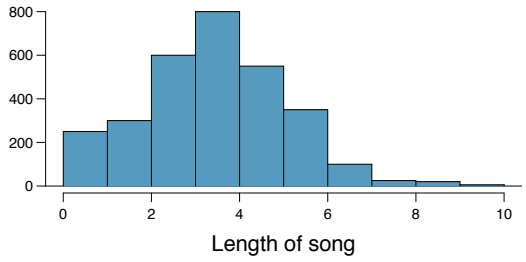
4.41 Songs on an iPod. Suppose an iPod has 3,000 songs. The histogram below shows the distribution of the lengths of these songs. We also know that, for this iPod, the mean length is 3.45 minutes and the standard deviation is 1.63 minutes.



(a) Calculate the probability that a randomly selected song lasts more than 5 minutes.

(b) 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?

(c) 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.

(a) What is the probability that the area covered by a can of spray paint is more than 27 square feet?

(b) 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?

(c) What is the probability that you can cover a 540 square feet area using 20 cans of spray paint?

(d) 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.[44]

(a) What are the hypotheses for evaluating if the average spam emails per day has changed from 2004 to 2009.

(b) 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?

(c) 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.

(d) Would you expect a confidence interval for the difference between the two population means to contain 0? Explain your reasoning.

[44] Rockbridge, 2009 National Technology Readiness Survey SPAM Report.

4.44 Nearsighted. It is believed that nearsightedness affects about 8% of all children. In a random sample of 194 children, 21 are nearsighted.

(a) Construct hypotheses appropriate for the following question: do these data provide evidence that the 8% value is inaccurate?

(b) What proportion of children in this sample are nearsighted?

(c) Given that the standard error of the sample proportion is 0.0195 and the point estimate follows a nearly normal distribution, calculate the test statistic (the Z-statistic).

(d) What is the p-value for this hypothesis test?

(e) What is the conclusion of the hypothesis test?

4.45 Spam mail percentages. The National Technology Readiness Survey sponsored by the Smith School of Business at the University of Maryland surveyed 418 randomly sampled Americans, asking them how often they delete spam emails. In 2004, 23% of the respondents said they delete their spam mail once a month or less, and in 2009 this value was 16%.

(a) What are the hypotheses for evaluating if the proportion of those who delete their email once a month or less has changed from 2004 to 2009?

(b) What is the point estimate for the difference between the two population proportions?

(c) A report on the survey states that the observed decrease from 2004 to 2009 is statistically significant. Explain what this means in context of the hypothesis test and the data.

(d) Would you expect a confidence interval for the difference between the two population proportions to contain 0? Explain your reasoning.

4.46 Unemployment and relationship problems. A USA Today/Gallup poll conducted between 2010 and 2011 asked a group of unemployed and underemployed Americans if they have had major problems in their relationships with their spouse or another close family member as a result of not having a job (if unemployed) or not having a full-time job (if underemployed). 27% of the 1,145 unemployed respondents and 25% of the 675 underemployed respondents said they had major problems in relationships as a result of their employment status.

(a) What are the hypotheses for evaluating if the proportions of unemployed and underemployed people who had relationship problems were different?

(b) The p-value for this hypothesis test is approximately 0.35. Explain what this means in context of the hypothesis test and the data.

4.47 Practical vs. statistical. Determine whether the following statement is true or false, and explain your reasoning: "With large sample sizes, even small differences between the null value and the point estimate can be statistically significant."

4.48 Same observation, different sample size. Suppose you conduct a hypothesis test based on a sample where the sample size is $n = 50$, and arrive at a p-value of 0.08. You then refer back to your notes and discover that you made a careless mistake, the sample size should have been $n = 500$. Will your p-value increase, decrease, or stay the same? Explain.