Problem 1: Assume that you have a random sample of 40 burgers the weight of which is normally distributed with sample mean 24.9 grams. Assume that you know that the population standard deviation of all burgers is 1.2 grams.

1. Calculate the 95% confidence interval and the 90% confidence interval for the mean burger weight. Be sure to include the relevant formulas.
   * The 95% confidence interval for *μ* may be computed with the formula . Here, =24.9,, *σ* =1.2, and *n* = 40. Plugging these values into the formula above, we get: 24.9 = 24.9.3719 = (24.53, 25.27).
   * Similarly, the formula for the 90% confidence interval for *μ* is . When we plug in the values , σ = 1.2, and *n* = 40, we obtain 24.9 = 24.9.3121 = (24.59, 25.21).
2. Interpret these results.
   * The 95% CI of (24.53, 25.27) means that the true population mean *μ* is between 24.53 and 25.27 with 95% confidence. To take it one step back, a 95% confidence interval (CI) means that if we draw many samples of size *n* = 40 from the population of burgers, then 95% of the samples will yield confidence intervals which contain the true population mean *μ*, and the remaining 5% of the samples will yield confidence intervals which do not. We do not know whether our sample is one of the 95% or one of the 5% - but if our sample is one of the 5%, we’re pretty unlucky. Said differently, the probability that the 95% CI constructed from our (random) sample contains the true population mean is 0.95, and the probability that the 95% CI constructed from our sample doesn’t contain the true population mean is 0.05. Note that this previous statement doesn’t mean that P(24.53 < μ < 25.27) = 0.95. In fact, the statement P(24.53 < μ < 25.27) = 0.95 is absolutely wrong, because every sample would have a different sample mean and therefore would yield different confidence interval bounds.
   * The 90% CI can be interpreted in a similar way.
   * In addition, note that the 95% CI is wider than the 90% CI. That is because a wider interval contains more possible values for *μ*, and as the width increases, so does our confidence that the true population mean *μ* is within that interval.
3. Assume you want to find the sample size which will yield a CI of width 5. What sample size should you use?
   * To find the sample size that would yield a 95% confidence interval with width 5, we would use the formula , where = = 1.96. Specifically, = 0.8851 ≈ 1.
   * Intuitively, the confidence interval width of 5 is so wide that most observations will be within that interval. But in confidence interval terms, the width of a 95% confidence interval will be (even less than) 5 even if our sample contains only 1 burger. So we can say that 95% of the confidence intervals constructed from all samples containing just 1 burger will contain the true population mean.

Problem 2: Assume burger weight is not normally distributed, but you have a sample of 100 burgers which weigh 25.1 grams. You don't know the population standard deviation, but the sample standard deviation is 1.8 grams.

1. Can you use the Central Limit Theorem (CLT) here? Why or why not?
   * Yes, because *n* = 100 > 40. Recall that the CLT states that the distribution of the sample mean will be (approximately) normal regardless of the distribution from which *X1*, *X2*,…, *Xn* come.
2. Calculate the 90% confidence interval for the mean burger weight.
   * To calculate the 90% confidence interval for mean burger weight *μ*, we use the following formula:. Plugging in the values, we get =25.10.2961 = (24.80, 25.40).
3. Interpret the confidence interval.
   * The interpretation from 1.b above can be adapted here.
4. In your own words, explain what happens when you use the population standard deviation as opposed to the sample standard deviation in the confidence interval calculations.
   * In most situations, we do not have the population standard deviation *σ* available (because *σ* is generally computed once *μ* is known and not before, and if we know *μ* there is no need to construct a confidence interval for it). Thus, in most situations, we use the value of the sample standard deviation, *s*, as the estimate of *σ*. However, because *S* is a random variable, its value *s* will vary from sample to sample just like the value of the sample mean, ; therefore, substituting *s* for *σ* in the formula does add more variability. However, this variability is small when *n* is large (say, greater than 40), and the random variable still has an approximately standard normal distribution. (Because of the CLT, this holds regardless of the distribution of *X*.) On the other hand, when *n* is small (say, less than 40), *s* is no longer a good estimate of *σ*, and the random variable has a T distribution with degrees of freedom. (Note that this holds only when *X* is normal.) Note that as *n* increases, the T distribution approaches the standard normal distribution.
5. Assume you have another 100 burger sample where the mean is 32 grams and the standard deviation is 1.0 grams. Calculate the 90% confidence interval for the mean given the data for this sample, and compare it to your response to (b) above. Do the intervals intersect? What does this indicate about the population mean?
   * The confidence interval is computed as = =32.00.2961 = (31.84, 32.16).
   * This interval and the interval in 2.b above do not overlap, and are quite far from it. While it is possible that these two samples come from the same population (i.e., populations with the same mean weight *μ*), it is very unlikely. If we use the value of *s* = 1.8 as an estimate of *σ*, then we can compute the standard error of the mean (i.e., the standard deviation of the sample mean ) to be 1.8/√100 = 0.18. In that case, 31.84 (the lower bound of the CI (31.84, 32.16)) and 25.40 (the upper bound of the CI (24.80, 25.40)) will be = 35.78 standard deviations apart. You will see (when we cover independent samples T-tests) that because of this, we will be able to reject the null hypothesis that the two samples come from the same population.