6.4: Exercises

You buy seeds of one particular species to plant in your garden, and the information on the seed packet tells you that, based on years of experience with that species, the mean number of days to germination is 22, with standard deviation 2.3 days.

What is the population and variable in that information? What parameter(s) and/or statistic(s) are they asserting have particular values? Do you think they can really know the actual parameter(s) and/or statistic's(s') value(s)? Explain.

You plant those seeds on a particular day. What is the probability that the first plant closest to your house will germinate within half a day of the reported mean number of days to germination – that is, it will germinate between 21.5 and 22.5 after planting?

You are interested in the whole garden, where you planted 160 seeds, as well. What is the probability that the average days to germination of all the plants in your garden is between 21.5 and 22.5 days? How do you know you can use the Central Limit Theorem to answer this problem – what must you assume about those 160 seeds from the seed packet in order for the CLT to apply?

You decide to expand your garden and buy a packet of different seeds. But the printing on the seed packet is smudged so you can see that the standard deviation for the germination time of that species of plant is 3.4 days, but you cannot see what the mean germination time is.

So you plant 100 of these new seeds and note how long each of them takes to germinate: the average for those 100 plants is 17 days.

What is a 90% confidence interval for the population mean of the germination times of plants of this species? Show and explain all of your work. What assumption must you make about those 100 seeds from the packet in order for your work to be valid?

What does it mean that the interval you gave had *90% confidence*? Answer by talking about what would happen if you bought many packets of those kinds of seeds and planted 100 seeds in each of a bunch of gardens around your community.

An SRS of size 120 is taken from the student population at the very large Euphoria State University [ESU], and their GPAs are computed. The sample mean GPA is 2.71 . Somehow, we also know that the population standard deviation of GPAs at ESU is .51 . Give a confidence interval at the 90% confidence level for the mean GPA of all students at ESU.

You show the confidence interval you just computed to a fellow student who is not taking statistics. They ask, "Does that mean that 90% of students at ESU have a GPA which is between a and b?" where a and b are the lower and upper ends of the interval you computed. Answer this question, explaining why if the answer is *yes* and both why not and what is a better way of explaining this 90% confidence interval, if the answer is *no*.

The recommended daily calorie intake for teenage girls is 2200 calories per day. A nutritionist at Euphoria State University believes the average daily caloric intake of girls in her state to be lower because of the advertising which uses underweight models targeted at teenagers. Our nutritionist finds that the average of daily calorie intake for a random sample of size $n = 36$ of teenage girls is 2150.

Carefully set up and perform the hypothesis test for this situation and these data. You may need to know that our nutritionist has been doing studies for years and has found that the standard deviation of calorie intake per day in teenage girls is about 200 calories.

Do you have confidence the nutritionist's conclusions? What does she need to be careful of, or to assume, in order to get the best possible results?

The medication most commonly used today for post-operative pain relieve after minor surgery takes an average of 3.5 minutes to ease patients' pain, with a standard deviation of 2.1 minutes. A new drug is being tested which will hopefully bring relief to patients more quickly. For the test, 50 patients were randomly chosen in one hospital after minor surgeries. They were given the new medication and how long until their pain was relieved was timed: the average in this group was 3.1 minutes. Does this data provide statistically significant evidence, at the 5% significance level, that the new drug acts more quickly than the old?

Clearly show and explain all your set-up and work, of course!

The average household size in a certain region several years ago was 3.14 persons, while the standard deviation was .82 persons. A sociologist wishes to test, at the 5% level of significance, whether the mean household size is different now. Perform the test using new information collected by the sociologist: in a random sample of 75 households this past year, the average size was 2.98 persons.

A medical laboratory claims that the mean turn-around time for performance of a battery of tests on blood samples is 1.88 business days. The manager of a large medical practice believes that the actual mean is larger. A random sample of 45 blood samples had a mean of 2.09 days. Somehow, we know that the population standard deviation of turn-around times is 0.13 day. Carefully set up and perform the relevant test at the 10% level of significance. Explain everything, of course.
