Recall:

- $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2) \Rightarrow \bar{X} \sim N(\mu, \sigma^2/n)$, or

  $$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

- $X_1, \ldots, X_n$ *iid* with common mean μ and common variance σ^2, **then for large** n we apply the central limit theorem to say that $\Rightarrow \bar{X} \approx N(\mu, \sigma^2/n)$, or

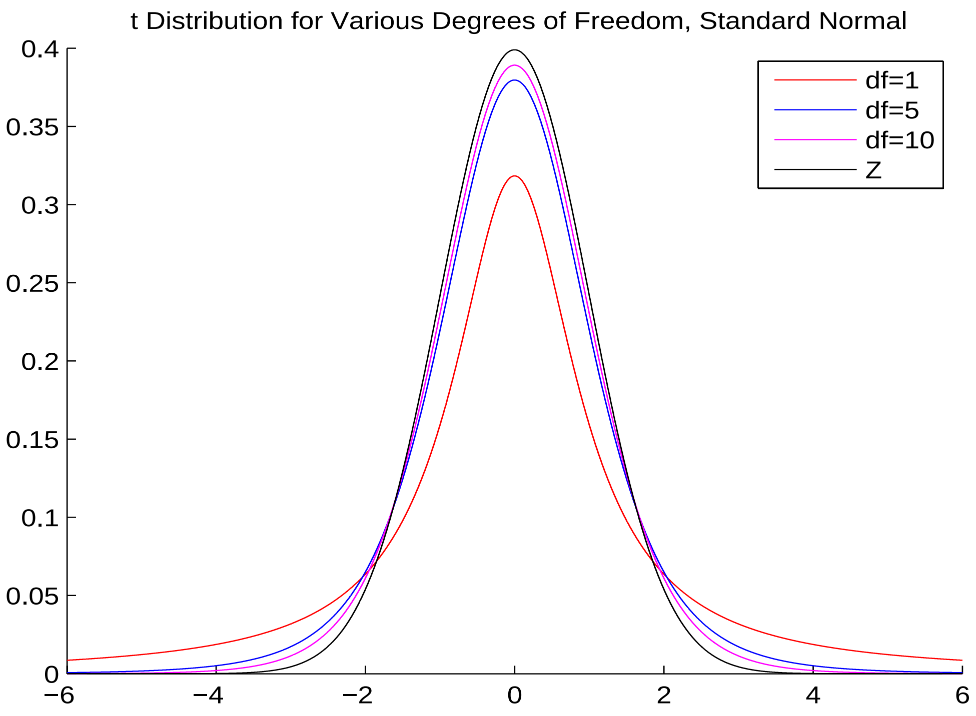
  $$\frac{\bar{X} - \mu}{s/\sqrt{n}} \approx N(0,1)$$

- ("Small Sample" case) $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2) \Rightarrow$

  $$\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t_{df \equiv (n-1)}$$

  where the density of t is given as

  $$h_k(t) = \frac{\Gamma[(k+1)/2]}{\Gamma(k/2)\sqrt{\pi k}}(1 + \frac{t^2}{k})^{-(k+1)/2}$$



Examples

1. Suppose a population is normally distributed with variance $\sigma^2 = 5$. Sampling you obtain the data

   7.83 1.67 10.6 11.4 4.26 15.9 15.9

   Calculate a 90% confidence interval for the population mean μ and explain what such an interval represents.

2. A population is exponentially distributed. You with to estimate it's mean and obtain a sample of size $n = 50$ with sample mean $\bar{x} = 0.230$ and sample standard deviation $s = 0.222$. Form a 99% confidence interval for the population mean μ.

3. Scores on an IQ test were obtained for a simple random sample of Applied Probability students:

   117 131 111 153 118

   Calculate a 90% confidence interval for the true mean of IQ scores from this population if

   (a) We use the traditional estimate of $\sigma = 15$.

   (b) We do not presume to know the real value of σ but use s instead.

Recall:

1. A random variable X is said to have a Gamma distribution with parameters α and β if its probability density function is given as

$$f_X(x) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & if x > 0 \\ 0 & else \end{cases}$$

2. If $Z \sim N(0,1)$ we showed that the distribution of $Y \equiv Z^2$ is distributed as a Gamma distribution with $\alpha = 1/2$ and $\beta = 2$.

3. Using moment generating functions we showed that if $X_1, X_2, \ldots, X_n$ are independent, identically distributed Gamma random variables with $\alpha_1, \alpha_2, \ldots, \alpha_n$ and with common β, then $X \equiv X_1 + X_2 + \ldots + X_n$ is distributed as Gamma with $\alpha = \alpha_1 + \alpha_2 + \ldots + \alpha_n$ and β.

4. Thus, if $Z_1, Z_2, \ldots, Z_n$ are standard normal random variables, if we define $X \equiv Z_1^2 + Z_2^2 + \ldots + Z_n^2$ then $X \sim Gamma(\alpha = n/2, \beta = 2)$. We say such a random variable has a χ^2 distribution with $2\alpha = n$ degrees of freedom.

It is easy to show that if X_1 and X_2 are independent χ^2 random variables, and if X_1 has ν_1 degrees of freedom, and if *the sum* $X_1 + X_2$ has ν degrees of freedom, then X_2 must have $\nu_2 = \nu - \nu_1$ degrees of freedom.

We are now ready for a fundamental result. Suppose $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$. Then

1. $\bar{X}$ and S^2 are independent

2. and

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

Recall that $\bar{X} \equiv \sum_{i=1}^n X_i/n$ and $S^2 \equiv \sum_{i=1}^n (X_i - \bar{X})^2/(n-1)$.

Use the above to calculate an $(1-\alpha) * 100\%$ confidence interval for the population variance σ^2 based upon the sample variance s^2.

Example: Using the IQ data above, construct 95% confidence intervals for σ^2 and σ.

117 131 111 153 118

Estimating a Population Proportion We have seen that, according to the central limit theorem, when $X_1, X_2, \ldots, X_n$ are independent, identically distributed random variables (with finite first and second moments), then their sum is approximately normally distributed for large n. As a practical matter many texts will tell you that the CLT applies when n is larger than 30.

Recall that the binomial random variable counts the number of successes on n independent Bernoulli trials of an experiment which has outcomes $SUCCESS$ or $FAILURE$. That is, when conducting n trials, we have a simple random sample from a dichotomous population. Applying the central limit theorem, the total number of successes will have approximately a normal distribution (careful: discrete versus continuous!). Similarly, the proportion of successes when conducting many trials has approximately a normal distribution.

Define the sample proportion as $\hat{p} = X/n$. Derive a confidence interval for the population proportion p using the sample proportion $\hat{p}$.

1. Easy case

2. More algebra