Stochastic Simulation
Generation of random variables
Discrete sample space

Bo Friis Nielsen

Applied Mathematics and Computer Science

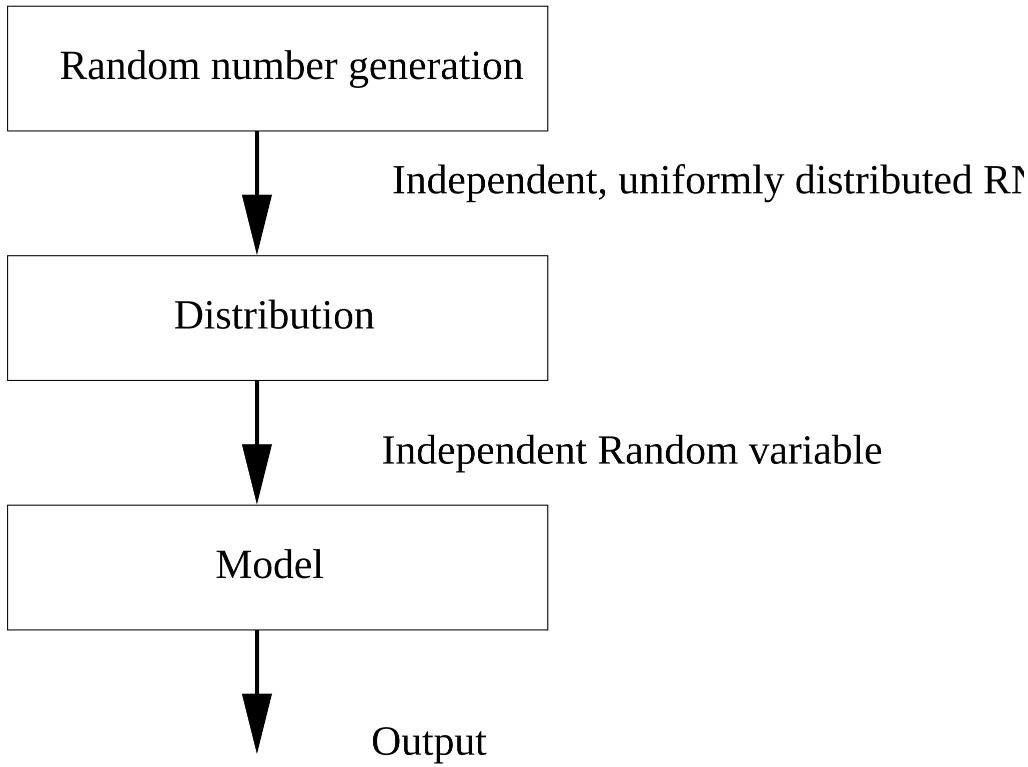
Technical University of Denmark

2800 Kgs. Lyngby – Denmark

Email: bfn@imm.dtu.dk

Insert bfnBuilds and bfnBlobds

Plan W1.1-2


Random number generation
Independent, uniformly distributed R
Distribution
Independent Random variable
Model
Output

Random variables

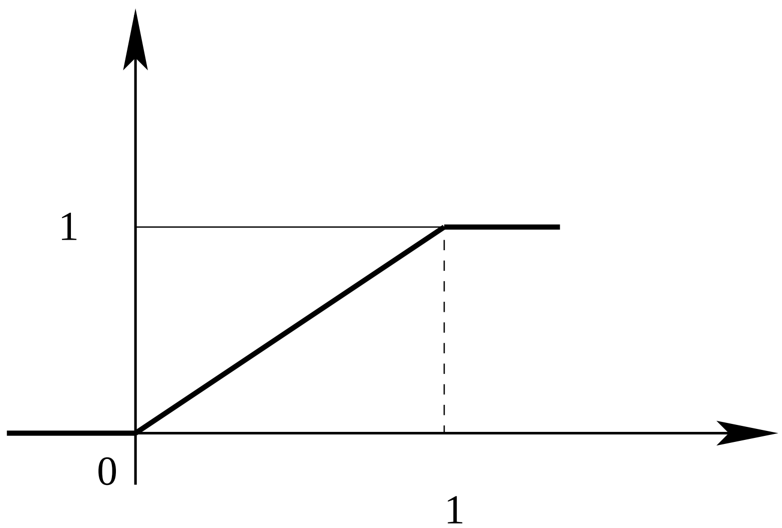
Aim

- The scope is the generation of **independent** random variables X_1, X_2, ... X_n with a **given distribution**, $F_x(x)$, (or probability density function [pdf]).
- We assume we have access to a supply (U_i) of random numbers, independent samples from the uniform distribution on $]0,\ 1[$.
- Our task is to transform U_i into X_i.

Uniform distribution I

Our norm distribution or building brick, $U(0,1)$

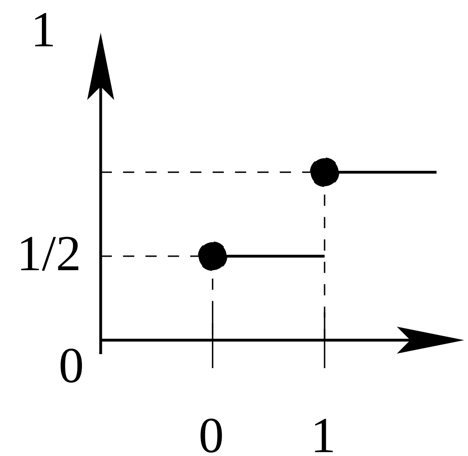
$$f(x) = 1 \quad F(x) = x \quad \text{for} \quad 0 \leq x \leq 1$$

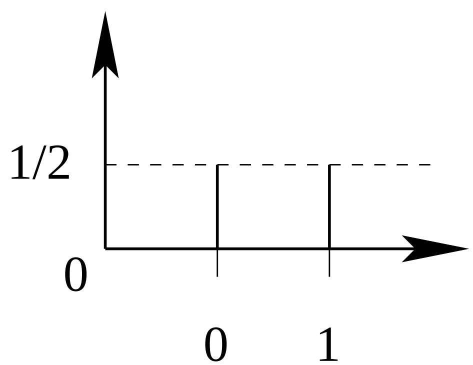


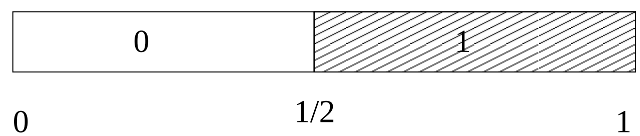
$$\mathbb{E}(X) = \frac{1}{2} \quad \mathbb{V}(X) = \frac{1}{12}$$

Coin

or uniform distribution







$$X = 0,\ 1$$

$$P(X=i) = \frac{1}{2}$$

$$X := \left(U > \frac{1}{2}\right) \qquad X = \lfloor (2U) \rfloor$$

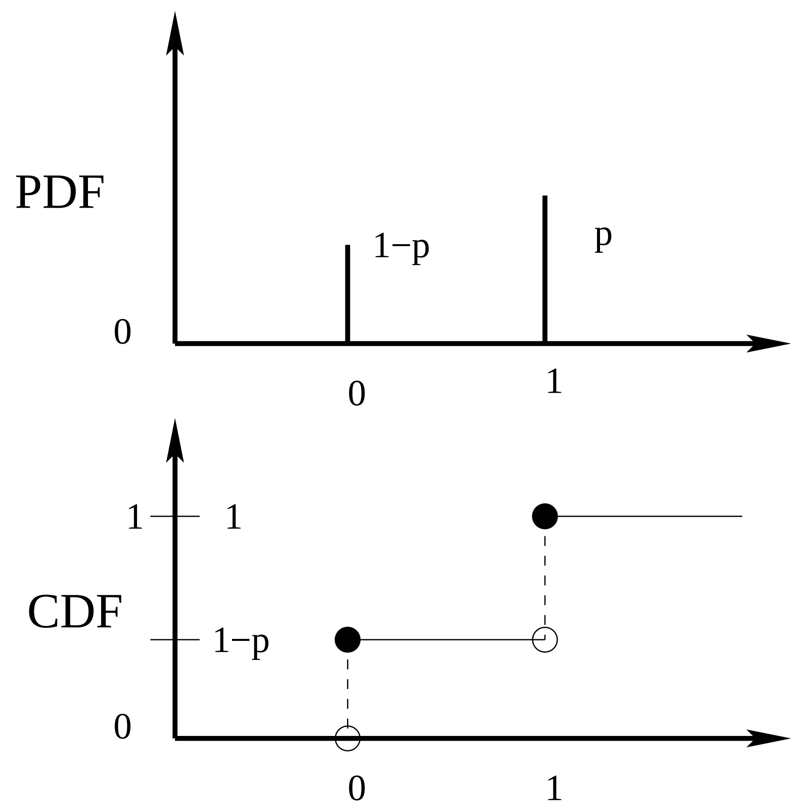
Bernoulli trial

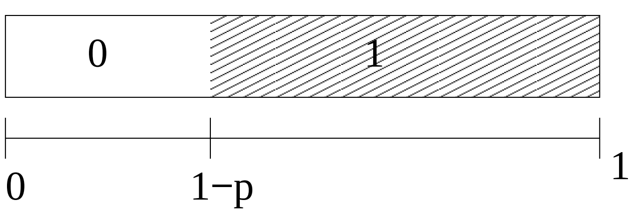
Toss a coin with $P(X=1)=p$ and $P(X=0)=1-p$.

$$F(x) = P\left\{\ X \le x\ \right\}$$

$$X = U > 1-p$$

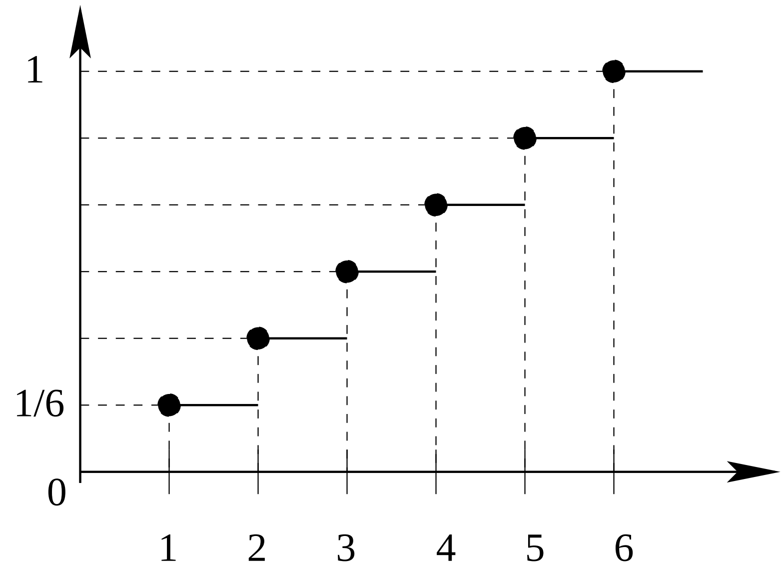
$$\begin{aligned} X &= 0 \qquad 0 \le U \le 1-p \\ X &= 1 \qquad 1-p < U \le 1 \end{aligned}$$

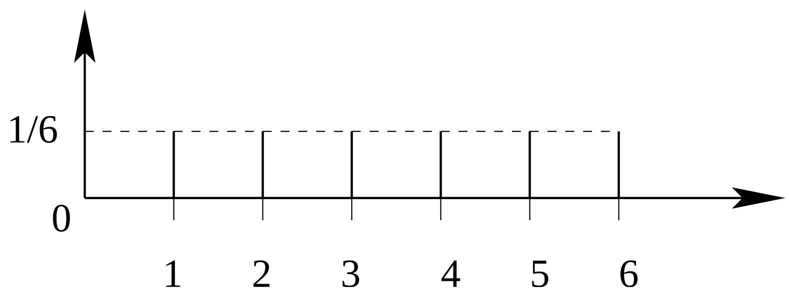


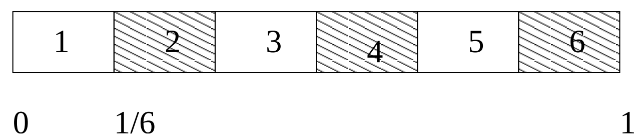


A fair die

or uniform distribution







$$X = 1,\ 2,\ \dots 6$$

$$P(X = i) = 1/6$$

$$X = \lfloor (6U) \rfloor + 1$$

Can be generalized $6 \to n$.

Discrete distribution - direct (crude) method

Suppose X can take n distinct values $x_1 < x_2 < \ldots x_n$ with

$$p_i = P(X = x_i), \quad i = 1,\ 2,\ \ldots\ ,\ n$$

Then X takes the value x_i with probability p_i if U falls in an interval with length p_i. That is if

$$\sum_{j=1}^{i-1} p_j < U \leq \sum_{j=1}^{i} p_j$$

or

$$\boxed{X = x_i \quad if \quad F(x_{i-1}) < U \leq F(x_i)}$$

Geometric distribution, $NB(1,p)$

The discrete time version of waiting time. Memory-less.

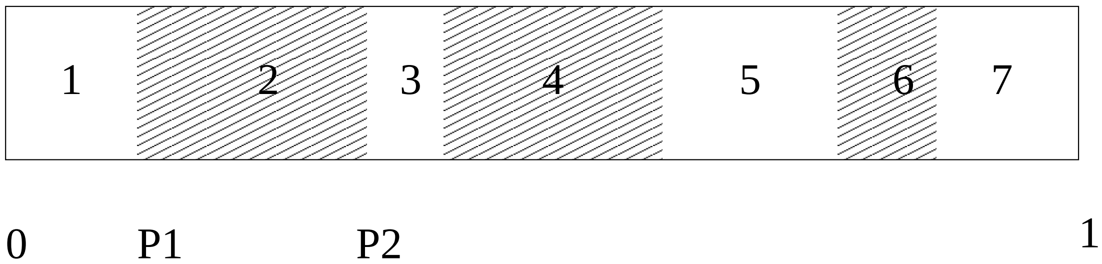
$$
\begin{aligned}
f(n) &= P(X=n) = (1-p)^{n-1}p \quad n = 1,\ 2,\ \ldots \\
F(n) &= P(X \le n) = 1-(1-p)^n
\end{aligned}
$$

$$X = n \quad if \quad F(n-1) < U \le F(n) \qquad 1-(1-p)^{n-1} < U \le 1-(1-p)^n$$

$$n-1 < \frac{\log(1-U)}{\log(1-p)} \le n$$

$$\boxed{X = \left\lfloor \left(\frac{\log(U)}{\log(1-p)} \right) \right\rfloor + 1}$$

Discrete distribution II



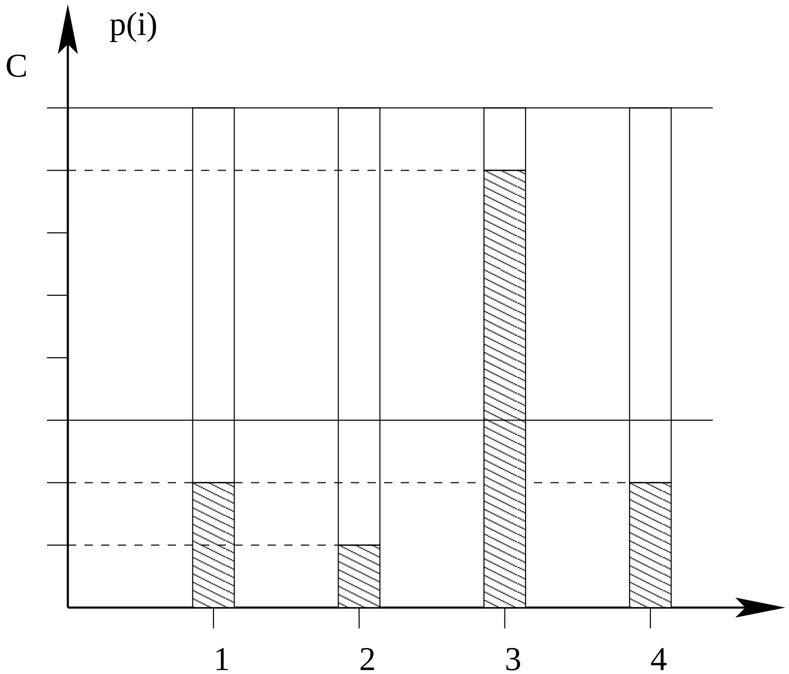
1. Generate U
2. Find the interval k which U belong to. $P_{k-1} < U \leq P_k$
3. output x_k

- Linear search ($\mathbb{E}(X)$)
- Rearrangement of intervals
- Binary search
- Indexed search

Rejection Method

Simple rejection More optimistic: acceptance method.

Assume $P(X = i) = p_i$ for $i = 1,\ 2,\ \ldots\ n$.



Let $c > p_i$ (then $p_i/c < 1$).

1. $I = 1 + \lfloor (n * U_1) \rfloor$
2. output I if $U_2 < p_I/c$.

frequency for i : $\dfrac{\frac{1}{n}\frac{p_i}{c}}{\sum \frac{1}{n}\frac{p_i}{c}} = p_i$

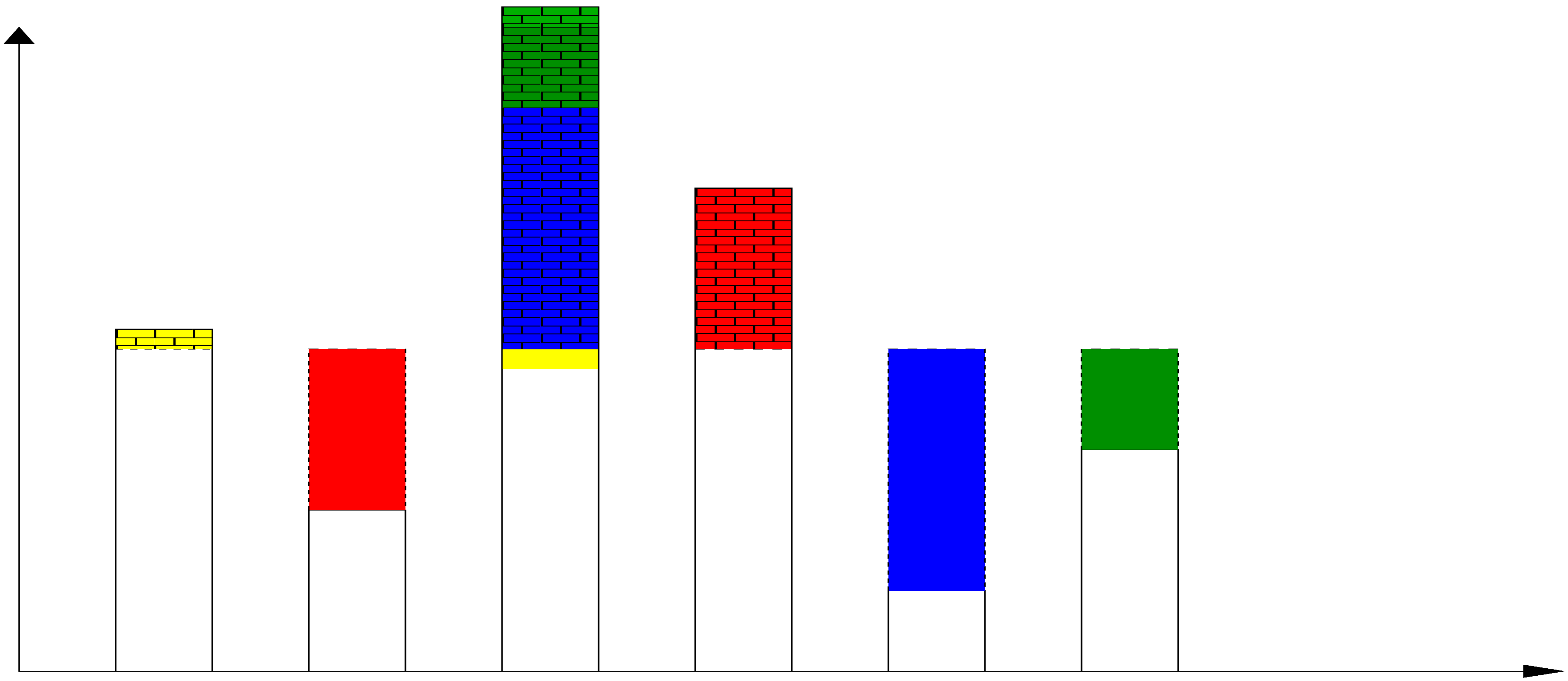
Alias method

- A method for generating discrete random variates of general type
- From discrete uniform to general discrete
- Generate one random number
- One comparison
- Potentially one table look-up
- Drawback: Complex set-up procedure

A six-point distribution

$$P\{X=1\} = \frac{17}{96} \qquad P\{X=2\} = \frac{1}{12} \qquad P\{X=3\} = \frac{1}{3}$$

$$P\{X=4\} = \frac{1}{4} \qquad P\{X=5\} = \frac{1}{24} \qquad P\{X=6\} = \frac{11}{96}$$



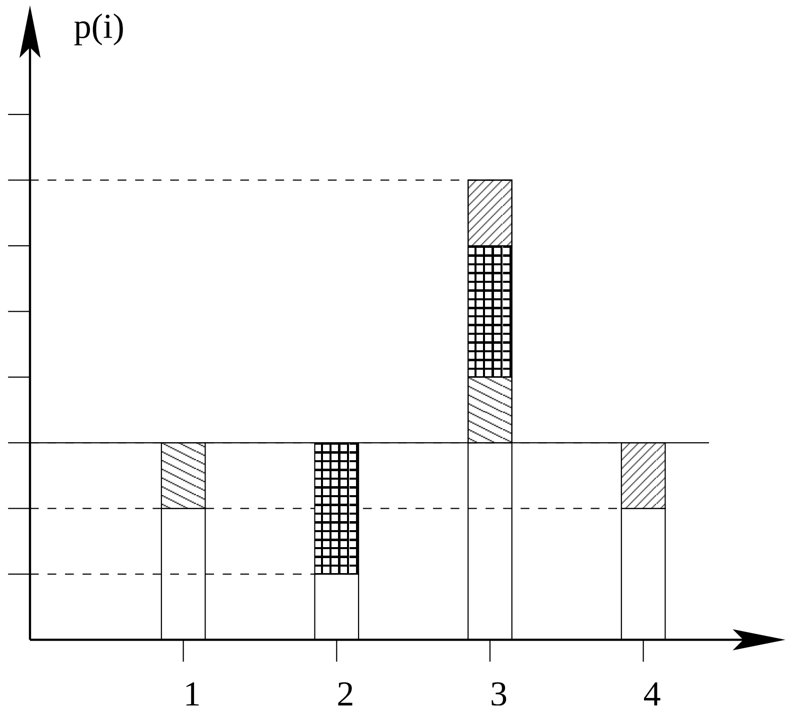
Alias method

- Setup procedure
  - Generate the table of $F(I)$-values, (which part of the mass belongs to I itself.
  - Generate the table of $L(I)$-values, (the alias of class I
- Method at run time
  - Generate I:$I = \lfloor n * U_1 \rfloor$
  - Test against $F(I)$.If $U_2 \leq F(I)$ then return $X = I$ else return $X = L(I)$. The L, F tables for the six-point distribution

$$F(1) = 1 \quad F(2) = \frac{1}{2} \quad F(3) = \frac{15}{16} \quad F(4) = 1 \quad F(5) = \frac{1}{4} \quad F(6) = \frac{11}{16}$$

$$L(1) = 1 \quad L(2) = 4 \quad L(3) = 1 \quad L(4) = 4 \quad L(5) = 3 \quad L(6) = 3$$

Alias Method



On setup: generate F and L.

1. $I = 1 + \lfloor (n * U_1) \rfloor$
2. output I if $U_2 < F(I)$ else $L(I)$.

The Alias tables

Generate F and L.

Pseudo code. p is a vector containing the probabilities.

1. `F=n*p` ($F = 1$ is equivalent for the uniform dist.)
2. `G=find(F>=1)` and `S=find(F<=1)`
3. `while ~isempty(S),`
   (a) `k=G(1)` and `j=S(1)`
   (b) `L(j)=k` and `F(k)=F(k)-(1-F(j))`
   (c) `if F(k)<1-eps then G(1)=[]` and `S=[S k]`
   (d) `S(1)=[]`

Rejection Method

General method

Aim: We will generate X with probabilities $p_i = P\{X = i\}$.

Assume Y with probabilities $q_i = P\{Y = i\}$ is easily generated and $C > \frac{p_i}{q_i}$ for all $i = 1, \ldots n$.

1. Generate Y with probability q_ii and let $j = Y$.
2. Generate U_2. Output $X = Y$ if $U_2 < \frac{p_j}{C q_j}$ else reject.

Probability for $X = i$:

$$q_i \, \frac{p_i}{C q_i} = \frac{p_i}{C}$$

Frequency of output:

$$\frac{p_i/C}{\sum p_i/C} = p_i$$

Excercise 2

Discrete random variables

In the excercise you can use a build in procedure for generating uniform random variables. Compare the results obtained in simulations with expected results. Use histograms (and tests).

- Choice a value for p in the geometric distribution and simulate 10,000 outcomes.
- Simulate the 6 point distribution with

X	1	2	3	4	5	6
p_i	7/48	5/48	1/8	1/16	1/4	5/16

  - by applying a direct (crude) method
  - by using the the rejction method
  - by using the Alias mehtod