Outline

1. Mean and SD
   - Mean and SD
   - Expected Value and SD of Binomial Random Variable

2. Normal Approximation
   - Normal Approximation for Binomial Probabilities
   - When Can One Use Normal Approximation
TV World Example, revisited

Recall that, in TV World Example, historical data shows that 20% of TV buyers at TV World purchase extended warranty. $X$ extended warranties were sold along with the 300 TV sets sold last quarter. Then $X \sim \text{binomial}(n = 300, p = 0.20)$. The expected number of extended warranties sold last quarter is around

$$\text{expected value } \approx 60 \quad \text{give or take } 7 \quad \text{SD}$$
Expected Value and SD
of a binomial random variable

If \( X \sim binomial(n, p) \) then

\[
\mu = E(X) = \text{Expected Value (or Mean or Average)} = E(X) = np = (\text{sample size}) \times (\text{prob. of success})
\]

\[
\sigma_X = SD(X) = \sqrt{npq} = \sqrt{(\text{sample size}) \times (\text{prob. of success}) \times (\text{prob. of failure})}.
\]
Expected Value and SD

TV World example

The number of warranties sold last quarter

\( X \sim \text{binomial}(n = 300, p = 0.20) \)

\[
\mu = np = 300 \times 0.20 = 60 \quad \text{and} \quad \sigma = \sqrt{npq} = \sqrt{300 \times 0.20 \times 0.80} \approx 7.
\]

Sometimes more (than 60), sometimes less.
By 7 (more or less), on average.
Expected Value and SD

5-question quiz example

Recall that, by pure guessing, the number of correct answers $X \sim \text{binomial}(n = 5, p = 0.20)$. If someone guesses all questions randomly then on average, he/she will get $\mu = 5 \times 0.20 = 1$ give or take $\sigma = \sqrt{5 \times 0.20 \times 0.80} \approx 0.9$ correct answers.
Expected Value and SD

Campus Video example

5% of videos rented at Campus Video incur a late rental fee. If 700 videos were rented last week, the number of videos that will incur late rental fees will be around __ give or take ____
Expected Value and SD

Campus Video example

5% of videos rented at Campus Video incur a late rental fee. If 700 videos were rented last week, the number of videos that will incur late rental fees will be around

\[35 \text{ give or take } \pm \]

\[\mu = 700 \times 0.05 = 35\]
5% of videos rented at Campus Video incur a late rental fee. If 700 videos were rented last week, the number of videos that will incur late rental fees will be around

35 give or take 5.77

\[
\mu = 700 \times 0.05 = 35
\]

\[
\sigma = \sqrt{700 \times 0.05 \times 0.95} = 5.77
\]
iClicker Question 10.1

A study was conducted concerning the use of gloves among the nurses with 15 years or more experience. The study showed that only $\frac{1}{6}$ of these nurses wear gloves during vascular access procedures. For a sample of $n=36$ nurses with 15 years or more experience, the number of nurses wear gloves during vascular access procedures is expected to be

A. 36
B. 6
C. 0
D. 15
E. $-1$
1 Mean and SD
   - Mean and SD
   - Expected Value and SD of Binomial Random Variable

2 Normal Approximation
   - Normal Approximation for Binomial Probabilities
   - When Can One Use Normal Approximation
Binomial Probability Histogram

Binomial(n=15, p=0.5)

$P(X=x) = \text{bar height}$

$P(X=x) = \text{bar area}$

Note: if probability of a success is 0.5, the shape is symmetric about $n/2$. 

Binomial(n=14, p=0.5)

$P(X=x)$

Note: if probability of a success is 0.5, the shape is symmetric about $n/2$. 

JC Wang (WMU)
Note that

- The binomial distribution is symmetric when $p = 0.5$
  Consequently, each rectangle in the probability histogram is centered at an integer with a width of 1. This is also true when $p \neq 0.5$.

- Hence, for integer $a$, $a = 0, \ldots, n$, $P(X = a) = P(a - 0.5 < X < a + 0.5) = \text{area of rectangle centered at } a \approx \text{area under normal curve between } a - 0.5 \text{ and } a + 0.5$ where normal curve $\sim N(\mu = np, \sigma = \sqrt{npq})$.

- So, we have (next slide):
Binomial Probability Histogram and Normal Curve

X ~ Binomial(n=30, p=0.4)

np = 12, √npq = 2.68

Binomial Probability

Y ~ Normal(μ = 12, σ = 2.68) Probability

- actual probability
- approximate probability

\[ P(X=14) = \text{area of rectangle over 14} \approx P(13.5 < Y < 14.5) = P(0.56 < Z < 0.93) = 0.8238 - 0.7123 = 0.1115 \]

\[ P(X \leq 10) = \text{area of rectangles over 10 and to its left} \approx P(Y < 10.5) = P(Z < -0.56) = 1 - 0.7123 = 0.2877 \]

\[ P(X \geq 16) = \text{area of rectangles over 16 and to its right} \approx P(Y > 15.5) = P(Z > 1.31) = 1 - 0.9049 = 0.0951 \]
An Example on Using Normal Approximation

Suppose a student in an introductory Statistics course has not been attending class this semester but decides to take the exam anyway. If he randomly guesses on each of the 25 questions, then he has a 1 out of 5 chance of getting a correct answer, since it is a multiple choice exam with choices a, b, c, d, or e. How many questions should the student expect to get correct on this exam, give or take by how many questions?

$$X = \# \text{ of correct answers} \sim X \sim \text{binomial}(n = 25, p = 0.20).$$

$$\text{mean} = \mu = 25 \times 0.20 = 5, \quad \text{SD} = \sigma = \sqrt{25 \times 0.20 \times (1 - 0.20)} = 2.$$
An Example on Using Normal Approximation

cont’d

What is the probability that the student will score lower than a “C” (15 or fewer correct answers)?

\[ P(X \leq 15) = P(X < 15.5) \]
\[ \approx P\left(Z < \frac{15.5 - 5}{2}\right) \]
\[ = P(Z < 5.25) \approx 1 \quad (\text{why?}) \]
An Example on Using Normal Approximation

cont’d

What is the probability that the student will get a “C” or better (16 or more correct answers)?

\[
P(X \geq 16) = P(X > 15.5) \\
\approx P \left( Z > \frac{15.5 - 5}{2} \right) \\
= P(Z > 5.25) \approx 0 \text{(why?)}
\]
An Example on Using Normal Approximation
cont’d

What is the probability that the student will answer 8 or more questions correctly?

\[ P(X \geq 8) = P(X > 7.5) \]
\[ \approx P\left(Z > \frac{7.5 - 5}{2}\right) \]
\[ = P(Z > 1.25) \]
\[ = 1 - P(Z \leq 1.25) \]
\[ = 1 - 0.8944 \]
\[ = 0.1056. \]
An Example
when normal approximation is inappropriate

An example when normal approximation is inappropriate
binomial(20, 0.1) which is right skewed
A Rule of Thumb
when normal approximation is appropriate

average number of successes $> 5$
\[ np > 5 \]

and

average number of failures $> 5$
\[ nq > 5 \]
A study was conducted concerning the use of gloves among the nurses with 15 years or more experience. The study showed that only \( \frac{1}{6} \) of these nurses wear gloves during vascular access procedures. For a sample of \( n=18 \) nurses with 15 years or more experience, is normal approximation appropriate to approximate a binomial probability?

A. No
B. Yes
C. insufficient information to judge.