S&DS 241 Lecture 14

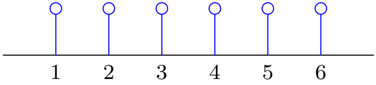
Uniform distribution, exponential distribution, random number generation

B-H 5.2,5.3,5.5

Discrete uniform distribution

"Uniform" means "equally likely".

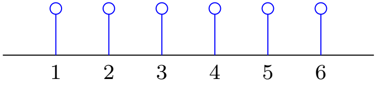
- Fair die: PMF is flat



Discrete uniform distribution

"Uniform" means "equally likely".

- Fair die: PMF is flat

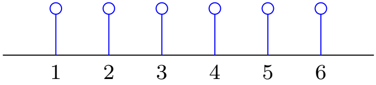


- How to extend this to continuous random variables?

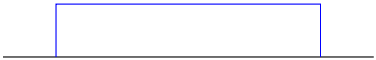
Discrete uniform distribution

"Uniform" means "equally likely".

- Fair die: PMF is flat



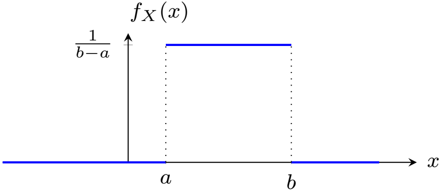
- How to extend this to continuous random variables? PDF is flat



Uniform distribution

A continuous random variable X is said to be uniformly distributed in the interval (a, b), denoted by $X \sim \text{Unif}(a, b)$, if it has the following PDF:

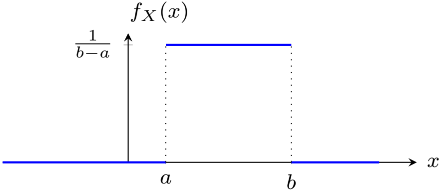
$$\boxed{f_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{else} \end{cases}}$$



Uniform distribution

A continuous random variable X is said to be uniformly distributed in the interval (a, b), denoted by $X \sim \text{Unif}(a, b)$, if it has the following PDF:

$$\boxed{f_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{else} \end{cases}}$$

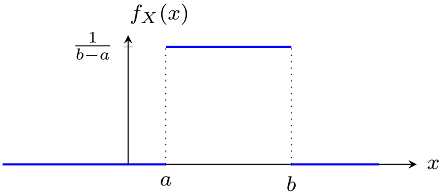


- Continuous analog of "equally likely"
- Support of X: $[a, b]$

Uniform distribution

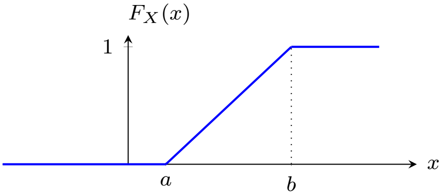
A continuous random variable X is said to be uniformly distributed in the interval (a, b), denoted by $X \sim \text{Unif}(a, b)$, if it has the following PDF:

$$\boxed{f_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{else} \end{cases}}$$



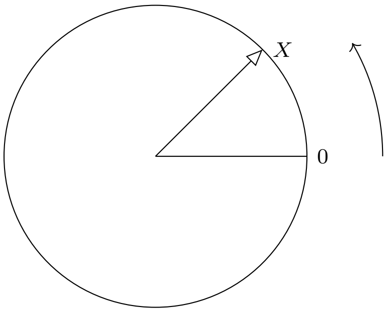
- Continuous analog of "equally likely"
- Support of X: $[a, b]$
- CDF:

$$F_X(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a < x < b \\ 1 & x \geq b \end{cases}$$



Example from last lecture

- Angle of a randomly spinned wheel

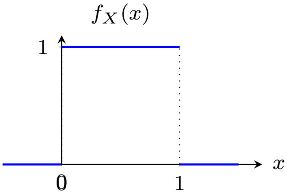


Then $X \sim \text{Unif}(0, 2\pi)$

Unif$(0, 1)$

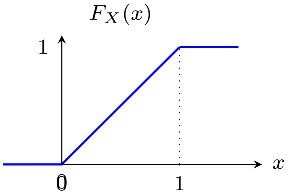
- PDF

$$f_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{else} \end{cases}$$



- CDF:

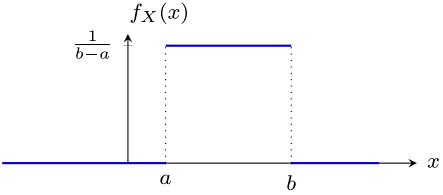
$$F_X(x) = \begin{cases} 0 & x \leq 0 \\ x & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}$$



Mean and variance

Let $X \sim \text{Unif}(a, b)$. Then

$$E(X) = \frac{a+b}{2} = \text{center of interval}$$



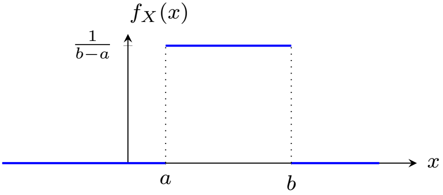
Mean and variance

Let $X \sim \text{Unif}(a, b)$. Then

$$E(X) = \frac{a+b}{2} = \text{center of interval}$$

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

$$\text{SD}(X) = \frac{b-a}{2\sqrt{3}} \propto \text{length of interval}$$



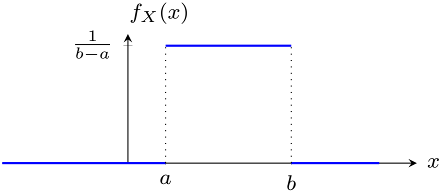
Mean and variance

Let $X \sim \text{Unif}(a, b)$. Then

$$E(X) = \frac{a+b}{2} = \text{center of interval}$$

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

$$\text{SD}(X) = \frac{b-a}{2\sqrt{3}} \propto \text{length of interval}$$



Proof.

- $E(X) = \frac{1}{b-a}\int_a^b x dx = \frac{1}{b-a}\frac{b^2-a^2}{2} = \frac{a+b}{2}$

□

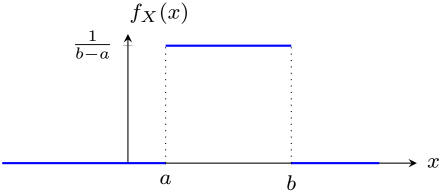
Mean and variance

Let $X \sim \text{Unif}(a,b)$. Then

$$E(X) = \frac{a+b}{2} = \text{center of interval}$$

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

$$\text{SD}(X) = \frac{b-a}{2\sqrt{3}} \propto \text{length of interval}$$



Proof.

- $E(X) = \frac{1}{b-a}\int_a^b x dx = \frac{1}{b-a}\frac{b^2-a^2}{2} = \frac{a+b}{2}$
- $\text{Var}(X) = E[(X-\frac{a+b}{2})^2] = \frac{1}{b-a}\int_a^b (x-\frac{a+b}{2})^2 dx = \frac{1}{b-a}\int_{(a-b)/2}^{(b-a)/2} y^2 dy = \frac{1}{b-a}\frac{2}{3}(\frac{b-a}{2})^3 = \frac{(b-a)^2}{12}$

□

Example

Let $X \sim$ Unif$(0, 2)$. For $t \in (0, 1)$, find

$$P(X \leq t \mid X \leq 1)$$

Example

Let $X \sim$ Unif$(0, 2)$. For $t \in (0, 1)$, find

$$P(X \leq t \mid X \leq 1)$$

- $P(X \leq t \mid X \leq 1) = \frac{P(X \leq t)}{P(X \leq 1)} = \frac{F_X(t)}{F_X(1)} = \frac{t/2}{1/2} = t$

Example

Let $X \sim$ Unif$(0, 2)$. For $t \in (0, 1)$, find

$$P(X \leq t \mid X \leq 1)$$

- $P(X \leq t \mid X \leq 1) = \frac{P(X \leq t)}{P(X \leq 1)} = \frac{F_X(t)}{F_X(1)} = \frac{t/2}{1/2} = t$
- Interpretation: You go to lunch with your friend, say Tom, who will arrive at some time uniformly distributed between noon and 2pm. Suppose he calls you and says he will arrive by 1pm. Then his arrival time is uniform between noon and 1pm.

Example

Let $X \sim \text{Unif}(0,2)$. For $t \in (0,1)$, find

$$P(X \leq t \mid X \leq 1)$$

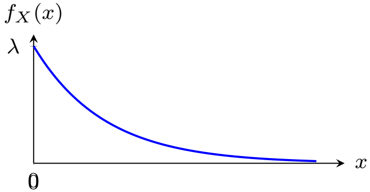
- $P(X \leq t \mid X \leq 1) = \frac{P(X \leq t)}{P(X \leq 1)} = \frac{F_X(t)}{F_X(1)} = \frac{t/2}{1/2} = t$
- Interpretation: You go to lunch with your friend, say Tom, who will arrive at some time uniformly distributed between noon and 2pm. Suppose he calls you and says he will arrive by 1pm. Then his arrival time is uniform between noon and 1pm.
- Mathematically: conditioned on the event $\{X \leq 1\}$, X is uniformly distributed on $(0,1)$.

Exponential distribution

Exponential distribution

A continuous random variable X is said to be exponentially distributed with parameter $\lambda > 0$, denoted by $X \sim \text{Expo}(\lambda)$, if it has the following PDF:

$$\boxed{f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{else} \end{cases}}$$

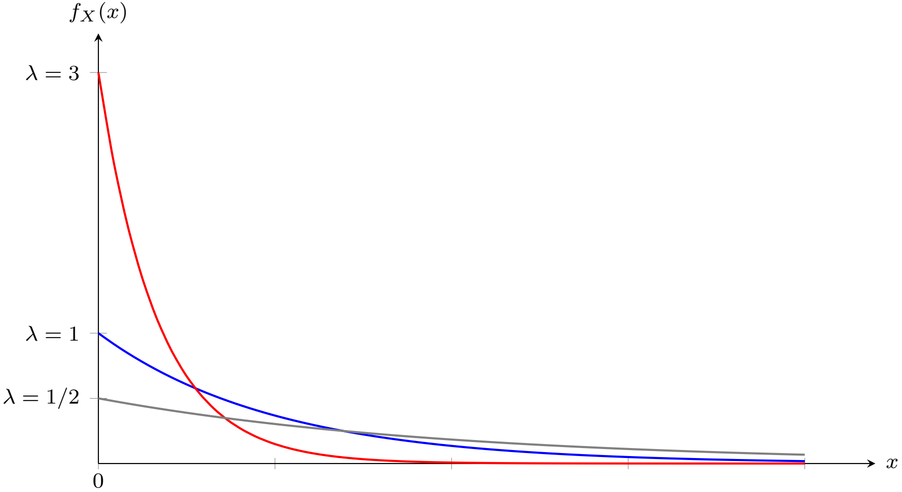


Usefulness

Exponential distribution is often used to model time:

- The time until a radioactive particle decays
- Interarrival time of calls to a call center
- Time till the next accident, etc...

Qualitative effect of the parameter

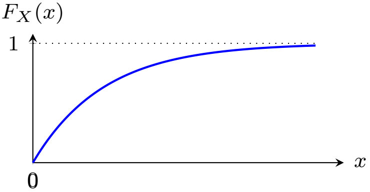


Observation

- $\lambda \uparrow \implies$ PDF concentrated near zero and decays faster
- $\lambda \downarrow \implies$ PDF more spread out and decays slower

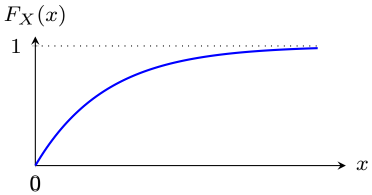
CDF

$$\boxed{F_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}}$$



CDF

$$\boxed{F_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}}$$



- From PDF to CDF:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \int_0^x \lambda e^{-\lambda t} dt = -e^{-\lambda t}\Big|_0^x = 1 - e^{-\lambda x}$$

- Tail probability:

$$P(X > x) = 1 - F_X(x) = e^{-\lambda x}$$

which becomes smaller if λ increases

Mean and Variance

Let $X \sim \text{Expo}(\lambda)$. Then

$$\boxed{\begin{aligned} E(X) &= \frac{1}{\lambda} \\ \text{Var}(X) &= \frac{1}{\lambda^2} \end{aligned}}$$

Mean and Variance

Let $X \sim \text{Expo}(\lambda)$. Then

$$\boxed{\begin{aligned} E(X) &= \frac{1}{\lambda} \\ \text{Var}(X) &= \frac{1}{\lambda^2} \end{aligned}}$$

Recall integration by parts:
$\int_a^b f(x)g'(x)dx = f(x)g(x)\big|_a^b - \int_a^b f'(x)g(x)dx$.

$$E(X) = \int_0^\infty x\lambda e^{-\lambda x}dx = \underbrace{-xe^{-\lambda x}\Big|_0^\infty}_{0} + \underbrace{\int_0^\infty e^{-\lambda x}dx}_{1/\lambda} = \frac{1}{\lambda}$$

Mean and Variance

Let $X \sim \text{Expo}(\lambda)$. Then

$$\boxed{\begin{aligned} E(X) &= \frac{1}{\lambda} \\ \text{Var}(X) &= \frac{1}{\lambda^2} \end{aligned}}$$

Recall <u>integration by parts</u>:
$\int_a^b f(x)g'(x)dx = f(x)g(x)\Big|_a^b - \int_a^b f'(x)g(x)dx$.

$$E(X) = \int_0^\infty x\lambda e^{-\lambda x}dx = \underbrace{-xe^{-\lambda x}\Big|_0^\infty}_{0} + \underbrace{\int_0^\infty e^{-\lambda x}dx}_{1/\lambda} = \frac{1}{\lambda}$$

and

$$E(X^2) = \int_0^\infty x^2\lambda e^{-\lambda x}dx = \underbrace{-x^2e^{-\lambda x}\Big|_0^\infty}_{0} + \underbrace{\int_0^\infty 2xe^{-\lambda x}dx}_{2E(X)/\lambda} = \frac{2}{\lambda^2}$$

Memoryless property of geometric distribution

Recall from Lec 8: $L \sim \mathsf{Geom}(p)$

$$P(L = k + \ell | L \geq k) = P(L = \ell), \quad k, \ell \geq 0.$$

Interpretation: Having failed k times already, the probability that one fails another ℓ times is the same as failing ℓ times from the fresh start, as if the past is "forgotten."

Memoryless property of exponential distribution

Let $X \sim \text{Expo}(\lambda)$. Then

$$P(X > s + t | X > t) = P(X > s), \quad s, t > 0$$

Proof.

$P(X > s + t | X > t) = \frac{P(X > s+t)}{P(X > t)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} = P(X > s).$ $\square$

Memoryless property of exponential distribution

Let $X \sim \text{Expo}(\lambda)$. Then

$$P(X > s + t | X > t) = P(X > s), \quad s, t > 0$$

Proof.

$P(X > s + t | X > t) = \frac{P(X > s+t)}{P(X > t)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} = P(X > s).$ $\square$

Therefore exponential distribution can be a useful model for

- Interarrival time between spam emails, visits to a website, accidents, etc

Memoryless property of exponential distribution

Let $X \sim \text{Expo}(\lambda)$. Then

$$P(X > s + t | X > t) = P(X > s), \quad s, t > 0$$

Proof.

$P(X > s + t | X > t) = \frac{P(X > s+t)}{P(X > t)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} = P(X > s).$ □

Therefore exponential distribution can be a useful model for

- Interarrival time between spam emails, visits to a website, accidents, etc

but not a good model for

- Lifetime of a human being, a light bulb, etc

since it does not account for the effect of "**getting worn out**"

How to generate random variables?

Why do we need to generate random numbers?

Why do we need to generate random numbers?

Why do we need to generate random numbers?

Why do we need to generate random numbers?

- Lottery
  - Raise revenue and tax
  - Incentivize

Why do we need to generate random numbers?

- Lottery
  - Raise revenue and tax
  - Incentivize
- Statistical applications
  - For example, polling (Lec 12), randomized clinical trial

Why do we need to generate random numbers?

- Lottery
  - Raise revenue and tax
  - Incentivize
- Statistical applications
  - For example, polling (Lec 12), randomized clinical trial
- Simulations
  - Test whether a system is noise resilient
  - Test whether an estimator is accurate (on synthetic dataset)

Why do we need to generate random numbers?

- Lottery
  - Raise revenue and tax
  - Incentivize
- Statistical applications
  - For example, polling (Lec 12), randomized clinical trial
- Simulations
  - Test whether a system is noise resilient
  - Test whether an estimator is accurate (on synthetic dataset)
  - Computer games (procedural generation)

Why do we need to generate random numbers?

- Lottery
  - Raise revenue and tax
  - Incentivize
- Statistical applications
  - For example, polling (Lec 12), randomized clinical trial
- Simulations
  - Test whether a system is noise resilient
  - Test whether an estimator is accurate (on synthetic dataset)
  - Computer games (procedural generation)
- Algorithmic needs
  - Monte Carlo tree search (Lec 1)
  - Cryptography

Why do we need to generate random numbers?

- Lottery
  - Raise revenue and tax
  - Incentivize
- Statistical applications
  - For example, polling (Lec 12), randomized clinical trial
- Simulations
  - Test whether a system is noise resilient
  - Test whether an estimator is accurate (on synthetic dataset)
  - Computer games (procedural generation)
- Algorithmic needs
  - Monte Carlo tree search (Lec 1)
  - Cryptography

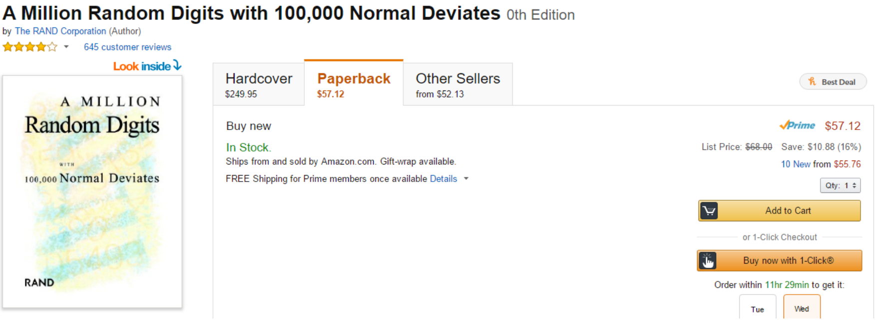
Question: How to generate a random variable of a specific distribution, e.g., uniform?


TOUR OF ACCOUNTING
OVER HERE WE HAVE OUR RANDOM NUMBER GENERATOR.
scottadams@aol.com
www.dilbert.com
NINE NINE NINE NINE NINE NINE
© 2001 United Feature Syndicate, Inc.
ARE YOU SURE THAT'S RANDOM?
THAT'S THE PROBLEM WITH RAN-DOMNESS: YOU CAN NEVER BE SURE.

Random number generator: Physical methods

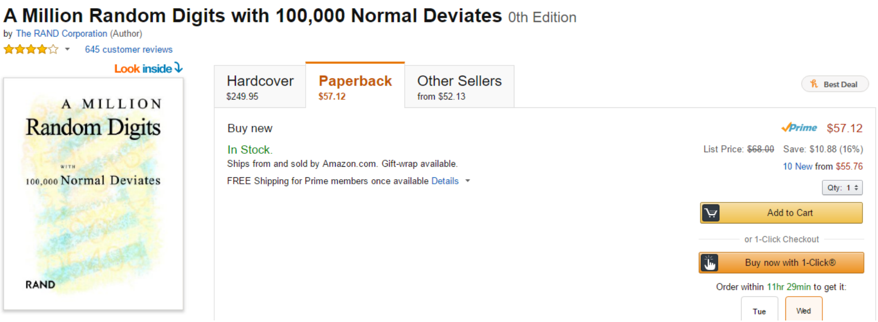
1. Hardware-based: coin, die, roulette
2. Physics-based: thermal noise, cosmic background radiation, quantum...

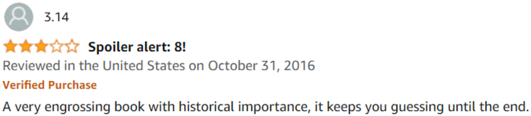
RAND company



A more serious review by the famous statistician John Tukey: https://www.jstor.org/stable/166772.

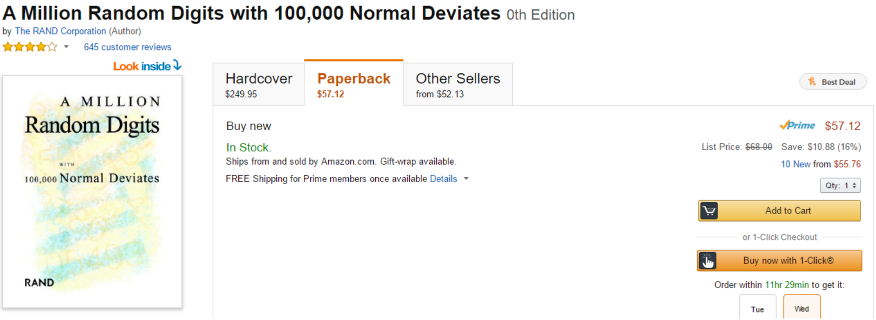
RAND company





A more serious review by the famous statistician John Tukey: https://www.jstor.org/stable/166772.

RAND company



3.14

Spoiler alert: 8!

Reviewed in the United States on October 31, 2016

Verified Purchase

A very engrossing book with historical importance, it keeps you guessing until the end.

Wait for the audiobook version

By R. Rosini on October 19, 2006

Format: Paperback

While the printed version is good, I would have expected the publisher to have an audiobook version as well. A perfect companion for one's Ipod.

A more serious review by the famous statistician John Tukey: https://www.jstor.org/stable/166772.

Methodology

... *The random digits in this book were produced by rerandomization of a basic table generated by an electronic roulette wheel.*

---

http://www.rand.org/pubs/monograph_reports/MR1418/index2.html
https://en.wikipedia.org/wiki/List_of_random_number_generators
https://www.wsj.com/articles/rand-million-random-digits-numbers-book-error-11600893049

Random number generator: Computerized methods

Pseudorandom number generator (PRNG): deterministic methods whose output appears random

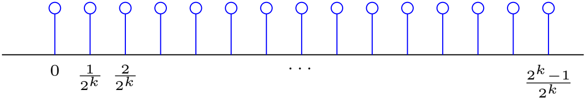
- Simplest method: LCG
  - $X_{n+1} = (aX_n + b) \mod m$
  - X_0: "seed"
- Most programming languages have a `rand` function based on more sophisticated algorithms and a `setseed` function
- For most of the applications, PRNG suffices

Uniform random variable

- For a k-bit number in binary representation:

$$X = 0.X_1X_2\dots X_k$$

flip a fair coin for each bit. Then X is discrete uniform:



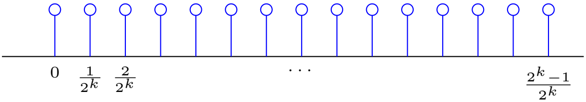
When k is large, this is very close to Unif$(0,1)$

Uniform random variable

- For a k-bit number in binary representation:

$$X = 0.X_1 X_2 \dots X_k$$

flip a fair coin for each bit. Then X is discrete uniform:



When k is large, this is very close to Unif$(0,1)$

- Many PRNG produce random variables close to uniform

Next question

Given uniform random variables, how to generate random variables with other distributions?

Next question

Given uniform random variables, how to generate random variables with other distributions?

B-H §5.3: "Universality of the Uniform"

Quantile transformation

- Let F be the CDF of a continuous distribution
  - Example: for Expo(1), $F(x) = 1 - e^{-x}$

Quantile transformation

- Let F be the CDF of a continuous distribution
  - Example: for Expo(1), $F(x) = 1 - e^{-x}$
- Let F^{-1} denote the inverse CDF, called the quantile function
  - Example: for Expo(1), $F^{-1}(u) = -\ln(1-u)$

Quantile transformation

- Let F be the CDF of a continuous distribution
  - Example: for Expo(1), $F(x) = 1 - e^{-x}$
- Let F^{-1} denote the inverse CDF, called the quantile function
  - Example: for Expo(1), $F^{-1}(u) = -\ln(1-u)$
- Fact: Let $U \sim$ Unif$(0, 1)$. Then the CDF of

$$X = F^{-1}(U)$$

  is given by F.

Proof.

$P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x).$ $\square$

Conversely

- Given X with CDF F,

$$U = F(X)$$

  is distributed as Unif$(0, 1)$.

Conversely

- Given X with CDF F,
$$U = F(X)$$
is distributed as $\text{Unif}(0, 1)$.
- To summarize:
$$X \underset{F^{-1}}{\overset{F}{\rightleftharpoons}} U$$

Conversely

- Given X with CDF F,

$$U = F(X)$$

  is distributed as Unif$(0, 1)$.

- To summarize:

$$X \underset{F^{-1}}{\overset{F}{\rightleftharpoons}} U$$

- Furthermore: given X with CDF F, generate Y with CDF G?

$$\begin{tikzcd} X \arrow[r, "F"] \arrow[dr, "G^{-1}\circ F"'] & U \arrow[d, "G^{-1}"] \\ & Y \end{tikzcd}$$

  that is

$$Y = G^{-1}(F(X))$$

Example

Given $X \sim$ Expo(1), how to generate Expo(3):

Example

Given $X \sim$ Expo(1), how to generate Expo(3):

- Recall CDF of Expo(λ): $1 - e^{-\lambda x}$. Then

$$F(x) = 1 - e^{-x}$$
$$G(x) = 1 - e^{-3x}, \quad G^{-1}(u) = -\frac{1}{3}\ln(1-u)$$

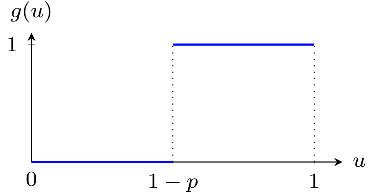
- Assembling everything:

$$Y = G^{-1}(F(X)) = -\frac{1}{3}\ln(1-(1-e^{-X})) = X/3$$

How to generate Bernoulli based on uniform?

How to generate Bernoulli based on uniform?

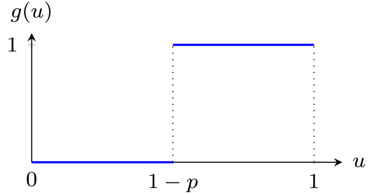
$$g(u) = \begin{cases} 0 & 0 < u < 1-p \\ 1 & 1-p < u < 1 \end{cases}$$



- Then $X = g(U) \sim \text{Bern}(p)$

How to generate Bernoulli based on uniform?

$$g(u) = \begin{cases} 0 & 0 < u < 1-p \\ 1 & 1-p < u < 1 \end{cases}$$

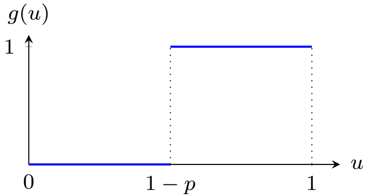


- Then $X = g(U) \sim \text{Bern}(p)$
- Verify:

$$P(X=0) = P(g(U)=0) = P(0 < U < 1-p) = 1-p$$
$$P(X=1) = P(g(U)=1) = P(1-p < U < 1) = p$$

How to generate Bernoulli based on uniform?

$$g(u) = \begin{cases} 0 & 0 < u < 1-p \\ 1 & 1-p < u < 1 \end{cases}$$



- Then $X = g(U) \sim \text{Bern}(p)$
- Verify:

$$P(X=0) = P(g(U)=0) = P(0 < U < 1-p) = 1-p$$
$$P(X=1) = P(g(U)=1) = P(1-p < U < 1) = p$$

- This g can be viewed as inverse of CDF[1]
- Clearly such g is not unique

[1] For F with jumps, we define $F^{-1} = \min\{x : F(x) \geq u\}$.