DS203: Programming in Data Science
IE605: Engineering Statistics

Introduction to Probability and Statistics
Lecture 03

Manjesh K. Hanawal

27th August 2020

Previous Lecture:

- Random Variable (RVs)
- Discrete and Continuous RVs
- Cumulative density functions (CDFs)
- Probability Density functions (PDFs)
- Examples of discrete RVs
- Examples of Continuous RVs

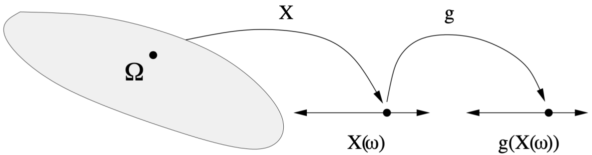
This Lecture:

- Functions of random variable
- Generate samples from a given distribution

Function of random variable

For any $g : \mathbb{R} \to \mathbb{R}$ and a RV X on Ω. We can define

$$\boxed{Y = g(X), \text{ i.e., for all } w \in \Omega, Y(w) = g(X(w))}$$



Example 1: Absolute error. $Y = |X|$
Example 2: Hinge loss. $Y = \max\{0, X\}$
Example 3: Linear function. $Y = aX + b$ for some $a, b \in \mathbb{R}$

Distribution of function of RVs

Let $Y = g(X)$ and F_X is the CDF of X. What is cdf of Y?

- $F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(w : g(X(w)) \leq y)$
- Can be expressed as $F_Y(y) = P(X \in \mathcal{A})$
- Set $\mathcal{A}$ depends on g and y.

Example: (X is Discrete case) PMF of Y:
$P_Y(y) = P(Y = y) = P(g(X) = y) = \sum_{x:g(x)=y} P_X(x)$

Example: (Continuous case) PDF of Y:
Obtain $F_Y(y)$ for all $y \in \mathbb{R}$ and then differentiate.

$$E[Y] = E[g(X)] = \int g(x) f_X(x) dx$$

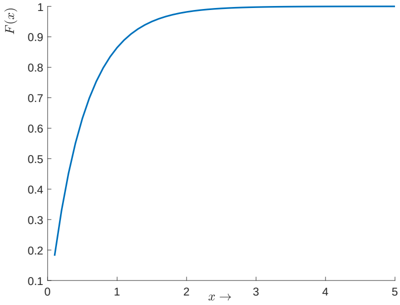
Law of The Unconscious Statistician (LOTUS!)

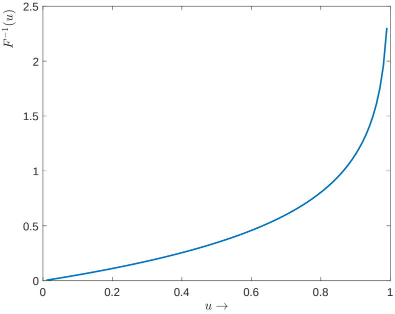
Simulation of Given Distribution

A CDF F is given. How to generate samples with CDF F?

Let $U \sim Unif(0,1)$.

- If F is continuous, define $X = g(U)$ where $g(u) = F^{-1}(u)$
- Claim: X has CDF F
- $P(X \le x) = P(F^{-1}(U) \le x) = P(U \le F(x)) = F(x)$

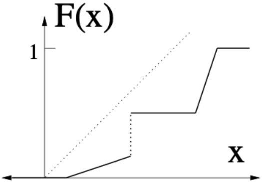


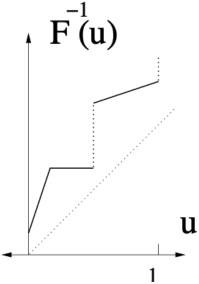


Simulation of Given Distribution contd...

F is not continuous

- Define $g(u) := F^{-1}(u) = \min\{x : F(x) \leq u\}$ for $0 < u < 1$
- for any x, u, $F^{-1}(u) \leq x$ if and only if $u \leq F(u)$ (verify!)
- Define $X = g(U)$. Then $P(X \leq x) = F_X(x)$.





How to generate Uniform RVs?

- Linear Congruential Generator (LCG) ($x_i = a_0 + a_1 x_{i-1} \mod M$)
- Multiplicative Recursive Generator (MRG)
- Lagged Fibonacci Generator (LFG)
- Inverse Congruential Generator (IVG)
- Linear Feedback Shift Register (LFSR)
- Pseudo Random Number Generators