Statistics 322, Statistical Computing, Home Work 4

Posted: Sept. 17, 1993 Due: Sept. 26, 1993

This exercise is intended to develop a simple rejection method algorithm for simulating i.i.d. random variables from the βeta density:

$$f_{\alpha,\beta}(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha}(1-x)^{\beta-1} 1_{(0,1)}(x),$$

where $\alpha > 0,\ \beta > 0$.

1. In the following, it will be useful to evaluate this density function. Write and test (you choose your own tests here) a MATLAB function called `fbeta(mx,alpha,beta)`, which evaluates this density, given scalars `alpha`, `beta` > 0, and a matrix `mx` of arguments (of course returning 0 for those entries not between 0 and 1).

   Hint: `beta` is a useful command (but CAREFUL about variable names!).

In the case $\alpha, \beta \geq 1$, a very simple rejection method would start with random pairs uniformly distributed on the rectangle $(0,1) \times (0, \max_x f_{\alpha,\beta}(x))$.

2 Why won't this work for $\alpha < 1$ or $\beta < 1$?

Write a function for generating a column vector of `nrows` i.i.d. βeta(α,β) psuedo random variables, by this method. Turn in:

3. A listing of the function (extra points will be given for good readability).

   Hint: the command `find` is especially useful for selecting certain entries of a matrix.

4. A table which checks the accuracy of the function, by generating 1000 observations, for each of the pairs $(\alpha, \beta) = (1,2), (12,4), (20,1), (40,40)$. For each data set, compare the theoretical values of $P[X \leq x]$ for $x = 0.25,\ 0.5,\ 0.75$, with the endpoints of a simple 95% Confidence Interval for this quantity, based on the empirical value.

   Hint: for the theoretical values, you can try the command `betacdf`. This didn't work for me, so I did my own using the "incomplete beta function", `betainc`.

5. A small table which shows the number of seconds required to generate each one of 10 data sets of 4000 each, as well as the total time for all 10 data sets, for the rather extreme values of $\alpha = 1,\ \beta = 60$.

6. Explain why the algorithm is very slow in the case of (5).