STAT 703/J703
January 18th, 2005
-Lecture 3-

Instructor: Brian Habing
Department of Statistics
LeConte 203
Telephone: 803-777-3578
E-mail: habing@stat.sc.edu

Today

Section 8.4: Method of Moments
- Estimating α and λ for the Gamma
- Q-Q plots
- The Parametric Bootstrap
- Consistency

A <u>sampling distribution</u> is the probability distribution of a statistic.

In general we want a sampling distribution that is as close as possible to the corresponding parameter.

8.4 – Method of Moments (cont.)

Use a number of moments equal to the number of parameters that need to be estimated, and set the sample moments equal to the distribution's moments.

Recall that…

$\mu_1 = E(X) = \mu$

$\mu_2 = E(X^2) = Var(X)+(E(X))^2 = \sigma^2 + \mu^2$

$$\hat{\mu}_1 = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x} \qquad \hat{\mu}_2 = \frac{\sum_{i=1}^{n} x_i^2}{n} = \frac{\sum_{i=1}^{n} (x_i - \bar{x} + \bar{x})^2}{n}$$

$$= \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2 + n\bar{x}^2}{n} = \hat{\sigma}^2 + \bar{x}^2$$

Consider a Gamma distribution where:

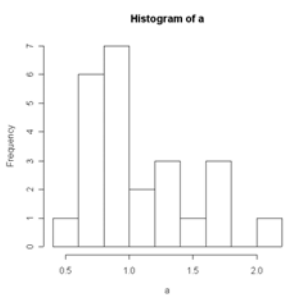
$$\mu = \alpha/\lambda$$

$$\sigma^2 = \alpha/\lambda^2$$

Example: Discrimination parameters for a law school admissions test.

0.52208	0.61226	0.61651
0.67259	0.68124	0.70027
0.79531	0.80179	0.85638
0.87090	0.88407	0.90651
0.95291	0.99212	1.08418
1.09365	1.23861	1.36625
1.36719	1.57871	1.61840
1.67781	1.77927	2.02504


```
> hist(a)
```

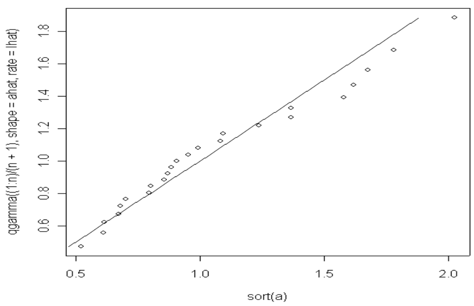



```
> mean(a)
[1] 1.070585

> (n-1)/n*var(a)
[1] 0.1691372
```


<u>Question 1:</u> Does the gamma with these parameters seem to match our data?

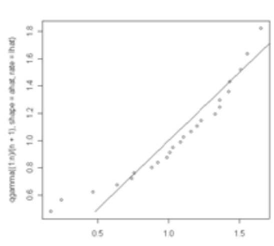
A quantile-quantile plot of our data against $F^{-1}(i/(n+1))$ could be used to see.

```
xbar<-mean(a)
sigma2hat<-(n-1)/n*var(a)
lhat<-xbar/sigma2hat
ahat<-xbar^2/sigma2hat
n<-length(a)

plot(sort(a),qgamma((1:n)/(n+1),
  shape=ahat,rate=lhat))
lines(qgamma((1:n)/(n+1),shape=ahat,
  rate=lhat),qgamma((1:n)/(n+1),
  shape=ahat,rate=lhat))
```

Say we repeated the plot with

a<-rnorm(24,1.07,sqrt(.1677))



Question 2: How accurate are the estimates?

If we had the actual α and λ we could get a large number of samples of size 24 from that distribution and calculate the estimates for each one.

Since we don't have the true α and λ the best we can do is to use the estimates instead.

This is called a parametric bootstrap.


```
nsamples<-100000
x<-
  rgamma(n*nsamples,shape=ahat,
  rate=lhat)
x<-matrix(x,ncol=n)
xbar.dist<-apply(x,1,mean)
s2h.dist<-
  (n-1)/n*apply(x,1,var)
lhat.dist<-xbar.dist/s2h.dist
ahat.dist<-xbar.dist^2/s2h.dist
```

Unfortunately we generally have no way of knowing exactly how well the method of moment estimators will behave in general.

We do, however, know that they are consistent.

Definition: Let $\hat{\theta}_n$ be an estimate of θ for a sample of size n. $\hat{\theta}_n$ Is said to be a consistent estimator of θ if it converges to θ in probability.

That is, if for any $\varepsilon>0$,

$$P(|\hat{\theta}_n - \theta| > \varepsilon) \rightarrow 0 \text{ as } n \rightarrow \infty$$