# STAT 333- Section 3.1
# The Binomial Test

• 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*,**

- The null hypothesis will be:

$$H_0: p = p^*$$

- The alternative hypothesis could be one of:

| Two-tailed | Lower-tailed | Upper-tailed |
|---|---|---|
| $H_1: p \neq p^*$ | $H_1: p < p^*$ | $H_1: p > p^*$ |

• The test statistic is $T$ = Number of "success" out of the n trial.

**Note :**

• The null distribution of $T$ is simply the Binomial distribution with parameters n and $p^*$

• 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**:

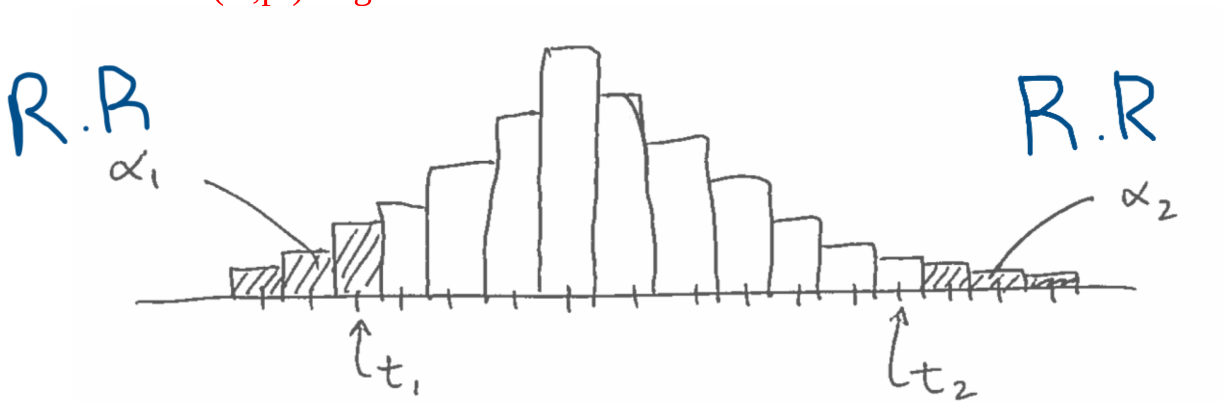
We reject $H_0$ if $T$ is very small or very large.

$$\text{Reject } H_0 \text{ if } T \leq t_1 \text{ or } T > t_2.$$

**• How to pick the numbers $t_1$ and $t_2$?**

**Picture of null distribution:**

Binomial(n ,$p^*$) might looks like:




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

$$P(y \leq t_1) = \alpha_1 \quad \text{and} \quad P(y \leq t_2) = 1- \alpha_2 \quad ( \; P(y > t_2) = \alpha_2)$$

Where , $\alpha_1 + \alpha_2 \leq \alpha$

- Note we need P(Type I error) $\leq \alpha$ ($\alpha$ :stated(nominal) significant level)

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

  $2x[\min\{P(y \leq t_{obs}), P(y \geq t_{obs})\}]$

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

Use table A3 (or computer) to find to find P-value

- **Lower-tailed test:**

  We reject $H_0$ if $T$ is very small.

  Reject $H_0$ if $T \leq t$.

- We pick the critical value $t$ such that

  $P(y \leq t) \approx \alpha$ , where $y \sim Bin(n, p^*)$

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

  $P(y \leq t) \leq \alpha$

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

  $P(y \leq t_{obs})$

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

- **Upper-tailed test:**

  We reject $H_0$ if $T$ is very large .

  Reject $H_0$ if $T > t$.

- We pick the critical value t such that

  $P(y \leq t) \approx 1- \alpha$

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

  $P(y \leq t) \geq 1- \alpha$ , So that $P(y > t) \leq \alpha$

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

  $P(y \geq t_{obs})$

where $Y \sim$ 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 $\alpha = 0.05$, is the new pill significantly better than the standard?

P= Probability of eliminating headache with 1 hour for new pill.

1)Hypotheses:

$H_0: P \leq 0.6$

$H_1: P > 0.6$

( $n=17$ , $p^*=0.6$ )

2)Decision rule: Reject $H_0$ if $T > 13$

$P(y \leq 12) = 0.8740$

$P(y \leq 13) = 0.9536$

3)Test statistic $T = 14 > 13$ , So we reject $H_0$

4) P-value = P(Type I error) = P( Reject $H_0$ \ $H_0$ true)

$= P( T > 14)$

$= 1- P( T \leq 13)$

$= 1- 0.9536 = 0.0464$

Since , P-value $\leq \alpha$

$0.0464 < 0.05$

Then , we reject $H_0$

5)Conclusion:

We conclude that at $\alpha$ =0.05 that the new pill has a significantly higher success probability than 0.6 ( standard pill's probability)

**On computer:** Use binom.test function in R (see example code on course web page)

**R Code:**

```
binom.test(x=14, n=17, p=0.60, alternative="greater")
```

**Output of R:**

Exact binomial test

data: 14 and 17
**number of successes = 14**, number of trials = 17, **p-value = 0.04642**
alternative hypothesis: true probability of success is greater than 0.6
**95 percent confidence interval:**
**0.6043586 1.0000000**
sample estimates:
probability of success
0.8235294

**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?

P =True proportion passing from this school.

1) Hypotheses:

$H_0: P = 0.35$

$H_1: P \neq 0.35$

( $n=19$ , $p^* = 0.35$ )

2) Decision rule: Reject $H_0$ if

$T \leq 2$ or $T > 11$

From table A3,

$P(Y \leq 2) = 0.0170 \longrightarrow \alpha_1 = 0.0170$

$P(Y \leq 11) = 0.9886 \longrightarrow \alpha_2 = 0.0114$

$\alpha_1 + \alpha_2 = 0.0284 \leq \alpha$

3) Test statistic $T = 8$

Note that , $8 \nleq 2$ or $8 \ngtr 11$

So, we do not reject $H_0$

4) P-value =

$= 2 \times [\min\{ P(T \leq 8), P(T \geq 8)\}]$

$= 2 \times [\min\{ 0.8145, (1- 0.6656)\}] = 2 \times [\min\{0.8145, 0.3344\}]$

$= 2 \times 0.3344 = 0.6688$

5) Conclusion:

We accept $H_0$: $P = 0.35$. At $\alpha = 0.05$, we can't conclude that the passing probability for this school differs from 0.35.

**On computer**: Use binom.test function in R (see example code on course web page)

**R Code:**

```
binom.test(x=8, n=19, p=0.35, alternative="two.sided")
```

**output:**

```
                    Exact binomial test
data:  8 and 19
number of successes = 8, number of trials = 19, p-value = 0.6312
alternative hypothesis: true probability of success is not equal to 0.35
95 percent confidence interval:
 0.2025214 0.6650022
sample estimates:
probability of success
         0.4210526
```

## 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.

**R Code:**

```
# The Clopper-Pearson CI for p (note 95% is the default confidence level):

binom.test(x=8, n=19, alternative="two.sided")$conf.int
[1] 0.2025214 0.6650022
attr(,"conf.level")
[1] 0.95
```

**# Using a 98% confidence level:**

```
binom.test(x=8, n=19, alternative="two.sided", conf.level=0.98)$conf.int
[1] 0.1732737 0.7019535
attr(,"conf.level")
[1] 0.98
```

**<u>Example 2 again</u>**:
Find a 95% CI for the probability that a random student for this school passes the exam.

n =19 , Y=8

Table A4:

95% CI for p : (0.203 , 0.665)

• Using R, find a 98% CI for *p*.

98% CI for p : (0.173 , 0.702)

**<u>Example 1 again</u>**: Find a 90% CI for the proportion of headaches relieved by the new pill.

n =17 , Y= 14

90% CI for p : (0.604 , 0.950)

<u>Table A4:</u>
• Using R, find a 90% <u>one-sided lower confidence bound</u> for *p*.

$p \geq 0.648$

[0.648 , 1] one -sided CI

• **<u>Note:</u>**

The Clopper-Pearson method guarantees coverage probability of <u>at least</u> 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 <u>could</u> have coverage probability <u>less than</u> the nominal level.

**R Code:**

```
# The Clopper-Pearson 90% CI for p:

binom.test(x=14, n=17, alternative="two.sided", conf.level=0.90)$conf.int

# A 90% lower confidence bound:

binom.test(x=14, n=17, alternative="greater", conf.level=0.90)$conf.int
```

**Output :**

```
binom.test(x=14, n=17, alternative="greater", conf.level=0.90)$conf.int
[1] 0.6481312 1.0000000
attr(,"conf.level")
[1] 0.9

binom.test(x=14, n=17, alternative="two.sided", conf.level=0.90)$conf.int
[1] 0.6043586 0.9501018
attr(,"conf.level")
```