The Normal Approximation to the Binomial Distribution
Sec. 7.4

Recall: Binomial Random Variables

You are in a binomial situation if…

1. Some simple procedure is being performed a fixed amount of times (n times)
2. Each time the simple procedure is performed, the outcome fits in one of 2 categories (success or failure)
3. The problem asks you to find the probability of obtaining a certain number of successes (x) when the procedure is performed n times.

Recall: Binomial Random Variables

Things to write down for every binomial problem

BINOMIAL

<u>About performing the simple procedure once</u>

success = one of the 2 categories that the outcomes of the simple procedure can belong to (the one the questions asks you to find the probability for)
failure = the other of the 2 categories that the outcomes of the simple procedure can belong to
p = probability of success when performing the simple procedure once
q = probability of failure when performing the simple procedure once

Recall: Binomial Random Variables

Things to write down for every binomial problem

BINOMIAL

<u>About performing the simple procedure *n* times</u>

n = number of times the procedure is run in the question

X = total number of successes out of the *n* trials

Formulas for a Binomial Random Variable

Probability Formula: $P(X = x) = {}_{n}C_{x}\, p^{x} q^{n-x}$

Expected Value Formula: $\mu = np$

Variance Formula: $\sigma^2 = npq$

Standard Deviation Formula: $\sigma = \sqrt{npq}$

Normal Approximation to a Binomial Random Variable

Suppose X has a binomial distribution. If $npq \geq 10$ then the distribution of X is approximately normal with $\mu = np$ and $\sigma = \sqrt{npq}$

Why do we care?
If the number of terms in a binomial problem is large, using the binomial probability formula makes the problem a long problem. Using the normal approximation gives you the answer much quicker (but it's an approximation)

The Continuity Correction

Going from binomial to normal alone will not give correct answers because binomial is discrete and normal is continuous.

To fix the problem, you must add or subtract 0.5 from the number(s) in the question.

How do you know if you should add or subtract 0.5?

1. Rewrite the problem notation so that it involves (some of) the symbols $\leq$ and/or $\geq$ (and not < or >)
2. Add or subtract 0.5, whichever makes the area bigger

Ex 1 (Hw. #21, sec 7.4, pg. 392, modified): **On-Time Flights**

According to American Airlines, their flights are on time 90% of the time. If 150 American Airlines' flights are randomly selected,

a) What is the probability that exactly 130 flights are on time?

b) What is the probability that at least 130 flights are on time?

c) What is the probability that more than 130 flights are on time?

d) What is the probability that fewer than 125 flights are on time?

e) What is the probability that at most 125 flights are on time?

f) What is the probability that between 125 and 135 flights, inclusive, are on time?

g) What is the probability that between 125 and 135 flights, exclusive, are on time?

h) What is the probability that between 125 and 135 flights, including 125 and not including 135, are on time?