**جلسه پنجم**

**مریم شمسائی**

**سرفصل**

* **فرضیه پوچ و جایگزین**
* **خطای نوع اول و خطای نوع دوم**
* **توان آزمون و حجم اثر**
* **عملیات اعمال آزمون Z و محاسبه حجم اثر در R**

**پژوهش جلسه**

**سوال پژوهشی[[1]](#footnote-1)** : سوالی که در یک پژوهش تعریف می­شود، مثال: آیا هوش در افراد با آسپرگر[[2]](#footnote-2) بالا است یا خیر؟

**فرضیه پژوهشی[[3]](#footnote-3) :** فرضیه اولیه یک تحقیق چیزی است که فرض می شود وجود دارد، مثال: در افرادی که آسپرگر دارند، هوش بالاتر است. در واقع فرضیه پژوهش یک جواب مثبت به سوال پژوهش است.

**نمونه[[4]](#footnote-4) پژوهش:** افرادی هستند که پژوهش روی آنها انجام می­شود مثال : افرادی که اتیسم از نوع آسپرگر دارند.

**مقایسه :** برای بررسی فرضیه تحقیق، نیازمند یک مقایسه هستیم. حالا سوال بعدی این است که: این مقایسه باید با چه چیزی باشد؟ یک پاسخ می­تواند جامعه افراد عادی باشد.

**انتخاب ابزار :** لازم است یک ابزار برای آن چیزی که قصد سنجش آن را داریم، تعیین کنیم.

**نکته:** به صورت کلی در این پژوهش میانگین هوش در افراد عادی برابر با 100 و انحراف استاندارد آن برابر با 16 است.

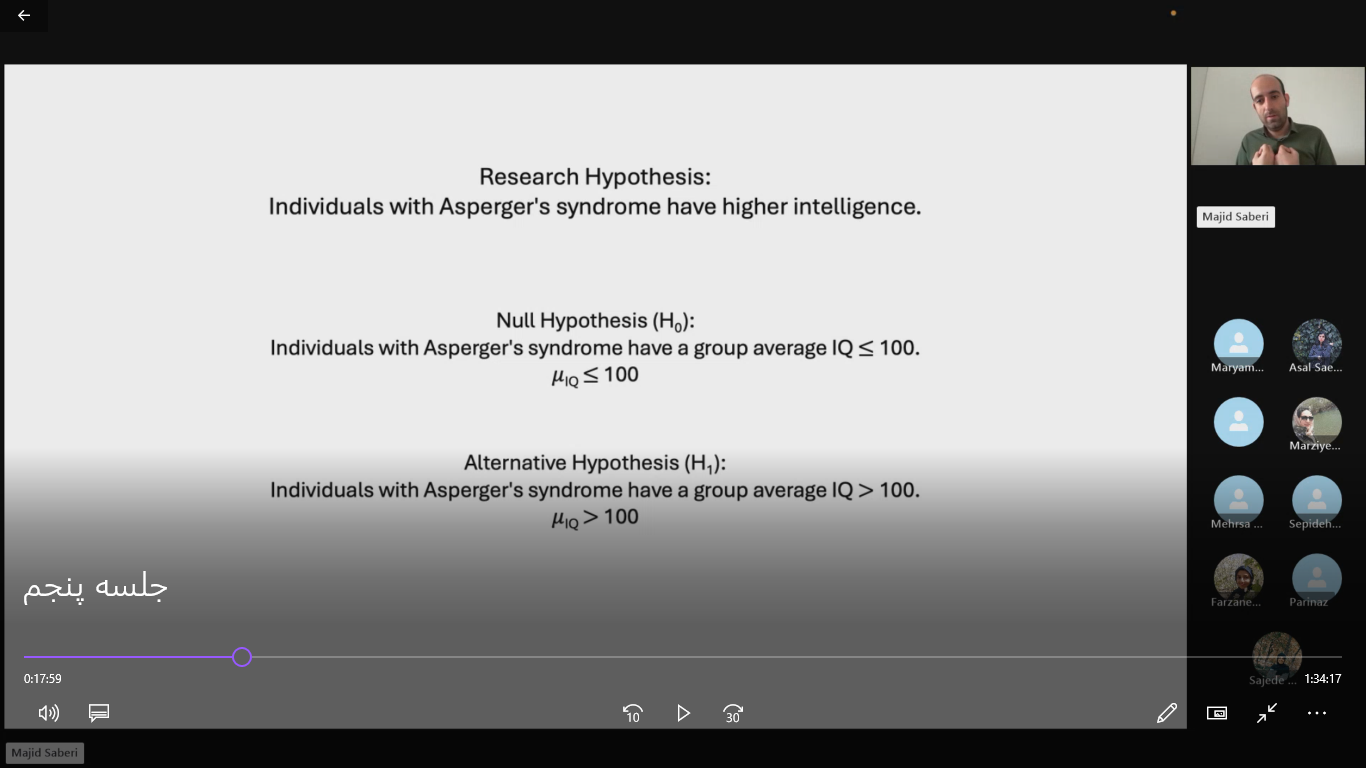
**آموزش مفاهیم**

**فرضیه پوچ [[5]](#footnote-5)و فرضیه جایگزین[[6]](#footnote-6)**

فرضیه پوچ و فرضیه جایگزین دو مفهوم مهم در آمار و تحقیقات علمی هستند. فرضیه پوچ معمولا به عنوان فرضیه صفر یا فرضیه عدم تأثیر نیز شناخته می‌شود. این فرضیه معمولا اظهار می‌کند که هیچ تفاوت یا تأثیر معنی‌داری بین دو گروه یا پدیده وجود ندارد. به عبارت دیگر، فرضیه پوچ اظهار می‌کند که هر گونه تفاوت یا تأثیر مشاهده شده به اتفاق افتاده است. فرضیه جایگزین معمولا بیانگر وجود تفاوت یا تأثیر معنی‌دار بین دو گروه یا پدیده است. به عبارت دیگر، این فرضیه بیان می‌کند که تفاوت یا تأثیر مشاهده شده نسبت به فرضیه پوچ، واقعی و قابل توجه است.

یک روش علمی برای اثبات یک فرضیه این است که فرض کنیم هرگز نتوانیم هیچ چیزی را ثابت کنیم. بنابراین، با در نظر گرفتن مخالف چیزی که می‌خواهیم، شروع به رد آن می‌کنیم و با رد کردن آن، به ما قدرت استدلال در مورد آن چیز می‌دهد. این فرض پوچ دقیقاً مخالف فرضیه مورد نظر ما است، به عبارت دیگر، این فرض اظهار می‌کند که متوسط هوش آسپرگر‌ها 100 یا کمتر از 100 است و ما تلاش می‌کنیم برای رد این فرض، مخالف آن را پیدا کنیم. فرض صفر دقیقاً برعکس فرض تحقیق است و فرض جایگزین ما، فرضیه‌ای است که ما انتظار داریم، یعنی افراد آسپرگر هوش بالای 100 دارند. بنابراین:

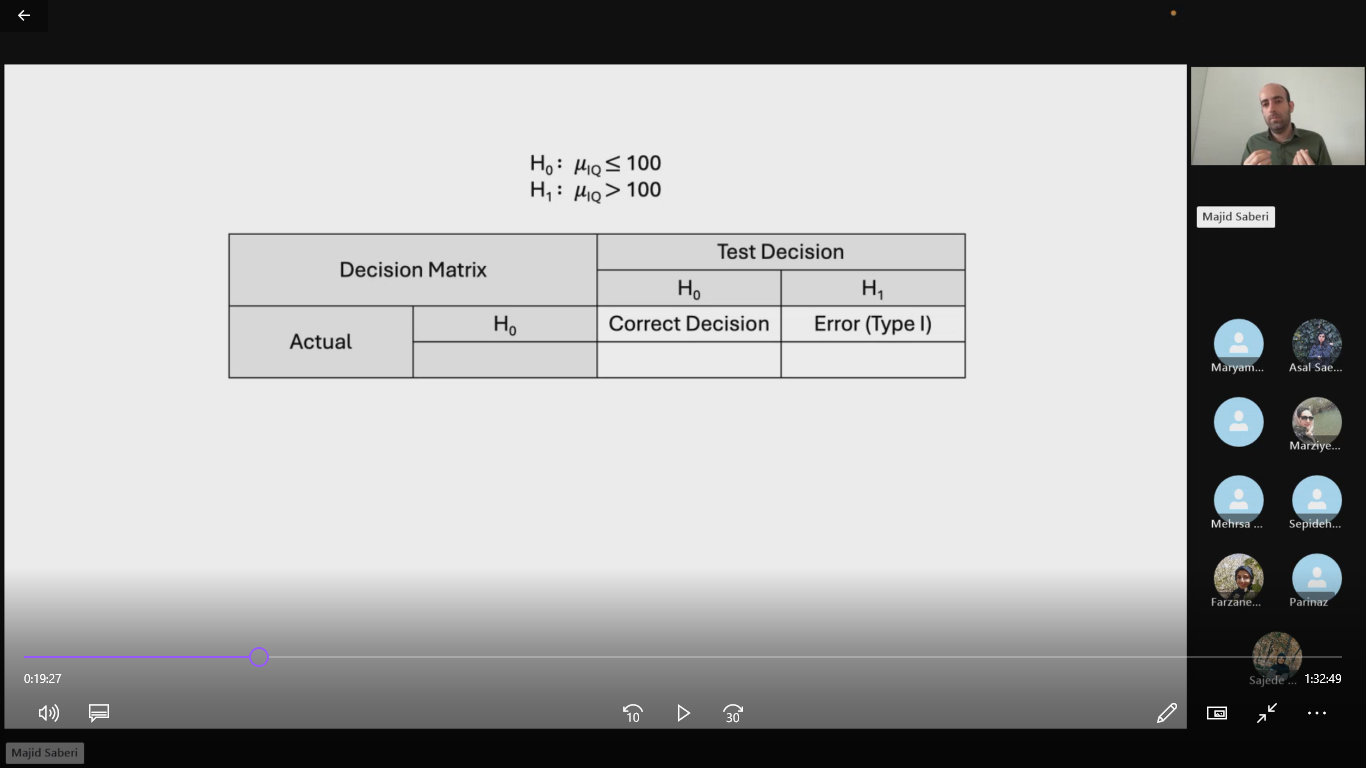
در پژوهش این جلسه مطابق شکل 1، فرض 1 (H1)مشابه فرضیه پژوهش است. فرض صفر (H0)، دقیقا مخالف فرضیه پژوهش و فرضیه 1 است و ما در تلاش هستیم تا فرض صفر را رد کنیم.



شکل 1

**خطای نوع اول**

شکل 2 را در نظر بگیرید.



شکل 2

یک چیزی ممکن است در واقعیت باشد یا نباشد و یک چیزی را در تست به آن می­رسیم. حالا فرض کنیم واقعیت ما H0  است. یعنی افرادی که آسپرگر دارند هوش کمتر از 100 دارند( هوش برابر با 100 و یا کمتر دارند). در نتیجه تست ما دو اتفاق ممکن است رخ دهد:

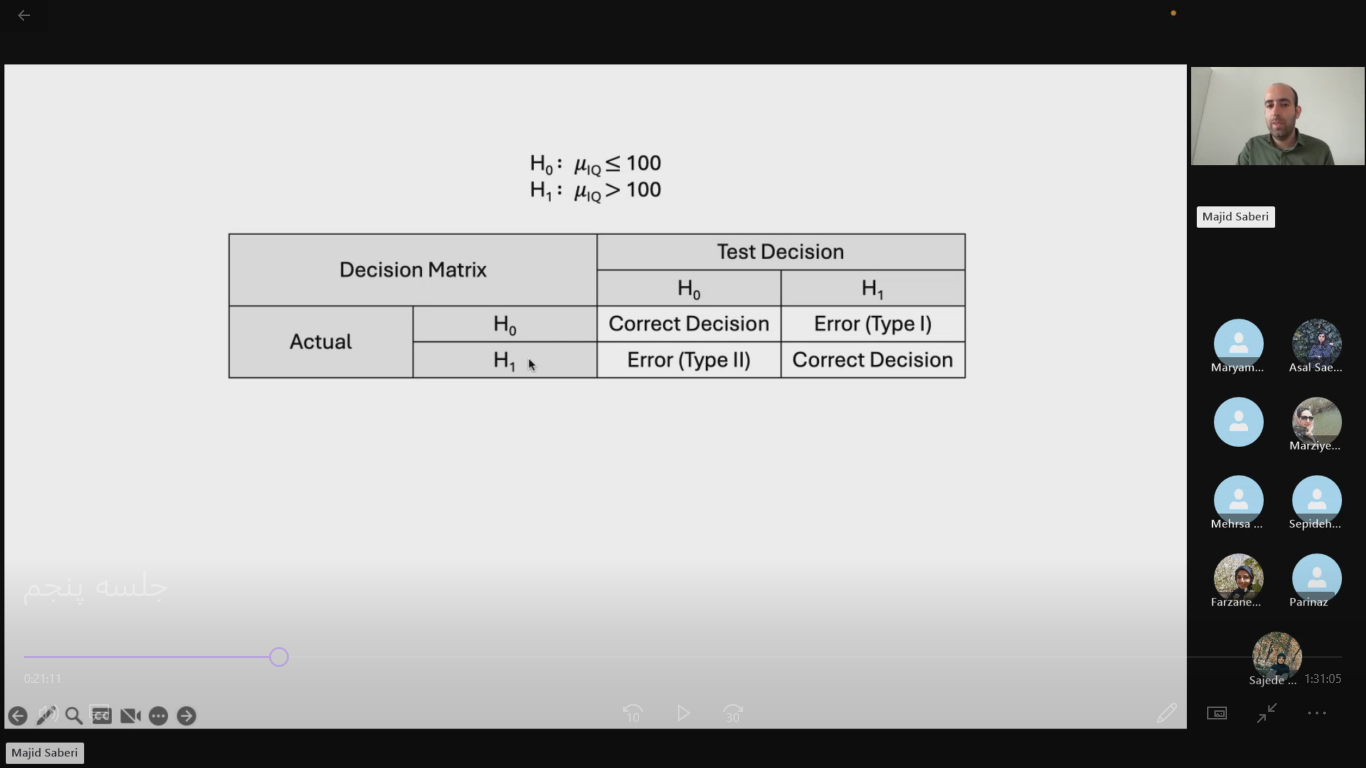
1. بله، میانگین هوش افراد آسپرگر کمتر از 100 است (برابر یا کمتر از 100 است).

2. خیر، میانگین هوش افراد آسپرگر بالاتر از 100 است(یعنی H1)

* اگر در تست نتیجه، H0 را تایید کند، تصمیم­گیری صحیح [[7]](#footnote-7) است (با فلش رنگ سبز در شکل 2 نمایش داده شده است).
* اگر که در تست نتیجه، فرض H0 را رد کند، ولی در واقعیت H0 صحیح باشد، خطای نوع اول[[8]](#footnote-8) اتفاق می­افتد (با فلش قرمز رنگ در شکل 2 نمایش داده شده است).

**خطای نوع دوم**

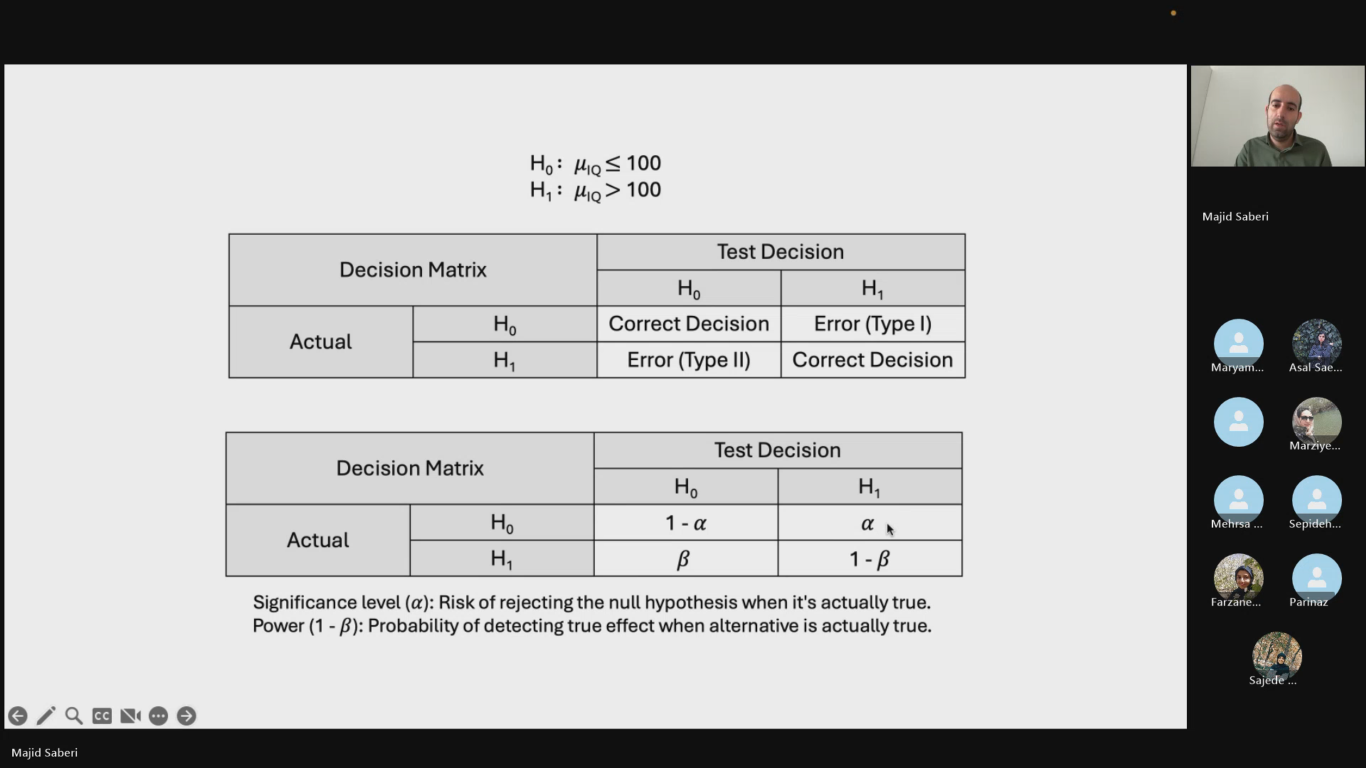
شکل 3 را در نظر بگیرید.



شکل 3

* اگر واقعیت H1 باشد و تست ما هم تایید کند که H1 برقرار است، تصمیم گیری صحیح است (با فلش رنگ سبز در شکل 3 نمایش داده شده است).
* اگر H1برقرار باشد و ما در تست به این نتیجه برسیم که H0 برقرار است در آنجا خطای نوع دوم[[9]](#footnote-9) خواهیم داشت (با فلش قرمز رنگ در شکل 3 نمایش داده شده است).

حالا ما می­توانیم برای هر کدام احتمال­­های متفاوتی اتخاذ کنیم. شکل 4 را در نظر بگیرید:



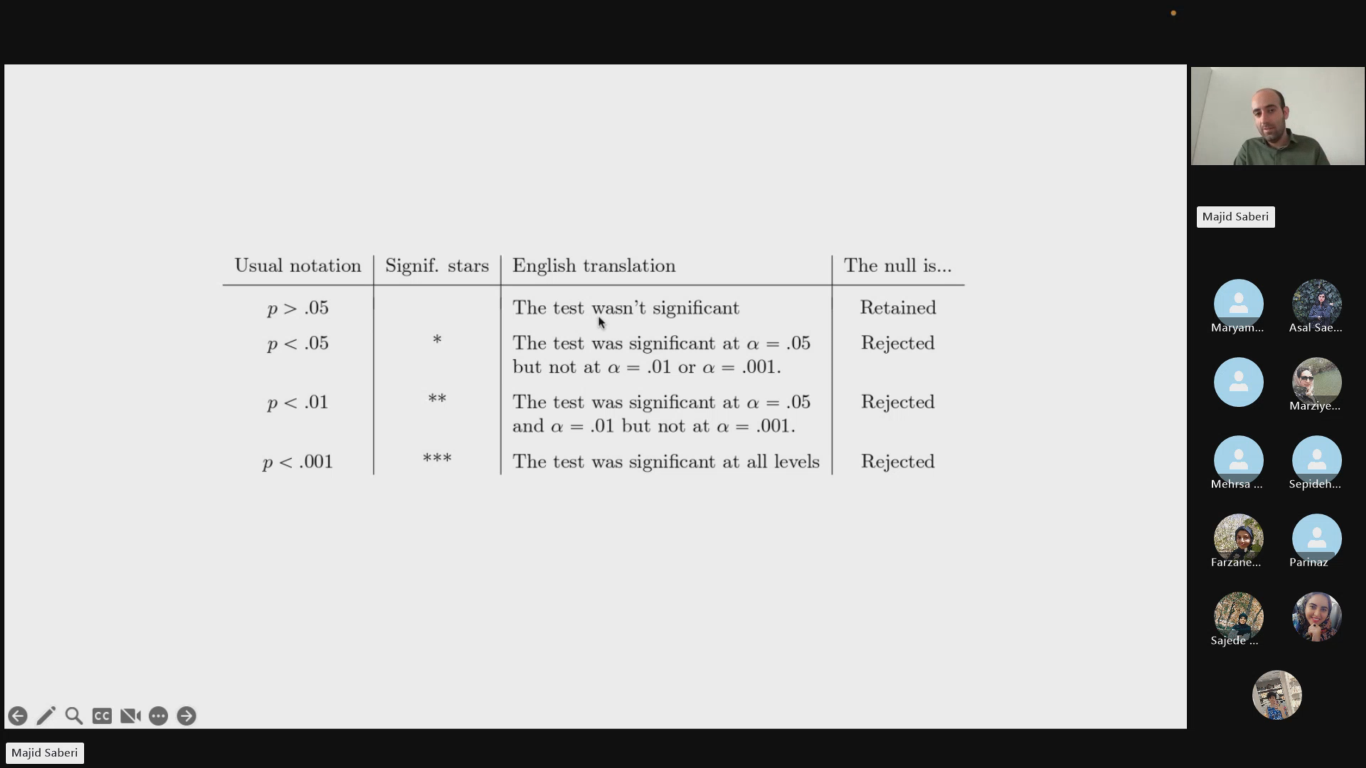
شکل 4

**ɑ**: همان سطح معناداری است. اگر ما H0 را که برقرار بود، ردش کردیم ( یعنی H1) ، درصد و احتمالی برای خطا داریم. اگر آلفا برابر با 0.25باشد، 0.25-1 برابر با ɑ - 1است یعنی 0.75

**ꞵ**: اگر H1 واقعا برقرار باشه و H0 برقرار شود، خطای بتا رخ خواهد داد. اگر بتا برابر با 0.25 باشد، ꞵ - 1برابر با 0.75که این مقدار یعنی توان [[10]](#footnote-10).

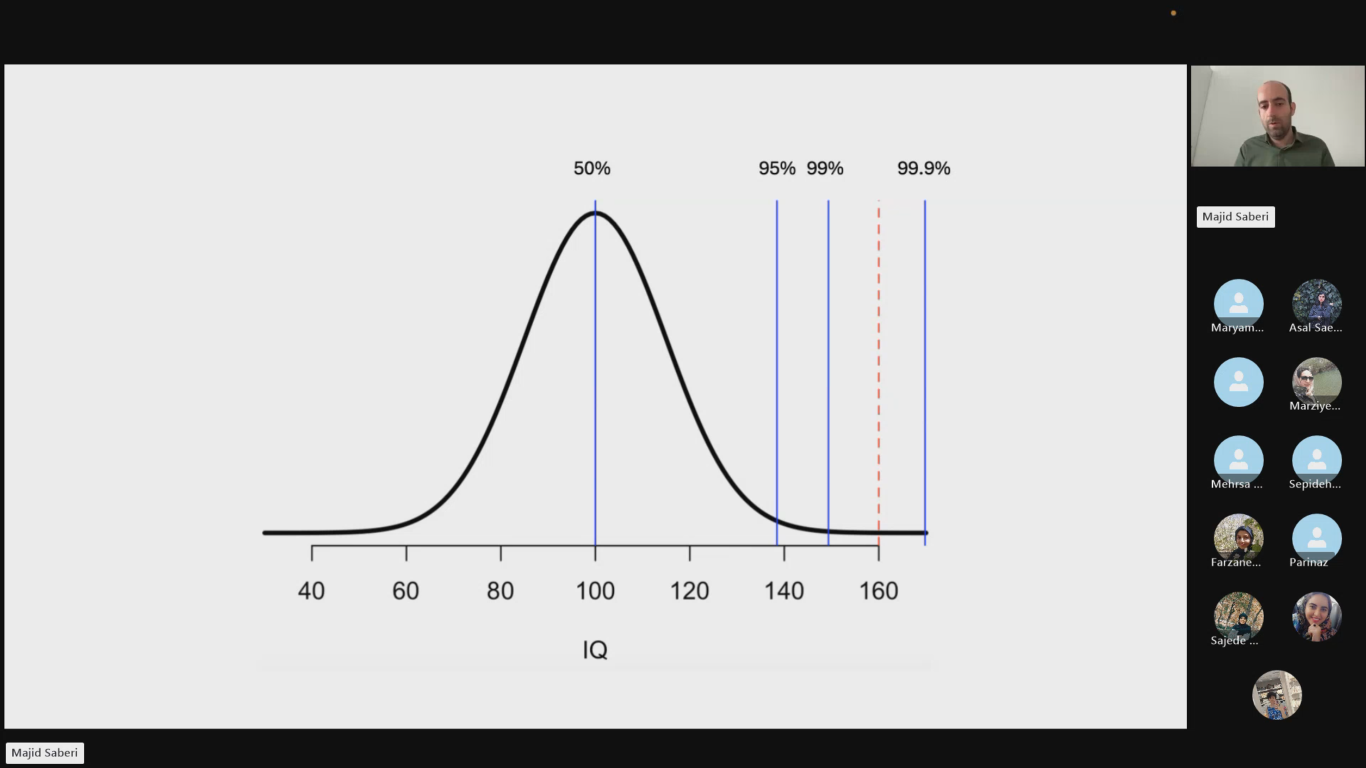
**نکته:** آلفا هر چه کوچتر و توان هرچه بزرگتر باشد، بهتر است.

**P-value:** با رد فرض صفر، ما احتمال خطای اشتباه رد کردن را ارزیابی می‌کنیم. هرچه این مقدار کمتر باشد، نتیجه رد فرض صفر با اطمینان بیشتری صورت گرفته است**.**



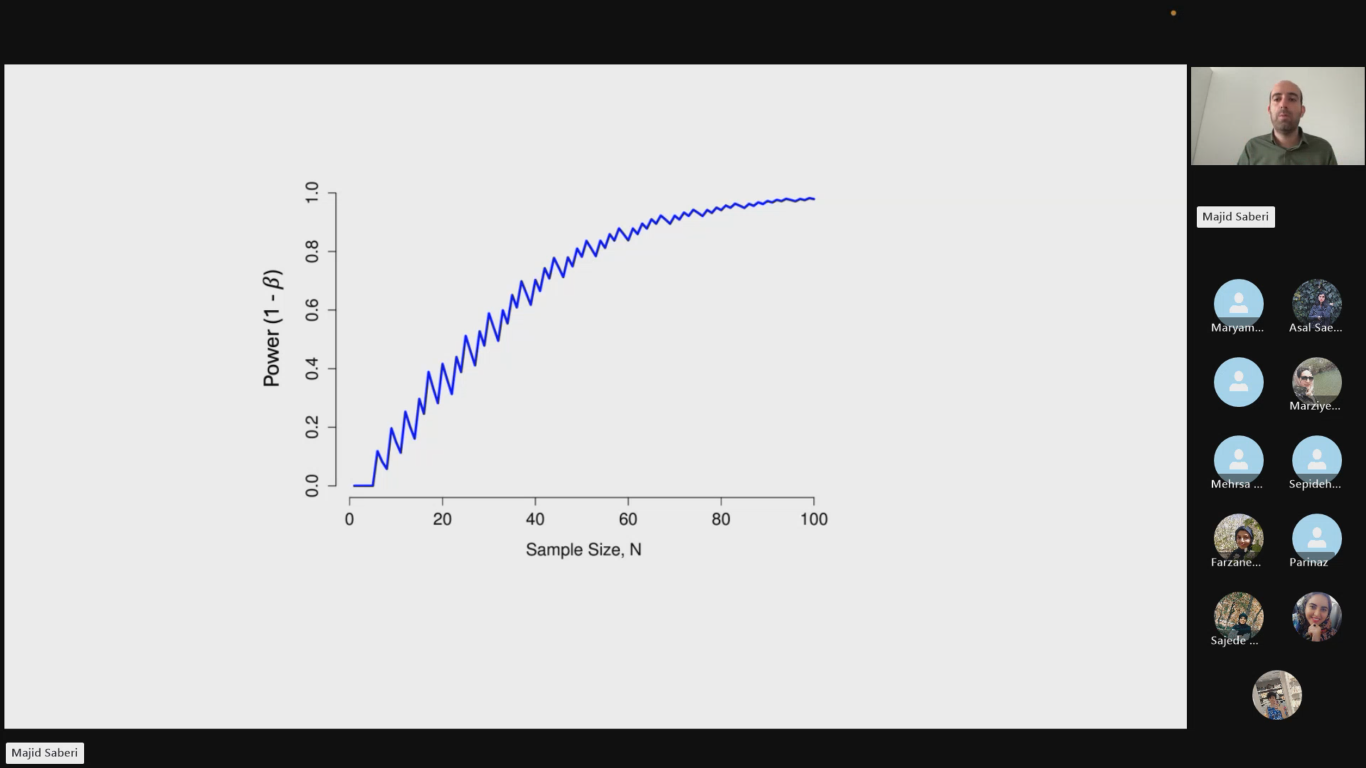
شکل 5

شکل 6 را در نظر بگیرید. این تابع توزیع هوش افراد سالم است.



شکل 6

نصف افراد دارای امتیاز هوش کمتر از 100 هستند. شکل نشان‌دهنده مقادیر مختلف مانند 95%، 99% و 99.9% است. خط قرمز به امتیاز 160 اشاره دارد. این امتیاز در سطوح 95% و 99% به صورت معنادار تلقی می‌شود، اما در سطح 99.9% معنادار نیست.



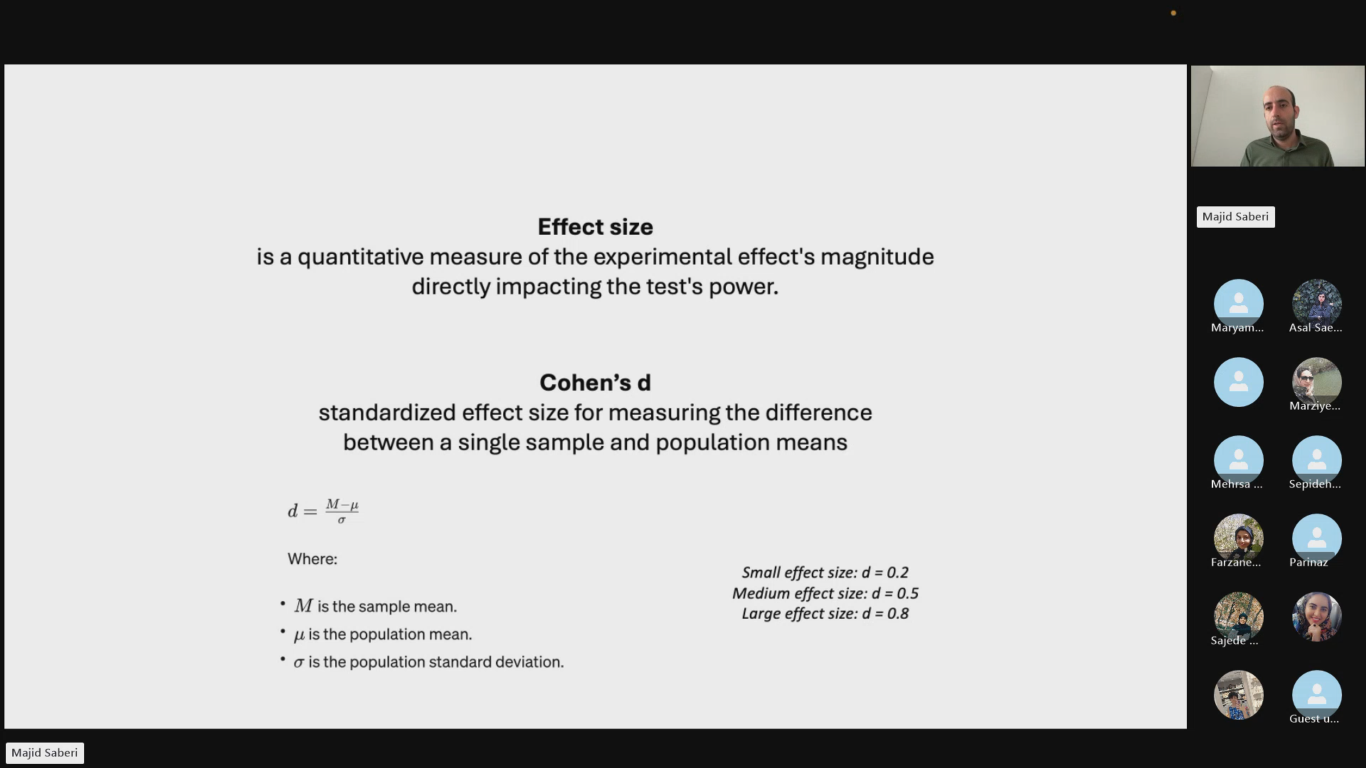
شکل 7

**نکته**: شکل 7 را در نظر بگیرید، توجه داشت که فقط سطح معناداری مهم نیست، بلکه توان نیز اهمیت دارد. با افزایش حجم نمونه، توان آزمون نیز افزایش می‌یابد.

**حجم اثر**[[11]](#footnote-11): حجم اثر یک ابزار کمی است که میزان تفاوت بین دو گروه را نشان می‌دهد. این ابزار معمولاً برای مقایسه یک گروه آسپرگر با جامعه کلی استفاده می‌شود. حجم اثر نشان‌دهنده اندازه تفاوت بین این دو گروه است. برای اندازه‌گیری حجم اثر، می‌توان از روش کوهن دی [[12]](#footnote-12)استفاده کرد که به مقایسه دو گروه کمک می‌کند و نشان می‌دهد که تفاوت بین آن‌ها چقدر است. این ابزار می‌تواند به ما کمک کند تا به دقت بیشتری بفهمیم که آیا تفاوت بین دو گروه واقعی است یا نه. از این رو، حجم اثر یک پارامتر مهم برای تفسیر نتایج است. با توجه به شکل 8:

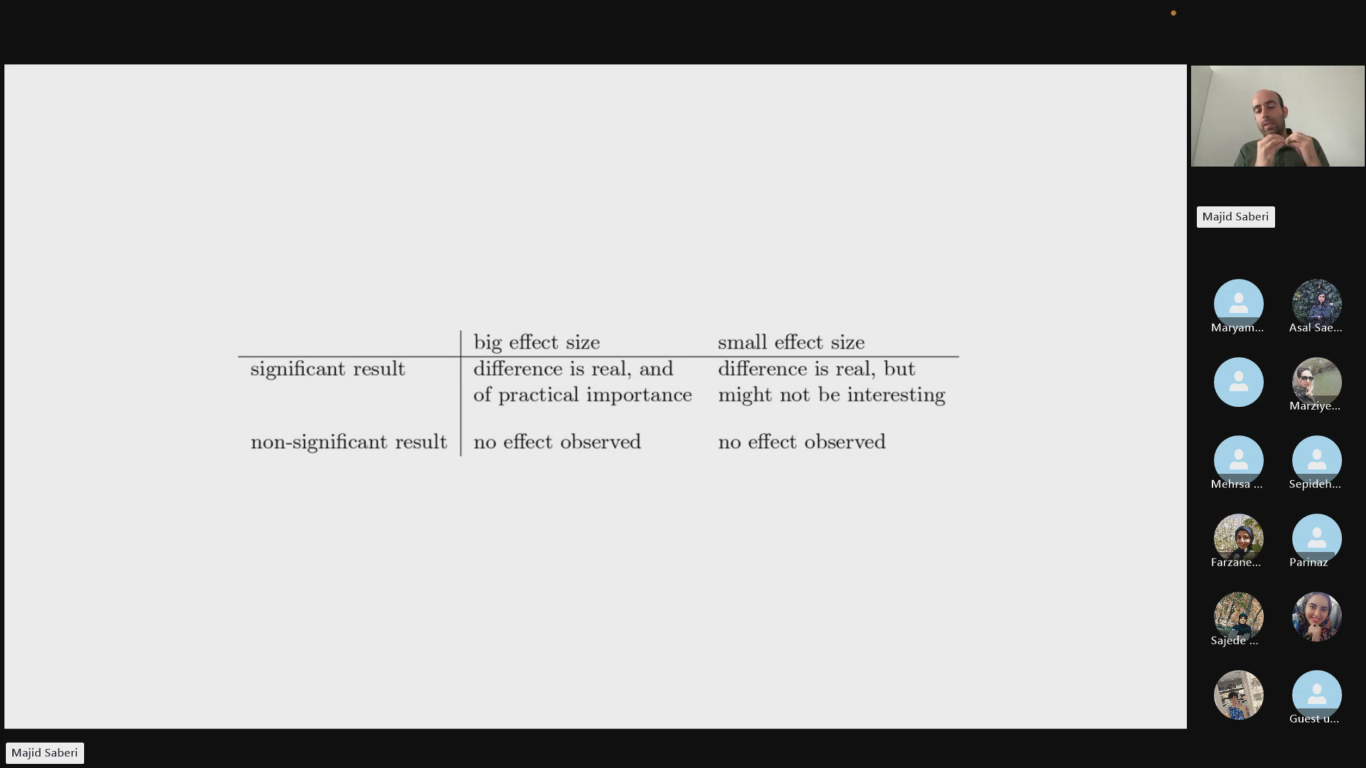
* حجم اثر کوچک: 0.2
* حجم اثر متوسط: 0.5
* حجم اثر بزرگ: 0.8

**نکته:** حجم اثر هرچه بزرگتر باشد، بهتر است.



شکل 8

شکل 9 را در نظر بگیرید. نتیجه تست ما ممکن است معنادار باشد یا نباشد. حجم اثر ما ممکن است کوچک یا بزرگ باشد. . حجم اثر با توان متناسب است.



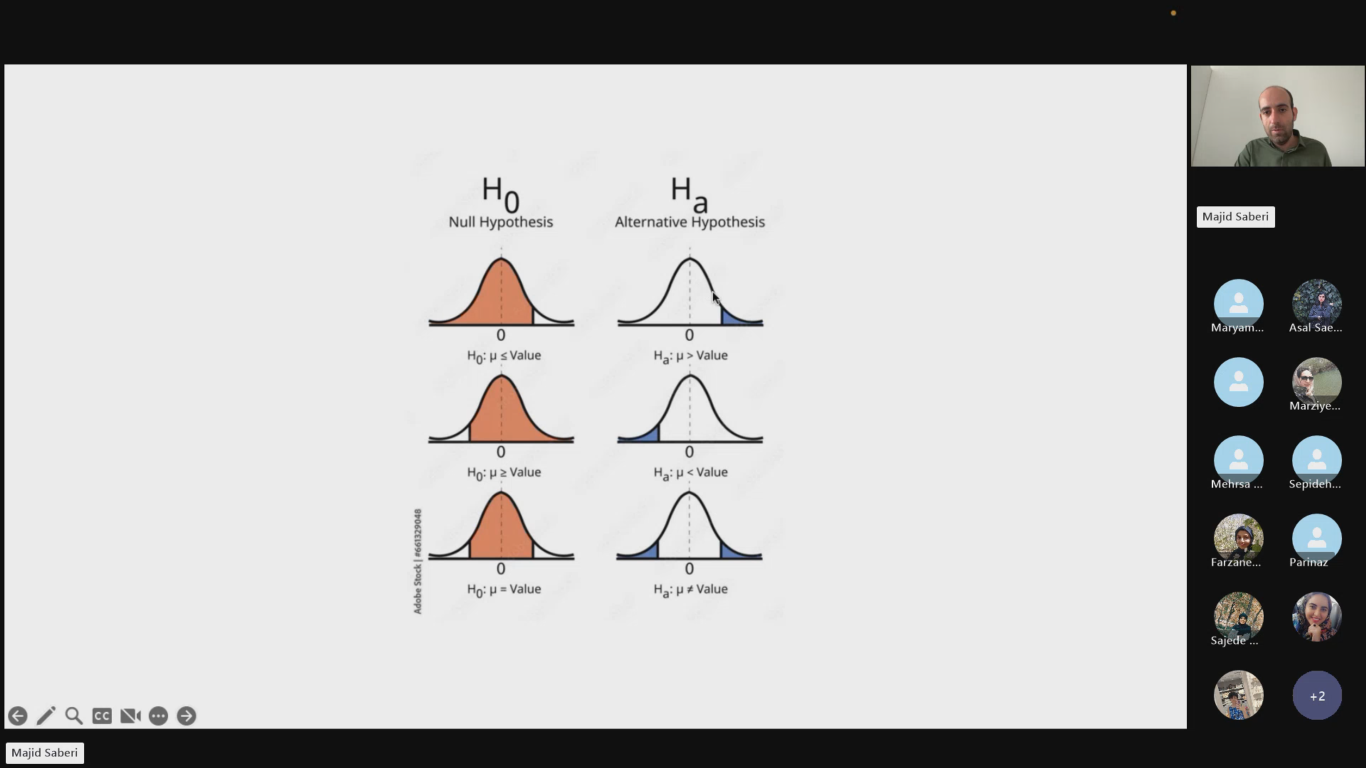
شکل 9

* اگر تفاوت معنادار بود و حجم اثر هم بزرگ بود، ما میگوییم تفاوت وجود دارد و این تفاوت واقعی و مهم است.
* اگر تفاوت معنادار بود و حجم اثر کوچک بود، ما می­گوییم تفاوت وجود دارد و واقعی است اما خیلی جذاب نیست.
* اگر تفاوت معنادار نبود دیگر هیچی مهم نیست.

**فرضیه یک طرفه و دو طرفه:**

* وقتی می­گوییم فرضیه یک طرفه یعنی همان مثالی که زدیم فرضیه 1 ما این بود که میانگین هوش افراد آسپرگر بالاتر از 100 است.
* برای فرضیه دو طرفه، فرضیه 1 ما این است که میانگین هوش افرادی که آسپرگر دارند با افراد جامعه تفاوت دارد. یعنی میانگین هوش افراد آسپرگر یا بالاتر یا پایین تر است به عبارتی برابر با 100 نیست.

نکته: در انتخاب فرضیه یک طرفه یا دو طرفه دقت کنید چون P Value را تحت تاثیر قرار می­دهد.



شکل 10

**آموزش در R**

فایل این جلسه با نام Session5-Studentversion در اختیار قرار گرفته است.

##Stats Course

#Instructor: Majid Saberi

#Session 5

#Mar 30th, 2024

##Clearing workspace

rm(list = ls())

##Loading required packages

library(TeachingDemos)

library(effsize)

##Setting the path

setwd("/Users/majid/Projects/StatsCourse/Session5/") #set to your path

##IQ Sample

pheno <- read.csv("Phenotypic.csv")

pheno[pheno == "-9999"] <- NA

#FQI for Asperger's syndrome

FIQ\_asperger <- pheno$FIQ[] #edit it

#Emitting NA values

FIQ\_asperger <- FIQ\_asperger[] #edit it

#Estimating sample average

mean[FIQ\_asperger] #edit it

#Examination of Z test

z.test(FIQ\_asperger,stdev = 15,mu = 100)

#Examination of Cohen's d effect size

cohen.d(FIQ\_asperger,mu = 100,f=NA)

# Assignment: For each DSM-5 group, define the null and alternative hypotheses.

Examine Z statistics and corresponding effect sizes.#

##Clearing workspace

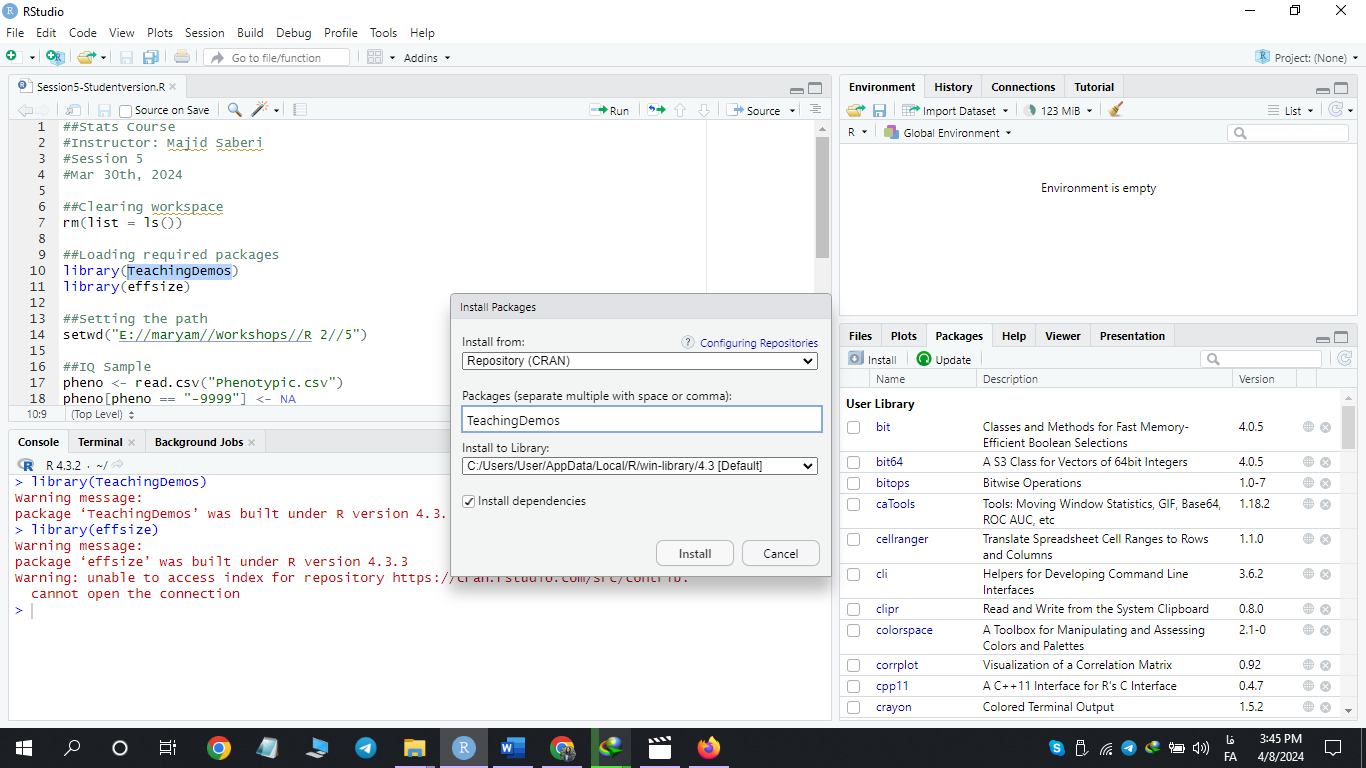
7 rm(list = ls())

##Loading required packages

10 library(TeachingDemos)

11 library(effsize)

خط 7، workspace را پاک می­کند. در خط 10 و 11، دو کتابخانه مورد نیاز باید لود کنید. با توجه به شکل 11 ، دو کتابخانه را نصب کنید. از سمت راست بر روی Install کلیک کنید و نام کتابخانه را بنویسید:



شکل 11

##Setting the path

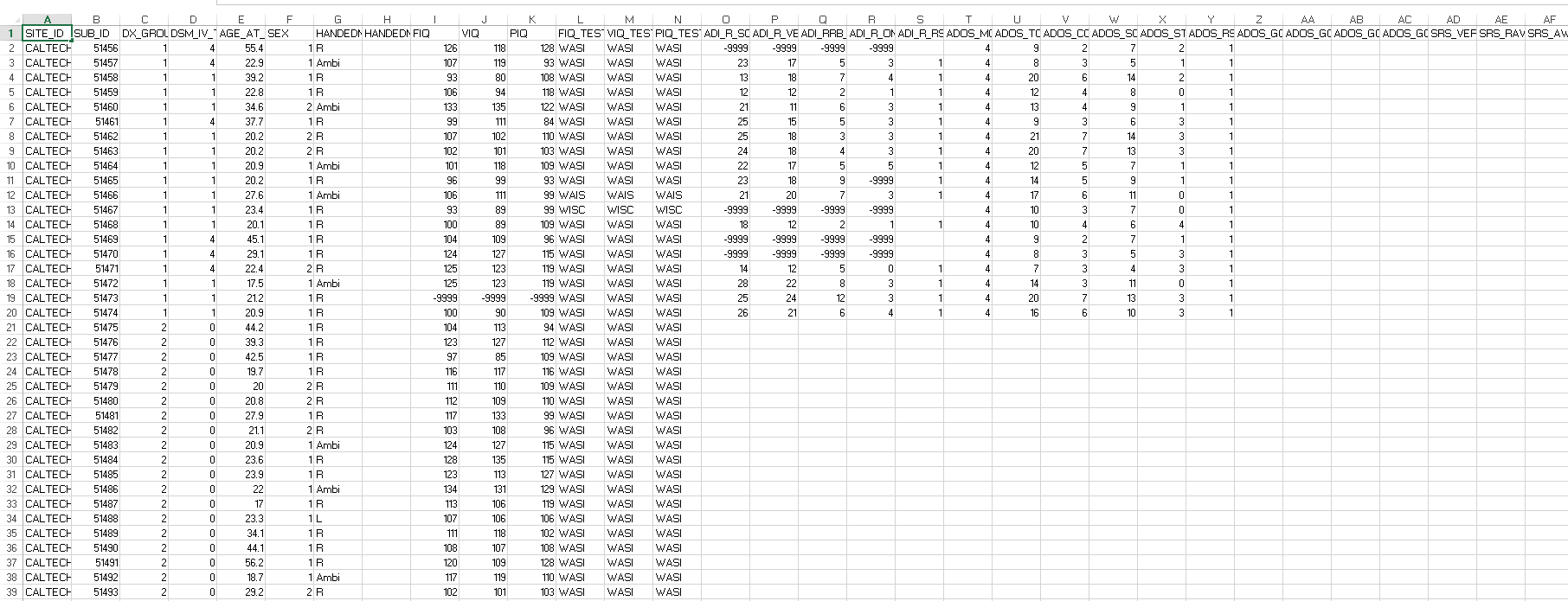
14 setwd("/Users/majid/Projects/StatsCourse/Session5/") #set to your path

##IQ Sample

17 pheno <- read.csv("Phenotypic.csv")

18 pheno[pheno == "-9999"] <- NA

در خط 14، مسیر فایل اکسل مشخص می­شود و در خط 17 فایل اکسل Phenotypic در اختیار قرار داده شده مطابق شکل 12 لود می­شود.



شکل 12

#FQI for Asperger's syndrome

21 FIQ\_asperger <- pheno$FIQ[] #edit it

#Emitting NA values

24 FIQ\_asperger <- FIQ\_asperger[] #edit it

21 FIQ\_asperger <- pheno$FIQ[pheno$DSM\_IV\_TR == 2]

24 FIQ\_asperger <- FIQ\_asperger[FIQ\_asperger != NA]

در خط 21، عدد 2، نشانگر افراد آسپرگر در ستون DSM\_IV\_TR است. شرطی گذاشتیم که FIQ تنها افرادی که آسپرگر هستند را انتخاب کند. در خط 24 ، می­توانیم تمامی NA را حذف کنیم. از راه دیگری هم می­توانیم NA ها را حذف کنیم.

[is.na (FIQ\_asperger)] یعنی فقط NA را انتخاب کنیم. اگر علامت نقیض (!) پشت سر آن بگذاریم، برعکس می­شود.

24 FIQ\_asperger <- FIQ\_asperger[!is.na (FIQ\_asperger)]

در خط 28، ما با خطایی مواجه می­شویم، آن را اصلاح کنید.

#Estimating sample average

28 mean[FIQ\_asperger] #edit it

#Examination of Z test

31 z.test(FIQ\_asperger,stdev = 15,mu = 100)

#Examination of Cohen's d effect size

34 cohen.d(FIQ\_asperger,mu = 100,f=NA)

میانگین هوش افراد آسپرگر عدد 112.1163 است.

در خط 31، یک آزمون Z می­گیریم. این آزمونی است که می­توانید میانگین نمونه را با جامعه مقایسه کنید.

نتایج آزمون Z :

**One Sample z-test**

data: FIQ\_asperger

z = 7.4908, n = 86.0000, Std. Dev. = 15.0000, Std. Dev. of the sample mean = 1.6175,

p-value = 6.846e-14

alternative hypothesis: true mean is not equal to 100

95 percent confidence interval:

108.9461 115.2865

sample estimates:

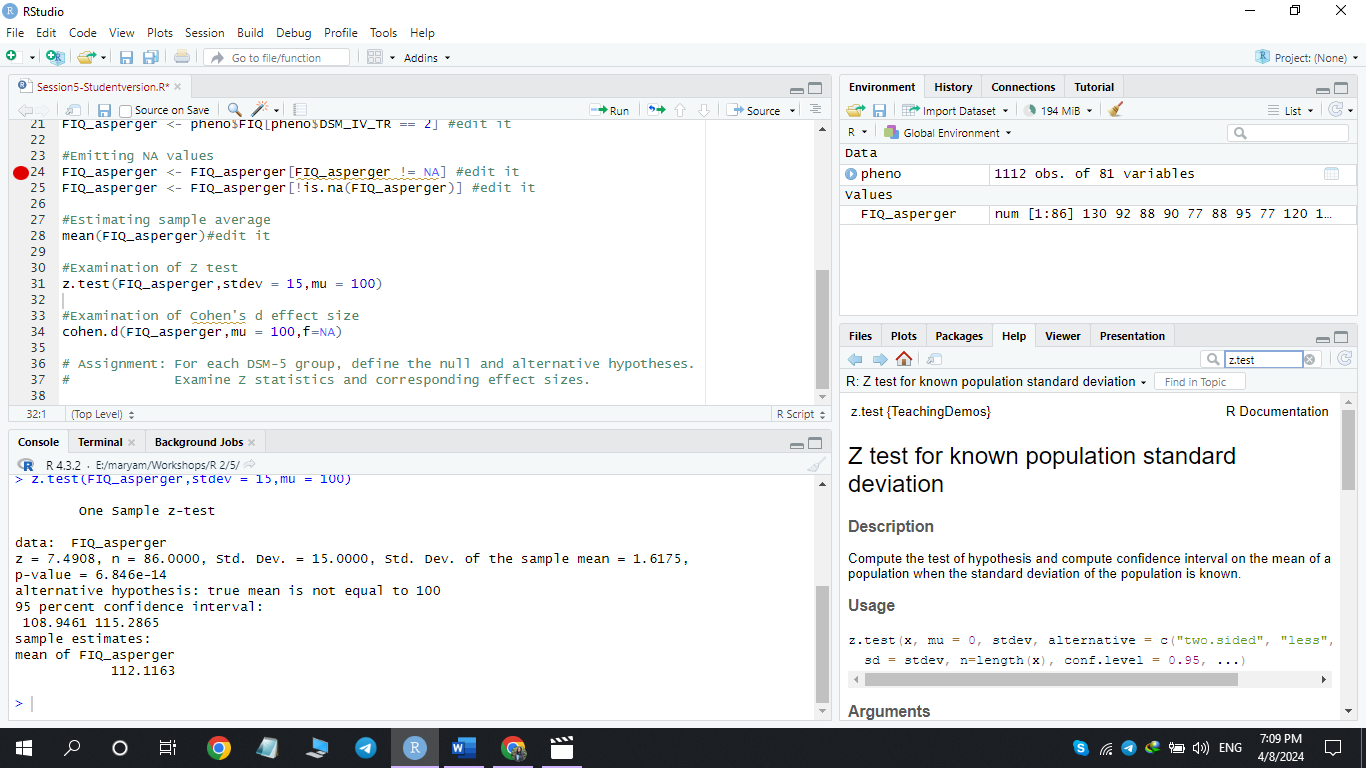
mean of FIQ\_asperger

112.1163

همانطور که می­بینید فرضیه ما دو طرفه است و ما نمی­دانیم میانگین نمونه بزرگتر یا کوچکتر است.

alternative hypothesis: true mean is not equal to 100

پس برای اینکه متوجه شویم هر فانکشن چطور عمل می­کند، در قسمت Help، نام تابع را بنویسید.



شکل 12

به صورت پیش­فرض alternative، دو طرفه است. اما دو حالت دیگر هم دارد که می­تواند بزرگتر و کمتر باشد.

alternative = "two.sided", "less", "greater"

پس در خط 31، اصلاح می­کنیم.

31 z.test(FIQ\_asperger,stdev = 15,mu = 100, alternative = "greater")

نتایج آزمون Z:

**One Sample z-test**

data: FIQ\_asperger

z = 7.4908, n = 86.0000, Std. Dev. = 15.0000, Std. Dev. of the sample mean = 1.6175,

p-value = 3.423e-14

alternative hypothesis: true mean is greater than 100

95 percent confidence interval:

109.4557 Inf

sample estimates:

mean of FIQ\_asperger

112.1163

حالا می­بینیم که مقدار P Value تغییر کرد و می­توانم استنباط کنیم که میانگین هوش نمونه از میانگین هوش جامعه بالاتر است.

z = 7.4908 ، آماره است

n = 86.0000درجات آزادی است

Std. Dev. = 15.0000 ، انحراف معیار جامعه است

Std. Dev. of the sample mean = 1.6175، انحراف معیار نمونه است.

برای محاسبه حجم اثر، خط 34 را اجرا کنید. حجم اثر، متوسط است.

34 cohen.d(FIQ\_asperger,mu = 100,f=NA)

نتایج:

**Cohen's d (single sample)**

d estimate: 0.7668271 (medium)

Reference mu: 100

95 percent confidence interval:

lower upper

0.3225466 1.2111077

**کد اصلاح شده نهایی:**

##Stats Course

#Instructor: Majid Saberi

#Session 5

#Mar 30th, 2024

##Clearing workspace

rm(list = ls())

##Loading required packages

library(TeachingDemos)

library(effsize)

##Setting the path

setwd("E://maryam//Workshops//R 2//5")

##IQ Sample

pheno <- read.csv("Phenotypic.csv")

pheno[pheno == "-9999"] <- NA

#FQI for Asperger's syndrome

FIQ\_asperger <- pheno$FIQ[pheno$DSM\_IV\_TR == 2] #edit it

#Emitting NA values

FIQ\_asperger <- FIQ\_asperger[FIQ\_asperger != NA] #edit it

FIQ\_asperger <- FIQ\_asperger[!is.na(FIQ\_asperger)] #edit it

#Estimating sample average

mean(FIQ\_asperger)#edit it

#Examination of Z test

z.test(FIQ\_asperger,stdev = 15,mu = 100, alternative = 'greater')

#Examination of Cohen's d effect size

cohen.d(FIQ\_asperger,mu = 100,f=NA)

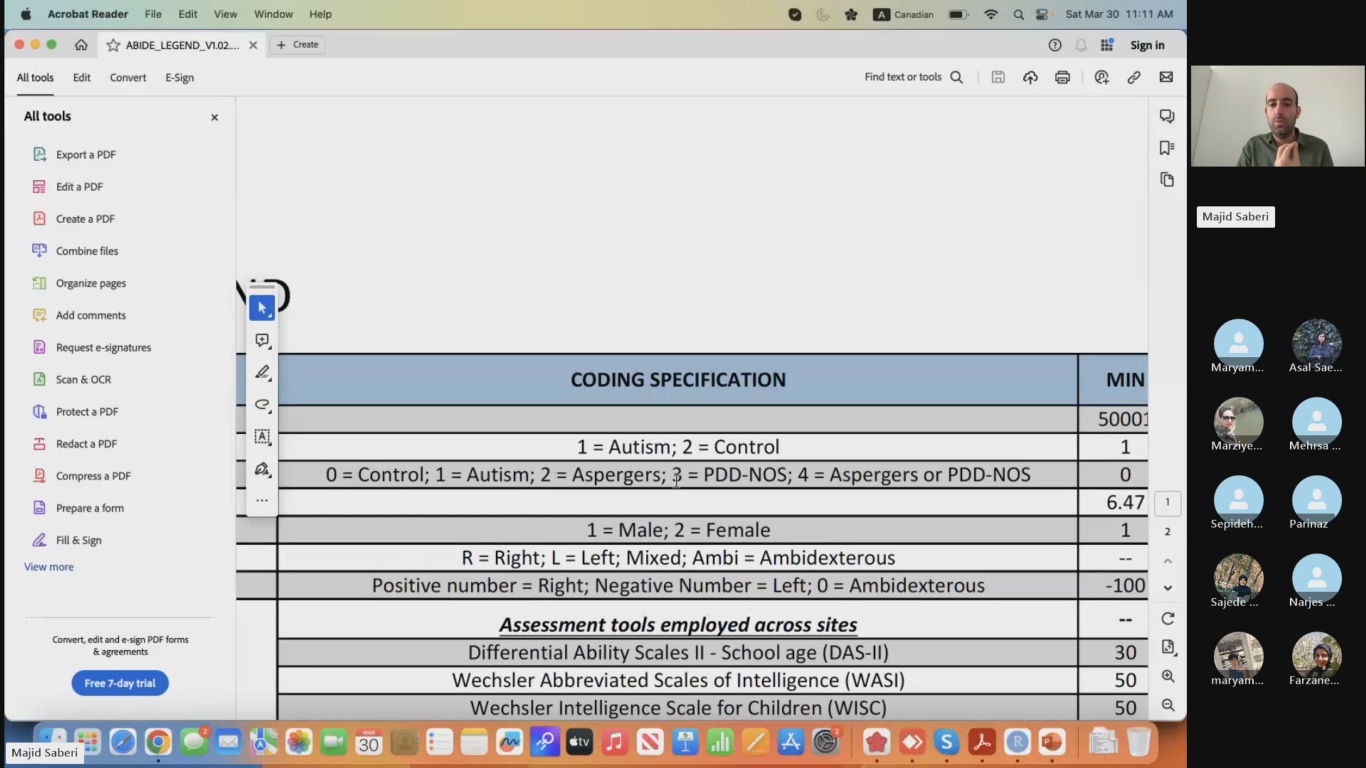
**تمرین**

Assignment: For each DSM-5 group, define the null and alternative hypotheses. Examine Z statistics and corresponding effect sizes.

همانطور که در شکل 13 می­بینید ما 5 برچسب (گروه کنترل، اتیسم، آسپرگر و ...) داریم، به ازای هر 4 تا برچسب از آزمون Z استفاده کنید و نتایج را با سه alternative hypothesis مختلف به دست آورید.

alternative = "two.sided", "less", "greater"

و بعد حجم اثر را محاسبه کنید.



شکل 13

**نکته**: ما از روی مقدار آماره، یعنی Z می­توانیم تشخیص دهیم که میانگین نمونه ما بزرگتر یا کوچکتر از جامعه است. اگر Z مثبت بود یعنی میانگین نمونه ما بزرگتر از جامعه است ولی اگر منفی شد یعنی میانگین ما کوچکتر از میانگین جامعه است.

1. Research Question [↑](#footnote-ref-1)
2. Asperger's syndrome [↑](#footnote-ref-2)
3. Research Hypothesis [↑](#footnote-ref-3)
4. Sample [↑](#footnote-ref-4)
5. Null Hypothesis [↑](#footnote-ref-5)
6. Alternative Hypothesis [↑](#footnote-ref-6)
7. Correct Decision [↑](#footnote-ref-7)
8. Error(Type I) [↑](#footnote-ref-8)
9. Error (Type II) [↑](#footnote-ref-9)
10. Power [↑](#footnote-ref-10)
11. Effect Size [↑](#footnote-ref-11)
12. Cohen’s D [↑](#footnote-ref-12)