Chapter 9 Review SOLUTIONS

1. A group of students used pennies to learn about sampling distributions. In the penny activity, each student took samples from a population of 1,000 pennies and were asked to calculate the mean age of their sample. The ***population*** of pennies had a distribution with a shape that was skewed left.

   a. The population of pennies have mean age of 1990.143 and a standard deviation of 11.652. Which symbols would we use to represent these values? _μ___ and __σ___

   b. Each student was first asked to take a sample of size 4, and asked to calculate the average age of their sample. Each student's mean value was placed on a distribution which is called a _sampling__ distribution of _$\bar{x}$ _. The mean value of this distribution should be close to __μ__ and the standard deviation of this distribution should be close to _σ/√4___. The shape of this distribution should be ____skewed left____________.

   c. Each student was then asked to take a sample of size 10, and asked to calculate the average age of their sample. Each student's mean value was again placed on a distribution which is called a _sampling__ distribution of _$\bar{x}$ _. The mean value of this distribution should be close to ___μ___ and the standard deviation of this distribution should be close to __σ/√10 _. The shape of this distribution was probably ____skewed left____________.

   d. Each student was then asked to take a sample of size 30, and asked to calculate the average age of their sample. Each student's mean value was placed on a distribution which is called a _sampling__ distribution of _$\bar{x}$ _. The mean value of this distribution should be close to ___μ__ and the standard deviation of this distribution should be close to __σ/√30 _. The shape of this distribution was probably ____normal____________.

   e. This penny activity was used to show the _ Central Limit_ theorem. When an SRS of sample size *n* is taken from a population with a mean of _μ__ and standard deviation of __σ_, then the mean value ($\mu_{\bar{x}}$) of the sampling distribution is close to __μ__ and the standard deviation is close to _σ/√n__. The _ Central Limit Theorem states that when *n* is small , the sampling distribution will have a shape similar to the population distribution, and As *n* is becomes greater than 30, the shape of the sampling distribution will be approximately _ normal_______.

   f. The _ Central Limit_ Theorem helps us to calculate probabilities because the sampling distribution is approximately normal_ when the sample size is large_ enough.

2. A study of rush-hour traffic in San Francisco counts the number of people in each car entering a freeway at a suburban interchange. Suppose that this count has mean 1.5 and standard deviation 0.75 in the population of all cars that enter at this interchange during rush hours.

   a. Traffic engineers estimate that the capacity of the interchange is 700 cars per hour. According to the central limit theorem, what is the approximate distribution of the mean number of persons $\mu_{\bar{x}}$ in 700 randomly selected cars at this interchange?

      The CLT affects the shape of the sampling distribution of means. Because 700 ≥ 30, this distribution will be approximately normal.

   b. What is the probability that 700 cars will carry more than 1075 people?

The event that 700 cars carry more than 1075 people ($X_1 + X_2 + ... X_{700} > 1075$) is the same as the event that the average number of persons in 700 cars is greater than 1075/700 = 1.535714.

P(Xbar > 1.535714) = normcdf(1.535714, E99, 1.5, .75/√700) = .1039

or P(Xbar > 1075) = normcdf(1075 , E99, 1.5* 700 , .75*700/√700) = .1039

3. A bottling company uses a filling machine to fill plastic bottles with cola. The bottles are supposed to contain 300 milliliters (ml). In fact, the contents vary according to a normal distribution with mean $\mu = 298$ ml and standard deviation $\sigma = 3$ ml. *Note this is already a normal dist.*

   a. What is the probability that an individual bottle contains less than 295 ml? *not a sampling dist.*
   P(X< 295) = normalcdf(-E99, 295, 298, 3) = .15866

   b. What is the probability that the mean contents of the bottles in a six-pack is less than 295 ml?
   this is a sampling dist of means with n = 6
   P(Xbar< 295) = normalcdf(-E99, 295, 298, 3/√6) = .0071

4. The number of flaws per square yard in a type of carpet material varies with mean 1.6 flaws per square yard and standard deviation 1.2 flaws per square yard. The population distribution cannot be normal because a count takes only whole-number values. An inspector studies 200 square yards of the material, records the number of flaws found in each square yard, and calculates $\bar{x}$, the mean number of flaws per square yard inspected. What is the probability that the mean number of flaws exceeds 2 per square yard?
   Though the number of flaws is discrete, the means are continuous. N ≥2000, n≥ 30 therefore CLT applies.
   P (xbar > 2) = normcdf(2, E99, 1.6, 1.2/√200) = .000001

5. The insurance company sees that in the entire population of homeowners, the mean loss from fire is $\mu = 250$ and the standard deviation $\sigma = \$300$. The distribution of losses is strongly skewed right: many policies have \$0 loss, but a few have large losses. If the company sells 10,000 policies, what is the probability that the average loss will be greater than \$260? N ≥ 100,000, n = 10,000 > 30; CLT

   P(Xbar > 260) = normalcdf(260, E99, 250, 3) = 4.3 x 10^{-4}

Not a sampling distribution of sample means problem.

6. Sarah and Betty both run the 400 meter dash in track. Sarah's average time is 71 seconds with a standard deviation of 6 seconds. Betty's times have a mean of 70 seconds with a standard deviation of 8 seconds. The distribution of times for both runners are approximately normal.

   a. What is the probability that Betty runs under 65 seconds in her next race?

   P(B< 65) = normalcdf(-E99, 65, 70, 8) = .265985

   b. What is the probability that Sarah runs under 65 seconds in her next race?

   P(S< 65) = normalcdf(-E99, 65, 71, 6) = .158653

   c. If they were to run an 800 meter relay (each runs 400 meters), what is the probability that it will take them longer than 150 seconds? What assumption do you have to make to do this problem? Independence
   This is a linear transformation so need to use addition formulas for means and standard deviations (Chap 7).
   $\mu_{S+B} = 71 + 70 = 141$; $\sigma_{S+B} = \sqrt{6^2 + 8^2} = 10$
   P(S+B > 150) = normalcdf(150, E99, 141, 10) = .184

   d. If they were to race, what is the probability that Betty wins? What assumption do you have to make to do this problem?
   $\mu_{B-S} = 70 - 71 = -1$; $\sigma_{B-S} = \sqrt{6^2 + 8^2} = 10$
   P(B < S) or P(B – S < 0) = normalcdf(- E99, 0, -1, 10) = .54

7. In an experiment comparing two weight loss plans, 100 subjects were randomly assigned to two groups, A and B. The mean weight loss in group A was 10 pounds with a standard deviation of 8 pounds and the mean weight loss in group B was 7 pounds with a standard deviation of 11 pounds. The distributions of weight loss are approximately normal. *X represents pounds lost. Note this is already a normal dist*

   a. What is the probability that a randomly selected person from group A *lost* weight?
   $P(X_A > 0)$ = normalcdf(0, E99, 10, 8) = .8994

   b. What is the probability that a randomly selected person from group B *gained* weight?
   $P(X_B < 0)$ = normalcdf(-E99, 0, 7, 11) = .2623

   c. If you were to randomly select 3 people from group A, what is the probability that they lost a total of 25 pounds or more? *This is a linear transformation so need to use addition formulas for means and standard deviations (Chap 7).*
   $P(X+X+X \geq 25)$ = normalcdf(25, E99, 10+10+10, $\sqrt{(64+64+64)}$) = .6409
   OR: n = 3, and 25/3 is the mean of the sample, so use sampling distribution of sample means: $P(\bar{X} \geq 25/3)$ = normalcdf(25, E99, 10, $8/\sqrt{3}$) = .6409

   d. If you were to randomly select one person from each group, what is the probability that the person from group B lost more weight? $\sigma_{B-A} = \sqrt{(8^2 + 11^2)} = \sqrt{185}$
   $P(X_B - X_A > 0)$ = normalcdf(0, E99, 7 – 10, $\sqrt{185}$) = .4127

   e. What is the probability that a randomly selected person from group A lost at least 5 more pounds than a randomly selected person from group B?
   $P(X_A - X_B \geq 5)$ or $P(X_B - X_A \leq -5)$ = normalcdf(–E99, –5, 7 – 10, $\sqrt{185}$) = .4415

   f. In parts c-e, what assumption did you have to make? Is it reasonable?
   independence

A linear transformation problem, but can use sampling distribution of sample means.

8. In an elevator at a local hotel, there is a maximum weight limit of 1500 pounds. Suppose that the weights of females who stay at this hotel are approximately normal with a mean of 142 pounds and a standard deviation of 30 pounds. The weights of males who stay at this hotel are approximately normal with a mean of 177 pounds and a standard deviation of 40 pounds.

   a. What is the probability that a group of 10 women exceed the weight limit?
   P(10W > 1500) = normcdf(1500, E99, 1420, 94.86833) = .19954
   $\sigma_{10M} = \sqrt{(10 * 30^2)} = 94.86833$
   OR: n = 10, use sampling distribution of sample means:
   $P(\bar{X}_M > 150)$ = normcdf (150, E99, 142, $30/\sqrt{10}$) = .19954

   b. What is the probability that a group of 10 men exceeds the weight limit?
   P(10M > 1500) = normcdf (1500, E99, 1770, 126.4911) = .9836
   $\sigma_{10M} = \sqrt{(10 * 40^2)} = 126.4911$
   OR: n = 10, use sampling distribution of sample means:
   $P(\bar{X}_M > 150)$ = normcdf (150, E99, 177, $40/\sqrt{10}$) = .9836

   c. What is the probability that a group of 4 men and 5 women exceed the weight limit?
   P(4m+5w > 1500) = normcdf (1500, E99, 1418, 104.4) = .216
   $\sigma_{10M} = \sqrt{(4(40)^2 + 5(30)^2)} = 104.4$
   OR: also true $\sigma_{10M} = \sqrt{[(40*4/\sqrt{4})^2+(30*5/\sqrt{5})^2]} = 104.4$

   d. If you were to randomly select 2 males, what is the probability that their weights are more than 10 pounds apart?
   $\sigma_{M-M} = \sqrt{(40^2 + 40^2)} = 56.568543$ P(m1 – m2 > 10) = normcdf (10, E99, 0, 56.568543) = .4298

   final answer: P(m1 – m2 > 10) or P(m2 – m1 > 10) = .8596