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

1. Standardize.

$$Z = \frac{X-\mu}{\sigma_X} = \frac{X-10}{\sqrt{3}}$$

2. $P\{X < 10-\sqrt{3}\} = P\left\{\frac{X-10}{\sqrt{3}} < -1\right\} = P\{Z < -1\}$

$$= \Phi(-1)$$

$$= 1 - Q(-1) = Q(1)$$

X is uniform (a,b) $\quad \mu_X = \frac{a+b}{2} = 10$

$a = 10-k \quad b = 10+k$.

Recall $W \sim Uni(0,1)$ $\quad \sigma_W^2 = \frac{1}{12}$.

$\sigma_X^2 = 3 = \sigma_W^2 \times 36$

$X = cW + d$.

$\sigma_X^2 = \sigma_W^2 \times c^2$

$c = 6$.

$b-a=6$

$a=7, b=13$.

$X \sim Uni(7,13)$

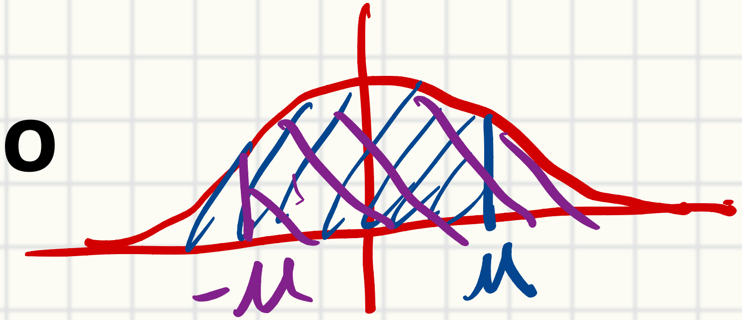
$P\{X < 8.27\}$

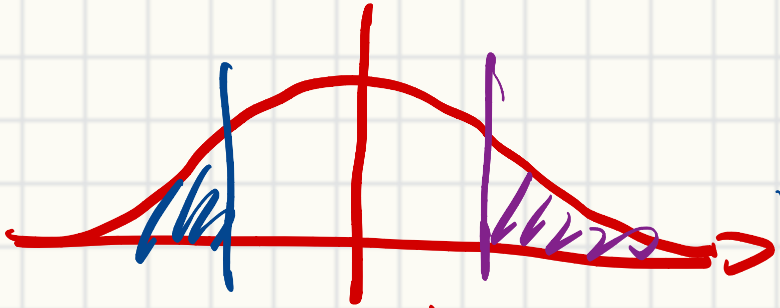
$= \int_{-\infty}^{8.27} f_X(u)\,du$

$= \int_{7}^{8.27} \frac{1}{6}\,du$

$= \frac{1.27}{6}$

Slido





Choose all the correct answers

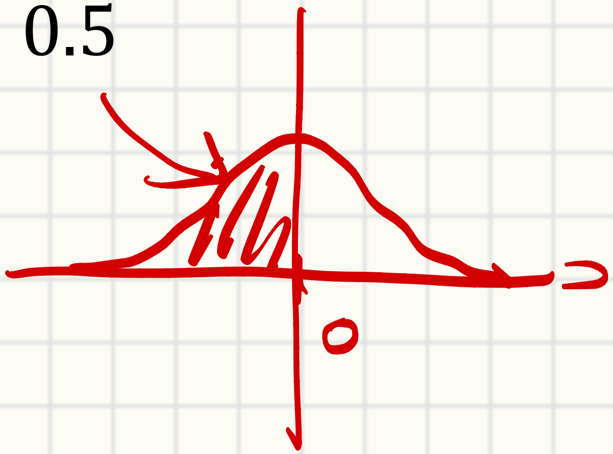
(a) $\Phi(u) + Q(u) > 1$

$F_X + F_X^c = 1.$

(b) $\Phi(u) - Q(-u) = 0$

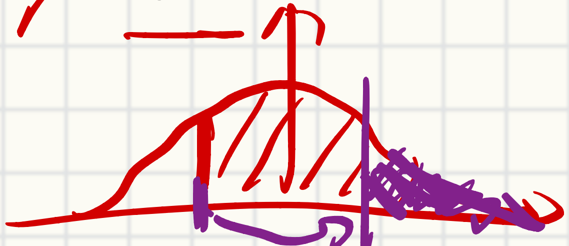
(c) $\Phi(0.5) + Q(-0.5) > 0$

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



(e) $Q(x)$ is monotonically increasing

decreasing



#4093788

Central Limit Theorem and Gaussian Approximation

Central Limit Theorem (CLT)

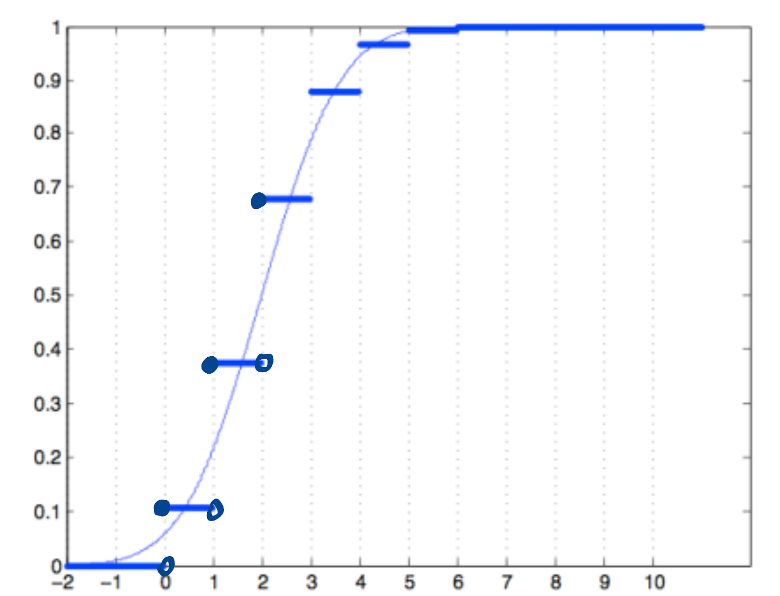
If many independent random variables are added together, and if each of them is small in magnitude compared to the sum, then the sum X has an approximately Gaussian distribution $\tilde{X}$.

- $P\{X \le v\} \approx P\{\tilde{X} \le v\}$
- E.g., $X \sim Binomial(n, p)$ when np and $n(1-p)$ are not small — Sum of n Bernoulli
  - $(n, p) = (10, 0.2)$
  - $\mu_X = np = 2$
  - $\sigma_X^2 = np(1-p) = 1.6$

$\tilde{X} = N(2, 1.6)$

  - What if np is small? Poisson

or p is small



Gaussian Approximation

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

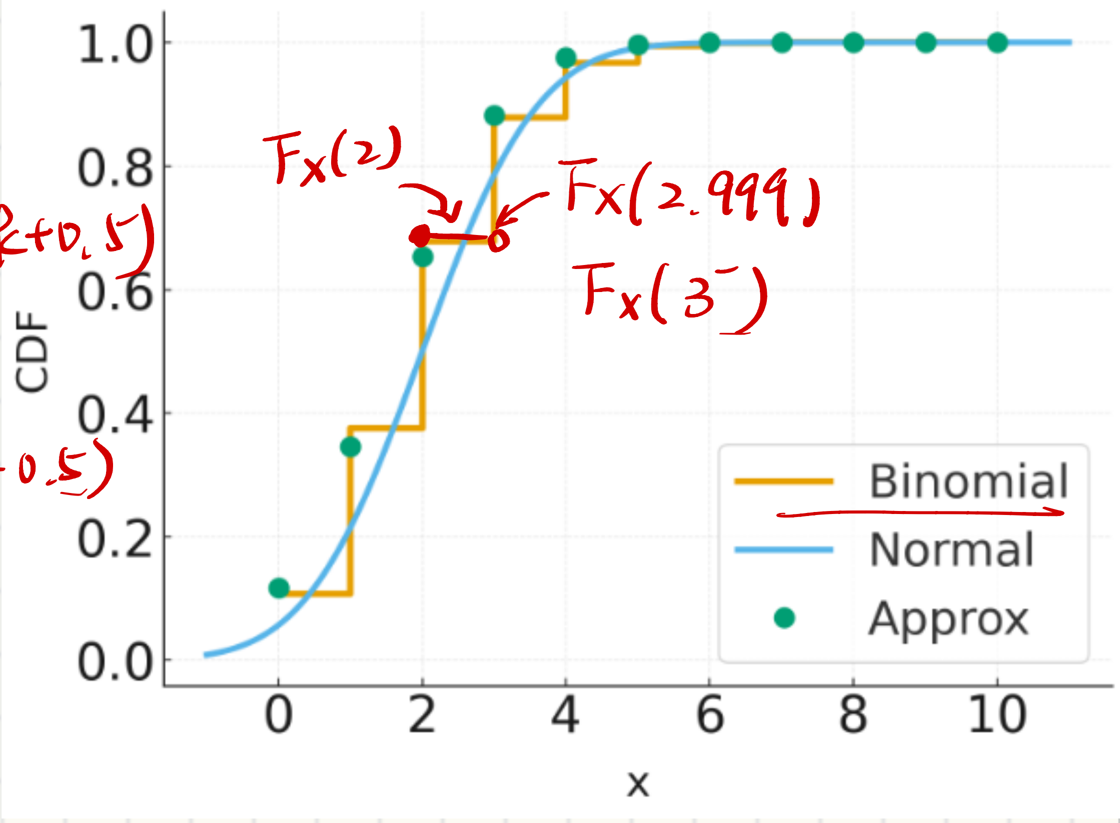
- $F_X(2) = F_X(2.1) = F_X(2.9) = \sum_{k=0}^{2} P_X(k)$
- $F_{\tilde{X}}(2) \neq F_{\tilde{X}}(2.9)$
- How should we approximate?

$= P\{X < k+1\}$

- $P\{X \le k\} \approx P\{\tilde{X} \le k+0.5\} = F_{\tilde{X}}(k+0.5)$ ①
- $P\{X < k\} \approx P\{\tilde{X} \le k-0.5\} = F_{\tilde{X}}(k-0.5)$
- $P\{X \ge k\} \approx$
- $P\{X > k\} \approx$

② $P\{X \le k-1\} \approx P\{\tilde{X} \le k-1+0.5\}$

$= F_{\tilde{X}}(k-0.5)$



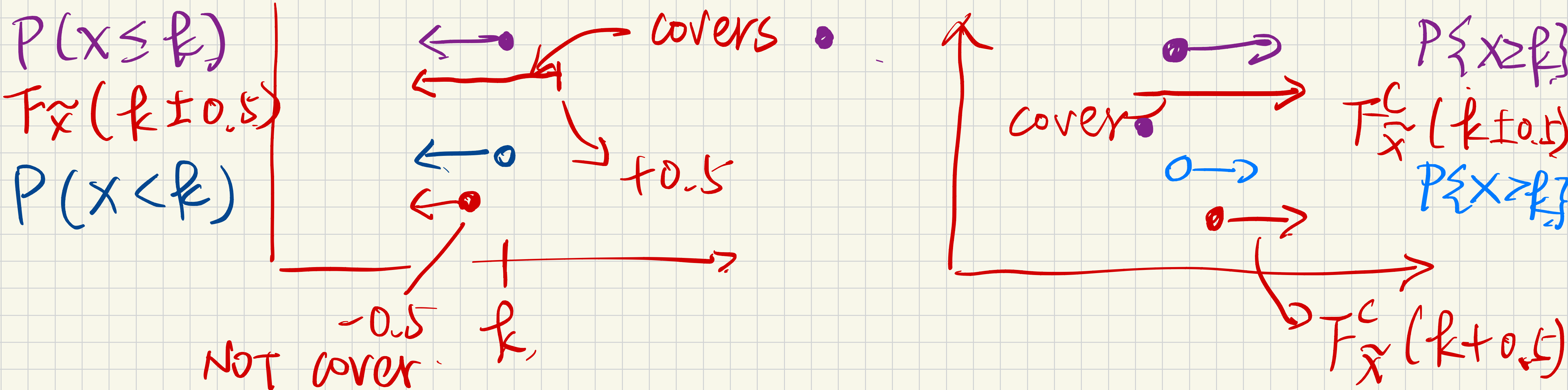
$$P\{X \ge k\} = 1 - P\{X < k\}$$

$$\approx 1 - F_{\tilde{X}}(k-0.5) = F^{c}_{\tilde{X}}(k-0.5)$$

$$P\{X > k\} = 1 - P\{X \le k\} = 1 - F_{\tilde{X}}(k+0.5)$$

$$\approx F^{c}_{\tilde{X}}(k+0.5)$$

Visualization → Interval should cover the condition.



Standardized Binomials

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

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

$X \sim N(np, np(1-p)$

DeMoivre-Laplace limit theorem (First CLT version)

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

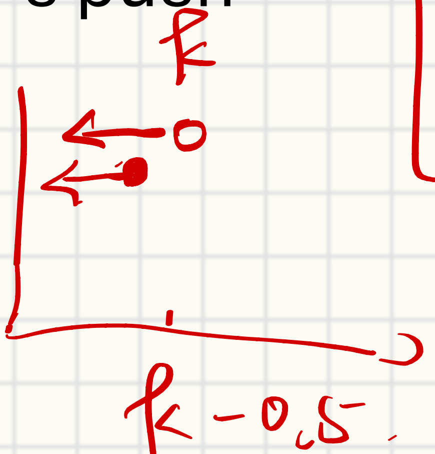
$\tilde{Z}$

Slido

$P\{X<70\} \simeq P\{\tilde{X} \le 69.5\} = P\{\frac{\tilde{X}-np}{\sqrt{np(1-p)}} \le c\}$

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?



#2750821

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

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

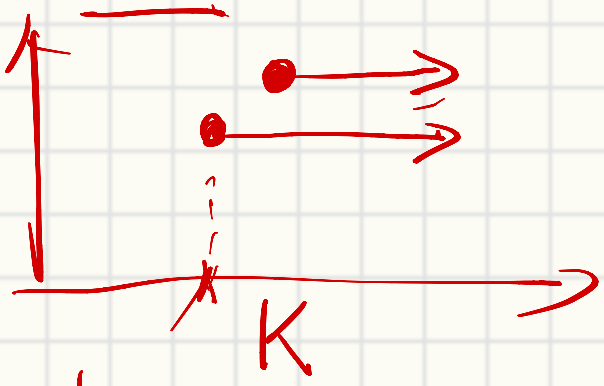
(c) $\Phi\left(\frac{70.5 - \mathrm{np}}{\sqrt{np(1-p)}}\right)$

(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. $P\{X \geq K\} \approx 0.01 = Q(2.325)$

- $\mu_X = 500$, $\sigma_X = \sqrt{np(1-p)} = \sqrt{250} \approx 15.8$
- $P\{X \geq K\} = P\{\tilde{X} \geq K - 0.5\}$
- What if $n = 1000000$?



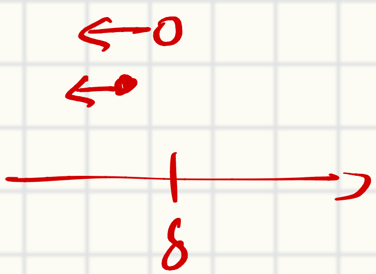
$$P\left\{\frac{\tilde{X} - 500}{15.8} \geq \frac{K - 500.5}{15.8}\right\} = 0.01$$

$$P\{Z \geq C\} = Q(2.325) \qquad \frac{K - 500.5}{15.8} = 2.325$$

$K = 537.26$

$$\frac{K - 500000.5}{500} = 2.325 \qquad K = 501163$$

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

$$P\{|\tfrac{X}{n} - p| < \delta\} \approx P\{|\tfrac{\tilde{X}}{n} - p| \le \delta - 0.5\} = P\{|\tfrac{\tilde{X} - np}{\sqrt{np(1-p)}}| \le \quad\}$$

$$c = \delta\sqrt{\frac{n}{p(1-p)}}$$

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

$$= \Phi(c) - \Phi(-c)$$

- What if $p = 0.1$?