THE UNIVERSITY OF MEMPHIS

COMP7/8150: Fundamentals of Data Science

Week 4: Statistical Inference / Hypothesis Testing

Max H. Garzon

Standard Methodology

- **Problem Definition/goal**
  - Identify/specify goals of the data analysis
  - commit to specific deliverables
- **Data pre-processing**
  - Identify appropriate data
  - Acquire data (gather, lookup, understand)
- **Data processing**
  - Identify methods (gather, cleanse, store)
  - Carry out the analysis (patterns, trends, predictions?)
- **Data post-processing**
  - Visualize and present
  - Deploy and evaluate. Iterate, if necessary

Learning Objectives

- To identify the concept of a "hypothesis Test" for making inferences and pre-conditions to run it
- To identify the computations needed to run a h-test when making inferences with both large (normal distribution) and small sample sizes (student-t distribution)
- To identify commands in a programming environments to run it

Learning Objectives

- To identify and contrast the concept *hypothesis testing* as a statistical decision making tool.
- To identify and characterize the statistical procedures used for hypothesis testing.
- To identify the concept of “p-value” as a tool to assess the (quality of the) decision and how to calculate it.
- To identify the test “chi-squared” as a tool to measure goodness-of-fit.

CI and tests of hypothesis

- In the previous module, we are mainly interested in the value of a parameter (interval) estimation.
- Often times, these parameter are calculated in order to make decisions between two competing hypotheses to make a decision.
- Can make this inference directly with a hypothesis test.

Hypothesis Testing (HT): Example 1

- Need decide between two possibilities:
  is Cassandra "clairvoyant" or not? (seems to be)
  - She can tell any card drawn at random (1/52) and shown the reverse of it
  - She cannot.

- Can show C: 25 cards and see if she can tell them
  Some successes/failures are not sufficient evidence she is/not. Where to draw the line?

- A decision must be made: yes or no?

.. HT: Example 3

- A steel company is concerned that the mean strength μ of their steel produced meet the minimum government standards. They need to decide between two possibilities:

  – The mean strength μ does not meet the minimum standard.

  – The mean strength μ exceeds the minimum standard.

Parts of a Statistical Test

http://www.youtube.com/watch?v=abjHpJ36pIE&feature=related

- **Alternative [or Research] hypothesis,**
  H_a: A claim (statement) to be decided upon based on some evidence (data)
- **Null hypothesis**
  H_0: The opposite claim (default statement to be assumed to be true until proven otherwise. It is usually stated first.)
- **Test statistic and its *p*-value**
  A single statistic calculated from the sample that can be used to reject or not reject H_0.

Decision of a statistical test

- **Rejection region**
  a region and a rule are used to decide, depending on where the value of the test statistic falls, whether the null hypothesis H_0 should be rejected.
- **Conclusion**
  Either "Reject H_0" or "Failed to reject H_0", with a pre-specified significance level.
- Usually, the significance level is set at
  $\alpha = .01$ or $\alpha = .05$.

.. Decision of by p-value

- *p*-value
  It is defined as the *probability of observing, just by chance, a test statistic as extreme or even more extreme than what we've actually observed*.
  If H_0 is rejected, the p-value is the actual probability that we have made an incorrect decision.
- If the p-value is smaller than the **preassigned significance level, α,** then H_0 is rejected.

p-value and its implications

http://www.youtube.com/watch?v=ZFXy_UdlQJg&feature=related

- A (very) small p-value is a (strong) indication that the null hypothesis (H_0) is unlikely to be true.
- For a small p-value, if H_0 were true, it is very unlikely one should observe such extreme events. Hence, the only reasonable explanation left is that H_0 is not true.
- There are other ways to do a t-test.

p-value and its computation

- To compute the p-value, we need to know
  - its alternative hypothesis (one-sided or two-sided alternative test)
  - sampling distribution about its test statistic (z, t, …) (under H_0)

  Like for the estimation problem, its sample distribution is known in many cases.

Common procedures for test statistics

http://www.youtube.com/watch?v=abjHpJ36pIE&feature=related

- Let θ be the parameter of interest in a statistical model (such as the population mean). Under the null hypothesis H_0 , depending on the sample size, we can use either z (large n) or t (small n) statistics below:

$$z = \frac{\hat{\theta} - \theta_0}{SE(\theta_0)}$$

$$t = \frac{\hat{\theta} - \theta_0}{SE(\theta_0)}$$

- Since a small sample t-test is a more conservative test than a z-test, we will consider only the t-test.

Summary on hypothesis testing for large sample size

Summary on hypothesis testing (large sample)

Parameter	Test Statistics
μ	$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$
π	$z = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$
$\mu_1 - \mu_2$	$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$
$\pi_1 - \pi_2$	$z = \frac{\hat{\pi}_1 - \hat{\pi}_2}{\sqrt{\hat{\pi}(1-\hat{\pi})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

Summary of small sample tests

http://www.youtube.com/watch?v=JlfLnx8sh-o

Summary on hypothesis testing (small sample)

Parameter	Test Statistics	Degrees of freedom
μ	$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$	$(n-1)$
$\mu_1 - \mu_2$ (equal variances)	$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$(\mathrm{n}_1 + n_2 - 2)$
$\mu_1 - \mu_2$ (unequal variances)	$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	(hard to find)
$\mu_1 - \mu_2$ (paired)	$t = \frac{\bar{d} - D_0}{s_d/\sqrt{n}}$	$(n-1)$

Polled variance $s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}$

Function t.test() in R

We can use t.test to perform a variety of t-tests (for small sample) in R but there is no z.test.

- **Description**

  Performs one and two sample t-tests on vectors of data.

- **Usage**

- ```
  t.test(x, y = NULL, alternative = c("two.sided", "less", "greater"), mu = 0, paired = FALSE, var.equal = FALSE, conf.level = 0.95, ...)
  ```
  ```

Case 3HT: Testing the difference $\mu_1 - \mu_2$

- For two quantitative populations with unknown means μ_1, μ_2 and standard deviations σ_1 and σ_2 .
- We take a random sample of sizes n_1 and n_2 from the two populations, compute their sample means $\bar{x}_1, \bar{x}_2$ and sample standard deviations s_1, s_2 .
- We are interested in hypothesis testing about the difference $\mu_1 - \mu_2$ when the sample sizes are small.

.. Case 3HT: Testing the difference $\mu_1 - \mu_2$

- To test $H_0: \mu_1 - \mu_2 = D_0$, where D_0 is a constant, usually 0.
- If we cannot assume equal variance assumption, the test statistic used is the same one used for large sample inference.
- However, hard to find the appropriate number d.f.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

.. Case 3HT: Testing the difference $\mu_1 - \mu_2$

- Under the additional assumption of equal variance, we first compute "pooled variance"

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

- We then compute that the test statistic $t = \dfrac{\bar{x}_1 - \bar{x}_2 - D_0}{\sqrt{s^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$

has a *t* distribution with certain degrees of freedom.

One-sample vs two-sample inference

- For certain designs of the experiment, assumption of independent samples is intentionally violated.

- For example, for a matched-pairs design (e.g. before vs. after, twins), it is unreasonable to assume independence between “two samples”.

One-sample inference

Paired-difference test.

- Paired experiment can eliminate unwanted variability in the experiment.
- We can analyze only the differences,

$$d_i = x_{1i} - x_{2i}$$

to see if there is a difference in the two population means, $\mu_1 - \mu_2$.

The Paired-Difference Test

One sample pair t-test

To test the hypothesis for $H_0 : \mu_1 - \mu_2 = 0$ for paired sample, we test $H_0 : \mu_d = \mu_1 - \mu_2 = 0$ and we use the test statistic

$$t = \frac{\bar{d}}{s_d/\sqrt{n}},$$

where

- n is the number of pairs
- $\bar{d}$ is the sample mean of the difference
- s_d is the sample s.d. of the difference

Hence, we can perform the test statistic using $t(df = n - 1)$ distribution

Example of paired design

- One Type A and one Type B tire are randomly assigned to each of the rear wheels of several cars.
- The pairs of responses are not independent because measurements are taken on the same car.
- We like to compare the average tire wear for types A and B using a test of hypothesis.

$$
\begin{aligned}
\mathrm{H}_0 &: \mu_1 - \mu_2 = 0 \\
\mathrm{H}_a &: \mu_1 - \mu_2 \neq 0
\end{aligned}
$$

Hypothesis testing in R

- **Description**

  Performs one and two sample t-tests on vectors of data.

  ```
  t.test(x, y = NULL, alternative =
  c("two.sided", "less", "greater"),
  mu = 0, paired = FALSE, var.equal =
  FALSE, conf.level = 0.95, ...)
  ```

- For the type A&B tire test, you would run:

  ```
  > t.test(x, y, alternative =
  "two.sided“, paired = TRUE)
  ```

Statistical Inference about variance(s)

http://www.youtube.com/watch?v=FJ4jkCpz_Wc

- If the primary parameter of interest is the population variance σ^2, the test statistic to be constructed will rather follow a Chi-Square distribution.

- If the primary parameter of interest is the equality of two population variances σ_1^2 and σ_2^2 the test statistic to be constructed will rather follow an F distribution.

Chi-Square distribution and F-distribution

- Both Chi-Square distribution and F-distribution have important applications in other areas of Statistics.
- Chi-Square distribution will be used in contingency tables.
- F- distribution will be used in ANOVA (Analysis of Variance) table.
- We will study their distributions after we describe the corresponding examples.

Inference Concerning a Population Variance

- If the primary parameter of interest is the population variance σ^2, we choose a random sample of size *n* from a normal distribution.
- The sample variance s^2 can be used in its standardized form:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

which has a Chi-Square distribution with *n* - 1 degrees of freedom.

Inference Concerning Two Population Variances

- We can make inferences about the ratio of two population variances σ_1^2 / σ_2^2 .
- Two independent random samples of size n_1 and n_2 from normal distributions.
- If the two population variances are equal, the statistic

$$F = \frac{s_1^2}{s_2^2}$$

has an F distribution with $df_1 = n_1 - 1$ and $df_2 = n_2 - 1$ degrees of freedom.

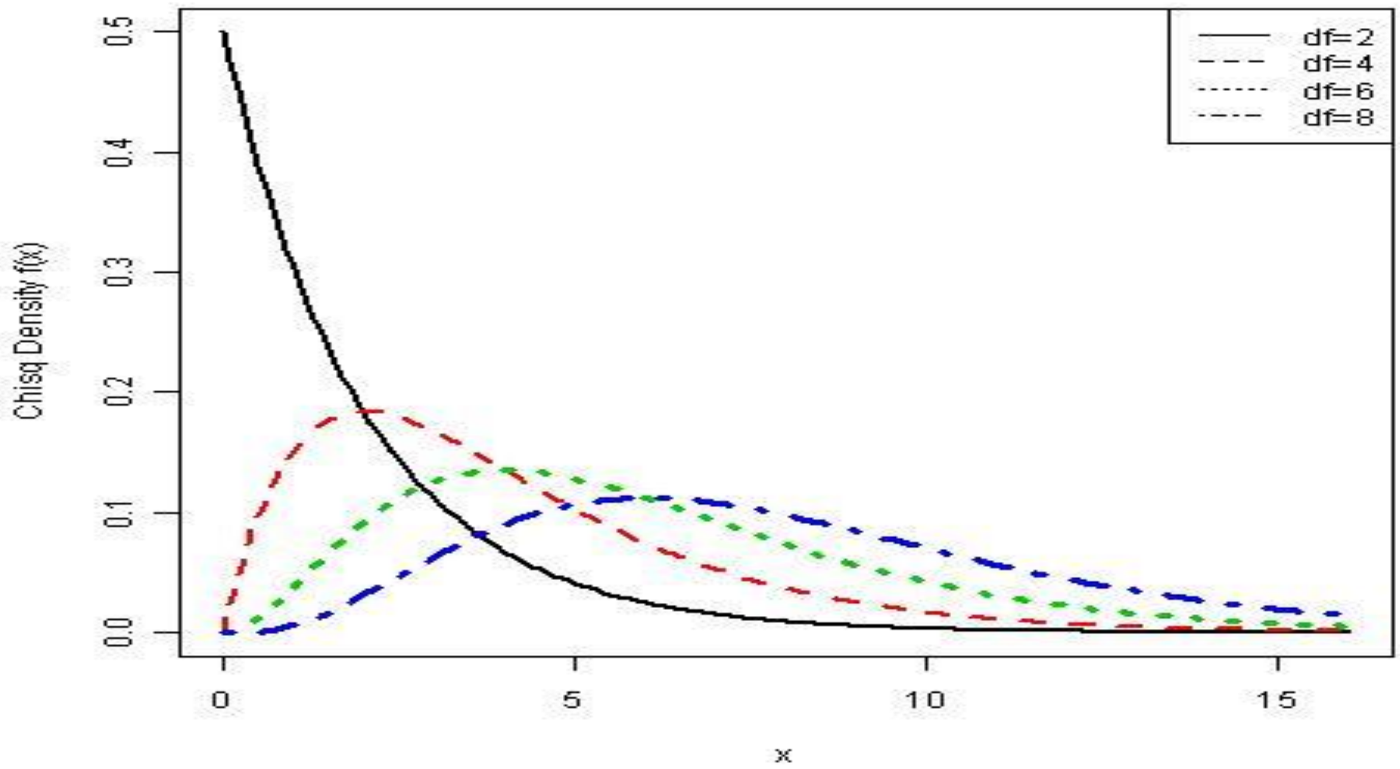
Chi-squared distribution

χ^2 distribution

$X \sim \chi^2(v)$ if and only if $X \sim Gamma(v/2, 2)$.

1. In R: df=v, ncp=0 (default) distribution function=$chisq$.
2. $E(X) = v$.
3. $Var(X) = 2v$

Plot of chisq densities



Chi square in R

```
x <- c(0:160)*0.1
y1 <- dchisq(x,2)
y2 <- dchisq(x,4)
y3 <- dchisq(x,6)
y4 <- dchisq(x,8)
y <- cbind(y1,y2,y3,y4)
matplot(x, y, type="l", ylab="Chisq Density f(x)", lwd=c(2.5,2.5,3.5,3.5))
legend("topright", c("df=2","df=4","df=6","df=8"), lty=c(1,2,3,4))
```

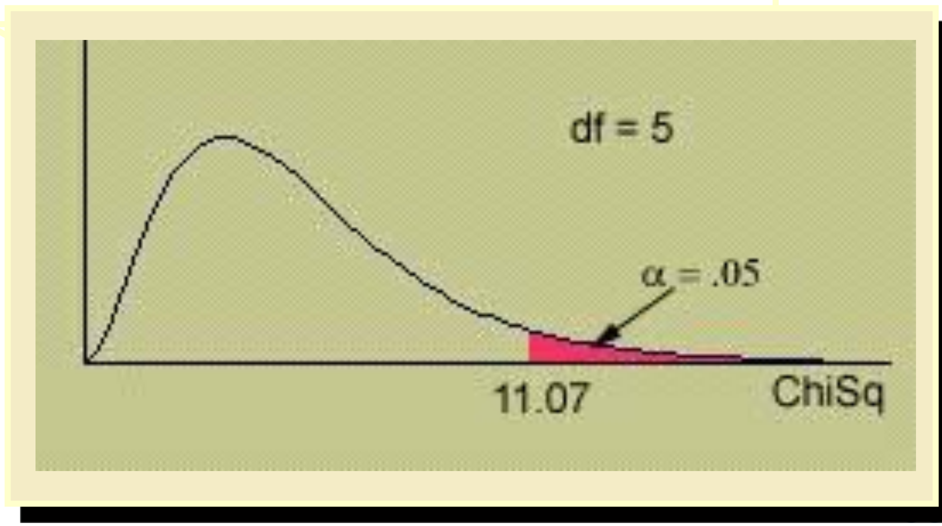
Percentiles of the chi-square distribution in R

- `> options(digits=4)`
- `> p <- c(0.01, 0.025, 0.05, 0.1, 0.9, 0.95, 0.975, 0.99)`
- `> qchisq(p, 5)`

```
[1]  0.5543  0.8312  1.1455  1.6103  9.2364 11.0705
12.8325 15.0863
```

- `> qchisq(p, 10)`

```
[1]  2.558  3.247  3.940  4.865 15.987 18.307 20.483
23.209
```



Interval Estimation of a Population Variance

To construct confidence interval for σ^2, we use the fact that

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

Let A and B be the lower and upper $\alpha/2$ percentiles of χ^2_{n-1} distribution.

The confidence interval for σ^2 is

$$\frac{(n-1)s^2}{B} < \sigma^2 < \frac{(n-1)s^2}{A}.$$

Hypothesis Testing about Pop Variance

To test the hypothesis for $H_0 : \sigma^2 = \sigma_0^2$, we use the fact that

$$\frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2_{n-1}.$$

Hence, we can use the test statistic using χ^2 distribution

$$\frac{(n-1)s^2}{\sigma_0^2}.$$

F-distribution

F distribution

$$Y = \frac{X_1/v_1}{X_2/v_2},$$

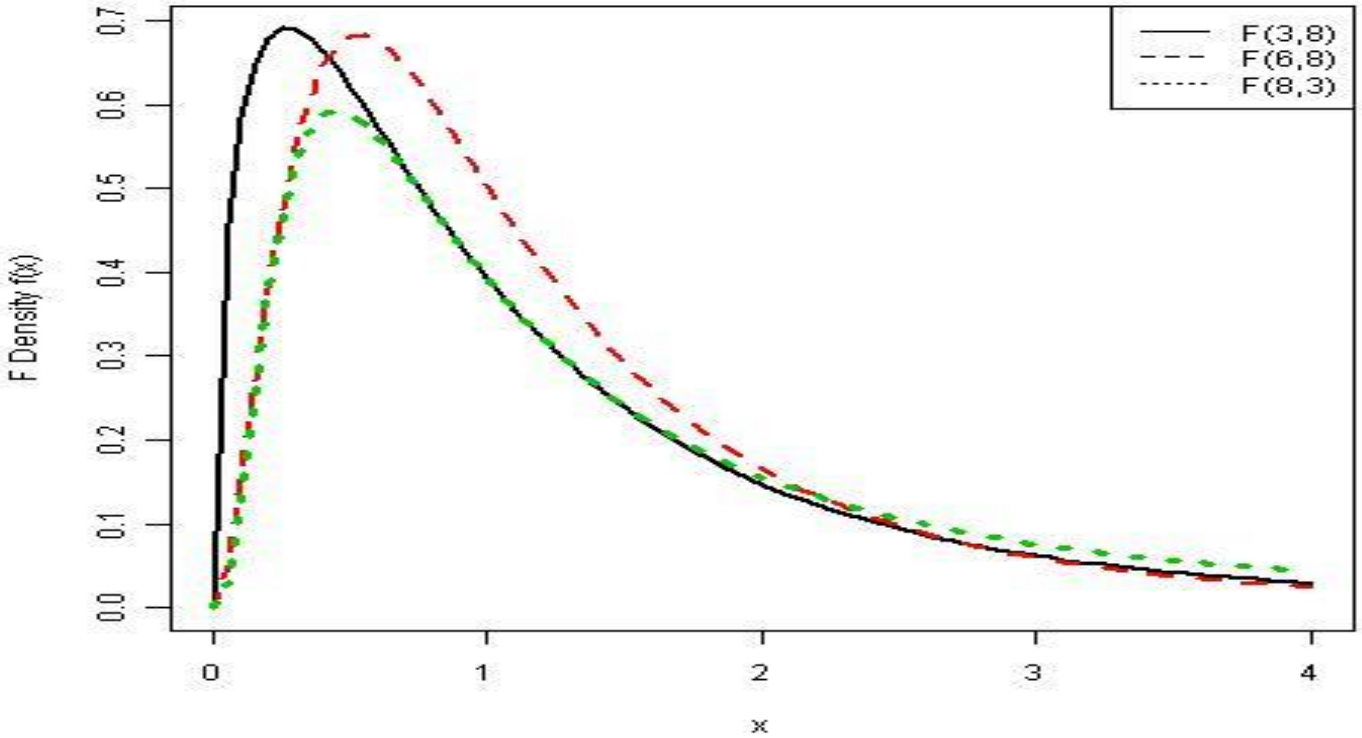
where $X_1 \sim \chi^2(v_1)$ independent of $X_2 \sim \chi^2(v_2)$.

1. In R: df1=v_1, df2=v_2, ncp=0 (default), distribution function=f.
2. $E(X) = \frac{v_2}{v_2-2}$. (when $v_2 > 2$).
3. $Var(X) = 2\left(\frac{v_2}{v_2-2}\right)^2 \frac{v_1+v_2-2}{v_1(v_2-4)}$ (when $v_2 > 4$).

Connection with t distribution:

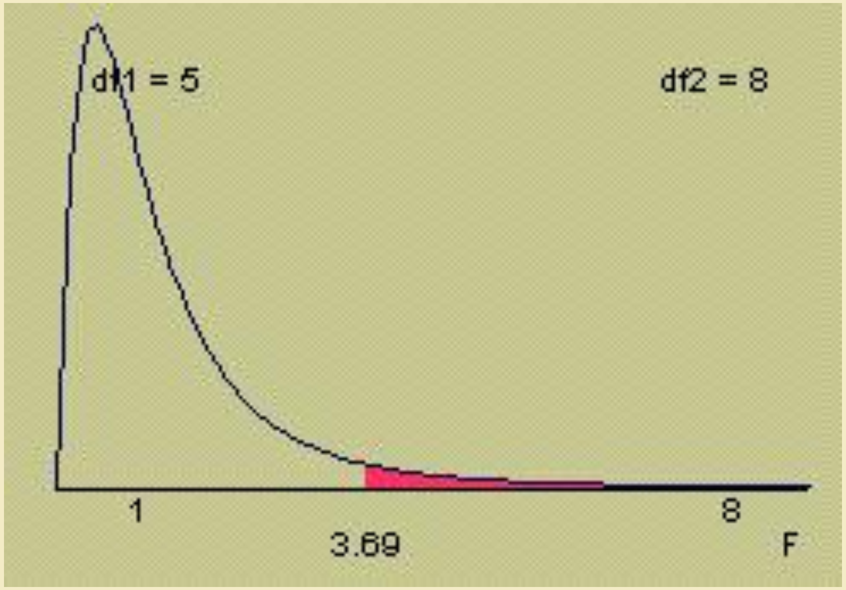
If $X \sim t(v)$, then $Y = X^2 \sim F(1, v)$. (why ?)

Plot of various F-distributions



Percentiles of an F-distribution

```
> options(digits=2)> p <- c(0.01, 0.025, 0.05, 0.1, 0.9,
> .95, 0.975, 0.99)
> qf(p, 5, 8)
  [1] 0.097 0.148 0.208 0.299 2.726 3.687 4.817 6.632
```



For example, the value of F that cuts off .05 in the upper tail of the distribution with $df_1 = 5$ and $df_2 = 8$ is F =3.69.

Confidence Interval for Ratio of Two Population Variances

To construct confidence interval for σ_1^2/σ_2^2, we use the fact that

$$\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim F_{n_1-1,n_2-1}.$$

Let A and B be the lower and upper $\alpha/2$ percentiles of F_{n_1-1,n_2-1} distribution.

The confidence interval for σ_1^2/σ_2^2 is

$$\frac{1}{B}\frac{s_1^2}{s_2^2} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{1}{A}\frac{s_1^2}{s_2^2}.$$

Hypothesis Testing for Ratio of Two Population Variances

To test the hypothesis for $H_0 : \sigma_1^2 = \sigma_2^2$, we use the fact that

$$F = \frac{s_1^2}{s_2^2} \sim F_{n_1-1, n_2-1}.$$

Hence, we can use the test statistic using F distribution

$$F = \frac{s_1^2}{s_2^2}.$$

Summary of small sample tests

Parameter	Test Statistic	Degrees of Freedom
μ	$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$	$n - 1$
$\mu_1 - \mu_2$ (equal variances)	$t = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$	$n_1 + n_2 - 2$
$\mu_1 - \mu_2$ (unequal variances)	$t \approx \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$	
$\mu_1 - \mu_2$ (paired samples)	$t = \dfrac{\bar{d} - \mu_d}{s_d/\sqrt{n}}$	$n - 1$
σ^2	$\chi^2 = \dfrac{(n-1)s^2}{\sigma_0^2}$	$n - 1$
σ_1^2/σ_2^2	$F = s_1^2/s_2^2$	$n_1 - 1$ and $n_2 - 1$

Multinomial Experiment

- Let us start with a simple experiment of drawing 100 random digits (0,1,….,9). Suppose that the count for all digits are summarized as:

0	1	2	3	4	5	6	7	8	9
13	16	9	10	9	7	8	12	7	9

- In R
  - x <- trunc(10 * runif(100))
  - table(x)
- Here’s an interesting problem (to be solved later):

  Test the hypothesis that all digits

.. Multinomial Experiment

- The experiment has ***n* identical trials.**
- Each trial results in **one of *k* categories.**
- The probability of falling into category i on a single trial is p_i and **remains constant**.

$$\boldsymbol{p_1 + p_2 + \ldots + p_k = 1}.$$

- The trials are **independent**.

.. Multinomial Experiment

- Observation from the experiment: the number of outcomes in each category, $\boldsymbol{O_1, O_2, \ldots O_k}$ with $\boldsymbol{O_1 + O_2 + \ldots + O_k = n}$.
- We are interested in testing whether or not the data, $\boldsymbol{O_1, O_2, \ldots O_k}$ is consistent with the hypothesis

$$H_0: p_i = p_{i0} \text{ for } i=1,2,\ldots,k.$$

- A popular test statistic is the chi-squared statistic (or goodness-of-fit)

Pearson's Chi-Square Statistic

- We want to use sample information to test if the values of the p_i's are equal to some specified values c_i, i.e., $p_0 = c_0$, … $p_9 = c_9$.
  - for example whether they are equally likely, i.e., $p_i = c_i = 1/k$.
- The **expected number** of times that outcome *i* will occur is $\boldsymbol{E_i = n}c_i$.
- If the **observed cell counts, $\boldsymbol{O_i}$,** are too far from what we hypothesized under H_0, it is more likely that H_0 should be rejected.

.. Pearson's Chi-Square Statistic

- Test statistics: Pearson's chi-square statistic:

$$X^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

- If H_0 is true, the differences $|O\text{-}E|$ will be small, but large when H_0 is false.
- Reject H_0 for large values of X^2 using the chi-square distribution with degrees of freedom is k-1.

The Goodness of Fit Test

- A single categorical variable is measured, and exact numerical values are specified for each of the cell probability $p_i = p_{i0}$
- the expected cell counts are $E_i = np_{i0}$
- Degrees of freedom: $df = k-1$

$$\text{Test statistic}: X^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Chi-square test in R

Pearson's Chi-squared Test for Count Data

- **Description**
- chisq.test performs chi-squared contingency table tests and goodness-of-fit tests.
- **Usage**
- chisq.test(x, ...)
- #default parameter for $p_i=1/k$
- #read pages 114-118 for some examples.

..Chi-square test in R

- Returning to our previous experiment using R:

```
> table(x)
x
 0   1   2   3   4   5   6   7   8   9
13  16   9  10   9   7   8  12   7   9
> chisq.test(table(x)) # NOT 'chisq.test(x)'!
```

- ```
   Chi-squared test for given probabilities
  data:  table(x) X-squared = 7.4, df = 9, p-value
= 0.5955
  ```
  ```

.. Chi-square test in R

- `X-squared = 7.4, df = 9, p-value = 0.5955`
- Since the p-value is > 0.05, the data failed to provide evidence to reject H_0.
- We will verify the test via direct computation in R:

```
> Oi <- table(x); Ei <- 100*0.1;
> sum((Oi-Ei)^2/Ei)
[1] 7.4
> 1-pchisq(7.4,9)
[1] 0.5955485
```

Contingency Tables: A Two-Way Classification

- Consider **two qualitative variables** (usually one as explanatory variable X, and other as response variable Y) to cross-classify the sampled data.
  - X:treatment (Y/N) and Y:disease outcome(Y/N)
  - X: gender(F/M) and Y:voting preference(D/R)
  - X: seatbelt use(Y/N) and Y:fatality(Y/N)
- Summarize the data by counting the **observed** number of outcomes in each of the intersections of category levels in a **contingency table.**

2x2 contingency table example

- We want to study the relationship (effect), if any, between seatbelt use (Y/N) and fatality (Y/N) in a traffic accident.
- *200 traffic reports are classified a the following 2x2 table :*

	Fatality(N)	Fatality(Y)
Seat belt(N)	20	20
Seat belt(Y)	75	25

.. 2x2 contingency table example

- The goal of the previous study is whether the seat belt use (Y/N) can affect the outcome of fatality (Y/N).
- One way to analyze this is to compute the fatality rates among subpopulation not using seat belt (0.50=20/40) and those using seat belt (0.25=25/100). Then treat it as the problem of comparing two binomial populations.
- We can use contingency table to analyze the data with two (or more) possible outcomes.

	Fatality(N)	Fatality(Y)
Seat belt(N)	20	20
Seat belt(Y)	75	25

Example simulated in R

```
> # both x and y are generated from binomial distn with the same p's
> x <- rbinom(40, 1, 0.5)
> y <- rbinom(100, 1, 0.5)
> Tx <- table(x)
> Ty <- table(y)
> Txy <- rbind(Tx,Ty); Txy
    0  1
Tx 23 17
Ty 51 49
> chisq.test(Txy)
        Pearson's Chi-squared test with Yates' continuity correction
data:  Txy
X-squared = 0.2587, df = 1, p-value = 0.611
```

can't reject H_0

Another simulation example

- > #different binomial distribution w/ different p's

```
> x <- rbinom(40, 1, 0.5)
> y <- rbinom(100, 1, 0.75)
> Tx <- table(x)
> Ty <- table(y)
> Txy <- rbind(Tx,Ty); Txy
    0  1
Tx 20 20
Ty 27 73
> chisq.test(Txy)
        Pearson's Chi-squared test with Yates' continuity correction
data:  Txy
X-squared = 5.7853, df = 1, p-value = 0.01616
```

reject H_0

Example for comparing two multinomial experiments

- Consider three simulation samples (x,y,z) on generating random digits (0,1,…,9) with the following R codes:

```
> x <- trunc(10*runif(100))
> y <- trunc(10*runif(200))
> z <- trunc(10*runif(150))
> Tx <- table(x)
> Ty <- table(y)
> Tz <- table(z)
> Txyz <- rbind(Tx,Ty, Tz); Txyz
    0  1  2  3  4  5  6  7  8  9
Tx 11 18  9 10  9  9  9  8  7 10
Ty 17 23 19 16 18 20 23 31 15 18
Tz 19 17 17 16 15 11  9 16 13 17
```

.. Example for comparing two multinomial experiments

Suppose that we are given another data:

- Txyz

	0	1	2	3	4	5	6	7	8	9
Tx	11	9	12	7	11	11	11	9	11	8
Ty	26	16	22	12	24	27	20	16	22	15
Tz	12	17	16	11	20	14	16	14	13	17

- Q: Can we test the hypothesis that three samples (Tx, Ty, Tz) have the same method to produce random digits ?
  - we can run chisq.test(Txyz) in R as before.

I x *J* Contingency Table

- It has *I* rows and *J* columns: *IxJ* total cells.
- We would like to see the relationship between the two classification variables.

	1	2	…	*J*
1	O_{11}	O_{12}	…	O_{1J}
2	O_{21}	O_{22}	…	O_{2J}
…	…	…	…	….
I	O_{I1}	O_{I2}	…	O_{IJ}

Mechanics of Chi-Square Test of Independence

H_0: classification variables are independent
H_a: classification variables are dependent

- Observed cell counts are O_{ij} for row i and column j.
- Expected cell counts are $E_{ij} = np_{ij}$
  - ✓Under $H_0 : p_{ij} = p_i p_j$ and we can use the same chi-square test for goodness-of-fit test.

$$X^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Other topics

- We will delay the discussion on ANOVA model which is given in a later module.
- The following topics will not be discussed. You can read the textbook for the method, application and procedure.
  - Multiple comparison (page 122)
  - Response curves (page 125)
  - Data with nested structure (page 127)
  - Re-sampling methods (page 128)

Questions?