# Lecture: Simulation from Some Standard Distributions

Min-ge Xie

Department of Statistics & Biostatistics,
Rutgers University

## Simulation from a distribution

**Assumption:**

Suppose we know how to generate (psudo) random numbers $r_{1,\ldots,}r_k$. So define $r_i^* = r_i \; mod \; M$. Then, we have

$$X \equiv \frac{r_i^*}{M} \sim U(0,1).$$

**Our start point**:

In this class, we assume that we know how to generate from $U(0,1)$.

## Simulation from a distribution

I. Simulation from an exponential distribution: $X \sim \exp(\lambda)$.

- Density function $f(x) = \lambda e^{-\lambda x} \mathbf{1}_{\{x>0\}}$
- **Algorithm**:
  1. Simulate $u \sim U(0,1)$
  2. Compute $x = -\frac{1}{\lambda} \log(u)$
- **Claim**: $x \sim \exp(\lambda)$.
- **Why**? For any $t$,

$$\begin{aligned} P(X \le t) &= P\left(-\tfrac{1}{\lambda}\log(U) \le t\right) = P(\log(U) \ge -\lambda t) \\ &= P\left(U \ge e^{-\lambda t}\right) = 1 - e^{-\lambda t} = \textstyle\int_0^t f(x)\,dx \end{aligned}$$

## Simulation from a distribution

II. Simulation from a known distribution function $F(\cdot)$:

$X \sim F(x)$.

- Known cumulative distribution function $F(t)$:
  - $X \sim F(x) \Rightarrow F(X) \sim U(0,1)$
  - $U \sim U(0,1) \Rightarrow X = F^{-1}(U) \sim F(t)$
- **Algorithm**:
  1. Simulate $u \sim U(0,1)$
  2. Compute $x = F^{-1}(u)$
- **Claim** $X \sim F(\cdot)$.
- **Why**? For any $t$,

$$P(X \leq t) = P\left(F^{-1}(U) \leq t\right) = P(U \leq F(t)) = F(t)$$

- *Remark*: The previous exponential example is a special case of this method. (Hint: find the c.d.f of *exp*($\lambda$).)

## Simulation from a distribution

III. Simulation from the standard normal distribution: $X \sim N(0, 1)$.

- The cumulative distribution function of $N(0, 1)$:

$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) ds$$

  - If $\Phi^{-1}$ can be easily computed, then we can use the previous algorithm.

- We have some faster approaches — introduce instead two other methods.

## Simulation from a distribution

Method A – A simple approximation method

- **Algorithm**
  1. Simulate $u_1, \ldots, u_{12} \overset{i.i.d.}{\sim} U(0,1)$
  2. Compute $x = \sum_{i=1}^{12} u_i - 6$
- **Claim**: $X \sim N(0,1)$!

## Simulation from a distribution

**Why?** By the cent ral limit theorem (CLT):

- By CLT, any sample mean

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(E(X_1), \frac{var(X_1)}{n}\right),$$

  when $n$ is very large.

- In the i.i.d uniform $U(0,1)$ case, we take $n = 12$ $\Longrightarrow \sum_{i=1}^{12} u_i$ is very close to $N(6,1)$.
  - Remember that $E(u_1) = 1/2$ and $var(u_1) = 1/12$

## Simulation from a distribution

**Why?** By the cent ral limit theorem (CLT):

- By CLT, any sample mean

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(E(X_1), \frac{var(X_1)}{n}\right),$$

  when $n$ is very large.

- In the i.i.d uniform $U(0,1)$ case, we take $n = 12$

  $\Longrightarrow \sum_{i=1}^{12} u_i$ is very close to $N(6,1)$.

  - Remember that $E(u_1) = 1/2$ and $var(u_1) = 1/12$
  - Important: The sample mean from i.i.d $U(0,1)$ converges to normal very fast!!

Method B – Box-Muller transformation

- **Algorithm**
  1. Simulate $u_1 \sim U(0,1)$ and calculate $\theta = 2\pi u_1 \sim U(0, 2\pi)$
  2. Simulate $u_2 \sim U(0,1)$ and calculate $e = -2\log(u_2) \sim \chi^2_2, r = \sqrt{e}$
  3. Calculate $x = r\cos\theta$ and $y = r\sin\theta$.
- **Claim**: Both $x$ and $y \sim N(0,1)$!

## Simulation from a distribution

- **Why?**

  Suppose $x$ and $y$ are independent $N(0,1)$ random variables. Then, their joint density is

$$f(x,y) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2} = \frac{1}{2\pi} \exp\left(-\frac{1}{2}\left(x^2+y^2\right)\right).$$

  Take the Box-Muller transformation

$$\begin{cases} r &= \sqrt{x^2+y^2} \\ \cos\theta &= \frac{x}{\sqrt{x^2+y^2}} \end{cases}$$

## Simulation from a distribution

Then,

$$f(r,\theta) = \frac{1}{2\pi}\exp\left(-\frac{r^2}{2}\right)|det(\begin{smallmatrix}\cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta\end{smallmatrix})| = \frac{1}{2\pi}\exp\left(-\frac{r^2}{2}\right) r.$$

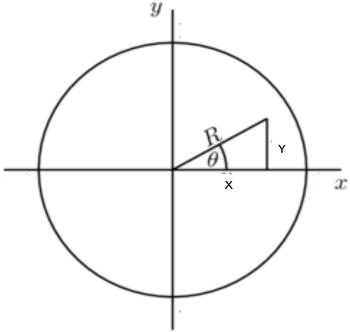
So

$$\theta \sim U(0,2\pi), \quad r^2 \sim \chi^2, \quad \text{and } \theta \text{ is independent of } r^2.$$

(since the above joint density formula is a product of the density of $U(0,2\pi)$ [for $\theta$] and the density of sqrt-$\chi^2$ [for $r$]).

- The inverse of the about calculation is also true. This leads to the algorithm.

# Simulation from a distribution




$$R^2 = X^2 + Y^2$$
$$\cos\theta = \frac{X}{R}$$
$$\sin\theta = \frac{Y}{R}$$

Box-Muller transform from Cartesian coordinates to Polar coordinates

## Simulation from a distribution

IV. Simulation from Binomial and Multinomial distributions

Method A – Simulation from a Bernoulli distribution

$X \sim Bernoulli\,(p)$

- **Algorithm**:
  1. Simulate $u \sim U\,(0,1)$
  2. Set $x = \mathbf{1}_{\{u<p\}}$
- **Claim**: $X \sim Bernoulli\,(p)$
- **Why?**

  The $X$ can only be 0 or 1 with $P(X=1) = P(U<p) = p$.

  $\implies \quad X \sim Bernoulli\,(p)$.

## Simulation from a distribution

Method B – Simulation from a binomial distribution

$X \sim Binomial\,(n, p)$

- **Algorithm**:
  1. Simulate $u_1, \ldots, u_n \overset{i.i.d.}{\sim} U\,(0, 1)$
  2. Set $x = \mathbf{1}_{\{u_1 < p\}} + \ldots + \mathbf{1}_{\{u_n < p\}}$.
- **Claim**: $X \sim Binomial\,(n, p)$
- **Why?** A sum of Bernoulli is binomial.

  (You can also easily get a rigorous math proof of the result yourselves. )

## Simulation from a distribution

Method C – Simulate $X \sim Multinomial\,(n; p_1, \ldots, p_k)$, where $0 < p_i < 1$, $i = 1, \ldots, k$ and $p_1 + \ldots + p_k = 1$.

- **Algorithm**: (illustrate only the $k = 3$ case)
  <u>Case A</u>: $n = 1$ (“Multi-noulli" distribution):
  1. Simulate $u \sim U\,(0, 1)$
  2. Set $(x_1, x_2, x_3) = \left(\mathbf{1}_{\{0<u\leq p_1\}}, \mathbf{1}_{\{p_1<u\leq p_1+p_2\}}, \mathbf{1}_{\{p_1+p_2<u\leq 1\}}\right)$
- **Claim**: $X \sim Multinomial\,(1; p_1, p_2, p_3)$

## Simulation from a distribution

- **Algorithm**: (continue...)
  Case B: General case with $n \geq 1$:
  1. Simulate $u_1, \ldots, u_n \overset{i.i.d.}{\sim} U(0,1)$
  2. Set

$$(x_1, x_2, x_3) = \left( \sum_{i=1}^{n} \mathbf{1}_{\{0 < u_i \leq p_1\}}, \sum_{i=1}^{n} \mathbf{1}_{\{p_1 < u_i \leq p_1 + p_2\}}, \sum_{i=1}^{n} \mathbf{1}_{\{p_1 + p_2 < u_i \leq 1\}} \right)$$

$$= \left( \sum_{i=1}^{n} \mathbf{1}_{\{0 < u_i \leq p_1\}}, \sum_{i=1}^{n} \mathbf{1}_{\{p_1 < u_i \leq p_1 + p_2\}}, n - x_1 - x_2 \right)$$

- **Claim**: $X \sim Multinomial\,(n; p_1, p_2, p_3)$

## Simulation from a distribution

- **Why?** Extension of Bernoulli and binomial cases:
  - Case A – we can easily show $P(X_1 = 1) = p_1$, $P(X_2 = 1) = p_2$, $P(X_3 = 1) = 1 - p_1 - p_2$ and $X_1 + X_2 + X_3 = 1$. So the first claim holds.
  - Case B – it is just a sum of "multinoulli" (case A) $\Longrightarrow$ The second claim holds.

## Simulation from a distribution

V. Simulation from a Poisson distribution with mean $\lambda$:

$X \sim \mathit{Poisson}\,(\lambda)$

- Probability mass function (“density")

$$f\,(x) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad \text{for} \quad x = 0, 1, \ldots$$

Method A: Utilize the connection between $\mathit{Poisson}\,(\lambda)$ and $\mathit{Exponential}(\lambda)$.

- **Algorithm**:
  1. Set $k = 0, s = 0$
  2. $\mathit{while}\,(s \leq 1)$

     simulate $u \sim U(0,\ 1)$ and compute $t = -\dfrac{\log(u)}{\lambda} \sim \exp(\lambda)$;

     set $k = k + 1$ and $s = s + t$
  3. return $x = k - 1$

## Simulation from a distribution

- **Claim**: $X \sim \mathit{Poisson}(\lambda)$
- **Why?**
  - Poisson $x$ is the number of events in $[0,1]$ when the time between consecutive events are i.i.d. *Exponential*$(\lambda)$.
  - Thus, $X$ is the integer $k$ such that $T_1 + \ldots + T_k \leq 1$ but $T_1 + \ldots + T_{k+1} > 1$.

Issue: This algorithm can be slow.

## Simulation from a distribution

Method B: Inverse transform method.

- **Recall**: $U \sim U(0,1) \Rightarrow X = F^{-1}(U) \sim F(x)$
- **Question**: In this discrete case, how can we define $X = F^{-1}(U)$?
  - $F(x) = \sum_{k=0}^{x} P(X = k) = \sum_{k=0}^{x} f(k)$ is the cumulative distribution function of $Poisson(\lambda)$.
  - $U \sim U(0,1)$
- **Answer**: Find/(Search for) the smallest integer $X$ such that $F(X) \geq U$.

## Simulation from a distribution

- **Algorithm**:
  1. Set $f = e^{-\lambda}$, $F = f$, $x = 0$;
     Generate $u \sim U(0,1)$;
  2. While $(F \leq u)$, set
     $x = x + 1,$

     $f = f\lambda/x,$

     $F = F + f;$
  3. return $x$
- **Claim**: $X \sim \mathit{Poisson}(\lambda)$.

## Simulation from a distribution

- **Why?**
  - In the algorithm, we used the formula $f(k+1) = f(k)\lambda/(k+1)$. Or, in a more familiar form, $P(X = k+1) = P(X = k)\lambda/(k+1)$.
  - $P(X \leq x) = P(f(0) + f(1) + ... + f(x) \geq U)$
    $= f(0) + f(1) + ... + f(x) = F(x)$.
- Some consider this a better algorithm (but less intuitive) (only simulate $u \sim U(0,1)$ once).

## Simulation from a distribution

VI. Simulation from a noncentral $\chi^2$ distribution with noncentral parameter $\lambda$ and degrees of freedom $k$: $X \sim \chi_k^2(\lambda)$.

- Density function (when $\lambda = 0$)

$$f(x) = \frac{1}{2^{k/2}\Gamma(k/2)} e^{x/2} x^{k/2-1}, \quad \text{for} \quad x > 0.$$

  (More complicated when $\lambda \neq 0$)

  - Connection to normal:

$$\chi_k^2(\lambda) = \left(Z_1 + \sqrt{\lambda}\right)^2 + Z_2^2 + \ldots + Z_k^2$$

    where $Z_i$'s are independent samples of the standard normal distribution $Z_i \sim N(0, 1)$.

## Simulation from a distribution

- **Idea**: Utilize its relationship with the standard normal.
- **Algorithm**:
  1. Simulate $z_1, \ldots, z_k \sim N(0, 1)$
  2. Set $x = \left(z_1 + \sqrt{\lambda}\right)^2 + z_2^2 + \ldots + z_k^2$
  3. Return $x$
- **Claim**: $X \sim \chi_k^2(\lambda)$.
- **Why?** Obvious... [You can prove yourselves. ]

VII. Simulation from a double exponential distribution with mean $\mu$ and scale $\theta$: $X \sim DE(\mu;\theta)$.

- Density function

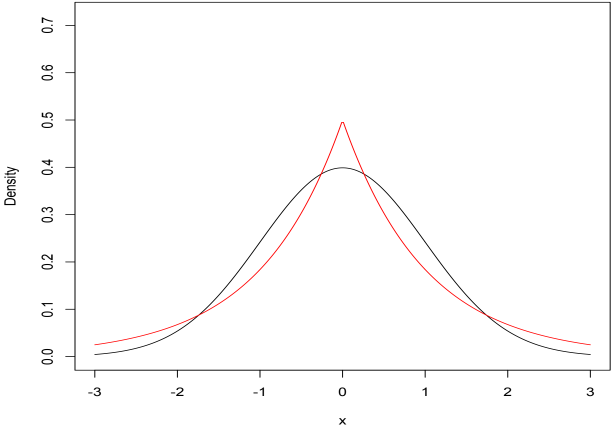
$$f\left(x\right) = \frac{1}{2\theta} exp\left(-\frac{|x-\mu|}{\theta}\right)$$

- Cumulative distribution function:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \frac{1}{2} + \frac{1}{2}\text{sgn}(x-\mu)\left(1 - e^{-\frac{|x-\mu|}{\theta}}\right).$$

- Inverse of the cumulative distribution function.

$$F^{-1}(u) = \mu - \theta\text{sgn}(u - \frac{1}{2})\log\left\{1 - 2\left|u - \frac{1}{2}\right|\right\}.$$




The density curve (in red) of the standard DE distribution (with parameters $\mu = 0$ and $\theta = 1$).

## Simulation from a distribution

Method A: Use inverse transformation method (omit here - you should try it out yourselves)
Method B: Utilize its relationship with the exponential distribution.

- **fact**: When centered at 0, it is a mirror image of exponential distribution $Exp(\theta)$.

$$f_{DE}(x; \mu = 0, \theta) = \frac{1}{2\theta} exp\left(\frac{x}{\theta}\right) \mathbf{1}_{(x<0)} + \frac{1}{2\theta} exp\left(-\frac{x}{\theta}\right) \mathbf{1}_{(x\geq 0)}$$

## Simulation from a distribution

- **Algorithm**:
  1. Simulate $u_1, u_2 \sim U(0, 1)$
  2. Set $z = -\theta \log(u_1)$
  3. If $(u_2 < 0.5)$, set $x = \mu + z$; else, set $x = \mu - z$.
  4. Return $x$
- **Claim**: $X \sim DE(\mu; \theta)$.
- **Why?** Can you show it yourselves?

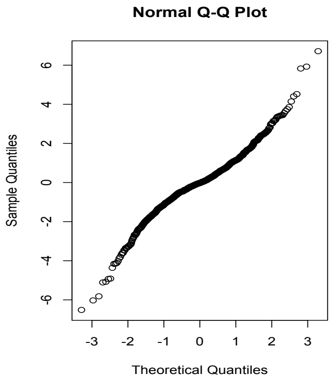
## Simulation from a distribution

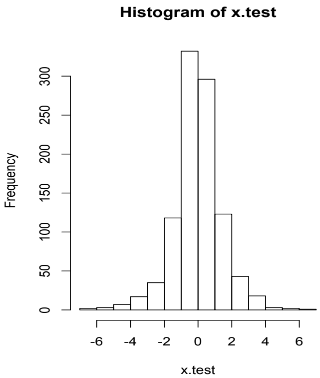
```
##### Example 1: Simulate from a standard double
##### exponential (DE) distribution

rv.de = function(n) {
    x.out = rep(0,n)
    for (ii in 1:n) {
        u1 = runif(1); u2 = runif(1)
        x = - log(u1)
        if (u2 < 0.5) {
        x = -x
        }
    x.out[ii] = x
    }
    x.out
}
```

# Simulation from a distribution

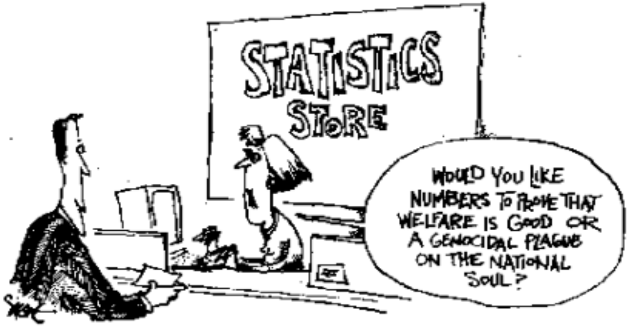
```
> ############
> ## Testing the code
> ############
> x.test = rv.de(1000)
>
> ## summary statistics
> mean(x.test); var(x.test); summary(x.test)
[1] -0.0119083
[1] 1.969701
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-6.51600 -0.66780 -0.02213 -0.01191  0.69840  6.71700
>
> ## Some plots:
> par(mfrow = c(1,2))
> qqnorm(x.test)
> hist(x.test)
```

# Simulation from a distribution




**Observations**: Symmetric around 0 and heavier tails than standard normal (To make it more visible, you can superimpose the corresponding normal distribution line/curve on top of the two figures)




By Signe Wilkinson, Philadelphia Daily News, Cartoonists & Writers Syndicate