Econ 5023: Statistics for Decision Making

UHypothesis Testing (II): Actual Examples

Le Wang

December 2, 2019

Questions behind a Statistical Hypothesis:

1. How is economy? Are we in a recession?
2. Is a job training program effective?
3. etc.

Questions

$\xrightarrow{\text{Step 1: Quantification}}$ **Population Parameter**

$\xrightarrow{\text{Step 2: Conversion}}$ **Statistical Hypothesis (Null vs Alternative Hypotheses)**

$\xrightarrow{\text{Step 3:}}$ **Test Statistic (with a known distribution)**

$\xrightarrow{\text{Step 4:}}$ **Data Collection and Calculation of Test Statistic**

$\xrightarrow{\text{Step 5:}}$ **Evaluate the "likeliness" or "unlikeliness" of the Test Statstic**

$\xrightarrow{\text{Step 6:}}$ **Answer Your Question or Interpret the Results**

Let's work through some examples.

1. One-dimension case
    1.1 Deviation case (Various Z or T-tests)
    1.2 Ratio Case (F-tests)
2. Multi-dimension case (Test of Normality)

Questions: Are we in a normal economy? (non-normal can be in a recession or expansion)

1. **Step 1: Quantification** Wages? Mean (log Hourly) wages (μ) (Population Parameter)

2. **Step 2: Statistical Hypothesis**

$$H_0 : \mu = 2 \quad \text{v.s } H_a : \mu \neq 2$$

1. **Step 1: Quantification** Wages? Mean (log Hourly) wages (μ) (Population Parameter)

2. **Step 2: Statistical Hypothesis**

$$H_0 : \mu = 2 \quad \text{v.s } H_a : \mu \neq 2$$

Problem:

We can collect data and calculate sample averages, $\overline{X}$, which we know would be a good estimator of the population mean, μ. But we also know that even we have the best data/sample,

$$\overline{X} \neq \mu$$

In other words, even if $\mu = 2$, we probably won't have

$$\overline{X} = 2$$

Instead, we may have

$$\overline{X} = 1.5 \quad \text{or } \overline{X} = 4 \quad \text{or } \overline{X} = 5$$

Problem:

We can collect data and calculate sample averages, $\overline{X}$, which we know would be a good estimator of the population mean, μ. But we also know that even we have the best data/sample,

$$\overline{X} \neq \mu$$

In other words, even if $\mu = 2$, we probably won't have

$$\overline{X} = 2$$

Instead, we may have

$$\overline{X} = 1.5 \quad \text{or } \overline{X} = 4 \quad \text{or } \overline{X} = 5$$

Problem:

We can collect data and calculate sample averages, $\overline{X}$, which we know would be a good estimator of the population mean, μ. But we also know that even we have the best data/sample,

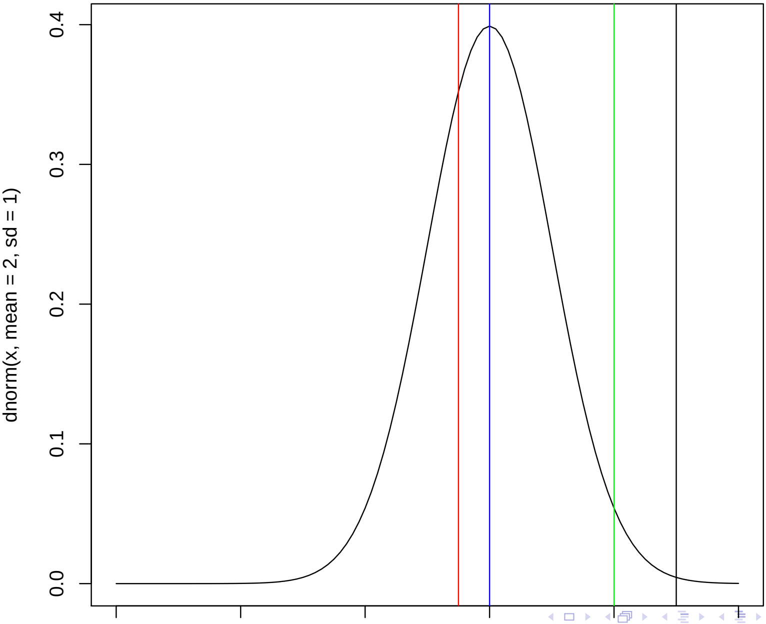
$$\overline{X} \neq \mu$$

In other words, even if $\mu = 2$, we probably won't have

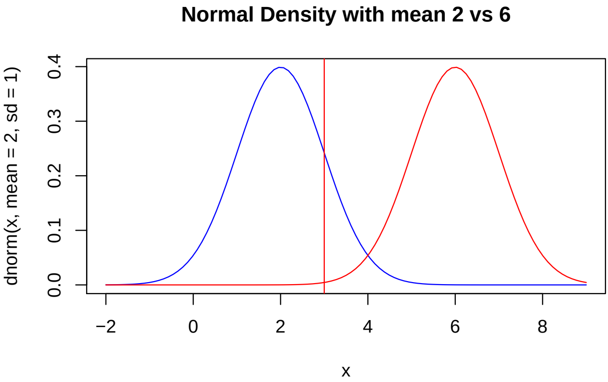
$$\overline{X} = 2$$

Instead, we may have

$$\overline{X} = 1.5 \quad \text{or } \overline{X} = 4 \quad \text{or } \overline{X} = 5$$


dnorm(x, mean = 2, sd = 1)
0.0
0.1
0.2
0.3
0.4

When we observe $\overline{X} = 3$, what would be the truth? It can come from a normal distribution from any mean, just with different probabilities.



I think this is probably one of the more confusing concepts in statistics or econometrics:

We will never know the truth, we can only say that some is more likely than the other.

Then, the question becomes:

If the truth were $\mu = 2$, how likely would we observe $\overline{X} = 3$?

Or,

If the truth were $\mu = 2$, how likely would we observe the discrepancy $\overline{X} - \mu = 1$?

We need to be able to evaluate whether or not the discrepancy is large enough! **The Context**

Step 3: Test Statistic with a known Distribution

$$\overline{X} - \mu$$

From Central Limit Theorem, we know that

$$\begin{aligned}
\overline{X} &\sim N(\mu, \frac{\sigma^2}{N}) \\
\frac{(\overline{X} - \mu)}{\sqrt{\frac{\sigma^2}{N}}} &\sim N(0, 1) \\
\frac{(\overline{X} - \mu)}{\frac{\sigma}{\sqrt{N}}} &\sim N(0, 1)
\end{aligned}$$

From Central Limit Theorem, we know that

$$\begin{aligned} \overline{X} &\sim N(\mu, \frac{\sigma^2}{N}) \\ \frac{(\overline{X} - \mu)}{\sqrt{\frac{\sigma^2}{N}}} &\sim N(0,1) \\ \frac{(\overline{X} - \mu)}{\frac{\sigma}{\sqrt{N}}} &\sim N(0,1) \end{aligned}$$

From Central Limit Theorem, we know that

$$\begin{aligned}
\overline{X} &\sim N(\mu, \frac{\sigma^2}{N}) \\
\frac{(\overline{X} - \mu)}{\sqrt{\frac{\sigma^2}{N}}} &\sim N(0, 1) \\
\frac{(\overline{X} - \mu)}{\frac{\sigma}{\sqrt{N}}} &\sim N(0, 1)
\end{aligned}$$

Step 4: Data Collection and Calculation of Test Statistic
Example:

$$\overline{X} = 2.043414, \quad N = 1,000, \quad s.d. = 4$$

Our hypothesis:

$$H_0 : \mu = 2 \quad H_A : \mu \neq 2$$

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

$$\frac{(2.043414 - 2)}{\frac{4}{\sqrt{1000}}} \sim N(0,1)$$

Step 4: Data Collection and Calculation of Test Statistic
Example:

$$\overline{X} = 2.043414, \quad N = 1,000, \quad s.d. = 4$$

Our hypothesis:

$$H_0 : \mu = 2 \quad H_A : \mu \neq 2$$

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

$$\frac{(2.043414 - 2)}{\frac{4}{\sqrt{1000}}} \sim N(0,1)$$

Step 4: Data Collection and Calculation of Test Statistic
Example:

$$\overline{X} = 2.043414, \quad N = 1,000, \quad s.d. = 4$$

Our hypothesis:

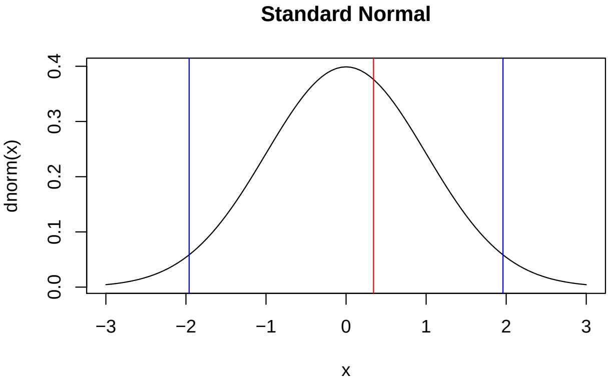
$$H_0 : \mu = 2 \quad H_A : \mu \neq 2$$

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

$$\frac{(2.043414 - 2)}{\frac{4}{\sqrt{1000}}} \sim N(0,1)$$

Suppose that we have a data set of 1000 observations (from a normal distribution with standard deviation of 4)

```
(mean(x)-2)/(4/sqrt(1000))

## [1] 0.3432183
```

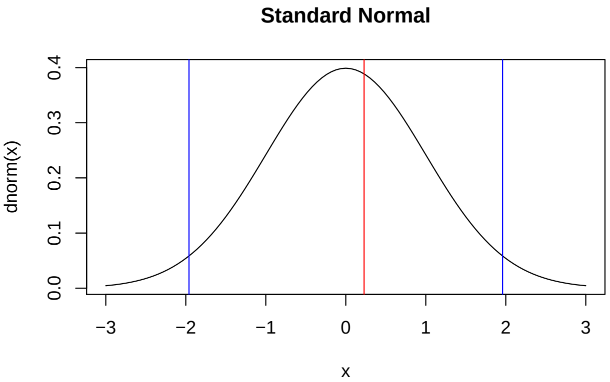


What are the blue lines?

```r
qnorm(.025)

## [1] -1.959964

qnorm(.975)

## [1] 1.959964
```

What if the standard deviation is instead 6?

```
(mean(x)-2)/(6/sqrt(1000))

## [1] 0.2288122
```



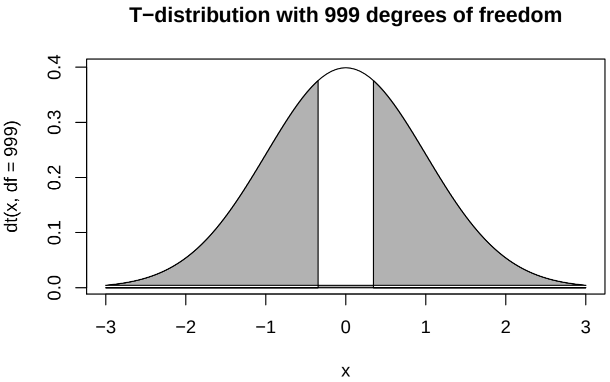
Of course, we never know the standard deviation and it has to be estimated as well. But it's fine!

$$\frac{(\overline{X} - \mu)}{\frac{\widehat{\sigma}}{\sqrt{N}}} \sim t_{N-1}$$

.

```
(mean(x)-2)/(sd(x)/sqrt(1000))

## [1] 0.3460184

qt(.025, df=999)

## [1] -1.962341

qt(.975, df=999)

## [1] 1.962341
```

We fail to reject at 5 percent significance level.



How to calculate this area

```
t<-(mean(x)-2)/(sd(x)/sqrt(1000))
pt(-abs(t),df = 999)*2


## [1] 0.7294018
```

```r
# Conduct t-test using a built-in command
t.test(x,mu=2)
```

```
##
##  One Sample t-test
##
## data:  x
## t = 0.34602, df = 999, p-value = 0.7294
## alternative hypothesis: true mean is not equal to 2
## 95 percent confidence interval:
##  1.797204 2.289624
## sample estimates:
## mean of x
##  2.043414
```

```r
# DIY!
t<-(mean(x)-2)/(sd(x)/sqrt(1000))
# Notice my sign here? Why?
pt(-abs(t),999)*2
```

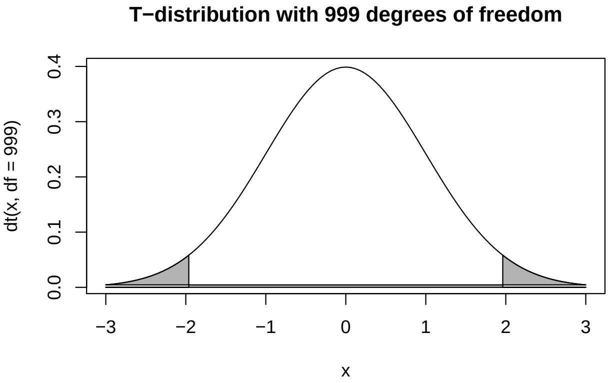
```
## [1] 0.7294018
```

Finally,

Unlikely $\neq$ Cannot be true

So, when we say that it is likely or unlikely, or when we fail to reject or reject the null hypothesis, we are always making some mistakes!

The significance level is exactly the probability that we may be wrong when we are rejecting the null hypothesis when it is indeed true.



Types of Errors

1. Type I error: when the null hypothesis is **true**, we reject it.
2. Type II error: when the null hypothesis is **wrong**, we fail to reject it.

Let's look at many many similar examples related to either means or coefficient estimators

Example 1: Tests of Population Proportion [p.536]

$\Pr[X = x] = \pi$

$\mathbb{E}[1(X = x)] = \pi$

$p = \frac{\sum 1(X_i = x)}{N} \sim N(\pi, \frac{Variance(1(X=x))}{N})$

We can show that $Variance(1(X = x)) = \pi(1 - \pi)$

Test statistic:

$$\frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{N}}} \sim N(0, 1)$$

Example 2: Tests whether a coefficient is different from a number [p.449]

We can show that

$$\widehat{\beta} \sim N(\beta, Variance(\beta))$$

Of course,

$$\frac{\widehat{\beta} - \beta}{\sqrt{Variance(\beta)}} \sim N(0, 1)$$

Example 3: Tests of Equal Means [p.351]

We want to test

$$\mu_1 = \mu_2$$

equilvalent to

$$\mu_1 - \mu_2 = 0$$

We can then test whether

$$\overline{X}_1 - \overline{X}_2 = 0$$

We can show that under certain conditions

$$\overline{X}_1 - \overline{X}_2 \sim N(0, \frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2})$$

Then,

Test statistic:

$$\frac{\overline{X}_1 - \overline{X}_2 - 0}{\sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}}} \sim N(0, 1)$$

If we replace the variances with sample variances, then everything becomes a t-distribution, again!

Case 2: Another form of Deviation

Question: Which country has a larger income inequality, America or China?

Step 1: Quantification – What would measure income inequality? (Population parameter)

Gini coefficient, Entropy measures (in my work), or Variance

Step 2: Statistical Hypothesis

$$H_0 : \sigma_1^2 = \sigma_2^2$$
$$H_a : \sigma_1^2 > \sigma_2^2$$

We can collect data and calculate sample variances, s_1^2 and s_2^2.

We need to compare these two sample variances, and then further compare this comparison to a benchmark value so that we know whether such deviation is large or small.

It is all Hypothesis Testing about!!

We need to compare these two sample variances, and then further compare this comparison to a benchmark value so that we know whether such deviation is large or small.

We need to compare these two sample variances, and then further compare this comparison to a benchmark value so that we know whether such deviation is large or small.

How about $s_1^2 - s_2^2$?

We cannot use it as our test statistic! We do not know the distribution and hence have no value to compare!

We need to compare these two sample variances, and then further compare this comparison to a benchmark value so that we know whether such deviation is large or small.

What should we do!?

We don't know too many things about sample variances so far! but we know one thing: <u>Under certain conditions</u>

$$\frac{(N-1)s^2}{\sigma^2} \sim \chi^2_{N-1}$$

Why is this useful?

$$\frac{\frac{(N_1-1)s_1^2}{\sigma_1^2}/(N_1-1)}{\frac{(N-1)s^2}{\sigma^2}/(N_2-1)} \sim \frac{\frac{\chi^2_{N_1-1}}{N_1-1}}{\frac{\chi^2_{N_2-1}}{N_2-1}} =$$

$$\frac{\frac{s_1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_2^2}} \sim$$

Under the null: $H_0 : \sigma_1^2 = \sigma_2^2$

$$\frac{s_1^2}{s_2^2} \sim$$

$$\frac{\frac{(N_1-1)s_1^2}{\sigma_1^2}/(N_1-1)}{\frac{(N-1)s^2}{\sigma^2}/(N_2-1)} \sim \frac{\frac{\chi^2_{N_1-1}}{N_1-1}}{\frac{\chi^2_{N_2-1}}{N_2-1}} = F_{N_1-1,N_2-1}$$

$$\frac{\frac{s_1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_2^2}} \sim F_{N_1-1,N_2-1}$$

Under the null: $H_0 : \sigma_1^2 = \sigma_2^2$

$$\frac{s_1^2}{s_2^2} \sim F_{N_1-1,N_2-1}$$

$$
\begin{aligned}
\frac{\frac{(N_1-1)s_1^2}{\sigma_1^2}/(N_1-1)}{\frac{(N-1)s^2}{\sigma^2}/(N_2-1)} &\sim \frac{\frac{\chi^2_{N_1-1}}{N_1-1}}{\frac{\chi^2_{N_2-1}}{N_2-1}} = F_{N_1-1,N_2-1} \\
\frac{\frac{s_1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_2^2}} &\sim F_{N_1-1,N_2-1}
\end{aligned}
$$

Under the null: $H_0 : \sigma_1^2 = \sigma_2^2$

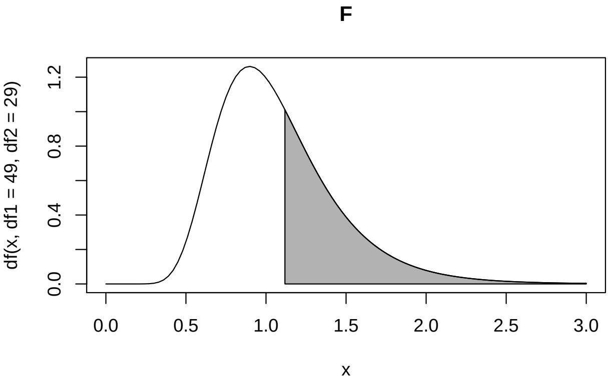
$$
\frac{s_1^2}{s_2^2} \sim F_{N_1-1,N_2-1}
$$

$$\frac{s_1^2}{s_2^2} \sim F_{N_1-1,N_2-1}$$

(We put the larger sample variance in the numerator, so that we only need to deal with the right-tail probability)

```
set.seed(123456)
x <- rnorm(50, mean = 0, sd = 1)
y <- rnorm(30, mean = 1, sd = 1)

# Test Statistic
F<- var(x)/var(y)

# Calculate p-value yourself
1-pf(F,df1=49,df2 = 29)


## [1] 0.3804075
```

Conclusion: p-value is much greater than 5 percent. It is not very extreme. We fail to reject the null hypothesis!


F
df(x, df1 = 49, df2 = 29)
0.0
0.4
0.8
1.2
0.0
0.5
1.0
1.5
2.0
2.5
3.0
x

$$\frac{s_1^2}{s_2^2} \sim F_{N_1-1,N_2-1}$$

(We put the larger sample variance in the numerator, so that we only need to deal with the right-tail probability)

```
# Test Statistic
F<- var(x)/var(y)

# Calculate p-value yourself
1-pf(F,df1=49,df2 = 29)
```

```
## [1] 0.3804075
```

```
var.test(x, y, alternative = "greater")
```

```
##
##  F test to compare two variances
##
## data:  x and y
## F = 1.118, num df = 49, denom df = 29, p-value = 0.3804
## alternative hypothesis: true ratio of variances is greater than 1
## 95 percent confidence interval:
##  0.6290288       Inf
## sample estimates:
## ratio of variances
##           1.118028
```

Command `var.test(x,y,ratio=1,alternative = c("two.sided", "less", "greater"), conf.level = 0.95, ...)`:

1. x,y: Normally distributed data sets
2. ratio: Hypothesized ratio of x/y, default is 1
3. alternative: alternative hypothesis, including "two.sided","greater","less"

Multiple-dimension Tests: Test of Normality

Question: How do we know whether or not my data are drawn from a normal distribution?

What are the features of a normal distribution that stand out to you?

Step 1: What are the features of a normal distribution that stand out to you?

1. Symmetry (Skewness $= 0$)
2. Normal tail (kurtosis $= 3$)

$$\Longrightarrow$$

Step 2: Null hypothesis

$$H_0 : s = 0 \text{ and } k = 3$$

$$H_a : s \neq 0 \text{ or } k \neq 3$$

We can of course calculate sample skewness and kurtosis, but how do I know if they are indeed not different from 0 and 3, respectively?

If we are only interested in testing $s = 0$, what will you do?

If data were drawn from a normal distribution, skewness and excess kurtosis should jointly equal zero, and so does the test statistic!

We have the following facts: (variances are approximiations)

$$
\begin{aligned}
\widehat{s} &\sim N(s, 6/N) \\
\widehat{k} &\sim N(k, 24/N)
\end{aligned}
$$

Jarque-Bera (JB) Test of Normality

$$\frac{(\widehat{s}-0)^2}{6/N}+\frac{(\widehat{k}-3)^2}{24/N}\sim\chi^2_2$$

where N is the number of observations.

Let's simulate a datset from the standard normal

```
set.seed(123456)
x<-rnorm(1000)
```

$$\frac{\widehat{s}^2}{6/N} + \frac{(\widehat{k}-3)^2}{24/N}$$

```
library(moments)
s<-skewness(x)
k<-kurtosis(x)
s
```

```
## [1] -0.09701971
```

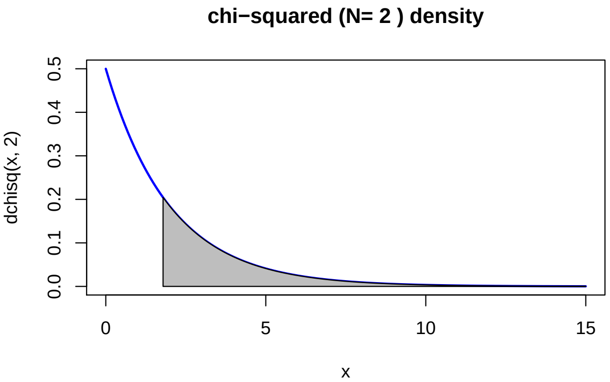
```
k
```

```
## [1] 3.072783
```

```
# Construct the test-statistic ourselves
JBtest <- s^2/(6/1000) + (k-3)^2/(24/1000)
JBtest
```

```
## [1] 1.789525
```

Same old business..... Is 1.7895245 are a large (extreme) number?



How to calculate this number?

```
1-pchisq(JBtest,df=2)

## [1] 0.4087048
```

There are 40.8704754 percent of the values greater than the JB test statistic

$\Longrightarrow$ Our test statistic is relatively small

$\Longrightarrow$ Deviations from skewness $= 0$ and kurtosis $= 3$ are small

$\Longrightarrow$ so is the deviation from normality.

$\Longrightarrow$ We thus fail to reject the normality!