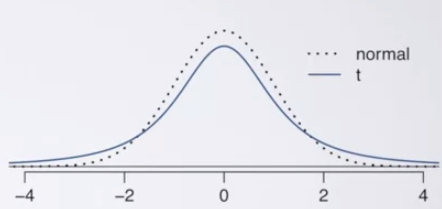
31 Stats Video Lecture – t Distribution

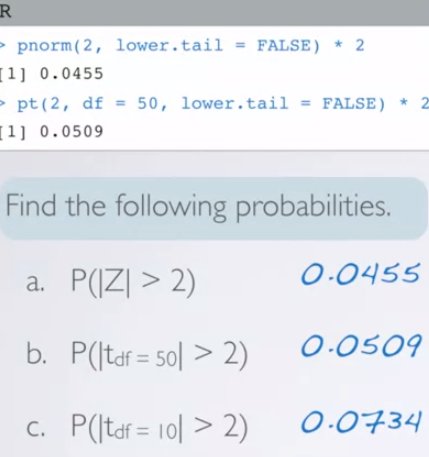
Intro

t dist is useful a distribution when the population mean is unknown

1. Why have a large sample?
   1. as long as observations are independent and the population dist is not very skewed, a large sample ensures:
      1. the sampling distribution of the mean is nearly normal
      2. the estimate of the standard error is reliable
   2. large sampling is not feasible for some experiments such as in a lab setting or experimenting on an endangered species. What, then, will work on small sample AND large samples?
2. t-distribution is used when the standard error is not reliable and sigma is not known
3. Properties of t-dist



* 1. bell shaped
  2. thicker tail than normal dist
  3. observations are more liklely to fall beyond 2   
     SDs from the mean than the normal dist
  4. allows for larger standard errors
  5. extra thick tails are helpful for mitigating the effect  
     of a less reliable estimate for the standard error of  
     the sampling dist
  6. always centered at zero
  7. instead of a mean and standard deviation, the t dist only has one parameter, the “degree of freedom” df – this determines the thickness of the tails.
     1. as the DF increases, the shape of the t-dist approaches the normal curve

1. t-statistic for inference
   1. for inference on a single mean where sigma is unknown
   2. the T statistic is much like the Z score:
   3. find the p-value as the probability of observed or more extreme outcomes given the null hypothesis is true.
      1. use a one or two tail area pvalue based on the alternative hypothesis
      2. use R to calculate
   4.  < P(Z) is basically the pvalue
   5. degrees of freedom are based on sample size. The lower the sample size, the harder it is to reject the null hypothesis