ECE 313 / MATH 362 PROBABILITY WITH ENGINEERING APPLICATIONS

LECTURE 25 : THE CENTRAL LIMIT THEOREM AND THE GAUSSIAN APPROXIMATION

- TOPICS TO COVER (BASED ON CH 3.6)

  → THE CENTRAL LIMIT THEOREM AND THE GAUSSIAN APPROXIMATION

→ THE CENTRAL LIMIT THEOREM AND THE GAUSSIAN APPROXIMATION

MAIN IDEA: IF MANY INDEP. RVS ARE ADDED TOGETHER, AND IF EACH OF THEM IS SMALL IN MAGNITUDE COMPARED TO THE SUM, THEN THE SUM HAS AN APPROXIMATELY GAUSSIAN DISTRIBUTION.

THAT IS, IF THE SUM IS X WITH

$$EX = \mu \quad \text{AND} \quad \mathrm{Var}\, X = \sigma^2,$$

AND IF $\tilde{X} \sim N(\mu, \sigma^2)$

THEN THE CDF OF X AND THE CDF OF $\tilde{X}$ ARE APPROXIMATELY THE SAME, I.E.,

$$F_X(c) := P(X \le c) \approx F_{\tilde{X}}(c) := P(\tilde{X} \le c) \qquad \text{(GAUSSIAN APPROXIMATION)}$$

AN IMPORTANT SPECIAL CASE OF THE GAUSSIAN APPROXIMATION

$X :=$ SUM OF n INDEP. BERNOULLI RVS, EACH WITH PARAMETER p

$\Rightarrow X \sim \mathrm{BINOMIAL}(n, p)$

EXAMPLE: $X \sim \text{BINOMIAL}(10, .2)$

$EX = 2$ AND $\text{Var}\,X = 1.6$

LET $\tilde{X} \sim N(2, 1.6)$. WE PLOT THE CDFS OF X AND $\tilde{X}$ AS FOLLOWS:

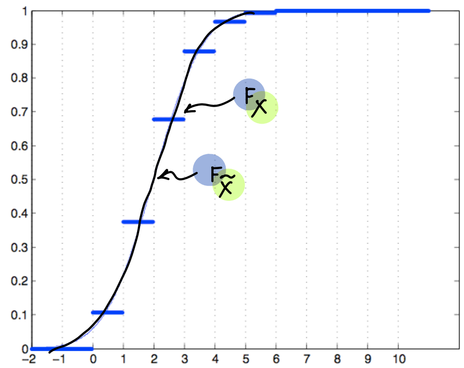

Figure 3.12: The CDF of a binomial random variable with parameters $n = 10$ and $p = 0.2$, and the Gaussian approximation of it.

NOTE THAT F_X AND $F_{\tilde{X}}$ ARE CLOSE BUT NOT EVERYWHERE AS F_X IS PIECEWISE CONSTANT (STEP FUNCTION) WITH JUMPS AT INTEGER POINTS, AND $F_{\tilde{X}}$ IS CONTINUOUS.

NOTICE, HOWEVER THAT F_X AND $F_{\tilde{X}}$ ARE PARTICULARLY CLOSE WHEN THE ARGUMENT IS HALFWAY BETWEEN TWO INTEGERS, I.E.,

$$F_X(k+0.5) \approx F_{\tilde{X}}(k+0.5) \; ; \; k \text{ INTEGER}$$

HOW? : TABLES!

FOR EXAMPLE: $F_X(2.5) = 0.6678 \approx F_{\tilde{X}}(2.5) = 0.6537$

SINCE X IS AN INTEGER-VALUED RV, $F_X(2) = F_X(2.5) \approx F_{\tilde{X}}(2.5)$

WHY?

SO, $F_{\tilde{X}}(2.5)$ IS A FAIRLY GOOD APPROXIMATION OF $F_X(2)$. IN PARTICULAR, ITS A BETTER APPROXIMATION THAN $F_{\tilde{X}}(2)$ IS — IN THIS CASE $F_{\tilde{X}}(2) = 0.5$.

IN GENERAL, WHEN X IS AN INTEGER-VALUED RV, THE GAUSSIAN APPROXIMATION WITH CONTINUITY CORRECTION IS :

$$F_X(k) \approx F_{\tilde{X}}(k + 0.5)$$

WHY SUBTRACTION ?

SIMILARLY, WE HAVE: $P\{X \geq k\} \approx P\{\tilde{X} \geq k - 0.5\}$

FURTHERMORE, THE APPROXIMATION IMPROVES AS n INCREASES.

FOR EXAMPLE: $X \sim$ BINOMIAL $(30, .2)$

$EX = 6$ AND $\text{Var}\, X = 4.8$

LET $\tilde{X} \sim N(6, 4.8)$. WE PLOT THE CDFS OF X AND $\tilde{X}$ AS FOLLOWS:

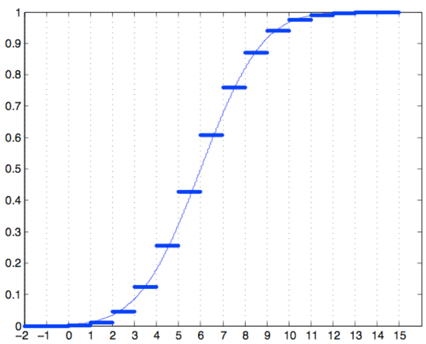

Figure 3.13: The CDF of a binomial random variable with parameters $n = 30$ and $p = 0.2$, and the Gaussian approximation of it.

FORMAL STATEMENT OF THE (FIRST) CENTRAL LIMIT THEOREM BY DEMOIVRE AND LAPLACE

n : POSITIVE INTEGER AND $0 < p < 1$

$S_{n,p} \sim$ BINOMIAL (n, p) , I.E., $E\, S_{n,p} = np$ AND $\mathrm{Var}\, S_{n,p} = np(1-p)$

$\Rightarrow$ STANDARDIZED VERSION OF $S_{n,p} = \dfrac{S_{n,p} - np}{\sqrt{np(1-p)}}$

FOR A FIXED $p \in (0,1)$ AND ANY CONSTANT c :

$F_{S_{n,p}}(c) \longrightarrow \Phi(c)$ AS $n \longrightarrow \infty$ (CONVERGENCE IN DISTRIBUTION)

EQUIVALENTLY : $\lim\limits_{n \to \infty} P\left\{ \dfrac{S_{n,p} - np}{\sqrt{np(1-p)}} \leq c \right\} = \Phi(c)$

ALSO $1 - F_{S_{n,p}}(c) \longrightarrow 1 - \Phi(c) := Q(c)$ AS $n \longrightarrow \infty$

EQUIVALENTLY : $\lim\limits_{n \to \infty} P\left\{ \dfrac{S_{n,p} - np}{\sqrt{np(1-p)}} \geq c \right\} = Q(c)$

EXAMPLE : (a) A FAIR COIN IS FLIPPED A THOUSAND TIMES. (INDEPENDENTLY)

X : # OF TIMES HEADS SHOWS

USE THE GAUSSIAN APPROX. WITH CONT. CORRECTION TO FIND K SUCH THAT

$P\{X \geq K\} \approx 0.01$

SOLUTION: $X \sim \text{BINOMIAL}(n = 1000, p = 0.5)$

$$\Rightarrow \quad EX = 500$$

$$\sqrt{\operatorname{Var} X} = \sqrt{1000(0.5)(0.5)} = \sqrt{250} \approx 15.8$$

X IS INTEGER-VALUED

$$\therefore \quad P\{X \geq K\} = P\{\tilde{X} \geq K - 0.5\}$$

$$= P\left\{\frac{\tilde{X} - 500}{15.8} \geq \frac{K - 0.5 - 500}{15.8}\right\}$$

$$\approx Q\left(\frac{K - 0.5 - \mu}{\sigma}\right)$$

FROM TABLES, $Q(2.235) \approx 0.01$

$\therefore$ WE WANT K SUCH THAT $\frac{K - 0.5 - 500}{15.8} = 2.235$

$$\Rightarrow \quad K = 537.26$$

$$\Rightarrow \quad K = 537 \text{ OR } 538 \text{ SHOULD WORK}$$

INTERPRETATION: IF A FAIR COIN IS FLIPPED A THOUSAND TIMES, THERE IS A 1% CHANCE TO GET HEADS MORE THAN 53.7% OF THE TIMES!

(b) REPEAT, BUT NOW ASSUME THE COIN IS FLIPPED A MILLION TIMES!

SOLUTION: EXERCISE!