

احتمالا همه می‌دانید که «علوم اعصاب شناختی»^۱ یکی از شاخه‌های علم تجربی هست. روند کاری ما برای انجام یک تحقیق در حوضه علوم اعصاب شناختی هم مثل بقیه شاخه‌های علوم تجربی از یک سری مراحل تشکیل شده. همه ما از دبستان با این مراحل آشنا هستیم که به صورت خیلی مختصر عبارت است از: (۱) طرح سؤال یا فرضیه، (۲) جمع‌آوری اطلاعات، (۳) پردازش اطلاعات و (۴) نتیجه‌گیری. این چند مورد خیلی ساده و ابتدایی چیزی است که در هر تحقیق کوچک و بزرگی ما باهاش سروکار داریم و بدون وجود هر یک از این مراحل کار ما ناقص و ارزشش کمتر می‌شود. حالا بیاید باهم این مراحل رو یکی‌یکی بریم جلو و ردپای آمار رو توی هر کدومش پیدا کنیم.

فرض کنید ما می‌خواهیم با سؤالات زیر یک کار تحقیقاتی رو شروع کنیم:

آیا با اضافه شدن وزن یک جسم تشخیص تغییرات وزن آن سخت‌تر می‌شود؟

آیا تشخیص جهت حرکت برای ما در مرکز و اطراف بینایی متفاوت است؟

آیا تشخیص مقصد و جهت حرکت دست کاملا به دیدن دست وابسته است؟

آیا استرس موش بر روی مدت زمان خواب موش تاثیر دارد؟

آقا قد دخترها و پسرهای در بازه سنی ۱۵ تا ۱۷ سال متفاوت است؟

همچنین علم آمار را می‌شود این‌گونه تعریف کرد: مطالعه به منظور یافتن بهترین روش برای جمع‌آوری اطلاعات، پردازش آن‌ها و نتیجه‌گیری از اطلاعات به دست. همان‌گونه که می‌بینید تعریف علم آماری ارتباط تنگاتنگی با چهار مرحله تحقیق دارد. در ادامه ما قدم‌به‌قدم بر مبنای چهار مرحله‌ای که گفته شد این سؤالات را بررسی می‌کنیم تا ببینیم کار ما به عنوان کسی که وظیفه انجام کارهای آماری را در یک گروه تحقیقاتی دارد چی هستش. اما برای هر قدم ما نیاز به مفاهیمی داریم که قبلش توی من‌های جدادشه، مثل زیر، با اون‌ها آشنا می‌شیم.

جامعه^۲: یک گروه کامل از چیزهایی که حداقل همه‌ی آن‌ها یک خصوصیت مشترک داشته باشند. توجه کنید که جمعیت فقط به انسان‌ها محدود نمی‌شود. برای مثال حیوانات، مشاغل، ساختمان‌ها، وسایل نقلیه موتوری، مزارع، روزهای شنبه، نمرات آزمون و تمام اشیاء یا رویدادها.

متغیر^۳: هر خصوصیت، عدد، یا کمیتی است که قابل اندازه‌گیری یا شمارش باشد. متغیر باید پتانسیل متفاوت بودن بین اعضای جامعه را داشته باشد. برای مثال: سن، جنسیت، درآمد و هزینه‌ها، کشور تولد، مخارج کارخانه، نمرات کلاس، رنگ چشم، نوع وسیله نقلیه و غیره.

متغیرها با توجه به تفاوت در نوع بررسی، اندازه و نمایششان به دسته‌های متفاوتی تقسیم‌بندی می‌شوند که در زیر آمده است.

¹ [Cognitive neuroscience](#)

² Population

³ Variable

(۱) **متغیرهای کمی (عددی)**^۴: مقادیر عددی‌ای دارند که یک کمیت قابل اندازه‌گیری را توصیف می‌کند. توصیف‌های مانند «چه اندازه؟» و «چه تعداد؟». برای مثال قد، زمان، سن، دما، تعداد و تعداد فرزندان در یک خانواده همگی متغیر کمی هستند.

(۲) **متغیرهای رسته‌ای**^۵: مقادیری دارد که یک کمیت یا خصوصیت را توصیف می‌کنند. توصیف‌هایی مانند «چه نوعی؟» و «چه دسته‌ای؟». برای نمرات دبستانی (یعنی خیلی خوب، خوب، متوسط، قابل قبول)، اندازه لباس (یعنی کوچک، متوسط، بزرگ، فوق‌العاده بزرگ)، نگرش (یعنی کاملاً موافق، موافق، مخالف، به شدت مخالف)، جنس، نوع تجارت، رنگ چشم و دین همگی متغیرهای رسته‌ای هستند.

توجه داشته باشید که در اکثر موارد متغیرهای رسته‌ای قابل تبدیل به متغیرهای کمی هستند.

هر سؤال تحقیقاتی یک **جامعه هدف** دارد که باید در مرحله اول به خوبی و کاملاً شناسایی کنید بعد از اون هم باید **متغیرهایی** از جامعه هدف که در سؤال تحقیق موردنظر است رو پیدا کنید. با پیدا کردن این متغیرها می‌دونیم که چه اطلاعاتی رو باید جمع‌آوری کنیم. برای مثال در زیر این کار رو برای سؤال‌های اول جزوه انجام‌شده (زیر متغیرها خط کشیدم):

آیا با اضافه شدن وزن یک جسم تشخیص تغییرات وزن آن سخت‌تر می‌شود؟

وزن و تغییرات وزن هر دو یک متغیر از نوع کمی هستند.

آیا تشخیص جهت حرکت برای ما در مرکز و اطراف بینایی متفاوت است؟

هر دو متغیر این سوال از نوع متغیر رسته‌ای هستند. که می‌شود آن‌ها رو به متغیر کمی هم تبدیل کرد.

آیا تشخیص مقصد و جهت حرکت دست کاملاً به دیدن دست وابسته است؟

هر دو متغیر این سوال هم از نوع متغیر رسته‌ای هستند.

آیا استرس موش بر روی مدت زمان خواب موش تاثیر دارد؟

در اینجا استرس موش از نوع رسته‌ای و مدت زمان خواب از نوع کمی است.

آقا قد دخترها و پسرهای در بازه سنی ۱۵ تا ۱۷ سال متفاوت است؟

قد دخترها و قد پسرها دو متغیر کمی این سوال هستند.

حالا که جامعه و متغیرهای سؤال رو پیدا کردید می‌تونیم بریم سراغ مرحله دوم از روند تحقیق که میشه «جمع‌آوری اطلاعات». الان با دانستن متغیرها ما می‌دونیم که چه اطلاعاتی را باید جمع‌آوری کنیم اما آیا عملی و منطقی است که این اطلاعات از کل جامعه جمع‌آوری کنیم؟ جواب خیر هستش. چون معمولاً به خاطر بزرگ بودن جامعه هدف نمیشه از همه‌ی جامعه اطلاعات

⁴ Numeric variables

⁵ Categorical variables

جمع‌آوری کرد. در این شرایط ما مجبور هستیم یک بخش کوچک‌تر از کل جامعه را انتخاب کنیم که به آن **نمونه**^۶ می‌گویند. در متن زیر درباره **نمونه‌گیری**^۷ توضیح داده شده است.

نمونه‌گیری همان‌گونه که از نامش پیداست به معنی اتخاذ یک نمونه کوچک‌تر از جامعه است، در این کار باید توجه شود که نمونه تا جای ممکن بیانگر کل جامعه باشد تا در مرحله نتیجه‌گیری درباره کل جامعه دچار خطا نشویم. موارد که می‌تواند در نمونه‌گیری مشکل‌زا باشد در زیر آمده است.

- (۱) انتخاب دسته‌ای از جامعه^۸ به صورت غیر تصادفی.
- (۲) انتخاب افراد داوطلب از جامعه به عنوان نمونه.
- (۳) انتخاب نمونه فقط از راحت‌ترین و در دسترس‌ترین موارد جامعه.

الان آماده هستیم تا متغیرهایی که از سؤال استخراج کردیم رو روی تمام نمونه‌هایمان اندازه‌گیری کنیم، برای این کار دو روش وجود دارد **مشاهده**^۹ و **آزمایش**^{۱۰} که در ادامه با تفاوت‌های این دو روش آشنا می‌شویم.

ثبت اطلاعات (مقادیر متغیرها) همه نمونه‌ها بدون انجام هیچ‌گونه دخالتی در آن‌ها را **مشاهده** گویند. اما هنگامی که ما در نمونه مداخله‌ای انجام می‌دهیم و سپس به ثبت اطلاعات می‌پردازیم **آزمایش** انجام داده‌ایم. این مداخله معمولاً به صورت دسته‌بند کردن نمونه به صورت تصادفی به چند دسته‌ی متفاوت با توجه به تعداد متغیرها سپس مبتلا کردن هر دسته به یکی از حالات یک (یا چند) متغیر و سپس اندازه‌گیری (یک یا چند) متغیر دیگر است. برای روشن شدن بهتر مطلب به تصویر ۱ دقت کنید.

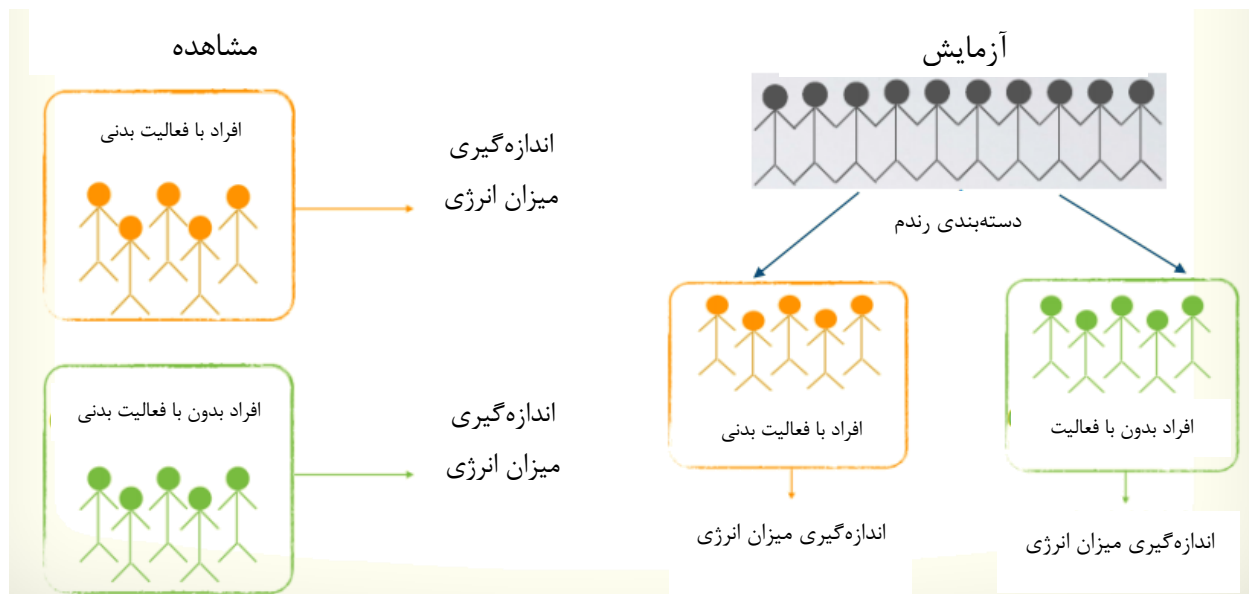
^۶ Sample

^۷ Sampling

^۸ همیشه دقت داشته باشید این جامعه به جامعه انسان‌ها لزومت اشاره ندارد.

^۹ Observation

^{۱۰} Experiment



نصیر ۱- فرض کنید سؤال تحقق ما این باشد که میزان انرژی در افرادی که فعالیت بدنی دارند با افراد بدون فعالیت بدنی متفاوت است یا نه؟

تفاوت مهم آزمایش و مشاهده این هستش که در مشاهده ما به ارتباط^{۱۱} پدیده‌ها پی می‌بریم و نمی‌تونیم بگیم که یکی علت اتفاق دیگر هستش. برای پیدا کردن رابطه‌ی علت و معلولی دو پدیده ما نیاز به آزمایش کردن داریم. مثلاً یک فروشنده بستنی مشاهده می‌کند که با افزایش تعداد افرادی که عینک آفتابی دارند تعداد خریدارهای بستنی هم بالا میره. احتمالاً اگر اینجا ما بگیم علت بالا رفتن خریدارهای بستنی بیشتر شدن آدم‌هایی که عینک آفتابی می‌زنند هستش خنده‌دار به نظر بیاد. درواقع اون چیزی که علت هردوی این پدیده‌ها هست و دیده نشده تابش بیشتر خورشید و افزایش دمای هواست. یا مثلاً گفته میشه افزایش خوردن فست‌فود در جوانی باعث سرطان در پیری میشه. اگر دقت کنید حرف هم از یک مشاهده آمده و کلمه «باعث» اشتباه به‌کاررفته شده، شاید کلی دلیل دیگه مثل سبک زندگی، تحرک کم، زندگی پرسترس این وسط دخیل باشه و هر کدام در ایجاد سرطان مؤثر باشند (مسلماً نتیجه این حرف من تأیید خوردن فست‌فود نیست! گفتم اشتباه نکنید) یا برای مثال با نشان دادن یک محرک بینایی جایی از مغز فعال شد نمی‌توان به‌سادگی گفت تنها علت فعال شدن آن بخش مغز آن محرک بینایی است.

حالا ما آمادگی داریم تا یک آزمایش طراحی کنیم و اطلاعات رو جمع‌آوری کنیم (برای این کار به دانش هم گروهی‌هاتون که توی کارگاه سایکوفیزیک و الکترونیک بودند احتمالاً نیاز خواهید داشت). به این اطلاعات جمع‌آوری شده داده^{۱۲} می‌گویند پس از این مرحله میریم سراغ پردازش اطلاعات به‌دست‌آمده (در این مرحله هم شاید داده‌ها نیاز به پردازش اولیه^{۱۳} داشت باشند که از اطلاعات هم گروهی‌هاتون توی کارگاه پردازش سیگنال می‌تونید استفاده کنید).

کاری که ما برای پردازش داده‌ها مون انجام می‌دیم، انجام محاسبات ریاضی روی آن‌هاست. این محاسبات عمدتاً ساده هستند و خودمان هم می‌تونیم انجام بدیم، ولی معمولاً به دلیل حجم بالای اطلاعات برای انجامشون به ماشین حساب قوی نیاز داریم. ماشین حساب ما نرم‌افزار «MATLAB» هستش. در ادامه دستورات که ما برای پردازش داده نیاز داریم اومده. ولی برای یادگیری بهتر می‌تونید از منابع معرفی شده در جزوه سیگنال استفاده کنید.

¹¹ Correlation

¹² Data

¹³ Data pre-processing

پردازش داده‌های مثال زیر را انجام می‌دهیم. اما تکنیک‌های استفاده‌شده در مثال مختص خود مثال نیست و در بقیه موارد هم قابل استفاده‌ست.

سؤال زیر رو قبلا دید:

« آقا قد دخترها و پسرهای در بازه سنی ۱۵ تا ۱۷ سال متفاوت است؟ »

برای پاسخ دقیق به این سؤال باید قد کل دخترها و پسرهای در بازه سنی ۱۵ تا ۱۷ رو داشته باشید (جدی) ولی چیزی که مشخصه این اطلاعات رو ما نمیتونیم به دست بیاریم، پس چی کار می‌کنیم؟ آفرین، نمونه‌گیری. حالا فرض کنید نمونه زیر روداریم و شرایط نمونه‌گیری درست رو رعایت کردیم (تو فکر کن یه درصدا!).

دخترها	۱۶۵	۱۷۰	۱۷۱	۱۷۵	۱۷۰	۱۷۵	۱۶۵	۱۷۱	۱۶۰	۱۸۰	۱۷۲	۱۷۰	۱۷۳	۱۶۸	۱۶۴	۱۵۷	۱۶۳
پسرها	۱۷۵	۱۷۲	۱۷۵	۱۷۴	۱۷۲	۱۸۰	۱۷۵	۱۷۶	۱۷۸	۱۷۵	۱۷۸	۱۸۰	۱۷۴	۱۵۵	۱۷۸	۱۷۸	۱۸۲

برای این‌که بتونیم با استفاده از نمونه درباره جامعه نظر بدیم اول باید داده‌های نمونه رو به‌درستی تحلیل کنیم. یعنی برای مثالمون باید ابتدا مشخص کنیم که آیا قد دخترها با پسرها متفاوت است یا نه. متفاوت بودن قد یک نفر از دیگری قابل فهم هست. ولی تفاوت قد دو گروه آدم از هم چه معنایی می‌دهد؟ یا به‌طور کلی متفاوت بودن یا نبودن دودسته از مقادیر چه معنایی داره؟

در یک حالت جواب این سؤال ساده است حالتی که همه پسرها هم قد و همه دخترها هم قد باشند چون می‌تونیم یک عدد را به‌عنوان شاخص قد تمام دخترها و یک عدد را به‌عنوان قد تمام پسرها در نظر بگیریم و اون دو رو باهم مقایسه کنیم. ولی متأسفانه این اتفاق تقریباً محاله. اما در شرایطی که این‌جوری نیست (قد‌های هر گروه یکسان نیست) باز ما با روش‌هایی به دنبال پیدا یک مقدار به‌عنوان نماینده تمام قد‌های گروه و مقدار دیگری برای نشان دادن دقت این نماینده هستیم.

گفتیم که ما کار پردازش داده‌ها رو با نرم‌افزار MATLAB می‌دهیم پس در اولین قدم باید داده‌های بالارو در متلب وارد کنیم. برای این کار دستورات زیر را می‌نویسیم. (این دستورات و دستورات بعدی رو می‌تونید برای داده‌های آزمایش خودتون با کمی اصلاح استفاده کنید)

```
clear;
clc;
close all;

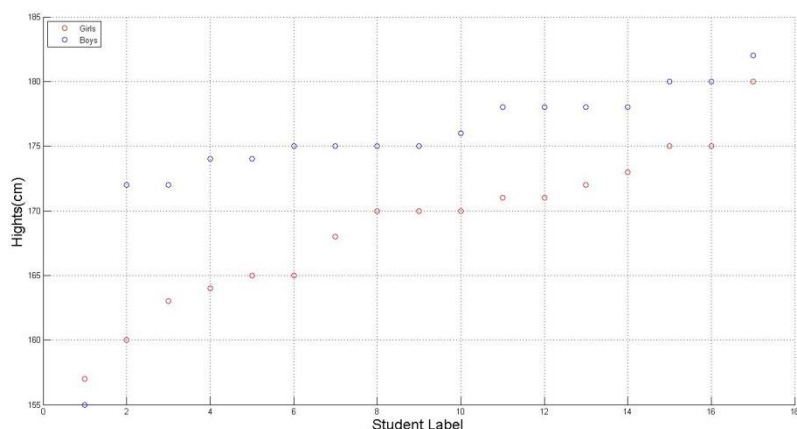
Hights = [163, 157, 164, 168, 173, 170, 172, 180, 160, 171, 165, 175,...
          170, 175, 171, 170, 165; 182, 178, 178, 155, 174, 180, 178, 175,...
          178, 176, 175, 180, 172, 174, 175, 172, 175];
```

حاصل اجرای خطوط بالا یک جدول با دو سطر و هفده ستون (ماتریس دو در هفده) هستش. هدف نهایی ما توش این مرحله از پردازش پی بردن به آینه که آیا قد دخترها و پسرها متفاوت است یا نه. اولین کارمون داده‌ها را مرتب می‌کنیم و در یک نمودار نقطه‌ای نمایش می‌دهیم. تا بهتر بتونیم این داده‌ها رو ببینیم و حسی نسبت به تفاوت یا عدم تفاوت اون‌ها داشته باشیم. دستورات زیر این کار رو برای ما انجام می‌دهند.

```

hold on
grid on
Student_Labels = 1: length(Hights);
plot(Student_Labels, sort(Hights(1,:)), 'or');
plot(Student_Labels, sort(Hights(2,:)), 'ob');
xlabel('Student Label', 'fontsize',18);
ylabel('Hights(cm)', 'fontsize',18)
legend('Girls', 'Boys','Location','NorthWest')

```



خروجی این دستورات برای دادهای ما نمودار زیره

از دیدن این نمودار یک حس قوی به ما دست می‌دهد که قد پسرها با دخترها متفاوت و قد پسرها بیشتره. اما این کافی نیست و ما به دنبال اعدادی برای نشان دادن این تفاوت هستیم.

علت تلاش ما برای پیدا کردن اعدادی برای نشان دادن مقدار نماینده و پی بردن به تفاوت دو گروه با استفاده از آن اعداد و بسنده نکردن به نمودار این است که، اولاً برای متوقف نشدن پردازش و جلو رفتن آن نیاز به عدد داریم تا با استفاده از ابزار آماری و ریاضیاتی بتوانیم این اعداد را تحلیل کنیم در ضمن اعدادی که در ادامه به آنها می‌ریسیم در جامعه علمی جاافتاده هستند و با استفاده از این اعداد ما می‌توانیم با یک‌زبان مشترک حرفمان را حرف بزنیم و دلایل دیگر 😊

در گام بعدی هدف ما پیدا کردن عدد نماینده (شاخص) قدهای هر گروه است. در حالتی که همه‌ی قدها یکسان باشند دیدن به‌سادگی میشه یکی از قدها را به‌عنوان نماینده کل گروه انتخاب کرد اما هر مقدار قدها از یکسان بودن دور بشند (دور بودن یعنی چی؟) کار ما سخت‌تر و انتخاب ما نادقیق‌تر میشه. این شاخص (نماینده) درواقع نشان‌دهنده‌ی نقطه تجمع داده‌هاست. به همین دلیل به این عدد «**شاخص تجمعی**» دسته اعداد می‌گیریم. در آمار روش‌های زیادی برای پیدا کردن این شاخص وجود داره. اما

شاخصی که معمولاً استفاده می‌شود بهش می‌گویند میانگین^{۱۴}. شاخصی که همگی با اون آشنایید و رابطه‌ی محاسبش نیاز به توضیح نداره.

به دودسته‌ی عددهای زیر دقت کنید،

$$A := \{-100, -10, -1, 1, 10, 100\}$$

$$B := \{-2, -1, 0, 0, 1, 2\}$$

میانگین هر دودسته صفره اما آیا صفر برای هر دودسته به یک اندازه نماینده‌ی خوبیه؟ از اعداد واضح که صفر به مراتب نماینده بهتری برای دسته B هستش تا A چون که اعداد دسته A پراکندگی بیشتری نسبت به B دارد. از مثال واضح که پراکندگی داده‌ها دقت شاخص تجمعی رو معین می‌کند. پس یک معیار برای پراکندگی داده‌ها هم باید داشته باشیم که به اون «شاخص پراکندگی» می‌گیم. چندین تا شاخص پراکندگی معروف وجود داره ولی ما اینجا فقط «انحراف معیار^{۱۵}» رو معرفی می‌کنیم.

هدف انحراف معیار پیدا کردن میانگین فاصله داده‌ها از معیار تجمعی مورد استفاده (در اینجا میانگین) است. رابطه انحراف معیار در زیر آمده است.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

تعداد $n :=$
اعضای مجموعه
 $\bar{x} :=$ میانگین
یکی از اعضای مجموعه $x_i :=$

برای نشون دادن انحراف معیار یک میانگین (یا شاخص تجمعی دیگر) در نمودار دوپاره خط موازی محوری که میانگین‌ها در اون قرار دارند و به طول انحراف معیار رسم می‌کنیم.

دستورات زیر میانگین و انحراف معیار دو گروه را حساب میکنه و اون را در نمودار میله‌ای نشون می‌ده.

¹⁴ Mean

¹⁵ Standard deviation

```

B_avg = mean(Hights(2,:));
G_avg = mean(Hights(1,:));

B_Hight_s = sort(Hights(2,:));
G_Hight_s = sort(Hights(1,:));

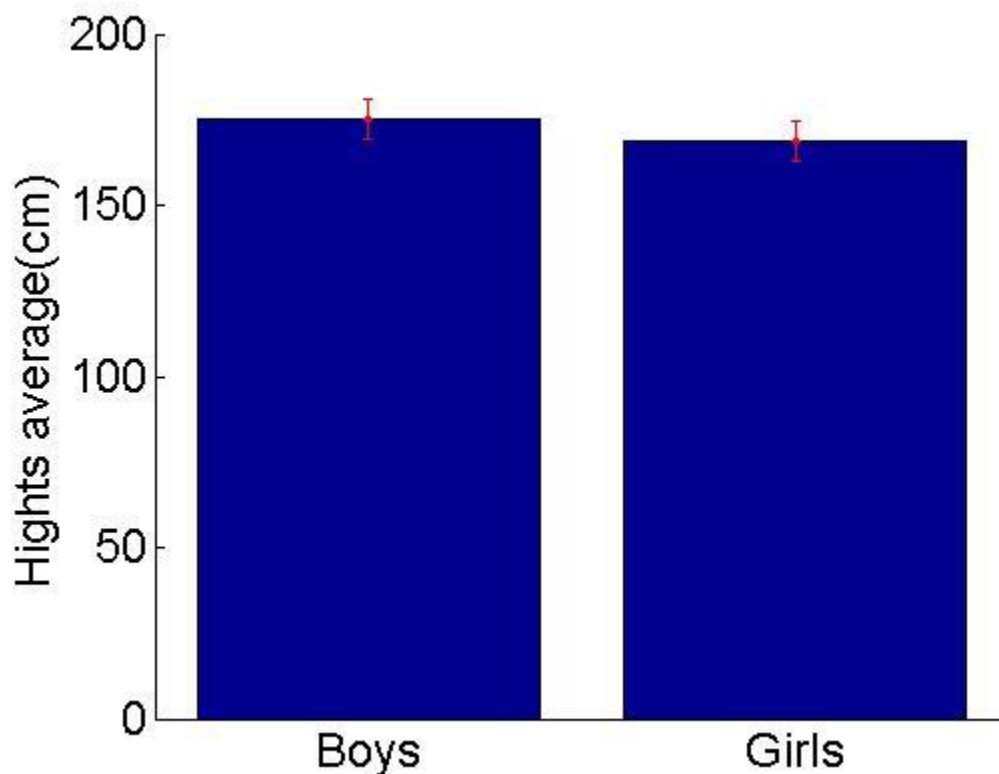
N_B = 1 : length(Hights(2,:));
N_G = 1 : length(Hights(1,:));

S_Boys = std2(Hights(2,:));
S_Girls = std2( Hights(1,:));

hold on
bar([1,2], [B_avg, G_avg])
errorbar([B_avg, G_avg], [S_Boys, S_Girls], '. r')
ylabel('Hights average(cm)', 'fontsize',18);
set(gca, 'XTick',[1 2], 'XTickLabel',{ 'Boys' , 'Girls'}, 'fontsize',18);

```

نتیجه اجرای دستورات بالا نمودار زیره.



حال ما برای هرکدام از گروه دخترها و پسرها دو عدد به دست آوردیم، میانگین و انحراف از معیار. میانگین نشون دهنده قدیه که می‌تونیم همه اعضای گروه رو با این قد فرض کنیم و انحراف یه جوری دقت این فرض رو نشون میده. درنتیجه هر چی تفاضل میانگین‌ها بیشتر باشد و مقدار انحراف از معیارها کمتر، داده‌های دودسته بیشتر از هم متفاوت هستند. پس می‌تونیم شاخصی برای

نشون دادن این تفاوت دودسته از داده‌ها تعریف کنیم که برابره با قدر مطلق تفاضل میانگین‌ها تقسیم‌بر مجموع انحراف معیارهای دودسته. که اسم این شاخص رو t می‌ذاریم.

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sigma_1 - \sigma_2}$$

\bar{x}_1 : میانگین دسته اول

σ_1 : انحراف معیار دسته اول

\bar{x}_2 : میانگین دسته دوم

σ_2 : انحراف معیار دسته دوم

شاخصی t را می‌تونیم برای هر دودسته‌ای از اعداد محاسبه کنیم و از تعریف t مشخصه که هر چه t بزرگ باشد دودسته باهم متفاوت‌ترند. مثلاً شاخص t برای داده‌های ما برابره با.

$$t_0 = 0.5411$$

حالا سؤال مهم اینه که این t یعنی دسته‌ها متفاوتند یا نه؟ و می‌تونیم این نتیجه رو به جامعه گسترش بدیم؟ برای پاسخ به این سؤال‌ها ما از یه سری آزمون آماری استفاده می‌کنیم که خروجیش یک‌دفعه اینه که آیا دودسته متفاوت هست یا نه؟ این‌که آزمون آماری چی جووری کار میکنه توی ویدیوهای کلاس کامل توضیح داده میشه و متن زیر یاد می‌ده چه شکلی از این آزمون آماری استفاده کنیم.

فرایند نتیجه‌گیری درباره جامعه از نمونه را **استنباط آماری**^{۱۶} می‌گویند، در این فرایند ما نتیجه‌گیری قطعی نداریم و استنباطی موردقبول ماست که احتمال خطای پایینی (مثلاً کمتر از یک درصد) داشته باشد. در آمار آزمون‌های متفاوتی برای پیدا کردن این احتمال خطا وجود دارد که با توجه به داده‌ها و کاری که مدنظر ما است یکی را انتخاب می‌کنیم. به آن‌ها **آزمون فرض آماری**^{۱۷} می‌گویند. این آزمون‌ها **فرض‌های آماری** را که مفهومی انتزاعی است می‌سنجند. در هر عمل مقایسه ما با دو فرض صفر و یک سرکار داریم که h_1 و h_0 نمایششان می‌دهیم. این دو فرض در زیر آمده است.

فرض صفر (h_0): هیچ اتفاقی نیفتاده است. یا مقدار متغیر مورد بررسی در دو جامعه‌ای که دودسته از داده‌های ما از آن آمده است یکسان است. که برای سؤال ما می‌شود قد دخترها و پسرها یکسان است.

فرض یک (h_1): اتفاقی افتاده است. یا مقدار متغیر مورد بررسی در دو جامعه‌ای که دودسته از داده‌های ما از آن آمده است متفاوت است. که برای سؤال ما می‌شود قد دخترها و پسرها متفاوت است.^{۱۸}

پس اولین مرحله تشکیل فرض‌های آماری است پس از تشکیل دو فرض آماری ما آزمون مناسب رو انتخاب می‌کنیم (مثلاً آزمون t) و آن را روی دودسته داده‌ای که داریم اجرا می‌کنیم (این کار را ما با مطلب انجام

¹⁶ Statistical inference

¹⁷ Statistical hypothesis testing

¹⁸ همان‌گونه که می‌بیند فرض یک نقیض منطقی فرض صفر است.

می‌دهیم) و خروجی این تست آماری به ما می‌گوید هر فرض با چه احتمالی درست است. جمع این دو احتمال یک می‌شود

پس حالا برای پاسخ سوالمون (این که قد دخترها و پسرها متفاوت یا نه؟) باید یک آزمون آماری روی داده‌ها مون اعمال کنیم که ما اینجا از $ttest2$ ^{۱۹} استفاده می‌کنیم. مثل زیر:

```
[flag, p_value] = ttest2(Hights(1,:), Hights(2,:), .01);
```

دو ورودی اول دستور $ttest2$ مشخصه و ورودی سوم مقدار بیشینه احتمال درستی فرض صفره تا بشه نتیجه گرفت فرض یک درسته (اگر درستی فرض صفر از این بیشتر بود نمی‌تونیم درباره فرض یک نظری بدیم). بعد از اجرای این خط اگر فرض یک درست باشه $flag$ برابر با یک می‌شه و اگر نتونیم درباره اون نظر داد $flag$ برابر صفر می‌شود. همچنین p_value (p_value) را در گوگل سرچ کنید!) احتمال درستی فرض صفر را می‌دهه. یا معادلاً احتمال غلط بودن فرض یک.

پیروز باشید

علیرضا ابوالقاسمی دهاقانی

ar.abolghasemi@ut.ac.ir

^{۱۹} در حالتی که دو دسته داده مربوط به یک گروه باشند (اینجا مربوط به دو گروه پسرها و دخترها است) ما از دستور $ttest$ استفاده می‌کنیم که دستورش در متلب کاملاً شبیه $ttest2$ است فقط باید اسم آن را تغییر دهید.