Mathematical Modeling of Infectious Diseases, UCSF

Homework 5

Exercise 0. What is the generating function of a Bernoulli trial with success probability p? Use the generating function to find its mean and variance. If you add up fixed N of them (independent, identical), what is the generating function? ∎

$$g(u) = E[u^X] = u^0(1-p) + u^1 p = 1 - p + up$$

Mean

```
sage: g = 1 - p + u*p
sage: gprime = diff(g,u).subs(u=1).simplify()
sage: gprime
p
```

Variance

```
sage: gdoubleprime = diff(diff(g,u),u).subs(u=1).simplify()
sage: gdoubleprime
0
sage: gdoubleprime + gprime - gprime^2
-p^2 + p
```

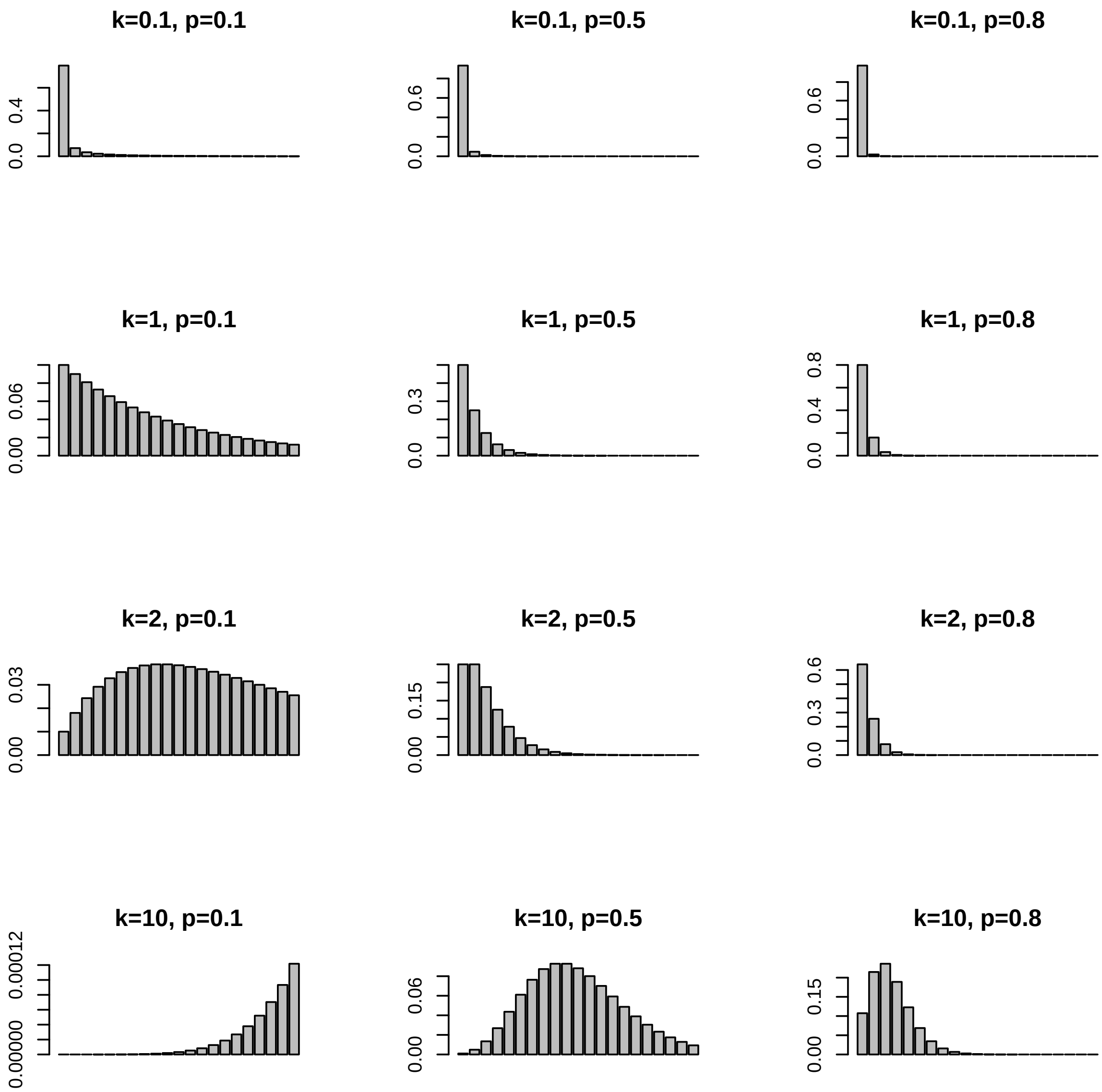
N independent and identical

$$h(u) = (1-p+up)^N$$

Exercise 1. Use R's function `dnbinom` to plot some negative binomial plots; pick $k = 0.1$, $k = 1$, $k = 2$, $k = 10$; try $p = 0.1$, $p = 0.5$, $p = 0.8$; go out further than 20 if you have to (plot out far enough that the tiny probabilities don't show up any more on barplot. Play with it; turn 2 or 3 plots in. ∎

```
kks <- c(0.1,1,2,10)
pps <- c(0.1,0.5,0.8)
par(mfrow=c(4,3))
for(kk in kks){
        for(pp in pps){
        barplot(dnbinom(0:20,kk,pp),main = paste("k=",kk,", ","p=",pp,sep=""))
```

```
        }
}
```



Exercise 2. Suppose that a random variable takes the value 2 with probability 1. Write the generating function, and from it, find the mean and variance. ∎

$g(u) = E[u^X] = u^2$

$g'(u) = 2u$

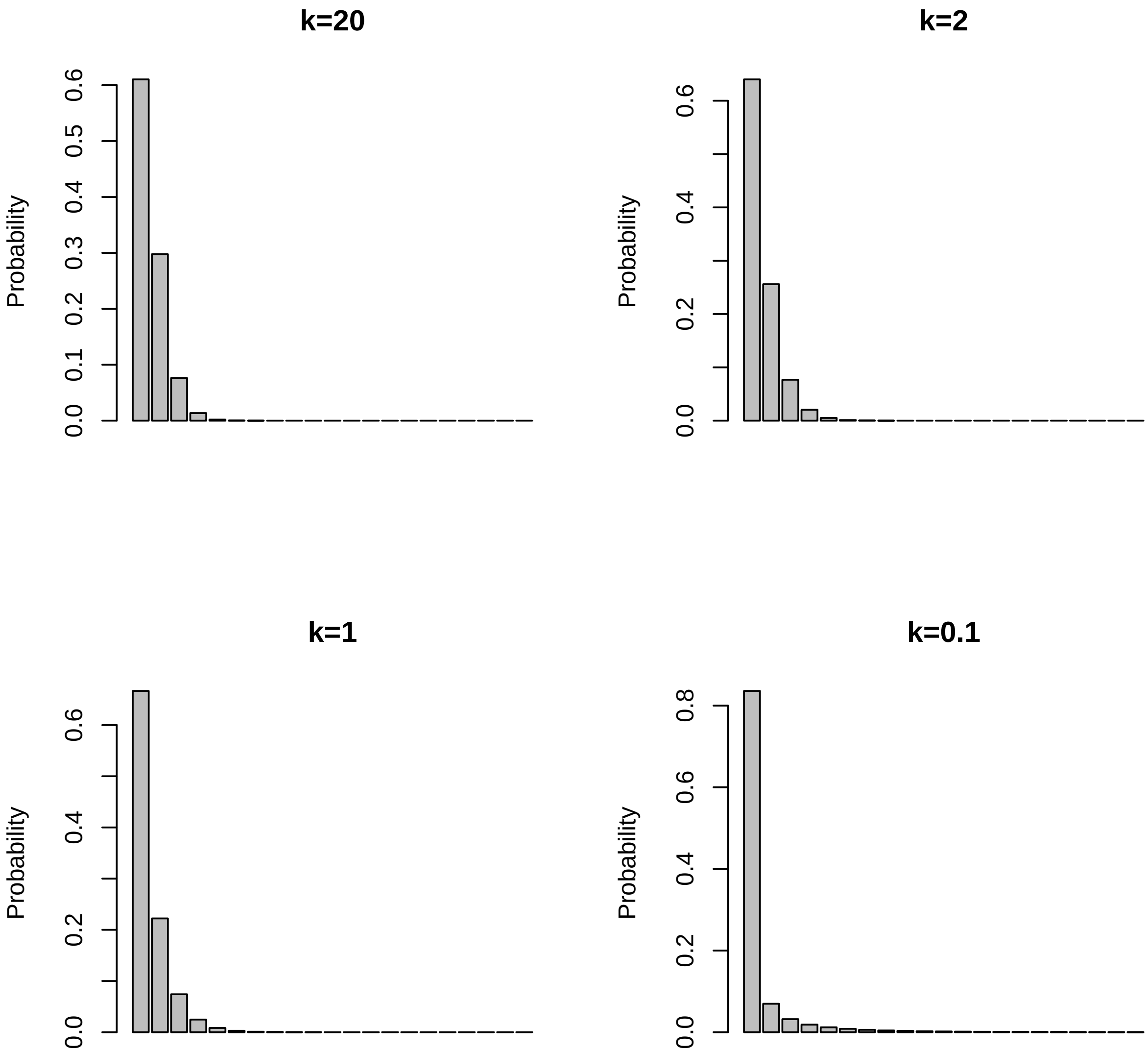
$g''(u) = 2$

Mean: 2

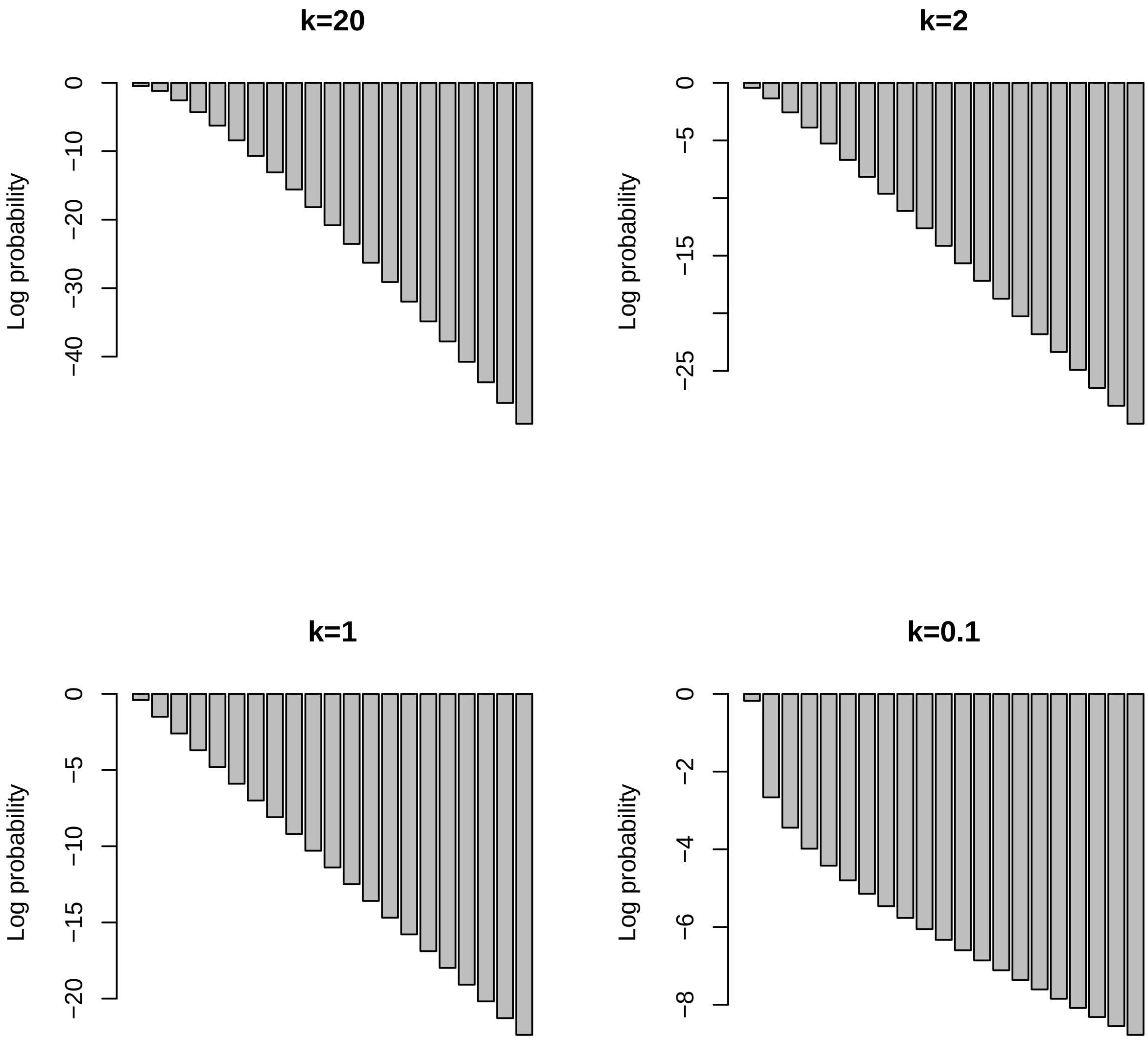
Variance: $2 + 2 - 2^2 = 0$

Exercise 3. Make barplots of the negative binomial distribution with mean 2 for $k = 20$, $k = 2$, $k = 1$, and $k = 0.1$. Holding the mean constant, how can you have more superspreading by changing k? ∎

```
kks <- c(20, 2,1,0.1)
mu <- 0.5
par(mfrow=c(2,2))
for(kk in kks){
        barplot((dnbinom(0:20,kk,mu=mu)),main = paste("k=",kk,sep=""),ylab="Probability")
}
```




```
for(kk in kks){
        barplot(log(dnbinom(0:20,kk,mu=mu)),main = paste("k=",kk,sep=""),ylab="Log probability")
}
```



More super spreading because the large k is the heavier the tail is. This is particularly obvious on the log-scale.

Exercise 4. Assume now the aggregation parameter is 0.45, and compute the confidence interval (approximately). ∎

```r
kk <- 0.45
gwsim <- function(n,cumul,rr,agg,curit=0,maxit=1024,mxclus=4096) {
        if (rr<0) {error("invalid r") }
        if (curit>maxit || cumul>=mxclus) { cumul
        } else if (n <= 0) { cumul
        } else {
        nc <- sum(rnbinom(n,size=agg,mu=rr))
        gwsim(nc,cumul+n,rr,agg,curit+1,maxit)
} }
gwsim.rep <- function(reps,n,cumul,rr,agg) { values <- rep(NA,reps)
for (ii in 1:reps) {
values[ii] <- gwsim(n,cumul,rr,agg) }
values
}
do.gwsim <- function(val,agg,reps=65536,nstart=42) { ans <- rep(0,reps)
for (ii in 1:reps) {
ans[ii] <- gwsim(nstart,0,rr=val,agg=agg) }
  sim.cluster.size <- ans/nstart
  rr = 1 - (1/sim.cluster.size)
  quantile(rr,c(0.025,0.975))
}
do.gwsim(0.5,0.45)

##      2.5%     97.5%
## 0.2500000 0.6769231
```

Exercise 5. How long did this assignment take you? ∎