Lecture 6: Generating random variables – transformations

Ciaran Evans

Logistics

- Nice work on HW 1!
- Extra office hours today, 12pm - 1pm and 2pm - 4pm
- Upcoming due dates:
  - HW 2: Friday 10am
  - HW 1 resubmissions: next Monday (February 3)
  - Challenge 1 (command line): next Friday (Feb 7) 10am
- New challenges released:
  - Challenge 2: Mersenne Twister (due Feb 28)
  - Challenge 3: Generating Poisson random variables (due Feb 28)
- Remember: you do not have to do all challenges! Pick and choose ones of interest

Previously

Methods to generate $U \sim Uniform(0, 1)$:

- Linear congruential generator
- Mersenne twister
- lots of other variants and alternatives

Now we want to generate random variables with *other* distributions.

Question: If I have a uniform random variable, how can I get other random variables?

use the inverse cdf !

Inverse transform method (continuous r.v.)

Let F be the cdf of a continuous r.v., and suppose F is invertible

Let $U \sim$ Uniform$(0,1)$, and

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

Then $P(X \le t) = F(t)$

i.e. $X \sim F$

Pf: HW 1

Example

Suppose we want to generate $X \sim Exponential(\theta)$ $\quad \theta > 0$

$$F(x) = 1 - e^{-\theta x}$$

$$\Rightarrow F^{-1}(u) = -\frac{1}{\theta}\log(1-u)$$

So: to generate $X \sim Exponential(\theta)$

1) $U \sim Uniform(0,1)$

2) $X = -\frac{1}{\theta}\log(1-U)$

(or $-\frac{1}{\theta}\log(U)$ since U & $1-U$ have same distribution)

Example

Generating $X \sim \textit{Exponential}(1)$:

```
# generate 1000 samples
u <- runif(1000)
x <- -log(u)
```

← simulate uniform

← transform inverse cdf

Question: How can I check that the samples match the desired *Exponential*(1) distribution?

one option: calculate empirical cdf, compare to true Exponential(1) cdf

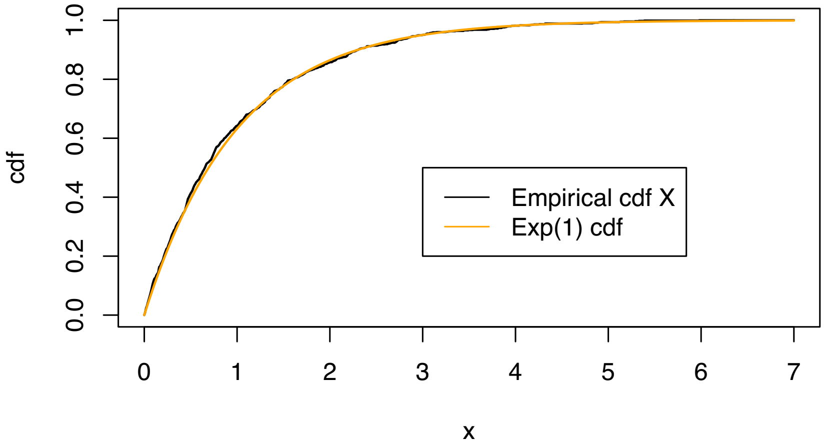
(visually or a hypothesis test e.g. Kolmogorov-Smirnov (KS) test)

Example

Generating $X \sim Exponential(1)$:

looks good!

```r
# generate 1000 samples
u <- runif(1000)
x <- -log(u)
```



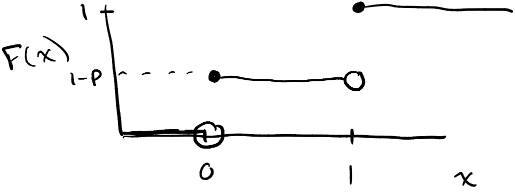
Discrete case

$p(X=1) = p$

$p(X=0) = 1-p$

Suppose that we want to generate $X \sim \mathit{Bernoulli}(p)$

$$F(x) = P(X \leq x) = \begin{cases} 0 & x < 0 \\ 1-p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

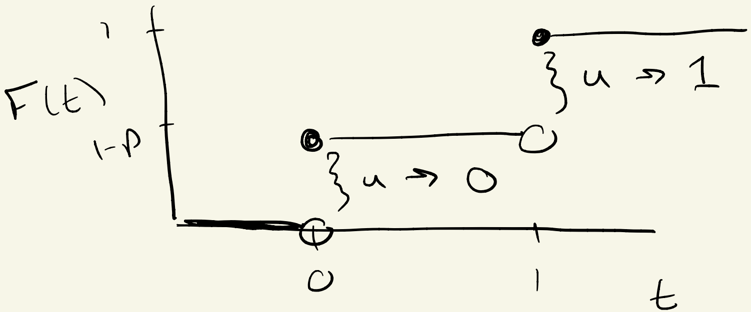


Challenge: F is <u>not</u> invertible!

So: how should we define F^{-1}?

Solution: $F^{-1}(u) = \inf\{t : F(t) \geq u\}$ $\quad \inf(0,1) = 0$

"minimum"



for all $t \in [0,1)$

$F(t) \geq 1-p$

$F(t) = 1-p \qquad t \in [0,1)$

$F(t) = 1 \qquad t \in [1,\infty)$

Let $u \in [0, 1-p]$

$\inf\{t : F(t) \geq u\} = 0$

if $u \leq 1-p$, then $F(t) \geq 1-p \geq u$ for any $t \in [0,1)$

$\to$ smallest t for which $F(t) \geq u$ $= 0$

$$F^{-1}(u) = \begin{cases} 0 & \text{if } u \in [0, 1-p] \\ 1 & \text{if } u \in (1-p, 1] \end{cases}$$

$P(X=0) = P(u \in [0,1-p])$
$= 1-p$

$P(X=1) = P(u \in (1-p,1])$
$= 1-(1-p) = p$

Summarize inverse transform method

Let X be some r.v. with cdf F

0) Define $F^{-1}(u) = \inf\{ t : F(t) \geq u \}$

1) Generate $U \sim \text{Uniform}(0,1)$

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

Then $X \sim F$

Generating a Normal random variable

Suppose we want to simulate $X \sim N(0, 1)$

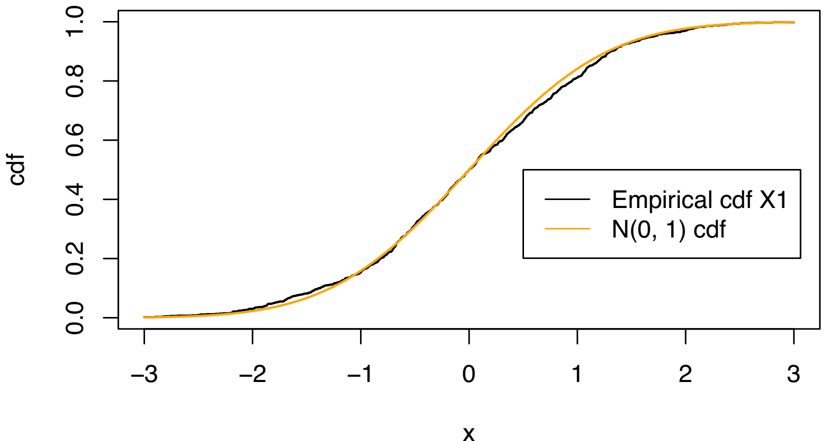
$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\}$$

$$F_X(t) = \text{?}$$

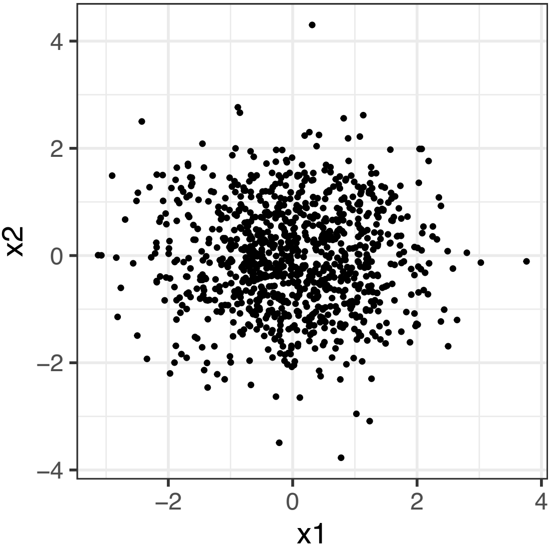
Box-Muller Transformation

Box-Muller in practice

```r
# generate 1000 samples
u1 <- runif(1000)
u2 <- runif(1000)
x1 <- sqrt(-2*log(u1)) * cos(2*pi*u2)
x2 <- sqrt(-2*log(u1)) * sin(2*pi*u2)
```



Box-Muller in practice



Other Normals

Suppose that $Z \sim N(0,1)$. How do I get $X \sim N(\mu, \sigma^2)$?

A few other transformations

- If $X \sim N(\mu, \sigma^2)$, then $e^X \sim Lognormal(\mu, \sigma^2)$
- If $Z_1, ..., Z_k \overset{iid}{\sim} N(0, 1)$, then

$$\sum_{i=1}^{k} Z_i^2 \sim ?$$

- If $V_1 \sim \chi^2_{d_1}$ and $V_2 \sim \chi^2_{d_2}$ are independent, then

$$\frac{V_1/d_1}{V_2/d_2} \sim ?$$

Your turn

Practice questions on the course website:

https://sta379-s25.github.io/practice_questions/pq_6.html

- Practice with inverse transform method
- Start in class. You are welcome to work with others
- Practice questions are to help you practice. They are not submitted and not graded
- Solutions are posted on the course website