

Computer Applications for Engineers

ET 601

Asst. Prof. Dr. Prapun Sukksompong

prapun@siit.tu.ac.th

Generating Random Numbers (SUPPL)



Office Hours: (BKD 3601-7)

Wednesday 9:30-11:30

Wednesday 16:00-17:00

Thursday 14:40-16:00

Ex: Building your own cdfplot

```
function h = mycdfplot(X)
n = length(X);
cdfest = (1:n)/n;
x = sort(X);
h = stairs([x(1) x],[0 cdfest]);
```

```
close all; clear all;
n = 20;
U = rand(1,n);

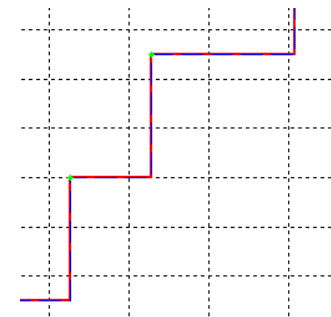
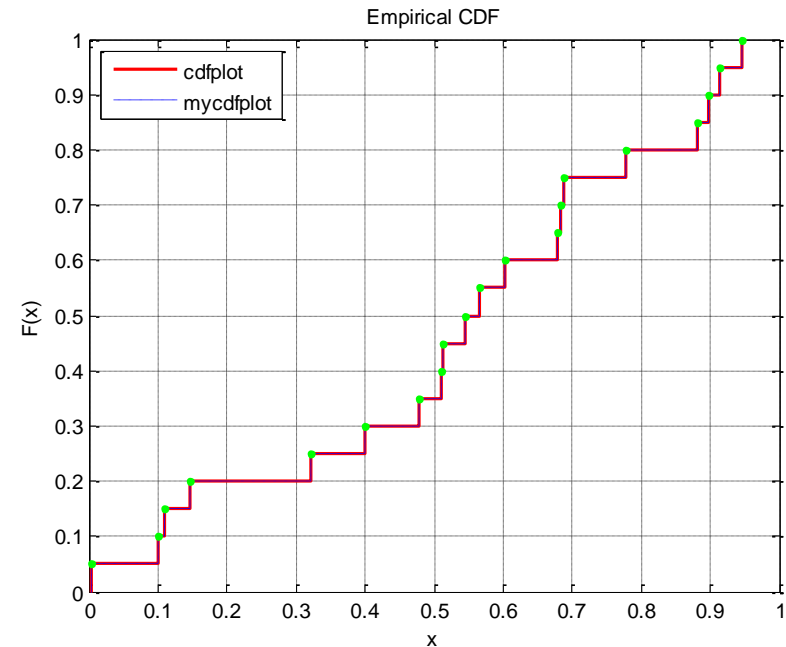
set(cdfplot(U), 'Color', 'r', 'LineWidth', 1.5)

hold on

set(mycdfplot(U), 'Color', 'b', 'LineStyle', '--')

cdfest = (1:n)/n;
plot(sort(U), cdfest, '.g')

legend('cdfplot', 'mycdfplot', 'Location', 'NorthWest')
```



Setting via Plot Tools

The screenshot displays the MATLAB interface with the 'File' menu open. The menu items include: New, Open... (Ctrl+O), Close (Ctrl+W), Save (Ctrl+S), Save As..., **Generate Code...**, Import Data..., Save Workspace As..., Preferences..., Export Setup..., Print Preview..., Print... (Ctrl+P), and Exit MATLAB (Ctrl+Q). The plot area shows an 'Empirical CDF' with a step function. The x-axis is labeled 'x' and ranges from 0.2 to 1.0. The Property Editor at the bottom shows the following settings:

- Display Name: mycdfplot
- X Data Source: auto
- Y Data Source: (empty)
- Plot Type: Stairs
- Line: --- 0.5
- Marker: none 6.0

Teal arrows point to the MATLAB icon in the top-left corner, the 'File' menu, and the 'Generate Code...' option in the menu.

Learning from automated codes

```
1 function createfigure(X1, Y1, X2, Y2, X3, Y3)
2 %CREATEFIGURE(X1, Y1, X2, Y2, X3, Y3)
3 % X1: vector of x data
4 % Y1: vector of y data
5 % X2: stairs x
6 % Y2: stairs y
7 % X3: vector of x data
8 % Y3: vector of y data
9
10 % Auto-generated by MATLAB on 25-Feb-2014 09:36:47
11
12 % Create figure
13 figure1 = figure;
14
15 % Create axes
16 axes1 = axes('Parent', figure1);
17 box(axes1, 'on');
18 grid(axes1, 'on');
19 hold(axes1, 'all');
20
21 % Create plot
22 plot(X1, Y1, 'Parent', axes1, 'LineWidth', 1.5, 'Color', [1 0 0], ...
23      'DisplayName', 'cdfplot');
```

Using Excel for Statistical Analysis

- In addition to its spreadsheet functions, Excel provides a number of standard statistical and graphing procedures.
- Excel is not recommended for statistical analysis, beyond very basic descriptive statistics and getting a feel for your data.
- Microsoft attempted to implement the **Wichmann-Hill (1982) RNG** in Excel 2003 and failed; it did not just produce numbers between zero and unity, it would also produce negative numbers.
 - Microsoft issued a patch for Excel 2003 and Excel 2007 that incorrectly fixed the problem
 - In 2008, McCullough and Heiser showed that whatever RNG it is that Microsoft has implemented in these versions of Excel, it is not the Wichmann-Hill RNG.
 - Microsoft has failed twice to implement the dozen lines of code that define the Wichmann-Hill RNG.

Wichmann–Hill RNG

- Wichmann and Hill (1982)
- Approximate period of 2^{43}
 - of the order of 10^{12}
- Combination of three MCGs.
- The “seeds” I_X , I_Y and I_Z are initialized to take on integer values between 1 and 30,000.

```
 $I_X := 171 \times (I_X \bmod 177) - 2 \times (I_X \div 177);$   
 $I_Y := 172 \times (I_Y \bmod 176) - 35 \times (I_Y \div 176);$   
 $I_Z := 170 \times (I_Z \bmod 178) - 63 \times (I_Z \div 178);$   
if  $I_X < 0$  then  
     $I_X := I_X + 30\,269;$   
if  $I_Y < 0$  then  
     $I_Y := I_Y + 30\,307;$   
if  $I_Z < 0$  then  
     $I_Z := I_Z + 30\,323;$   
 $W := I_X/30269.0 + I_Y/30307.0 + I_Z/30323.0;$   
return  $W - \lfloor W \rfloor;$ 
```

$$x_i \equiv 171x_{i-1} \pmod{30269},$$

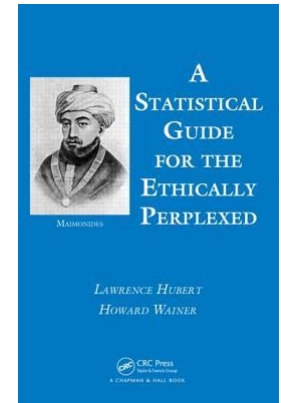
$$y_i \equiv 172y_{i-1} \pmod{30307},$$

$$z_i \equiv 170z_{i-1} \pmod{30323},$$

$$u_i = \left(\frac{x_i}{30269} + \frac{y_i}{30307} + \frac{z_i}{30323} \right) \pmod{1}.$$

References

- Section 7.7 in “A Statistical Guide for the Ethically Perplexed” by Lawrence Hubert and Howard Wainer, 2013
- McCullough, B.A. and David A. Heiser. 2008. “On the accuracy of statistical procedures in Microsoft Excel 2007.” *Computational Statistics and Data Analysis* 52: 4570–4578
 - Microsoft Excel 2003: (2005), *Computational Statistics & Data Analysis*, 49, pp. 1244-1252
 - Microsoft Excel '97: (1999), *Computational Statistics & Data Analysis*, 31, pp. 27-37
- Yalta, A. Talha. 2008. “The accuracy of statistical distributions in Microsoft Excel 2007.” *Computational Statistics and Data Analysis* 52: 4579–4586



Monte Carlo Algorithms

- There are two main directions in the development and study of Monte Carlo algorithms.
 1. **Monte Carlo simulation**: a method for solving probabilistic problems using some kind of simulations of random variables (or random fields.)
 - The algorithms just follow the corresponding processes and phenomena.
 2. **Monte Carlo numerical algorithms**
 - Can be used for solving deterministic problems by modeling random variables or random fields.
 - The main idea is to construct some artificial random process and to prove that the mathematical expectation of the process is equal to the unknown solution of the problem or to some functional of the solution.