DEPARTMENT OF STATISTICS, Penn State University
Statistics 514 (Spring 2007)

Part of HW 4:

- Suppose that a particle of initial size y_0 is subjected to repeated impacts, that on each impact a proportion, X_i, of the particle is left, and that the X_i are modeled as independent random variables having the same distribution (iid). After the first impact the size of the particle is $Y_1 = X_1 y_0$, and after n impacts the size is $Y_n = X_n X_{n-1} \dots X_2 X_1 y_0$. We are interested in $E(Y_n)$. On the course web page you will find the data from an experiment in which a particle was subjected to 50 impacts. Using the data, estimate $E(Y_n)$ and the standard error of the estimate of $E(Y_n)$. Also form a 95% bootstrap CI for $E(Y_n)$. Let the value of y_0 be $10,000,000$ units.

- **Solution** We are going to use non-parametric bootstrap: Resample from $x_1, \dots, x_{50}$ for B times (let B be some large number, say 1000) allowing for repetitions. Calculate the corresponding Y_n for every bootstrap sample. Take the mean and standard deviation, that is the estimator of the $E(Y_n)$ and the standard error of the estimate of $E(Y_n)$. The 2.5th and the 97.5th percentile of the 1000 bootstrap Y_ns will be the 95% bootstrap CI for $E(Y_n)$.

- **Implementation** You can do the coding in any computer language you are familiar with, but here is a sample R code you could use:

```
boots=function(x,B){
   est=1:B
   est0=prod(x)*10^7
   for (i in 1:B){
      newx=sample(x,50,replace=T)
      est[i]=prod(newx)*10^7
   }
   hist(est)
   meann=mean(est)
   se=sqrt(var(est)+(mean(est)-est0)^2)
   q1=quantile(est,.025)
   q2=quantile(est,.975)
   cat("95\% CI is", c(q1,q2),"\n")
   list(mean=meann, se=se, CI=c(q1,q2))
}
```

where x is the target data set and B is the number of bootstrap samples. In order to run this program, first you need to copy it to an R-command window and then type:

```
boots(x,1000)
```

and you will see a histogram as well as the asked mean, se and the CI.