CS368 MATLAB Programming

Lecture 9

Young Wu

Based on lecture slides by Michael O’Neill and Beck Hasti

March 30, 2022
Random Choice

Quiz

- Select a random choice.
- A : A
- B : B
- C : C
- D : D
- E : E
Pseudo Randomness

Math

- Truly random numbers are difficult or impossible to generate.
- A sequence of pseudo-random numbers is a deterministic sequence with complicated pattern that looks random to users who do not know the pattern therefore cannot predict the next number in the sequence.
Random Number Generator

Math

- A simple pseudo-random number generator is the Linear Congruential Generator (LCG).

1. Start with a seed $x_0$.
2. Compute $x_{n+1} = (ax_n + c) \mod m$ for some $m, a, c$ unknown to the user.

- The resulting sequence $\frac{x_0}{m}, \frac{x_1}{m}, \ldots$ is approximately uniformly distributed between 0 and 1, including 0, not including 1.
- For example, Java uses $m = 2^{48}, a = 25214903917, c = 11$.
- MATLAB uses another more complicated algorithm called Mersenne Twister.
Random Number Generation

Code

- `rand()` generates a uniform random number between 0 and 1, including 0, not including 1.
- `rand() * u` generates a uniform random number between 0 and `u`, including 0, not including `u`.
- `rand() * (u - l) + l` generates a uniform random number between `l` and `u`, including `l`, not including `u`.
- `rand(n, m)` creates an `n × m` matrix of random numbers between 0 and 1.
Integer Random Numbers

Code

- \texttt{randi}(n) \textit{generates a uniform random integer between 1 and n, including 1 and n.}
- \texttt{randi}([l, u]) \textit{generates a uniform random integer between l and u, including l and u.}
- \texttt{randperm}(n) \textit{generates a random permutation of 1:n.}
- \texttt{randperm}(n, k) with \( k \leq n \) \textit{generates a random sample from 1:n of size k, sampled without replacement.}
A discrete random variable is a random variable that takes on a finite (or countable) number of values with positive probabilities.

The probability that the random variable $X$ takes on value $x$ in $\{1, 2, 3, \ldots\}$ is denoted by $f(x) = \mathbb{P}\{X = x\}$. The function $f$ is called the probability mass function.

A random number generated based on the probabilities specified by $f$ is called a realization of the $X$. 
Cumulative Distribution Functions, Discrete

Math

- The cumulative probability that the random variable $X$ takes on value less than $x$ is denoted by

$$F(x) = \Pr \{ X \leq x \} = \sum_{i=1}^{x} \Pr \{ X = i \}.$$  

The function $F$ is called the cumulative distribution function (CDF).

- The CDF can be efficiently computed using a for loop.

1. $F(1) = \Pr \{ X = 1 \}$.

2. $F(x) = F(x-1) + \Pr \{ X = x \}, x > 1$. 
The inverse transform sampling (also called CDF inversion method) can be used to generate a realization of the random variable $X$.

1. Generate $u \sim \text{Uniform}(0, 1)$.
2. Compute CDF of $X$, call it $F(x)$.
3. Find the largest $x$ such that $F(x) < u$. 
For Loop, Review

Code

Use a for loop to compute the CDF based on the probabilities stored in a vector \( p \), where \( p_i = \mathbb{P}\{X = i\}, i = 1, 2, \ldots, n \).

1. \( f = \text{zeros}(n); \)
2. \( f(1) = p(1); \)
3. \( \text{for } t = 1:n \)
4. \( f(t) = f(t - 1) + p(t); \)
5. \( \text{end} \)

\textit{cumsum}(x) finds the same CDF.
A while loop is used when the loop stops after an unknown number of iterations until some condition is met.

Use a while loop to compute the inverse of the CDF stored in a vector $f$, where $f_i = \mathbb{P}\{X \leq i\}$.

1. $t = 1$;
2. while $f(t) \leq x$
3. \hspace{.5cm} $t = t + 1$;
4. end

$\sum(f \leq x) + 1$ or $\text{find}(x < f, 1)$ finds the same inverse CDF.
Random Variable, CDF

Quiz

- \( \text{cumsum}([0.3, 0.4, 0.3]) \)
- \( A : 1 \)
- \( B : 0.7 \quad 1 \)
- \( C : 0.3 \quad 0.7 \quad 1 \)
Random Variable, Realization

Quiz

1. \( \text{cdf} = \text{cumsum}([0.3, 0.4, 0.3]); \)
2. \( \text{rng}(0); \ u = \text{rand}(); \)  \% it has value \( u = 0.8147 \)
3. \( \text{sum(cdf} \leq u) \) + 1
   - \( A : 0 \)
   - \( B : 1 \)
   - \( C : 2 \)
   - \( D : 3 \)
Random Variable, Realization

Quiz

1. $cdf = \text{cumsum}([0.3, 0.4, 0.3])$
2. $\text{rng}(1); \ u = \text{rand}(); \ % \ it \ has \ value \ u = 0.4170$
3. $\text{sum}(cdf \leq u) + 1$
   - $A: 0$
   - $B: 1$
   - $C: 2$
   - $D: 3$
Random Variable, Continuous

Math

- A continuous random variable is a random variable that takes on uncountably infinite number of values.
- The (theoretical) probability that the random variable is equal to any number is 0.
- Inverse transform sampling can also be used to generate a random value of the variable.
The CDF of a distribution $X$ taking values $(-\infty, \infty)$ is given by $F(x) = \mathbb{P}\{X \leq x\}$.

The derivative of this function is called the probability density function $f(x) = F'(x)$, meaning $F(x) = \int_{-\infty}^{x} f(\hat{x}) \, d\hat{x}$.

More details in the next next lecture.
Simulation
Math

- Direct computation of the probability of an event is sometimes difficult, and simulation can be used to approximate this probability by repeating the same random process a large number of times and find the fraction of times the event occurs.
Simulation, Equal

Quiz

(Estimate the probability that the values from two dice are the same.)

\[ \approx 0.1667 \]

\[ B: \text{mean}(\text{rand}(1, 1000) \times 6) == \text{rand}(1, 1000) \times 6 \]

\[ C: \text{mean}(\text{randi}(6, 1, 1000)) == \text{randi}(6, 1, 1000)) \]
Simulation, Sum
Quiz

(Estimate the probability that the values from two dice sum up to an odd number.)

\( \approx 0.5 \)

\( B : \text{mean} (\text{mod} (\text{randi} (12, 1, 1000), 2) == 1) \)

\( C : \text{mean} (\text{mod} (\text{randi} (6, 1, 1000) + \text{randi} (6, 1, 1000), 2) == 1) \)
Simulation, Geometric
Quiz

• (Estimate the average number of times it takes to throw a die until it lands six.)
• $\approx 6$

1 $s = \text{zeros}(1, 1000)$;
2 \hspace{1cm} \text{for } t = 1:1000
3 \hspace{1cm} B:\quad \text{while} \quad \text{randi}(6) == 6
4 \hspace{1cm} C:\quad \text{while} \quad \text{randi}(6) \sim= 6
5 \hspace{1cm} s(t) = s(t) + 1;
6 \hspace{1cm} \text{end}
7 \hspace{1cm} \text{end}; \quad \text{mean}(s) + 1
Simulation, Geometric Again

Quiz

(Estimate the average number of times it takes to throw a loaded die with probabilities $[0.1 \ 0.2 \ 0.4 \ 0.2 \ 0 \ 0.1]$ until it lands six.)

$\approx 10$

```matlab
s = zeros(1, 1000);
for t = 1:1000
    B: while rand() <= 0.9
    C: while rand() > 0.9
    s(t) = s(t) + 1;
end
end; mean(s) + 1
```
Simulation, Geometric General

Quiz

1. \( s = \text{zeros}(1, 1000); \)
2. \( \text{for } t = 1:1000 \)
   - The following conditions are more general (with \( \text{cdf} = [0.1 \ 0.2 \ 0.4 \ 0.2 \ 0 \ 0.1] \))
     - \( \text{while } \text{sum} (\text{cumsum}(\text{cdf}) \leq \text{rand}()) + 1 = 6 \)
     - \( \text{while } \text{find} (\text{rand}() \leq \text{cumsum}(\text{cdf}), 1) = 6 \)
3. \( s(t) = s(t) + 1; \)
4. \( \text{end} \)
5. \( \text{end}; \text{mean}(s) + 1 \)
Reproducibility

Code

- The same code can produce a different output every time it is executed.
- In order to make a simulation reproducible, the best practice is to always set a seed at the beginning of the simulation using \texttt{rng(seed)}. 
Blank Slide