dt(theta, df = df)

hist(log(w), breaks = S, xlim = c(-25, 1))
```



(b) Estimate $E[\theta|y]$ and $Var(\theta|y)$ using importance sampling. Compare to the true values.

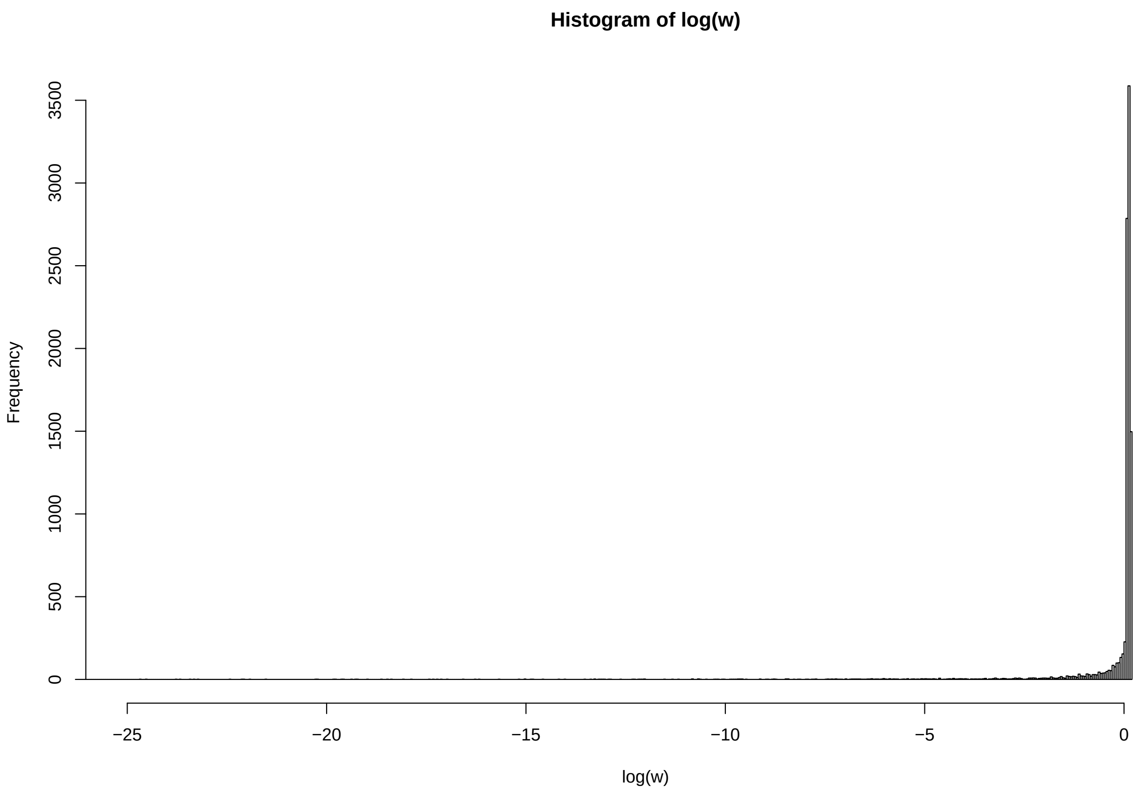
```r
theta.expected <- sum(theta * w) / sum(w)
theta.variance <- sum(((theta - theta.expected)^2) * w) / sum(w)
```

$$E[\theta|y] = 0.0947607$$
$$Var(\theta|y) = 0.9912126$$

■

(c) Repeat the first two parts of the question using $S = 10000$.

```
S <- 10000
df <- 3
theta <- rt(S, df)

w <- dnorm(theta, 0, 1) / dt(theta, df = df)

hist(log(w), breaks = S, xlim = c(-25, 1))
```




```
theta.expected <- sum(theta * w) / sum(w)
theta.variance <- sum(((theta - theta.expected)^2) * w) / sum(w)
```

$$E[\theta|y] = -0.0085259$$
$$Var(\theta|y) = 0.9914361$$

■

(d) Using the sample obtained in the previous part, compute an estimate of the effective sample size.

```r
w_tilde <- w / sum(w)
Seff <- 1 / sum(w_tilde^2)
```

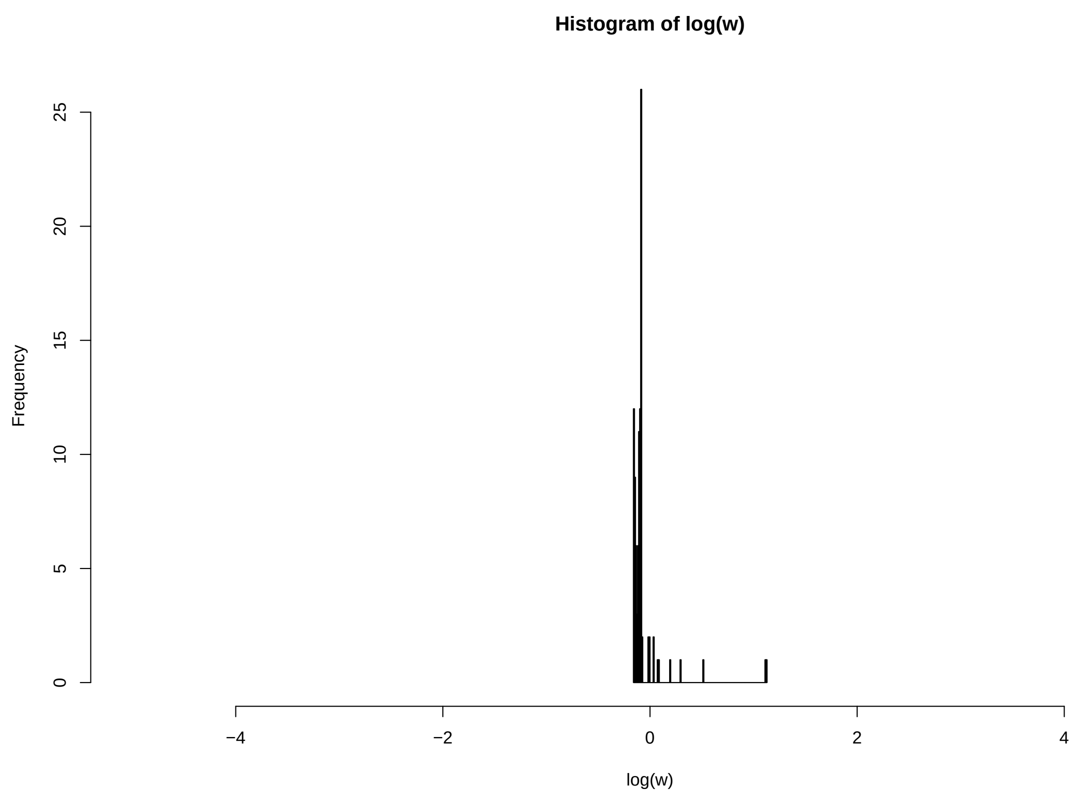
The effective sample size is 9216.854558.

■

2. Repeat Question 1, swapping the distributions used for $p(\theta|y)$ and $g(\theta)$.

(a) Draw a sample of size $S = 100$ from $g(\theta)$ and compute the importance ratios. Plot a histogram of the log importance ratios.

```r
S <- 100
df <- 3
theta <- rnorm(S, 0, 1)

w <- dt(theta, df = df) / dnorm(theta, 0, 1)

hist(log(w), breaks = S, xlim = c(-5, 5))
```



■

(b) Estimate $E[\theta|y]$ and $Var(\theta|y)$ using importance sampling. Compare to the true values.

```
theta.expected <- sum(theta * w) / sum(w)
theta.variance <- sum(((theta - theta.expected)^2) * w) / sum(w)
```

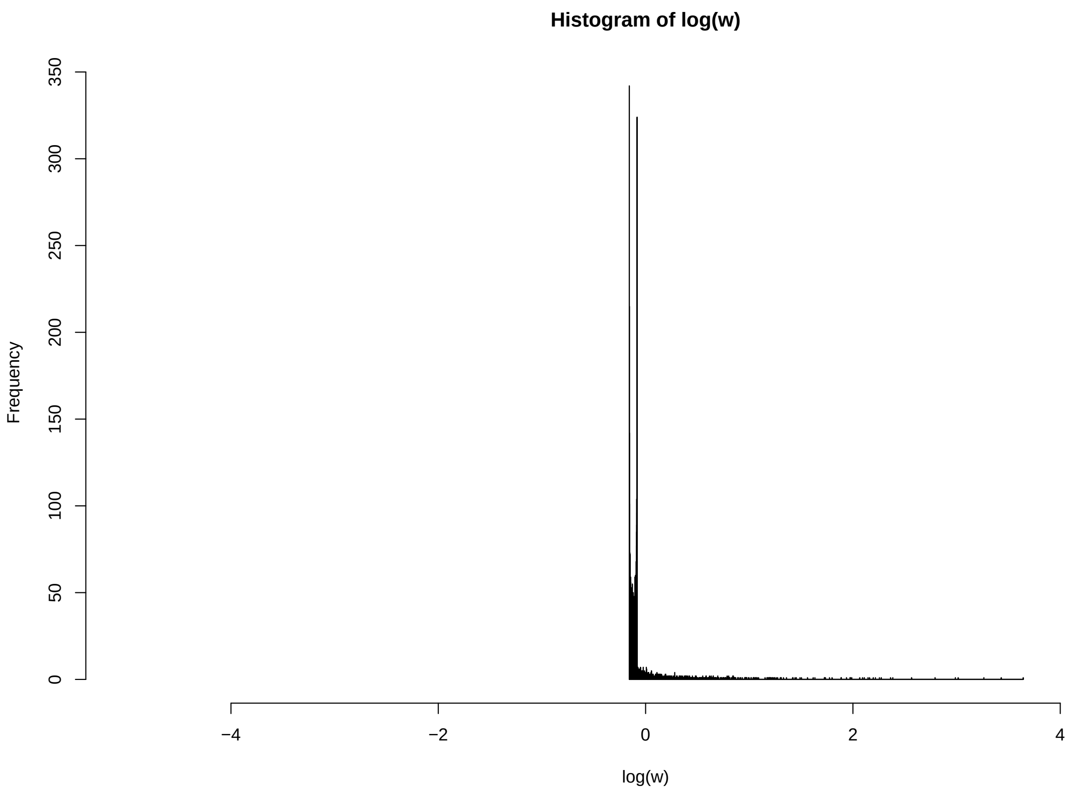
$$E[\theta|y] = 0.067346$$
$$Var(\theta|y) = 1.2887199$$

The true mean and variance are $\mu = 0$ and $\sigma^2 = 1$. The estimates are off by 0.067346 and -0.2887199, respectively, which are reasonably close.

■

(c) Repeat the first two parts of the question using $S = 10000$.

```
S <- 10000
df <- 3
theta <- rnorm(S, 0, 1)

w <- dt(theta, df = df) / dnorm(theta, 0, 1)

hist(log(w), breaks = S, xlim = c(-5, 5))
```




```
theta.expected <- sum(theta * w) / sum(w)
theta.variance <- sum(((theta - theta.expected)^2) * w) / sum(w)
```

$$E[\theta|y] = 0.0433438$$
$$Var(\theta|y) = 1.5689019$$

The true mean and variance are $\mu = 0$ and $\sigma^2 = 1$. The estimates are off by 0.0433438 and -0.5689019, respectively, which are reasonably close.

■

(d) Using the sample obtained in the previous part, compute an estimate of the effective sample size.

```
w_tilde <- w / sum(w)
Seff <- 1 / sum(w_tilde^2)
```

The effective sample size is 6236.7340525.

■

3. **Use the Metropolis algorithm to obtain simulated values of *alpha* and *beta* from the Bioassay example in Section 3.7. Be sure to define your starting points and your jumping rule. Compute with log-densities (see page 261). Run your simulations long enough for approximate convergence. Include a plot of the last 25% of your simulated values overlaying the contour plot.**

■