18.06 - Spring 2004 - Problem Set 5
SOLUTION TO CHALLENGE PROBLEM

March 19, 2004

- **Work:** book section 4.3 (exercises 4, 9, 12, 26 and 27).
- **Work:** book section 4.4 (exercises 3, 7, 15, 18, 24 and 36).
- **Work:** book section 5.1 (exercises 3, 8, 12, 15, 28 and 33).

Challenge Problem with MATLAB This problem fits the best (least squares) straight line to m random observations : use $b = randn(1, m)$. Repeat N times to get average values of C and D in the best line $C + Dt$. (The values of $t_1, \ldots, t_m$ should be just $1 : m$.)

1. For a fixed size m (try $m = 100$), how quickly (what power of N) does the average of C, D over N experiments approach its limit?
   HINT: You might take logarithms so that average $\approx N^{-p}$ tells you the power p.
2. As you increase m, how does the average behave as function of m?
   HINT: Again I am looking for the exponent q in average $\approx m^{-q}$.

Solution:

```
function [C,D]=sim1(N,m)

x=1:m;
C=0; D=0;
for i=1:N
  b=randn(1,m);
  p=polyfit(x,b,1);
  C=C+p(2);
  D=D+p(1);
end
C=C/N;
D=D/N;
```

1.
```
m=1000;
Ns=[100,200,400,800,1600,3200,6400,12800];

C=[]; D=[];
for N=Ns
  [C1,D1]=sim1(N,m);
  C=[C,C1];
  D=[D,D1];
end

% create a log-log plot
loglog(Ns,abs(C),Ns,abs(D)),legend('Mean(C)','Mean(D)'),xlabel('N')
pfitC=polyfit(log(Ns),log(abs(C)),1); slopeC=pfitC(1)
pfitD=polyfit(log(Ns),log(abs(D)),1); slopeD=pfitD(1)
```

2.
```
ms=[100,200,400,800,1600,3200];
N=10000;

C=[]; D=[];
for m=ms
  [C1,D1]=sim1(N,m);
  C=[C,C1];
  D=[D,D1];
end

% create a log-log plot
loglog(ms,abs(C),ms,abs(D)),legend('Mean(C)','Mean(D)'),xlabel('m')
pfitC=polyfit(log(ms),log(abs(C)),1); slopeC=pfitC(1)
pfitD=polyfit(log(ms),log(abs(D)),1); slopeD=pfitD(1)

% we are sure p=1/2. We think q might be 1 for C and 2 for D.
```