Many studies can be classified as **binomial experiments**.

**Characteristics of a binomial experiment**

1. The experiment consists of a number (denoted \( n \)) of identical trials.
2. There are only two possible outcomes for each trial – denoted “Success” (\( O_1 \)) or “Failure” (\( O_2 \)).
3. The probability of success (denoted \( p \)) is the same for each trial. (Probability of failure = \( q = 1 - p \).)
4. The trials are independent.

**Example 1**: We want to estimate the probability that a pain reliever will eliminate a headache within one hour.

**Example 2**: We want to estimate the proportion of schools in a state that meet a national standard for excellence.

**Example 3**: We want to estimate the probability that a drug will reduce the chance of a side effect from cancer treatment.

- Consider a specific value of \( p \), say \( p^* \) where \( 0 < p^* < 1 \).

- For a test about \( p \), our null hypothesis will be:
• The alternative hypothesis could be one of:

  Two-tailed     Lower-tailed     Upper-tailed

• The test statistic is \( T = \)

• The null distribution of \( T \) is simply the ____________ distribution with parameters

• Table A3 tabulates this distribution for selected parameter values (for \( n \leq 20 \)).

• For examples with \( n > 20 \), a normal approximation may be used, or better yet, a computer can perform the exact binomial test even with large sample sizes.

  **Decision Rules**

• Two-tailed test: We reject \( H_0 \) if \( T \) is very ________ or very ________.

  Reject \( H_0 \) if \( T \leq t_1 \) or \( T > t_2 \).
• How to pick the numbers $t_1$ and $t_2$?

**Picture of null distribution:**

• From Table A3, using $n$ and $p^*$, find $t_1$ and $t_2$ such that

$$\alpha_1 + \alpha_2 \leq \alpha.$$  

• Note we need $P(\text{Type I error}) \leq \alpha$.

• The P-value of the test, for an observed test statistic $T_{obs}$, is defined as:

$$Y \sim \text{Binomial}(n, p^*).$$

• Lower-tailed test: We reject $H_0$ if $T$ is very ________.

Reject $H_0$ if $T \leq t$. 
• We pick the critical value $t$ such that

• From Table A3, using $n$ and $p^*$, find $t$ such that

• The P-value of the test, for an observed test statistic $T_{\text{obs}}$, is:

$$\text{where } Y \sim \text{Binomial}(n, p^*).$$

• Upper-tailed test: We reject $H_0$ if $T$ is very ________.

Reject $H_0$ if $T > t$.

• We pick the critical value $t$ such that

• From Table A3, using $n$ and $p^*$, find $t$ such that

• The P-value of the test, for an observed test statistic $T_{\text{obs}}$, is:

$$\text{where } Y \sim \text{Binomial}(n, p^*).$$
Example 1: The standard pain reliever eliminates headaches within one hour for 60% of consumers. A new pill is being tested, and on a random sample of 17 people, the headache is eliminated within an hour for 14 of them. At $\alpha = .05$, is the new pill significantly better than the standard?

Hypotheses:

Decision rule: Reject $H_0$ if

Test statistic $T =$

P-value =

Conclusion:

On computer: Use `binom.test` function in R (see example code on course web page)
Example 2: In the past, 35% of all high school seniors have passed the state science exit exam. In a random sample of 19 students from one school, 8 passed the exam. At $\alpha = .05$, is the probability for this school significantly different from the overall probability?

Hypotheses:

Decision rule: Reject $H_0$ if

Test statistic $T = \frac{X - np}{\sqrt{npq}}$

$P$-value =

Conclusion:

On computer: Use `binom.test` function in R (see example code on course web page)
Interval Estimation of $p$

- The binomial distribution can be used to construct exact (even for small samples) confidence intervals for a population proportion or binomial probability.

- The Clopper-Pearson CI method inverts the test of $H_0$: $p = p^*$ vs. $H_1: p \neq p^*$.

- This CI consists of all values of $p^*$ such that the above null hypothesis would not be rejected, for our given observed data set.

**Example 2:**
- You can verify that a $p^*$ of 0.40 would not be rejected based on our exit-exam data.

- So 0.40 would be inside the CI for $p$.

- But a value for $p^*$ like 0.90 would have been rejected, so the CI for $p$ would not include 0.90.

- In general, finding all the values that make up the CI requires a table or computer.

- Table A4 gives two-sided confidence intervals (either 90%, 95%, or 99% CIs) for $p$ when $n \leq 30$.

- For larger samples, for one-sided CIs, or for other confidence levels, the `binom.test` function in R gives the Clopper-Pearson CI.
Example 2 again: Find a 95% CI for the probability that a random student for this school passes the exam.

Table A4:

- Using R, find a 98% CI for $p$.

Example 1 again: Find a 90% CI for the proportion of headaches relieved by the new pill.

Table A4:

- Using R, find a 90% one-sided lower confidence bound for $p$.

- Note: The Clopper-Pearson method guarantees coverage probability of at least the nominal level. It may result in an excessively wide interval. The Wilson score CI approach (use prop.test in R) typically gives shorter intervals, but could have coverage probability less than the nominal level.
Section 3.2: The Quantile Test

• Assume that the measurement scale of our data are at least ordinal. Then it is of interest to consider the quantiles of the distribution.

Case I: Suppose the data are continuous. Then the $p^{*}$th quantile is a number $x^{*}$ such that

• Consider testing the null hypothesis that the $p^{*}$th quantile is some specific number $x^{*}$, i.e.,

• If we denote $P(X \leq x)$ by $p$, then we see this is the same null as in the binomial test, and we can conduct the test in the same way.

• Assume the data are a random sample (i.i.d. random variables) measured on at least an ordinal scale.

• Which test statistic we use will depend on the alternative hypothesis. Consider

$$T_1 =$$

$$T_2 =$$
• Note $T_1 \geq T_2$, and if none of the data values equal the number $x^*$, then:

• The null distribution of the test statistics $T_1$ and $T_2$ is again ______________.

Three Possible Sets of Hypotheses

Two-tailed test:

H$_0$: $H_1$: 

Decision rule:

where $\alpha_1 + \alpha_2 \leq \alpha$.

• The P-value of the test is:

where $Y \sim \text{Binomial}(n, p^*)$.

“Quantile greater than” alternative:

H$_0$: $H_1$: 
Decision rule:

• The P-value of the test is:

where \( Y \sim \text{Binomial}(n, p^*) \).

“Quantile less than” alternative:

\( H_0: \) \hspace{2cm} \( H_1: \)

Decision rule:

• The P-value of the test is:

where \( Y \sim \text{Binomial}(n, p^*) \).
Example: Suppose the upper quartile (0.75 quantile) of a college entrance exam is known to be 193. A random sample of 15 students’ scores from a particular high school are given on page 139. Does the population upper quartile for this high school’s students differ from the national upper quartile of 193? Use $\alpha = 0.05$.

$H_0$: 

$H_1$: 

Decision Rule (using Table A3):

Observed test statistics:

P-value:

Conclusion:
Example: Suppose the median (0.5 quantile) selling price (in $1000s) of houses in the U.S. from 1996-2005 was 179. Suppose a random sample of 18 house sale prices from 2011 is 120 500 64 104 172 275 336 55 535 251 214 1250 402 27 109 17 334 205. Has the population median sale price decreased from 179? Use α = 0.05.

H₀: 

H₁:

Decision Rule (using Table A3):

Observed test statistic:

P-value:

Conclusion:

- See R code on course web page for examples using the quantile.test function.
Confidence Interval for a Quantile

• Recall $X^{(1)} \leq X^{(2)} \leq \ldots \leq X^{(n)}$ are called the ordered sample, or order statistics.

• The order statistics can be used to construct an exact CI for any population quantile.

• Suppose the desired confidence level is $1 - \alpha$ (e.g., 0.90, 0.95, 0.99, etc.).

• In Table A3, use the column for $p^*$ (quantile desired).

• In Table A3, find a probability near $\alpha/2$ (call this $\alpha_1$).

• The corresponding $y$ in Table A3 is then called $r - 1$.

• Then find a probability near $1 - \alpha/2$ (call this $1 - \alpha_2$).

• The corresponding $y$ in Table A3 is then called $s - 1$.

• Then the pair of order statistics $[X^{(r)}, X^{(s)}]$ yields a CI for the $p^*$th population quantile.

• This CI will have confidence level at least $1 - \alpha_1 - \alpha_2$ (exactly $1 - \alpha_1 - \alpha_2$ if the data are continuous).
**Example** (House prices): Find an exact CI with confidence level at least 95% for the population median house price in 2011.

**Example** (House prices): Find an exact CI with confidence level at least 95% for the population 0.80 quantile of house prices in 2011.

- See R code on course web page for examples using the quantile.interval function.
Comparison of the Quantile Test to Parametric Tests

• The quantile test is valid for data that are __________ or ______________, whereas the one-sample t-test about the mean requires that data be ______________.

• So the quantile test is more applicable.

• Suppose our distribution is continuous and symmetric. Then: population median = population mean.

• So the quantile test about the median is testing the same thing as the t-test about the mean.

• Which is more efficient? Depends on true population distribution:

<table>
<thead>
<tr>
<th>Population</th>
<th>A.R.E. of quantile test to t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td></td>
</tr>
<tr>
<td>Uniform (light tails)</td>
<td></td>
</tr>
<tr>
<td>Double exponential</td>
<td></td>
</tr>
<tr>
<td>(heavy tails)</td>
<td></td>
</tr>
</tbody>
</table>

• See R code on course web page for power functions of quantile test and t-test for various population distributions.