

In the name of the absolute power and the absolute knowledge

1



Social and Cognitive Robotics

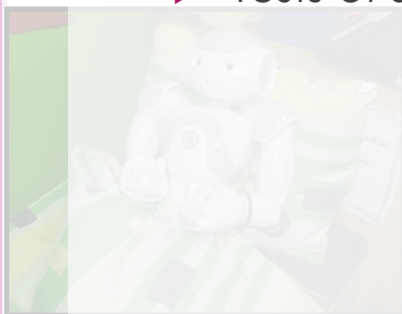
Chapter 4: Design human-robot interactions and evaluation of the interventions

Alireza Taheri, PhD, Assistant Professor
artaheri@sharif.edu

Head of the Social and Cognitive Robotics Lab.,
Center of Excellence in Design, Robotics, and Automation
School of Mechanical Engineering,
Sharif University of Technology, Tehran, Iran

Lecture Schedule

- ▶ Introduction, definitions, and basic concepts
- ▶ **Chapter 4: Design human-robot interactions and evaluation of the interventions**
- ▶ Design interaction perspective of social robots
- ▶ Artificial intelligence
- ▶ The necessity of this chapter
- ▶ Planning, Designing, Executing HRI Study
- ▶ Structure of papers for HRI studies
- ▶ Design human-robot interactions and evaluation of the interventions
- ▶ Tests of Statistical Significance



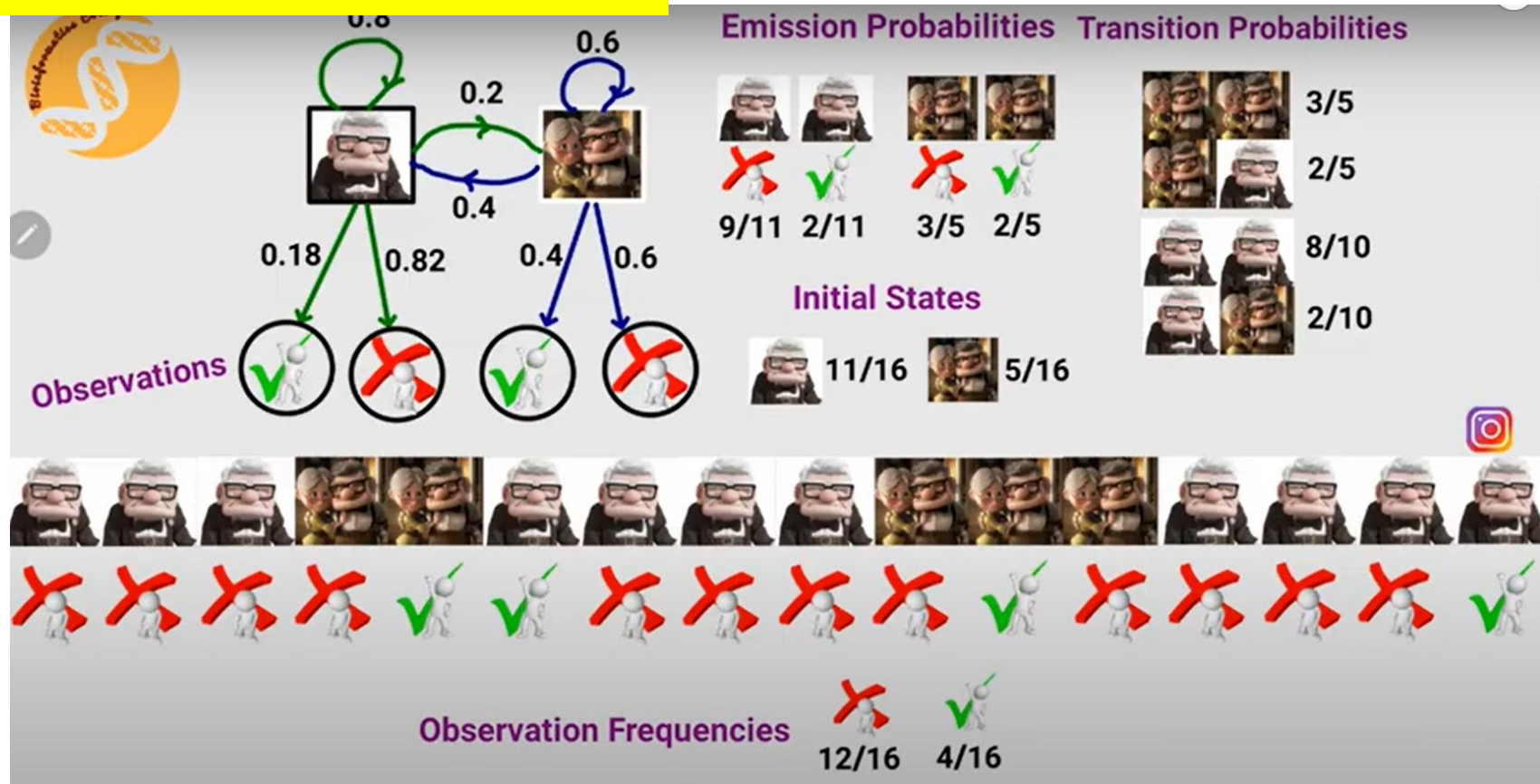
With Special Thanks to

- ▶ Kerstin Dautenhahn
- ▶ Saeed Khodaygan



Previously ...

مثالی از نحوه مناسبه پارامترهای مدل مففی مارکوف با داشتن



The necessity of this chapter



Our Approach in Social Robotics

پناه صنعتی شریف

مدرسی n باره بر مهم ترین پیام این درس ...

شناسایی یک نیاز در جامعه (عمدتاً برای کودکان)

انجام مطالعات میدانی و
مداخلات بالینی
در حوزه
بررسی مقبولیت،
آموزش،
توانبخشی و ...

Clinical Interventions

پیاده سازی
الگوریتم های کنترلی و
هوش مصنوعی
جهت توانمندسازی ربات
ها
برای نیل به اهداف تعیین
شده

Control and Machine Learning

خرید،
و یا
طراحی و ساخت
ربات های اجتماعی-
شناختی
با توجه به نیازهای مورد
نظر

Design and Fabrication



به بهانه کنفرانس بین المللی رباتیک اجتماعی (ICSR)



<https://www.icsr2022.it/>

Planning, Designing, Executing HRI Study



Planning a study

- ▶ What is your **main research question**?
- ▶ How many **participants** do I need?
 - ▶ What **type** of study will be conducted (**within-subjects, between-subjects, mixed-model, etc.**)?
 - ▶ How many **groups** will be in the study?
 - ▶ How many participants **per group** will be required?
 - ▶ How will participants be **recruited** for the study?
- ▶ What **type** of **robots** will be used?
 - ▶ **How many** different types of robots will be used?
- ▶ What types of **tasks** will the participants perform or observe (**Study Protocol**)?
- ▶ What methods of **evaluation** will be used?
 - ▶ self-assessments, behavioral measures, interviews, psychophysiological measures, task performance, etc.?



Planning a study

- ▶ What type of **equipment** will be needed?
 - ▶ computers, measurement equipment, recording devices, etc.?
 - ▶ What type of **environment** and space is required to conduct the study?
 - ▶ field, laboratory, virtual, etc.
 - ▶ How will **contingencies** and failures be **handled**?
 - ▶ What type of **assistance** is needed to conduct the study?
- ▶ So ... ready to **start**, now??
- ▶ **Not yet** ...!!! You need to ...
- ▶ Pass Human Subject Protection Training
 - ▶ Get research permission



Planning a study

- ▶ Do not forget
 - ▶ Mandatory training tutorial Human research ethics
 - ▶ Submit ethics application
 - ▶ Ethics application focusses primarily on experimental protocol (what happens to the participants, how you recruit them, etc.)
 - ▶ Only after approval, you are allowed to recruit participants
 - ▶ While waiting for approval, you can still work on the technical implementation details of the study



Human Subject Protection Training

COLLABORATIVE INSTITUTIONAL TRAINING INITIATIVE (CITI PROGRAM) COURSEWORK REQUIREMENTS REPORT*

* NOTE: Scores on this Requirements Report reflect quiz completions at the time all requirements for the course were met. See list below for details. See separate Transcript Report for more recent quiz scores, including those on optional (supplemental) course elements.

- **Name:** Alireza Taheri (ID: 4920566)
- **Email:** Alireza.Taheri@du.edu
- **Institution Affiliation:** University of Denver (ID: 1294)
- **Institution Unit:** Electrical and Computer Engineering (CEC)
- **Phone:** 7203528128

- **Curriculum Group:** Responsible Conduct of Research Course
- **Course Learner Group:** RCR FOR Social, Behavioral, Educational, and Humanities
- **Stage:** Stage 1 - RCR
- **Description:** This course is for investigators, staff and students with an interest or focus in Social and Behavioral research. This course contains text, embedded case studies AND quizzes.

- **Report ID:** 16573542
- **Completion Date:** 07/12/2015
- **Expiration Date:** 07/11/2019
- **Minimum Passing:** 80
- **Reported Score*:** 100

REQUIRED AND ELECTIVE MODULES ONLY	DATE COMPLETED	SCORE
Responsible Conduct of Research (RCR) Course Introduction (ID:1522)	07/06/15	No Quiz
Research Misconduct (RCR-Basic) (ID:16604)	07/09/15	5/5 (100%)
Data Management (RCR-Basic) (ID:16600)	07/09/15	5/5 (100%)
Authorship (RCR-Basic) (ID:16597)	07/10/15	5/5 (100%)
Peer Review (RCR-Basic) (ID:16603)	07/11/15	5/5 (100%)
Mentoring (RCR-Basic) (ID:16602)	07/11/15	5/5 (100%)
Conflicts of Interest (RCR-Basic) (ID:16599)	07/12/15	5/5 (100%)
Collaborative Research (RCR-Basic) (ID:16598)	07/12/15	5/5 (100%)
Research Involving Human Subjects (RCR-Basic) (ID:13566)	07/12/15	5/5 (100%)
Responsible Conduct of Research (RCR) Course Conclusion (ID:1043)	07/12/15	No Quiz
University of Denver (ID:12728)	07/12/15	No Quiz

For this Report to be valid, the learner identified above must have had a valid affiliation with the CITI Program subscribing institution identified above or have been a paid Independent Learner.

CITI Program
Email: citisupport@miami.edu
Phone: 305-243-7970
Web: <https://www.citiprogram.org>

10/25/2016

Record Of Completion



Yale University

Human Subject Protection Training - Yale web-based training program Training Certificate of Completion

This is to certify that

Alireza Taheri

*has completed the Yale University Training in Human Subject Protections on
10/25/16.*

*This course provides information on federal and University requirements
related to the ethical treatment of human research participants.*



IRB Permission

University of Denver Institutional Review Board IRB Application Form

Last edited by: Mohammad Mahoor

Last edited on: February 18, 2014

[\[click for checklist\]](#)

- | | |
|--|---|
| ☐ Ceded review | ☐ HIPAA |
| ☐ Non-DU researcher | ☐ Federal funds |
| | ☐ Schools research |

[472157-2] Evaluating the Use of Humanoid Robots at Improving Social Behaviors and Facial Expressions of Children with Autism



PARENTAL CONSENT TO PARTICIPATE IN A RESEARCH STUDY

**Exploring the Effect of Teaching Music on Imitation and Joint Attention skills
of Children through a Xylophone Player Humanoid Robot**
Dr. Mohammad Mahoor
DU IRB Protocol Number: 779333-1

PURPOSE OF THE STUDY

You are being asked to give your permission for your child to be in a research study. The purpose of this study is to evaluate the use of a xylophone player humanoid robot on teaching music to children. We propose to evaluate the improvement in imitation, joint attention, and social skills of children through our robot-assisted intervention sessions and understand how children respond and interact in dyadic/triadic (Robot-Child and Robot-Child-Parent/Therapist) interactions. The parent will be asked to attend approximately 6 (half of) the sessions to play turn-taking games with the child and the robot. You are being asked to give permission for your child and yourself to be in this research study to provide bases on if this intervention works and how it compares to children with disability (e.g. Autism).

NUMBER OF STUDY PARTICIPANTS

If you agree to allow your child to participate in this study, he/she will be one of 20 people in this research study.




A Fun Filled Research Study
Aimed to Engage Children
on the Autism Spectrum and
Help Improve Social Skills

NOW ACCEPTING
PARTICIPANTS!



**UNIVERSITY of
DENVER**

University of Denver RoboAssist
Autism Research Study

Contact Us

For more information please
contact our research coordinator
Sophia Silver (720) 939-2187

roboassist@du.edu

Dr. Mohammad Mahoor
(303) 871-3745 or
mmahoor@du.edu

**COMPENSATION
WILL BE
PROVIDED**

**This research is
funded by the
National Science
Foundation**

To learn More Visit Our Website
<http://roboassist.wix.com/roboassist>

Or like our Facebook page:
University of Denver Autism
Humanoid Robot Research Study

Let your child explore the world of robotics
while learning social communication skills

**This study examines the use of humanoid robots at
improving social skills such as eye gaze attention
and facial expressions of children with ASD.**

Children are eligible if they:

- Are between 7-17 years
- Have been diagnosed with High Functioning Autism, Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS), or Asperger Syndrome.
- We are also looking for children that have not been diagnosed with Autism to serve as controls for the study. Anyone is Welcome!

Participants will do the following experiments:

1. Play with Robot
2. Facial Expressions Imitation
3. Joint Attention
4. Communicative Skill Training



Research permit forms

[مصوبات کمیته](#)[فرایند بررسی طرح های پژوهشی](#)[فرم ها](#)[اعضا](#)[اساسنامه و آیین نامه ها](#)[اهداف](#)[درباره](#)

Research permit request forms and instructions:

- Request for Research Permit, Form 1.  
- Request for Research Permit, Form 2  
- Occupational Health Form for the Researchers of Animal Labs  
- Instructions for How to Apply for Research Permit from SCS Ethics Committee (IPM) 
- Check List 

Supporting forms:

- Subject Consent Letter [an example template]  
- General Health Questionnaire [an example template]  
- Test Information and the Subjects' Rights [an example template] 



Research permit forms



کمیته اخلاق و ایمنی
پژوهشکده علوم شناختی، پژوهشگاه دانش‌های بنیادی



« SCS Research Ethics Committee of IPM »

Approval Statement Investigator's profile	Name: Alireza Taheri		
	Place of work: Institute for Research in Fundamental sciences (IPM)		
	Email address: artaheri@sharif.edu		
Research Title	Design and Fabrication of a Robotic Arm to be Controlled by EEG Signals		
Research Category	Human research: Healthy individuals ☒ Patients ☐ Psychophysics: ☐ EEG: ☒ fMRI: ☐ Others: ☐ Animal Experiment: Behavioral study: ☐ Invasive Rec.: ☐ Non-invasive Rec.: ☐ Surgery: Headpost: ☐ Recording Chamber: ☐ Craniotomy: ☐ Others: ☐ Animal species: Rat: ☐ Cat: ☐ Non-human primates: ☐ Others: ☐ Awake experiment: ☐ Anesthetized experiment: ☐		
	Ref. No.	Ref. No.SCS.	Ref. NO.SCS.REC.
	Date	Receipt date	Meeting date
	Decision	Approved ☐ Revision ☐	
	Approval Statement		

فرم رضایت نامه

- لطفا تمام اطلاعات موجود در فرم را به دقت مطالعه کنید.
- در صورت داشتن هر گونه سوال یا ابهام در مورد پروژه از افراد مسئول انجام پروژه سوالتان را بپرسید.

عنوان علمی مطالعه: طراحی و ساخت یک بازوی مکانیکی هدایت پذیر توسط سیگنال‌های EEG

شماره ی مجوز کمیته اخلاق پروژه:

(انجام آزمایش بدون مجوز کمیته اخلاق مجاز نیست)

مرکز مسئول انجام پژوهش: پژوهشگاه دانشهای بنیادی- پژوهشکده علوم شناختی

محل انجام مطالعه: بزرگراه ارتش- روبروی آراج- پژوهشگاه دانشهای بنیادی - پژوهشکده علوم شناختی

پژوهشگر اصلی پروژه: علیرضا طاهری ، آزمایشگاه مهندسی عصبی، پژوهشگاه دانشهای بنیادی. آدرس ایمیل:

artaheri@sharif.edu

نام فرد شرکت کننده:

تاریخ

جنسیت:

تولد:

- من تایید می کنم که اطلاعات در مورد طرح پژوهشی و اهداف انجام این طرح پژوهشی توسط محققان مسئول انجام پروژه به صورت کاملی و همین طور در فرم اطلاعات را دریافت کرده ام.
- من به صورت داوطلبانه در این مطالعه شرکت می کنم و شرایط شرکت در مطالعه که در فرم اطلاعات پژوهش به من داده شده را می پذیرم. من به اندازه ی کافی برای تصمیم گیری در مورد شرکت در پروژه زمان دارم.
- من در مورد سولاتم در مورد پروژه از محققان مسئول پروژه پاسخ های رضایت بخش گرفته ام. من فرم اطلاعات شرکت در مطالعه و



فرم اطلاعات طرح پژوهشی و حقوق آزمایش شوند

عنوان رسمی پروژه: طراحی و ساخت یک بازوی مکانیکی هدایت پذیر توسط سیگنال‌های EEG

عنوان کوتاه: کنترل ربات توسط سیگنال های مغز

با تشکر از شرکت شما در این آزمایش این نوشتار اطلاعاتی در مورد طرح آزمایش و حقوق شما در اختیار شما خواهند گذاشت.

۱- هدف انجام این طرح

هدف اصلی ما در این پروژه، کنترل حرکات یک ربات بازو توسط سیگنال های مغزی EEG کاربران می باشد. در این راستا، دو هدف فرعی ما عبارتند از: (۱) طراحی و ساخت یک بازوی مکانیکی حداقل ۳ درجه آزادی به گونه ای که چند حرکت محدود را از طریق سیگنال های EEG تحت کنترل مغز انسان انجام دهد، (۲) اطلاع از فعالیت بخش های مختلف مغز حین تکان دادن دست ها و یا تجسم کردن انجام تعدادی حرکت متداول و روزمره و به کارگیری الگوریتم های مختلف پردازش سیگنال و هوش مصنوعی در شناسایی و دسته بندی حرکات انجام انجام شده توسط کاربر. جهت راستی آزمایی سیگنال های مغزی، از دستبندهای مابو حاوی سنسورهای EMG (ثبت سیگنال های الکتریکی از ماهیچه ها) نیز استفاده می شود که این ادوات نیز مانند داده برداری EEG



Subjects' Rights

لازم است بدانید!

▶ آزمودنی (Subjects, Participants): انسان ها و سایر جاندارانی که آزمایش های یک پژوهش بر روی آنها انجام می شود

موارد مهم نقض حقوق آزمودنی ها:

▶ ارائه اطلاعات نادرست در مورد آزمودنی ها توسط پژوهشگر و موسسه ای که پژوهش در آن انجام می شود.

▶ عدم ارائه اطلاعات یا ارائه اطلاعات نادرست به آزمودنی ها در ارتباط با نوع آزمایش ها و اثرات آنها.

▶ اعمال اجبار جهت مشارکت آزمودنی ها در پژوهش.

▶ عدم ایجاد فرصتهای مناسب برای آزمودنی ها جهت سرباز زدن از ادامه پژوهش در صورتی که پژوهش پرخطر باشد.

▶ عدم رعایت رازداری در حفظ اطلاعات خصوصی آزمودنی ها و افشای هویت آنها.

▶ عدم رعایت حریم خصوصی آزمودنی ها.

▶ استفاده از اطلاعات مربوط به آزمودنی ها برای اهدافی خارج از حیطه پژوهش.

▶ عدم رعایت استانداردهای مربوط به سلامت و ایمنی آزمودنی ها.

▶ استفاده از افراد آسیب پذیر به عنوان آزمودنی در پژوهش، تنها به دلیل موقعیت نامناسب ایشان



Consent vs. Assent form

لازم است بدانید ۱۲!

- ▶ **Consent** may only be given by individuals who have reached the **legal age of consent**.
- ▶ **Assent** is the agreement of someone **not able to** give legal consent to participate in the activity.
 - ▶ Work with **children** or **adults not capable of giving consent** requires the consent of the parent or legal guardian and the assent of the subject.



HRI Considerations

- ▶ یک موضوع (بخوانید معضل) جدی:
- ▶ همان ابتدای مطالعه باید همه چیز مشخص باشد!!
- ▶ دنبال یافتن پاسخ برای چه پرسش (های) پژوهشی هستید؟
- ▶ سناریوی دقیق اجرای پژوهش تان چیست؟
- ▶ چگونه اثربخشی مداخلات خود را ارزیابی می کنید؟ (ابزارها و روش ها)
- ▶ ...



Research Ethics

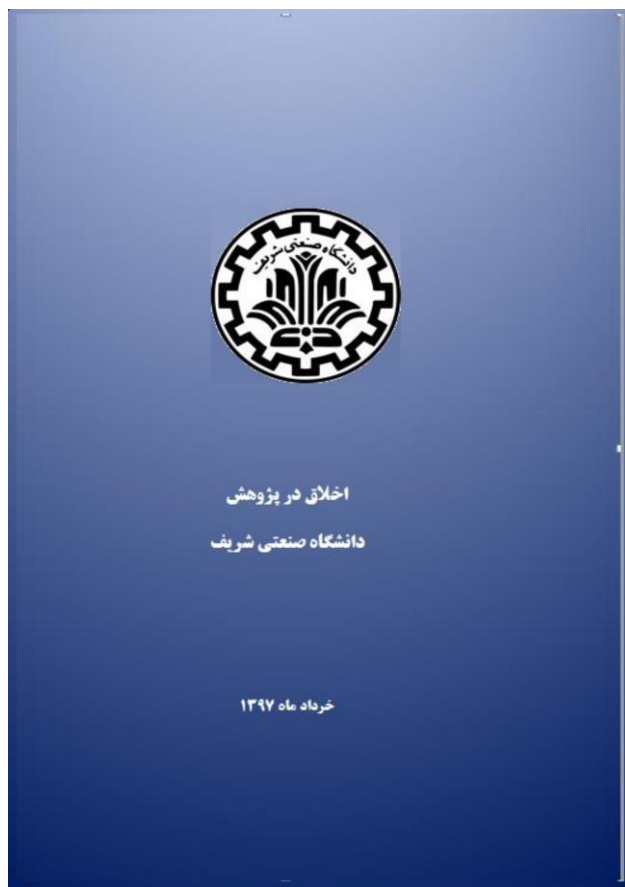
For your information

آیا می دانید؟ (منظور لازم است بدانید ۳ ... !)

▶ نوشتار «اخلاق در پژوهش»

▶ منتشر شده توسط کارگروه سلامت در پژوهش دانشگاه صنعتی شریف

▶ معرفی معیارهای رعایت اخلاق در پژوهش و آموزش و ترویج آن در میان پژوهشگران و دانشجویان تحصیلات تکمیلی



Structure of papers for HRI studies



Structure of papers for HRI studies

An overview

- ▶ Abstract
- ▶ Introduction
- ▶ Related Work
- ▶ Research Questions
- ▶ Robotics Implementation
- ▶ Methodology (or Experiment)
- ▶ Results
- ▶ Discussion
- ▶ Conclusions
- ▶ Future work
- ▶ Acknowledgements
- ▶ Conflict of Interest, etc.
- ▶ References



Structure of papers for HRI studies

▶ 1 Introduction

- ▶ May be split into two sections: **Introduction** and **Background**
 - ▶ for a journal paper, this is suitable since we have to describe the background literature in great detail

▶ 2 Related Work

- ▶ This may be a subsection of the introduction, or stand alone, this depends

▶ 3 Research Questions

یکی از مهم ترین تفاوت ها با مقالات متداول حوزه مهندسی

- ▶ May include further more detailed/precise **research hypotheses** and **expectations**
 - ▶ if hypotheses can be defined based on **previous findings/theories** in the literature
 - ▶ if not (more **exploratory** work) we only have research questions that were driving the research



Structure of papers for HRI studies

به صورت متعارف، بخش ابتدایی مقاله شما باید حداقل شامل سه پاراگراف زیر باشد:

➤ (۱) اهمیت موضوع و حوزه مورد بررسی شما

➤ (۲) اهم پژوهش های انجام شده دیگران در این حوزه به همراه بیان خلاصه یافته ها و حتی اشاره به موارد لحاظ نشده آن ها (البته با بیان احترام آمیز)

➤ (۳) تعریف دقیق مسئله (In this paper, ...):

➤ بیان پرسش (های) پژوهشی، فرضیه ها، تفاوت برجسته با کارهای گذشتگان و ...

➤ **توجه:** صرف انجام یک پژوهش به عنوان **اولین بار** در دنیا/منطقه/کشور، دلیلی بر اهمیت بالا و تضمین پذیرش آن نیست!!

1 Introduction

Children with autism have impaired reactions to real world events and often avoid social interactions and communications [1]. Since 2000, it has been widely shown that using social robots in autism treatment can significantly increase the impact of intervention sessions for children with ASD¹ [2–19]. In this type of research, the impact of the robot-assisted games/scenarios on children's performance as well as their different behavioral skills are investigated over time using different qualitative/quantitative assessment tools (such as video coding of the sessions, questionnaires, etc.) [3–9]. Improvement in imitation [10, 11], joint attention [2–4, 11–14], communication and social interaction skills

In this research, we have designed a set of robot-assisted therapeutic games and run them on 3 pairs of Iranian male individuals with autism during 12 intervention paired-group sessions as a case study. What makes this study different is that 2 out of the 3 pairs of participants are special cases which are rarely investigated in robotic-based studies in the literature; one pair is a set of twins with autism (includes one high- and one low-functioning 7-year-old subject); and the second pair are two (15- and 10-year-old) siblings with high-functioning autism. The strength of investigating twins and siblings are the control advantage provided by these special cases, including the same parents, food, clothing and educa-

and c) the human therapists. The purpose of this study is to find preliminary exploratory answers for the following research questions: a) Does a social robot (as a teacher's assistant) have the ability to teach music (i.e., notes and rhythms) to children with autism?; b) Can robot-assisted interventions affect social and cognitive skills in children with autism through music education?; c) What is the impact of robot-assisted music education intervention on the stress of the subjects' parents during the program?; d) How acceptable is it to have a robot as a co-teacher in music education for the human therapists?



Structure of papers for HRI studies

Participants

► Variables

- **Independent variable**: the variables that are manipulated by the experimenter and is of interest.
- **Dependent variables**: the behavior that is evaluated as the outcome of the experiment.
- Some examples:

Independent	Dependent
The type of medication you gave to a patient	Improvements in health
The type of food you feed to an animal	Weight gain



Structure of papers for HRI studies

Control condition

- ▶ **Experimental condition**: the group(s) in an experimental design that **receive** the experimental condition being evaluated in a study.
- ▶ **Control condition**: one of the groups in an experimental design that **does not receive** the experimental condition being evaluated in a study.
 - ▶ e.g. evaluating **robot tutor A** vs. **robot tutor B** (*two different sets of behaviors, independent variable*), measuring progress in teaching children a second language
 - ▶ Children assigned to robot tutor A or B are considered as the experimental group.
 - ▶ Children who are working with a human tutor are considered as the control group.



Structure of papers for HRI studies

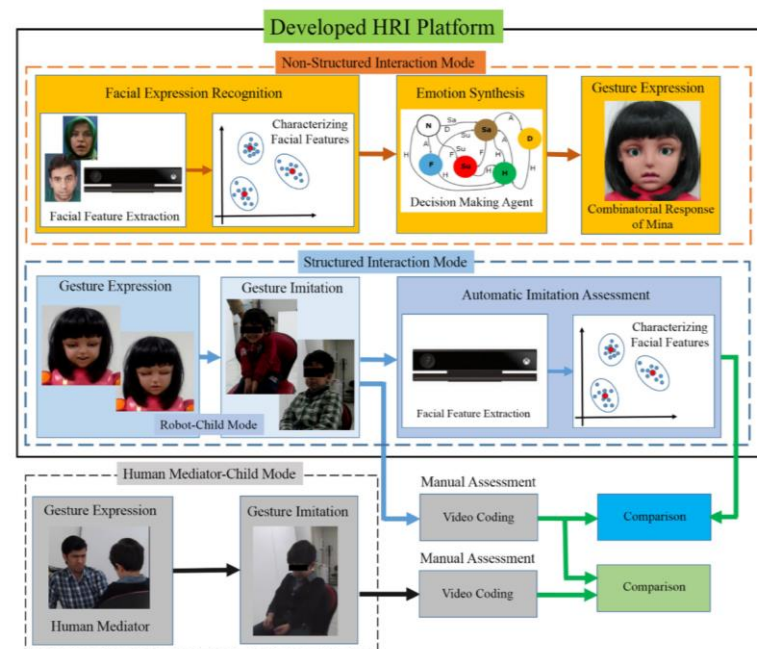
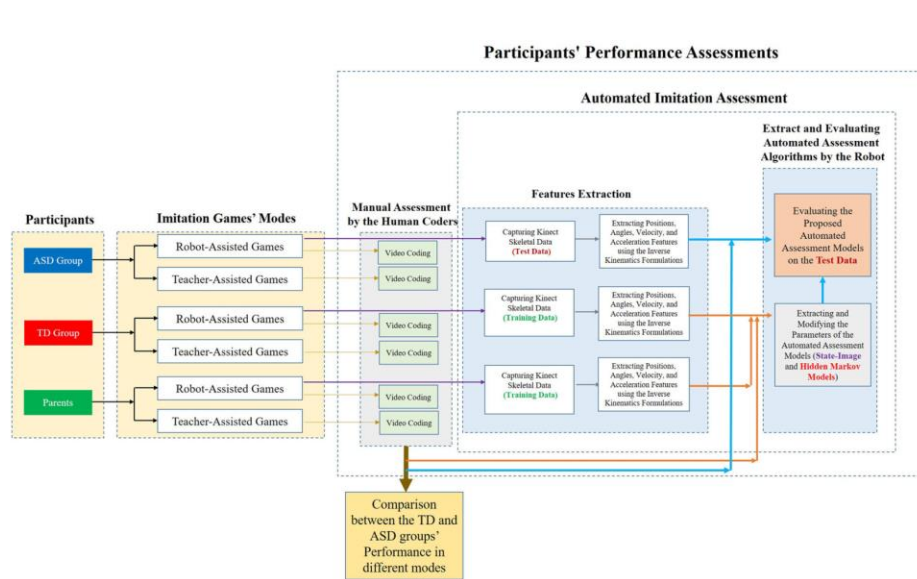
Designs

- ▶ **Between-participants design**: Participants are placed in different groups with each group experiencing different experimental conditions.
 - ▶ e.g. one group interacts only with robot A, the other group only interacts with a different type of robot B
 - ▶ **Sample**: a portion or subset of a population that you recruit for your experiment
 - ▶ Try to be as inclusive as you possibly can within the constraints given (and usually there are many)
- ▶ **Within-participants design**: each participant is exposed to all levels of the independent variable(s).
 - ▶ e.g. all participants interact with robot A and B (counterbalanced)
 - ▶ **Counterbalance**: a procedure used in within-participants designs that changes the order variables are presented to control for sequence/order effects.
- ▶ **Mixed-model factorial design**: include both between-subjects and within-subjects design components



Structure of papers for HRI studies

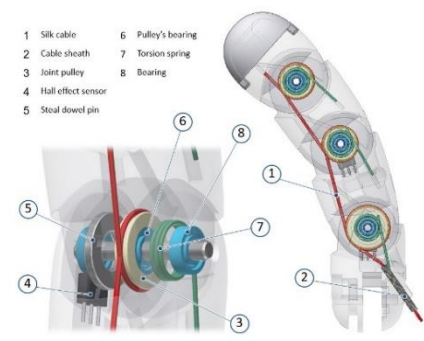
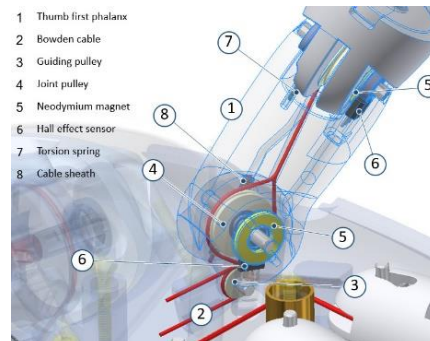
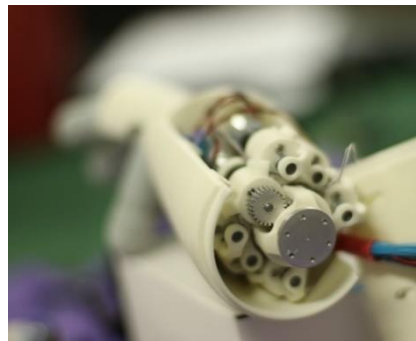
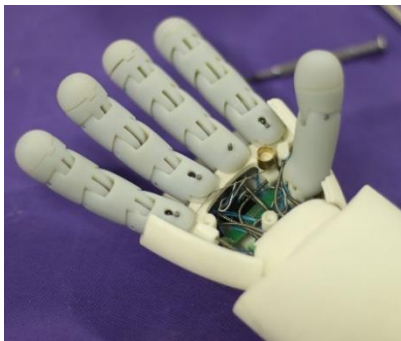
بیان یک ذائقه شفاهی (۱): افزودن فهرست تصویری از مراحل انجام پژوهش و مسیر آنالیز داده ها در انتهای مقدمه (در صورت امکان)



Structure of papers for HRI studies

► 4 Robotics Implementation

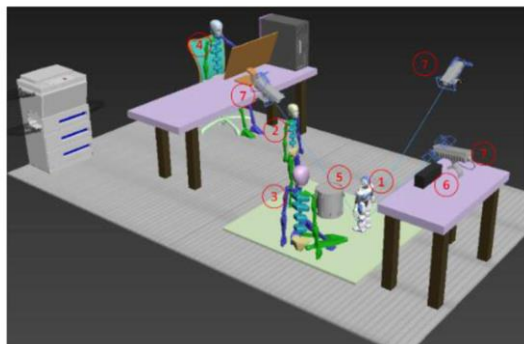
- This section describes **hardware** and **software** and applies if part of the work is a **new robot design/new control architecture**, etc.
- if not (i.e. if we use **an existing robot** with an existing program), then this section is **not necessary** and the robot can just be described briefly in the “Artifacts” section below.
- This section can have subsections on hardware, software, behavior control/learning architecture, etc.



Structure of papers for HRI studies

► 5 Method (Experiment)

- 5.1.1 **Setting** (scenario, location, etc.): this section may also be called “Overall design”
- 5.1.2 Participants: this section may also be called “Sample” (“participants” is preferred if you talk about humans)
- 5.1.3 Artifacts (what robots have been used and how were they controlled, if you use Wizard-of-Oz you may include a subsection on robot control)



(a)



(b)

- 1 Robot
- 2 Autistic child
- 3 Therapist
- 4 Robot operator
- 5 Musical instrument
- 6 Kinect
- 7 Camera
- 8 Screen



Structure of papers for HRI studies

► 5 Method (Experiment)

- 5.1.1 Setting (scenario, location, etc.): this section may also be called "Overall design"
- 5.1.2 **Participants**: this section may also be called "Sample" ("participants" is preferred if you talk about **humans**)
- 5.1.3 Artifacts (what robots have been used and how were they controlled, if you use Wizard-of-Oz you may include a subsection on robot control)

Table 1. The participants' details.

#	Name	Age	Autism severity
1	P1-Ben	6 years and 3 months	High-functioning autism, with hyperactivity; MLU ^a >6
2	P2-Moh	6 years and 1 month	High-functioning autism; MLU>6
3	P3-Rad	6 years and 5 months	High-functioning autism, with verbal deficits; MLU<6
4	P4-Roh	6 years and 6 months	Low-functioning autism, with intellectual disability, poor verbal skills, and lack of eye-contact; MLU<2



Structure of papers for HRI studies

► 5 Method (Experiment)

- 5.1.1 Setting (scenario, location, etc.): this section may also be called "Overall design"
- 5.1.2 Participants: this section may also be called "Sample" ("participants" is preferred if you talk about humans)
- 5.1.3 **Artifacts** (what robots have been used and how were they controlled, if you use **Wizard-of-Oz** you may include a subsection on robot control)



Structure of papers for HRI studies

Study Protocol

- ▶ The **study protocol** involves determining **exactly** how the study will proceed **from start to finish** once a participant arrives.
 - ▶ It is a **detailed description** of **instructions** that will be provided to the participant, what assessments will be done and in what order, what tasks the participant will perform, the timing of events, recording of information, how the data and personal information will be handled, and where this information will be stored for security purposes.
- ▶ This is **necessary** for completing the **IRB paperwork** required for human studies and for maintaining privacy.
- ▶ **Trial runs** of experiments should be conducted until the study can be executed smoothly from start to finish. This is the only way to determine where problems can and likely will occur.
- ▶ It is important that **once the study begins with participants** that the study protocol is discussed with each participant as part of the instruction process as well as providing this information as part of the informed consent form participants will sign.



Structure of papers for HRI studies

Robot Control Paradigms

- ▶ **Wizard of Oz**: the most classical method
 - ▶ robots are remotely controlled, unbeknown to the participants, by a human operator
- ▶ **Fully autonomous**
- ▶ **Between the mentioned methods**
 - ▶ e.g. supervised autonomy



پس از انجام این رساله، معتقدیم که جهت‌گیری فعالیت‌های مهندسی بهره‌گیری از ربات‌ها در آموزش و توانبخشی می‌بایست به سمت توانمندسازی ربات‌ها در دو مقوله‌ی «ارزیابی خودکار» و «توانبخشی تطبیقی» باشد. این‌که ربات (ماشین)، تا حد امکان قابلیت نمره‌دهی و ارزیابی خودکار فعالیت‌های کودک و گزارش آن به معلم/درمانگر انسانی در طول جلسات را داشته باشد و همچنین پروتکل درمانی را منطبق بر شرایط و پیشینه‌ی کودک تعیین نماید.

۲۱۷

یک دیدگاه که همپنان در مسیر تعریف موضوعات پژوهشی، راهگشاست!

علیرضا طاهری، رساله دکتری، ۱۳۹۶



Structure of papers for HRI studies

Robot Control Paradigms

- ▶ Wizard of Oz: the most classical method
- ▶ robots are remotely controlled unbeknownst to the participants



۲۱۷



Structure of papers for HRI studies

Participants

- ▶ Method of recruiting participant?
- ▶ How many participants do I need?
 - ▶ An a priori **Power Analysis** is a statistical calculation that can be performed to determine the appropriate number of participants needed to obtain accurate and reliable results
 - ▶ based on the **number of groups** in the study, the **alpha level** (typically $\alpha = .05$), the **expected or calculated effect size**, and a certain level of **statistical power** (commonly 80%).
 - ▶ There are power analysis **tables** in the appendices of most **statistical textbooks** that will provide group size values. Additionally, there is **software** available online that will assist with this type of calculation (e.g., **G*Power3**);



Structure of papers for HRI studies

- ▶ 5 Method (Experiment) (cont.)
 - ▶ 5.1.4 Procedures and measures
 - ▶ What happens in the experiment, what measures are used (e.g. video data, questionnaires, interviews etc.).
 - ▶ Alternatively, instead of “Procedures and measures” you can have a subsection called “Instruments” (describing the measures) and a separate subsection called “Procedure”
 - ▶ 5.1.5 Data collection (optional)
 - ▶ May be included in 5.1.4, depends on how extensive the methods were for data collection, whether new data collection methods had been developed etc.)
 - ▶ 5.1.6 Coding and Reliability (optional, needed when doing e.g. behavioral data analysis)
 - ▶ 5.1.7 Statistical approach (optional, if statistical methods for data analysis have been used)



Structure of papers for HRI studies

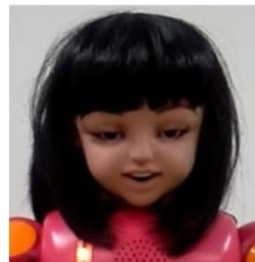
► 5 Method (Experiment) (cont.)

► 5.1.4 Procedures and measures

- What happens in the experiment, what measures are used (e.g. video data, questionnaires, interviews etc.).

Table 2. The designed music-based therapy interventions schedule.

Week #	Outline	Games' modes and descriptions	Purpose of the game
2	Orientation session	After conducting the Pre-test in week two, Nima robot was introduced to the child, showing robot's greeting, singing, dancing, etc. No bars were removed from the xylophone.	Familiarize the child with the robot/class' environment
3	Strike the drum and/or xylophone bars using mallets	3a Child should hit an arbitrary bar after each robot's strike.	-Improve fine motor skills -Visual pursuit
		3b Some of the bars are removed. Repeat 3a with fewer bars on the xylophone.	-Learn how to make sound with the instruments -Joint attention
		3c Repeat 3a: Child should use both of hands in order to play the instruments.	
		3d Repeat 3b: Child should use both of hands in order to play the instruments.	
4	Matching the colors	4a Child imitates the robot's strike considering bars' color (preferable colors: red, blue and yellow).	-Colors' recognition -Joint attention -Instruction perception
		4b The robot strikes a bar and child points to the same color on the board/screen (and vice versa).	
		4c The robot's eye LED color changes and child has to hit the same color on the xylophone.	
5	Hitting bars while counting	5a Imitating the robot's strikes considering the number of played notes in order. The child/robot should count "one-two" while hitting the bars.	-Improve verbal short -Term memory -Numbers' perception and counting
		5b Repeat 5a while saying "Baam-Baam"	



(a)



(b)



(e)



(f)



Structure of papers for HRI studies

Common assessment tools for evaluation of HRI studies

- ▶ Primary methods of **evaluation** (*from a social interaction perspective*):
 - ▶ 1. Self assessments
 - ▶ Paper or computer-based scales, **questionnaires** or surveys, **Personal assessment of robots, situation, tasks**, etc.
 - ▶ Problems with validity and corroboration
 - ▶ **Biases (!)**
 - ▶ 2. Interview
 - ▶ 3. Behavioral measures
 - ▶ **Interrater reliability in behavioral data**: the **degree of agreement** among raters. It is a score of how much homogeneity or consensus exists in the ratings given by various judges.
 - ▶ **Second coder** who **does not know** about the codes percentage, e.g. **20-100%**, statistical interrater reliability tests
 - ▶ 4. Psychophysiological measures
 - ▶ The primary advantage for using psychophysiological measurements is that participants **cannot consciously manipulate** the activities of their autonomic nervous system (**heart rate, EEG, ECG, etc.**).
 - ▶ 5. Task performance metrics
 - ▶ These metrics are designed to measure **how well** a person/team **performs** or **completes** tasks.
- ▶ Each comes with **advantages** and **disadvantages**. Disadvantages can overcome by using **combinations of methods** for evaluation.



Structure of papers for HRI studies

Common assessment tools for evaluation of HRI studies

▶ Questionnaire examples in the field of social robotics:

▶ The Godspeed Questionnaire Series,

<https://www.bartneck.de/2008/03/11/the-godspeed-questionnaire-series/>

- ▶ The Godspeed questionnaire categorizes perception using the following: (1) anthropomorphism, (2) animacy, (3) likability, (4) perceived intelligence, and (5) perceived safety.

▶ Unified Theory of Acceptance and Use of Technology (UTAUT)

https://en.wikipedia.org/wiki/Unified_theory_of_acceptance_and_use_of_technology

- ▶ The UTAUT aims to explain user intentions to use an information system and subsequent usage behavior. The theory holds that there are four key constructs: 1) performance expectancy, 2) effort expectancy, 3) social influence, and 4) enabling conditions.

ملاحظات

- ✓ پرسشنامه های تخصصی فراوانی در حوزه های مختلف وجود دارد که لازم است توسط محقق شناسایی و انتخاب شوند.
- ✓ تلاش کنید تا حد امکان از پرسشنامه های ترجمه شده موجود بهره گیرید.
- ✓ طراحی پرسشنامه های اختصاصی اگرچه شدنی است، اما نیاز به رعایت نکات و جزئیات مربوط به خود را دارد.



Structure of papers for HRI studies

Common assessment tools for evaluation of HRI studies

Primary

1. S

2. I

3. E

4. F

5. T

6. C

7. A

8. P

9. R

10. S

11. T

12. C

Each of these tools is used for evaluation of HRI studies.



of robots, situation,

it is a score of how

ater reliability tests

participants cannot
e, EEG, ECG, etc.).

ompletes tasks.

ercome by using

Structure of papers for HRI studies

► 5 Method (Experiment) (cont.)

► 5.1.4 Procedures and measures

- What happens in the experiment, what measures are used (e.g. video data, questionnaires, interviews etc.).

2.7. Assessment tools

To find preliminary answers for the research questions of this study, three types of measuring instruments, i.e., human assessments, questionnaires, and interviews, were used at different times:

- a2) To investigate music learning ability: 1) Stambak's Rhythmic Structures Reproduction test [31], which contains 21 (easy to hard level) rhythmic tasks in which the participant should reproduce the patterns on a drum after hearing (and not seeing) them was conducted.

- b) Three valid questionnaires in autism research. All of the questionnaires were filled in by the parents four times.

- b1) *Gilliam Autism Rating Scale (GARS)* [33] is a questionnaire for estimating autism severity, covering subscales: Stereotyped behaviors, communication, social interactions, and developmental disturbances;
- b2) *Autism Social Skills Profile (ASSP)* [34], is a comprehensive tool to measure social behaviors of an individual with autism from 6 to 17 years old. ASSP consists of 45 items. Three subscales, i.e., social interaction, social reciprocity, and detrimental social behaviors, can be measured with this questionnaire;
- b3) *Parenting Stress Index-Short Form (PSI-SF)* [35] is a tool to measure parenting stress.



Structure of papers for HRI studies

- ▶ 5 Method (Experiment) (cont.)
 - ▶ 5.1.4 Procedures and measures
 - ▶ What happens in the experiment, what measures are used (e.g. video data, questionnaires, interviews etc.).

3.1.1 Gaze-Shifting

The gaze-shifting score is the total number of the participant's gazes toward the robot and/or the other individuals (i.e. the therapist, parents, or his playmate) in the room during the intervention sessions. The purposes of studying this Visual Attention-based parameter are to find out (a) whether there is correlation between the rate of the total gaze-shifts of different participants/paired-groups during the treatment time, and (b) the gaze-shifting scores' trend toward the robot (i.e. the item with the most important role in our therapeutic sessions). In order for the data to be comparable, the scores

3.1.2 Joint Attention (JA)

In this study, the video coders scored the following behaviors as the Joint Attention (JA) scores of the participants: (a) following the pointing/gaze-direction of the robot or other individuals in the room correctly, or gaze-shifting between a person and other situation/object; which is known as the Responding to Joint Attention (RJA) score, and (b) showing an object, or pointing to far/near points was called the Pointing score. The total JA score is the sum of RJA and Pointing activities that have been observed in the interventions' movies. The scores have been normalized/reported by dividing the number of correct tasks to the game time (i.e. number per minute). In this research, joint attention behaviors occurred more often in the Kinect-based recognition game, pointing game, and playing the virtual/real xylophone game.



Structure of papers for HRI studies

▶ 6 Results

- ▶ Usually with subsections, just presenting the quantitative and/or qualitative results without much interpretation

اساسی ترین سوال این فصل: چگونه با یک تحلیل آماری معتبر نشان دهیم که مطالعه ما اثربخش بوده است؟
تا انتهای فصل با ما همراه باشید

▶ 7 Discussion

- ▶ Interpretation of the results in light of the research questions/hypotheses, also mention limitations

▶ 8 Conclusions

- ▶ Brief summary of main results, “what have we learnt”, how can other researchers benefit from this work

▶ 9 Future work

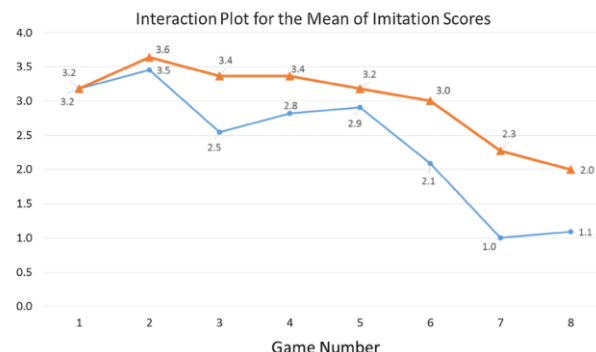
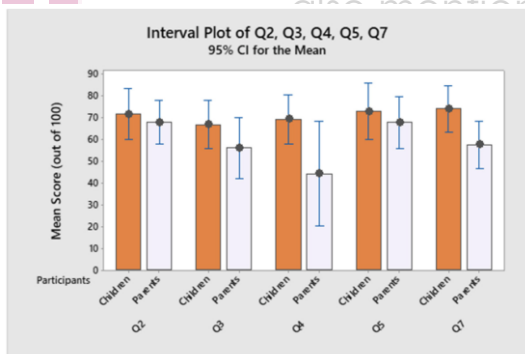


Structure of papers for HRI studies

6 Results

Table 7 Analysis of Variance for imitation scores of the eleven collaborating children using Minitab software

Source	DOF	F-value	p value	Partial η^2
Factor1: Imitation Mode	1	11.81	0.001	0.069
Factor2: Action Number	7	8.45	0.000	0.270
Interaction of Imitation Mode and Action Number	7	0.75	0.632	0.032



9 Future work

Item	Dependent Variable	Condition1: Individual Participants P1 to P6			Condition2: Participants in three paired groups: "Twins", "Siblings", and "Classmates"		
		Participant	Treatment Time	Interaction of the factors	Participant Groups	Treatment Time	Interaction of the factors
Gaze-Shifting	Total Gaze Score	0.000	0.079	0.078	0.622	0.110	0.021
	Gaze toward the Robot	0.201	0.010	0.146	0.327	0.004	0.006
	Gaze Percentage toward the Robot	0.001	0.038	0.319	0.017	0.024	0.146
	Gaze toward the Humans	0.000	0.401	0.112	0.475	0.578	0.106
Joint-Attention (JA)	Pointing Score	0.310	0.003	0.391	0.083	0.001	0.301
	Responding to JA Score	0.932	0.008	0.831	0.703	0.005	0.338
	Total Score of JA	0.449	0.000	0.264	0.115	0.000	0.281
Imitation	Imitation Success Rate	0.483	0.910	0.932	0.288	0.878	0.756
Behavioral Problem	Maladaptive Behaviors Score	0.000	0.698	0.425	0.471	0.676	0.213
Verbal	Total VC Score	0.000	0.038	0.200	0.002	0.046	0.169
Communication (VC)	Amount of VC with Robot	0.042	0.017	0.052	0.051	0.021	0.113
	VC Percentage with Robot	0.023	0.030	0.028	0.647	0.250	0.736



Structure of papers for HRI studies

- ▶ **Acknowledgements**
 - ▶ **Do not forget** to mention sponsors as well as colleagues who are not authors but e.g. gave practical help with the experiment or the analysis
- ▶ **Conflict of Interest, etc.**
- ▶ **References**

Acknowledgements This research was supported by the “Iranian National Science Foundation (INSF)” (<http://en.insf.org/>). We also appreciate the “Dr. AliAkbar Siassi Memorial Grant Award” for their complementary support of the Social & Cognitive Robotics Laboratory.

Authors' contributions All authors contributed to the study conception and design, material preparation, data collection and analysis were performed by Salar Basiri and Alireza Taheri. The first draft of the manuscript was written by Salar Basiri and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Author Alireza Taheri has received research grants from the “Iranian National Science Foundation (INSF)” (Gant No. 98025100).

Data availability All data from this project (videos of the sessions, results of the questionnaires, scores of performances, etc.) are available in the archive of the Social & Cognitive Robotics Laboratory.

Code availability All of the codes are available in the archive of the Social & Cognitive Robotics Laboratory. In case the readers need the codes, they may contact the corresponding author.

Declarations

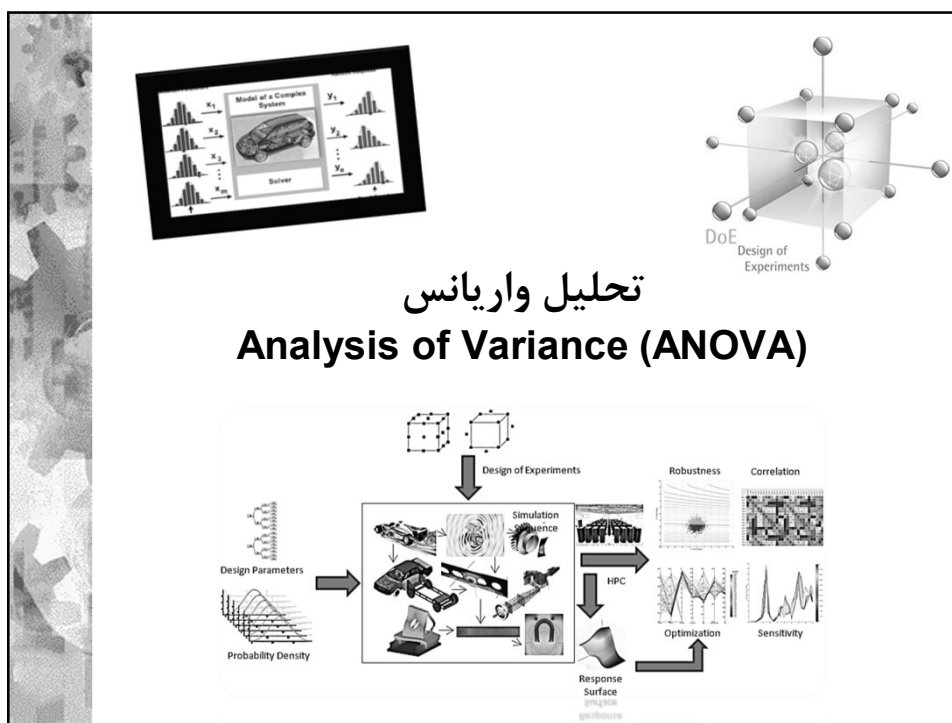
Conflict of interest The authors Salar Basiri, Ali Meghdari, Mehrdad Boroushaki, and Minoo Alemi declare that they have no conflict of interest.

Ethics approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Ethical approval for the protocol of this study was provided by the Iran University of Medical Sciences (#IR.IUMS.REC.1395.95301469).

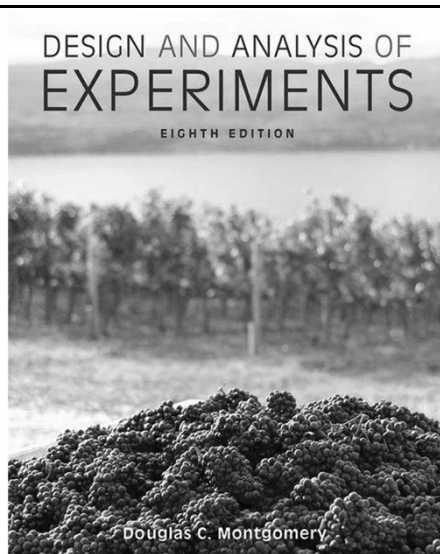


Tests of Statistical Significance



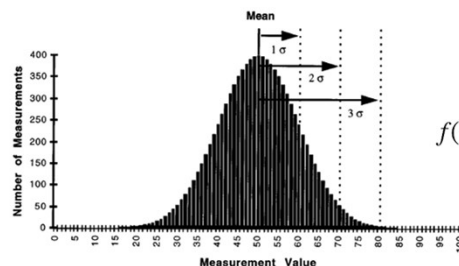


مرجع اصلی



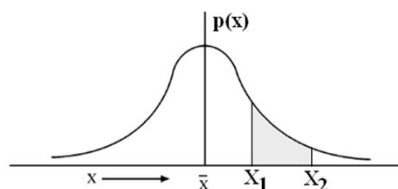
توزیع نرمال

■ توزیع احتمال بسیاری از پدیده ها و اتفاقات نرمال است و یا نرمال مدل می شود.



$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(1/2)[(y-\mu)/\sigma]^2}$$

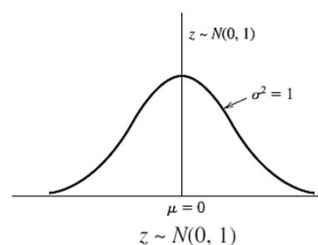
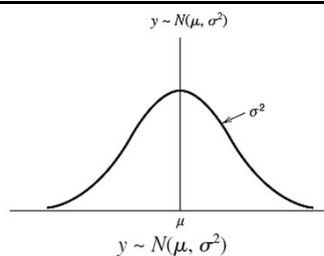
■ توزیع نرمال تنها به مقدار متوسط و انحراف از معیار وابسته است.



$$P(-\infty < X < +\infty) = \int_{-\infty}^{+\infty} p(x) dx = 1$$

$$P(X_1 \leq X \leq X_2) = \int_{X_1}^{X_2} p(x) dx$$

توزیع نرمال استاندارد



$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(1/2)[(y-\mu)/\sigma]^2} \xrightarrow{z = \frac{y-\mu}{\sigma}} f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

$$P(\mu - \sigma \leq Y \leq \mu + \sigma) = \int_{\mu-\sigma}^{\mu+\sigma} p(y) dy = 0.68$$

$$P(-1 \leq Z \leq 1) = \int_{-1}^1 p(z) dz = 0.68$$

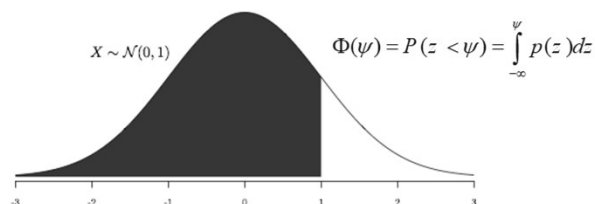
$$P(\mu - 2\sigma \leq Y \leq \mu + 2\sigma) = \int_{\mu-2\sigma}^{\mu+2\sigma} p(y) dy = 0.95$$

$$P(-2 \leq Z \leq 2) = \int_{-2}^2 p(z) dz = 0.95$$

$$P(\mu - 3\sigma \leq Y \leq \mu + 3\sigma) = \int_{\mu-3\sigma}^{\mu+3\sigma} p(y) dy = 0.997$$

$$P(-3 \leq Z \leq 3) = \int_{-3}^3 p(z) dz = 0.997$$

تابع توزیع تجمعی (Cumulative Distribution Function)



	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015

نمونه برداری و توزیع های نمونه ای

- میانگین نمونه (*Sample mean*): برآوردی از میانگین نامعلوم.

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$$

- واریانس نمونه (*Sample variance*): برآوردی از واریانس نامعلوم.

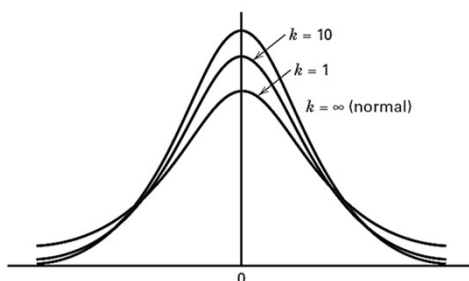
$$S^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - 1}$$

- انحراف از استاندارد نمونه (*Sample standard deviation*):

برآوردی از انحراف از استاندارد نامعلوم.

$$S = \sqrt{S^2}$$

تابع توزیع احتمال t - student

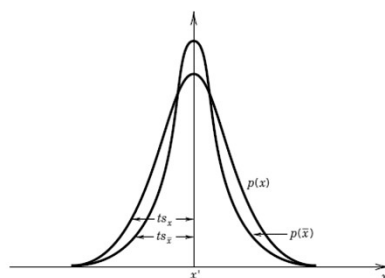
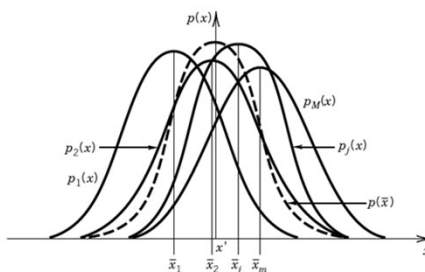


t - student distribution with k degrees of freedom

$$k = n - 1$$

$$f(t) = \frac{\Gamma[(k+1)/2]}{\sqrt{k\pi}\Gamma(k/2)} \frac{1}{[(t^2/k) + 1]^{(k+1)/2}} \quad -\infty < t < \infty$$

انحراف از معیار میانگین ها (Standard Deviation of the Means)

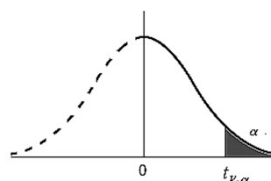


$$s_{\bar{x}} = \frac{s_x}{\sqrt{N}}$$

تابع توزیع احتمال *t-student*

Coverage factor;

$$t = \frac{\bar{y} - \mu}{S/\sqrt{n}}$$



Percentage Points of the *t* Distribution^a

$\alpha \backslash \nu$	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
120	0.254	0.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	0.253	0.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

مقایسه آماری نتایج آزمایش ها

آزمون فرض (Hypothesis Testing)

- یک فرض آماری ادعایی در مورد یک یا چند جمعیت مورد مطالعه است که ممکن است درست یا نادرست باشد.
- برای بررسی فرض، یک آزمایش تصادفی انجام می‌دهیم، اگر نتیجه آزمایش تفاوت "معنی داری" با آنچه که انتظار داریم، داشته باشد فرض را رد می‌کنیم و در غیر این صورت آن را رد نمی‌کنیم (می‌پذیریم).
- ارائه یک فرض صفر و فرض مقابل؛
 H_0 : Null فرض صفر
 H_1 : Alternative فرض مقابل
- بررسی رد یا عدم رد فرض صفر؛
- اگر دلایل کافی برای رد شدن فرض صفر وجود داشته باشد، فرض صفر رد می‌شود.
- اگر دلایل کافی برای رد شدن فرض صفر وجود نداشته باشد، فرض صفر رد نمی‌شود.

روشهای مقایسه آماری نتایج آزمایش ها

۲- بازه اطمینان (confidence interval)

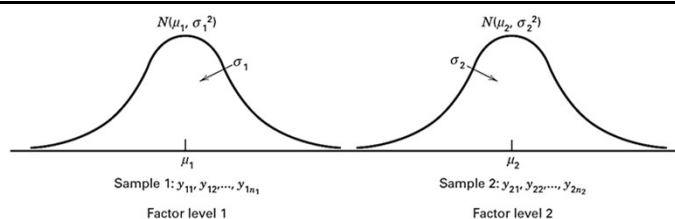
- هرچند آزمون فرض، ابزاری بسیار مفید در نتیجه گیری و تصمیم گیری است، ولی ارائه گر تمام نتایج مورد نظر نمی باشد.
- گاهی ما نتیجه آزمون فرض را می دانیم ولی نمی دانیم تغییرات متغیر مورد بررسی با چه درصد اطمینان و در چه محدوده ای قرار دارد.
- در روش بازه اطمینان، محدوده های تغییرات متغیرها در قالب بازه هایی با درصد اطمینان مشخص، بیان می شود.

$$\bar{y} - u_p \leq \mu_y \leq \bar{y} + u_p$$

$$\mu_y = \bar{y} \pm u_p \quad (\% p)$$

آزمون فرض

(Hypothesis Testing)



آزمون فرض:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

انواع خطاها:

$$\alpha = P(\text{I خطای نوع}) = P(H_0 \text{ رد} \mid H_0 \text{ درست})$$

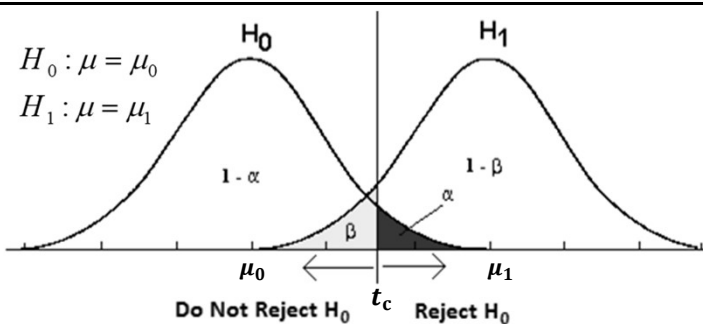
$$\beta = P(\text{II خطای نوع}) = P(H_0 \text{ عدم رد} \mid H_0 \text{ نادرست})$$

هر دو نوع خطا بایستی در حد امکان مینیموم باشند.

معمولاً استفاده از تعریف توان آزمون (*Power of the test*) ساده تر است:

$$1 - \beta = P(H_0 \text{ نادرست} \mid H_0 \text{ رد})$$

مفاهیم α و β در آزمون فرض



α : احتمال رد فرض وقتی صحیح است.

β : احتمال قبول فرض وقتی غلط است.

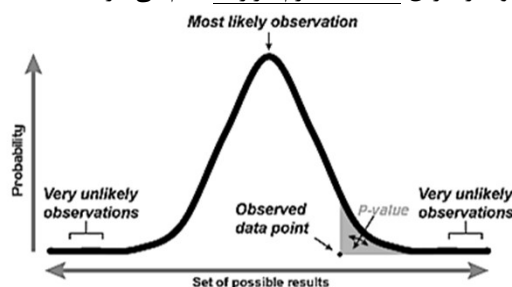
$1 - \alpha$: احتمال قبول فرض وقتی صحیح است.

$1 - \beta$: احتمال رد فرض وقتی غلط است.

		Reality	
		True	False
Research	Fail to reject H_0	Correct ($1 - \alpha$)	Type II error β
	Reject H_0	Type I error α	Correct ($1 - \beta$)

مقدار p

- مقدار بحرانی یا آستانه رد فرض صفر (سطح معناداری آزمون) است.
- کوچکترین سطح α که منجر به معنادار شدن اختلاف می گردد.
- با درصد اطمینان $100 \times (1 - p)$ می توان فرض صفر را رد کرد.
- مقدار p احتمال شانس بودن رد فرض صفر را نشان می دهد.
- محاسبه مقدار p همواره کار ساده ای نیست که غالباً با سعی و خطا قابل حصول است و معمولاً از طریق محاسبات نرم افزاری انجام می گردد.



A p-value (shaded green area) is the probability of an observed (or more extreme) result arising by chance

- Pregnancy Test Example

H_0 : person is not pregnant vs. H_A : person is pregnant

False Positive

Type I Error



False Negative

Type II Error



- Courtroom Example

H_0 : the defendant is innocent vs. H_A : the defendant is guilty

Type I Error: sentencing an innocent person to jail

Type II Error: letting a guilty person go free.

$$\alpha = P(\text{Type I Error}) \quad \beta = P(\text{Type II Error}) \rightarrow 1 - \beta = \text{Power}$$



تعداد تکرار آزمایش مورد نیاز جهت انجام آزمون فرض

- یکی از سوالاتی که همواره مطرح است این است که؛
- ✓ تعداد تکرار در آزمایش جهت انجام یک آزمون فرض قابل قبول چقدر است؟
- محاسبه تعداد تکرار آزمایش جهت انجام آزمون آماری فرض، وابسته به پاسخ سوالات زیر است؛
- ۱- رزولوشن (میزان اختلاف معنی دار) در آزمایش چقدر است؟
- ۲- پراکندگی (انحراف از معیار) آزمایش چقدر است؟
- ۳- خطاهای مطلوب نوع اول α (سطح معناداری آزمون) و نوع دوم β چقدر است؟

مثال

تعداد تکرار آزمایش جهت انجام آزمون فرض

- ✓ فرض کنید ادعا شده است عمر نوعی باتری لیتیوم ۱۰۰۰ ساعت است.
- ✓ طی مجموعه آزمایش هایی می خواهیم درستی این ادعا را بررسی کنیم.
- ✓ با چه تعداد تکرار آزمایشی می توان با ۹۵ درصد معناداری آزمون، با حداقل رزولوشن ۰/۵ ساعت این ادعا را بررسی کرد.
- ✓ بر اساس تجربیات گذشته انحراف از استاندارد عمر باتری ها ۲ ساعت است.
- حل: ابتدا به سوالات زیر پاسخ می دهیم؛



$$H_0: \mu_0 = 1000h \quad \text{؟ آزمایش چقدر است؟}$$

$$H_1: \mu_1 = 999.5h$$

$$\text{Res} = 0.5h$$

۲- پراکندگی (انحراف از معیار) آزمایش چقدر است؟

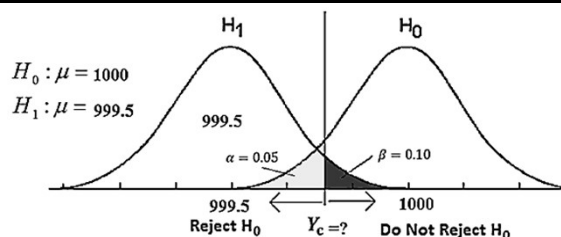
$$\sigma = 2h \quad \text{مطابق با تجربه:}$$

۳- خطاهای نوع اول α (سطح معناداری آزمون) و نوع دوم β چقدر است؟

$$\alpha = 0.05 \quad \text{(مطابق صورت مساله) و فرض می کنیم: } \beta = 0.10$$

مثال

تعداد تکرار آزمایش جهت انجام آزمون فرض



$$Z_\alpha = \frac{|Y_c - \mu|}{\sigma/\sqrt{n}}$$

$$\frac{1000 - 999.5}{2/\sqrt{n}} = 1.645$$

▪ معادله خطای نوع اول (α):

از طریق جدول $Z_{0.05} = 1.64$

▪ معادله خطای نوع دوم (β):

$$\frac{Y_c - 999.5}{2/\sqrt{n}} = 1.282$$

از طریق جدول $Z_{0.10} = 1.282$

▪ از حل دستگاه دو معادله و دو مجهول:

✓ تعداد تکرار مورد نیاز: $n = 136.9 \approx 137$

✓ حد بحرانی مقایسه: $\bar{Y}_c = 999.7$

در آزمایش ۱۳۷ نمونه اگر عمر میانگین کمتر از ۹۹۹.۷ ساعت بود فرض رد و در غیر این صورت فرض پذیرفته است.

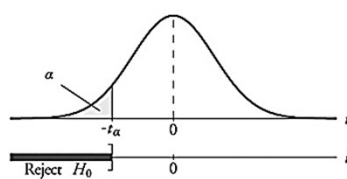
$$H_0: \mu = 1000$$

$$H_1: \mu < 1000$$

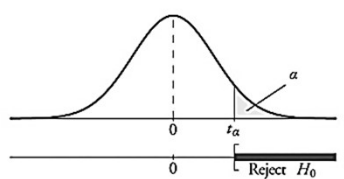
انواع فرض های مقابل

$$H_0: \mu = \mu_0$$

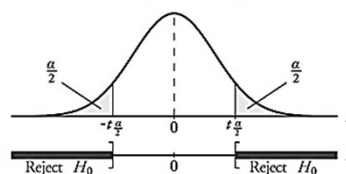
$$H_1: \mu < \mu_0$$



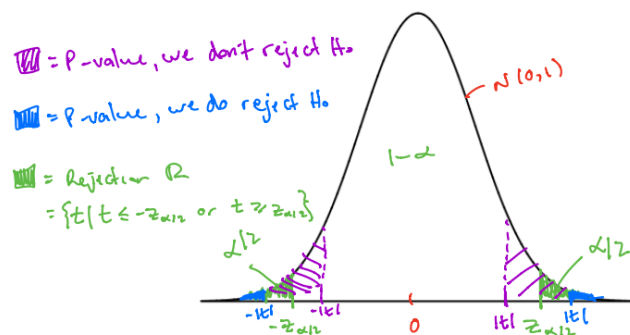
$$H_1: \mu > \mu_0$$



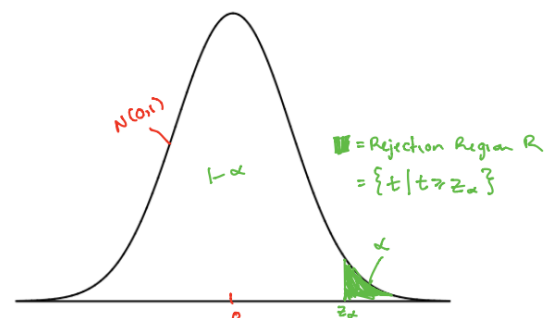
$$H_1: \mu \neq \mu_0$$



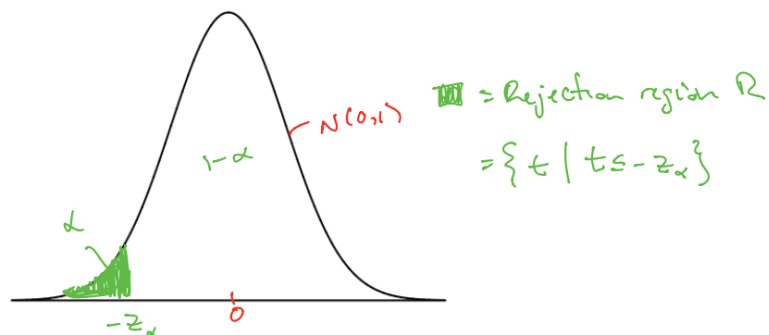
$$H_0: \theta_1 = \theta_2 \text{ vs. } H_A: \theta_1 \neq \theta_2$$



$$H_0: \theta_1 \leq \theta_2 \text{ vs. } H_A: \theta_1 > \theta_2$$

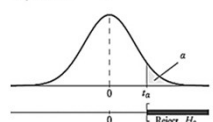
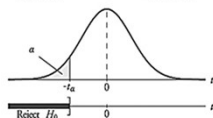
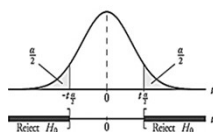


$$H_0: \theta_1 \geq \theta_2 \text{ vs. } H_A: \theta_1 < \theta_2$$



آزمون t جامعه آماری با واریانس معلوم: مقایسه میانگین با مقدار مشخص

معیار رد	آماره آزمون	آزمون فرض
$ Z_0 > Z_{\alpha/2}$		$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$
$Z_0 < -Z_\alpha$	$Z_0 = \frac{\bar{y} - \mu_0}{\sigma / \sqrt{n}}$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$
$Z_0 > Z_\alpha$		$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$



مثالی از آزمون t جامعه آماری با واریانس معلوم: مقایسه میانگین با مقدار مشخص

مثال

- می دانیم انحراف معیار قطر میله های فولادی $\sigma = 0.0001 \text{ in}$ است. متوسط قطر یک نمونه تصادفی مشتمل بر 10 میله را برابر 0.2545 in بدست آورده ایم. مایلیم این فرض را که میانگین واقعی قطر این میله ها برابر 0.255 in است را آزمون کنیم؟

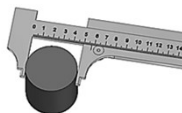
$$H_0: \mu = 0.255$$

$$H_1: \mu \neq 0.255$$

REJECTED

حل:

- مجموعه آماری دارای توزیع نرمال است:



$$Z_0 = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{0.2545 - 0.255}{0.0001 / \sqrt{10}} = -15.81$$

- از طریق جدول توزیع نرمال برای فرض صفر در سطح 95%:

$$Z_{\alpha/2} = Z_{0.025} = 1.96$$

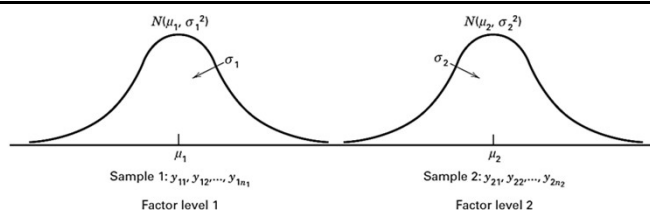
- از آنجا که $|Z_0| (= |-15.81|) > Z_{0.025} (= 1.96)$ بنابراین فرض صفر رد می شود؛

$$\mu \neq 0.255$$

آزمون t جامعه آماری با واریانس نامعلوم: مقایسه میانگین با مقدار مشخص

معیار رد	آماره آزمون	آزمون فرض
$ t_o > t_{\alpha/2, n-1}$		$H_o: \mu = \mu_o$ $H_1: \mu \neq \mu_o$
$t_o < -t_{\alpha, n-1}$	$t_o = \frac{\bar{y} - \mu_o}{S / \sqrt{n}}$	$H_o: \mu = \mu_o$ $H_1: \mu < \mu_o$
$t_o > t_{\alpha, n-1}$		$H_o: \mu = \mu_o$ $H_1: \mu > \mu_o$

آزمون t دو جامعه آماری با واریانس های معلوم : مقایسه میانگین ها



معیار رد	آماره آزمون	آزمون فرض
$ Z_o > Z_{\alpha/2}$		$H_o: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$
$Z_o < -Z_{\alpha}$	$Z_o = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$H_o: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$
$Z_o > Z_{\alpha}$		$H_o: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$

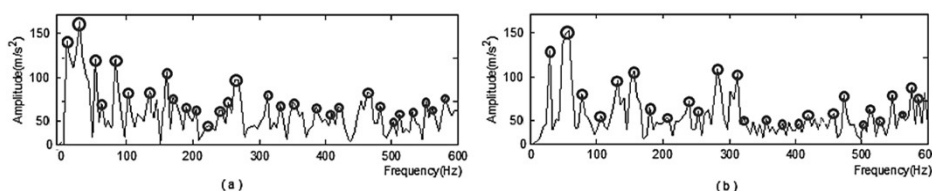
مثال آزمون t دو جامعه آماری با واریانس های معلوم : مقایسه میانگین ها

مثال

▪ دو گروه مجزا، جهت تخمین دامنه شتاب در تستهای ارتعاشی نتایج زیر را بدست آورده اند:

مشخصات	گروه ۱	گروه ۲
تعداد تست	90	100
میانگین نمونه	76 m/s^2	81 m/s^2
انحراف از معیار نمونه	8.2 m/s^2	7.6 m/s^2

□ در سطح 98% آیا اختلاف میانگین های دو جامعه آماری معنی دار است؟



مثال: بازه اطمینان جهت مقایسه دو جامعه آماری

مثال

▪ حل: الف) دو جامعه آماری به اندازه کافی بزرگ است و می توان نرمال فرض کرد.

$$H_0 : \mu_1 = \mu_2$$

▪ آزمون فرض:

$$H_1 : \mu_1 \neq \mu_2$$

$$Z_0 = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{76 - 81}{\sqrt{\frac{(8.2)^2}{90} + \frac{(7.6)^2}{100}}} = -4.34$$

▪ درصد اطمینان 98% معادل با خطای $\alpha = 0.02$ است. از طریق جدول خواهیم داشت:

$$Z_{\alpha/2} = Z_{0.01} = 2.326$$

▪ از آنجاکه $|Z_0| > Z_{0.01}$ فرض برابری میانگین ها با 98% اطمینان رد می شود.

اختلاف میانگین های نتایج دو گروه معنی دار است.

آزمون t دو جامعه آماری با واریانس های نامعلوم : مقایسه میانگین ها

معیار رد	آماره آزمون	آزمون فرض
	اگر $\sigma_1^2 = \sigma_2^2$	
	$t_e = \frac{\bar{y}_1 - \bar{y}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ $S_p = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$ $v = n_1 + n_2 - 2$	$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$
$ t_e > t_{\alpha/2, v}$		
	اگر $\sigma_1^2 \neq \sigma_2^2$	
	$t_e = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$ $v = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}}$	$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 < \mu_2$ $H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 > \mu_2$
$t_e < -t_{\alpha, v}$ $t_e > t_{\alpha, v}$		

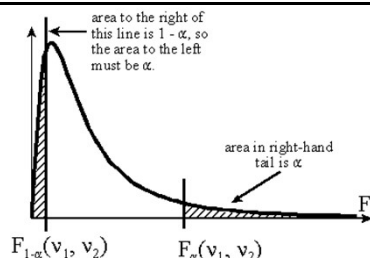
واریانس های واقعی (σ_i^2): نامعلوم

واریانس های نمونه (S_i^2): معلوم

آزمون برابری دو واریانس

- در بسیاری از موارد، واریانس های واقعی (σ_i) نامعلوم و واریانس های نمونه (S_i) معلوم هستند.
 - در این شرایط باید برابری واریانس های واقعی از طریق آزمون مقایسه واریانس های نمونه (S_i) انجام گردد.
- $$H_0 : \sigma_1^2 = \sigma_2^2$$
- $$H_1 : \sigma_1^2 \neq \sigma_2^2$$
- در این شرایط باید نسبت واریانس های نمونه ها محاسبه گردد:
- $$F_o = \frac{S_1^2}{S_2^2}$$
- به دلیل اینکه نسبت واریانس ها دارای توزیع F است، آزمون فرض مقایسه واریانس ها تحت این توزیع بررسی و تحلیل می شود.

مقادیر توزیع F



$$F_{1-\alpha}(v_1, v_2) = 1/F_{\alpha}(v_2, v_1)$$

Percentage Points of the F Distribution

		Degrees of Freedom for the Numerator (v_1)																		
v_2		1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
		1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33	
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49	
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13	
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76	
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10	
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72	
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47	
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29	
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16	
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06	
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29	
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19	
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.60	

آزمون برابری دو واریانس

آزمون فرض	آماره آزمون	معیار رد
$H_0: \sigma_1^2 = \sigma_2^2$	$F_0 = \frac{S_1^2}{S_2^2}$	$F_0 > F_{\alpha/2, n_1-1, n_2-1}$
$H_0: \sigma_1^2 \neq \sigma_2^2$		$F_0 > F_{1-\alpha/2, n_1-1, n_2-1}$
$H_0: \sigma_1^2 = \sigma_2^2$	$F_0 = \frac{S_2^2}{S_1^2}$	
$H_0: \sigma_1^2 < \sigma_2^2$		$F_0 > F_{\alpha, n_1-1, n_2-1}$
$H_0: \sigma_1^2 = \sigma_2^2$	$F_0 = \frac{S_1^2}{S_2^2}$	
$H_0: \sigma_1^2 > \sigma_2^2$		$F_0 > F_{\alpha, n_1-1, n_2-1}$

مثال

مثالی از آزمون برابری دو واریانس

- مهندسی واریانس ذاتی (خطای رندوم) دو سیستم اندازه گیری را بررسی می کند.
- تعداد نمونه های $n_1 = 11$ و $n_2 = 10$ و واریانس ها $S_1^2 = 14.5$ و $S_2^2 = 10.8$.
- وی گمان می کند که سیستم ۱ واریانس بیشتری نسبت به سیستم ۲ دارد.
- با طرح ریزی یک آزمون فرض مناسب وی را یاری کنید؟
- حل: ابتدا آزمون فرضی به صورت زیر طرح می کنیم؛

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

$$H_1 : \sigma_1^2 > \sigma_2^2$$

- سپس نسبت واریانس ها را محاسبه می کنیم؛

$$F_0 = \frac{S_1^2}{S_2^2} = \frac{14.5}{10.8} = 1.34$$

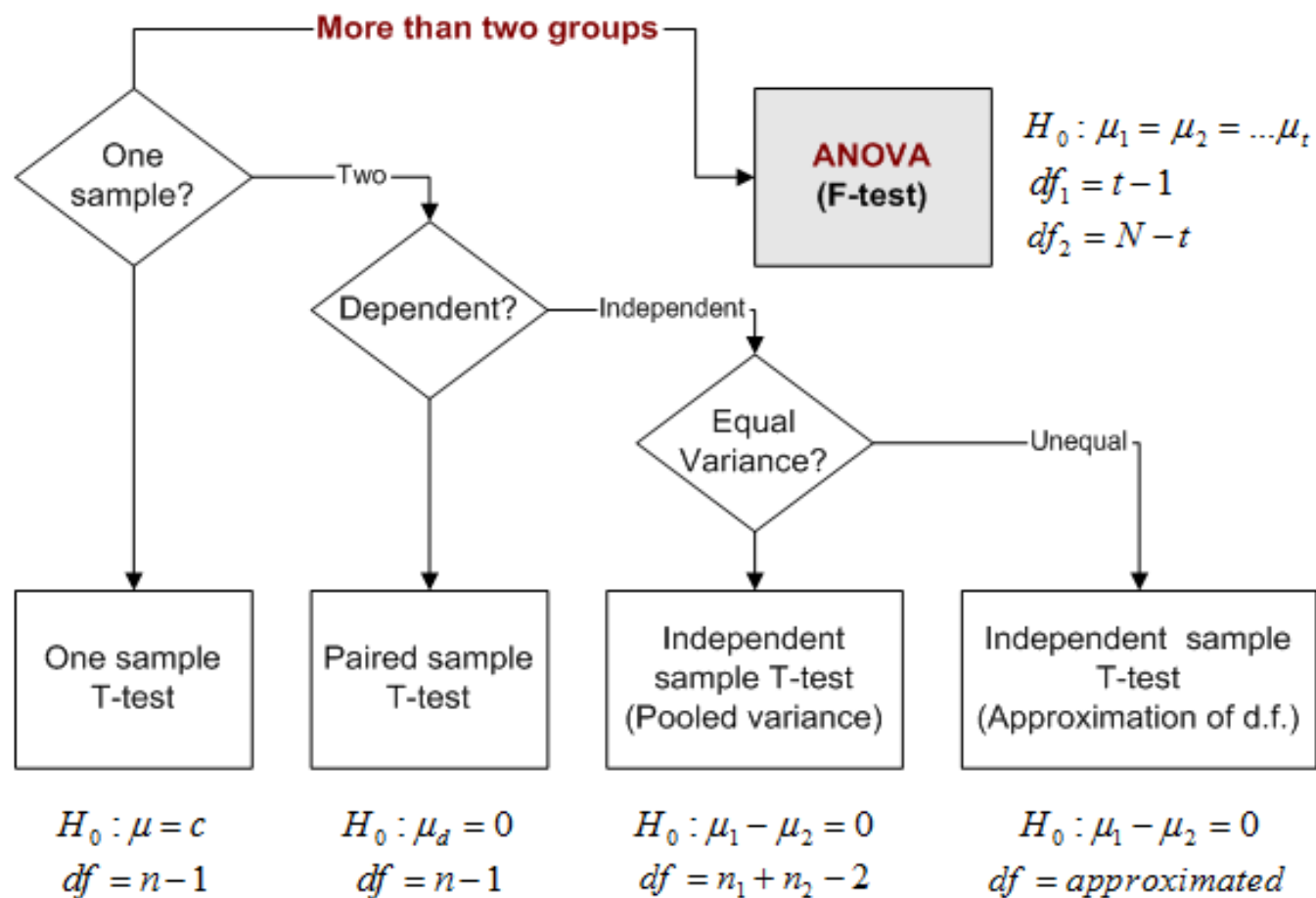
		Degrees of Freedom for the Numerator (ν_1)																		
ν_2	ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
		161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.39	19.40	19.41	19.43	19.45	19.46	19.47	19.48	19.49	19.50	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.77	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.97	5.95	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.72	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.07	4.05	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.65	3.63	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.33	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.14	4.28	3.89	3.66	3.51	3.40	3.32	3.26	3.21	3.17	3.15	3.10	3.04	2.97	2.94	2.90	2.86	2.83	2.79	2.75
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.96	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.83	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.56	2.49	2.45	2.41	2.37	2.33	2.29	2.24
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.40	2.36	2.32	2.28	2.24	2.20	2.15
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.33	2.29	2.25	2.21	2.17	2.13	2.08
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.27	2.23	2.19	2.15	2.11	2.07	2.02
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.22	2.18	2.14	2.10	2.06	2.01	1.96
17	4.45	3.59	3.20	2.97	2.81	2.70	2.62	2.55	2.49	2.45	2.38	2.31	2.23	2.17	2.13	2.09	2.05	2.01	1.96	1.91
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.13	2.09	2.05	2.01	1.96	1.91	1.86
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.10	2.06	2.02	1.98	1.93	1.88	1.83
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.06	2.02	1.98	1.93	1.88	1.83	1.78
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.04	2.00	1.96	1.91	1.86	1.81	1.76
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.01	1.97	1.93	1.88	1.83	1.78	1.73

✓ مقدار F در آزمون فرض مربوطه با ۹۵ درصد اطمینان (۵ درصد خطا): $F_{0.05,10,9} = 3.14$

✓ از آنجاکه $F_0(1.34) < F_{0.05,10,9}(3.14)$ فرض صفر قابل رد نیست (پذیرفته می شود).

✓ گمان مهندس در مورد وجود اختلاف معنی دار در واریانس (خطای رندوم) ذاتی دو دستگاه صحیح نیست.

اختلاف واریانس های دو دستگاه معنی دار نیست.



بازه اطمینان جهت تخمین میانگین یک جامعه آماری

- جامعه آماری نرمال با واریانس معلوم یا جامعه آماری با اندازه بزرگ ✓
تخمین میانگین جامعه آماری با درصد اطمینان $(1 - \alpha)\%$ ؛

❖ فرم بازه ای:

$$\bar{x} - Z_{\alpha/2} \frac{\sigma_n}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\alpha/2} \frac{\sigma_n}{\sqrt{n}}$$

$Z_{\alpha/2}$: از طریق جدول توزیع نرمال

σ_n : انحراف از معیار

❖ فرم تکرانسی:

$$\mu = \bar{x} \pm Z_{\alpha/2} \frac{\sigma_n}{\sqrt{n}}$$

بازه اطمینان جهت تخمین میانگین یک جامعه آماری

- جامعه آماری محدود با واریانس نمونه معلوم ✓
تخمین میانگین جامعه آماری با درصد اطمینان $(1 - \alpha)\%$ ؛
- ❖ فرم بازه ای:

$$\bar{x} - t_{\alpha/2, n-1} \frac{S_n}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2, n-1} \frac{S_n}{\sqrt{n}}$$

$t_{\alpha/2, n-1}$: از طریق جدول توزیع t

S_n : انحراف از معیار نمونه

❖ فرم تکرانسی:

$$\mu = \bar{x} \pm t_{\alpha/2, n-1} \frac{S_n}{\sqrt{n}} \quad (1 - \alpha)\%$$

مثال

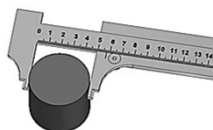
مثال: بازه اطمینان

می دانیم انحراف معیار قطر میله های فولادی $\sigma = 0.001 \text{ in}$ است. متوسط قطر یک نمونه تصادفی مشتمل بر 10 میله را برابر 1.1525 in بدست آورده ایم.

مطلوب است یافتن بازه اطمینان 90% برای اندازه میانگین واقعی قطر این میله ها؟

✓ حل: مجموعه آماری دارای توزیع نرمال با واریانس معلوم است:

$$\mu = \bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad (90\%)$$



✓ از طریق جدول توزیع نرمال برای فرض صفر در سطح 90%:

$$Z_{0.05} = 1.645$$

✓ بنابراین؛

$$\mu = 0.2545 \pm 1.645 \left(\frac{0.001}{\sqrt{10}} \right) (90\%) \rightarrow 1.1520 < \mu < 1.1530$$

بازه اطمینان جهت مقایسه دو جامعه آماری

▪ دو جامعه آماری نرمال یا به اندازه کافی بزرگ

✓ تعداد نمونه ها، میانگین های نمونه و واریانس های دو جامعه آماری معلوم است.

✓ اختلاف میانگین های دو جامعه با درصد اطمینان $(1 - \alpha)\%$ ؛

❖ فرم بازه ای:

$$(\bar{y}_1 - \bar{y}_2) - Z_{\alpha/2} \sigma_p \leq \mu_1 - \mu_2 \leq (\bar{y}_1 - \bar{y}_2) + Z_{\alpha/2} \sigma_p$$

$$\sigma_p = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$Z_{\alpha/2}$: از طریق جدول توزیع نرمال

❖ فرم تلرانسی:

$$\mu_1 - \mu_2 = (\bar{y}_1 - \bar{y}_2) \pm Z_{\alpha/2} \sigma_p$$

بازه اطمینان جهت مقایسه دو جامعه آماری

- دو جامعه آماری با واریانس های نامعلوم (دو جامعه آماری محدود)
- ✓ تعداد نمونه ها، میانگین های نمونه و واریانس های نمونه دو جامعه آماری معلوم است.
- ✓ اختلاف میانگین های دو جامعه با درصد اطمینان $1 - \alpha$ ؛
- ❖ فرم بازه ای:

$$(\bar{y}_1 - \bar{y}_2) - t_{\alpha/2, n_1+n_2-2} \hat{S}_P \leq \mu_1 - \mu_2 \leq (\bar{y}_1 - \bar{y}_2) + t_{\alpha/2, n_1+n_2-2} \hat{S}_P$$

$$\hat{S}_P = \left(\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \right) \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

از طریق جدول توزیع t : $t_{\alpha/2, n_1+n_2-2}$

❖ فرم تفراسی:

$$\mu_1 - \mu_2 = \bar{y}_1 - \bar{y}_2 \pm t_{\alpha/2, n_1+n_2-2} \hat{S}_P$$

مثال: بازه اطمینان جهت مقایسه دو جامعه آماری

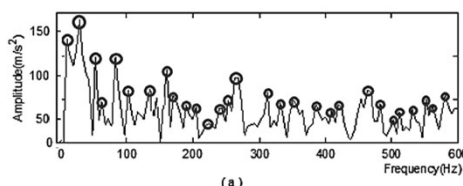
مثال

- دو گروه مجزا، جهت تخمین دامنه شتاب در تستهای ارتعاشی نتایج زیر را بدست آورده اند:

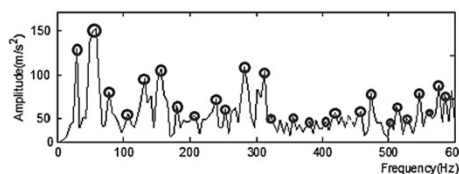
مشخصات	گروه ۱	گروه ۲
تعداد تست	90	100
میانگین	76 m/s^2	81 m/s^2
انحراف از معیار	8.2 m/s^2	7.6 m/s^2

❑ بازه اطمینان 98% را برای اختلاف بین میانگین های دو دسته نتایج بیابید؟

❑ در سطح 98% آیا اختلاف میانگین های دو جامعه آماری معنی دار است؟



(a)



(b)

مثال

مثال: بازه اطمینان جهت مقایسه دو جامعه آماری

- حل: الف) دو جامعه آماری به اندازه کافی بزرگ است و می توان نرمال فرض کرد:

$$\mu_1 - \mu_2 = (\bar{y}_1 - \bar{y}_2) \pm Z_{\alpha/2} \sigma_p$$

$$\bar{y}_1 - \bar{y}_2 = 76 - 81 = -4.8$$

$$\sigma_p = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{(8.2)^2}{90} + \frac{(7.6)^2}{100}} = 1.15$$

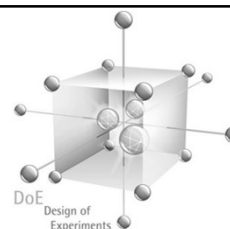
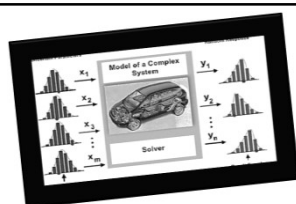
- درصد اطمینان 98% معادل با خطای $\alpha = 0.02$ است. از طریق جدول خواهیم داشت؛

$$Z_{\alpha/2} = Z_{0.01} = 2.33$$

$$\mu_1 - \mu_2 = -4.8 \pm 2.7$$

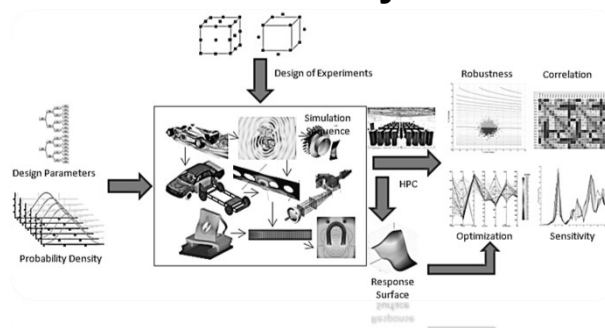
$$-7.5 \leq \mu_1 - \mu_2 \leq -2.2$$

- میانگین نتایج گروه 2 حداقل به اندازه $2.2m/s^2$ بزرگتر از میانگین نتایج گروه 1 است.
- ب) چون بازه اطمینان شامل صفر نیست، فرض برابری میانگین ها با 98% اطمینان رد می شود.
- اختلاف میانگین های نتایج دو گروه معنی دار است.



تحلیل واریانس: طرح تک عاملی

Analysis of Variance (ANOVA): One Way



Example: Instagram Ad Frequency

- Suppose that you are a data scientist at Instagram, and you are interested in running an experiment to learn about how user engagement is influenced by ad frequency.
- Currently users see an ad every 8 posts in their social feed, but, in order to increase ad revenue, your manager is pressuring your team to show an ad every 5 posts
 - Condition 1 – 7:1 Ad Frequency
 - Condition 2 – 4:1 Ad Frequency
- You are justifiably nervous about this change and you worry that this will substantially decrease user engagement and hurt the overall user experience.
- The metric of interest you choose to optimize for is μ = average session time (where y = the length of time a user engages with the app – in minutes). ↑



$$H_0: \mu_1 \leq \mu_2 \text{ vs. } H_A: \mu_1 > \mu_2$$

• The data summaries are:

$$- n_1 = 500, \hat{\mu}_1 = \bar{y}_1 = 4.92, \hat{\sigma}_1 = s_1 = 0.96$$

$$- n_2 = 500, \hat{\mu}_2 = \bar{y}_2 = 3.05, \hat{\sigma}_2 = s_2 = 0.99$$

↑

↑

↑

your boss
believes this

you believe this

$$\bullet \text{ F-test: } t = \frac{\sigma_1^2}{\sigma_2^2} = \frac{0.96^2}{0.99^2} = 0.938$$

$$\text{p-value} = P(T \leq 0.938) + P(T \geq 1/0.938) \text{ where } T \sim F_{\substack{\uparrow \\ n_1-1}, \substack{\uparrow \\ n_2-1}} = 0.4720$$

This p-value is much larger than any ordinary significance level so we do not reject $H_0: \sigma_1^2 = \sigma_2^2$. And so we continue with the student t-test.

$$\bullet \text{ Student's t-test: } \hat{\sigma}^2 = \frac{499 \times 0.96^2 + 499 \times 0.99^2}{998} = 0.9793^2$$

$$t = \frac{4.92 - 3.05}{0.9793 \sqrt{\frac{1}{500} + \frac{1}{500}}} = 30.1$$

where $T \sim t_{(998)}$

$$\text{p-value} = P(T \geq 30.1) \\ = 1.84 \times 10^{-142} \\ \approx 0$$

this is much smaller than a typical significance level and so we reject $H_0: \mu_1 \leq \mu_2$ and conclude that increasing ad frequency significantly reduces average session duration



مدل میانگین ها (Means Model)

- فرض کنید a سطح از یک عامل را در یک آزمایش با یکدیگر مقایسه کنیم:
داده های مربوط به یک آزمایش یک عاملی

تیمار (سطح)	مشاهدات				مجموع	میانگین (μ_i)
۱	y_{11}	y_{12}	...	y_{1n}	$y_{1.}$	$\bar{y}_{1.}$
۲	y_{21}	y_{22}	...	y_{2n}	$y_{2.}$	$\bar{y}_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	...	y_{an}	$y_{a.}$	$\bar{y}_{a.}$
					$y_{..}$	$\bar{y}_{..}$

- مدل میانگین ها (Means Model)

$$y_{ij} = \mu_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, n \end{cases}$$

y_{ij} : مشاهده ij
 μ_i : میانگین سطح i عامل
 ε_{ij} : مؤلفه خطای تصادفی

مدل اثرات (Effect Model)

- مدل میانگین ها با استفاده از رابطه $\mu_i = \mu + \tau_i$ به صورت زیر قابل بیان است:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, n \end{cases}$$

μ : میانگین کل
 τ_i : اثر سطح i

به مدل حاصل "مدل اثرات" گفته می شود.

- مؤلفه های خطای تصادفی دارای میانگین صفر و واریانس σ^2 هستند؛

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

$$\mu_i \sim N(\mu + \tau_i, 0)$$

$$y_{ij} \sim N(\mu + \tau_i, \sigma^2)$$



$$y_{ij} = \mu_i + \varepsilon_{ij}$$

$$\mu_i = \mu + \tau_i$$

تحلیل واریانس و آزمون فرض

- هدف: انجام آزمون های فرض برای مقایسه میانگین های حاصل از سطوح مختلف است.
- فرض: متغیرهای تصادفی نرمال و مستقل با واریانس یکسان.
- آزمون فرض برای مقایسه a سطح از یک عامل:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_a$$

$$H_1: \mu_i \neq \mu_j \text{ (زوج } (i,j) \text{ حداقل به ازای یک زوج)}$$

داده های مربوط به یک آزمایش یک عاملی

تیمار (سطح)	مشاهدات				مجموع	میانگین (μ_i)
1	y_{11}	y_{12}	...	y_{1n}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	...	y_{2n}	$y_{2.}$	$\bar{y}_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	...	y_{an}	$y_{a.}$	$\bar{y}_{a.}$
					$y_{..}$	$\bar{y}_{..}$

تجزیه مجموع مربعات

- تغییرپذیری کل (مجموع مربعات کل):
 - واژه تحلیل واریانس از تجزیه تغییرپذیری کل به دو مولفه گرفته شده است.
- $$SS_T = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2$$
- ✓ a : تعداد سطوح عامل
 - ✓ n : تعداد تکرار
 - ✓ $N = an$: تعداد کل آزمایش ها
 - اگر SS_T بر درجه آزادی یعنی $N - 1$ تقسیم شود، واریانس نمونه کل را نتیجه می دهد.

$$\sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = n \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2$$

می توان اثبات کرد:

بنابراین می توان نوشت:

$$SS_T = SS_{Treatments} + SS_E$$

تجزیه مجموع مربعات

$$SS_T = SS_{Treatments} + SS_E$$

SS_T : مجموع مربعات کل

$SS_{Treatments}$: مجموع مربعات تیمارها

SS_E : مجموع مربعات خطا

$$SS_T = \sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 - \frac{y_{..}^2}{N}$$

$$N = an$$

$$SS_{Treatments} = \frac{1}{n} \sum_{i=1}^a y_{i.}^2 - \frac{y_{..}^2}{N}$$

مجموع مربعات خطا از تفاضل SS_T و $SS_{Treatments}$ قابل محاسبه است:

$$SS_E = SS_T - SS_{Treatments}$$

تجزیه مجموع مربعات

▪ درجات آزادی مجموع مربعات:

SS_T دارای $N-1$ درجه آزادی است.

$SS_{Treatments}$ دارای $a-1$ درجه آزادی است.

SS_E دارای $N-a$ درجه آزادی است.

$$a(n-1) = an - a = N - a$$

▪ واریانس های سطوح و خطا:

$$MS_{Treatment} = \frac{SS_{Treatment}}{a-1}$$

$$MS_E = \frac{SS_E}{N-a}$$

تجزیه مجموع مربعات

- مولفه های خطای تصادفی دارای میانگین صفر و واریانس σ^2 هستند.
- بنابراین واریانس خطاها (MS_E) σ^2 را برآورد می کند.
- اگر میانگین سطوح برابر باشد ($\mu_i = \mu_j$)؛
- آنگاه واریانس سطوح ($MS_{Treatments}$) برابر MS_E یعنی σ^2 است.
- اگر میانگین سطوح برابر نباشد ($\mu_i \neq \mu_j$)؛
- آنگاه واریانس سطوح ($MS_{Treatments}$) بزرگتر MS_E یعنی σ^2 است.
- لذا آزمون برابری میانگین سطوح می تواند با مقایسه $MS_{Treatments}$ و MS_E انجام گردد.

$$F_0 = \frac{MS_{Treatments}}{MS_E}$$

آزمون مقایسه میانگین ها با روش تحلیل واریانس

- آزمون مقایسه میانگین ها:
 $H_0: \mu_1 = \mu_2 = \dots = \mu_a$
 $H_1: \mu_i \neq \mu_j \text{ (} i, j \text{ زوج)}$ حداقل به ازای یک زوج

جدول تحلیل واریانس برای مدل یک عاملی با اثرات ثابت				
F_0	میانگین مربعات	درجه های آزادی	مجموع مربعات	منبع تغییرات
$F_0 = \frac{MS_{Treatments}}{MS_E}$	$MS_{Treatments}$	$a - 1$	$SS_{Treatments}$	تیمارها
	MS_E	$N - a$	$SS_E = SS_T - SS_{Treatments}$	خطا
		$N - 1$	$SS_T = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2$	کل

- F_0 دارای توزیع F با درجه آزادی صورت $a - 1$ و مخرج $N - a$ است.
- اگر $F_0 > F_{\alpha, a-1, N-a}$ باشد، فرض H_0 رد و تفاوت معناداری بین میانگین ها وجود دارد.
- اگر $F_0 < F_{\alpha, a-1, N-a}$ باشد، فرض H_0 رد نمی شود و تفاوتی بین میانگین ها وجود ندارد.
- برای تعیین احتمال خطای رد شدن فرض H_0 می توان $p - value$ را تخمین زد.

جدول توزیع F

		F_{α, ν_1, ν_2}																			
ν_2	ν_1	Degrees of Freedom for the Numerator (ν_1)																			
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞		
Degrees of Freedom for the Denominator (ν_2)	1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3	
	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50	
	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53	
	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63	
	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37	
	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67	
	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23	
	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93	
	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71	
	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54	
	11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.41	
	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30	
	13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.22	
	14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.15	
	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07	
	16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.02	
	17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96	
	18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92	
	19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88	
	20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84	
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81		
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.77		
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.75		
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73		
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71		
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69		
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67		
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65		
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64		
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62		
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.53		
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39		
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25		
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.08		

مثال

مثال: آزمون مقایسه میانگین ها با روش تحلیل واریانس

- بررسی اثر قدرت فرکانس رادیویی (RF) در چهار سطح و پنج تکرار بر نرخ حک کاری پلازما بر روی مدارهای مجتمع.



قدرت RF (W)	نرخ حک کاری مشاهده شده					مجموع $\sum y_i$	میانگین $\bar{y}_i$
	۱	۲	۳	۴	۵		
۱۶۰	۵۷۵	۵۴۲	۵۳۰	۵۳۹	۵۷۰	۲۷۵۶	۵۵۱/۲
۱۸۰	۵۶۵	۵۹۳	۵۹۰	۵۷۹	۶۱۰	۲۹۳۷	۵۸۷/۴
۲۰۰	۶۰۰	۶۵۱	۶۱۰	۶۳۷	۶۲۹	۳۱۲۷	۶۲۵/۴
۲۲۰	۷۲۵	۷۰۰	۷۱۵	۶۸۵	۷۱۰	۳۵۳۵	۷۰۷/۵
						$\sum y_{..} = ۱۲۳۵۵$	$\bar{y}_{..} = ۶۱۷/۷۵$

مثال

مثال: آزمون مقایسه میانگین ها با روش تحلیل واریانس

▪ محاسبات دستی مربوط به میانگین مجموع مربعات:

$$SS_T = \sum_{i=1}^r \sum_{j=1}^s y_{ij}^2 - \frac{y_{..}^2}{N}$$

$$= (575)^2 + (542)^2 + \dots + (710)^2 - \frac{12355}{20} = 7220.9/75$$

$$SS_{Treatments} = \frac{1}{n} \sum_{i=1}^s y_{i.}^2 - \frac{y_{..}^2}{N}$$

$$= \frac{1}{5} [(2756)^2 + \dots + (3535)^2] - \frac{12355}{20} = 6687.0/55$$

$$SS_E = SS_T - SS_{Treatments}$$

$$= 7220.9/75 - 6687.0/55 = 533.9/20$$

مثال

مثال: آزمون مقایسه میانگین ها با روش تحلیل واریانس

منبع تغییرات	مجموع مربعات	درجه های آزادی	میانگین مربعات	F_0	مقدار P
قدرت RF	66,870/55	3	22290/18	$F_0 = 66/18$	< 0.01
خطا	5339/20	16	333/70		
کل	72209/75	19			

▪ با فرض $\alpha = 0.05$ توجه به $F_{\alpha, a-1, N-a}$ از طریق جدول مربوطه:

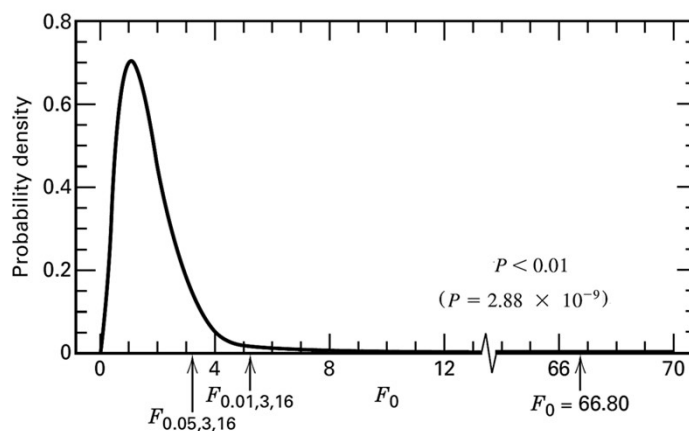
$$F_{0.05, 3, 16} = 3/24$$

$$F_0 = 66/18$$

▪ چون $F_0 > F_{0.05, 3, 16}$ فرض H_0 رد و تغییر قدرت فرکانس رادیویی در نرخ حک کاری تاثیر معنا داری دارد.

مثال

مثال: آزمون مقایسه میانگین ها با روش تحلیل واریانس



مثال

بررسی کفایت مدل

- برای بررسی کفایت مدل مورد استفاده بایستی مفروضات بررسی گردد:

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

- مفروضات مدل: ۱- نرمال بودن ۲- مستقل بودن ۳- ثابت بودن واریانس ها.

- برای بررسی مفروضات می توان باقیمانده ها (*Residuals*) را بررسی کرد:

$$e_{ij} = y_{ij} - \hat{y}_{ij}$$

- $\hat{y}_{ij}$ برآوردی از مقدار مشاهده y_{ij} است؛

$$\hat{y}_{ij} = \hat{\mu} + \hat{\tau}_i = \bar{y}_{i.}$$

بنابراین؛

$$e_{ij} = y_{ij} - \bar{y}_{i.}$$

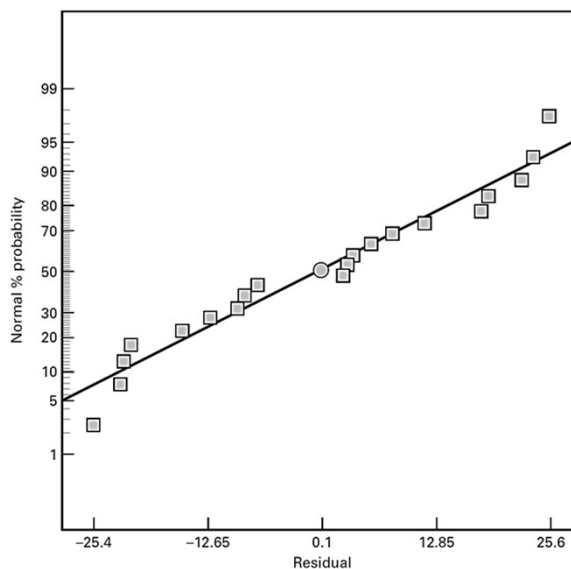
مثال

بررسی فرض نرمال بودن

تشخیص داده های پرت با
استاندارد سازی باقیمانده ها

$$d_{ij} = \frac{e_{ij}}{\sqrt{MS_E}}$$

$$-3 \leq d_{ij} \leq +3$$

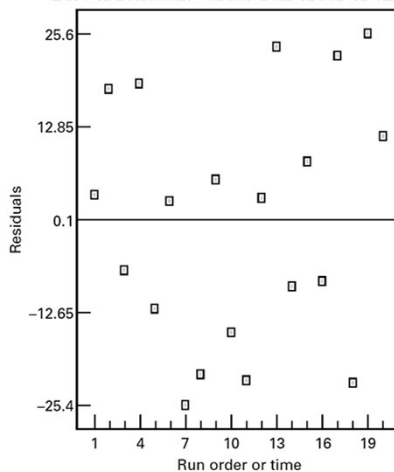


مثال

بررسی مستقل بودن مشاهدات

- اگر تغییرات در نمودار باقیمانده ها بر اساس ترتیب یا زمان آزمایش رفتار مشخص صعودی یا نزولی را دنبال نکنند مشاهدات دارای استقلال قابل قبولی هستند.

Plot of residuals versus run order or time



مثال

بررسی ثابت بودن واریانس مشاهدات

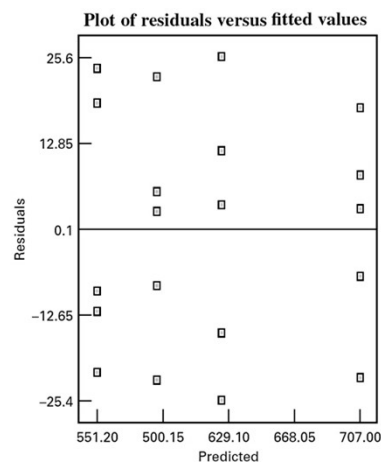
- اگر در نمودار باقیمانده ها بر اساس مقادیر برازش شده (مقادیر مورد انتظار در هر سطح $\hat{y}_{ij}$ یا میانگین هر سطح $(\bar{y}_{i.})$) تغییرات مشابه باشد، واریانس ها برابرند.
- برای بررسی برابری واریانس ها می توان از آزمون های فرض نیز استفاده کرد.

- در صورت عدم برابری واریانس ها می توان از برخی تبدیلات استفاده نمود؛

square root transformation $y_{ij}^* = \sqrt{y_{ij}}$ or $y_{ij}^* = \sqrt{1 + y_{ij}}$

logarithmic transformation $y_{ij}^* = \log y_{ij}$

arcsin transformation $y_{ij}^* = \arcsin \sqrt{y_{ij}}$



بازه اطمینان میانگین های سطوح مختلف

✓ تخمین میانگین های هر سطح با درصد اطمینان $(1 - \alpha)\%$ ؛

❖ فرم بازه ای:

$$\bar{y}_{i.} - t_{\alpha/2, N-a} \sqrt{\frac{MS_E}{n}} \leq \mu_i \leq \bar{y}_{i.} + t_{\alpha/2, N-a} \sqrt{\frac{MS_E}{n}}$$

$t_{\alpha/2, N-a}$: از طریق جدول توزیع t

MS_E : میانگین مربعات خطا

n : تعداد تکرار

❖ فرم تکراری:

$$\mu = \bar{y}_{i.} \pm t_{\alpha/2, N-a} \sqrt{\frac{MS_E}{n}} \quad (1 - \alpha)\%$$

مثال

مثال: بازه اطمینان میانگین های سطوح مختلف

▪ مطلوب است یافتن بازه اطمینان ۹۵٪ میانگین های سطوح (RF) در مساله پلاسما؟

$$N = 20$$

$$a = 4$$

$$n = 5$$

$$\alpha = 0.05 (95\%)$$

$$MS_E = 333.70$$

$$t_{\alpha/2, N-a} = t_{0.025, 16} = 2.120$$

$$\mu = \bar{y}_{i.} \pm t_{\alpha/2, N-a} \sqrt{\frac{MS_E}{n}} \quad (1-\alpha)\%$$

RF (W)	$\bar{y}_{i.}$	S_i	95% CI
160	551.20	20.02	(533.88, 568.52)
180	587.40	16.74	(570.08, 604.72)
200	625.40	20.53	(608.08, 642.72)
220	707.00	15.25	(689.68, 724.32)



CEJRA
Center of Excellence In Design, Robotics & Automation

Thanks for your attention

