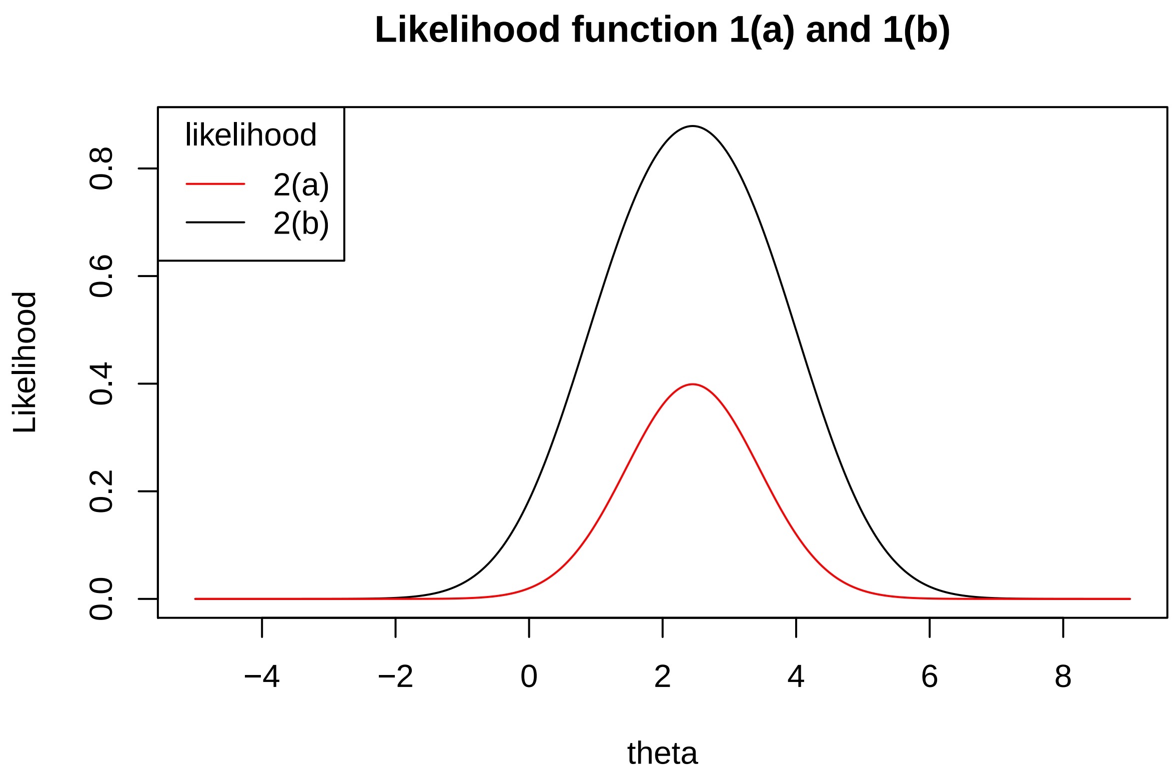
440a1

2c



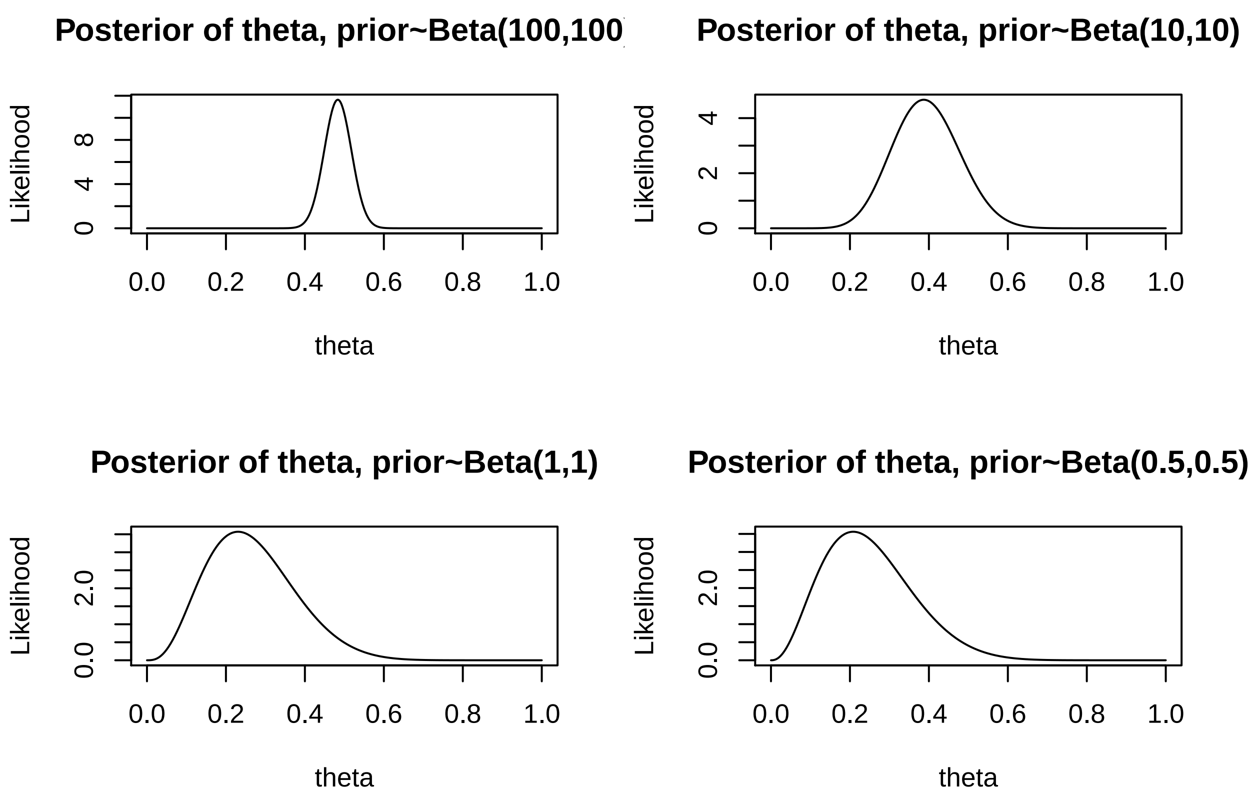
we observe the 2 MLE produces same $\hat{\theta}$, are both symetric around 2.45.

but 2(b) has a higher likelihood at every nontrival value θ than 2(a)

this is because it's easier (higher probability) to obtain a value with in a range than getting a specif value

3

```
pBeta=function(theta, n){
  dbeta(theta, 3+n, 10+n)
}
Theta=seq(0,1,0.001)
par(mfrow=c(2,2))
plot(Theta,pBeta(Theta,100),type="l",xlab="theta",ylab="Likelihood",main="Posterior of theta, prior~Beta
plot(Theta,pBeta(Theta,10),type="l",xlab="theta",ylab="Likelihood",main="Posterior of theta, prior~Beta
plot(Theta,pBeta(Theta,1),type="l",xlab="theta",ylab="Likelihood",main="Posterior of theta, prior~Beta(
plot(Theta,pBeta(Theta,0.5),type="l",xlab="theta",ylab="Likelihood",main="Posterior of theta, prior~Beta
```




```
Theta[which.max(pBeta(Theta,100))]
```

```
## [1] 0.483
```

```
Theta[which.max(pBeta(Theta,10))]
```

```
## [1] 0.387
```

```
Theta[which.max(pBeta(Theta,1))]
```

```
## [1] 0.231
```

```
Theta[which.max(pBeta(Theta,0.5))]
```

```
## [1] 0.208
```

5(b) confidence interval

```r
help(qgamma)
```

```
## starting httpd help server ... done
```

```r
qgamma(0.025, shape=219, rate=112)
```

```
## [1] 1.704943
```

```r
qgamma(0.975, shape=219, rate=112)
```

```
## [1] 2.222679
```

```r
qgamma(0.025, shape=68, rate=45)
```

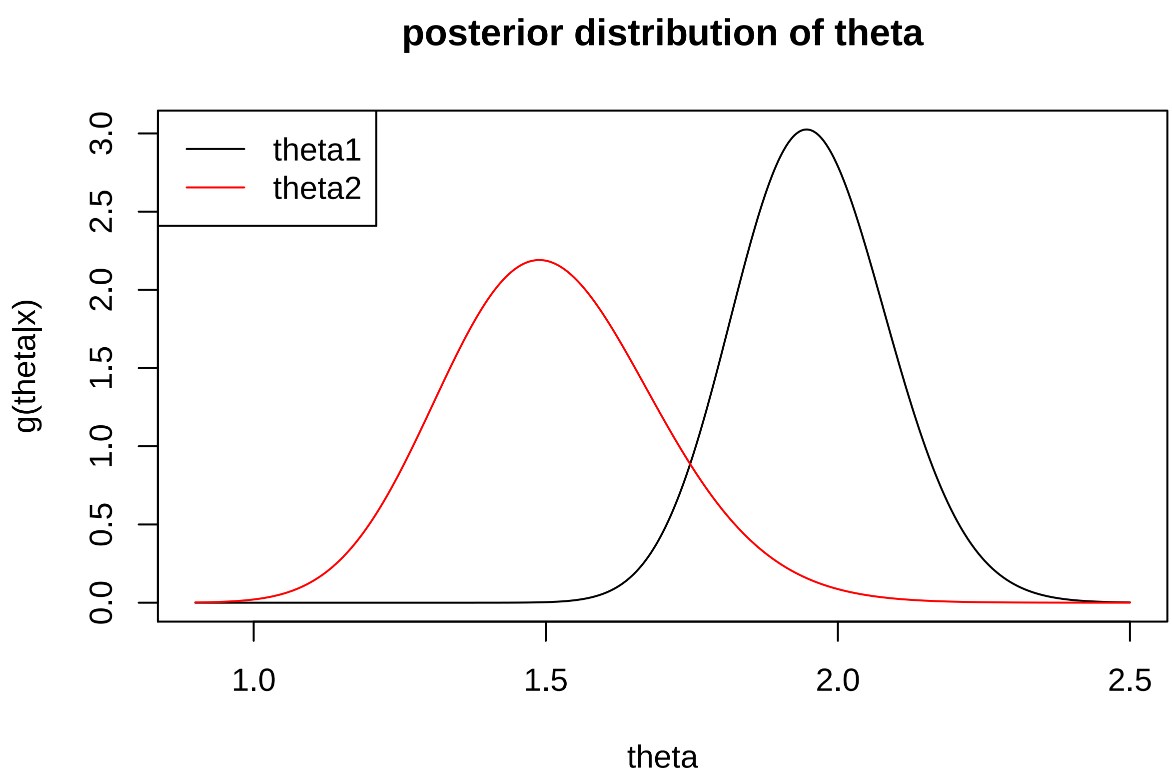
```
## [1] 1.173437
```

```r
qgamma(0.975, shape=68, rate=45)
```

```
## [1] 1.890836
```

5(c)

```r
Theta=seq(0.9,2.5,0.001)
plot(Theta,dgamma(Theta, 219, 112),type="l",xlab="theta",ylab="g(theta|x)",main="posterior distribution
points(Theta,dgamma(Theta, 68, 45),lwd=1,type="l",col="red")
legend("topleft",
       legend = c("theta1", "theta2"),
       lty = c(1, 1),
       col = c("black", "red"))
```



we conclude that θ_1 is likely to be larger than θ_2, both are very slightly left skewed, $\theta 2$ has a larger span than θ_1

5d

```
num_obs=1000000
# generate g(theta|x)
x=rgamma(num_obs,219, 112)
# generate g(theta|y)
y=rgamma(num_obs ,68, 45)
# number of trails that x>y
d=x>y
# computed probility of x>y
sum(d)/num_obs
```

```
## [1] 0.972419
```

so it's quite evident $\theta_1 > \theta_2$

5e

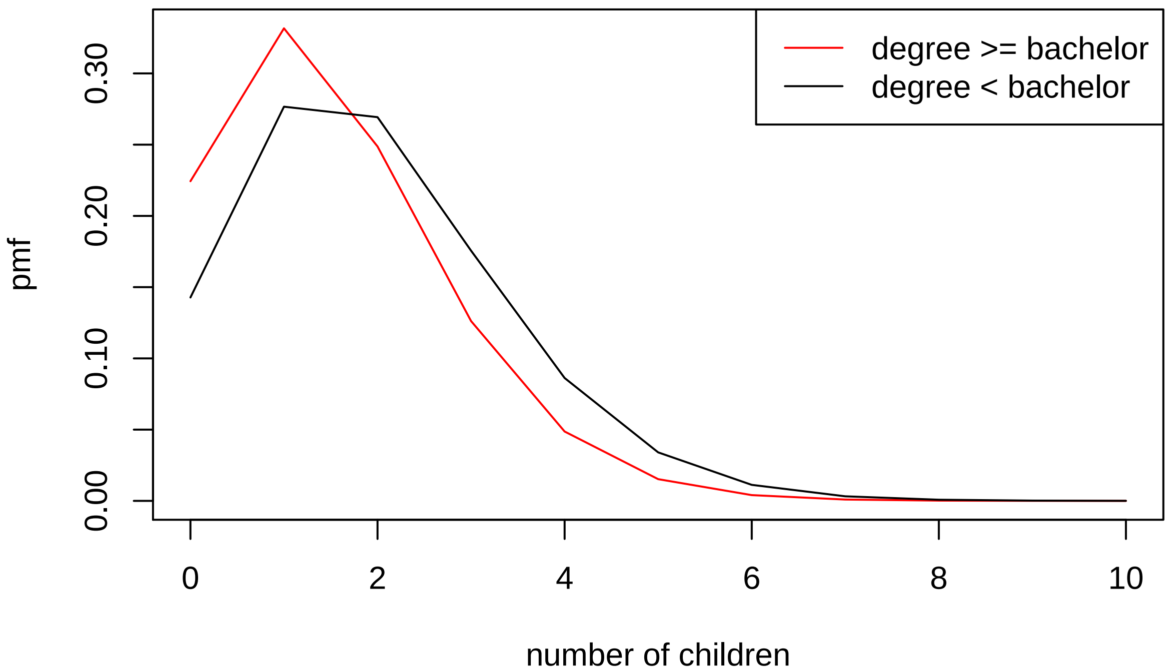
```
help("rnbinom")
X=rnbinom(num_obs,219,112/113)
Y=rnbinom(num_obs,68,45/46)
d=X>Y
sum(d)/num_obs
```

```
## [1] 0.482007
```

we see we expect 48.23% possibility of observing one without a bachelor degree have more children at 40 than one with a bachelor degree.

5f

```
Theta=seq(0,10,1)
plot(Theta,dnbinom(Theta, 68, 45/46),type="l",xlab="number of children",ylab="pmf",main="",col="red")
points(Theta,dnbinom(Theta, 219, 112/113),lwd=1,type="l",col="black")
legend("topright",
       legend = c("degree >= bachelor", "degree < bachelor"),
       lty = c(1, 1),
       col = c("red", "black"))
```



we observe number of child's pmf for those with bachelor degree is a leftshift of those without.

those with bachelor degree is likely to have smaller number (<=4) of children

$\theta_1 > \theta_2$ is a relationship between distribution parameters, it means the expected number of children is larger for people without bachelor degree.

$X^f > Y^f$ is just a comparison between 2 r.v. outcome. it means a random person without bachelor degree has more children than a random person with bachelor degree

6

```r
#a
sqrt(2*pi)
```

```
## [1] 2.506628
```

```r
find.theta.cv=function(n){
  x=runif(n,min=-10,max=10)
  theta=exp(-x^2/2)*20
  theta.star=(1-abs(x)/5)

  theta.hat=mean(theta)
  theta.star.hat=mean(theta.star)

  theta.star.real = 0

  cov.hat=sum((theta-theta.hat)*(theta.star-theta.star.hat))/(n*(n-1))
  var.hat.theta.start.hat=sum((theta.star-theta.star.hat)^2)/(n*(n-1))

  alpha=-cov.hat/var.hat.theta.start.hat

  theta.cv=theta+alpha*(theta.star-theta.star.real)
  mean(theta.cv)
}
means=c()
for (i in 1:50){
  means=c(means, find.theta.cv(500))
}
#average estimate
mean(means)
```

```
## [1] 2.519476
```

```r
#variance of estimate
var(means)
```

```
## [1] 0.03229462
```

```r
#b
find.theta.cv=function(n){
  x=runif(n,min=-10,max=10)
  theta=exp(-x^2/2)*20
  theta.star=(1-x^2/25)

  theta.hat=mean(theta)
  theta.star.hat=mean(theta.star)

  theta.star.real = -1/3

  cov.hat=sum((theta-theta.hat)*(theta.star-theta.star.hat))/(n*(n-1))
  var.hat.theta.start.hat=sum((theta.star-theta.star.hat)^2)/(n*(n-1))

  alpha=-cov.hat/var.hat.theta.start.hat

  theta.cv=theta+alpha*(theta.star-theta.star.real)
  mean(theta.cv)
```

```
}
means=c()
for (i in 1:50){
  means=c(means, find.theta.cv(500))
}
#average estimate
mean(means)
```

```
## [1] 2.536957
```

```
#variance of estimate
var(means)
```

```
## [1] 0.05474061
```

7

```r
#a
n=1000
quantil=runif(n,pnorm(1),1)
x=qnorm(quantil)
f=dnorm(x)*x^2
g=dnorm(x)/(1-pnorm(1))
mean(f/g)
```

```
## [1] 0.4247239
```

```r
sd(f/g)/sqrt(n)
```

```
## [1] 0.009101798
```

```r
#b
x0=rexp(n,1/2)
x=sqrt(x0+1)

f=dnorm(x)*x^2
g=x*exp((1-x^2)/2)

mean(f/g)
```

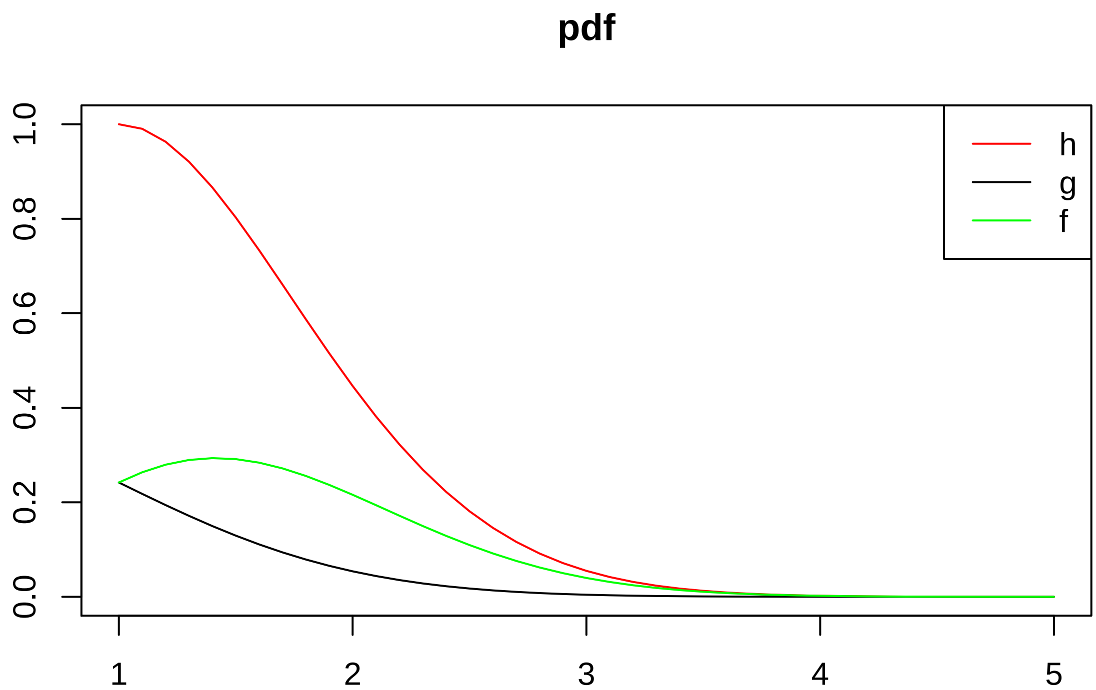
```
## [1] 0.4069952
```

```r
sd(f/g)/sqrt(n)
```

```
## [1] 0.004013599
```

```r
# c
x=seq(1,5,0.1)
f=dnorm(x)*x^2
g=dnorm(x)
h=x*exp((1-x^2)/2)

plot(x,h,type="l",xlab="",ylab="",main="pdf",col="red")
points(x,g,lwd=1,type="l",col="black")
points(x,f,lwd=1,type="l",col="green")
legend("topright",
       legend = c("h", "g", "f"),
       lty = c(1, 1,1),
       col = c("red", "black","green"))
```




```
# we see h looks more like f than g does.
```

8

```r
f=function(x){x^2*cos(x^2)}
#a
x=rexp(10000,25)
mean(f(x))
```

```
## [1] 0.003179886
```

```r
#b
x=runif(5000)
y=(f(qexp(x, 25))+f(qexp(1-x, 25)))/2
mean(y)
```

```
## [1] 0.00313356
```

```r
simple.mean=c()
anti.mean=c()
#c
for (i in (1:1000)){
  x=rexp(10000,25)
  simple.mean=c(simple.mean, mean(f(x)))

  x=runif(10000)
  y=(f(qexp(x, 25))+f(qexp(1-x, 25)))/2
  anti.mean=c(anti.mean, mean(y))
}
mean(simple.mean)
```

```
## [1] 0.003200009
```

```r
mean(anti.mean)
```

```
## [1] 0.003197004
```

```r
var(simple.mean)
```

```
## [1] 4.986208e-09
```

```r
var(anti.mean)
```

```
## [1] 1.937149e-09
```

we find the mean of 2 estimators are roughly the same, but antithetic does produce smaller variance in estimator

9

```r
monte=function(n){
  x=rgeom(n,4/5)
  y=5/4*(x^2+3)^(-7)
  yhat=sum(y)/n
  se=sqrt( (sum(y^2)/n-yhat^2)/n   )
  #95%CI

  c(yhat, 1.96*c(-1,1)*se+yhat)
}

#
findsize=function(n){
  while(1){
    x=rgeom(n,4/5)
    y=5/4*(x^2+3)^(-7)
    yhat=sum(y)/n
    se=sqrt( (sum(y^2)/n-yhat^2)/n   )
    #95%CI
    if(1.96*2*se >=0.0002){
      break
    }
    if(n<5){
      break
    }
    n=round(0.95*n, 0)
  }
  n


}
monte(1000)
```

```
## [1] 0.0004718751 0.0004592506 0.0004844997
```

```r
monte(10000)
```

```
## [1] 0.0004677569 0.0004637245 0.0004717893
```

```r
monte(100000)
```

```
## [1] 0.0004688901 0.0004676194 0.0004701607
```

```r
findsize(1000)
```

```
## [1] 16
```