Section 5.4 - Sampling Distributions and the Central Limit Theorem

1. **Sampling Distributions** - the probability distribution (of a statistic) formed by repeatedly taking samples of size n from same population.

2. Sampling Distribution of $\bar{x}$: If $\bar{x}$ is the mean of an SRS of size n drawn from a population with mean μ and standard deviation σ, then the sampling distribution of $\bar{x}$ has a mean of $\mu_{\bar{x}} = \mu$ and a standard deviation of $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$.

   - $\bar{x}$ is an **unbiased estimator** of μ, meaning that the mean of the $\bar{x}$'s is μ.
   - Sometimes $\bar{x}$ will fall below μ, and sometimes it will fall above μ.
   - Averages are less variable than individual observations.
   - The shape of the distribution of $\bar{x}$ depends on the shape of the population.
   - The results of large samples are less variable than the results of small samples.

3. Example: #1
   The number of lightning strikes on a square kilometer of open ground in a year has mean 6 and standard deviation 2.4. Counts of strikes on separate areas are independent. The National Lightning Detection Network uses automatic sensors to watch for lightning in an area of 10 square kilometers. What are the the mean and standard deviation of $\bar{x}$, the mean number of strikes per square kilometer?

4. **Central Limit Theorem:**
   If samples of size n ($n \geq 30$) are drawn from a population with mean μ and standard deviation σ, then the sampling distribution of $\bar{x}$'s is approximately normal with a mean $\mu_{\bar{x}} = \mu$ and standard deviation $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ (regardless of the shape of the population).

   If the population itself is normally distributed, then the sampling distribution of $\bar{x}$'s is also normal, regardless of the sample size.

5. Example #2: The scores of students on the ACT college entrance examination in 2001 had mean $\mu = 21.0$ and standard deviation $\sigma = 4.7$. The distribution of scores is only roughly Normal.

   (a) What is the approximate probability that a single student randomly chosen from all those taking the test scores 23 or higher?

   (b) Now take an SRS of 50 students who took the test. What are the mean and standard deviation of the sample mean score $\bar{x}$ of these 50 students? What is the approximate probability that the mean score $\bar{x}$ of these students is 23 or higher?

   (c) Which of your two Normal probability calculations is more accurate? Why?

6. Example #3:
   Rotor bearings are produced with mean weight $\mu = 1.64$ grams and standard deviation $\sigma = .03$g. For which sample size, $n = 10$ or $n = 50$, is it more likely that the sample mean will be within .02 grams of μ? Why?

7. Example #4:
   Let X be a normal variable with mean 100 and standard deviation 5. Suppose we take a random sample of $n = 25$ items and calculate the sample mean.
   (a) Find $P(90 < X < 110)$ (b) Find $P(90 < \bar{X} < 110)$
   (c) Why are the answers different?

8. Example #5:
   A sample of size $n = 100$ is randomly selected from a non-normal population with mean $\mu = 240$ and standard deviation $\sigma = 40$. Find $P(\bar{X} \geq 230)$.