Last lecture

Uniform Distribution (Ch 3.3)

Exponential Distribution (Ch 3.4)

- Memoryless property
- Connect to Geo(p)

Poisson process (Ch 3.5)

- Motivation
- Bernoulli process to Poisson process
- Definition
- Properties

Agenda

Gaussian (normal) Distribution (Ch 3.6.2)
- Example

The Central Limit Theorem and Gaussian Approximation (Ch 3.6.3)
- Definition
- CDF Approximation
- Examples

ML estimation for continuous RVs (Ch 3.7) – Will not be tested
- Definition
- Examples

Examples

Suppose $\mu_X = 10$ and $\sigma_X^2 = 3$. Compute $P\{X < 10 - \sqrt{3}\}$ if

- X is a Gaussian RV in terms of Q
- X is a uniform RV

(Hint: $10 - \sqrt{3} \approx 8.27$)

Slido

#4093788

Choose all the correct answers

(a) $\Phi(u) + \mathrm{Q}(u) > 1$

(b) $\Phi(u) - \mathrm{Q}(-u) = 0$

(c) $\Phi(0.5) + \mathrm{Q}(-0.5) > 0$

(d) $\Phi(0) = 0.5$

(e) $\mathrm{Q}(x)$ is monotonically increasing

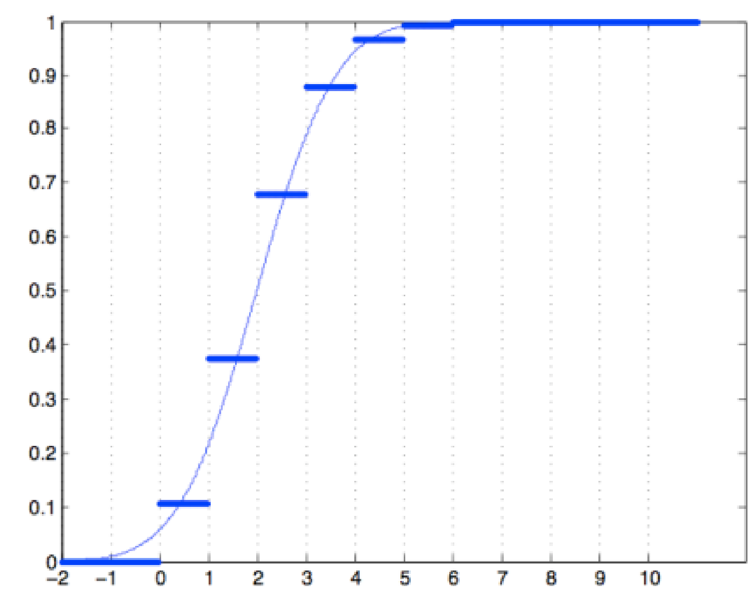
Central Limit Theorem and Gaussian Approximation

Central Limit Theorem (CLT)

If many independent random variables are added together, and if each of them is small in magnitude compared to the sum, then the sum X has an approximately Gaussian distribution $\tilde{X}$.

- $P\{X \leq v\} \approx P\{\tilde{X} \leq v\}$
- E.g., $X \sim Binomial(n, p)$ when np and $n(1-p)$ are not small
  - $(n, p) = (10, 0.2)$
  - $\mu_X =$
  - $\sigma_X^2 =$

  - What if np is small?

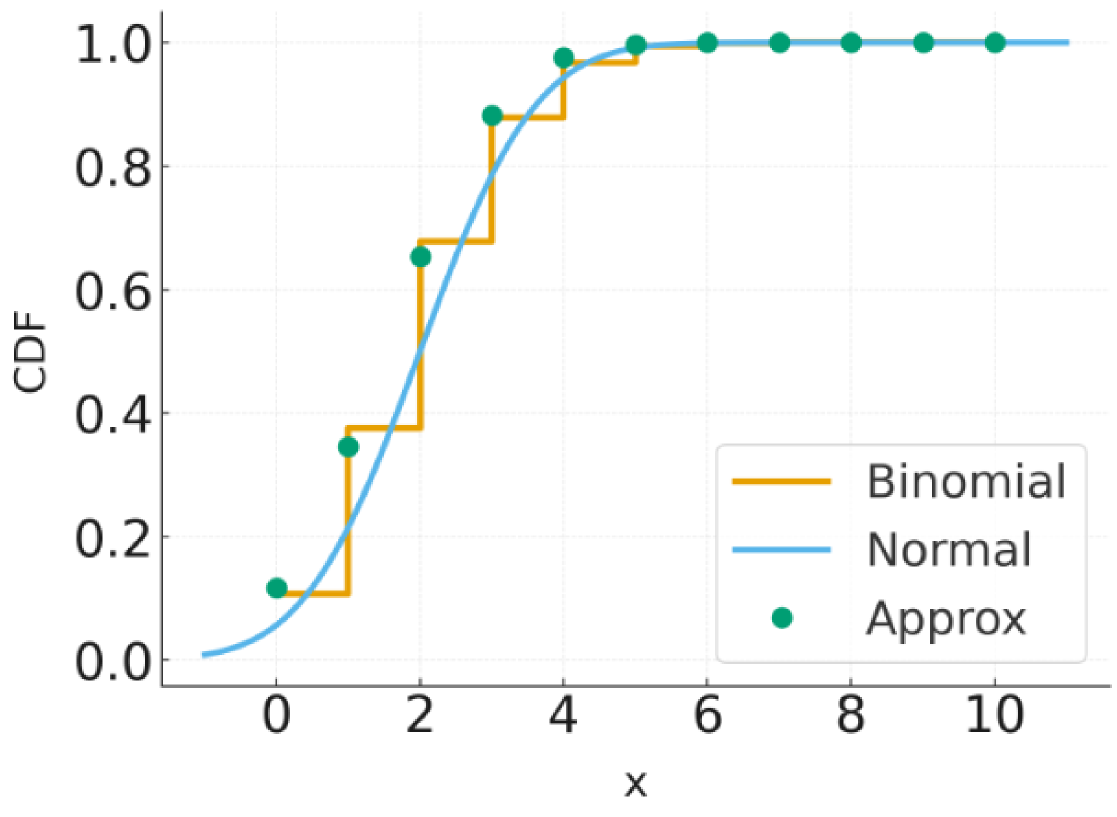


Gaussian Approximation

Approximate $X \sim Bin(10, 0.2)$ with $\tilde{X} \sim N(2, 1.6)$

- $F_X(2) = F_X(2.1) = F_X(2.9)$
- $F_{\tilde{X}}(2) \neq F_{\tilde{X}}(2.9)$
- How should we approximate?

- $P\{X \leq k\} \approx$
- $P\{X < k\} \approx$
- $P\{X \geq k\} \approx$
- $P\{X > k\} \approx$



Standardized Binomials

Standardized $S_{n,p} \sim Bin(n,p)$

- $X = \frac{S_{n,p} - np}{\sqrt{np(1-p)}}$

DeMoivre-Laplace limit theorem (First CLT version)

$$\lim_{n \to \infty} P\left\{\frac{S_{n,p} - np}{\sqrt{np(1-p)}} \le c\right\} = \Phi(c)$$

Slido

#2750821

Each user independently opens today's push notification with probability $p = 0.3$.

- You send it to $n = 200$ users
- Let X be the number of opens
- $P\{X < 70\}$ in terms of Φ?

$$(a)\ \ \Phi\left(\frac{70 - \mathrm{np}}{\sqrt{np(1-p)}}\right)$$

$$(b)\ \ \Phi\left(\frac{69.5 - \mathrm{np}}{\sqrt{np(1-p)}}\right)$$

$$(c)\ \ \Phi\left(\frac{70.5 - \mathrm{np}}{\sqrt{np(1-p)}}\right)$$

$$(d)\ \ \Phi\left(\frac{70}{\sqrt{np(1-p)}}\right)$$

Example

$X \sim Bin(n = 1000, p = 0.5)$, Using Gaussian approximation, find K s.t. $P\{X \geq K\} \approx 0.01 = \mathrm{Q}(2.325)$

- $\mu_X =$ $\quad$ $, \sigma_X = \sqrt{np(1-p)} = \sqrt{250} \approx 15.8$
- $P\{X \geq K\} =$
- What if $n = 1000000$?

Example

We want to estimate p with $\hat{p} = \frac{X}{n}$

- Find $P\{|\hat{p} - p| < \delta\}$ in terms of n, p, δ, and Φ

- Find δ w/ 99% confidence if $p = 0.5$, $n = 1000$. Given that $\Phi(2.58) \approx 0.995$

- What if $p = 0.1$?

ML estimation

Definition

Recall for discrete RV X, given observation u, ML is to find θ maximizing $p_\theta(u)$

- But for continuous RV, $p_\theta(u) = 0$
- Instead, ML maximize $f_\theta(u)$ because
  - $f_\theta(u) \approx \frac{1}{\epsilon} P\{u - \frac{\epsilon}{2} < X < u + \frac{\epsilon}{2}\}$
  - $\hat{\theta}_{ML}(u) \triangleq argmax_\theta \, f_\theta(u)$

Example

$X \sim Uni([0, b])$ where b is unknown. We want to estimate b with observation u

- $f_X(u) =$

- Extrema happens at

Functions of a random variable

Find CDF/ PDF of g(X)

Motivation – I know X follows some distribution

- but what about $Y = g(X)$?

1. Scope the problem - Find of X and Y, are they continuous or discrete?
2. Find $F_Y(c)$ from integrating $f_X(x)$ over $\{x: g(x) \leq c\}$
   - If Y is discrete, normally we can find pmf $p_Y(c)$
3. Get $f_Y = F_Y{}'$

Examples

1. Find support and continuity
2. $F_Y(c) = \int_{x:f(x)\le c} f_X(x)dx$
3. $f_Y = {F_Y}'$

RV X follows $f_X(u) = \frac{e^{-|u|}}{2}$ for $u \in R$. $Y = X^2$. Find f_Y, μ_Y and σ_Y^2

1. $F_Y =$

2. $P\{X \le c\} =$

3. $f_Y(c) =$