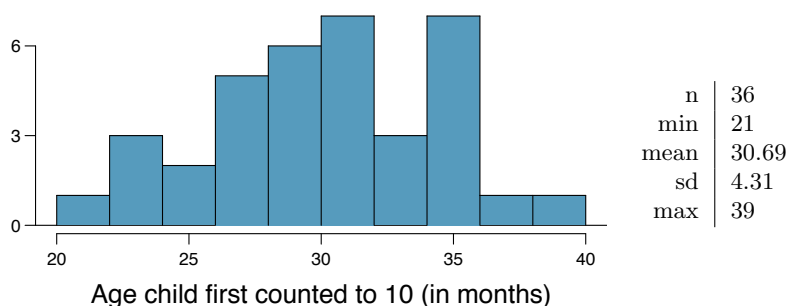


**4.24 Gifted children, Part I.** Researchers investigating characteristics of gifted children collected data from schools in a large city on a random sample of thirty-six children who were identified as gifted children soon after they reached the age of four. The following histogram shows the distribution of the ages (in months) at which these children first counted to 10 successfully. Also provided are some sample statistics.<sup>43</sup>



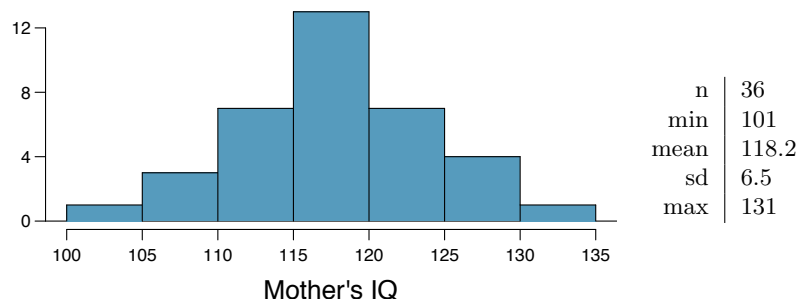
- Are conditions for inference satisfied?
- Suppose you read online that children first count to 10 successfully when they are 32 months old, on average. Perform a hypothesis test to evaluate if these data provide convincing evidence that the average age at which gifted children first count to 10 successfully is less than the general average of 32 months. Use a significance level of 0.10.
- Interpret the p-value in context of the hypothesis test and the data.
- Calculate a 90% confidence interval for the average age at which gifted children first count to 10 successfully.
- Do your results from the hypothesis test and the confidence interval agree? Explain.

**4.25 Waiting at an ER, Part III.** The hospital administrator mentioned in Exercise 4.13 randomly selected 64 patients and measured the time (in minutes) between when they checked in to the ER and the time they were first seen by a doctor. The average time is 137.5 minutes and the standard deviation is 39 minutes. She is getting grief from her supervisor on the basis that the wait times in the ER has increased greatly from last year's average of 127 minutes. However, she claims that the increase is probably just due to chance.

- Are conditions for inference met? Note any assumptions you must make to proceed.
- Using a significance level of  $\alpha = 0.05$ , is the change in wait times statistically significant? Use a two-sided test since it seems the supervisor had to inspect the data before she suggested an increase occurred.
- Would the conclusion of the hypothesis test change if the significance level was changed to  $\alpha = 0.01$ ?

<sup>43</sup>F.A. Graybill and H.K. Iyer. *Regression Analysis: Concepts and Applications*. Duxbury Press, 1994, pp. 511–516.

**4.26 Gifted children, Part II.** Exercise 4.24 describes a study on gifted children. In this study, along with variables on the children, the researchers also collected data on the mother's and father's IQ of the 36 randomly sampled gifted children. The histogram below shows the distribution of mother's IQ. Also provided are some sample statistics.



- Perform a hypothesis test to evaluate if these data provide convincing evidence that the average IQ of mothers of gifted children is different than the average IQ for the population at large, which is 100. Use a significance level of 0.10.
- Calculate a 90% confidence interval for the average IQ of mothers of gifted children.
- Do your results from the hypothesis test and the confidence interval agree? Explain.

**4.27 Working backwards, one-sided.** You are given the following hypotheses:

$$H_0 : \mu = 30$$

$$H_A : \mu > 30$$

We know that the sample standard deviation is 10 and the sample size is 70. For what sample mean would the p-value be equal to 0.05? Assume that all conditions necessary for inference are satisfied.

**4.28 Working backwards, two-sided.** You are given the following hypotheses:

$$H_0 : \mu = 30$$

$$H_A : \mu \neq 30$$

We know that the sample standard deviation is 10 and the sample size is 70. For what sample mean would the p-value be equal to 0.05? Assume that all conditions necessary for inference are satisfied.

**4.29 Testing for Fibromyalgia.** A patient named Diana was diagnosed with Fibromyalgia, a long-term syndrome of body pain, and was prescribed anti-depressants. Being the skeptic that she is, Diana didn't initially believe that anti-depressants would help her symptoms. However after a couple months of being on the medication she decides that the anti-depressants are working, because she feels like her symptoms are in fact getting better.

- Write the hypotheses in words for Diana's skeptical position when she started taking the anti-depressants.
- What is a Type 1 Error in this context?
- What is a Type 2 Error in this context?

**4.30 Testing for food safety.** A food safety inspector is called upon to investigate a restaurant with a few customer reports of poor sanitation practices. The food safety inspector uses a hypothesis testing framework to evaluate whether regulations are not being met. If he decides the restaurant is in gross violation, its license to serve food will be revoked.

- (a) Write the hypotheses in words.
- (b) What is a Type 1 Error in this context?
- (c) What is a Type 2 Error in this context?
- (d) Which error is more problematic for the restaurant owner? Why?
- (e) Which error is more problematic for the diners? Why?
- (f) As a diner, would you prefer that the food safety inspector requires strong evidence or very strong evidence of health concerns before revoking a restaurant's license? Explain your reasoning.

**4.31 Which is higher?** In each part below, there is a value of interest and two scenarios (I and II). For each part, report if the value of interest is larger under scenario I, scenario II, or whether the value is equal under the scenarios.

- (a) The standard error of  $\bar{x}$  when  $s = 120$  and (I)  $n = 25$  or (II)  $n = 125$ .
- (b) The margin of error of a confidence interval when the confidence level is (I) 90% or (II) 80%.
- (c) The p-value for a Z-statistic of 2.5 when (I)  $n = 500$  or (II)  $n = 1000$ .
- (d) The probability of making a Type 2 Error when the alternative hypothesis is true and the significance level is (I) 0.05 or (II) 0.10.

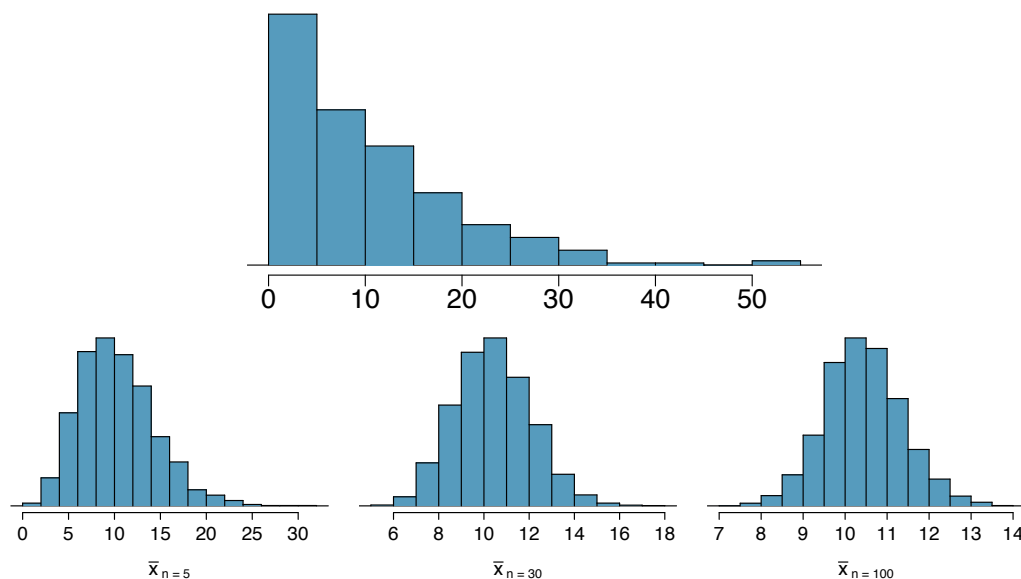
**4.32 True or false.** Determine if the following statements are true or false, and explain your reasoning. If false, state how it could be corrected.

- (a) If a given value (for example, the null hypothesized value of a parameter) is within a 95% confidence interval, it will also be within a 99% confidence interval.
- (b) Decreasing the significance level ( $\alpha$ ) will increase the probability of making a Type 1 Error.
- (c) Suppose the null hypothesis is  $\mu = 5$  and we fail to reject  $H_0$ . Under this scenario, the true population mean is 5.
- (d) If the alternative hypothesis is true, then the probability of making a Type 2 Error and the power of a test add up to 1.
- (e) With large sample sizes, even small differences between the null value and the true value of the parameter, a difference often called the effect size, will be identified as statistically significant.

### 4.6.4 Examining the Central Limit Theorem

**4.33 Ages of pennies.** The histogram below shows the distribution of ages of pennies at a bank.

- Describe the distribution.
- Sampling distributions for means from simple random samples of 5, 30, and 100 pennies is shown in the histograms below. Describe the shapes of these distributions and comment on whether they look like what you would expect to see based on the Central Limit Theorem.
- The mean age of the pennies is 10.44 years, with a standard deviation of 9.2 years. Using the Central Limit Theorem, calculate the means and standard deviations of the distribution of means from random samples of size 5, 30, and 100. Comment on whether the sampling distributions shown in part (b) agree with the values you compute.



**4.34 CLT.** Define the term “sampling distribution” of the mean, and describe how the shape, center, and spread of the sampling distribution of the mean change as sample size increases.

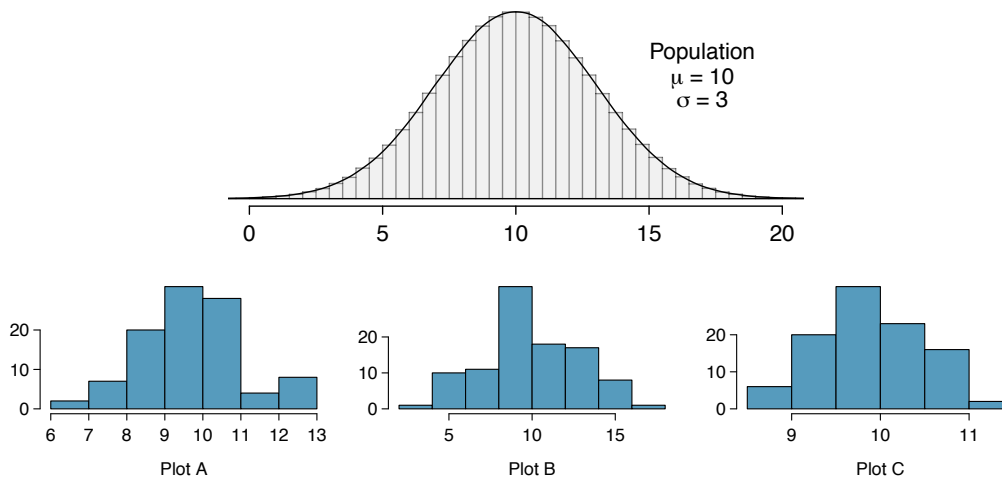
**4.35 Housing prices.** A housing survey was conducted to determine the price of a typical home in Topanga, CA. The mean price of a house was roughly \$1.3 million with a standard deviation of \$300,000. There were no houses listed below \$600,000 but a few houses above \$3 million.

- Is the distribution of housing prices in Topanga symmetric, right skewed, or left skewed? *Hint:* Sketch the distribution.
- Would you expect most houses in Topanga to cost more or less than \$1.3 million?
- Can we estimate the probability that a randomly chosen house in Topanga costs more than \$1.4 million using the normal distribution?
- What is the probability that the mean of 60 randomly chosen houses in Topanga is more than \$1.4 million?
- How would doubling the sample size affect the standard deviation of the mean?

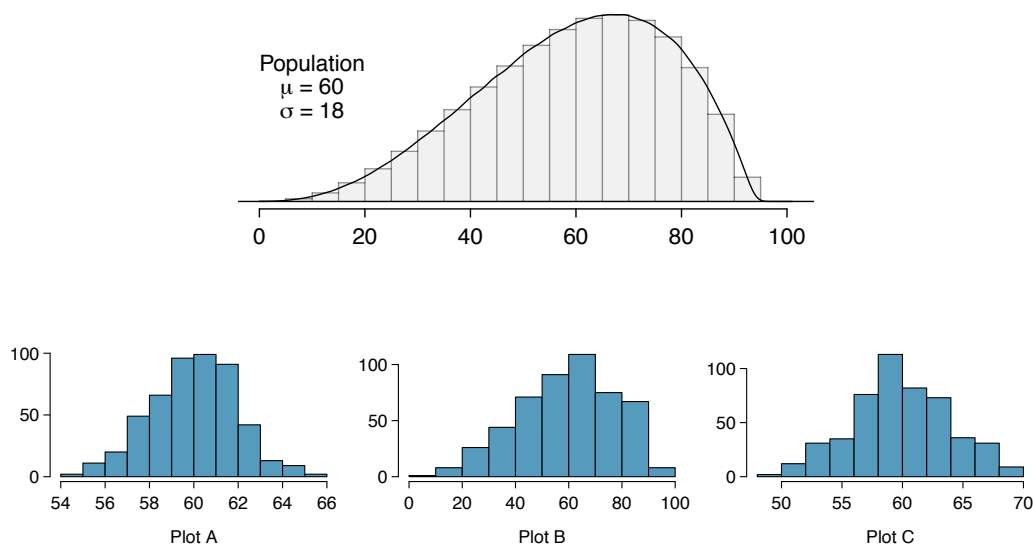
**4.36 Stats final scores.** Each year about 1500 students take the introductory statistics course at a large university. This year scores on the final exam are distributed with a median of 74 points, a mean of 70 points, and a standard deviation of 10 points. There are no students who scored above 100 (the maximum score attainable on the final) but a few students scored below 20 points.

- Is the distribution of scores on this final exam symmetric, right skewed, or left skewed?
- Would you expect most students to have scored above or below 70 points?
- Can we calculate the probability that a randomly chosen student scored above 75 using the normal distribution?
- What is the probability that the average score for a random sample of 40 students is above 75?
- How would cutting the sample size in half affect the standard deviation of the mean?

**4.37 Identify distributions, Part I.** Four plots are presented below. The plot at the top is a distribution for a population. The mean is 10 and the standard deviation is 3. Also shown below is a distribution of (1) a single random sample of 100 values from this population, (2) a distribution of 100 sample means from random samples with size 5, and (3) a distribution of 100 sample means from random samples with size 25. Determine which plot (A, B, or C) is which and explain your reasoning.



**4.38 Identify distributions, Part II.** Four plots are presented below. The plot at the top is a distribution for a population. The mean is 60 and the standard deviation is 18. Also shown below is a distribution of (1) a single random sample of 500 values from this population, (2) a distribution of 500 sample means from random samples of each size 18, and (3) a distribution of 500 sample means from random samples of each size 81. Determine which plot (A, B, or C) is which and explain your reasoning.



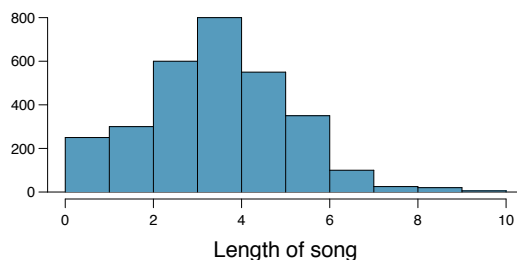
**4.39 Weights of pennies.** The distribution of weights of United States pennies is approximately normal with a mean of 2.5 grams and a standard deviation of 0.03 grams.

- What is the probability that a randomly chosen penny weighs less than 2.4 grams?
- Describe the sampling distribution of the mean weight of 10 randomly chosen pennies.
- What is the probability that the mean weight of 10 pennies is less than 2.4 grams?
- Sketch the two distributions (population and sampling) on the same scale.
- Could you estimate the probabilities from (a) and (c) if the weights of pennies had a skewed distribution?

**4.40 CFLBs.** A manufacturer of compact fluorescent light bulbs advertises that the distribution of the lifespans of these light bulbs is nearly normal with a mean of 9,000 hours and a standard deviation of 1,000 hours.

- What is the probability that a randomly chosen light bulb lasts more than 10,500 hours?
- Describe the distribution of the mean lifespan of 15 light bulbs.
- What is the probability that the mean lifespan of 15 randomly chosen light bulbs is more than 10,500 hours?
- Sketch the two distributions (population and sampling) on the same scale.
- Could you estimate the probabilities from parts (a) and (c) if the lifespans of light bulbs had a skewed distribution?

**4.41 Songs on an iPod.** Suppose an iPod has 3,000 songs. The histogram below shows the distribution of the lengths of these songs. We also know that, for this iPod, the mean length is 3.45 minutes and the standard deviation is 1.63 minutes.



- Calculate the probability that a randomly selected song lasts more than 5 minutes.
- You are about to go for an hour run and you make a random playlist of 15 songs. What is the probability that your playlist lasts for the entire duration of your run? *Hint:* If you want the playlist to last 60 minutes, what should be the minimum average length of a song?
- You are about to take a trip to visit your parents and the drive is 6 hours. You make a random playlist of 100 songs. What is the probability that your playlist lasts the entire drive?

**4.42 Spray paint.** Suppose the area that can be painted using a single can of spray paint is slightly variable and follows a nearly normal distribution with a mean of 25 square feet and a standard deviation of 3 square feet.

- What is the probability that the area covered by a can of spray paint is more than 27 square feet?
- Suppose you want to spray paint an area of 540 square feet using 20 cans of spray paint. On average, how many square feet must each can be able to cover to spray paint all 540 square feet?
- What is the probability that you can cover a 540 square feet area using 20 cans of spray paint?
- If the area covered by a can of spray paint had a slightly skewed distribution, could you still calculate the probabilities in parts (a) and (c) using the normal distribution?

### 4.6.5 Inference for other estimators

**4.43 Spam mail counts.** The 2004 National Technology Readiness Survey sponsored by the Smith School of Business at the University of Maryland surveyed 418 randomly sampled Americans, asking them how many spam emails they receive per day. The survey was repeated on a new random sample of 499 Americans in 2009.<sup>44</sup>

- What are the hypotheses for evaluating if the average spam emails per day has changed from 2004 to 2009.
- In 2004 the mean was 18.5 spam emails per day, and in 2009 this value was 14.9 emails per day. What is the point estimate for the difference between the two population means?
- A report on the survey states that the observed difference between the sample means is not statistically significant. Explain what this means in context of the hypothesis test and data.
- Would you expect a confidence interval for the difference between the two population means to contain 0? Explain your reasoning.

<sup>44</sup>Rockbridge, 2009 National Technology Readiness Survey SPAM Report.