I. Probability background

A. Common probability distributions (Fig. 1 has 3 of these)

1. (Standard) normal or Gaussian distribution
   a. The density for the normal, or Gaussian, distribution is
      \[ f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \]
   b. CDF \( F(x) = \int_{-\infty}^{x} f(y) \ dy \)
   c. Mean is 0.
   d. Distribution is continuous.

2. Uniform
   a. Density \[ f(x) = \begin{cases} \frac{1}{b-a} & \text{for } x \in [a, b) \\ 0 & \text{otherwise} \end{cases} \]
   b. CDF \[ F(x) = \begin{cases} 0 & \text{for } x \leq a \\ x/(b-a) & \text{for } x \in (a, b) \\ 1 & \text{for } x \geq b \end{cases} \]
   c. Variance is \( (b-a)^2/12 \).
   d. Distribution is continuous.

3. Exponential
   a. Density \( f(x) = \frac{1}{\lambda} e^{-x/\lambda} \) for \( x \geq 0 \)
   b. CDF \( F(x) = 1 - e^{-x/\lambda} \) for \( x \geq 0 \)
   c. Mean is \( \lambda \).
   d. Distribution is continuous.

4. Double Exponential (Laplace)
   a. Density \( f(x) = \frac{1}{2\beta} e^{-|x|/\beta} \)
   b. CDF \( F(x) = 1 - e^{-|x|/\beta} \) for \( x \geq 0 \)
   c. Mean is 0.
   d. Distribution is continuous.

5. Cauchy
   a. Density is \( 1/(\pi(1+z^2)) \). (Note typo in Higgins).
   b. CDF \( \frac{1}{\pi} \arctan(z) \)
   c. So upper and lower quartiles are \( \pm 1 \).
   d. Distribution is continuous.
   e. Variance is infinite.
   f. Curiously enough, the CDF can be expressed in terms of the F distribution CDF.
   g. Mean is \( \pi \).
   h. Median does not have a closed-form expression.
   i. Distribution is symmetric only if \( \pi = 1/2 \).

II. Normal-theory Inference background

A. Want to learn about \( \mu = E[X_j] \).

B. If
   1. \( X_1, \ldots, X_j, \ldots, X_n \) are iid, each has a finite variance
   2. \( \mu \) has a normal distribution

C. Central Limit Theorem (CLT) says \( \bar{X} = \sum_{j=1}^{n} X_j/n \) as \( n \to \infty \) is approximately \( N(\mu, \sigma^2/n) \).

D. To test the null hypothesis \( \mu = \mu_0 \) vs. the alternative \( \mu > \mu_0 \), reject the null hypothesis if \( \bar{X} > \mu_0 + z_{\alpha} \sigma/\sqrt{n} \).

E. CI is \( \{ \mu_0 : \text{null hypothesis with } \mu_0 \text{ is not rejected} \} \).
1. If \( \sigma \) is not known, shift to \( t \) approach.
   a. Substitute estimate \( s = \sqrt{\frac{\sum_{j=1}^{n}(X_j - \bar{X})^2}{n-1}} \).
   b. Compare to \( t \) distribution with \( n-1 \) degrees of freedom.

2. Empirical properties are similar for moderately large and moderately small samples.
   a. Normal size is as expected.
   b. Laplace is pretty close.
   c. Cauchy is conservative, \([\text{Mark C R]}\) \([\text{Mark B DS}]\)
      i. Which is paradoxical, because one might expect heavy tails to make it anti-conservative.
   d. Conservativeness in one direction and anti-conservativeness in other direction indicates skewness we know is present in exponential.
   e. Effect decreases as sample size increases.

III. Methods for inference on a location parameter that do not depend on knowing the rest of the parametric family

A. Suppose we have independent identically distributed random variables \( X_i \) for \( i = 1, \ldots, n \).

B. The only location parameter we’ll really get our heads around is the median.

1. Recall: Median \( \theta \) of random variable \( X_j \) is defined so that \( \Pr [X_j \geq \theta] = 1/2, \Pr [X_j \leq \theta] = 1/2 \).

C. To test whether a putative median value \( \theta_0 \) is the true value:

   1. Define new random variables \( Y_j = \begin{cases} 1 & \text{if } X_j \leq \theta_0 \\ 0 & \text{if } X_j > \theta_0 \end{cases} \)

   a. \( a, b \) so that \( \sum_{j=a}^{b-1} 5^n \leq 1 - \alpha \)
   b. Often largest \( a \) so that \( \sum_{j=a}^{b-1} 5^n < \alpha/2 \);
      \( a \approx n/2 - \sqrt{n/2 \alpha/2} \)
   c. Often smallest \( b \) so that \( \sum_{j=b}^{n} 5^n < \alpha/2 \);
      \( b \approx n/2 + \sqrt{n/2 \alpha/2} \)
   d. Reject null if \( T < a \) or \( T \geq b \).

5. Power of test is determined by \( \Pr [X_j < \theta_0] \) for values of \( \theta_0 \) not equal to the true median.
   a. Since \( \theta > \theta_0 \) if and only if \( \Pr [X_j < \theta_0] > 1/2 \),
      alternatives \( \theta > \theta_0 \) correspond to one sided alternatives \( \Pr [Y_j = 1] > 1/2 \).

6. Test is called sign test.
   a. Book calls it binomial test. \([\text{Mark D R]}\) \([\text{Mark D AS}]\)

D. Confidence interval

1. Remember: Confidence interval of level \( 1 - \alpha \) for parameter \( \theta \) is a set \((L, U)\) such that
   a. \( L \) and \( U \) depend on data.
   b. Such that for any \( \theta \), \( \Pr [L \leq \theta \leq U] \geq 1 - \alpha \).

2. Most general method: Invert a family of tests depending on null.
   a. For every possible null value \( \theta_0 \), find a test (ie., a rule \( \delta_0 \)) that takes a data set and returns "Reject null hypothesis that \( \theta = \theta_0 \), or "Do not reject null hypothesis") of size no larger than \( \alpha \).
   b. Then confidence interval is \( \{\theta | \delta_0(\text{data}) = "\text{Do not reject}" \} \).
   c. Sometimes isn’t really an interval, but in most cases it is.

2. Then under \( H_0 : \theta = \theta_0 \), \( Y_j \sim \text{Bin}(1/2, 1) \).
3. Only works if \( \Pr [X_j = \theta_0] > 0 \) ; assume this.
   a. It is usually easier to assess this assumption than it is for distributional assumptions.
4. Problem reduces to one of binomial testing.
   a. \( T = \text{number of observations less than } \theta_0 \).
   b. Pick \( \alpha \) so that \( T < a \) or \( T \geq b \).

6. Then find upper and lower end points of interval.

2. Rejected since \( \sum_{j=a}^{b-1} 5^n \geq 1 - \alpha \) with \( \alpha = 0.2 \).
3. Only works if \( \Pr [X_j = \theta_0] > 0 \) ; assume this.
   a. It is usually easier to assess this assumption than it is for distributional assumptions.
4. Problem reduces to one of binomial testing.
   a. \( T = \text{number of observations less than } \theta_0 \).
   b. Pick \( \alpha \) so that \( T < a \) or \( T \geq b \).