STA258H5

Al Nosedal, Asal Aslemand and Samir Hamdi

Winter 2024

INTRODUCTION TO HYPOTHESIS TESTING

Review:
CI's are a range of plausible values for a parameter.

Test of Hypothesis for One Mean

Hypothesis Tests

An inferential procedure to determine where there is sufficient evidence to suggest a condition for a population parameter using statistics from a sample

Attach a probability to the conclusion of a hypothesis test

Steps

1/ Decide on a level of significance (α)

2/ State the null hypothesis (H_0) and the alternative hypothesis (H_a) (H_1)

3/ Calculate the appropriate test statistic.

4/ Use the test statistic and a reference distribution to calculate a p-value.
(Also refer back to H_a)

5/ Compare p-value to α to make a conclusion

Note:
The definition of a p-value can be confusing, we will define it later.

1/ Decide on a level of significance (α)

Threshold for decision making

Depends on tolerence for consequences of errors, sample size, nature of the study, and variability.

Common values : 0.10, 0.05, 0.01

very common default

2/ State the null hypothesis (H_0) and the alternative hypothesis (H_a)

θ: parameter of interest

θ_0: Numerical value of the parameter of interest hypotesized under the null hypothesis.

$H_0: \theta = \theta_0 \quad (\theta \leq \theta_0)$ $\qquad H_a: \theta > \theta_0$

$H_0: \theta = \theta_0 \quad (\theta \geq \theta_0)$ $\qquad H_a: \theta < \theta_0$

one-sided (one-tailed)

$H_0: \theta = \theta_0$ $\qquad H_a: \theta \neq \theta_0$

two-sided (two-tailed)

Null (H_0) : Represents the $\underbrace{\text{current belief}}_{\text{status quo}}$ or the safe belief.

Alternative (H_a): " " research hypothesis (or what you are asked to test)

3/ Calculate an appropriate test statistic

Depends on the hypothesis test conducted and the information available

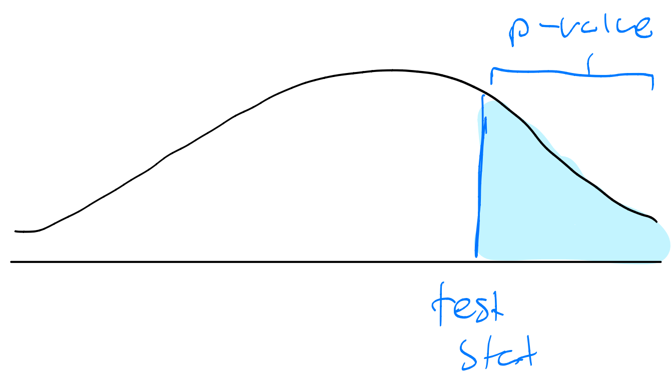
Skeleton of a Test statistic

$$\text{test statistic} = \frac{\underbrace{(\text{a statistic})}_{\text{a point estimate}} - \left(\begin{array}{c}\text{hypothesized value of} \\ \text{parameter under } H_0\end{array}\right)}{(\text{standard error of statistic})}$$

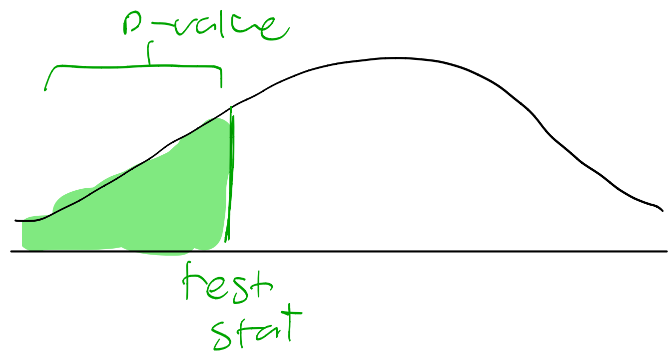
The test statistic follows a reference distribution (Z / t / F / χ^2)

4/ calculate p-value (use test stat, reference distribution and also refer back to Ha)

$H_0: \theta = \theta_0$ $H_a: \theta > \theta_0$



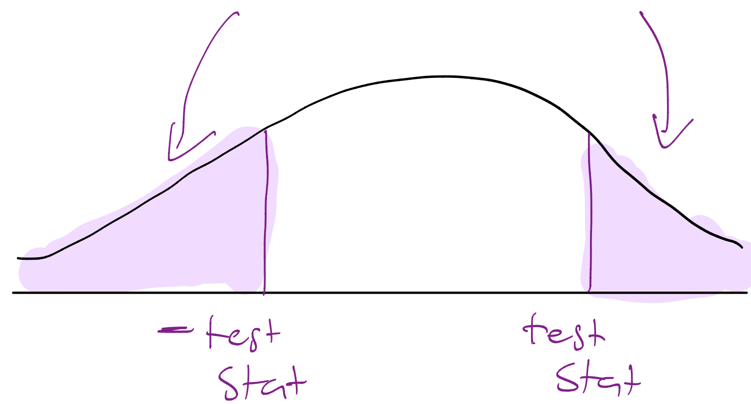
$H_0: \theta = \theta_0$ $H_a: \theta < \theta_0$



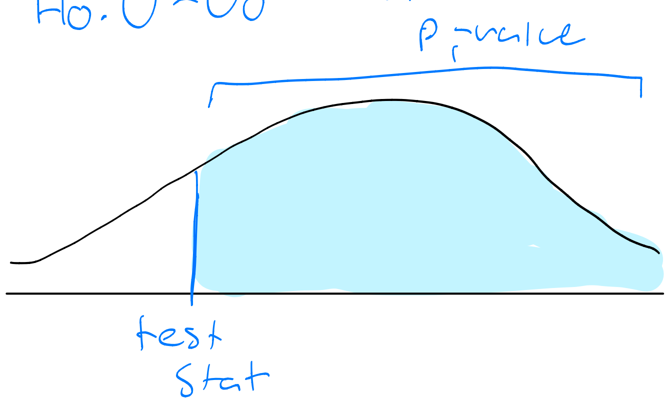
two-sided
two-tailed

$H_0: \theta = \theta_0$ $H_a: \theta \neq \theta_0$

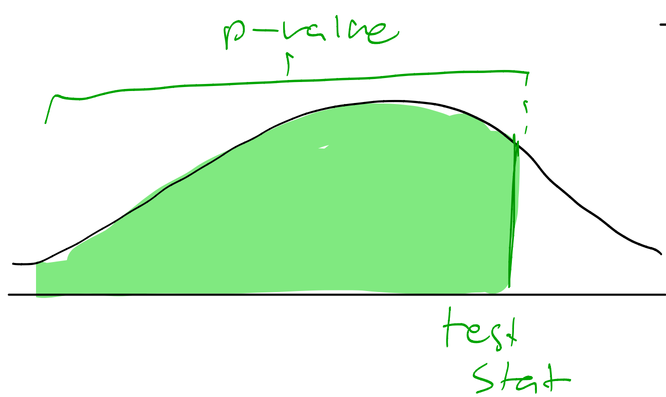
total shaded = p-value



$H_0: \theta = \theta_0$ $H_a: \theta > \theta_0$



$H_0: \theta = \theta_0$ $H_a: \theta < \theta_0$



5/ Compare p-value to level of significance α and make a conclusion.

p-value $< \alpha$ $\longrightarrow$ Sufficient evidence against H_0. The hyp. test rejects H_0 in favour of H_a

p-value $\geq \alpha$ $\longrightarrow$ Insufficient evidence against H_0.
Do not reject H_0
(Fail to reject H_0)

Note:
It is not good practice to give conclusion in the context of stating we accept H_0 or accept H_a

Sweetening colas

Diet colas use artificial sweeteners to avoid sugar. These sweeteners gradually lose their sweetness over time. Manufacturers therefore test new colas for loss of sweetness before marketing them. Trained tasters sip the cola along with drinks of standard sweetness and score the cola on a "sweetness score " of 1 to 10. The cola is then stored for a month at high temperature to imitate the effect of four months' storage at room temperature. Each taster scores the cola again after storage. This is a matched pairs experiment. Our data are the differences (score before storage minus score after storage) in the tasters' scores. The bigger these differences, the bigger the loss of sweetness.

Sweetening colas (cont.)

Suppose we know that for any cola, the sweetness loss scores vary from taster to taster according to a Normal distribution with standard deviation $\sigma = 1$. The mean μ for all tasters measures loss of sweetness, and is different for different colas.
The following are the sweetness losses for a new cola, as measured by 10 trained tasters: 2.0 0.4 0.7 2.0 -0.4 2.2 -1.3 1.2 1.1 2.3.
Are these data good evidence that the cola lost sweetness in storage?

Solution

$\mu=$ mean sweetness loss for the population of **all** tasters.

1. State hypotheses. $H_0 : \mu = 0$ vs $H_a : \mu > 0$
2. Test statistic. $z_* = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{1.02-0}{1/\sqrt{10}} = 3.23$
3. P-value. $P(Z > z_*) = P(Z > 3.23) = 0.0006$
4. Conclusion. We would very rarely observe a sample sweetness loss as large as 1.02 if H_0 were true. The small P-value provides strong evidence against H_0 and in favor of the alternative $H_a : \mu > 0$, i.e., it gives good evidence that the mean sweetness loss is not 0, but positive.

Simulation

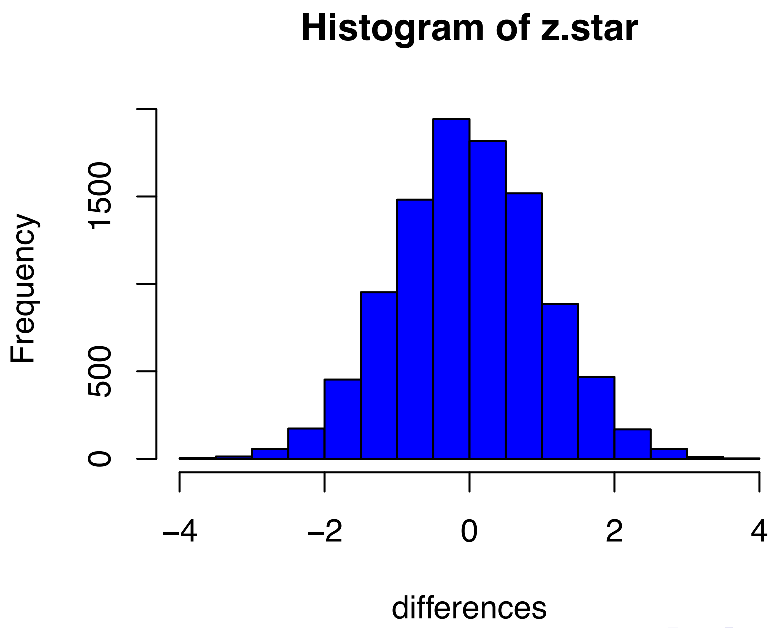
```
# n = sample size;
n<-10;
mu.zero<-0;
sigma<-1;
sigma.xbar<-sigma/sqrt(n);
# x bar = sample mean with 10 obs;
x.bar<-rnorm(1,mean=mu.zero,sd=sigma.xbar);
x.bar;

## [1] 0.3265859

# z.star = test statistic;
z.star<-(x.bar-mu.zero)/sigma.xbar;
z.star;

## [1] 1.032755
```

Another Simulation

```
# n = sample size;
n<-10;
mu.zero<-0;
sigma<-1;
sigma.xbar<-sigma/sqrt(n);
# x bar = sample mean with 10 obs;
x.bar<-rnorm(1,mean=mu.zero,sd=sigma.xbar);
x.bar;

## [1] -0.4571676

# z.star = test statistic;
z.star<-(x.bar-mu.zero)/sigma.xbar;
z.star;

## [1] -1.445691
```

10000 Simulations

```
n<-10;
mu.zero<-0;
sigma<-1;
sigma.xbar<-sigma/sqrt(n);
# x bar = sample mean with 10 obs;
# m = number of simulations;
m<-10000;
x.bar<-rnorm(m,mean=mu.zero,sd=sigma.xbar);

# z.star = test statistic;
z.star<-(x.bar-mu.zero)/sigma.xbar;
hist(z.star,xlab="differences",col="blue");
```


Histogram of z.star
Frequency
0
500
1500
−4
−2
0
2
4
differences

P-value

```
## P-value

p_value<-length(z.star[z.star>3.23])/m;

p_value

## [1] 8e-04
```

One Sample Hypothesis Test on the population mean (μ)

1/ when σ is known

$H_0: \mu = \mu_0$ (or $\mu \leq \mu_0$) $H_a: \mu > \mu_0$

$H_0: \mu = \mu_0$ (or $\mu \geq \mu_0$) $H_a: \mu < \mu_0$

$H_0: \mu = \mu_0$ $H_a: \mu \neq \mu_0$

Test statistic

$$z^* = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

Reference distribution: standard normal (Z)

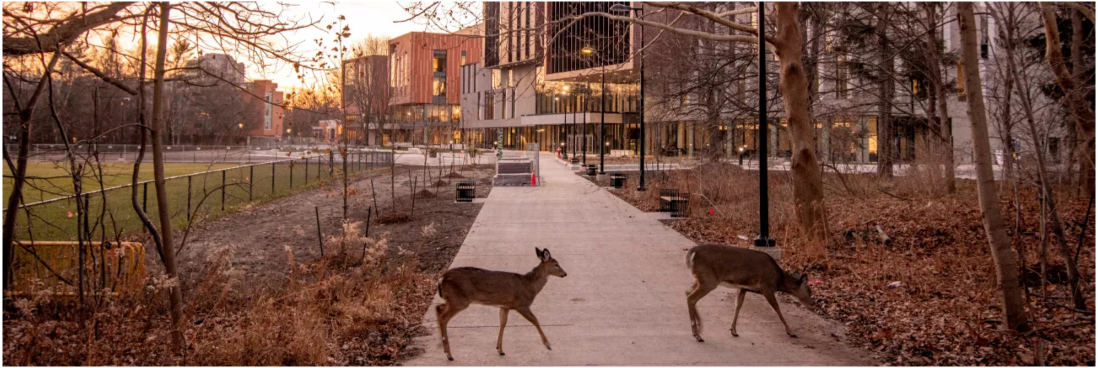
Example: UTM Deer

Deer are a common sight on the UTM campus.

Example: UTM Deer

Credit: University of Toronto at Mississauga

Example: UTM Deer



Credit: University of Toronto at Mississauga

Example: UTM Deer

Credit: University of Toronto at Mississauga

Example: UTM Deer

Credit: University of Toronto at Mississauga

doe : female deer

Example: UTM Deer

$n = 36$

Deer are a common sight on the UTM campus. Suppose an ecologist is interested in the average mass of adult white-tailed does (female deer) around the Mississauga campus to determine whether they are healthy for the upcoming winter. The ecologist captures a sample of 36 adult females around the UTM and measures the average mass of this sample to be 42.53 kg.

$\bar{x} = 42.53$

$H_0: \mu = 45$

From previous studies conducted in the area, the average mass of healthy does was reported to be 45 kg. Conduct a hypothesis test at the 5% significance level to determine whether the mass of does around UTM has decreased. Assume the standard deviation is known to be 5.25 kg.

$\alpha = 0.05$

$H_a: \mu < 45$

$\sigma = 5.25$ (σ known)

1/ Level of significance. $\alpha = 0.05$

2/ State the null and alternative hypotheses

$H_0: \mu = 45$ $\quad\quad$ $H_a: \mu < 45$

[or $H_0: \mu \geq 45$ (μ_0) $\quad\quad$ $H_a: \mu < 45$ (μ_0)]

3/ Calculate appropriate test statistic

$n = 36$ $\quad$ $\bar{x} = 42.53$ $\quad$ $\sigma = 5.25$ (σ known)

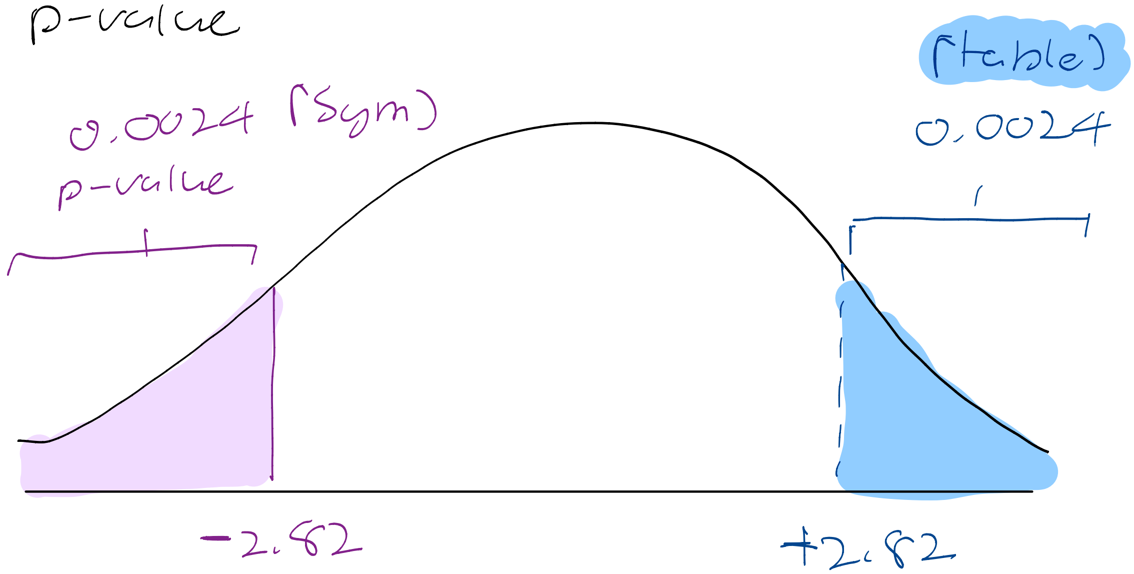
Since σ is known, the test statistic is

$$z^* = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{42.53 - 45}{5.25/\sqrt{36}} = -2.82$$

Reference distribution: Standard normal

4/ Calculate p-value

$H_a: \mu < 45$



p-value = 0.0024

5/ Compare p-value with level of significance α and make a conclusion

give a conclusion in context

$0.0024 < 0.05$

p-value $< \alpha$

There is sufficient evidence at the 5% level of significance to reject the null that does this winter weigh the same as in the past and we conclude the alternative that does this winter weigh less than 45 kg

---

Suppose we conduct the following hypothesis test for the same data

$H_0: \mu = 45$ $H_a: \mu \neq 45$

[or $H_0: \mu \leq 45$ (μ_0) $H_a: \mu > 45$ (μ_0)]

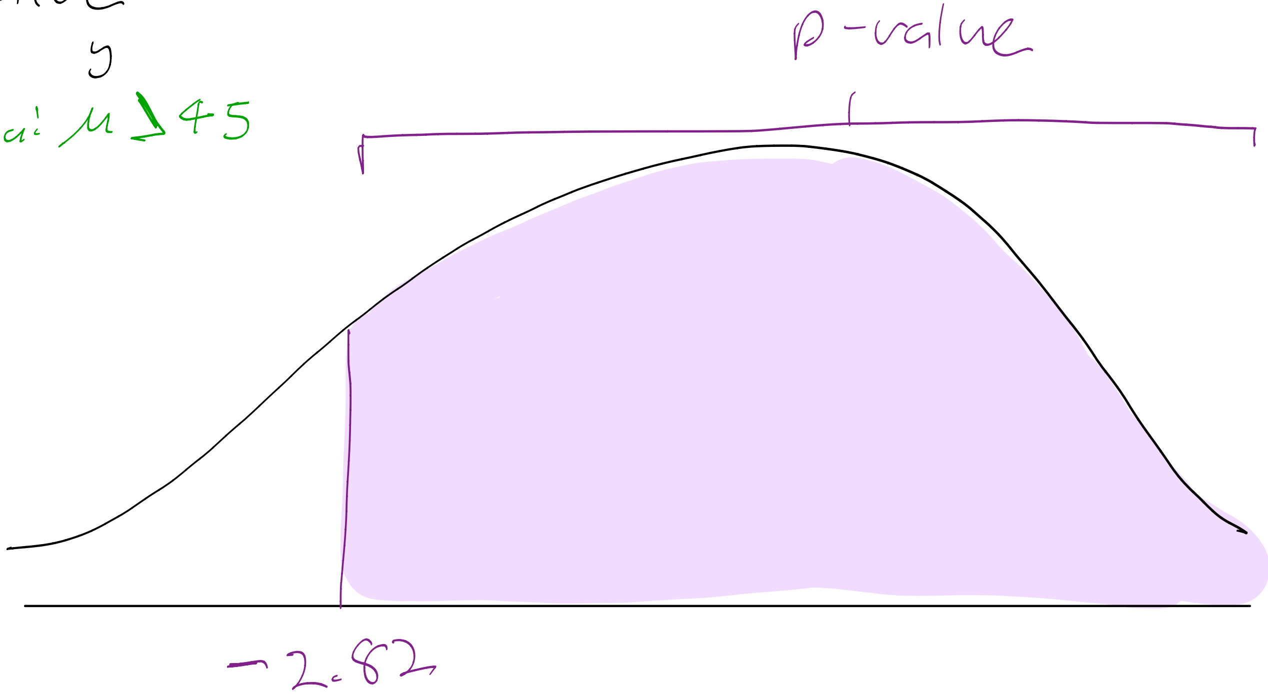
test statistic

$z = -2.82$ (does not change)

reference distribution : Z distrib

p-value

$H_a: \mu > 45$



Suppose we conduct the following hypothesis test for the same data

$H_0: \mu = 45$ $H_a: \mu \neq 45$

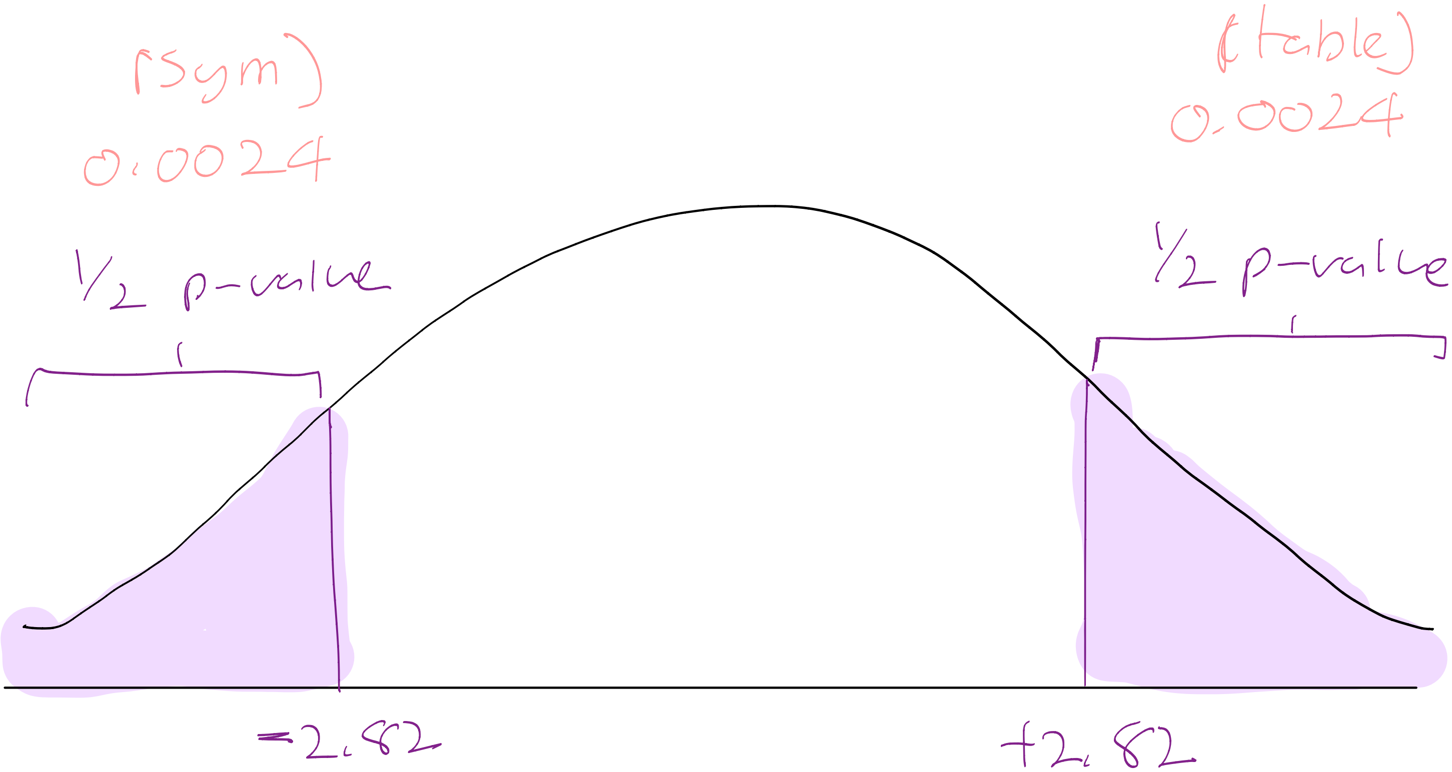
test statistic

$z = -2.82$ (does not change)

reference distribution : z distrib

p-value $H_a: \mu \neq 45$ two-sided (two-tailed)

(total shaded area = p-value)



left $\frac{1}{2}$ p-value = 0.0024

right $\frac{1}{2}$ p-value = 0.0024

p-value = 0.0048

or

$\frac{1}{2}$ p-value = 0.0024

$\times 2$ $\times 2$

p-value = 0.0048

Example: UTM Deer in R

Code Block

```
# Find test stat
> z_test_stat = (42.53 - 45) / (5.25 / sqrt(36))
> z_test_stat
[1] -2.822857

# Find the p-value
# Since the alternative is Ha : mu < 45
> p-value = pnorm(z_test_stat)
[1] 0.00237989
```

pnorm by default gives
area to the left

Example: UTM Deer in R

Code Block

```
# Using the BSDA library. install BSDA if it is not already installed.
# install.packages("BSDA")
> library(BSDA)
```

z test using summary stats

z.test(vector_of_data, ...)

```
> # Conduct the z-test with the zsum.test function
> zsum.test(mean.x = 42.53, sigma.x = 5.24, n.x = 36, mu = 45, alternative = "less")

 One-sample z-Test

data:  Summarized x
z = -2.8282, p-value = 0.00234
alternative hypothesis: true mean is less than 45
95 percent confidence interval:
       NA 43.96651
sample estimates:
mean of x
    42.53
```

? z.test

? zsum.test

Executives' blood pressures

The National Center for Health Statistics reports that the systolic blood pressure for males 35 to 44 years of age has mean 128 and standard deviation 15. The medical director of a large company looks at the medical records of 72 executives in this age group and finds that the mean systolic blood pressure in this sample is $\bar{x} = 126.07$. Is this evidence that the company's executives have a different mean blood pressure from the general population?

Suppose we know that executives' blood pressures follow a Normal distribution with standard deviation $\sigma = 15$.

Solution

$\mu =$ mean of the **executive population**.

1. State hypotheses. $H_0 : \mu = 128$ vs $H_a : \mu \neq 128$
2. Test statistic. $z_* = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{126.07 - 128}{15/\sqrt{72}} = -1.09$
3. P-value. $2P(Z > |z_*|) = 2P(Z > |-1.09|) = 2P(Z > 1.09) = 2(1 - 0.8621) = 0.2758$
4. Conclusion. More than 27% of the time, a simple random sample of size 72 from the general male population would have a mean blood pressure at least as far from 128 as that of the executive sample. The observed $\bar{x} = 126.07$ is therefore not good evidence that executives differ from other men.

Tests for a population mean

There are four steps in carrying out a significance test:

1. State the hypotheses.
2. Calculate the test statistic.
3. Find the P-value.
4. State your conclusion in the context of your specific setting.

Once you have stated your hypotheses and identified the proper test, you or your computer can do Steps 2 and 3 by following a recipe. Here is the recipe for the test we have used in our examples.

Z test for a population mean μ

Draw a simple random sample of size n from a Normal population that has unknown mean μ and known standard deviation σ. To test the null hypothesis that μ has a specified value, $H_0 : \mu = \mu_0$
calculate the **one-sample z statistic**

$$z_* = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

In terms of a variable Z having the standard Normal distribution, the P-value for a test of H_0 against
$H_a : \mu > \mu_0$ is $P(Z > z_*)$
$H_a : \mu < \mu_0$ is $P(Z < z_*)$
$H_a : \mu \neq \mu_0$ is $2P(Z > |z_*|)$

Example 1

Consider the following hypothesis test:
$H_0 : \mu = 20$
$H_a : \mu < 20$
A sample of 50 provided a sample mean of 19.4. The population standard deviation is 2.

a. Compute the value of the test statistic.

b. What is the p-value?

c. Using $\alpha = 0.05$, what is your conclusion?

Solution

a. Test statistic.

$z_* = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{19.4-20}{2/\sqrt{50}} = -2.1213$

b. P-value.

$P(Z < z_*) = P(Z < -2.1213) = 0.0169$

c. Conclusion.

Since P-value $= 0.0169 < \alpha = 0.05$, we reject $H_0 : \mu = 20$. We conclude that $\mu < 20$.

Example 2

Consider the following hypothesis test:
$H_0 : \mu = 25$
$H_a : \mu > 25$
A sample of 40 provided a sample mean of 26.4. The population standard deviation is 6.
a. Compute the value of the test statistic.
b. What is the p-value?
c. Using $\alpha = 0.01$, what is your conclusion?

Solution

a. Test statistic.

$z_* = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{26.4-25}{6/\sqrt{40}} = 1.4757$

b. P-value.

$P(Z > z_*) = P(Z > 1.4757) = 0.0700$

c. Conclusion.

Since P-value $= 0.0700 > \alpha = 0.01$, we CAN'T reject $H_0 : \mu = 25$. We conclude that we don't have enough evidence to claim that $\mu > 25$.

Example 3

Consider the following hypothesis test:
$H_0 : \mu = 15$
$H_a : \mu \neq 15$
A sample of 50 provided a sample mean of 14.15. The population standard deviation is 3.

a. Compute the value of the test statistic.

b. What is the p-value?

c. Using $\alpha = 0.05$, what is your conclusion?

Solution

a. Test statistic.

$z_* = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{14.15-15}{3/\sqrt{50}} = -2.0034$

b. P-value.

$2P(Z > |z_*|) = 2P(Z > |-2.0034|) = 2P(Z > 2.0034) = 0.0451$

c. Conclusion.

Since P-value $= 0.0451 < \alpha = 0.05$, we reject $H_0 : \mu = 15$. We conclude that $\mu \neq 15$.

Tests from confidence intervals

CONFIDENCE INTERVALS AND TWO-SIDED TESTS.

A level α two-sided significance test rejects a hypothesis $H_0 : \mu = \mu_0$ exactly when the value μ_0 falls outside a level $1 - \alpha$ confidence interval for μ.

Example 3 (again...)

The 95% confidence interval for μ in example 3 is:

$$\bar{x} \pm z_*(\frac{\sigma}{\sqrt{n}})$$

$$14.15 \pm 1.96(\frac{3}{\sqrt{50}})$$

$$(13.3184, 14.9815)$$

The hypothesized value $\mu_0 = 15$ in example 3 falls outside this confidence interval, so we reject $H_0 : \mu = 15$.

The one-sample t test

Draw an SRS of size n from a large population having unknown mean μ. To *test the hypothesis* $H_0 : \mu = \mu_0$, compute the *one-sample t statistic*

$$t^* = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

In terms of a variable T having the $t(n-1)$ distribution, the P-value for a test of H_0 against
$H_a : \mu > \mu_0$ is $P(T \geq t^*)$.
$H_a : \mu < \mu_0$ is $P(T \leq t^*)$.
$H_a : \mu \neq \mu_0$ is $2P(T \geq |t^*|)$.
These P-values are exact if the population distribution is Normal and are approximately correct for large n in other cases.

One Sample Hypothesis Tests on the Population Mean μ

1/ When σ is not known

$H_0: \mu = \mu_0$ (or $\mu \leq \mu_0$) $\qquad H_a: \mu > \mu_0$

$H_0: \mu = \mu_0$ (or $\mu \geq \mu_0$) $\qquad H_a: \mu < \mu_0$

$H_0: \mu = \mu_0 \qquad H_a: \mu_0$

Test statistic

$$t^* = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

sample st. dev

Reference distribution: t-distribution at $n-1$ degrees of freedom

Example: Hypothesis Testing about a Population Mean for Small Sample with Unknown σ

Laughter is often called "the best medicine"; studies have shown that laughter can reduce muscle tension and increase oxygenation of the blood. Researchers investigated the physiological changes that accompany laughter. 25 subjects (18-34 years old) watched film clips designed to evoke laughter. During the laughing period, researchers measured the heart rate (beats per minutes) of each subject and obtained: $\bar{x} = 73.5$, and $s = 6$. It is well known that mean restoring heart rate is 71 beats per minute. Is there evidence that the true mean heart rate during laughter exceeds 71 beats per minute? Use $\alpha = 0.05$.

$H_0: \mu = 71$

$H_a: \mu > 71$

Assumptions

- Independent Random Sample of 18-34 years old is taken from the population.
- Heart rate during laughter has a normal distribution.

Example (Slide 102)

$n = 25$ $\bar{x} = 73.5$ $s = 6$ (σ unknown)

1/ $\alpha = 0.05$

2/ $H_0: \mu = 71$ $H_a: \mu > 71$

[or $H_0: \mu \leq \underset{(\mu_0)}{71}$ $H_a: \mu > \underset{(\mu_0)}{71}$]

3/ Calculate test stat

Since σ is unknown

$$t^* = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{73.5 - 71}{6/\sqrt{25}} = 2.083$$

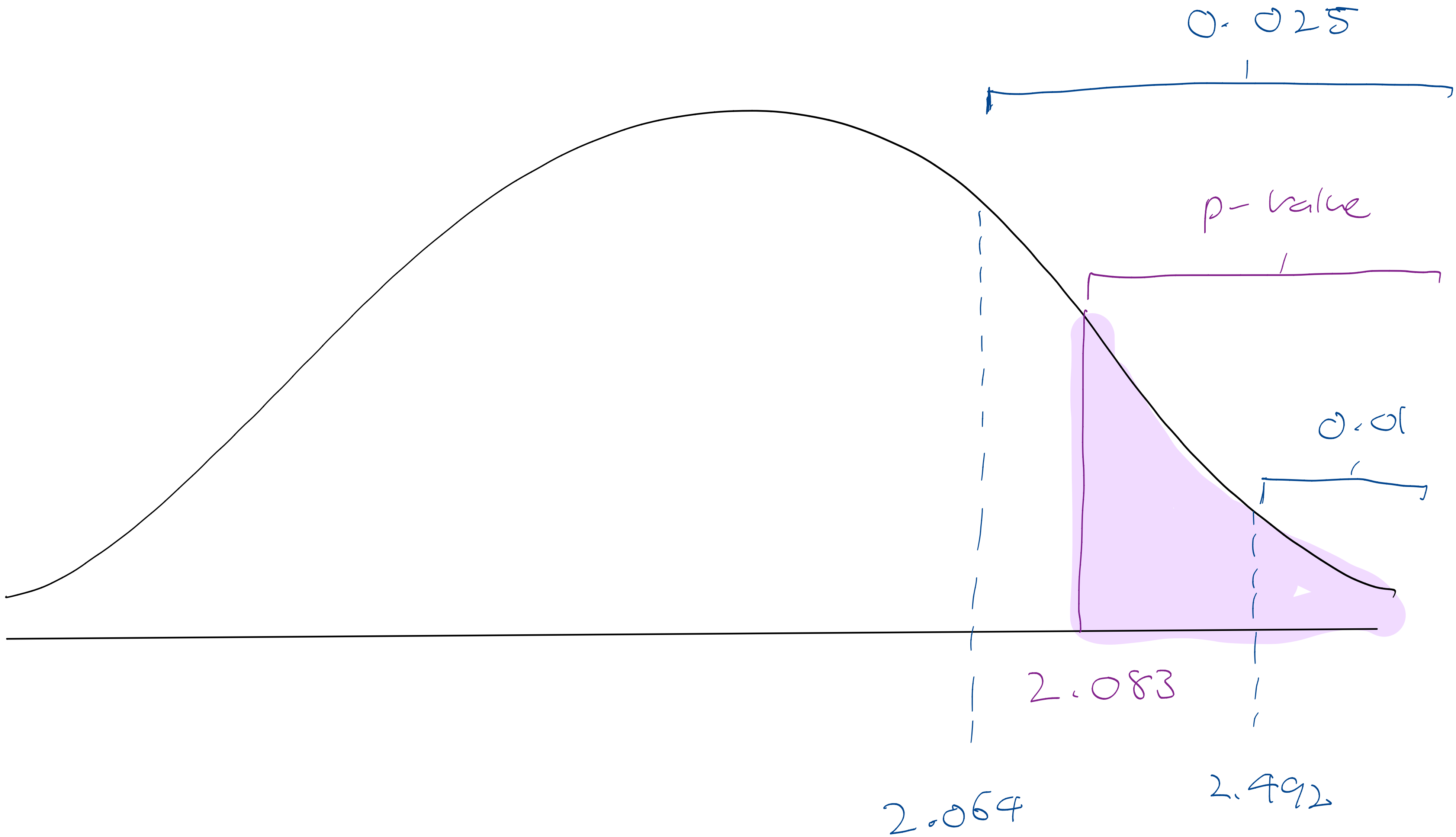
reference distribution:

t distribution at $n-1 = 25-1 = 24$ df

4/ p-value

t distrib at 24 df

$H_a: \mu > 71$



$0.01 < \text{p-value} < 0.025$

5/ make a conclusion $\alpha = 0.05$

$$0.01 < p\text{-value} < 0.025 < \underset{(\alpha)}{0.05}$$

$$p\text{-value} < \alpha$$

There is sufficient evidence at the 5% level of significance to reject the null that the mean is 71 bpm in favour of the alternative that the mean is greater than 71 bpm for people who are laughing.

Suppose we conducted the following test

$H_0: \mu = 71$ $\quad$ $H_a: \mu \neq 71$

calculate test stat

Since σ is unknown

$$t^* = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{73.5 - 71}{6/\sqrt{25}} = 2.083$$

reference distribution:

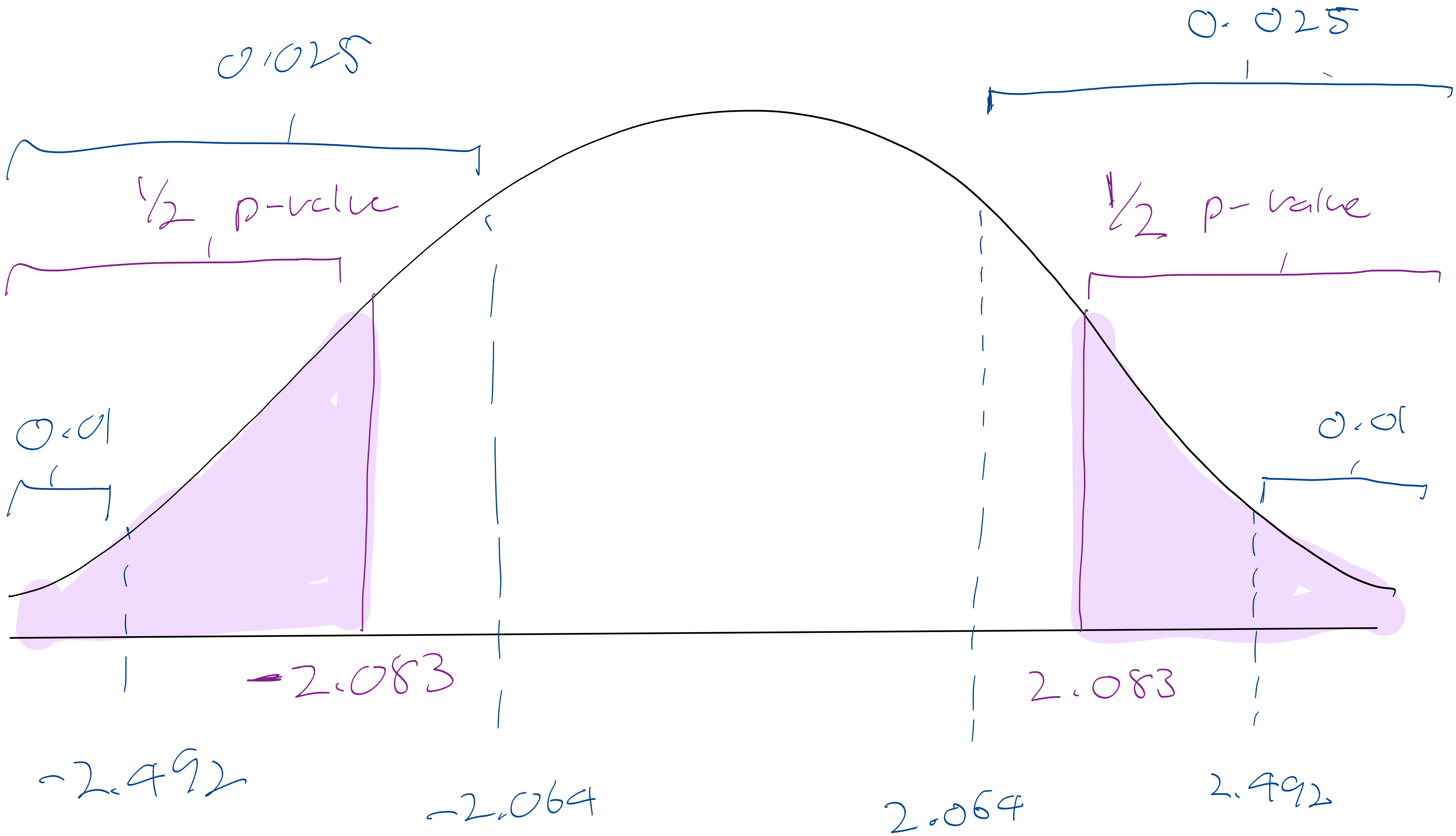
t distribution at $n-1 = 25-1 = 24$ df

p-value

p-value

t distrib at 24 df

$H_a: \mu \neq 71$



$0.01 < \frac{1}{2}\ p\text{-value} < 0.025$

$\times 2 \qquad \times 2 \qquad \times 2$

$0.02 < p\text{-value} < 0.05$

---

why is the following incorrect?

$H_0: \cancel{\bar{x}} = 10 \qquad H_a: \bar{x} > 10$

$\bar{x}$ is calculated and known (nothing to test)

Solution

1. State hypotheses. $H_0 : \mu = 71$ vs $H_a : \mu > 71$.
2. Test statistic. $t^* = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{73.5-71}{6/\sqrt{25}} = 2.083$
3. P-value. For $df = n-1 = 25-1 = 24$, using t table,

P-value$= P[t_{24} > t^*] = P[t_{24} > 2.083]$ is between 0.01 and 0.025

```
1-pt(2.083,df=24);

## [1] 0.02403853
```

4. Conclusion. Since P-value < 0.05, we reject H_0. We have evidence to indicate that the true mean heart rate during laughter exceeds 71 beats per minute.

Example

A researcher is asked to test the hypothesis that the average price of a 2-star (CAA rating) motel room has decreased since last year. Last year a study showed that the prices were Normally distributed with an average of \$89.50. A random sample of twelve 2-star motels has yielded the following information on room prices:
\$ 85.00, 92.50, 87.50, 89.90, 90.00, 82.50, 87.50, 90.00, 85.00, 89.00, 91.50, and \$87.50. If it is believed that the distribution of room prices is Normal, at the 5% level of significance, what conclusion should the researcher make?

Solution

Let μ be the **true** average price of a 2-star motel room.

1. State hypotheses. $H_0 : \mu = 89.5$ vs $H_a : \mu < 89.5$.
2. Test statistic. $t^* = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{88.1583-89.5}{2.9203/\sqrt{12}} = -1.5915$
3. P-value. For $df = 11$, using t table, P-value is between 0.05 and 0.10
4. Conclusion. Since P-value > 0.05, we fail to reject H_0. In other words, the evidence does not indicate that the mean price of a 2-star motel room has decreased this year.

```r
# Step 1. Entering data;

prices=c(85.00, 92.50, 87.50, 89.90, 90.00, 82.50,
87.50, 90.00, 85.00, 89.00, 91.50, 87.50);



# Step 2. Hypothesis test;

t.test(prices,alternative="less", mu=89.5);
```

```
##
##  One Sample t-test
##
## data:  prices
## t = -1.5915, df = 11, p-value = 0.0699
## alternative hypothesis: true mean is less than 89.5
## 95 percent confidence interval:
##      -Inf 89.67229
## sample estimates:
## mean of x
##  88.15833
```

Additional examples.

Is it significant?

The one-sample t statistic for testing
$H_0 : \mu = 0$
$H_a : \mu > 0$
from a sample of $n = 20$ observations has the value $t^* = 1.84$.
a) What are the degrees of freedom for this statistic?
b) Give the two critical values t from Table that bracket t^*. What are the one-sided P-values for these two entries?
c) Is the value $t^* = 1.84$ significant at the 5% level? Is it significant at the 1% level?
d) (Optional) If you have access to suitable technology, give the exact one-sided P-value for $t^* = 1.84$?

Solution

a) df = 20 -1 = 19.
b) $t^* = 1.84$ is bracketed by $t = 1.729$ (with right-tail probability 0.05) and $t = 2.093$ (with right-tail probability 0.025). Hence, because this is a one-sided significance test, $0.025 <$ P-value < 0.05.
c) This test is significant at the 5% level because the P-value < 0.05. It is not significant at the 1% level because the P-value > 0.01.

Solution d)

```r
1 - pt(1.84,df=19);

## [1] 0.04072234

# pt gives you the area to the left of 1.84
# for a T distribution with df =19;
```

Is it significant?

The one-sample t statistic from a sample of $n = 15$ observations for the two-sided test of
$H_0 : \mu = 64$
$H_a : \mu \neq 64$
has the value $t^* = 2.12$.
a) What are the degrees of freedom for t^*?
b) Locate the two-critical values t from Table that bracket t^*. What are the two-sided P-values for these two entries?
c) is the value $t^* = 2.12$ statistically significant at the 10% level? At the 5% level?
d) (Optional) If you have access to suitable technology, give the exact two-sided P-value for $t^* = 2.12$.

Solution

a) df = 15 -1 = 14.
b) $t^* = 2.12$ is bracketed by $t = 1.761$ (with two-tail probability 0.10) and $t = 2.145$ (with two-tail probability 0.05). Hence, because this is a two-sided significance test, $0.05 <$ P-value < 0.10.
c) This test is significant at the 10% level because the P-value < 0.10. It is not significant at the 5% level because the P-value > 0.05.

Solution d)

```
2*(1 - pt(2.12,df=14));

## [1] 0.05235683

# pt gives you the area to the left of 2.12
# for a T distribution with df =12;
```

Example

$H_0 : \mu = 12$
$H_a : \mu > 12$
A sample of 25 provided a sample mean $\bar{x} = 14$ and a sample standard deviation $s = 4.32$.

a. Compute the value of the test statistic.
b. Use the t distribution table to compute a range for the p-value.
c. At $\alpha = 0.05$, what is your conclusion?

Solution

a. $t_* = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{14-12}{4.32/\sqrt{25}} = 2.31$
b. Degrees of freedom $= n-1 = 24$.
P-value $= P(T > t_*) = P(T > 2.31)$
Using t-table, P-value is between 0.01 and 0.02.
Exact P-value $= 0.0149$ (using R).
c. Since P-value $< \alpha = 0.05$, we reject H_0.

Example

$H_0 : \mu = 18$
$H_a : \mu \neq 18$
A sample of 48 provided a sample mean $\bar{x} = 17$ and a sample standard deviation $s = 4.5$.

a. Compute the value of the test statistic.
b. Use the t distribution table to compute a range for the p-value.
c. At $\alpha = 0.05$, what is your conclusion?

Test of Hypothesis for One Proportion

Example: 100-Cup Challenge

A YouTuber goes to her nearest Tim Hortons and buys 100 empty cups. After rolling up the rims, she ends up with 12 winning cups out of the 100 she bought, all of them were food prizes.
If the probability of winning a food prize is supposed to be $\frac{1}{6}$, does she have evidence to claim that the probability of winning a food prize is less than $\frac{1}{6}$?

Probability Models

The **sample space S** of a random phenomenon is the set of all possible outcomes.
An **event** is an outcome or a set of outcomes of a random phenomenon. That is, an event is a subset of the sample space.
A **probability model** is a mathematical description of a random phenomenon consisting of two parts: a sample space S and a way of assigning probabilities to events.

Example

Rolling a fair die (random phenomenon). There are 6 possible outcomes when we roll a die.
The sample space for rolling a die and counting the pips is

$$S = \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$$

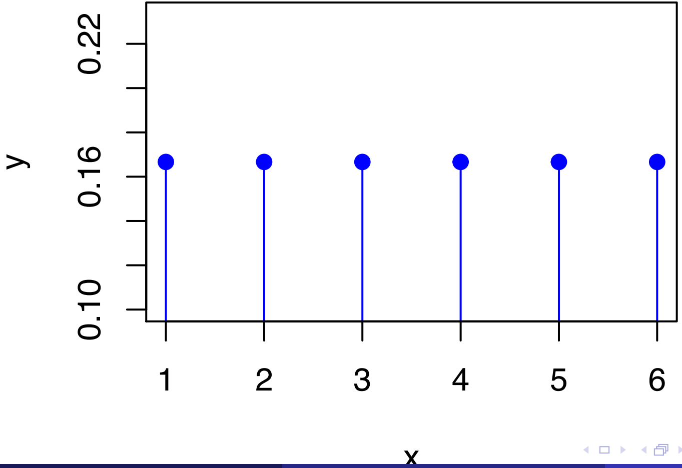
"Roll a 6" is an event, that contains one of these 6 outcomes.

Discrete Uniform Distribution

A random variable X has a **discrete uniform distribution** if each of the n values in its range, say, x_1, x_2, . . . , x_n has equal probability. Then,

$$f(x_i) = \frac{1}{n}$$

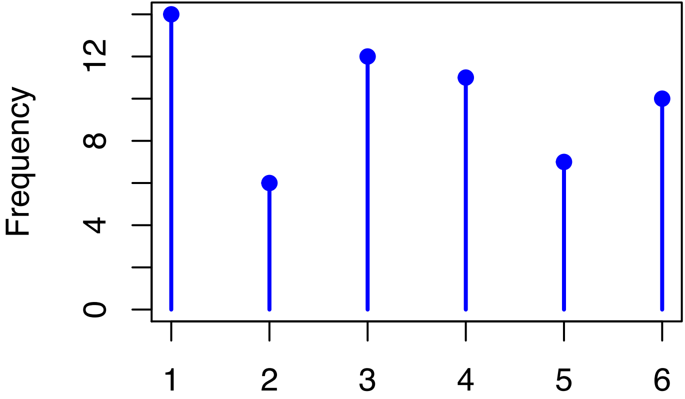
Probability mass function (pmf)



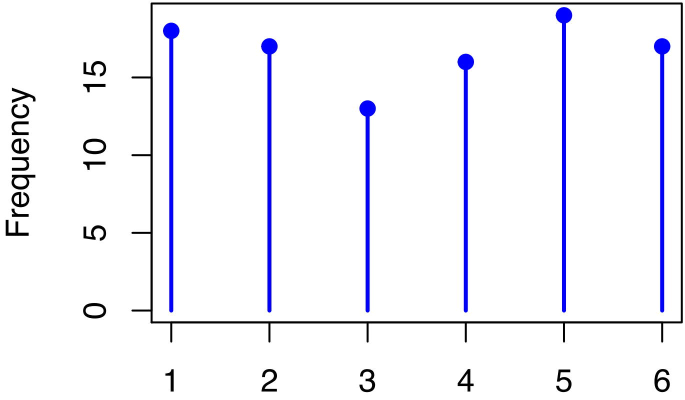
Six simulations

```r
die=c(1,2,3,4,5,6);

sample(die,1,replace=TRUE);

## [1] 2

sample(die,6,replace=TRUE);

## [1] 1 3 2 6 3 1
```

60 simulations



100 simulations



Random Variable

A **random variable** is a variable whose value is a numerical outcome of a random phenomenon.
The **probability distribution** of a random variable X tells us what values X can take and how to assign probabilities to those values.

The Binomial setting

- There are a fixed number n of observations.
- The n observations are all **independent**. That is, knowing the result of one observation tells you nothing about the other obsevations.
- Each observation falls into one of just two categories, which for convenience we call "success" and "failure".
- The probability of a success, call it p, is the same for each observation.

Example

Think of rolling a die n times as an example of the binomial setting. Each roll gives either a six or a number different from six. Knowing the outcome of one roll doesn't tell us anything about other rolls, so the n rolls are independent. If we call six a success, then p is the probability of a six and remains the same as long as we roll the same die. The number of sixes we count is a random variable X. The distribution of X is called a **binomial distribution**.

Binomial Distribution

A random variable Y is said to have a **binomial distribution** based on n trials with success probability p if and only if

$$p(y) = \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y}, \ \ y = 0, 1, 2, ..., n \text{ and } 0 \leq p \leq 1.$$

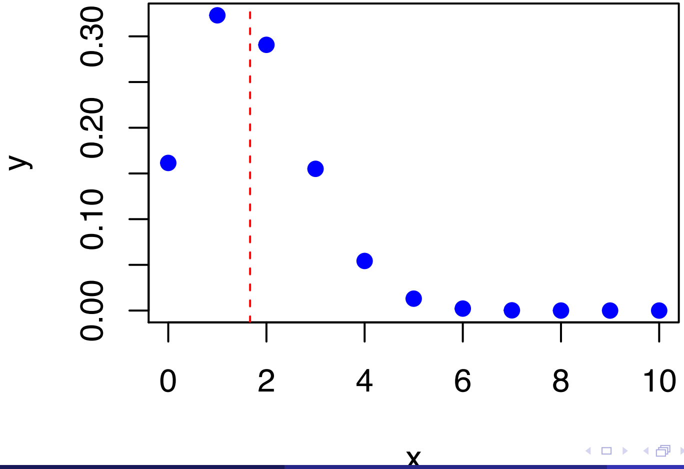
A few simulations

```
## Simulation: Binomial with n=10 and p=1/6.

rbinom(1,size=10,prob=1/6);

## [1] 3

rbinom(1,size=10,prob=1/6);

## [1] 1

rbinom(1,size=10,prob=1/6);

## [1] 0
```

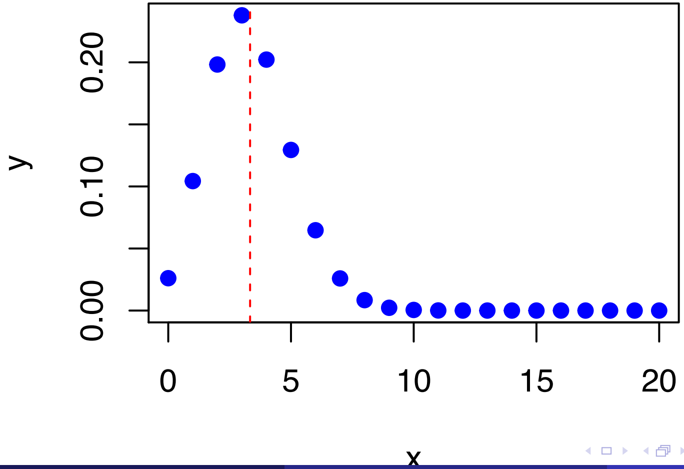
Probability Mass Function when n=10 and p=1/6

```
## Pmf: Binomial with n=10 and p=1/6.

x<-seq(0,10,by=1);

y<-dbinom(x,10,1/6);

plot(x,y,type="p",col="blue",pch=19);
```

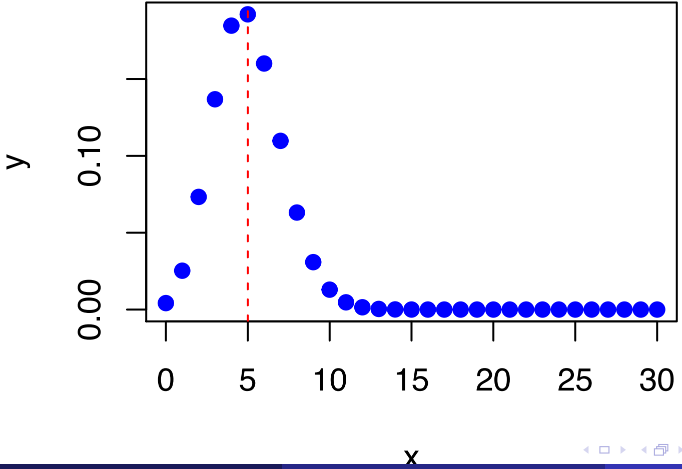
Probability Mass Function when n=10 and p=1/6



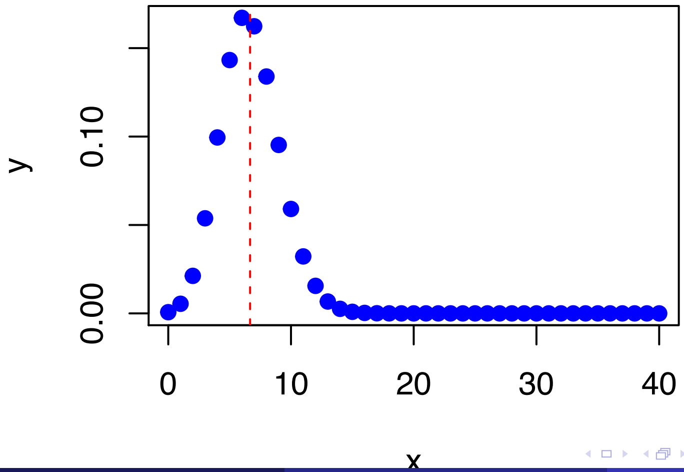
Pmf when n=20 and p=1/6



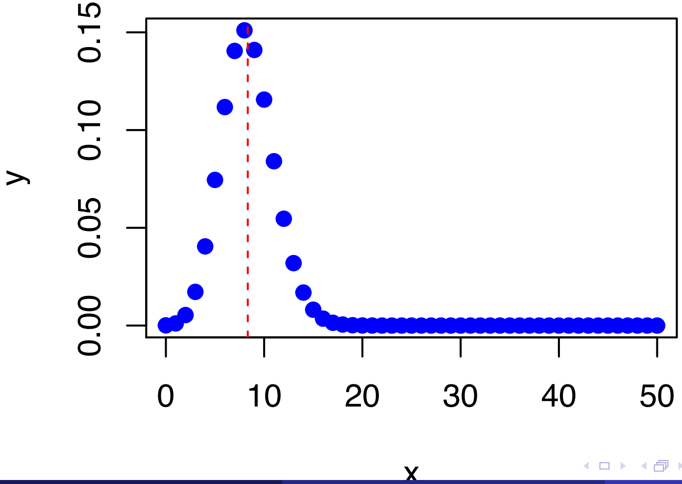
Pmf when n=30 and p=1/6



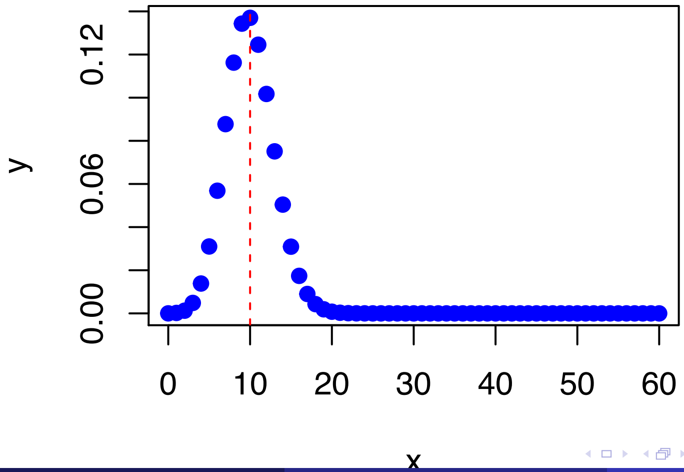
Pmf when n=40 and p=1/6



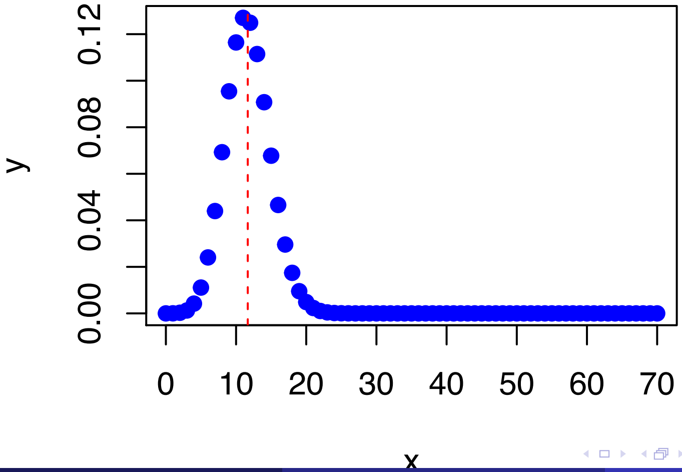
Pmf when n=50 and p=1/6



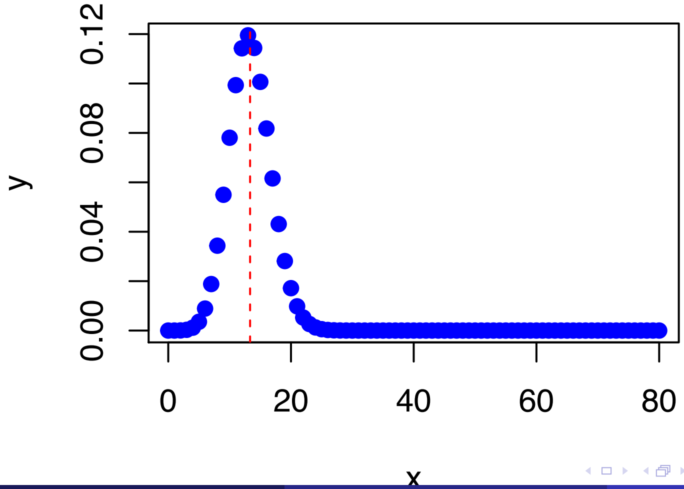
Pmf when n=60 and p=1/6



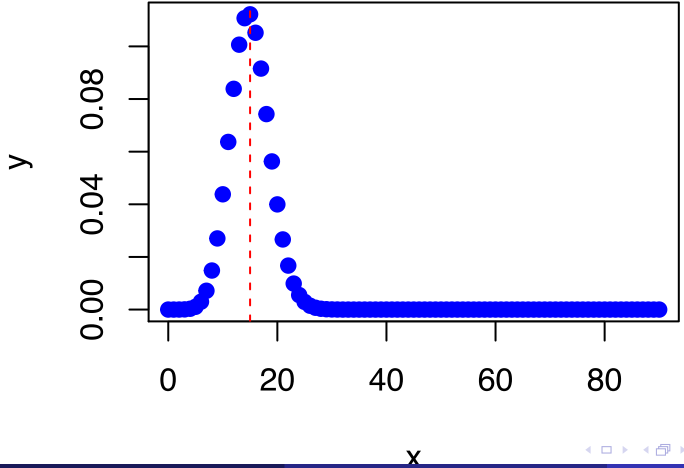
Pmf when n=70 and p=1/6



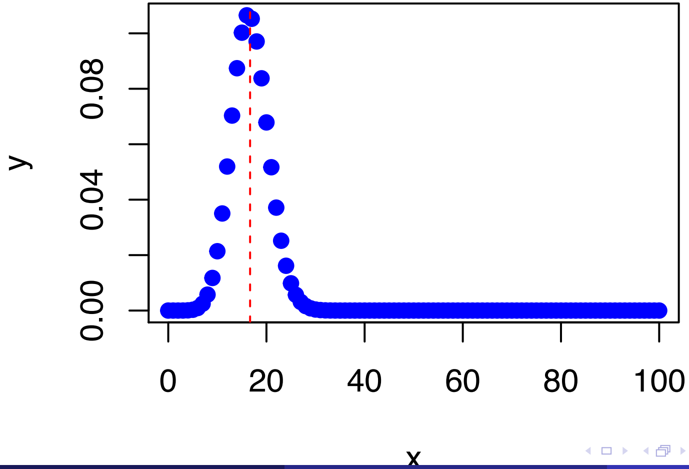
Pmf when n=80 and p=1/6



Pmf n=90 and p=1/6



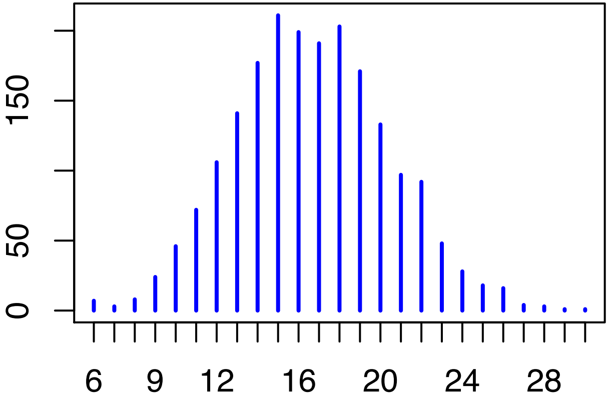
Pmf when n=100 and p=1/6



A few values from our pmf (n=100 and p=1/6)

```
dbinom(c(15,16,17,18),size=100,prob=1/6);

## [1] 0.10023663 0.10650142 0.10524847 0.09706247
```

Simulation: 2000 YouTuber, n=100, and p=1/6



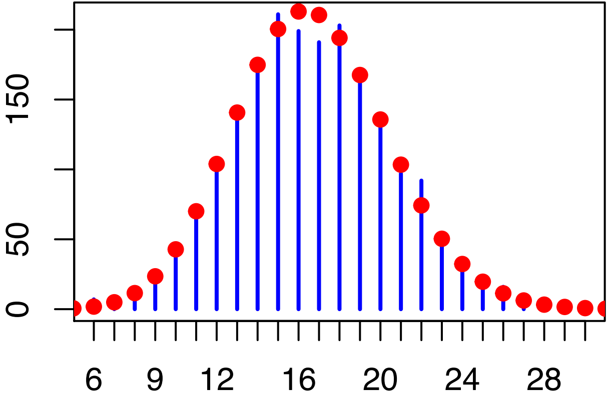
A few values from our simulation

```
## vec.prop
##    6    7    8    9   10   11   12
##    7    3    8   24   46   72  106
## [1] 266
## [1] 0.133
```

P-value

It turns out that our P-value for this simulation is:
0.133

Simulation vs Theoretical pmf



Sampling Distribution of a sample proportion

Draw an SRS of size n from a large population that contains proportion p of "successes". Let $\hat{p}$ be the **sample proportion** of successes,

$$\hat{p} = \frac{\text{number of successes in the sample}}{n}$$

Then:

- The **mean** of the sampling distribution of $\hat{p}$ is p.
- The **standard deviation** of the sampling distribution is

$$\sqrt{\frac{p(1-p)}{n}}.$$

Sampling Distribution of a sample proportion

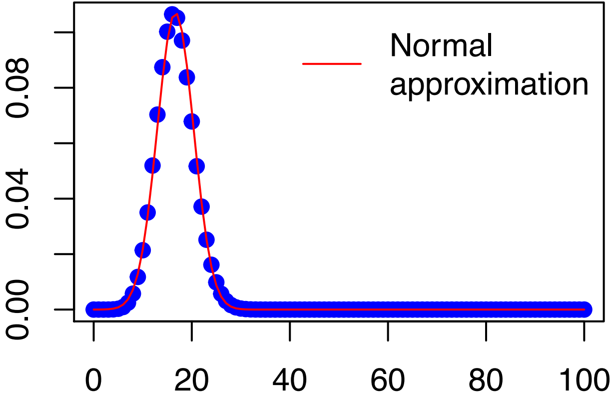
Draw an SRS of size n from a large population that contains proportion p of "successes". Let $\hat{p}$ be the **sample proportion** of successes,

$$\hat{p} = \frac{\text{number of successes in the sample}}{n}$$

Then:

- As the sample size increases, the sampling distribution of $\hat{p}$ becomes **approximately Normal.** That is, for large n, $\hat{p}$ has approximately the $N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$ distribution.

Binomial with Normal Approximation



Hypotheses Tests for a Proportion

To test the hypothesis $H_0 : p = p_0$, compute the z_* statistic,
$z_* = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$
In terms of a variable Z having the standard Normal distribution, the approximate P-value for a test of H_0 against
$H_a : p > p_0$: is : $P(Z > z_*)$
$H_a : p < p_0$: is : $P(Z < z_*)$
$H_a : p \neq p_0$: is : $2P(Z > |z_*|)$

Solution

Step 1. $H_0 : p = \frac{1}{6}$ vs $H_a : p < \frac{1}{6}$

$\hat{p} = \frac{12}{100} = 0.12$

Step 2. (Without continuity correction)

$$z_* = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.12 - 0.1667}{\sqrt{\frac{(0.1667)(1-0.1667)}{100}}} \approx -1.25$$

(WITH continuity correction)

$$P[X \leq 12] \approx P[X \leq 12.5] = P[\frac{X}{n} \leq 0.125] = P\left[\frac{\bar{X} - \mu}{\sigma/n} \leq \frac{0.125 - 0.1667}{0.0373}\right]$$
$$= P[Z \leq -1.1179]$$

Step 3. Using Normal table, P-value $=$

$P(Z < z_*) = P(Z < -1.1179) \approx 0.1314$

P-value is not small enough to provide evidence against H_0, we can't reject H_0. We conclude that there is not evidence to claim that probability of winning a food prize is less than $\frac{1}{6}$.

R Code

```
prop.test(12,100,p=1/6,alternative="less");

##
##  1-sample proportions test with continuity correction
##
## data:  12 out of 100, null probability 1/6
## X-squared = 1.25, df = 1, p-value = 0.1318
## alternative hypothesis: true p is less than 0.1666667
## 95 percent confidence interval:
##  0.0000000 0.1894571
## sample estimates:
##    p
## 0.12
```

Hypotheses Tests for a Proportion

To test the hypothesis $H_0 : p = p_0$, compute the z_* statistic,
$z_* = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$
In terms of a variable Z having the standard Normal distribution, the approximate P-value for a test of H_0 against
$H_a : p > p_0$: is : $P(Z > z_*)$
$H_a : p < p_0$: is : $P(Z < z_*)$
$H_a : p \neq p_0$: is : $2P(Z > |z_*|)$

Introduction to Hypothesis Testing (Significance Test)

Consider the following problem: In 1980s, it was generally believed that congenital abnormalities affect 5% of the nation's children. Some people believe that the increase in the number of chemicals in the environment in recent years has led to an increase in the incidence of abnormalities. A recent study examined 384 children and found that 46 of them showed signs of abnormality. Is this strong evidence that the risk has increased?

- The above statement serves as a hypothesis, moreover it is a Research Hypothesis.

A hypothesis is:

- a statement about a population.
- a predication that a parameter describing some characteristics of a variable (e.g., true proportion, p) takes a particular numerical value or falls in a certain range of values.

Introduction to Hypothesis Testing (Significance Test)

For conducting a Significance Test:

- Researchers (you) use data to summarize the evidence about a hypothesis.
- With data, you can compare the point estimates of parameters to the values predicted by the hypothesis.

Important Ideas about Hypothesis Testing

- All the hypothesis tests boils down to the same question: "Is an observed difference or pattern too large to be attributed to chance?"
- We measure "how large" by putting our sample results in the context of a sampling distribution model (e.g., Normal model, t distribution).

Specify Statistical Model

- To plan a statistical hypothesis test, specify the model you will use to test the null hypothesis and the parameter of interest.
- All model require assumptions, so you will need to state them and check any corresponding conditions.
- For example, if the conditions are satisfied, we can model the sampling distribution of the proportion with a Normal model. Otherwise, we cannot proceed with the test (we need to stop and reconsider).

Steps in conducting Hypothesis Testing

1. State the null and the alternative hypothesis.
2. Check the necessary assumptions.
3. Identify the test-statistic. Find the value of the test-statistic.
4. Find the p-value of the test-statistic.
5. State (if any) a conclusion.

Example of Hypothesis Testing for a Proportion

In 1980s, it was generally believed that congenital abnormalities affect 5% of the nation's children. Some people believe that the increase in the number of chemicals in the environment in recent years has led to an increase in the incidence of abnormalities. A recent study examined 384 children and found that 46 of them showed signs of abnormality. Is this strong evidence that the risk has increased?

Example: Hypothesis Testing for a Proportion (One-sided Test) - Step 1

Step 1. Set up the null and alternative hypothesis:

- The null hypothesis is the current belief: $H_0 : p = p_0$

In our example it would have a form: $H_0 : p = 0.05$

- The Alternative hypothesis is what the researcher(s) [you] want to prove: $H_a : p > p_0$

In our example it would have a form: $H_a : p > 0.05$
This means a one-sided test

- The goal here is to provide evidence against H_0 (e.g., suggest H_a).

You want to conclude H_a.
Try a Proof by Contradiction Assume H_0 is true ... and hope your data contradicts it

Example: Hypothesis Testing for a Proportion (One-sided Test) - Step 2

Step 2. Check the Necessary Assumptions:

- Independence Assumption: There is no reason to think that one child having genetic abnormalities would affect the probability that other children have them.
- Randomization Condition: This sample may not be random, but genetic abnormalities are plausibly independent. The sample is probably representative of all children, with regards to genetic abnormalities.
- 10% Condition: The sample of 384 children is less than 10% of all children.
- Success/Failure Condition: $np = (384)(0.05) = 19.2$ and $n(1-p) = (384)(0.95) = 364.8$ are both greater than 10, so the sample is large enough.

Example: Hypothesis Testing for a Proportion (One-sided Test) - Step 3

Step 3. Identify the test-statistics. Find the value of the test-statistic:
Since the conditions are met, assume H_0 is true:
The sampling distribution of $\hat{p}$ becomes **approximately Normal.** That is, for large n, $\hat{p}$ has approximately the $N\left(p_0, \sqrt{\frac{p_0(1-p_0)}{n}}\right)$ distribution.

$z_* = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.1198-0.05}{\sqrt{\frac{(0.05)(0.95)}{384}}} \approx 6.28$

Recall that $\hat{p} = \frac{46}{384} = 0.1198$.

The value of z^* is approximately 6.28, meaning that the observed proportion of children with genetic abnormalities is over 6 standard deviations above the hypothesized proportion $(p_0 = 0.05)$.

About the P-value of the Test-statistics

- P-value is a conditional probability.
- It is not the probability that H_0 (null hypothesis: current belief) is true.
- It is: P(observed statistic value [or even more extreme] — H_0]. Given H_0 (the null hypothesis), because H_0 gives the parameter values that we need to find required probability.
- P-value serves as a measure of the strength of the evidence against the null hypothesis (but it should not serve as a hard and fast rule for decision).
- If p-value $= 0.03$ (for example) all we can say is that there is 3% chance of observing the statistic value we actually observed (or one even more inconsistent with the null value).

P-value of the Test-statistics

- P-value is the chance (the proportion) of getting a, for instance, $\hat{p}$ as far as or further from H_0 than the value observed.
- P-value is the probability of getting at least something (e.g., sample proportion $\hat{p}$) more extreme (e.g., unusual, unlikely, or rare) than what we have already found (our observed value of $\hat{p}$) that provide even stronger evidence against H_0.
- The more extreme the z-score (large in absolute values) are the ones that denote farther departure of the observed value (e.g., our $\hat{p}$) from the parameter value (p_0) in H_0.
- In the one-sided test, e.g., $H_a : p > p_0$, p-value is one-tailed probability. This is the probability that sample proportion $\hat{p}$ falls at least as far from p_0 in one direction as the observed value of $\hat{p}$.
- In the two-sided test, e.g., $H_a : p \neq p_0$, p-value is two-tailed probability. This is the probability that sample proportion $\hat{p}$ falls at least as far from p_0 in either direction as the observed value of $\hat{p}$.

Example: Hypothesis Testing for a Proportion (One-sided Test) - Step 4

Step 4. Find the p-value of the test-statistic.
P-value $= P(Z > 6.28) \approx 0.000$ (better to report $p - value < 0.0001$)
Note: We find the area above Z of 6.28 since $H_a : p > 0.05$.
Meaning of this p-value:
If 5% of children have genetic abnormalities, the chance of observing 46 children with genetic abnormalities in a random sample of 384 children is almost 0.

P-values

The probability, computed assuming that H_0 is true, that the test statistic would take a value as extreme or more extreme than that actually observed is called the **P-value** of the test. The smaller the P-value, the stronger the evidence against H_0 provided by the data.

Small P-values are evidence against H_0, because they say that the observed result is unlikely to occur when H_0 is true. Large P-values fail to give evidence against H_0.

The P-value Scale

If P-value < 0.001, we have very strong evidence against H_0.
If $0.001 \leq$ P-value < 0.01, we have strong evidence against H_0.
If $0.01 \leq$ P-value < 0.05, we have evidence against H_0.
If $0.05 \leq$ P-value < 0.075, we have some evidence against H_0.
If $0.075 \leq$ P-value < 0.10, we have slight evidence against H_0.

Use p-value Method to Make a Decision (Reject or Fail to Reject Ho)

But how small is small p-value?
We would need to choose an α-level (significance-level): a number such that if:

- $P - value \leq \alpha$-level, we reject H_0; We can conclude H_a (we have evidence to support our claim). Often we phrase as a statistically significant result at that specified α-level.
- $P - value > \alpha$-level, we fail to reject H_0; We cannot conclude H_a (we have not enough evidence to support our claim; thus, H_0 is plausible - We do not accept H_0). Often we phrase as the result is not statistically significant at that specified α-level.
- The default α-level (significance-level) is typically $\alpha = 0.05$ (but it can be different based on the context of the study - it is usually not higher than 0.10).

Example: Hypothesis Testing for a Proportion (One-sided Test) - Step 5

Step 5. Give (if any) a conclusion.
p-value is less than 0.0001, which is less than $\alpha = 0.05$; We reject $H_0 : p = 0.05$, and conclude $H_a : p > 0.05$. Our result is statistically significant at $\alpha = 0.05$.
There is a very strong evidence that more than 5% of children have genetic abnormalities.

Example: Hypothesis Testing for a Proportion (One-sided Test) - All Steps

$H_0 : p = 0.05$ vs $H_a : p > 0.05$
$\hat{p} = \frac{46}{384} \approx 0.1198$
$z_* = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.1198 - 0.05}{\sqrt{\frac{(0.05)(0.95)}{384}}} \approx 6.28$
P-value $= P(Z > z_*) = P(Z > 6.28) \approx 1.747 \times 10^{-10}$
P-value $= < \alpha$, we reject $H_0 : p = 0.05$. Our result is statistically significant at $\alpha = 0.05$. There is a very strong evidence that more than 5% of children have genetic abnormalities.

R Code

```
prop.test(x=46, n = 384 ,p=0.05,alternative="greater",
correct=FALSE);

##
##  1-sample proportions test without continuity correction
##
## data:  46 out of 384, null probability 0.05
## X-squared = 39.377, df = 1, p-value = 1.747e-10
## alternative hypothesis: true p is greater than 0.05
## 95 percent confidence interval:
##  0.09516097 1.00000000
## sample estimates:
##         p
## 0.1197917
```

The p-value in the previous example was extremely small (less than 0.0001). That is a strong evidence to suggest that more than 5% of children have genetic abnormalities. However, it does not say that the percentage of sampled children with genetic abnormalities was "a lot more than 5%". That is, the p-value by itself says nothing about how much greater the percentage might be. The confidence interval provides that information.

95% CI for a Proportion

To assess the difference in practical terms, we should also construct a confidence interval:

$$\begin{aligned} 0.1198 \pm (1.96 \times 0.0166) \\ 0.1198 \pm 0.0324 \\ (0.0874, 0.1522) \end{aligned}$$

Interpretation: We are 95% Confident that the true percentage of children with genetic abnormalities is between 8.74% and 15.22%.

```
prop.test(x=46, n = 384, correct=FALSE);

##
##  1-sample proportions test without continuity correction
##
## data:  46 out of 384, null probability 0.5
## X-squared = 222.04, df = 1, p-value < 2.2e-16
## alternative hypothesis: true p is not equal to 0.5
## 95 percent confidence interval:
##  0.09102214 0.15609290
## sample estimates:
##         p
## 0.1197917
```

95% CI for p: (9.1%, 15.6%) - We are 95% confident that the true percentage of all children that have genetic abnormalities is between approximately 9.1% and 15.6%. Since both values of this CI are more than the hypothesized value of $p = 0.05$ (5%), we can further infer that this true percentage is more than 5%.

Do environmental chemicals cause congenital abnormalities?

We do not know that environmental chemicals cause genetic abnormalities. We merely have evidence that suggests that a greater percentage of children are diagnosed with genetic abnormalities now, compared to the 1980s.

More About P-values

- Big p-values just mean that what we have observed is not surprising. It means that the results are in line with our assumption that the null hypothesis models the world, so we have no reason to reject it.
- A big p-value does not prove that the null hypothesis is true.
- When we see a big p-value, all we can say is: we cannot reject H_0 (we fail to reject H_0) - we cannot conclude H_a (We have no evidence to support H_a).

Additional Examples.

Example

Consider the following hypothesis test:
$H_0 : p = 0.75$
$H_a : p < 0.75$
A sample of 300 items was selected. Compute the p-value and state your conclusion for each of the following sample results. Use $\alpha = 0.05$.
a. $\hat{p} = 0.68$
b. $\hat{p} = 0.72$
c. $\hat{p} = 0.70$
d. $\hat{p} = 0.77$

Solution a.

$z_* = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.68 - 0.75}{\sqrt{0.75(1-0.75)/300}} = -2.80$
Using Normal table, P-value $= P(Z < z_*) = P(Z < -2.80) = 0.0026$
P-value$<\alpha = 0.05$, reject H_0.

Solution b.

$z_* = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.72-0.75}{\sqrt{0.75(1-0.75)/300}} = -1.20$
Using Normal table, P-value $= P(Z < z_*) = P(Z < -1.20) = 0.1151$
P-value$> \alpha = 0.05$, do not reject H_0.

Solution c.

$z_* = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.70-0.75}{\sqrt{0.75(1-0.75)/300}} = -2.00$
Using Normal table, P-value $= P(Z < z_*) = P(Z < -2.00) = 0.0228$
P-value$< \alpha = 0.05$, reject H_0.

Example

Consider the following hypothesis test:
$H_0 : p = 0.20$
$H_a : p \neq 0.20$
A sample of 400 provided a sample proportion $\hat{p} = 0.175$.
a. Compute the value of the test statistic.
b. What is the p-value?
c. At the $\alpha = 0.05$, what is your conclusion?
d. What is the rejection rule using the critical value? What is your conclusion?

Solution

a. $z_* = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.175-0.20}{\sqrt{\frac{(0.20)(0.80)}{400}}} = -1.25$

b. Using Normal table, P-value $=$ $2P(Z > |z_*|) = 2P(Z > |-1.25|) = 2P(Z > 1.25) = 2(0.1056) = 0.2112$

c. P-value $> \alpha = 0.05$, we CAN'T reject H_0.

Problem

A study found that, in 2005, 12.5% of U.S. workers belonged to unions. Suppose a sample of 400 U.S. workers is collected in 2006 to determine whether union efforts to organize have increased union membership.

a. Formulate the hypotheses that can be used to determine whether union membership increased in 2006.

b. If the sample results show that 52 of the workers belonged to unions, what is the p-value for your hypothesis test?

c. At $\alpha = 0.05$, what is your conclusion?

Solution

a. $H_0 : p = 0.125$ vs $H_a : p > 0.125$
b. $\hat{p} = \frac{52}{400} = 0.13$
$z_* = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.13 - 0.125}{\sqrt{\frac{(0.125)(0.875)}{400}}} = 0.30$
Using Normal table, P-value =
$P(Z > z_*) = P(Z > 0.30) = 1 - 0.6179 = 0.3821$
c. P-value = > 0.05, do not reject H_0. We cannot conclude that there has been an increase in union membership.

R Code

```
prop.test(52,400,p=0.125,alternative="greater",
correct=FALSE);

##
##  1-sample proportions test without continuity correction
##
## data:  52 out of 400, null probability 0.125
## X-squared = 0.091429, df = 1, p-value = 0.3812
## alternative hypothesis: true p is greater than 0.125
## 95 percent confidence interval:
##  0.1048085 1.0000000
## sample estimates:
##    p
## 0.13
```

Problem

A study by Consumer Reports showed that 64% of supermarket shoppers believe supermarket brands to be as good as national name brands. To investigate whether this result applies to its own product, the manufacturer of a national name-brand ketchup asked a sample of shoppers whether they believed that supermarket ketchup was as good as the national brand ketchup.

Problem (cont.)

a. Formulate the hypotheses that could be used to determine whether the percentage of supermarket shoppers who believe that the supermarket ketchup was as good as the national brand ketchup differed from 64%.
b. If a sample of 100 shoppers showed 52 stating that the supermarket brand was as good as the national brand, what is the p-value?
c. At $\alpha = 0.05$, what is your conclusion?

Solution

a. $H_0 : p = 0.64$ vs $H_a : p \neq 0.64$
b. $\hat{p} = \frac{52}{100} = 0.52$
$z_* = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.52 - 0.64}{\sqrt{\frac{(0.64)(0.36)}{100}}} = -2.50$
Using Normal table, P-value $=$
$2P(Z > |z_*|) = 2P(Z > |-2.50|) = 2P(Z > 2.50) = 2(0.0062) = 0.0124$
c. P-value $= < 0.05$, reject H_0. Proportion differs from the reported 0.64.

R Code

```
prop.test(52,100,p=0.64,alternative="two.sided",
correct=FALSE);

##
##  1-sample proportions test without continuity correction
##
## data:  52 out of 100, null probability 0.64
## X-squared = 6.25, df = 1, p-value = 0.01242
## alternative hypothesis: true p is not equal to 0.64
## 95 percent confidence interval:
##  0.4231658 0.6153545
## sample estimates:
##    p
## 0.52
```

Problem

The National Center for Health Statistics released a report that stated 70% of adults do not exercise regularly. A researcher decided to conduct a study to see whether the claim made by the National Center for Health Statistics differed on a state-by-state basis.
a. State the null and alternative hypotheses assuming the intent of the researcher is to identify states that differ from 70% reported by the National Center for Health Statistics.
b. At $\alpha = 0.05$, what is the research conclusion for the following state: Wisconsin: 252 of 350 adults did not exercise regularly.

a. $H_0 : p = 0.70$ vs $H_a : p \neq 0.70$
b. Wisconsin $\hat{p} = \frac{252}{350} = 0.72$
$z_* = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.72 - 0.70}{\sqrt{\frac{(0.70)(0.30)}{350}}} = 0.82$
Using Normal table, P-value $=$
$2P(Z > |z_*|) = 2P(Z > |0.82|) = 2P(Z > 0.82) = 2(0.2061) = 0.4122$
c. P-value > 0.05, we don't have enough evidence to reject H_0. There is not enough evidence against the claim made by the National Center for Health Statistics.

Solution

a. $t_* = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{17-18}{4.5/\sqrt{48}} = -1.54$

b. Degrees of freedom $= n - 1 = 47$.

P-value $= 2P(T > |t_*|) = 2P(T > |-1.54|) = 2P(T > 1.54)$

Using t-table, P-value is between 0.10 and 0.20.

Exact P-value $= 0.1303$ (using R).

c. Since P-value $> \alpha = 0.05$, we CAN'T reject H_0.

Test of Hypothesis for One Variance

Hypothesis Tests for One Variance

- Data from a single normal population Independent observations
- Variance unknown
- Large or small sample

Hypothesis Test
$H_0 : \sigma^2 = \sigma_0^2$ vs $H_a : \sigma^2 \neq \sigma_0^2$ (or $\sigma^2 > \sigma_0^2$ or $\sigma^2 < \sigma_0^2$).
Assume H_0 is true, then

$$\text{Test statistic: } \chi_*^2 = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi_{n-1}^2$$

Decision rules:
$H_a : \sigma^2 \neq \sigma_0^2$.
Reject H_0 if $\chi^2_* > \chi^2_{n-1;\alpha/2}$ or if $\chi^2_* < \chi^2_{n-1;1-\alpha/2}$.
$H_a : \sigma^2 > \sigma_0^2$.
Reject H_0 if $\chi^2_* > \chi^2_{n-1;\alpha}$ or if $P[\chi^2_{n-1} > \chi^2_*]$ is too small.
$H_a : \sigma^2 < \sigma_0^2$.
Reject H_0 if $\chi^2_* < \chi^2_{n-1;1-\alpha}$ or if $P[\chi^2_{n-1} < \chi^2_*]$ is too small.
Note. This is NOT robust to departures from Normality.

Example

A company produces metal pipes of a standard length, and claims that the standard deviation of the length is at most 1.2 cm. One of its clients decides to test this claim by taking a sample of 25 pipes and checking their lengths. They found that the standard deviation of the sample is 1.5 cm. Does this undermine the company's claim? Use $\alpha = 0.05$.
Note. Assume length is Normally distributed.

Solution

$H_0 : \sigma^2 \leq 1.2^2$ vs $H_a : \sigma^2 > 1.2^2$.
$\chi^2_* = \frac{(n-1)s^2}{\sigma^2} = \frac{(25-1)1.5^2}{1.2^2} = 37.5$
P-value $= P[\chi^2_{24} > 37.5] \approx 0.0389$

```
1-pchisq(37.5,df=24);

## [1] 0.0389818
```

Conclusion

We reject $H_0 : \sigma^2 \leq 1.2^2$. We have evidence to indicate that the variance of the length of metal pipes is more than 1.2^2.

Test of Hypotheses concerning a Population Variance

Assumptions: $Y_1, Y_2, ..., Y_n$ constitute a random sample from a Normal distribution with $E(Y_i) = \mu$ and $V(Y_i) = \sigma^2$.

$H_0 : \sigma^2 = \sigma_0^2$

$$H_a : \begin{cases} \sigma^2 > \sigma_0^2 & \text{upper-tailed alternative} \\ \sigma^2 < \sigma_0^2 & \text{lower-tailed alternative} \\ \sigma^2 \neq \sigma_0^2 & \text{two-tailed alternative} \end{cases}$$

Test of Hypotheses concerning a Population Variance

Test statistic: $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$

$$\text{Rejection Region} \; : \begin{cases} \chi^2 > \chi^2_{\alpha} & \text{upper-tailed RR} \\ \chi^2 < \chi^2_{1-\alpha} & \text{lower-tailed RR} \\ \chi^2 > \chi^2_{\alpha/2} \; \textit{or} \; \chi^2 < \chi^2_{1-\alpha/2} & \text{two-tailed RR} \end{cases}$$

Example

A manufacturer of car batteries claims that the life of his batteries is approximately Normally distributed with a standard deviation equal to 0.9 year. If a random sample of 10 of these batteries has a standard deviation of 1.2 years, do you think that $\sigma > 0.9$ year? Use a 0.05 level of significance.

Solution

Step 1. State hypotheses.
$H_0 : \sigma^2 = 0.81$

$H_a : \sigma^2 > 0.81$

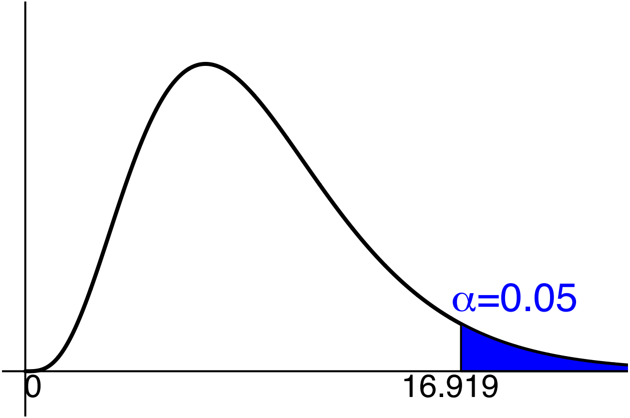
Solution

Step 2. Compute test statistic.
$S^2 = 1.44$, $n = 10$, and

$$\chi^2 = \frac{(9)(1.44)}{0.81} = 16$$

Solution

Step 3. Find Rejection Region.
From Figure and our table we see that the null hypothesis is rejected when $\chi^2 > 16.919$, where $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$ with $\nu = 9$ degrees of freedom.


α=0.05
0
16.919

Solution

Step 4. Conclusion.

The χ^2 statistic is not significant at the 0.05 level. We conclude that there is insufficient evidence to claim that $\sigma > 0.9$ year.