Econ. 517, Econometric Theory I — Fall 2001 — Chris Sims

Subway Exercise Answer*

Apparently most students were able to answer parts (a)-(c) of this exercise well, but most had trouble with part (d). Part (d) probably requires more Matlab cleverness than was appropriate for an exercise, though how it should be done in principle should not have been mysterious.

We have a joint pdf, formed from prior (in this case flat) and likelihood, proportional to

$$\prod_{i=1}^{8} \frac{e^{(c+\gamma/d_i)s_i}}{(1+e^{c+\gamma/d_i})^{n_i}},$$

where s_i is the number who took the subway at distance d_i and n_i is the total number sampled at distance d_i. It is fairly straightforward to evaluate this function at a matrix of points in c, γ space and then to plot its contours and to sum it horizontally and vertically to get functions proportional to the two marginal pdf's. In part (d), you were asked to determine the distribution of the proportional increase in the number of riders who take the subway, once all riders are uniformly spread over distances .1-.4, instead of over .1-.8 as at present.

The current ridership, as a proportion of the population, is

$$\frac{1}{8}\sum_{i=1}^{8} \frac{e^{c+\gamma/d_i}}{1+e^{c+\gamma/d_i}}.$$

The proportion riding after the change will be

$$\frac{1}{4}\sum_{i=1}^{4} \frac{e^{c+\gamma/d_i}}{1+e^{c+\gamma/d_i}}.$$

The ratio of these two quantities gives the proportional increase factor $f(c,\gamma)$. Since f is a function of the unknown parameters γ and c, we can evaluate it at all the points of our grid of c and γ values, by the same methods we used to evaluate the joint pdf at those values. The tricky question is, how to get from these two matrices of function values, $f(c,\gamma)$ and $p(c,\gamma)$, to a pdf for f.

A first step is to convert the matrix of f values into a vector and sort it, preserving the sort order as well as the sorted values. Then reorder the density matrix p into a vector, permuting it the same way the f vector was permuted. Then a plot of `c=cumsum(p)` against `f` provides a plot of the cdf. To get the pdf, we need to estimate the slope of this cdf. One could try just looking at `p(1:end-1)./(f(2:end)-f(1:end-1))`. In this case, though, the result is too erratic. Instead, we smooth over more points. See the attached listing of Matlab commands.

```
subwaydata=[.1  6  10
            .2  8  15
            .3  5  11
            .4  5  13
            .5  4  20
            .6  1  5
            .7  0  1
            .8  2  7];
n=subwaydata(:,3);
s=subwaydata(:,2);
d=subwaydata(:,1);
c=-3:.03:0;
gam=0:.004:.4;
%
%  Note that the problem statement's suggestion of (0,4)
%  as the range for gamma is appropriate only if the distances
%  are measured as 1:8, instead of .1:.1:.8
%
[C,G]=meshgrid(c,gam);
llh=zeros(size(C));
m=max(max(llh))
llh=llh-m;
p=exp(llh);
contour(c,g,p);
plot(c,sum(p))
plot(gam,sum(p'))
% Basically finished with parts (a)-(c).
f=zeros(size(C));
g=f;
for id=1:8
   f=f+exp(C+G/d(id))./(1+exp(C+G/d(id)));
end
for id=1:4
   g=g+exp(C+G/d(id))./(1+exp(C+G/d(id)));
end
f=f/8;g=g/4;
%
% f is proportion using subway before new stations
% g is proportion using subway after new stations
```

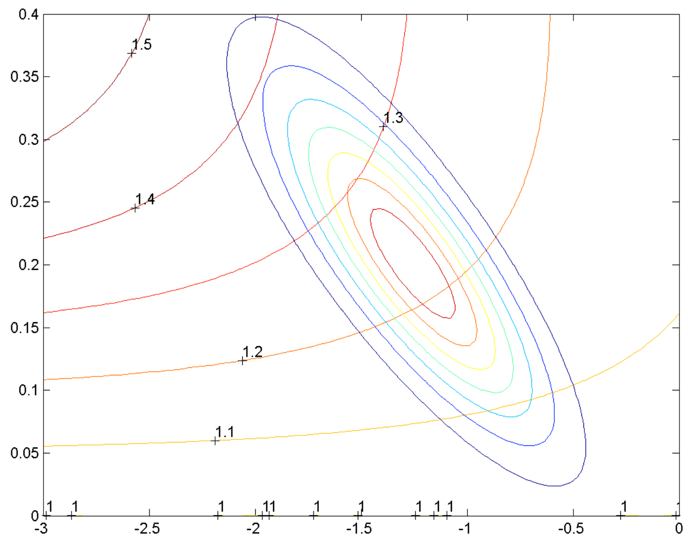

FIGURE 1. Contours of the pdf and of the ridership ratio

```
%
r=g./f;
[rs,irs]=sort(r(:));
ps=p(irs);
cs=cumsum(ps);
cs=cs/cs(end);
% cs is cdf of the increase ratio r.
h=contour(c,gam,r)
clabel(h)
hold
contour(c,gam,p)
% graph with pdf contours and r contours on same plot.
nb=75;
plot(rs(nb+1:end-nb),(cs(2*nb+1:end)-cs(1:end-2*nb))./...
   (-rs(1:end-2*nb)+rs(2*nb+1:end)))
grid
% probability per unit change in r, i.e. the pdf of r
% note that rs has about 10000 elements.  If we smooth less,
% using nb=50, say, results become very erratic.
```

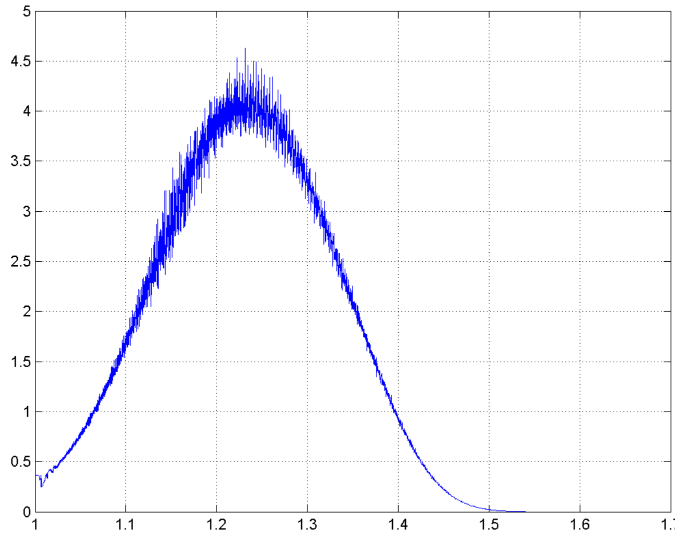

FIGURE 2. pdf for the ridership ratio