

Prior Sensitivity from Importance Sampling

Robert McCulloch

Prior Sensitivity from Importance Sampling

Robert McCulloch

3/25/2021

Prior Sensitivity from Importance Sampling

Let's do a simple example of reweighting draws from a posterior from one prior to get expectation with respect to a posterior from another prior.

We will use the simple

$[Y_i \sim \text{Bernoulli}(\theta)]$
model. So " (θ) " is " (p) ".

Data And Uniform Prior

We will assume that whether or not the Leafs score on a powerplay is iid Bernoulli.

Out of 25 powerplays, they have scored 7 times.

What is the probability they score on the next powerplay?

```
## data, Leafs scored on 7 out of 25 powerplays.  
y=7  
n=25
```

What prior should we use?

We will start with the uniform prior on (0,1) which is a Beta(1,1),

```
## g, g will be post from uniform prior  
ag = 1; bg=1 # g will be the post from the beta(ag,bg) conjugate prior, which is uniform
```

Lets' draw from the posterior and plot the prior and the posterior.

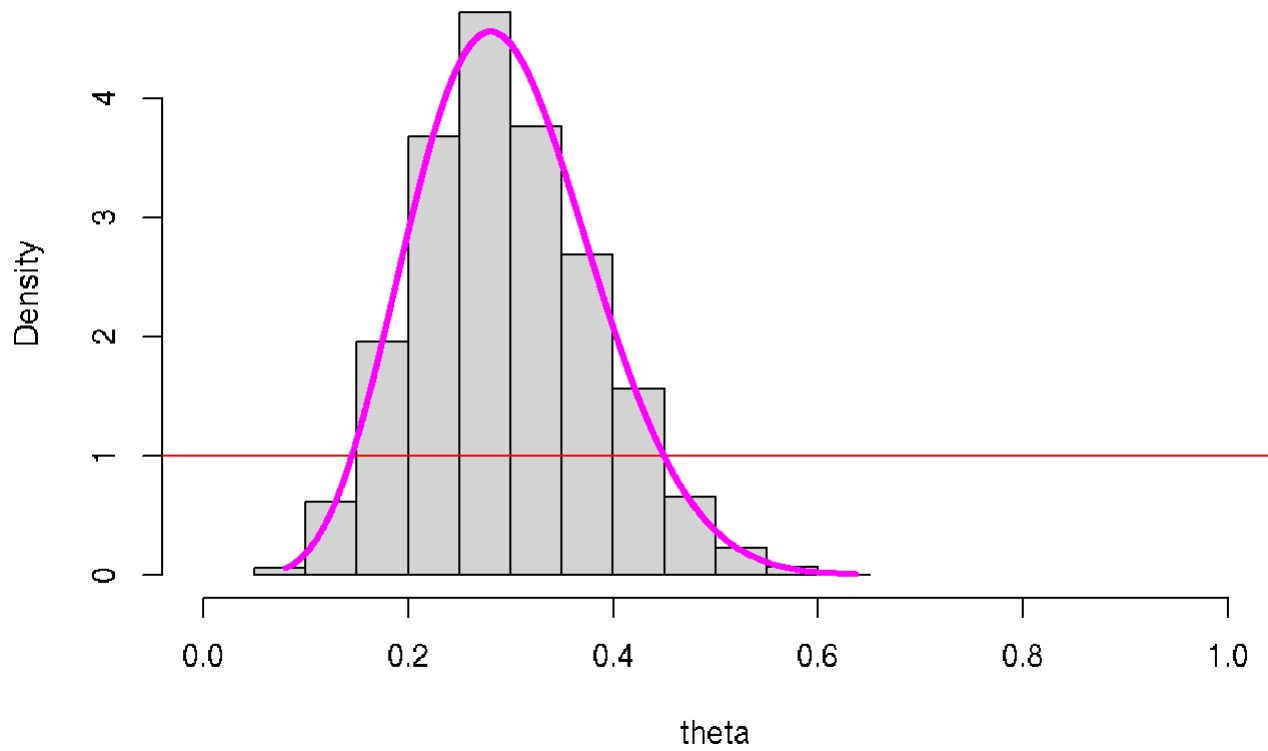
```
## draw from g  
set.seed(99)
```

```
m = 5000
theta = rbeta(m,ag+y,bg+n-y)

cat("my from g: ",mean(theta),"\n")
## my from g: 0.2960294
cat("check against formula:", (ag+y)/(ag+bg+n),"\n")
## check against formula: 0.2962963

hist(theta,prob=TRUE,xlim=c(0,1))
abline(h=1,col='red')
stheta = sort(theta)
lines(stheta,dbeta(stheta,ag+y,bg+n-y),col='magenta',lwd=3)
```

Histogram of theta



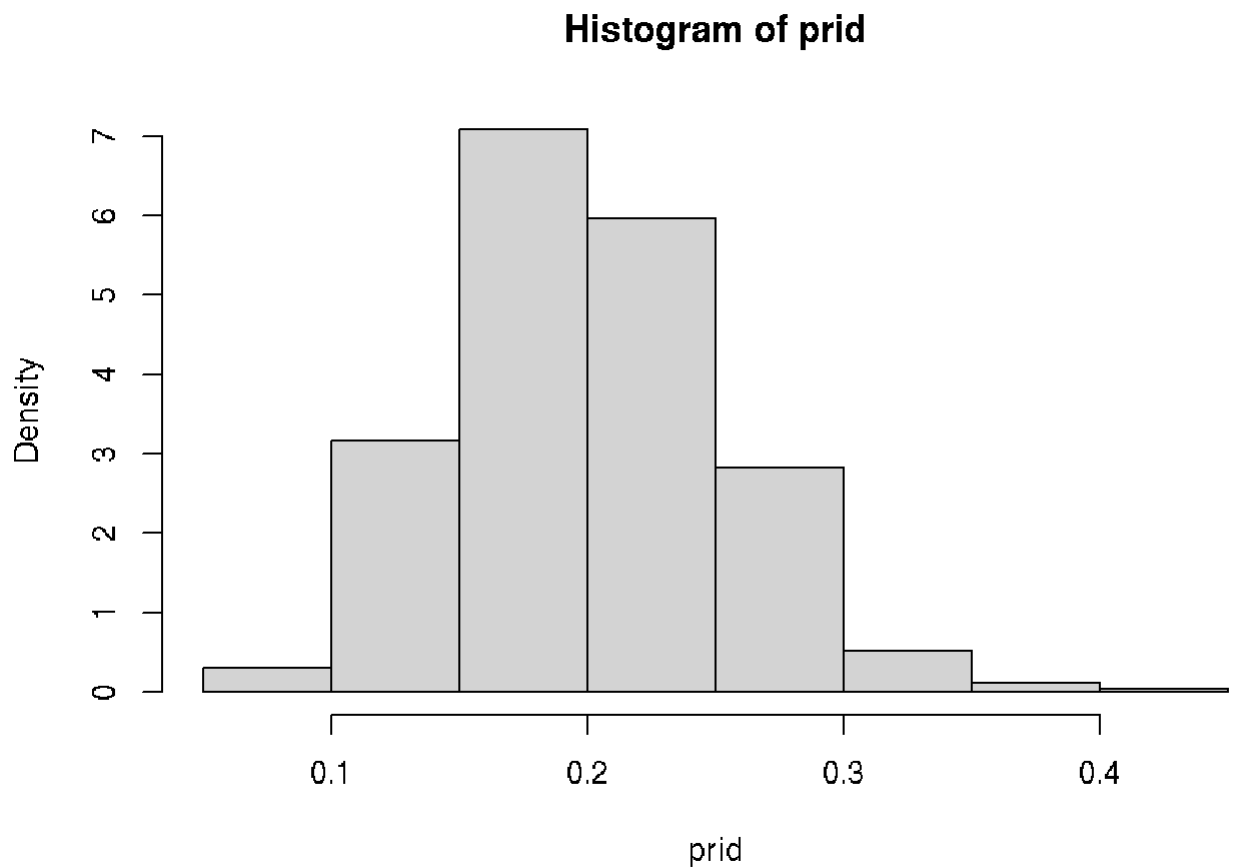
More Serious Prior

The uniform prior is ridiculous.

We know θ is much less than .5!!

Let's consider a more reasonable Beta prior.

```
## more serious beta prior
apri = 10
bpri = 40
prid = rbeta(1000,apri,bpri)
hist(prid,prob=TRUE)
```



Now, since we again are using a conjugate Beta prior, we can easily compute the posterior.

But let's compute the posterior mean using importance sampling and then check

our results against the correct result.

So, we will do prior sensitivity where our initial prior is the uniform (1,1) prior, and our new prior is the (10,40) prior. We have draws from the posterior from prior 1, and would like to compute the posterior mean under prior2, using the draws from the prior 1 posterior and importance sample.

The weights are just the ratios of the two priors in the normalized version of importance sampling!!

Remember the density of prior 1 is just 1.

In the unnormalized version I will have to use the ratio of the two posteriors.

```
## compute weights for adjust g prior to more serious prior
## remember, g(theta) = 1, so f/g is just f, where f is the more serious prior
#normalized weights
wnor = dbeta(theta,apri,bpri) #raw weights (note other prior is uniform), f/g
wnor = wnor/sum(wnor) #normalize weights

#unnormalized weights, ratio of posterior using (apri,bpri) and posterior using (ag,bg)
## note that for the unnormalized I use the posterior ratios, for the normalized I used the
w = dbeta(theta,apri+y,bpri+n-y)/dbeta(theta,ag+y,bg+n-y)
```

Ok, let's see how we do.

```
## importance sampling estimates
muhatnor = sum(wnor*theta)
muhat = sum(w*theta)/length(theta)

## since the prior is conjugate we can check
apost = apri+7; bpost = bpri + 18
mutrue = apost/(apost+bpost)

cat("muhat",muhat,"\n") #estimate for unnormalized importance sampling
## muhat 0.2287034
cat("muhatnor",muhatnor,"\n") #estimate from normalized importance sampling
## muhatnor 0.2269946
cat("mutrue",mutrue,"\n") # true posterior mean from conjugate formulas
## mutrue 0.2266667
```

Pretty good !!

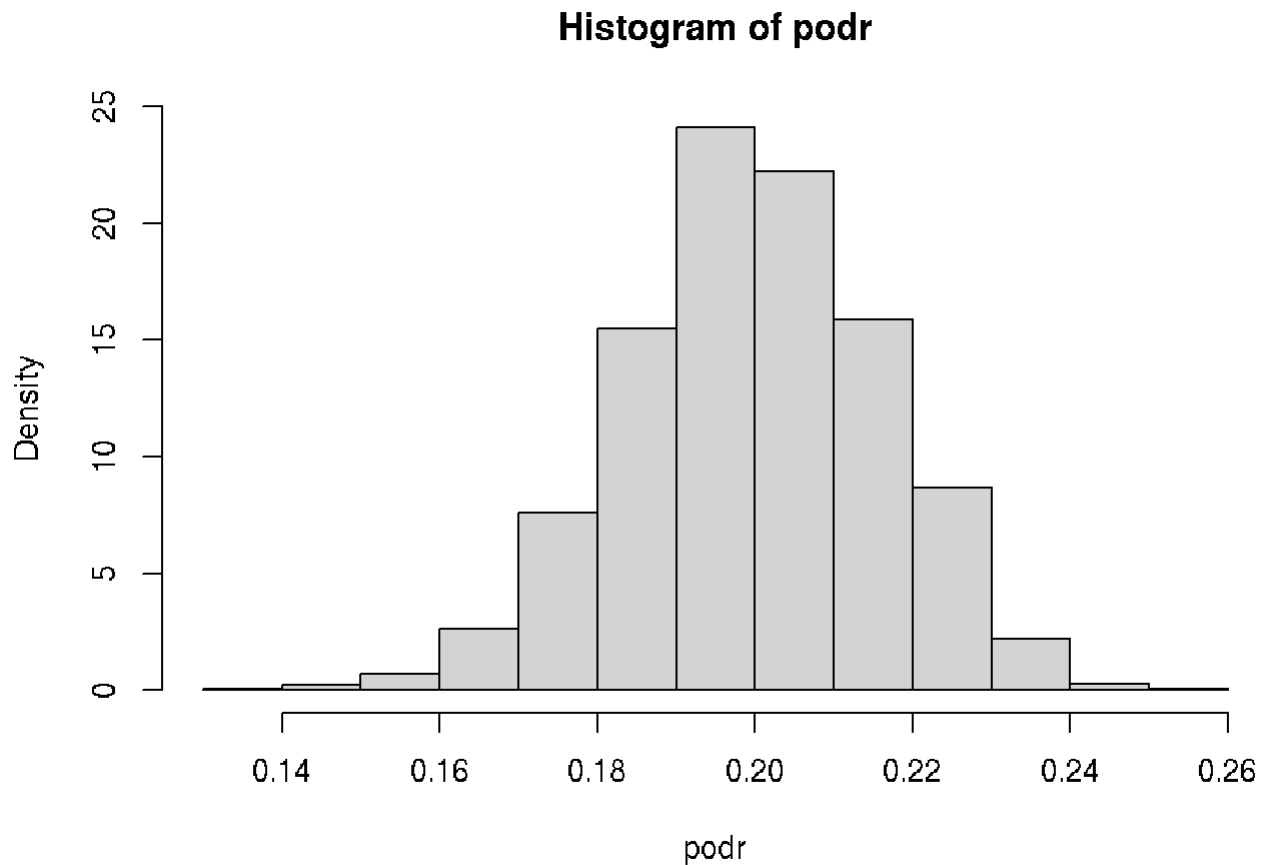
now let's do a non-conjugate prior

We will put an informative prior on the odds ratio $\theta/(1-\theta)$ and then compute the posterior mean under this prior using normalized importance sampling.

We get a prior on θ from the prior on the odds, but this prior is certainly not of the conjugate Beta form.

Recall, $o = \frac{\theta}{1-\theta}$, $\theta = \frac{o}{1+o}$

```
## now let's do a non-conjugate prior
## odds ratio prior, the odds ratio is normal.
mu = .25; sigma = .05/2.0
odr = rnorm(5000,mu,sigma) # draw from prior on odds
podr = odr/(1+odr) # draw from prior on p
hist(podr,prob=TRUE)
```



I want to get the posterior expectation of θ under this new prior using the draws from the uniform prior.

This is what we just did with the informative Beta prior, but this time we don't have a simple alternative way to compute the answer!!

Using the change of variable formula $p(\theta) = n(o(\theta)) |do/d\theta| = n(o(\theta)) \frac{1}{(1-\theta)^2}$

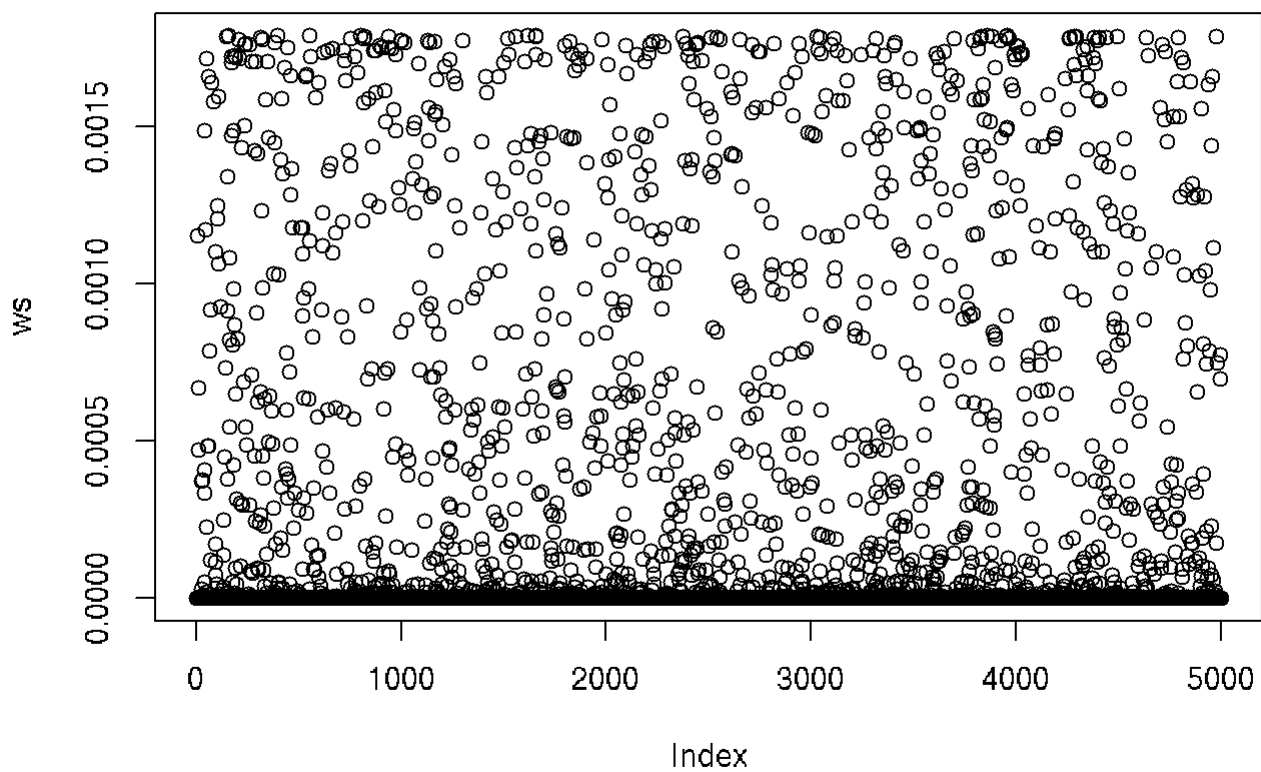
where $\phi(\cdot)$ is the normal density.

```
## normalized weights
wnor = dnorm(theta/(1-theta),mu,sigma)/(1-theta)^2
ws = wnor/sum(wnor)

##estimate
muhatnoropri = sum(ws*theta)
cat("muhatnoropri",muhatnoropri,"\n")
## muhatnoropri 0.2026163
```

That was easy!!

```
## have a look at the weights
plot(ws)
```



The zeros are a “waste” but we don’t have any crazy big weights. The weights

are just the value of our new prior.

But how do check our posterior mean?

In this simple case we can just discretize the problem to get a good idea of the true posterior in this non-conjugate case.

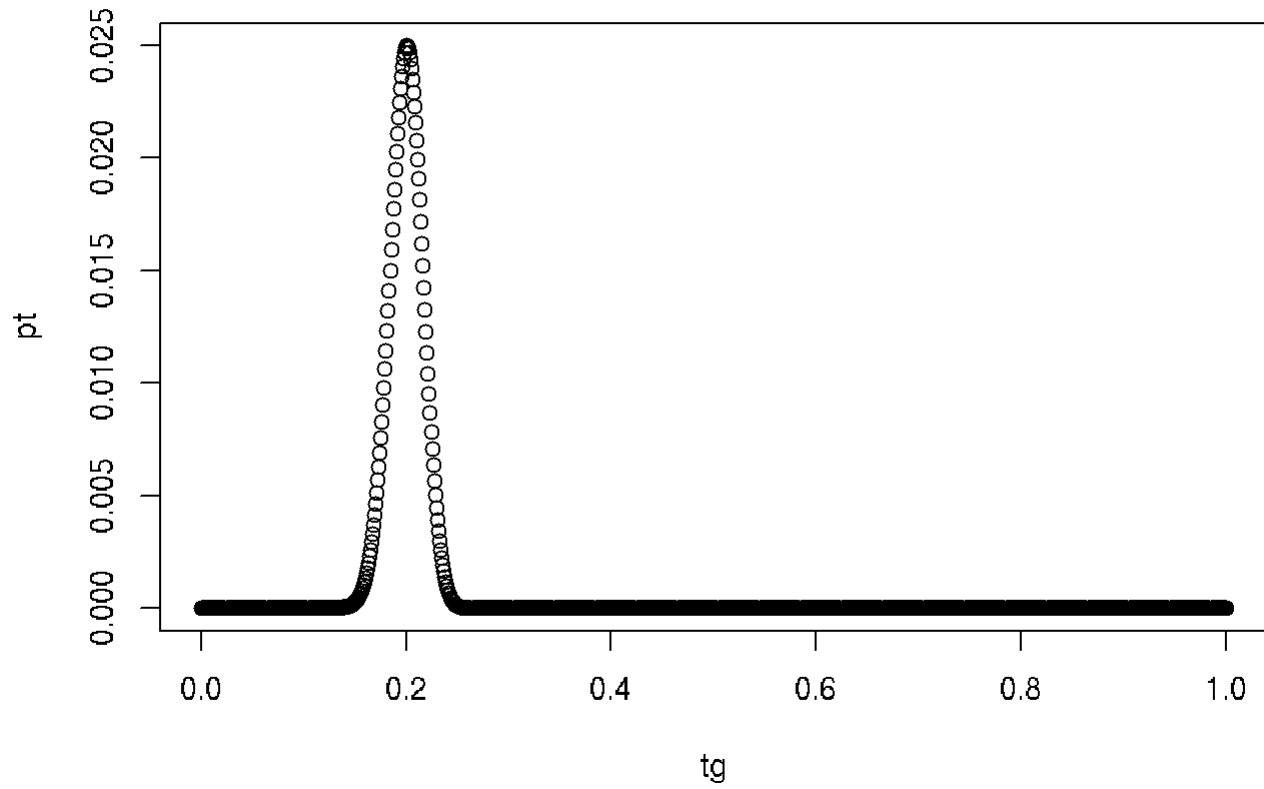
In a more complex example, you code very carefully, check it any way you can and pray.

Discretization of the Non-conjugate Prior

Discretize the prior.

```
#prior on grid
tg = seq(from =.00001, to=.99999,length.out=1000)
pt = dnorm(tg/(1-tg),mu,sigma)/(1-tg)^2
pt = pt/sum(pt)
plot(tg,pt)
title(main="odds prior on grid")
```

odds prior on grid

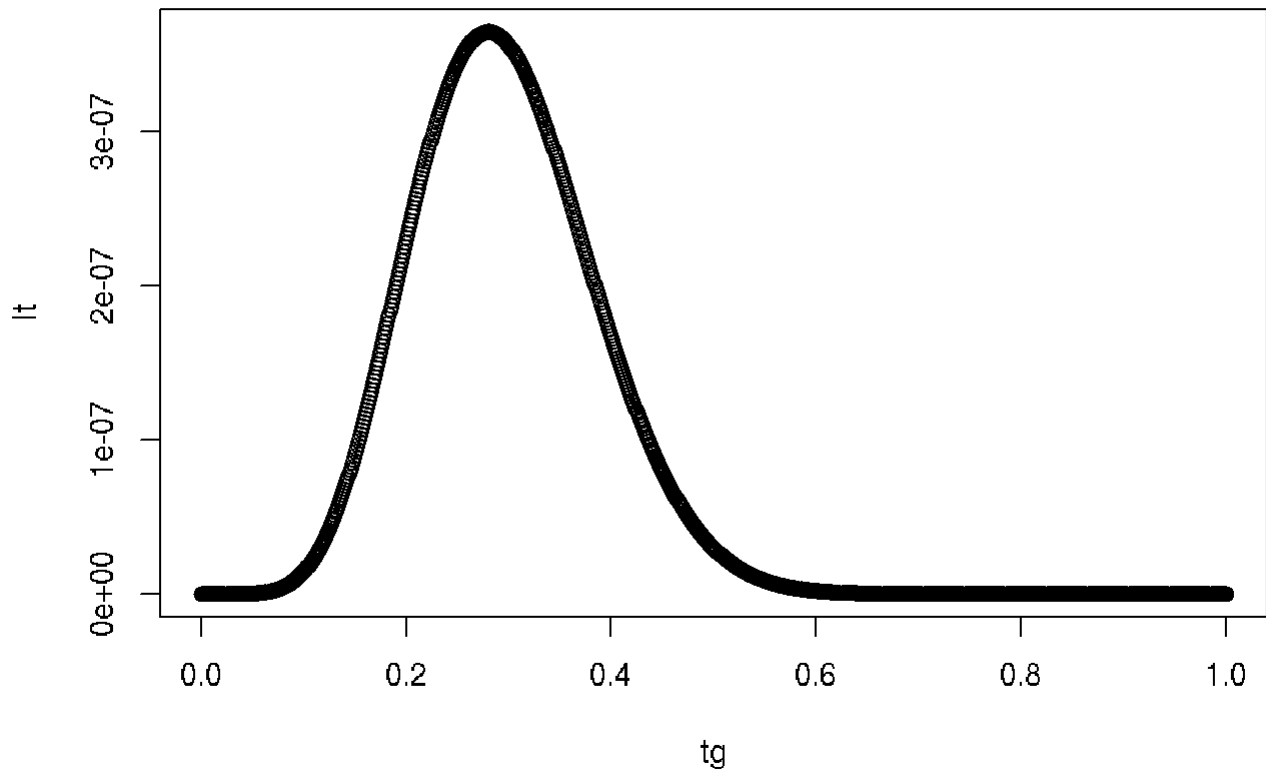


Evaluate the likelihood on the grid.

```
# likelihood on grid
## data, Leafs scored on 7 out of 25 powerplays.
y=7
n=25

lt = (tg^y)*((1-tg)^(n-y))
plot(tg,lt)
title(main="likelihood on grid")
```

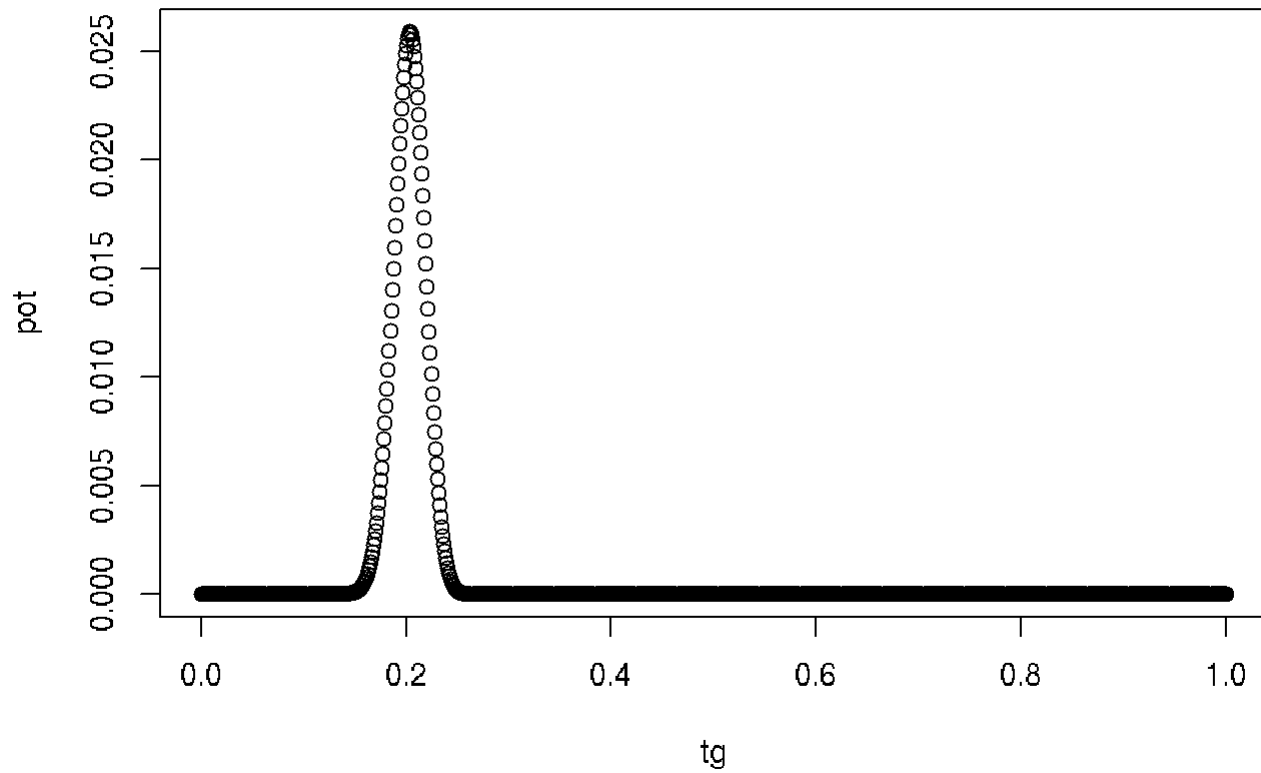
likelihood on grid



Compute the posterior on the grid.

```
## posterior on grid  
pot = pt*lt;  
pot = pot/sum(pot)  
plot(tg,pot)  
title(main="posterior on grid")
```

posterior on grid



Compute the posterior mean from the discretized solution.

```
#expected value  
Et = sum(pot*tg)  
cat("expected value from discretized version: ",Et,"\n")  
## expected value from discretized version: 0.2028056
```

which is the same as what we got from the importance sampling.

Plot the inference.

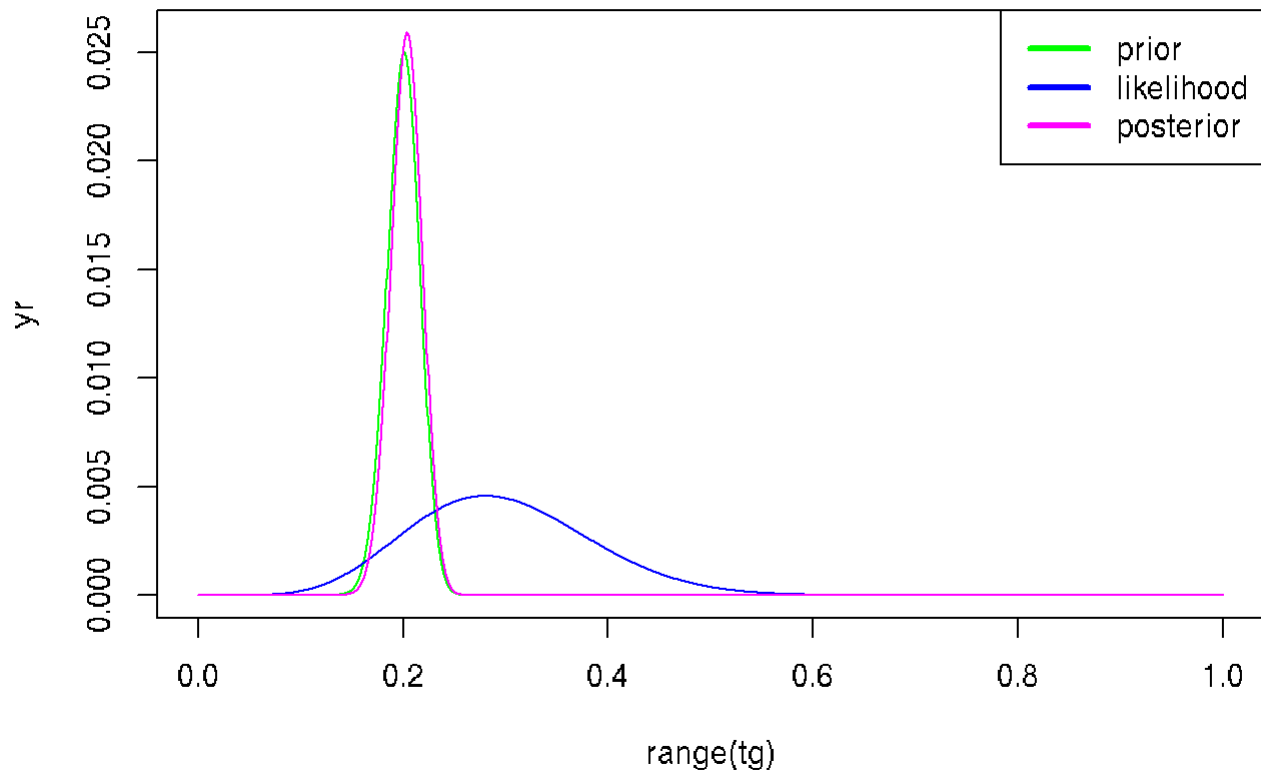
```
## plot inference  
ltnor = lt/sum(lt)  
  
yr = range(c(pt,pot,lt))  
plot(range(tg),yr,type="n")  
lines(tg,pt,col='green')
```

```

lines(tg, ltnor, col="blue")
lines(tg, pot, col="magenta")

legend("topright", legend=c("prior", "likelihood", "posterior"), col=c("green", "blue", "magenta"))

```



In this case, the prior is quite a bit more informative than the likelihood, which fully captures the information in the data.

SIR

Let's get draws from the posterior using Sampling-Importance-Resampling.

We just have to get iid draws from the theta values we drew from $g = \text{posterior}$ from uniform prior using the normalized weights for the odds ratio prior.

The draws are in `theta` and the weights are in `ws`.

I'll get posterior draws from SIR (`tdsir` below) and posterior draws from the discretized version (`tddis` below) and then compare them using a qqplot.

```
ndsir = 5000
tdsir = sample(theta, ndsir, replace = TRUE, prob = ws)
tddis = sample(tg, ndsir, replace=TRUE, prob=pot)
qqplot(tdsir, tddis)
abline(0,1, col="red")
```

