

Introduction to Statistical Data Science

October 1, 2018

1 Lecture 1

1.1 Normal Distribution

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

$$F(x) \equiv P(X \leq x)$$

Central Limit Theorem (for normal distribution)

The more random variables we average over, the closer the resulting distribution will be to the Normal distribution

Parameter: u The mean is the location parameter.

Parameter: σ^2 The variance is the scale parameter

1.2 Uniform Distribution

$$X \sim U[0, 1] \tag{1}$$

$$0 \leq x \leq 1 \tag{2}$$

$$p(x) = 1 \tag{3}$$

$$F(x) = P(X \leq x) = x \tag{4}$$

$$\tag{5}$$

Use uniform distribution to construct normal distribution

$$\mathbf{X} = \begin{bmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(n)} \end{bmatrix} \quad (6)$$

$$X^{(i)} \sim U[0, 1]$$

$$Y = \frac{1}{n} \sum_{i=1}^n X^{(i)}$$

subsampling \mathbf{X} vector to construct Y . The distribution of $Y_j \sim ?, j = 1, 2, 3, \dots, p$ would close to normal distribution.