

Math Statistics

Semiweekly HW 8

Ozaner Hansha

October 2, 2020

Question 1

Problem: Suppose we have a normally distributed population with a mean of 72 and unknown variance. If we take 10 independent samples and obtain a sample mean $\bar{x} = 80$ and a sample standard deviation of $s = 10$, use the t-distribution to give the probability of obtaining such a measurement (i.e. a measured mean this high, given this sample standard deviation).

Solution: Our desired probability is given by:

$$\begin{aligned} P(\bar{X} \geq 80 \mid S = 10) &= t_{\frac{\bar{x}-\mu}{s/\sqrt{n}}, n-1} \\ &= t_{\frac{80-72}{10/\sqrt{10}}, 10-1} \\ &= t_{2.5298, 9} \\ &\approx 0.01612 \end{aligned}$$

I'm not entirely sure what is meant by "a measured mean this high" in the question. I assume this is it.

Question 2

Problem: Suppose we are sampling from a normally distributed population. If we know the population variance, we can obtain a 95% confidence interval for the population mean using a standard normal distribution. If we don't know the population variance, we can obtain a 95% confidence interval for the population mean using the t-distribution. Which confidence interval will be narrower?

Solution: Recall that as the degrees of freedom, which in this case corresponds to the sample size, increases, the t-distribution becomes more and more narrow. This should be intuitive as more samples means less uncertainty in the population mean and thus a narrower distribution.

Also recall that as the degrees of freedom approach infinity, the t-distribution approaches the normal distribution.

Putting these two facts together, we see that the normal distribution is narrower than the t-distribution for *any* sample size since the normal distribution is the limit of the t-distribution getting narrower with greater sample sizes.

This should be intuitive as knowing the population variance eliminates any uncertainty that estimating it from the sample variance introduces, thus maximizing the narrowness of the distribution.

And so, to be explicit, the confidence interval using the population variance (i.e. the normally distributed population mean) will be narrower than that of the unknown population variance (i.e. the t-distributed population mean).