دانشگاه آزاد اسلامی زنجان
گروه مهندسی نقشه برداری

جزوه درس تئوری خطاها

تألیف، ترجمه و گرد آوری شده توسط:

دکتر فرید اسماعیلی

درس تئوری خطاها

جلسه اول

فرید اسماعیلی
Farid_63@yahoo.com
www.faridesm.ir

تماس با استاد از طریق پست الکترونیکی
مشاهده اطلاعیه ها، نمرات، دریافت فایل ها در وب سایت

سرفصل ها

تئوری خطاها

کد : ۲۲

تعداد واحد : ۲

نوع واحد : نظری

پیش‌نیاز : آمار و احتمالات مهندسی

هدف : آماده‌سازی دانشجویان جهت درک مفاهیم پردازش اطلاعات نقشه‌برداری و بررسی آماری آنها

سرفصلهای درس :

۱- مروری بر اصول و مفاهیم احتمالات و آمار : مفاهیم پایه، نمونه آماری، متغیر اتفاقی، مشخصات توزیع متغیر اتفاقی، واریانس و مفهوم دقت، تابع توزیع چندمتغیره، کوواریانس، همبستگی و استقلال، بیضی خطاها، ماتریس واریانس کوواریانس.

۲- تئوری اندازه‌گیری :

- فرآیندها و سری اطلاعاتی، فرآیند قطعی و اتفاقی
- مدل‌سازی آماری فرآیند اتفاقی : ممان‌های آماری، تابع چگالی احتمال، تابع اتوکوواریانس.
- کمیت قابل مشاهده : تعریف مشاهده، مفهوم دقت و صحت، مؤلفه‌های کمیت قابل مشاهده.
- خطای سیستماتیک (قوانین و نحوه برخورد با آن)
- خطای اتفاقی و روش برخورد با آن
- بررسی اشتباه و روش مشاهده

۳- انتشار خطاها :

- کلیات
- بررسی انتشار خطاها در ترازیابی و تعیین موقعیت
- ماتریس واریانس، کوواریانس، بیضی خطاها
- مفهوم هندسی ماتریس کوواریانس (بیضی خطاها)، جهت‌های ماکزیمم و مینیمم انحراف معیار.
- انتشار خطا در حالت مدل غیرخطی.

۴- تئوری برآورد : برآورد واریانس براساس نمونه، برآورد کمترین مربعات، برآورد ماکزیمم احتمال، مفهوم وزن، میانگین واریانس نمونه، ماتریس وزن و کمترین مربعات، آشنایی با اصول سرشکنی، سرشکنی پارامتریک، سرشکنی با معادلات شرط، نوسان مختصات.

۵- مروری بر مبانی جبر خطی و آشنایی با یک نرم‌افزار با قابلیت محاسبات ماتریسی مثل Matlab

به طور کلی در درس تئوری خطاها انتظار می رود که دانشجو با مباحث زیر آشنا شود :

- مفاهیم اندازه گیری و خطا
- مفاهیم اساسی مورد نیاز از آمار و احتمال بر اساس تئوری خطاها (فاصله اطمینان، امید ریاضی، ضریب همبستگی، توابع توزیع و...)
- قوانین پخش و انتشار خطاها
- نمایش هندسی پخش خطاها
-
- آزمون های پیش از سرشکنی
- آشنایی اولیه و عملی با مفاهیم جبر ماتریسی به کمک نرم افزار متلب

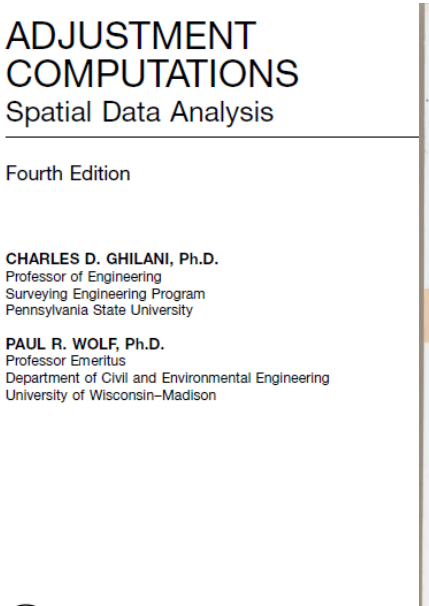
ارزشیابی :

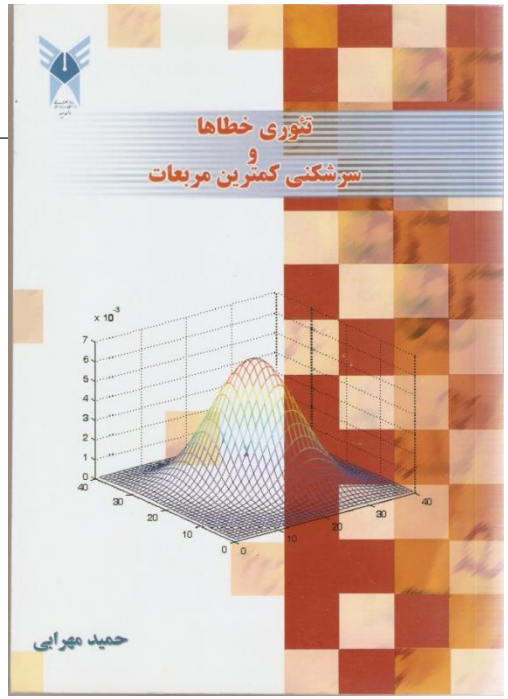
۲ نمره حل تمرین، فعالیت کلاسی (و البته حضور در کلاس!)

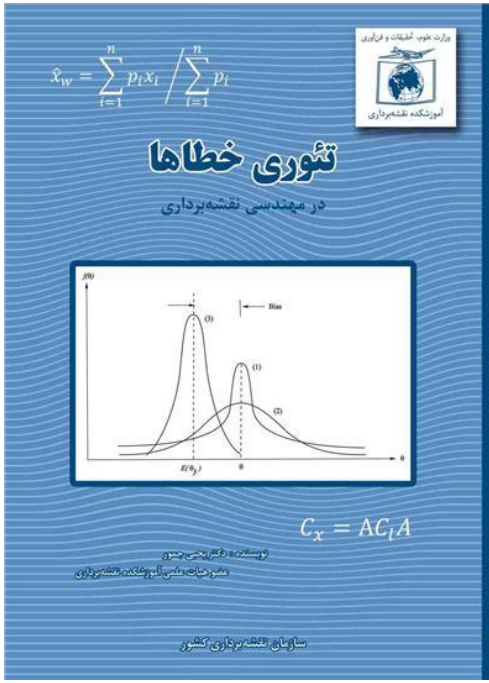
۵ نمره امتحان میان ترم

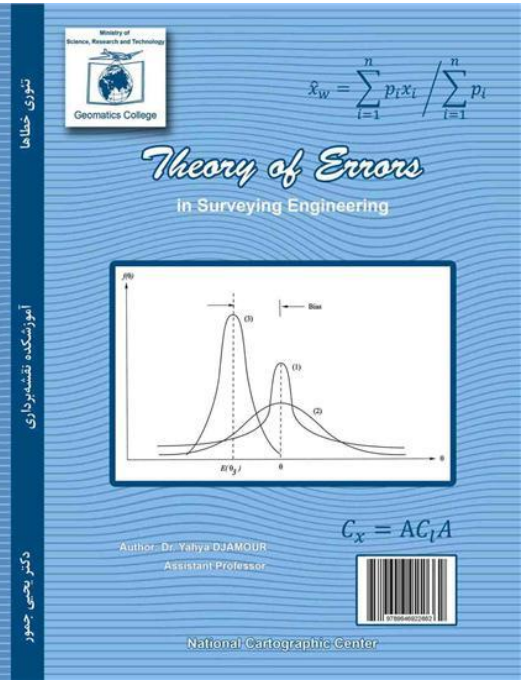
۱۳ نمره امتحان پایان ترم

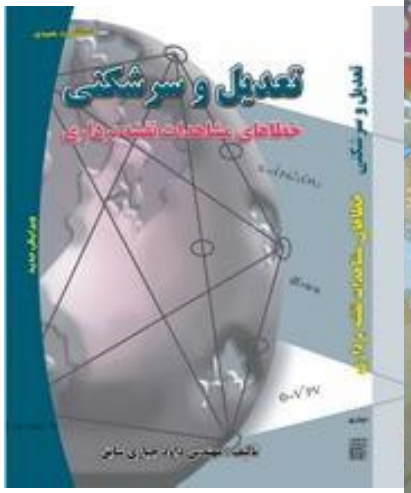
منابع

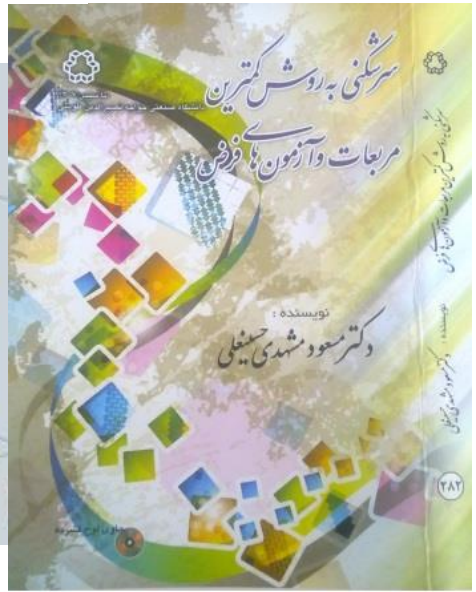

ADJUSTMENT COMPUTATIONS
Spatial Data Analysis
Fourth Edition
CHARLES D. GHILANI, Ph.D.
Professor of Engineering
Surveying Engineering Program
Pennsylvania State University
PAUL R. WOLF, Ph.D.
Professor Emeritus
Department of Civil and Environmental Engineering
University of Wisconsin–Madison
WILEY
JOHN WILEY & SONS, INC.

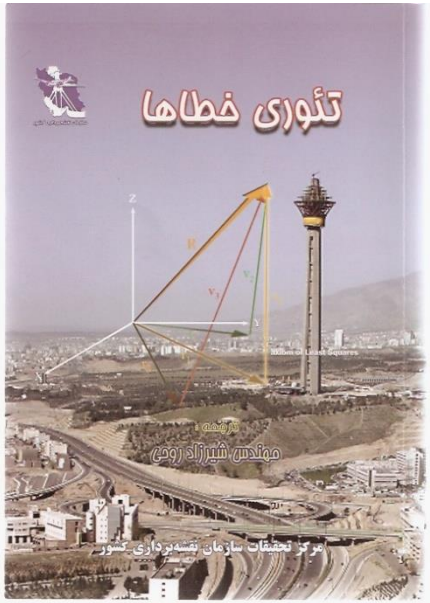

تئوری خطاها
و
سرشکنی کمترین مربعات
حمید مهرابی

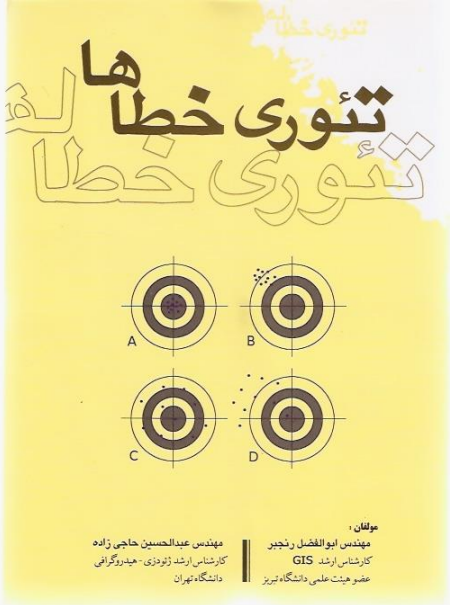

تئوری خطاها
در مهندسی نقشه‌برداری
$C_x = AC_lA$
سازمان نقشه‌برداری کشور


Theory of Errors
in Surveying Engineering
$C_x = AC_lA$
Author: Dr. Yahya OJAMOUR
Assistant Professor
National Cartographic Center


تعدیل و سرشکنی


سرشکنی به روش کمترین مربعات و آزمون های فرض
نویسنده:
دکتر مسعود مشهدی حسینلی


تئوری خطاها
مرکز تحقیقات سازمان نقشه‌برداری کشور


تئوری خطاها
A
B
C
D
مولفان:
مهندس ابوالفضل رنجبر
کارشناس ارشد GIS
عضو هیئت علمی دانشگاه تبریز
مهندس عبدالحسین حاجی زاده
کارشناس ارشد ژئودزی - هیدروگرافی
دانشگاه تهران

- شناخت ویژگی های زمین و محیط پیرامون زندگی بشر به دلیل ایجاد بستر مناسب زندگی، همواره مورد توجه بوده است.

- بررسی و تعیین ویژگی های محیط زندگی ما در قالب تعیین پارامتر های هندسی و فیزیکی ای همچون طول، زاویه، ارتفاع، فشار، وزن، مساحت، حجم، دما و ... انجام می پذیرد که مجهول می باشند.

- رشته مهندسی نقشه برداری در میان سایر رشته ها در دستیابی به شناخت از محیط پیرامون بشر از جایگاه ویژه ای برخوردار است. اندازه گیری های مختلف در حوزه زمان و مکان نقشی زیر بنایی در مسائل مهندسی نقشه برداری دارند.

- نقشه برداری مسطحه، نقشه برداری ارتفاعی، توپوگرافی، نقشه برداری زیرزمینی، نقشه برداری هوایی، نقشه برداری مسیر آبنگاری (هیدروگرافی)، نقشه برداری نظامی، نقشه برداری مشتقه (نقشه برداری موضوعی)، ، کارتوگرافی، نقشه برداری ثبتی املاکی، نقشه برداری ژئودزی

- در تمام پروژه های نقشه برداری به دنبال اندازه گیری و تعیین کمیت های مجهول هستیم. اما بسته به هدف و دقت مورد نظر، همواره یک نبود قطعیت در بر آورد کمیت های مجهول وجود دارد. دلیل نبود قطعیت وجود خطا در اندازه گیری ها و نیز مدل های مورد استفاده در مسائل مختلف مهندسی نقشه برداری است.

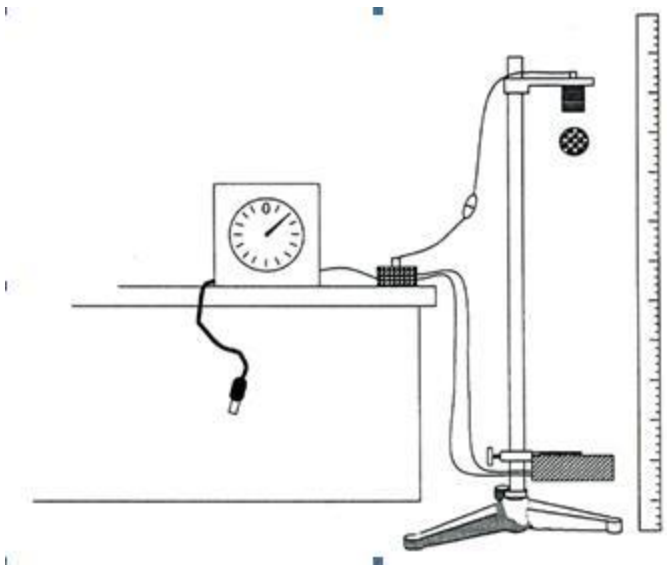
- با توجه به اینکه هیچگاه امکان دستیابی به اندازه گیری های بدون خطا و مدل سازی های کامل و بی نقص وجود ندارد، لذا بررسی رفتار خطاها در اندازه گیری ها و سپس در مدل سازی ها و نتایج حاصل از آن به ویژه در پروژه های حساس و دقیق ضروری است و باعث ایجاد مبحثی به نام **نظریه خطاها** می گردد.

- با توجه به اینکه مقدار واقعی یک کمیت عموماً قابل حصول نیست، تلاش می کنیم تا به دقیق ترین مقدار برای یک کمیت که به کمک مجموعه ای از اندازه گیری ها امکان پذیر می باشد، دست یافته و چگونگی بر آورد دقت این مقدار را تعیین کنیم.

- اختلاف بین کمیت اندازه گیری شده با مقدار واقعی آن کمیت را **خطا** می گویند.

– فرآیند های قطعی

در عمل بسیاری از پدیده های فیزیکی وجود دارند که داده های تولیدی آنها با دقتی قابل قبول با مدل های ریاضی صریح قابل بیان هستند.
برای مثال در سقوط آزاد یک جسم چنانچه y ارتفاع پیموده شده در زمان t و g شتاب گرانش زمین باشد بر اساس قوانین پایه فیزیک رابطه زیر را میتوان بیان نمود:



$$y = \frac{1}{2} g t^2$$

با در نظر گفتن دقتی منطقی، حرکت ماهواره در یک مدار مشخص به دور زمین، یا داده های فیزیکی مربوط به حرکت جسم در سامانه ساده جرم – فنر مثال هایی از پدیده های فیزیکی قطعی هستند.

– فرآیند های اتفاقی

-خیلی از پدیده های فیزیکی دیگر وجود دارند که داده هایی با ماهیت قطعی تولید نمی کنند و رفتار آنها به هیچ وجه قابل مدل سازی و پیش بینی نیست. در چنین فرآیند هایی هر مشاهده، از مشاهدات دیگر کاملاً مستقل بوده و به طور اتفاقی رخ می دهد. و به همین دلیل به آنها فرآیند های اتفاقی گفته می شود. در فرآیند های قطعی می توان رفتار پدیده های فیزیکی را بر اساس مدل های ریاضی برای آینده پیش بینی نمود؛ اما در فرآیند های اتفاقی چنین چیزی امکان پذیر نیست و باید از علم احتمال و آمار برای تجزیه و تحلیل آنها استفاده نمود.

- به عنوان مثال تغییر لحظه ای سطح آب دریا و یا دمای هوا را میتوان به عنوان مثالی برای پدیده های اتفاقی در نظر گرفت.

فرآیند های اتفاقی

ایستا :
فرآیند هایی با خواص آماری ثابت در طول زمان (مانند پراکندگی و میانگین نسبتاً ثابت)

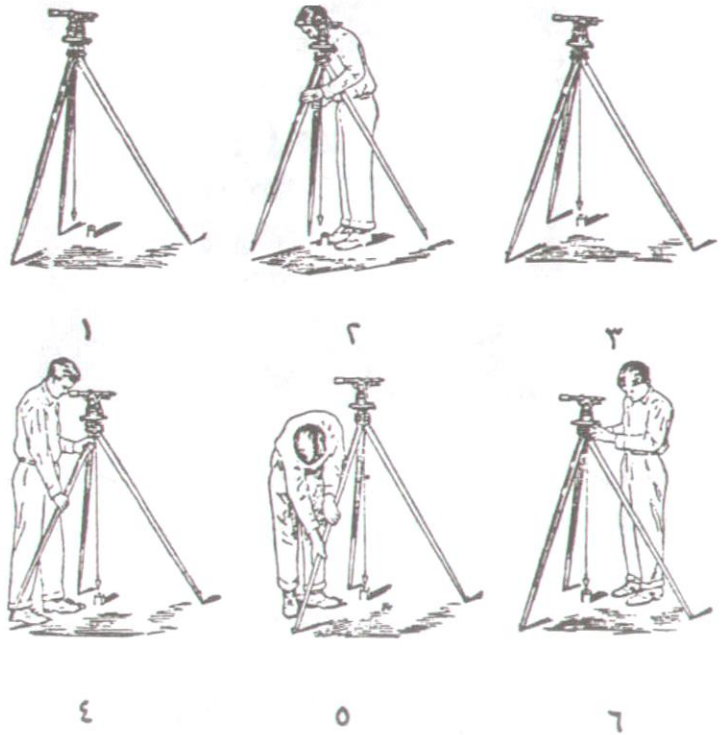
نا ایستا :
فرآیند هایی که خواص آماری با گذشت زمان در آنها ثابت نخواهد ماند.

آنچه ما در نقشه برداری اغلب با آنها سرو کار داریم پدیده های می باشند که یا واقعاً ایستا هستند و یا اینکه فرض ایستایی با تقریب بسیار خوب برای آنها اعتبار دارد.

مفهوم اندازه گیری و خطا

در اغلب اوقات افراد مختلف حتی متخصصین اندازه گیری ها نیز واژه های "اندازه گیری" و "مشاهده" با مفهوم یکسانی بکار می برند. اما در اصل باید توجه نمود که انجام هر اندازه گیری نیازمند یکسری مشاهدات است و هیچ اندازه ای حاصل نمی شود مگر قبل از آن چیزی مشاهده گردد. بنابراین عمل "مشاهده" مقدم بر "اندازه گیری" است و هر "اندازه گیری" حاصل حداقل یک "مشاهده" می باشد. از آنجا که در مهندسی نقشه برداری و برخی علوم و فنون دیگر وابستگی شدیدی بین مشاهدات و اندازه گیری ها وجود دارد، عبارات "اندازه گیری" و "مشاهده" اغلب به صورت معادل و مترادف استعمال می شوند.

ممکن است اندازه گیری در ذهن ما به صورت یک عمل منفرد تصور شود، ولی حقیقت موضوع کمی پیچیده تر از این است و ممکن است در یک نقشه برداری دقیق هر اندازه گیری شامل چندین مرحله اساسی نظیر استقرار (سانتراژ)، نشانه روی، تطابق و قرائت باشد. با این حال در پایان تمام این مراحل تنها یک مقدار عددی برای بیان "اندازه گیری" یا "مشاهده" کمیت مورد نظر ارائه می شود.



مفهوم اندازه گیری و خطا

برای مثال عمل نسبتاً ساده اندازه گیری یک فاصله 20 متری بین دو نقطه را با استفاده از یک متر نواری فولادی 30 متری در نظر بگیرید. برای اندازه گیری دقیق فاصله بین این دو نقطه مراحل اساسی زیر باید انجام شوند.

1- نوار متر در نزدیکی دو سر ابتدایی و انتهایی فاصله مورد نظر، بالای سطح زمین و تقریباً در امتداد مستقیم بین دو نقطه نگهداشته می شود.

2- روی نقطه اول و در مجاورت نوار متر، یک شاقول به حالت آویزان نگهداشته می شود.

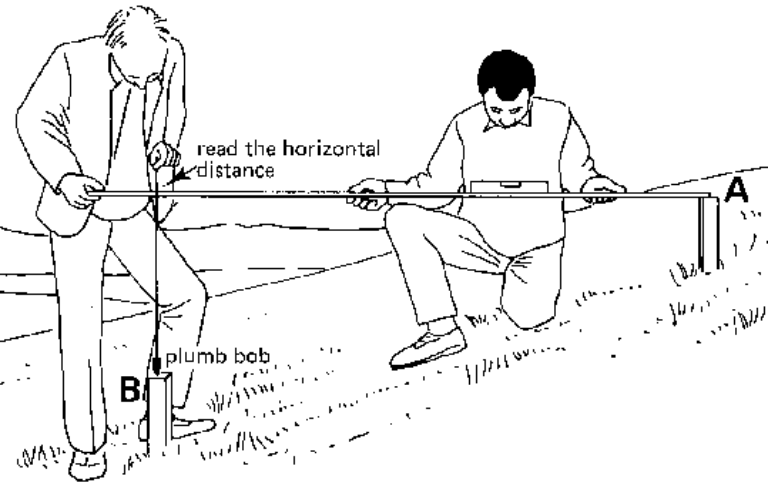
3- روی نقطه دوم و در مجاورت نوار متر، یک شاقول به حالت آویزان نگهداشته می شود.

4- در حالیکه هر دو شاقول روی دو نقطه حفظ می شوند، متر به حالت کشیده در می آید.

5- عدد مجاور نخ شاقول اول از روی متر قرائت و ثبت می شود.

6- عدد مجاور نخ شاقول دوم از روی متر قرائت و ثبت می شود.

7- با کم کردن دو قرائت مربوط به مراحل 5 و 6 فاصله بین دو نقطه بدست می آید.



ملاحظه می شود که برای چنین اندازه گیری ساده ای چندین مرحله اساسی شامل استقرار، کشش متر و قرائت انجام می گیرد که همه آنها برای بدست آوردن فاصله دقیق بین دو نقطه لازم و ضروری هستند. بنابراین اندازه گیری فاصله بین دو نقطه با روش متر کشی فقط یک قرائت واحد نیست و در واقع نتیجه چندین مرحله در فرآیند مشاهده و اندازه گیری کمیت مورد نظر است.

مراحل 1 تا 7 ممکن است هنوز تمام فرایند مربوط به بدست آوردن یک اندازه گیری قابل اعتماد از فاصله مورد نظر نباشند. به عبارت دیگر ممکن است برای خیلی از مقاصد، مقدار عددی حاصل از مرحله 7 مناسب باشد، اما برای مقاصدی دیگر اینطور نباشد. چنانچه نتیجه بهتر و دقیقتری برای فاصله بین دو نقطه مورد نظر باشد، می بایست علاوه بر تکرار مراحل اندازه گیری، مقدار حاصل از مرحله 7 برای موارد مختلفی همچون خطای طول واقعی متر، انبساط طولی، انکسار و کمانی شدن متر نیز تصحیح گردد.

مفهوم اندازه گیری و خطا

اندازه گیری، فرایندی است توأم با تغییرات که ممکن است به دلایل مختلفی اتفاق بیفتند. به عنوان مثال چنانچه فاصله بین دو نقطه چندین بار بوسیله یک متر فولادی اندازه گیری شود و در همان حال تغییرات دمایی نیز وجود داشته باشد، متناظر با تغییرات دمایی شاهد تغییر در طول متر فولادی و در نتیجه تغییر در قرائت اعداد خواهیم بود. چنانچه بجز تغییر دما هیچ عامل دیگری وجود نداشته باشد، در آن صورت تغییرات موجود در اندازه گیری ها نشان دهنده تغییرات دمایی خواهند بود. اما واقعیت این است که عوامل بسیار زیادی در تغییرات اندازه گیری ها مؤثر هستند که دما یکی از آنهاست و عوامل دیگری نظیر نقص و محدودیتهای دستگاه های اندازه گیری و محدودیت توانایی عامل مشاهده برای سانتراژ، نشانه روی، تطابق و قرائت وجود دارند. بنابراین به دلیل تغییرات کوچکی که در هر مرحله از اندازه گیری اتفاق می افتد، نتیجه هیچ مشاهده ای نمی تواند عیناً تکرار شود.

از آنجا که تمام اندازه گیری ها توأم با تغییرات هستند، لذا کمیت مورد نظر بطور کامل و قطعی قابل تعیین نیست. حالت ایده آل آن است که برای کمیت مورد نظر به یک مقدار کاملاً ثابت و بدون تغییر دست یابیم که به آن "مقدار واقعی" گویند. اما در حقیقت آنچه ما بدست می آوریم چیزی جز یک "برآورد" از مقدار واقعی نیست. با توجه به اینکه تغییر در اندازه گیری ها یا مشاهدات دارای ماهیتی اتفاقی است و یک امری طبیعی تلقی می شود، از لحاظ ریاضی اندازه گیری یا مشاهده می بایست بعنوان یک متغیر در نظر گرفته شود. چنین متغیری در علم آمار و احتمال، متغیر تصادفی تصادفی نامیده می شود که در مورد آن بیشتر صحبت خواهد شد.

اگر τ بیانگر مقدار واقعی یک کمیت (مانند طول یا زاویه) و x_i بیانگر یک اندازه گیری از آن باشد، آنگاه خطای اندازه گیری x_i بصورت زیر تعریف می شود.

$$\varepsilon_i = x_i - \tau$$

از آنجا که به طور معمول ما هرگز مقدار τ را نمی دانیم، لذا مقدار واقعی ε_i را نیز دست نیافتنی است. اما اگر بتوان برآورد خوبی از τ بدست آورد، آنگاه می توان از آن بعنوان یک معیار مناسب به منظور مطالعه تغییرات مشاهدات استفاده نمود. چنانچه برآورد τ را با $\hat{x}$ نمایش دهیم، آنگاه اختلاف بین برآورد $\hat{x}$ و مقدار مشاهده x_i را باقیمانده r_i نامیده و بصورت زیر تعریف می نمائیم.

$$r_i = x_i - \hat{x}$$

انواع خطاها

-خطا های بزرگ یا اشتباهات
-خطاهای سیستماتیک
-خطاهای اتفاقی یا تصادفی

عوامل مهم بروز این خطاها را در سه دسته زیر می توان بیان نمود:

- **عوامل طبیعی :** شامل کرویت زمین، شکست نور، وزش باد، تشعشع آفتاب، و تغییرات دمای هوا
- **عوامل دستگاهی :** شامل نقص دستگاه ها، تنظیم نبودن و یا پایین بودن ازرش تقسیمات آنها
- **عوامل انسانی :** شامل نارسایی حواس انسانی، کم دقتی، نداشتن تجربه و تسلط در کار

خطاهای بزرگ یا اشتباهات

خطاهای بزرگ نتیجهٔ اشتباهاتی هستند که نه لزوما ولی به‌طورعمده از بی‌توجهی، فراموشی و بی‌تجربگی عامل اندازه‌گیری ناشی‌می‌شوند. برای مثال، ممکن است عامل اندازه‌گیری اشتباهی به‌هدفی دیگر نشانه‌روی‌کند یا ممکن است در قرائت یا ثبت مشاهدهٔ خود دچار اشتباه شود، مثلاً ۴۱/۵۶ را به‌جای ۴۱/۶۵ ثبت‌کند. چنان‌چه عامل اندازه‌گیری فردی بی‌توجه و بی‌دقت باشد، بی‌شک اشتباهات زیادی در مشاهدات به‌وجودخواهدآمد. خوشبختانه اشتباه، معمولا بزرگ و قابل‌تشخیص است و با مراقبت می‌توان از وقوع آن جلوگیری‌کرد یا در صورت وقوع، محل آن‌را شناسایی و آن‌را حذف یا تصحیح‌نمود. به‌هرحال در برآورد کمیت‌های موردراندازه‌گیری، وجوداشتباهات غیرقابل‌قبول است و می‌باید پیش‌از برآورد و تجزیه‌وتحلیل نتایج، این‌گونه خطاهای بزرگ، شناسایی و از فهرست اندازه‌گیری‌ها حذف‌شوند.

معمولاً در پروژه‌های نقشه‌برداری، که توجه و تاکید زیاد بر سرعت انجام کار و جنبه‌های سودمندی و اقتصادی آن‌ها باشد، خطاهای بزرگ و اشتباهات نیز بیشتر اتفاق‌می‌افتند. از این‌رو روش‌های صحرایی مناسب به‌منظور کمک در کشف و تعیین اشتباهات وجوددارند که در ادامه به‌برخی از آن‌ها اشاره‌می‌شود:

- ✓ کنترل دقیق تمام قرائت‌ها روی نشانه‌های نقشه‌برداری
- ✓ تکرار قرائت‌ها و کنترل سازگاری معقول و منطقی بین آن‌ها
- ✓ تصدیق اعداد ثبت‌شده با مقیاس اندازه‌گیری‌ها
- ✓ تکرار کامل اندازه‌گیری‌ها به‌طورمستقل، و کنترل آن‌ها به‌منظور وجودسازگاری بین آن‌ها
- ✓ استفاده از ابزار کنترل هندسی و جبری ساده، مانند مجموع زاویای داخلی یک‌چندضلعی

اتخاذ روش‌های بالا برای جلوگیری از بروز اشتباهات بسیارمهم است ولی دوباره تاکیدمی‌شود چنان‌چه به‌هردلیل، اشتباهی‌رخ‌دهد، می‌باید پیش‌از استفاده از اندازه‌گیری‌ها از لیست مشاهدات حذف‌شود.

خطاهایی که بر اساس روندهای قطعی و تا حد امکان شناخته‌شده اتفاق‌افتاده و با مدل‌ها و روابط ریاضی بیان‌شدنی باشند، خطاهای سیستماتیک نامیده‌می‌شوند. برای مثال دربارهٔ تغییر طول یک‌متر نواری فولادی با ضریب انبساط طولی مشخص، یک‌رابطهٔ تابعی بین دما و تغییرطول نوار فولادی به‌صورت خطی ارائه‌شده‌است. بنابراین اگر طول متر نواری در دمای استاندارد معلوم‌باشد، تغییر طول نوار از حالت استاندارد، به‌دلیل انحراف از دمای استاندارد، به‌عنوان خطای سیستماتیک شناخته‌می‌شود.

ماهیت خطاهای سیستماتیک به‌گونه‌ای است که چنان‌چه اندازه‌گیری‌ها تحت‌شرایطی کاملاً یک‌سان تکرارشوند، خطاهای سیستماتیک نیز عیناً تکرارمی‌گردند و آشکارنمی‌شوند. برای مثال در تکرار اندازه‌گیری یک‌فاصله با متر نواری فولادی، چنان‌چه از متر فولادی و عاملان اندازه‌گیری یک‌سان تحت‌شرایط مشابه دما، کشش، شیب و. . . استفاده‌شود، در آن‌صورت، خطاهای سیستماتیک نیز عیناً تکرارخواهندشد.

اگر مقدار و علامت خطایی سیستماتیک در سراسر فرآیند اندازه‌گیری ثابت باقی‌بماند به‌آن خطای «ثابت» اطلاق‌می‌شود. سازوکاری که رفتار خطاهای سیستماتیک تحت‌آن‌شکل‌می‌گیرد، ممکن است به‌عامل اندازه‌گیری، وسیلهٔ اندازه‌گیری، شرایط فیزیکی و محیطی در زمان اندازه‌گیری یا هر نوع ترکیبی از این عوامل وابسته‌باشد.

ساختار غیرکامل دستگاه یا نقص در تنظیم دستگاه، خطای دستگاهی نامیده‌می‌شود که در زمرهٔ خطاهای سیستماتیک قرارمی‌گیرد. ساختار غیرکامل شامل مواردی هم‌چون نایک‌نواختی مقیاس درجه‌بندی و خروج از مرکزیت در سوارکردن قطعات دستگاه است. تنظیم دستگاهی ناقص نیز شامل مواردی هم‌چون عمودنبودن محور تلسکوپ زاویه‌یاب بر محور افقی دستگاه است.

ازآن‌جاکه عموما مشاهدات نقشه‌برداری در صحرا انجام‌می‌گیرند، از بسیاری عوامل فیزیکی و محیطی متأثراند. برای مثال دما، کشش، و شیب سطح زمین در مترکشی فواصل موثراند. درحالی‌که علاوه بر دما، رطوبت، و فشار هوا نیز در اندازه‌گیری فواصل با طول‌یاب‌های الکترونیک، اندازه‌گیری زوایا و ترازیابی مؤثراند. همهٔ این اثرات به‌صورت تابعی از عوامل مذکور قابل‌بیان هستند و بنابراین به‌عنوان خطاهای سیستماتیک طبقه‌بندی‌می‌شوند.

تمام منابع خطاهای سیستماتیک که تا کنون موردبحث قرارگرفتند، مستقیماً با عمل اندازه‌گیری مرتبط‌اند. اما خطاهای سیستماتیک از طریق ساده‌سازی مدل‌های هندسی و ریاضی انتخابی نیز ممکن‌است اتفاق بیفتند. برای مثال، اگر به‌جای یک‌مثلث کروی یک‌مثلث مسطحاتی برای اتصال سه‌ایستگاه نقشه‌برداری با فواصل چندکیلومتری مورداستفاده قرارگیرد، اضافه کرویت به‌عنوان یک‌خطای سیستماتیک ظاهرخواهدشد.

با توجه به‌مطالب مطرح‌شده، بایدگفت در پردازش و تجزیه‌وتحلیل اندازه‌گیری‌های نقشه‌برداری، باید تا حد امکان خطاهای سیستماتیک، شناسایی، و تصحیح‌شوند. برای این‌منظور دو راه وجوددارد:

-روش های صحرایی

-روش های محاسباتی

در روش‌های صحرایی با شناختی که از ماهیت هرخطای سیستماتیک وجوددارد روشی متناسب برای مقابله با آن‌اتخاذمی‌شود. برای مثال می‌توان به‌قراردادن ترازیاب در فاصله‌ای مساوی از دو شاخص اشاره‌کرد که باعث حذف خطاهای سیستماتیک انکسار، کرویت، و کلیماسیون می‌شود. در روش‌های محاسباتی یا دفتری نیز با داشتن مدل یا رابطهٔ ریاضی خطای سیستماتیک موردنظر، مانند خطاهای سیستماتیک ناشی از تغییرات جوّی، کمانی‌شدن متر نواری یا شیب زمین، مقادیر تصحیح مربوط محاسبه و به‌اندازه گیری‌ها اعمال‌می‌شود.

با فرض این‌که تمام اشتباهات شناسایی و حذف‌شوند و تصحیحات لازم نیز برای تمام خطاهای سیستماتیک شناخته‌شده اعمال‌گردند، هنوز برخی تغییرات در اندازه‌گیری‌ها دیده‌خواهندشد. این تغییرات باقی‌مانده، به‌طورعمده ناشی از شرایط محیطی و عوامل ناشناخته‌اند که دیگر دارای یک‌رابطهٔ مشخص ریاضی بر مبنای یک‌فرآیند قطعی نیست؛ اغلب با تکرار اندازه‌گیری‌ها خود را نشان‌می‌دهند و از لحاظ وقوع، نامرتب و از نظر بزرگی متغیراند، به‌طوری‌که در اندازه‌گیری مکرر به‌صورت نامعین به‌دنبال هم می‌آیند. بنابراین این نوع خطاها دارای رفتاری کاملا تصادفی هستند و لازم است بر همین‌اساس با آن‌ها برخوردشود.

پیش‌تر اشاره‌شد که یک‌اندازه‌گیری یا مشاهده از دیدگاه ریاضی یک‌متغیر است. ازآن‌جاکه هر اندازه‌گیری شامل مولفه‌های خطا با رفتار تصادفی است، لذا به‌وضوح می‌توان آن‌را یک«متغیر تصادفی» فرض‌نمود. بدیهی است که خطاهای تصادفی، خودشان متغیرهای تصادفی‌اند.

به‌خاطرداشته‌باشیم که با توجه به‌ماهیت این‌نوع خطاها، هیچ‌گاه نمی‌توانیم آن‌ها را به‌صفر برسانیم.

بر خلاف خطاهای سیستماتیک، که در برخورد با آن‌ها از مدل‌های ریاضی یا روابط تابعی استفاده‌می‌شود، برای خطاهای تصادفی از مدل‌های احتمال استفاده‌می‌شود.

پایان جلسه اول

درس تئوری خطاها

جلسه دوم

فرید اسماعیلی
Farid_63@yahoo.com
www.faridesm.ir

تماس با استاد از طریق پست الکترونیکی
مشاهده اطلاعیه ها، نمرات، دریافت فایل ها در وب سایت

- روشهایی که به وسیله آنها می توان اطلاعات جمع آوری شده را تنظیم، طبقه بندی و خلاصه نمود و آنها را به وسیله نمودار هایی نمایش داد به آمار توصیفی مرسوم است.

جمعیت مجموعهٔ تمام افراد یا اشیایی که مطالعات آماری در مورد یک یا چند صفت آنها در یک مکان و زمان معین انجام می‌گیرد به جمعیت موسوم است. هر یک از این افراد یا اشیا را یک عضو جمعیت می‌نامند و تعداد اعضای جمعیت را اندازهٔ جمعیت می‌نامند.

نمونه زیر مجموعه‌ای از جمعیت که طبق یک قاعده و ضابطهٔ خاصی برای مطالعهٔ صفتی از جمعیت انتخاب می‌شود را یک نمونه گویند. تعداد اعضای نمونه به اندازهٔ نمونه موسوم است.

داده‌ها در یک بررسی آماری، بایستی صفت مورد مطالعه را به صورت اعداد و ارقام نمایش دهیم. اگر صفت مورد مطالعه کمی، مانند وزن، حجم، درجهٔ حرارت و غیره باشد آنگاه این عمل به سادگی با اندازه‌گیری امکان‌پذیر است. اما اگر صفت مورد مطالعه کیفی، مانند گروه خون، شغل، رنگ چشم و غیره باشد آنگاه بایستی با یک قاعده معین این مسائل کیفی را با اعداد و ارقام نشان داد. در هر صورت این اعداد و ارقام را داده‌ها گویند که به دو صورت گسسته و پیوسته می‌باشند.

داده‌های گسسته داده‌هایی هستند که بین دو مقدار متصور آنها هیچ عدد دیگری وجود نداشته باشد، مانند تعداد فرزندان یک خانواده که شامل مقادیر ۰، ۱، ۲ و ... است و همچنین صفت شغل افراد که به آن مثلاً اعداد ۱، ۲، ۳ و ... را نسبت می‌دهیم و بین این مقادیر عدد دیگری در رابطه با صفت مورد نظر وجود ندارد. داده‌های پیوسته داده‌هایی هستند که بین هر دو مقدار متصور آنهاهمواره عدد دیگری وجود دارد، مانند وزن افراد که بین دو نفر با وزنهای نزدیک به هم همواره می‌توان فردی را با وزنی بین وزن دو فرد یاد شده در جمعیت یافت. از جمله داده‌های گسسته می‌توان داده‌های مربوط به صفات گروه خون، رنگ، نژاد، شغل، تعداد کالاهای تولیدی و غیره را برشمرد و از جمله داده‌های پیوسته می‌توان داده‌های مربوط به صفات وزن، طول قد، فشار گاز، قطر لوله تولیدی یک کارخانه و غیره را برشمرد.

- در نقشه برداری ما عموماً با کمیت ها و داده های پیوسته سرکار داریم.

نخستین گام در خلاصه کردن داده‌ها، طبقه‌بندی و تنظیم آنهادر یک جدول موسوم به جدول آماری است. یک جدول آماری بایستی به نحوی تنظیم شود که بتوان از آن به راحتی اطلاعات نهفته در داده‌ها را استخراج کرد. متداولترین جدول آماری جدول فراوانی است که در آن داده‌ها، تعداد موجود از هر داده و درصد موجود از هر داده مشخص می‌شود. بنابراین یک جدول فراوانی شامل موارد زیر است

الف– فراوانی و فراوانی نسبی فرض کنید n داده از k نوع (k طبقه) داشته باشیم و تعداد این داده‌ها در این k طبقه به ترتیب f_1 ، f_2 ، ... و f_k باشند. به f_1 ، f_2 ، ... و f_k فراوانی‌های طبقات و به $r_1=\frac{f_1}{n}$ ، $r_2=\frac{f_2}{n}$ ، ... و $r_k=\frac{f_k}{n}$ فراوانی‌های نسبی طبقات گوئیم. واضح است که

$$1\le f_i\le n \quad , \quad 0\le r_i\le 1 \quad , \quad \forall i=1,2,\dots,k$$

$$f_1+f_2+\dots+f_k=\sum_{i=1}^{k} f_i=n \quad , \quad r_1+r_2+\dots+r_k=\sum_{i=1}^{k} r_i=1$$

ب- فراوانی تجمعی و فراوانی تجمعی نسبی اگر در هر طبقه فراوانی آن طبقه و طبقات قبل از آن را با یکدیگر جمع کنیم فراوانی تجمعی آن طبقه به دست می‌آید و اگر در هر طبقه فراوانی نسبی آن طبقه و طبقات قبل از آن را با یکدیگر جمع کنیم فراوانی تجمعی نسبی آن طبقه حاصل می‌شود، یعنی

$$\text{فراوانی تجمعی طبقهٔ زام} = g_j = f_1 + f_2 + \ldots + f_j = \sum_{i=1}^{j} f_i$$

$$\text{فراوانی تجمعی نسبی طبقه زام} = s_j = r_1 + r_2 + \ldots + r_j = \sum_{i=1}^{j} r_i$$

واضح است که $g_k = n$ و $s_k = 1$

یک جدول فراوانی جدولی است که ستونهای آن شامل نوع داده‌ها (طبقات)، فراوانی، فراوانی نسبی، فراوانی تجمعی و فراوانی تجمعی نسبی داده‌ها باشد.

مثال صنعتگری چهار نوع قطعه A، B، C و D تولید می‌کند. اگر او در یک روز ۲۰ قطعه از این قطعات را به شرح زیر تولید کرده باشد

B, C, C, A, D, C, C, B, D, C, A, C, D, C, B, C, C, B, D, D

یک جدول فراوانی برای این قطعات تشکیل دهید.

حل ابتدا چهار نوع قطعه A، B، C و D را به ترتیب با اعداد ۱، ۲، ۳ و ۴ متناظر می‌کنیم. این اعداد تنها برای نامگذاری این قطعات است و نمی‌توان روی آنها چهار عمل اصلی حساب را انجام داد. سپس بوسیلهٔ شمارش، فراوانی هر قطعه را محاسبه کرده و با محاسبهٔ فراوانیِ نسبی، فراوانی تجمعی و فراوانی تجمعی نسبی، جدول فراوانی به صورت جدول ۱.۱ را به دست می‌آوریم.

نوع قطعه	x_i	f_i	r_i	g_i	s_i
A	۱	۲	۰/۱	۲	۰/۱
B	۲	۴	۰/۲	۶	۰/۳
C	۳	۹	۰/۴۵	۱۵	۰/۷۵
D	۴	۵	۰/۲۵	۲۰	۱
جمع		۲۰	۱/۰		

جدول ۱.۱ جدول فراوانی ۲۰ قطعهٔ تولیدی صنعتگر در یک روز

از این جدول اطلاعات زیادی را می‌توان استخراج کرد. برای مثال عدد ۰/۲ در ستون فراوانی نسبی به معنای آن است که ۲۰٪ از قطعات تولیدی آن روز از نوع B می‌باشند و عدد ۰/۳ در ستون فراوانی تجمعی نسبی به معنای آن است که ۳۰٪ از قطعات تولیدی آن روز از نوع A یا B می‌باشند.

مثال وزنهای ۴۰ قالب کره که به نزدیکترین عدد صحیح گرد شده‌اند به قرار زیر است

۳۳	۴۶	۳۸	۳۴	۵۱	۳۶	۴۷	۲۴	۳۵	۵۲
۴۰	۴۲	۴۰	۴۱	۳۴	۴۷	۵۰	۳۸	۳۶	۴۷
۲۱	۴۱	۳۷	۳۷	۳۵	۳۰	۳۲	۳۰	۲۹	۲۶
۴۳	۳۴	۳۱	۴۳	۲۳	۴۵	۳۵	۲۶	۳۰	۳۱

یک جدول فراوانی برای این داده‌ها تشکیل دهید.

حل داده‌ها از نوع پیوسته می‌باشند. در این حالت داده‌ها را به تعدادی رده (فاصله) با طول مساوی تقسیم کرده و در هر رده فراوانی داده‌ها را می‌شماریم. روش به دست آوردن تعداد رده‌ها و طول هر رده به ترتیب در زیر آورده شده است.

۱– برای به دست آوردن تعداد رده‌ها یک قاعدهٔ عمومی وجود ندارد و معمولاً تعداد رده‌ها را بین ۵ تا ۲۵ رده اختیار می‌کنند. یک قاعدهٔ مفید استفاده از دستور استورگس (Sturges) است که در آن تعداد رده k از رابطهٔ زیر محاسبه می‌شود

$$k = ۱ + ۳/۳۲۲ \log ۴۰$$

$$k = ۱ + ۳/۳۲۲ \log_{۱۰}^{n}$$

که در آن n تعداد کل داده‌هاست. چون حاصل این عدد اعشاری است آن را به عدد صحیح بزرگتر از آن گرد می‌کنند. بنابراین در این مثال داریم که

$$k = ۱ + ۳/۳۲۲ \log_{۱۰}^{۴۰} = ۶/۳۲۲ \simeq ۷$$

۲– با توجه به اینکه وزنها به نزدیکترین عدد صحیح گرد شده‌اند بنابراین عدد ۳۵ در داده‌ها در واقع عددی در فاصلهٔ (۳۴/۵–۳۵/۵) می‌باشد. عدد ۰/۵ را میزان تغییر پذیری مقادیر داده‌ها نامیده و به صورت زیر آن را محاسبه می‌کنیم

$$\text{میزان تغییر پذیری داده‌ها} = S = \frac{\text{واحد گرد شدهٔ داده‌ها}}{۲} = \frac{۱}{۲} = ۰/۵$$

۳– کوچکترین، بزرگترین و دامنهٔ واقعی داده‌ها را به صورت زیر محاسبه می‌کنیم

$$\min = \text{کوچکترین داده} - S = ۲۱ - ۰/۵ = ۲۰/۵$$

$$\max = \text{بزرگترین داده} + S = ۵۲ + ۰/۵ = ۵۲/۵$$

$$\text{دامنه} = R = \max - \min = ۵۲/۵ - ۲۰/۵ = ۳۲$$

۴- طول هر رده را از تقسیم دامنهٔ R بر تعداد رده k به دست می‌آوریم و عدد حاصل را که ممکن است دارای چند رقم اعشار باشد مطابق واحد گرد شدهٔ داده‌ها به عدد بالاتر گرد می‌کنیم. بنابراین

$$\text{طول رده} = w = \frac{R}{k} = \frac{۳۲}{۷} = ۴/۵۷۱۴ \approx ۵$$

چون داده‌ها به عدد صحیح گرد شده‌اند و اولین عدد صحیح بزرگتر از ۴/۵۷۱۴ عدد ۵ می‌باشد پس طول رده به ۵ گرد شده است.

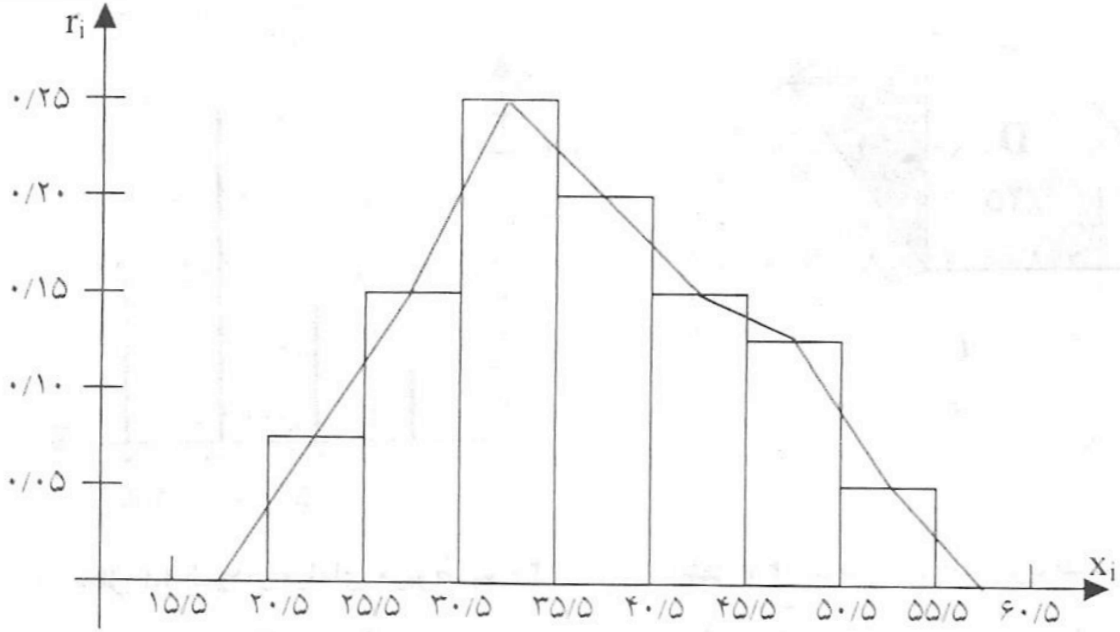
حال بایستی ۷ رده هر یک به طول ۵ را تشکیل دهیم. برای این منظور در جدول فراوانی یک ستون به نام رده‌ها ایجاد می‌کنیم و در آن اولین رده به طول ۵ را به صورت ۲۰/۵-۲۵/۵ (با شروع از min با طول w) در نظر گرفته و ردهٔ بعدی را به صورت ۲۵/۵-۳۰/۵ (شروع از انتهای ردهٔ قبل با طول w) در نظر می‌گیریم و این عمل را تا تشکیل ۷ رده ادامه می‌دهیم.

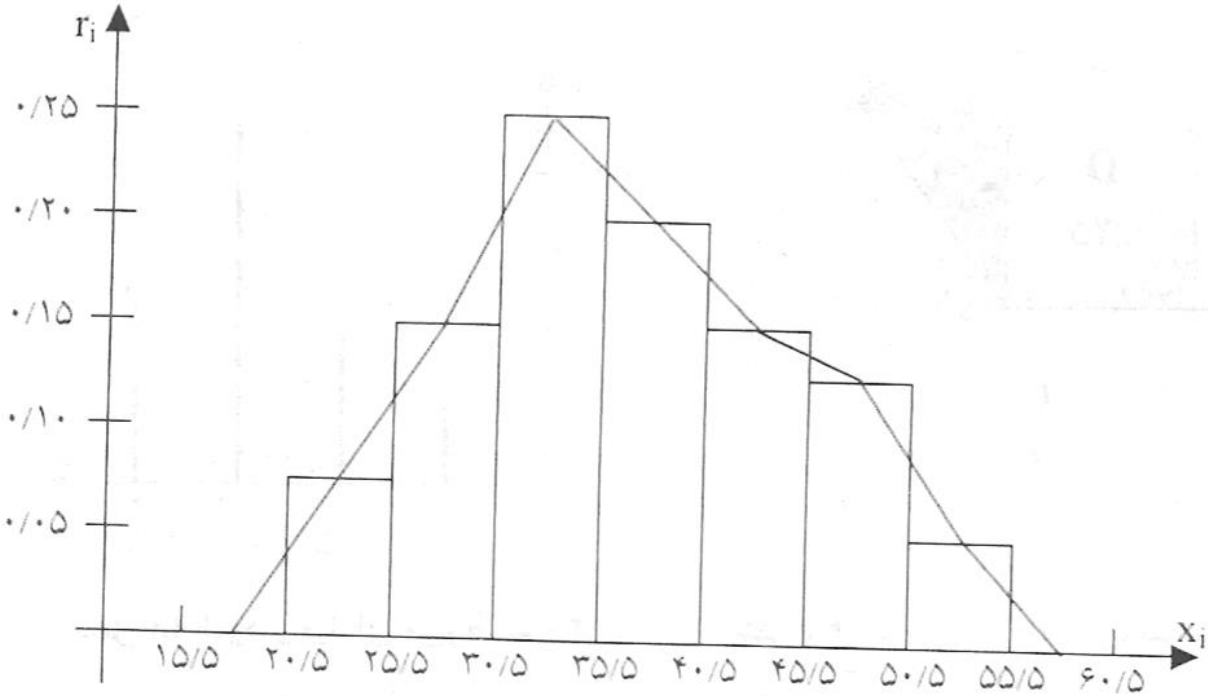
رده‌ها
۲۰/۵-۲۵/۵
۲۵/۵-۳۰/۵
۳۰/۵-۳۵/۵
۳۵/۵-۴۰/۵
۴۰/۵-۴۵/۵
۴۵/۵-۵۰/۵
۵۰/۵-۵۵/۵
جمع

فراوانی هر طبقه در جدول فراوانی از ستونی به نام خط و نشان استفاده می‌کنیم که در این ستون با خط زدن هر داده و مشخص کردن مکان آن در طبقهٔ مربوطه بوسیلهٔ یک نشان، فراوانی داده‌ها در طبقات مشخص می‌گردند. توجه کنید که اگر بوسیلهٔ انجام مراحل فوق آخرین رده دارای فراوانی صفر باشد، آن ردهٔ آخر را حذف می‌کنیم. در جدول فراوانی داده‌های پیوسته نقاط وسط رده‌ها را با x_i نمایش می‌دهیم و آن را نمایندهٔ رده می‌نامیم. با انجام عملیات گفته شده در بالا جدول فراوانی مربوط به وزن قالبهای کره به صورت جدول به دست می‌آید

رده‌ها	خط و نشان	x_i	f_i	r_i	g_i	s_i
۲۰/۵–۲۵/۵	|||	۲۳	۳	۰/۰۷۵	۳	۰/۰۷۵
۲۵/۵–۳۰/۵	卌 |	۲۸	۶	۰/۱۵	۹	۰/۲۲۵
۳۰/۵–۳۵/۵	卌 卌	۳۳	۱۰	۰/۲۵	۱۹	۰/۴۷۵
۳۵/۵–۴۰/۵	卌 |||	۳۸	۸	۰/۲۰	۲۷	۰/۶۷۵
۴۰/۵–۴۵/۵	卌 |	۴۳	۶	۰/۱۵	۳۳	۰/۸۲۵
۴۵/۵–۵۰/۵	卌	۴۸	۵	۰/۱۲۵	۳۸	۰/۹۵
۵۰/۵–۵۵/۵	||	۵۳	۲	۰/۰۵	۴۰	۱/۰۰
جمع			۴۰	۱/۰۰		

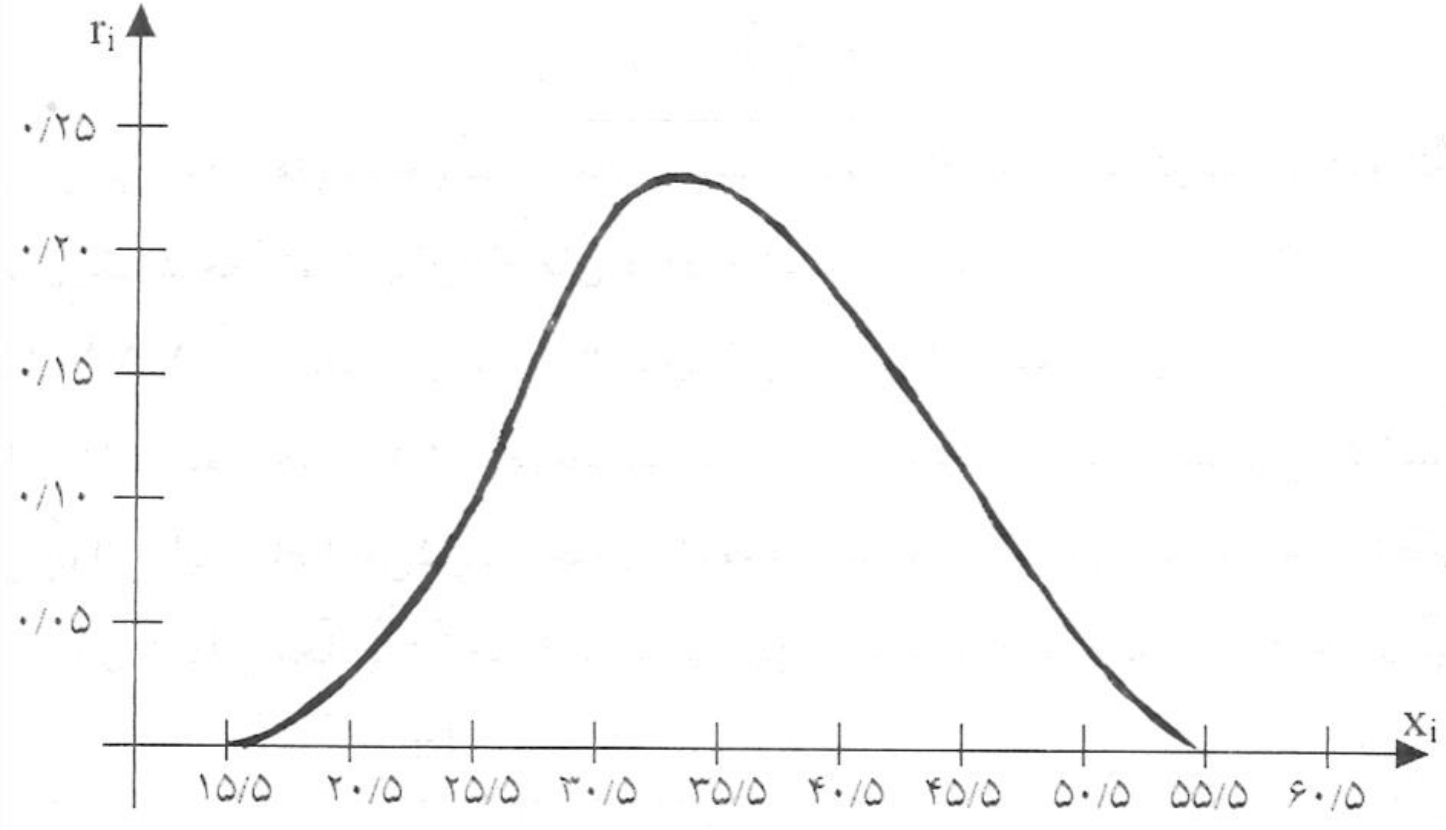
الف- هیستوگرام (نمودار ستونی) هیستوگرام نموداری متشکل از تعدادی مستطیل است که تعداد این مستطیل‌ها برابر تعداد رده‌های جدول فراوانی می‌باشد. قاعده هر مستطیل روی محور افقی قرار دارد و طول آن برابر طول واقعی رده است، که هر چه باشد آن را یک واحد در نظر می‌گیریم و مرکز آن نمایندهٔ رده است. ارتفاع هر مستطیل برابر فراوانی نسبی مربوط به آن رده است. برای مثال هیستوگرام مربوط به مثال در شکل رسم شده است. توجه کنید که چون عرض هر مستطیل برابر یک واحد در نظر گرفته شده و ارتفاع هر مستطیل برابر فراوانی نسبی ردهٔ مربوطه می‌باشد پس مجموع مساحت تمام مستطیل‌های هیستوگرام برابر یک واحد مربع است.





ب- چندبر فراوانی اگر وسط قاعده‌های بالای مستطیل‌های هیستوگرام را بوسیلهٔ خطوط مستقیم به طور متوالی به یکدیگر وصل کرده و ابتدای آن را به وسط ردهٔ ماقبل و انتهای آن را به وسط ردهٔ مابعد مستطیل‌های هیستوگرام وصل کنیم یک چند ضلعی بوجود می‌آید که آن را چندبر فراوانی گویند و مساحت زیر این چندبر فراوانی نیز یک واحد مربع است.

ج- منحنی فراوانی اگر تعداد داده‌ها زیاد و طول رده‌ها کوچک باشند در این صورت تعداد رده‌ها زیاد شده و در نتیجه تعداد اضلاع چندبر فراوانی نیز زیاد می‌شود و به یک منحنی نزدیک می‌شود. مساحت زیر این منحنی یک واحد مربع است و به آن منحنی فراوانی گویند.



مثال

یک‌طول حدود ۸۱۰ متری را، ۲۰۰ بار اندازه‌می‌گیریم. نتایج حاصل به‌عنوان یک«نمونهٔ نصادفی»[۲] درنظرگرفته‌می‌شود. فرض‌کنیم هیچ خطای بزرگی در اندازه‌گیری‌ها وجودندارد و خطاهای سیستماتیک نیز تا دقت بهتر از یک‌سانتی‌متر تصحیح‌شده‌اند. نتیجهٔ اندازه‌گیری‌های تصحیح‌شده از ۸۱۰/۱۱ متر تا ۸۱۰/۲۳ متر نوسان‌دارند که مطابق جدول ۱.۲ در ۱۳ دستهٔ[۳] یک‌سانتی‌متری به‌ترتیب از کم‌ترین مقدار تا بیشترین مقدار مرتب‌شده‌اند. چنان‌چه تعداد کل اندازه‌گیری‌ها را با N و تعداد دسته‌ها را با M نمایش‌دهیم، باید همواره $M \leq N$ باشد. تعداد اندازه‌گیری‌های واقع‌شده در هردسته را «فراوانی»[۴] می‌نامند و مجموع آن‌ها باید برابر تعداد کل اندازه‌گیری‌ها باشد.

مثال

جدول ۱.۲. فراوانی ۲۰۰ بار اندازه‌گیری یک‌طول

فراوانی	مقدار اندازه‌گیری (متر)
۱	۸۱۰/۱۱
۳	۸۱۰/۱۲
۷	۸۱۰/۱۳
۱۹	۸۱۰/۱۴
۲۰	۸۱۰/۱۵
۳۶	۸۱۰/۱۶
۳۸	۸۱۰/۱۷
۲۹	۸۱۰/۱۸
۲۴	۸۱۰/۱۹
۱۰	۸۱۰/۲۰
۱۱	۸۱۰/۲۱
۵	۸۱۰/۲۲
۲	۸۱۰/۲۳

به‌منظور ارزیابی و ترسیم پخش یا توزیع اندازه‌گیری‌ها برای هر دسته از کمیت «فراوانی نسبی»[1] استفاده‌می‌کنیم. فراوانی نسبی همان‌طور که از نامش پیداست با تقسیم فراوانی هردسته بر تعداد کل اندازه‌گیری‌ها به‌دست‌می‌آید.

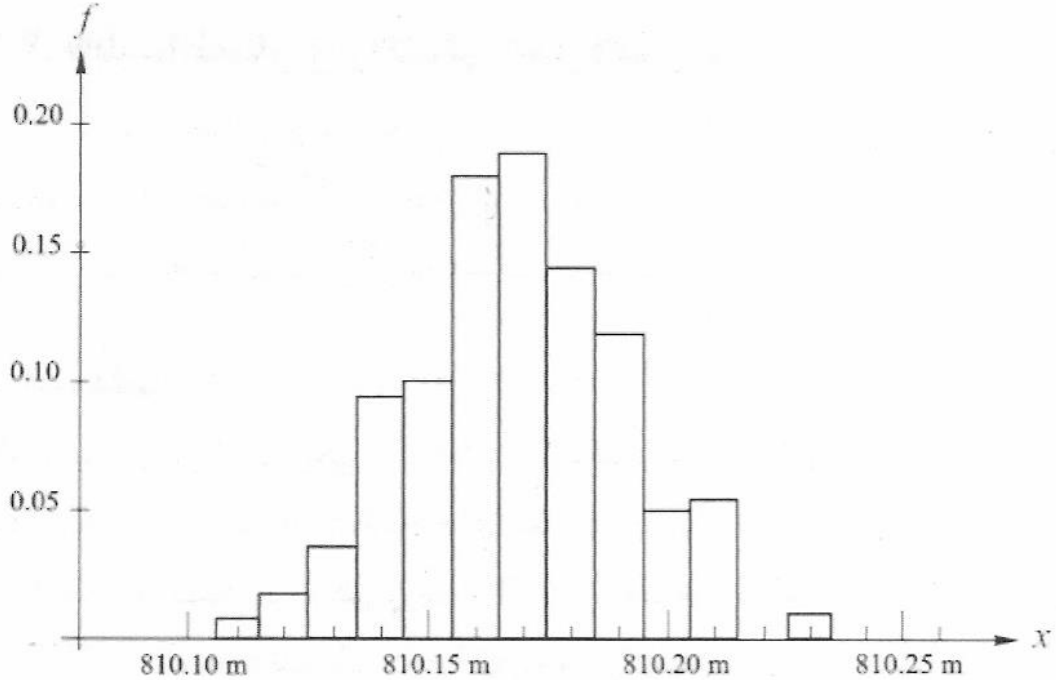
مثال

جدول ۲.۲. فراوانی نسبی ۲۰۰ بار اندازه‌گیری یک‌طول

فراوانی نسبی	مقدار اندازه‌گیری (متر)
۰/۰۰۵	۸۱۰/۱۱
۰/۰۱۵	۸۱۰/۱۲
۰/۰۳۵	۸۱۰/۱۳
۰/۰۹۵	۸۱۰/۱۴
۰/۱۰۰	۸۱۰/۱۵
۰/۱۸۰	۸۱۰/۱۶
۰/۱۹۰	۸۱۰/۱۷
۰/۱۴۵	۸۱۰/۱۸
۰/۱۲۰	۸۱۰/۱۹
۰/۰۵۰	۸۱۰/۲۰
۰/۰۵۵	۸۱۰/۲۱
۰/۰۲۵	۸۱۰/۲۲
۰/۰۱۰	۸۱۰/۲۳

بدیهی است با توجه به‌تعریف فراوانی نسبی، باید مجموع فراوانی‌های نسبی برابر با یک (عدد۱) شود. فراوانی نسبی به‌دست‌آمده در نگارهٔ ۱.۲ به‌صورت ستونی (یا مستطیلی) ترسیم‌شده‌است که به‌نمودار ستونی یا هیستوگرام معروف است.

مثال



نگارهٔ ۱.۲. نمودار ستونی فراوانی نسبی اندازه‌گیری طول

نگارهٔ ۱.۲ بیان‌گر یک‌نمونهٔ توزیع فراوانی در قالب نمودار ستونی (هیستوگرام) است که در آن‌عرض هرمستطیل بیان‌گر «فاصلهٔ دسته» و ارتفاع آن‌بیان‌گر فراوانی نسبی است. برای مثال، عرض بلندترین مستطیل در نگارهٔ ۱.۲ بیان‌گر دسته‌ای از اندازه‌گیری‌ها است که بین ۸۱۰/۱۶۵ متر تا ۸۱۰/۱۷۵ متر واقع‌شده‌اند و ارتفاع آن‌نشان‌دهندهٔ فراوانی نسبی متناظر با آن (۰/۱۹۰) است.

مطابق نگارهٔ ۱.۲ بیشترین فراوانی‌ها در حوالی مقدار ۸۱۰/۱۷ متر دیده‌می‌شود. چنان‌چه بتوان تعداد اندازه‌گیری‌ها را بی‌نهایت‌بار افزایش‌داد و فاصلهٔ دسته‌ها را بی‌نهایت‌کوچک درنظرگرفت، در آن‌صورت فراوانی‌های نسبی به‌حد پایدار نزدیک می‌شوند که این‌مقدارِ حدی، همان « احتمال»[1] است.

میانگین

مجموع داده ها تقسیم بر تعداد آنها را میانگین داده ها می گویند.

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

مثال : طولی 6 بار اندازه گیری شده و اعداد زیر بدست آمده است.

$$x_1 = 112.03m\,,\, x_2 = 112.00m\,,\, x_3 = 111.98m$$

$$x_4 = 112.03m\,,\, x_5 = 112.04m\,,\, x_6 = 111.98m$$

مطلوب است محاسبه:

الف) محتملترین (میانگین) مقدار این طول

$$\bar{x} = \frac{\sum x_i}{n} = \frac{x_1 + x_2 + x_3 + x_4 + x_5 + x_6}{6} = 112.01m$$

نما یا مد : داده ای که فراوانی آن از سایر داده های بیشتر باشد.

میانه

ب- میانه اگر داده‌ها را به طور غیر نزولی مرتب کنیم آنگاه عدد m را میانهٔ این داده‌ها گویند در صورتی که تقریباً نصف داده‌ها در سمت چپ و نصف داده‌ها در سمت راست این عدد قرار گیرند،

محاسبهٔ میانهٔ نمونه برای داده های گسسته

- مرتب کردن داده‌ها از کوچک به بزرگ
- اگر تعداد داده‌ها فرد باشد، میانه مقدار وسط می‌باشد.

داده‌ها: 2 8 3 4 1

داده‌های مرتب شده: 1 2 3 4 8

میانه

- اگر تعداد داده‌ها زوج باشد، میانه متوسط دو مقدار وسطی است

داده‌ها : 2 8 3 4 1 8

داده‌های مرتب شده: 1 2 3 4 8 8

میانه = (3+4)/2 = 3.5

محاسبهٔ میانهٔ نمونه برای داده های پیوسته

رده‌ها	خط و نشان	x_i	f_i	r_i	g_i	s_i
۲۰/۵–۲۵/۵	|||	۲۳	۳	۰/۰۷۵	۳	۰/۰۷۵
۲۵/۵–۳۰/۵	~~||||~~ |	۲۸	۶	۰/۱۵	۹	۰/۲۲۵
۳۰/۵–۳۵/۵	~~||||~~ ~~||||~~	۳۳	۱۰	۰/۲۵	۱۹	۰/۴۷۵
۳۵/۵–۴۰/۵	~~||||~~ |||	۳۸	۸	۰/۲۰	۲۷	۰/۶۷۵
۴۰/۵–۴۵/۵	~~||||~~ |	۴۳	۶	۰/۱۵	۳۳	۰/۸۲۵
۴۵/۵–۵۰/۵	~~||||~~	۴۸	۵	۰/۱۲۵	۳۸	۰/۹۵
۵۰/۵–۵۵/۵	||	۵۳	۲	۰/۰۵	۴۰	۱/۰۰
جمع			۴۰	۱/۰۰		

۱– ابتدا در جدول فراوانی نخستین رده‌ای که فراوانی تجمعی نسبی آن بزرگتر یا مساوی ۰/۵۰ باشد را در نظر می‌گیریم. این رده که میانه درون آن قرار دارد را ردهٔ میانه می‌نامیم. در این مثال ردهٔ میانه ۳۵/۵–۴۰/۵ می‌باشد.

$$m = L_{./۵} + \frac{(۰/۵n - g_{./۵})w}{f_{./۵}}$$

m = کران پائین ردهٔ میانه + x = ۳۵/۵ + $\frac{(۲۰-۱۹)۵}{۸}$ = ۳۵/۵ + ۰/۶۲۵ = ۳۶/۱۲۵

در این فرمول $L_{./۵}$ کران پائین ردهٔ میانه، n تعداد داده‌ها، $g_{./۵}$ فراوانی تجمعی ردهٔ قبل از ردهٔ میانه، $f_{./۵}$ فراوانی ردهٔ میانه و w طول رده می‌باشد.

بوسیلهٔ شاخص‌های تمرکز می‌توان میزان تمرکز داده‌ها را در یک عدد خلاصه کرد. اما یک مسئله مهم در ارتباط با داده‌های آماری، میزان تغییرات و پراکندگی آنها است، بدین معنی که اندازه‌گیری تا چه اندازه از فردی به فرد دیگر تغییر می‌کند. در این قسمت شاخص‌های پراکندگی را به عنوان معیاری برای سنجش میزان تغییرات داده‌ها معرفی می‌کنیم.

مثال نمرات ۱۰ دانش‌آموز در دو امتحان تستی درسهای ریاضی و زبان به صورت زیر است:

درس ریاضی : ۰,۴,۴,۱۲,۱۲,۱۲,۱۴,۱۶,۱۶,۲۰

درس زبان : ۸,۸,۹,۹,۱۲,۱۲,۱۲,۱۳,۱۳,۱۴

همان طور که دیده می‌شود در هر دو کلاس میانگین نمره $\bar{x}=۱۱$، میانه $m=۱۲$ و نما $M=۱۲$ می‌باشد، اما نحوهٔ تغییر پذیری نمرات در دو درس نسبت به میانگین ۱۱ یا میانه و نمای ۱۲ متفاوت می‌باشد. در درس ریاضی پراکندگی نمرات زیاد ولی در درس زبان پراکندگی نمرات کمتر است.

با توجه به مثال بالا برای اینکه بتوانیم این دو درس را با یکدیگر مقایسه کنیم بایستی از شاخص‌های دیگری بنام شاخص‌های پراکندگی استفاده کنیم

دقت، بیان‌گر میزان نزدیکی یا انطباق اندازه‌گیری‌های تکراری از یک‌کمیت واحد نسبت به‌هم‌دیگر است. اگر اندازه‌گیری‌های تکراری با تفاوت‌های بسیارجزیی در یک‌دسته قرارگیرند، از دقت بالا برخورداراند و اگر با تفاوت‌های زیاد از یک‌دیگر دورباشند، دقت پایین دارند. عموماً دقت بالا نتیجهٔ توجه و مراقبت زیاد عامل اندازه‌گیری، انتخاب ابزار دقیق اندازه‌گیری، و نیز به‌کارگیری روش‌های صحیح اندازه‌گیری است. پارامتر دقت با پراکندگی توزیع احتمال، بیان‌کردنی است. هرقدر شکل هندسی توزیع احتمال باریک‌تر باشد، بیان‌گر دقت بالاتر است. میانگین تفاوت اندازه‌گیری‌ها (x_i) با میانگین‌شان ($\bar{x}$) را می‌توان معیاری مناسب برای نمایش دقت دانست، اما ازآن‌جاکه مجموع این‌تفاوت‌ها برابر صفر است، متوسط آن‌ها نیز همواره صفر خواهدشد و بنابراین قابل‌استفاده نیست. برای حل این‌مشکل، از متوسط مجذور تفاوت اندازه‌گیری‌ها با میانگین‌شان استفاده‌می‌کنیم. این‌کمیت، همان واریانس نمونه است. برای نمونه‌ای nتایی از اندازه‌گیری‌ها، به‌عنوان برآوردی از واریانس واقعی σ_x^2، از رابطهٔ زیر به‌دست‌می‌آید:

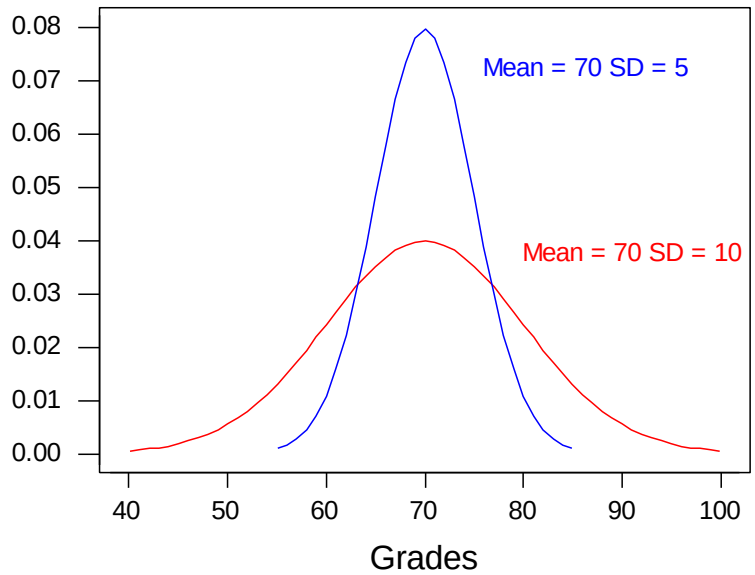
$$S_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

وریانس ، انحراف از معیار و یا دقت تعیین مشاهدات

ازآن‌جاکه واحد سنجش دقت باید از جنس کمیت موردراندازه‌گیری باشد، لذا معیار متداول برای نمایش دقت، جذر واریانس یعنی همان انحراف‌معیار[1] (σ) انتخاب‌می‌شود که به‌آن «خطای متوسط هندسی»[2] نیز اطلاق‌می‌گردد. واضح است برای دقت‌های بالاتر مقدار σ کوچک‌تر و برای دقت‌های پایین‌تر مقدار σ بزرگ‌تر است. چنان‌چه بخواهیم موضوع دقت را به‌صورت کمیتی نسبی موردمطالعه قراردهیم از نسبت دقت (σ) به‌مقدار اندازه‌گیری (l) به‌عنوان دقت نسبی ($e = \sigma / l$) استفاده‌می‌کنیم.

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

Bell-shaped curve



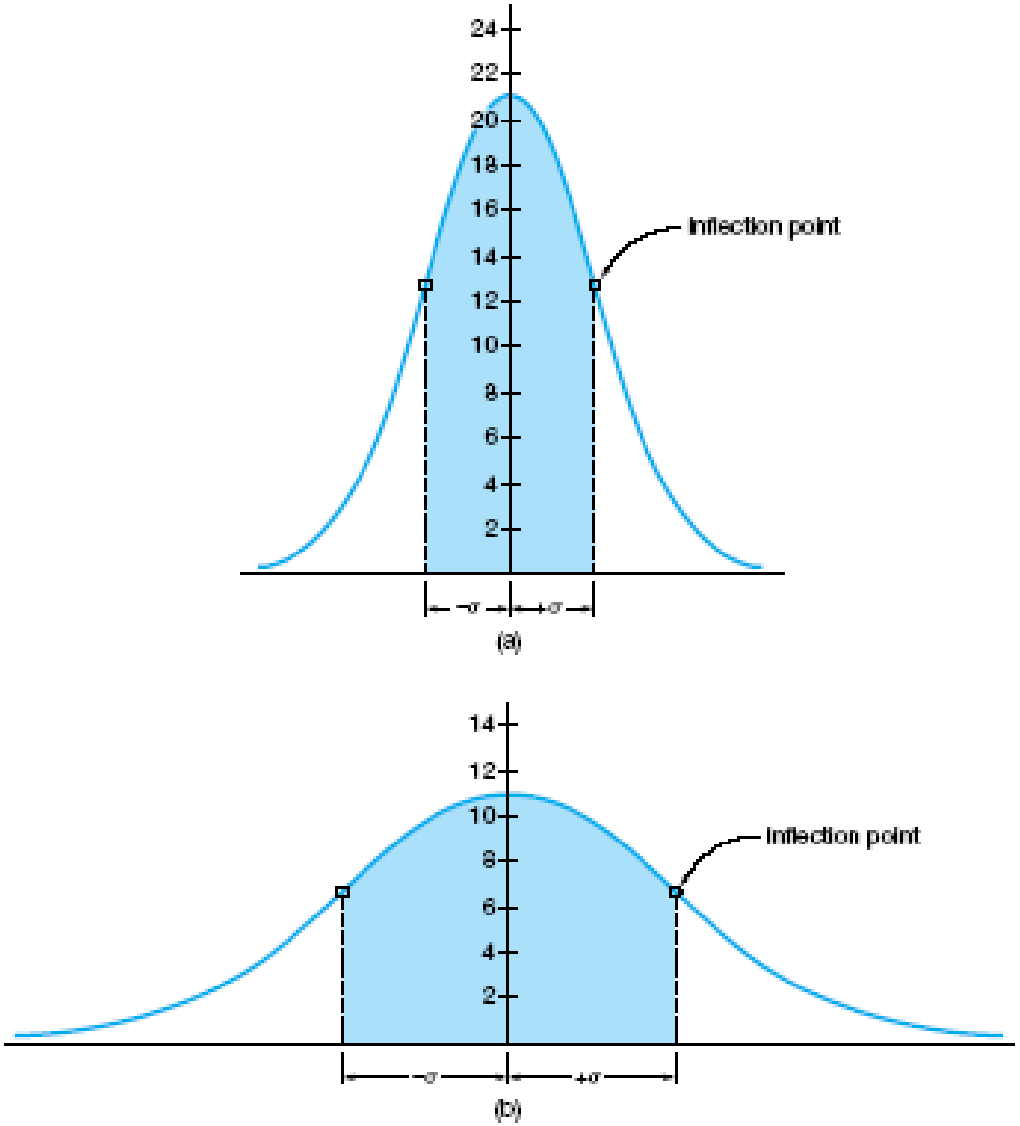


Figure 3–4 Normal distribution curves for: (a) increased precision, (b) decreased precision.

خطای معیار. خطای معیار درواقع به‌انحراف‌معیار میانگین نمونه‌ها اطلاق‌می‌گردد. برای یک‌نمونهٔ n تایی از اندازه‌گیری‌ها، خطای معیار از رابطهٔ زیر به‌دست‌می‌آید:

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

خطای متوسط حسابی. ازآن‌جاکه مجموع جبری باقی‌مانده‌های اندازه‌گیری‌ها نسبت به‌میانگین، صفر است ($\sum_{i=1}^{n}(x_i - \bar{x}) = 0$)، یکی از راه‌های بیان پراکندگی مشاهدات از یک‌دیگر به‌کارگیری قدرمطلق باقی‌مانده‌ها و محاسبهٔ میانگین آن‌ها، موسوم به‌خطای متوسط حسابی، به‌صورت زیر است:

$$E_a = \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|$$

خطای محتمل. این‌خطا، شاخصی دیگر از دقت اندازه‌گیری‌ها است که در گذشته بیشتر مورداستفاده بوده‌است. خطای محتمل بیان‌گر عددی است که قدرمطلق خطای ۵۰ درصد اندازه‌گیری‌ها از آن‌بزرگ‌تر است. این‌کمیت بر اساس مقدار انحراف‌معیار به‌صورت زیر تعریف‌می‌شود:

$$E_{50} = 0.6745\sigma$$

خطای ۹۰٪. با تعریفی مشابه، تعریف خطای محتمل، خطای ۹۰٪ نیز شاخصی دیگری از دقت اندازه‌گیری‌ها است که بیان‌گر حدودی است که ۹۰درصد خطاها در آن قرار می‌گیرند. این‌کمیت بر اساس مقدار انحراف‌معیار به‌صورت زیر تعریف‌می‌شود.

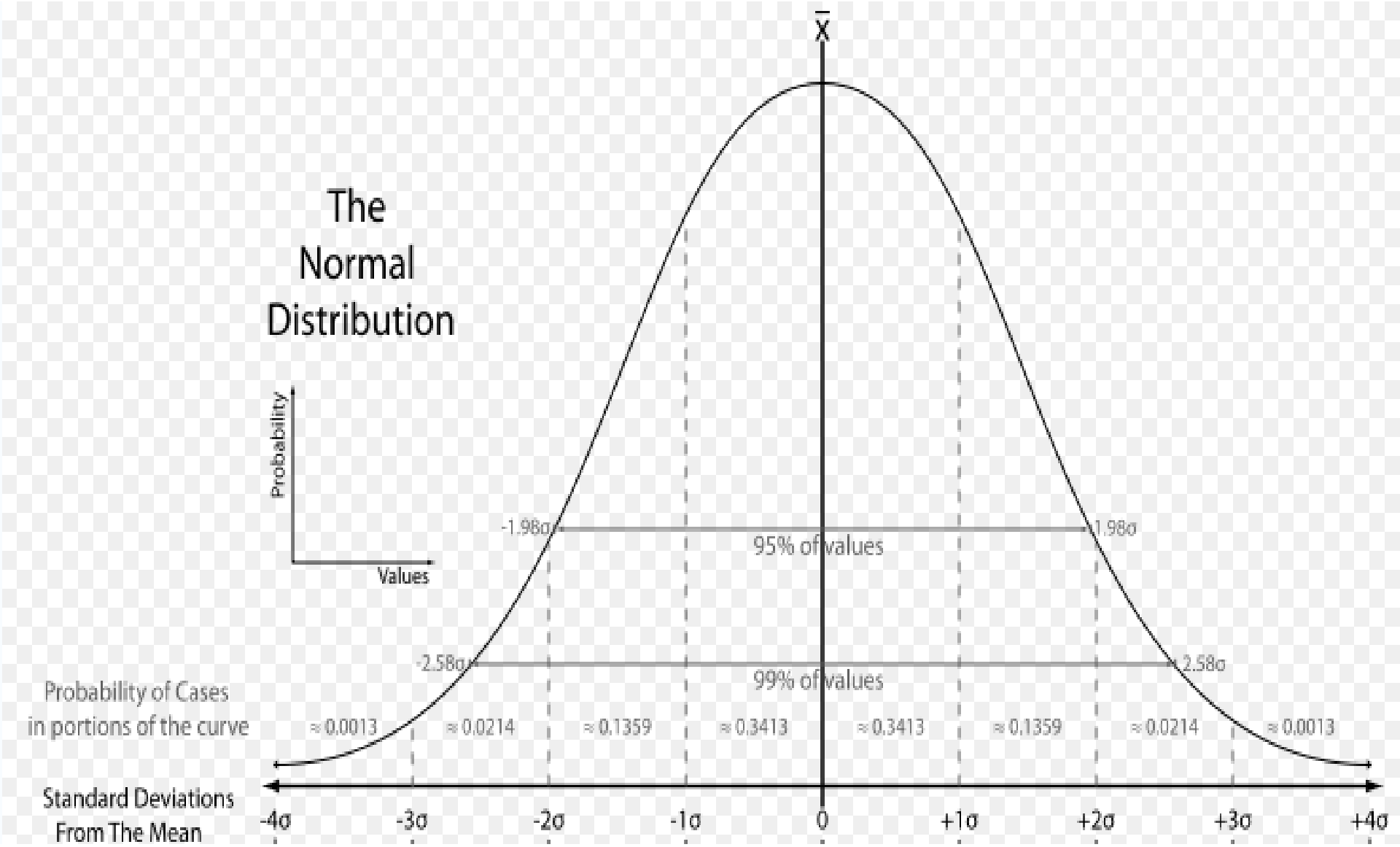
$$E_{90} = 1.6449\sigma$$

خطای حداکثر. از نظر آماری این‌کمیت بیان‌گر عددی است که تنها یک‌درصد از اندازه‌گیری‌ها قدرمطلق خطایشان از آن تجاوزنکند و چنان‌چه از آن بیشتر باشد معمولا آن‌اندازه‌گیری را به‌عنوان اندازه‌گیری اشتباه تلقی‌می‌کنند. مقدار این‌کمیت بر اساس مقدار انحراف‌معیار به‌صورت زیر تعریف‌می‌شود:

$$E_M = 2.675\sigma$$

در عمل برای سادگی در بسیاری از نیازهای روزمره نقشه‌برداری به‌جای ضریب ۲/۶۷۵ از عدد ۲/۵ استفاده‌می‌کنند.

برخی از معیار های بیان دقت مشاهدات



مثال۱: طولی 6 بار اندازه گیری شده و اعداد زیر بدست آمده است.

$$x_1 = 112.03m\,,\, x_2 = 112.00m\,,\, x_3 = 111.98m$$
$$x_4 = 112.03m\,,\, x_5 = 112.04m\,,\, x_6 = 111.98m$$

مطلوب است محاسبه:

ب) خطاهای ظاهری مشاهدات

$$v_i = a_1 - \overline{a}$$
$$v_1 = 112.03 - 112.01 = 0.02m$$
$$v_2 = 112.00 - 112.01 = -0.01m$$
$$v_3 = 111.98 - 112.01 = -0.03m$$
$$v_4 = 112.03 - 112.01 = 0.02m$$
$$v_5 = 112.04 - 112.01 = 0.03m$$
$$v_6 = 111.98 - 112.01 = -0.03m$$

پ) خطای متوسط هندسی

$$e_q = \pm\sqrt{\frac{0.0004+0.0001+0.0009+0.0004+0.0009+0.0009}{5}} = 0.0268$$

ت) خطای متوسط حسابی

$$e_a = \frac{|0.02|+|-0.01|+|-0.03|+|0.02|+|0.03|+|-0.03|}{6} = 0.023$$

ث) خطای ماکزیمم

$$e_{\max} = 2.5\times 0.0268 = 0.067m$$

چ) مشخص کردن داده های اشتباه در صورت وجود

برای این کار خطاهای ظاهری را با خطای max مقایسه می کنیم اگر هر کدام از این خطاها از خطای max بیشتر باشد، آن اندازه گیری، اندازه گیری اشتباه خواهد بود. با توجه به اینکه $v_i < e_{\max}$ هستند، لذا تمام مشاهدات قابل قبول می باشند. لازم به یادآوری می باشد که اگر مشاهده اشتباه داشتیم آن مشاهده اشتباه را از بین مشاهدات انجام شده، حذف کرده و دوباره باید از ابتدا میانگین و انحراف معیار مشاهدات باقیمانده را محاسبه کرد.

وریانس ، انحراف از معیار و یا دقت تعیین مشاهدات (مثال۲)

ـ طولی را ۴ بار اندازه گیری نموده و ارقام ۳۰|۲۲، ۳۰|۲۵، ۳۰|۱۵ و ۳۰|۱۸ متر قرائت گردید متوسط حسابی (میانگین) و خطاهای ظاهری را حساب کنید.

$$\bar{X}=\frac{\sum X_i}{n} \qquad \bar{X}=\frac{30.22+30.25+30.15+30.18}{4}=30.20\,m$$

$$\nu_i=\bar{X}-X_i$$

$$\nu_1=30.20-30.22=-0.02\,m$$

$$\nu_2=30.20-30.25=-0.05\,m$$

$$\nu_3=30.20-30.15=0.05\,m$$

$$\nu_4=30.20-30.18=0.02\,m$$

$$\sum_{i=1}^{n}\nu_i=-0.02-0.05+0.05+0.02=0$$

به علت اینکه میانگین اندازه‌ها به عنوان اندازه کمیت موردنظر قرارگرفت مجموع خطاهای ظاهری صفر شد.

ـ در تمرین بالا خطای متوسط هندسی و خطای ماکزیمم را محاسبه نمایید.

$$e_q=\pm\sqrt{\frac{\sum \nu_i^2}{n-1}} \qquad \sum \nu_i^2=(-0.02)^2+(-0.05)^2+(0.05)^2+(0.02)^2$$

$$\sum \nu_i^2=0.0058\ m^2$$

$$e_q=\pm\sqrt{\frac{0.0058}{4-1}}=\pm 0.043\ m$$

$$e_m=\pm 2.5\ e_q=\pm 2.5(0.043)=0.11\,m$$

با توجه به‌تعاریفی که برای شاخص‌های دقت ارائه‌شد، می‌توان رابطهٔ کلی زیر را برای شاخص دقت با هردرصد اطمینان موردنظر نوشت:

$$e_{1-\alpha} = k\sigma$$

k ضریبی است که با استفاده‌از تابع چگالی احتمال برای هردرصد اطمینان دل‌خواه قابل‌محاسبه‌است. با فرض تابع چگالی احتمال نرمال، جدول ۳.۲ برخی مقادیر ضریب k را نمایش‌می‌دهد.

جدول ۳.۲. سطوح اطمینان مختلف برای تابع چگالی احتمال نرمال در حالت یک‌بعدی

k	$(1-\alpha)\times$۱۰۰
۰/۶۷۴	۵۰/۰۰
۱/۰۰۰	۶۸/۳۰
۱/۶۴۵	۹۰/۰۰
۱/۹۶۰	۹۵/۰۰
۲/۰۰۰	۹۵/۴۰
۲/۵۷۶	۹۹/۰۰
۳/۰۰۰	۹۹/۷۰

فاصلهٔ اطمینان[۲] دامنه یا بازه‌ای از اندازه‌گیری‌هاست که انتظارمی‌رود اندازهٔ واقعی در آن‌واقع‌شود. معمولاً برای بیان هرفاصلهٔ اطمینانی از یک‌سطح اطمینان خاص استفاده‌می‌شود. برای مثال، فاصلهٔ اطمینان ۹۰٪، بازه‌ای است که احتمال وقوع اندازهٔ واقعی در آن ۹۰درصد است. با توجه به‌سطح اطمینان موردنظر رابطهٔ زیر را می‌توان برای ارائهٔ یک‌فاصلهٔ اطمینان نوشت:

$$\left[\bar{x} - k\sigma_{\bar{x}} \quad \bar{x} + k\sigma_{\bar{x}}\right]$$

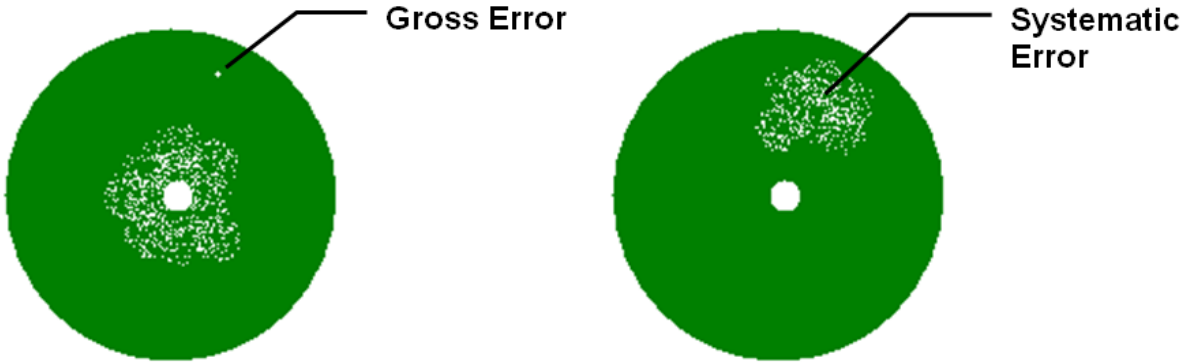
که در آن $\bar{x}$ میانگین حسابی، $\sigma_{\bar{x}}$ انحراف‌معیار میانگین حسابی، و k ضریب تعیین درصد اطمینان است که با استفاده از تابع چگالی احتمال به‌دست‌می‌آید. جدول ۳.۲ برخی از مقادیر مهم و پُرکاربرد ضریب k را نمایش‌می‌دهد.

جدول ۳.۲. سطوح اطمینان مختلف برای تابع چگالی احتمال نرمال در حالت یک‌بعدی

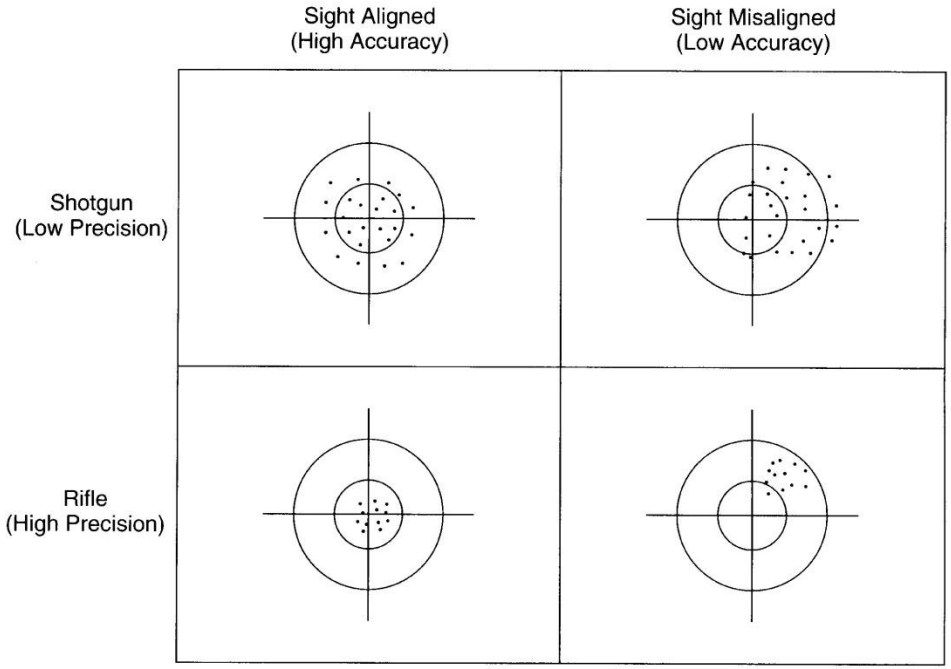
k	$(1-\alpha)\times$۱۰۰
۰/۶۷۴	۵۰/۰۰
۱/۰۰۰	۶۸/۳۰
۱/۶۴۵	۹۰/۰۰
۱/۹۶۰	۹۵/۰۰
۲/۰۰۰	۹۵/۴۰
۲/۵۷۶	۹۹/۰۰
۳/۰۰۰	۹۹/۷۰

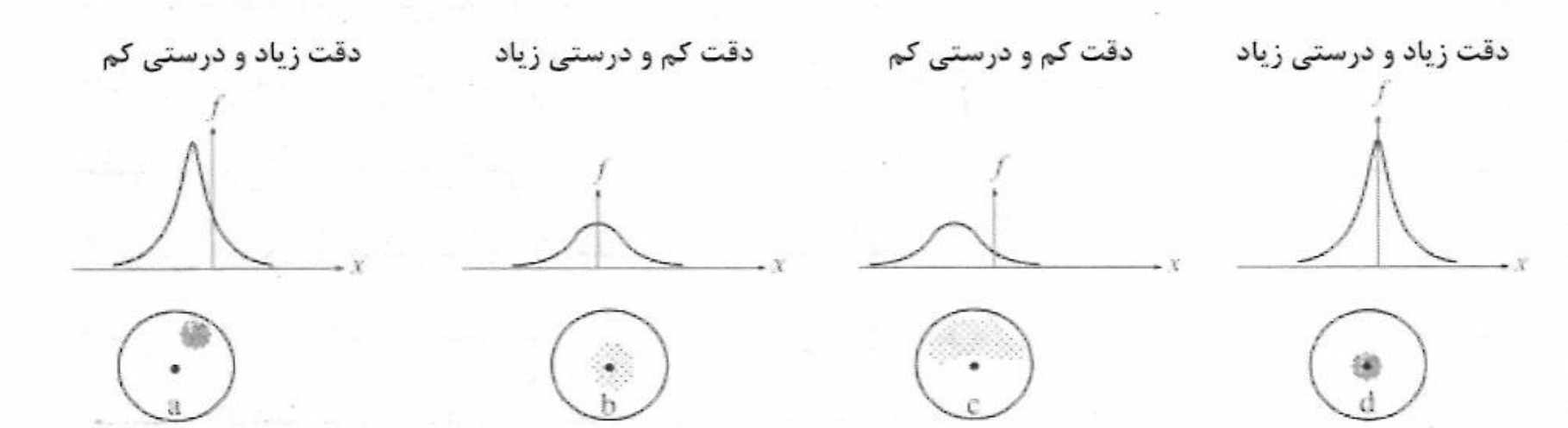
درستی یا صحت

درستی، بیان‌گر میزان نزدیکی یا انطباق اندازه‌گیری‌ها با مقدار واقعی کمیت مورداندازه‌گیری است. پارامترِ درستی به‌نوعی بیان‌گر وجود یا نبود خطای سیستماتیک در اندازه‌گیری‌ها است. بنابراین چنان‌چه هیچ اثر خطای سیستماتیک وجودنداشته‌باشد، میانگین اندازه‌گیری‌ها را می‌توان به‌عنوان برآوردی از مقدار واقعی به‌کاربرد. معمولا شاخص درستی اندازه‌گیری‌ها، انحراف آن‌ها از میانگین اندازه‌گیری‌هایی است که از ابزار دقیق‌تر و مطمئن‌تر به‌دست‌آمده‌باشند. بدیهی است این‌گونه شاخص‌ها نسبی‌اند و بسته به‌هدف موردنظر، می‌توان یکی را انتخاب نمود. برای مثال، طول حاصل از طول‌یاب‌های الکترونیک خیلی‌دقیق یا گیرنده‌های سامانهٔ تعیین‌موقعیت جهانی[1] (GPS) را می‌توان شاخصی مناسب برای سنجش درستی سایر اندازه‌گیری‌ها دانست.



مقایسه مفهوم دقت و صحت





نگارهٔ ۲.۲. تعبیر هندسی دقت و درستی در رفتار تابع چگالی احتمال و یک‌هدف تیراندازی

تمرین

برای افزایش دقت در اندازه گیری پارالاکس یک نقطه بر روی یک زوج عکس هوایی، فلاتین مارک ها را ۱۰ بار در نقطه مورد نظر بر روی زمین مماس کرده و و اعداد زیر مشاهده گردیده اند:

Rm1=4.890 mm	Rm2=4.920 mm	Rm3=4.900 mm	Rm4=4.870 mm	Rm5=4.930 mm
Rm6=4.960 mm	Rm7=4.990 mm	Rm8=4.980 mm	Rm9=4.970 mm	Rm10=4.940 mm

پارامتر های زیر را محاسبه کنید.
میانگین، خطاهای ظاهری، وریانس، خطای متوسط هندسی، خطای متوسط حسابی، خطای معیار، خطای محتمل، خطای ۹۰٪ ، خطای حداکثر یا ماکزیمم، بازه اطمینان ۹۰٪ برای مشاهدات.

پایان جلسه دوم

درس تئوری خطاها

جلسه سوم

فرید اسماعیلی
Farid_63@yahoo.com
www.faridesm.ir

تماس با استاد از طریق پست الکترونیکی
مشاهده اطلاعیه ها، نمرات، دریافت فایل ها در وب سایت

مجموعهٔ تمام نتایج متمایز ممکن در یک‌آزمایش تصادفی، فضای نمونه[1] نامیده‌می‌شود و هر نتیجهٔ متمایز یک پیش‌آمد[2] یا رُخ‌داد ساده[3] یا یک‌عنصر فضای نمونه خوانده‌می‌شود (آریانژاد و ذهبیون، ۱۳۷۱). در این‌جا فضای نمونه را با Ω نمایش‌می‌دهیم.

$$\Omega = \{u_1, u_2, u_3, \ldots, u_n\}$$

برای مثال، می‌توان نتایج حاصل از ریختن یک‌تاس را به‌صورت زیر نمایش‌داد:

$$\Omega = \{1,2,3,4,5,6\}$$

فضاهای نمونه به‌دودستهٔ کلی زیر تقسیم‌می‌شوند:

- فضای نمونهٔ گسسته[1]. اگر فضای نمونه‌ای دارای تعداد عناصر متناهی یا به‌طور شمارش‌پذیر متناهی باشد، آن‌را *فضای نمونهٔ گسسته* گوییم.

- فضای نمونهٔ پیوسته[2]. اگر فضای نمونه شامل تمام اعداد متعلق به‌یک‌فاصله باشد آن‌را *فضای نمونهٔ پیوسته* گوییم.

در یک‌فضای نمونه، به‌هرزیرمجموعه از آن،که بیان‌گر یک‌یا چندپیش‌آمد است، *رُخ‌داد*[3] با مفهوم تصادفی، می‌گویند. پیش‌آمدی که به‌طورقطع رخ‌می‌دهد (پیش‌آمد برابر با فضای نمونه) را *رُخ‌داد حتمی*[4] و پیش‌آمدی که هیچ گاه رخ‌نمی‌دهد (پیش‌آمد بدون عضو) را *رُخ‌دادِ ناممکن*[5] می‌نامند. برای مثال، در انداختن یک‌تاس، پیش‌آمدِ آمدنِ عدد ۷ پیش‌آمدی ناممکن و آمدن عدد ۱ تا ۶ پیش‌آمدی حتمی است.

احتمال وقوع یک‌پیش‌آمد به‌معنای شانس وقوع آن‌پیش‌آمد در انجام یک‌آزمایش تصادفی است. برای محاسبهٔ احتمال عبارات مختلفی ازجمله فراوانی نسبی و هم‌شانسی وجوددارد که هرکدام در برآورد احتمال می‌تواند مفید باشد. بنابرتعریف، تابع احتمال، تابعی از فضای نمونه به‌داخل مجموعهٔ اعداد حقیقی بین ۰ تا ۱ است. به‌عبارت دقیق‌تر احتمال، تابعی است که به‌هرپیش‌آمد از فضای نمونه، عدد حقیقی P(A) را به‌گونه‌ای نسبت‌می‌دهد که قوانین و شرایط زیر صدق‌کند: با فرض داشتن یک‌فضای نمونهٔ n تایی احتمـال وقوع یک‌پیش‌آمد عبارت‌است از نسبت پیش‌آمدهای دل‌خواه (kپیش‌آمد) به‌تمام پیش‌آمدهای ممکن (n پیش‌آمد) که رابطهٔ آن‌به‌صورت زیر نمایش‌داده‌می‌شود:

$$P(A) = \frac{\sum_{i=1}^{k}(u_i)}{\sum_{i=1}^{n}(u_i)}$$

مثال سکه‌ای را دو بار پرتاب می‌کنیم احتمال آوردن حداقل یک شیر را بیابید.

حل فضای نمونه حاصل از این آزمایش برابر است با $S = \{TT, TH, HT, HH\}$

اگر سکه سالم باشد آنگاه شانس رخداد شیر یا خط با هم برابر است و در نتیجه به هر کدام از نقاط فضای نمونه شانس یکسان w را نسبت می‌دهیم و بنابراین ۴w=۱ و یا $w=\frac{۱}{۴}$. پس مدل احتمال برای این آزمایش عبارت است از

S	TT	TH	HT	HH
احتمال	$\frac{۱}{۴}$	$\frac{۱}{۴}$	$\frac{۱}{۴}$	$\frac{۱}{۴}$

حال اگر A پیشامد مشاهده حداقل یک شیر باشد آنگاه $A=\{TH, HT, HH\}$ و در نتیجه

$$P(A) = \frac{۱}{۴} + \frac{۱}{۴} + \frac{۱}{۴} = \frac{۳}{۴}$$

مثال دو تاس را با هم پرتاب می‌کنیم و متغیر تصادفی X را برابر مجموع اعداد روی دو تاس مشاهده شده در نظر می‌گیریم.

الف- تابع احتمال X را به دست آورید.

ب- احتمال اینکه مجموع اعداد روی دو تاس حداکثر ۴ شود را بیابید.

ج- احتمال اینکه مجموع اعداد روی دو تاس بین ۶ و ۹ شود را بیابید.

حل الف- تکیه گاه X برابر $S_X = \{۲, ۳, ۴, ..., ۱۲\}$ است و برای مثال

$$P(X=۴) = P(\{(۱,۳),(۲,۲),(۳,۱)\}) = \frac{۳}{۳۶}$$

با محاسبه احتمالات مربوط به نقاط دیگر S_X، تابع احتمال X به صورت زیر به دست می‌آید.

x	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
$f_X(x) = P(X=x)$	$\frac{۱}{۳۶}$	$\frac{۲}{۳۶}$	$\frac{۳}{۳۶}$	$\frac{۴}{۳۶}$	$\frac{۵}{۳۶}$	$\frac{۶}{۳۶}$	$\frac{۵}{۳۶}$	$\frac{۴}{۳۶}$	$\frac{۳}{۳۶}$	$\frac{۲}{۳۶}$	$\frac{۱}{۳۶}$

جدول فوق یک تابع از S_X به اعداد حقیقی در فاصله [۰,۱] برقرار می‌کند که آن را با تابع $f_X(x)$ نمایش داده و به آن تابع احتمال متغیر تصادفی X گویند، بنابراین

$$f_X(x) = P(X=x)$$

این تابع یک تابع غیر منفی است و مجموع آن روی کلیه مقادیری که می‌تواند اختیار کند برابر ۱ است.

x	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
$f_X(x)=P(X=x)$	$\frac{۱}{۳۶}$	$\frac{۲}{۳۶}$	$\frac{۳}{۳۶}$	$\frac{۴}{۳۶}$	$\frac{۵}{۳۶}$	$\frac{۶}{۳۶}$	$\frac{۵}{۳۶}$	$\frac{۴}{۳۶}$	$\frac{۳}{۳۶}$	$\frac{۲}{۳۶}$	$\frac{۱}{۳۶}$

ب – $P(X\leq ۴)$ مورد سئوال است بنابراین از فرمول (۱.۳) داریم که

$$P(X\leq ۴)=f_X(۲)+f_X(۳)+f_X(۴)=\frac{۱}{۳۶}+\frac{۲}{۳۶}+\frac{۳}{۳۶}=\frac{۱}{۶}$$

ج – با استفاده از فرمول (۱.۳) داریم که

$$P(۶<X<۹)=f_X(۷)+f_X(۸)=\frac{۶}{۳۶}+\frac{۵}{۳۶}=\frac{۱۱}{۳۶}$$

تعریف تابع $f_X(x)=P(X=x)$ را تابع احتمال متغیر تصادفی X گویند هر گاه

۱ـ برای هر $x\in R$ داشته باشیم که $f_X(x)\geq ۰$

۲ـ $\sum_{x\in R} f_X(x)=۱$

تابع توزیع احتمال پیوسته

متغیر تصادفی که مجموعه مقادیر آن یک فاصله عددی یا اجتماع چند فاصله عددی باشد را متغیر تصادفی پیوسته گویند. در مورد متغیر تصادفی پیوسته بایستی توجه کرد که احتمال اینکه یک متغیر تصادفی پیوسته بخواهد فقط یک مقدار بخصوص از مجموعه مقادیرش را بگیرد برابر صفر است. به مثال زیر توجه کنید.

بنابراین توزیع احتمال یک متغیر تصادفی پیوسته را نمی‌توان به صورت یک جدول نمایش داد. در این حالت توزیع احتمال متغیر تصادفی پیوسته X را به صورت یک تابع $f_X(x)$ نمایش داده و آن را **تابع چگالی احتمال** X می‌نامند.

برای به دست آوردن تابع چگالی احتمال یک متغیر تصادفی پیوسته معمولاً ابتدا تابع توزیع آن را به دست می‌آورند، زیرا تابع توزیع احتمال را در فواصل محاسبه می‌کند و محاسبه این احتمالات در حالت پیوسته امکان‌پذیر است. تابع توزیع یک متغیر تصادفی پیوسته X به صورت $F_X(x)=P(X\leq x)$ تعریف می‌شود. با توجه به آنکه در حالت گسسته تابع توزیع به صورت $F_X(x)=P(X\leq x) = \sum_{t\leq x} f_X(t)$ محاسبه می‌شد، با قرار دادن تابع چگالی احتمال به جای تابع احتمال و تبدیل مجموع به انتگرال می‌توان تابع توزیع در حالت پیوسته را به طور مشابه و به صورت زیر محاسبه کرد:

$$\boxed{F_X(x)=P(X\leq x) = \int_{-\infty}^{x} f_X(t)dt} \qquad (۴.۳)$$

تابع $f_X(x)$ که در رابطه (۴.۳) صدق می‌کند را تابع چگالی احتمال و تابع $F_X(x)$ در این رابطه را تابع توزیع متغیر تصادفی پیوسته X گویند. همچنین با توجه به قضیه اساسی حساب از رابطه (۴.۳) نتیجه می‌شود که

$$\boxed{F_X'(x) = \frac{d}{dx}F_X(x) = f_X(x)} \qquad (۵.۳)$$

با توجه به روابط (۴.۳) و (۵.۳) با داشتن تابع توزیع X و مشتق گرفتن از آن به راحتی می‌توان تابع چگالی احتمال X را به دست آورد و برعکس با داشتن تابع چگالی احتمال X و انتگرال گرفتن از آن می‌توان تابع توزیع X را به دست آورد.

مثال نقطه‌ای را به تصادف از فاصله اعداد حقیقی [۰,۲] انتخاب می‌کنیم و متغیر تصادفی X را برابر نقطه انتخاب شده در فاصله [۰,۲] در نظر می‌گیریم.

تابع توزیع متغیر تصادفی پیوسته X را به دست آورده و از روی آن تابع چگالی احتمال X را محاسبه کنید و سپس این دو تابع را رسم کنید.

حل با توجه به نمودار روبرو و مفهوم تابع توزیع **تجمعی**

اگر $x < ۰$ آنگاه $F_X(x) = P(X \le x) = ۰$ و اگر $۰ \le x < ۲$ آنگاه

$$F_X(x) = P(X \le x) = \frac{\text{طول فاصله } [۰, x]}{\text{طول فاصله } [۰,۲]} = \frac{x}{۲}$$

و اگر $x \ge ۲$ آنگاه $F_X(x) = P(X \le x) = ۱$. بنابراین

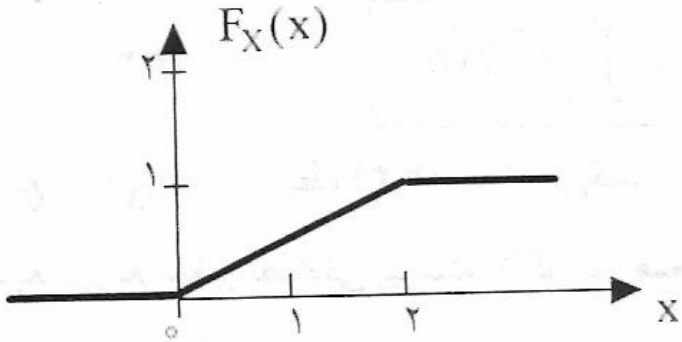
و اگر $x \geq ۲$ آنگاه $F_X(x) = P(X \leq x) = ۱$. بنابراین

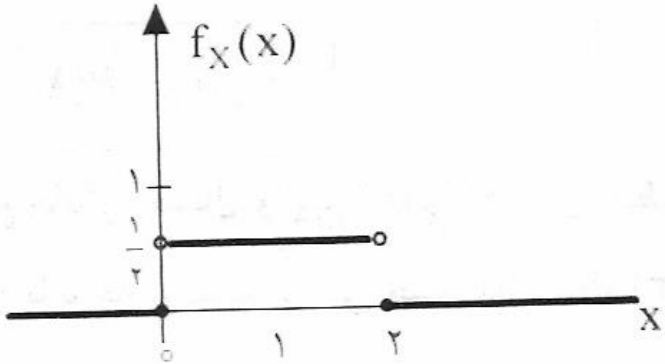
$$F_X(x) = P(X \leq x) = \begin{cases} ۰ & x < ۰ \\ \frac{x}{۲} & ۰ \leq x < ۲ \\ ۱ & ۲ \leq x \end{cases}$$

و در نتیجه

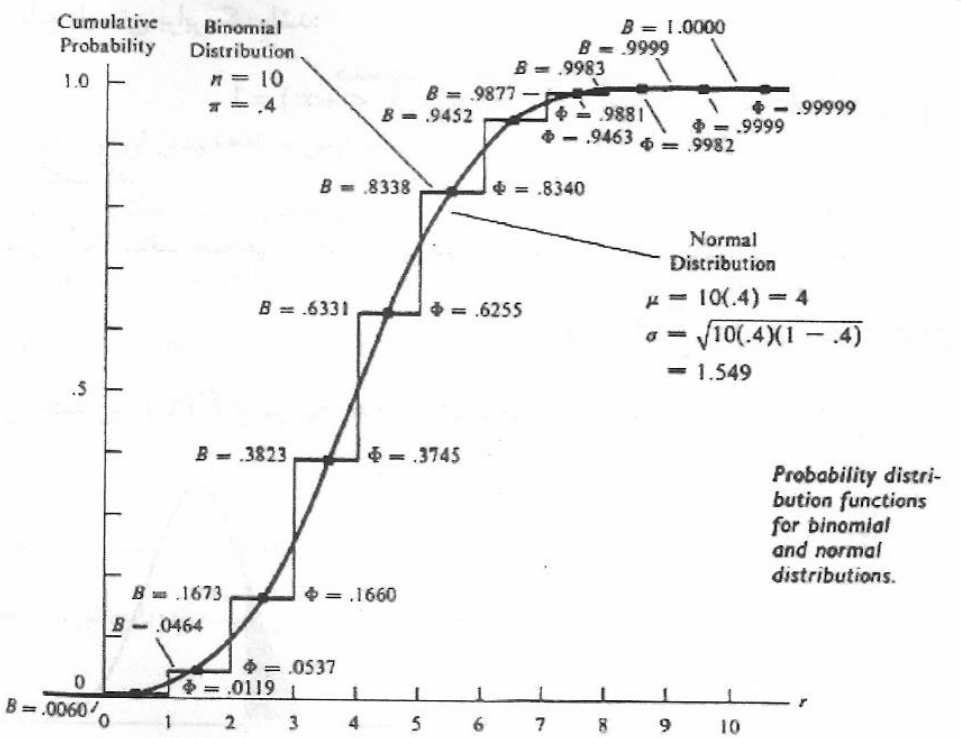
$$f_X(x) = F_X'(x) = \begin{cases} \frac{۱}{۲} & ۰ < x < ۲ \\ ۰ & \text{سایر نقاط} \end{cases}$$

نمودار این دو تابع در زیر رسم شده است. همان طور که دیده می‌شود نمودار تابع توزیع یک متغیر تصادفی پیوسته یک نمودار پیوسته است.



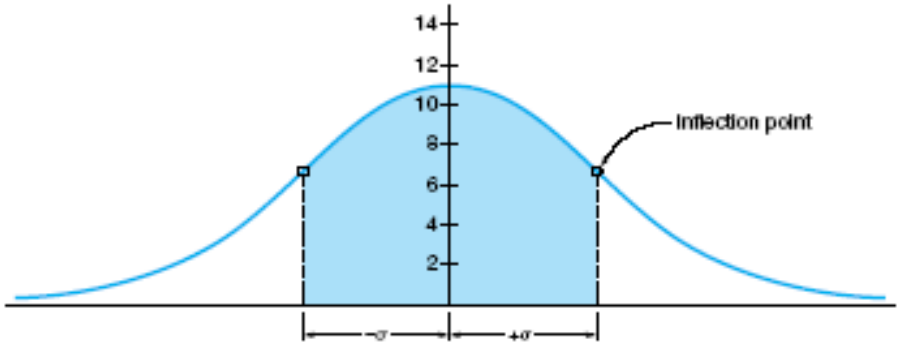


تابع توزیع احتمال پیوسته



Probability distribution functions for binomial and normal distributions.

شکل ۲-۲: تابع توزیع تجمعی برای توزیع نرمال



تابع توزیع احتمال تجمعی پیوسته ⟵ مشتق اول تابع توزیع احتمال تجمعی پیوسته و یا تابع چگالی احتمال پیوسته

خواص تابع چگالی احتمال و تابع توزیع با توجه به روابط (۴.۳) و (۵.۳) بین تابع چگالی احتمال و تابع توزیع می‌توان خواص زیر را برای تابع چگالی احتمال $f_X(x)$ بیان کرد.

الف- برای هر $x \in R$، $f_X(x) \geq 0$

ب- $\int_{-\infty}^{+\infty} f_X(x)\,dx = 1$

برای محاسبه احتمالات مربوط به یک متغیر تصادفی پیوسته می‌توان از رابطه زیر استفاده کرد

$$\boxed{P(a < X \leq b) = \int_a^b f_X(x)\,dx = F_X(b) - F_X(a)} \qquad (۶.۳)$$

که در آن a می‌تواند $-\infty$ و b می‌تواند $+\infty$ باشد.

مثال فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد

$$f_X(x)=\begin{cases}\dfrac{c}{x^{۲}} & ۱<x<۱۰\\ ۰ & \text{سایر نقاط}\end{cases}$$

الف – مقدار c را تعیین کنید.

حل الف – با توجه به خواص تابع چگالی احتمال بایستی $c\geq ۰$ و همچنین

$$۱=\int_{-\infty}^{+\infty} f_X(x)dx=\int_{-\infty}^{۱} f_X(x)\,dx+\int_{۱}^{۱۰} f_X(x)\,dx+\int_{۱۰}^{+\infty} f_X(x)\,dx$$

$$=۰+\int_{۱}^{۱۰}\frac{c}{x^{۲}}dx+۰=\frac{-c}{x}\Big|_{۱}^{۱۰}=-\frac{c}{۱۰}+c=\frac{۹}{۱۰}c$$

پس بایستی $c=\frac{۱۰}{۹}$ باشد

مثال فرض کنید جریمهٔ توقف اتومبیل در محل پارک ممنوع یک هزار تومان باشد و در یک شهر ۴۰ درصد افراد هیچگاه، ۳۰ درصد افراد یک‌بار، ۲۰ درصد افراد دوبار و ۱۰ درصد افراد سه‌بار در ماه به جهت پارک در محل پارک ممنوع جریمه شوند. به طور متوسط انتظار دارید که هر نفر در این شهر چه مبلغی را در ماه برای جریمه توقف اتومبیل در محل پارک ممنوع پرداخت نماید.

حل فرض کنید ۵۰۰ نفر از افراد این شهر را در نظر بگیریم. بر اساس درصدهای داده شده انتظار داریم که ۲۰۰ نفر آنها هیچ مبلغی، ۱۵۰ نفر آنها یک هزار تومان، ۱۰۰ نفر آنها دو هزار تومان و ۵۰ نفر آنها سه هزار تومان جریمه شوند. بنابراین انتظار داریم که ۵۰۰ نفر مبلغ زیر را (بر حسب هزار تومان) جریمه پرداخت کنند

$$(۵۰۰\times۰/۴۰\times۰)+(۵۰۰\times۰/۳۰\times۱)+(۵۰۰\times۰/۲۰\times۲)+(۵۰۰\times۰/۱۰\times۳)=۵۰۰$$

بنابراین بطور متوسط انتظار داریم که هر نفر $۱=\frac{۵۰۰}{۵۰۰}$ هزار تومان در ماه جریمه پرداخت نماید.

با توجه به عملیات مثال قبل دیده می‌شود که متوسط مبلغ جریمه یک هزار تومان ربطی به تعداد ۵۰۰ نفر ندارد، زیرا اگر تعداد را یک میلیون نفر نیز می‌گرفتیم در جواب تغییری حاصل

با توجه به عملیات مثال قبل دیده می‌شود که متوسط مبلغ جریمه یک هزار تومان ربطی به تعداد ۵۰۰ نفر ندارد، زیرا اگر تعداد را یک میلیون نفر نیز می‌گرفتیم در جواب تغییری حاصل نمی‌گردید. در حقیقت این مقدار یک عدد انتظاری (حدی) می‌باشد. در این مثال ما یک متغیر تصادفی X داریم که برابر مبلغ جریمه شخص در یک ماه بر حسب هزار تومان است و تابع احتمال آن به صورت زیر می‌باشد

x	۰	۱	۲	۳
$f_X(x)$	۰/۴۰	۰/۳۰	۰/۲۰	۰/۱۰

عدد زیر را که در حقیقت میانگین وزنی مبلغ جریمه می‌باشد امید ریاضی X یا میانگین X و یا مقدار مورد انتظار X می‌نامند و آنرا با نمادهای $E(X)$ یا μ یا μ_X نمایش می‌دهند.

$$\mu = E(X) = (۰ \times ۰/۴۰) + (۱ \times ۰/۳۰) + (۲ \times ۰/۲۰) + (۳ \times ۰/۱۰) = ۱$$

$$= \sum_{x=۰}^{۳} x f_X(x)$$

اگر متغیر تصادفی X پیوسته باشد امید ریاضی آن با تبدیل مجموع به انتگرال در فرمول بالا محاسبه می‌گردد.

تعریف فرض کنید X یک متغیر تصادفی با تابع احتمال یا چگالی احتمال $f_X(x)$ باشد. امید ریاضی X یا میانگین X به صورت زیر تعریف می‌شود

اگر X گسسته باشد	$E(X)=\sum_x x f_X(x)$
اگر X پیوسته باشد	$E(X)=\int_{-\infty}^{+\infty} x f_X(x)\,dx$

برخی قواعد پراستفاده برای عمل‌گر امیدریاضی به‌صورت زیر معرفی‌می‌شوند:

$$E(X+Y)=E(X)+E(Y)$$

$$E(cX)=cE(X)$$

$$E(XY)=E(X)E(Y)+E\left[(X-\mu_x)(Y-\mu_y)\right]$$

$$E[E[x]]=E[x]$$

در بعضی از مسائل نیاز به محاسبهٔ امید ریاضی تابعی از یک متغیر تصادفی X مانند g(X) داریم. به عنوان مثال g(X) می‌تواند X^2 یا $2X+3$ یا... باشد. برای محاسبهٔ امید ریاضی g(X) از قضیهٔ زیر استفاده می‌کنیم.

قضیهٔ ۱.۴ فرض کنید X یک متغیر تصادفی با تابع احتمال یا تابع چگالی احتمال $f_X(x)$ باشد. امید ریاضی تابع g(X) به صورت زیر به دست می‌آید

$$\boxed{\begin{array}{ll} E[g(X)] = \sum_x g(X) f_X(x) & \text{اگر X گسسته باشد} \\ E[g(X)] = \int_{-\infty}^{+\infty} g(X) f_X(x)\,dx & \text{اگر X پیوسته باشد} \end{array}} \qquad (۲.۴)$$

مثال فرض کنید متغیر تصادفی X دارای تابع احتمال زیر باشد

x	-۱	۰	۱	۲
$f_X(x)$	$۰/۱$	$۰/۱$	$۰/۵$	$۰/۳$

امید ریاضی تابع $g(X)=(X-۱)^۲$ را به دست آورید.

حل X یک متغیر تصادفی گسسته است. بنابراین از رابطهٔ (۲.۴) داریم که

$$E[(X-۱)^۲]=\sum_{x=-۱}^{۲}(x-۱)^۲ f_X(x)$$
$$=(-۱-۱)^۲(۰/۱)+(۰-۱)^۲(۰/۱)+(۱-۱)^۲(۰/۵)+(۲-۱)^۲(۰/۳)=۰/۸$$

به شرطی که میانگین جامعه برابر میانگین نمونه ها فرض شود

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

به شرطی که میانگین جامعه معلوم نباشد

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

$$E[g(X)] = \sum_{i=1}^{n} g(x_i)\,p(x_i)$$

به‌طورخاص اگر $g(X) = (X - \mu_x)^2$، امیدریاضی آن، معروف به‌واریانس است.

$$\mathrm{var}(X) = \sigma_x^2 = E\left[(X - \mu_x)^2\right] = \sum_{i=1}^{n}(X - \mu_x)^2\,p(x_i)$$

حال چنان‌چه X یک‌متغیر تصادفی پیوسته و با تابع چگالی $f(x)$ درنظرگرفته‌شود، عمل‌گر $\sum$ به $\int$ تبدیل‌شده و میانگین نظری و واریانس به‌صورت زیر تعریف‌می‌شوند.

$$\mu_x = E[X] = \int_{-\infty}^{\infty} x f(x)\,dx$$

$$\mathrm{var}(X) = \sigma_x^2 = E\left[(X - \mu_x)^2\right] = \int_{-\infty}^{\infty}(X - \mu_x)^2 f(x)\,dx$$

محاسبه واریانس با استفاده از فرمول زیر ساده‌تر می‌شود:

$$\sigma_x^{\ 2} = E[(x - E[x])^2]$$
$$= E[x^2 + E^2[x] - 2xE[x]]$$
$$= E[x^2] + E[E^2[x]] - E[2xE[x]]$$
$$= E[x^2] + E^2[x] - 2E[x]E[x]$$
$$= E[x^2] - E^2[x]$$

در نتیجه همیشه رابطه زیر برقرار می‌باشد:

$$\boxed{\sigma_x^{\ 2} = E[x^2] - E^2[x]}$$

مثال ۱) اگر بدانیم برای متغیر تصادفی X روابط $E[X(X-4)] = 5, E[X] = 2$ برقرار است، واریانس این متغیر تصادفی چقدر می باشد؟

$$E[X(X-4)] = E[X^2 - 4X] = 5 \Rightarrow E[X^2] - 4E[X] = 5 \Rightarrow E[X^2] = 13$$
$$\mathrm{var}(X) = E[X^2] - E^2[X] = 13 - 4 = 9$$

مثال ۶ -۳) طولی درسه سری وهرسری چندبار بصورت زیرا ندازه‌گیری شده‌است (تعداد اندازه‌گیریها دریک سری یکسان است) مطلوبست محاسبه میانگین ووا ریانس نمونه چندمتغیری

سری اول	سری دوم	سری سوم
102.120	102.122	102.119
102.125	102.126	102.120
102.124	102.124	102.120
102.119	102.124	102.121

حل

$$\begin{cases} M_1 = \sum_{i=1}^{4} \xi_i P(\xi_i) = E(\xi_i) = 1/4\,(102.120 + 102.125 + 102.124 . 102.119) = 102.122^{(m)} \\ P(\xi_1) = P(\xi_2) = P(\xi_3) = P(\xi_4) = 1/4 \end{cases}$$

پس :

$$S_1^2 = \sum (\xi_i - M)^2 P(\xi_i) = \sum_{i=1}^{4} \xi_i^2 P(\xi_i) - E(\xi_i)^2$$

$$S_1^2 = 1/4\left[(102.120 - 102.122)^2 + \cdots + (102.119 - 102.122)^2\right]$$

$$S_1^2 = 10^{-6}/4\,(4 + 9 + 4 + 9) = 6.5 \times 10^{-6}\ (m) = 6.5\ (mm)$$

بهمین ترتیب خواهیم‌داشت :

$$M_2 = 102.124\ (m) \longrightarrow S^2 = 2 \times 10^{-6} = 2\ (mm)$$

$$M_3 = 102.120\ (m) \longrightarrow S^2 = 0.5 \times 10^{-6} = 0.5\ (mm)$$

درنتیجه داریم :

$$M_i = (M_1, M_2, M_3) = (102.122^{(m)}, 102.124^{(m)}, 102.120^{(m)})$$

$$S_i^2 = (S_1^2, S_2^2, S_3^2) = (6.5^{(mm)}, 2^{(mm)}, 0.5^{(mm)})$$

به‌کمک مجموعه‌ای از زوج‌های (x_1,y_1) و (x_2,y_2) و... و (x_n,y_n) از یک متغیر برداری تصادفی مانند $(\bar{x},\bar{y})$ می‌توان علاوه‌بر محاسبهٔ واریانس حاشیه‌ای S_x^2 و S_y^2، کواریانس نمونه را از رابطهٔ زیر محاسبه نمود.

$$S_{xy} = \frac{1}{n-1}\sum(x_i - \bar{x})(y_i - \bar{y})$$

که در آن $\bar{x},\bar{y}$ میانگین‌های نمونه هستند.

ویا از طریق امید ریاضی :

$E\left[(x - E[x])(y - E[y])\right]$ را کواریانس متغیر تصادفی x, y می‌نامند و آنرا با نماد σ_{xy} یا cov(x,y) نشان داده می‌شود.

$$\sigma_{xy} = E[(x-\mu_x)(y-\mu_y)] = E[xy] - E[x]E[y]$$

اگر دو متغییر x,y کاملاً مستقل از هم باشند، در این صورت کواریانس بین آنها صفر خواهد بود.
و اگر دو متغییر کاملاً با هم برابر باشند کواریانس بین آنها برابر واریانس بین آنها خواهد بود.

مثال وریانس کواریانس

مثال : در چند نمونه روبرو میانگین، وریانس، کواریانس را محاسبه کنید.

$$D = \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} (1,2,3) \\ (3,2,1) \end{bmatrix} \qquad n_a = 3 \qquad n_b = 3$$

میانگین

$$M = \begin{bmatrix} m_a \\ m_b \end{bmatrix} = \begin{bmatrix} \frac{1}{n_a}\sum_{i=1}^{n_a}(a_i) \\ \frac{1}{n_b}\sum_{i=1}^{n_b}(b_i) \end{bmatrix} = \begin{bmatrix} \frac{1}{3}(1+2+3) \\ \frac{1}{3}(3+2+1) \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

وریانس

$$S^2 = \begin{bmatrix} s_a^2 \\ s_b^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{n_a}\sum_{i=1}^{n_a}(a_i - m_a)^2 \\ \frac{1}{n_b}\sum_{i=1}^{n_b}(b_i - m_b)^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3}((1-2)^2 + (2-2)^2 + (3-2)^2) \\ \frac{1}{3}((3-2)^2 + (2-2)^2 + (1-2)^2) \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

کواریانس

$$S_{a,b} = S_{b,a} = \frac{1}{n_{a\&b}}\sum_{k=1}^{n_a}(a_i - m_a) \times (b_i - m_b)$$

$$= \frac{1}{3}[(1-2)(3-2) + (2-2)(2-2) + (3-2)(1-2)] = \frac{1}{3}[-1-1] = -\frac{2}{3}$$

ماتریس وریانس کواریانس

$$\Sigma_D = \begin{bmatrix} S_a^2 & S_{ab} \\ S_{ba} & S_B^2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} \end{bmatrix}$$

با استفاده از مقادیر محاسبه شده برای واریانس وکوواریانس در یک سری مشاهدات ، ماتریسی بصورت زیرکه قطر اصلی آن واریانس ها وسایر اجزاء آن بصورت قرینه نسبت به قطر اصلی کوواریانس ها می باشند نوشته میشودکه آنرا ماتریس واریانس کوواریانس (یا فول ماتریس ویا بیس ماتریس) می نامند :

$$\Sigma_x = \begin{bmatrix} S_{11}^2 & S_{12} & S_{13} & \cdots & S_{1n} \\ S_{21} & S_{22}^2 & S_{23} & \cdots & S_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ S_{n1} & S_{n2} & S_{n3} & \cdots & S_{nn}^2 \end{bmatrix} \qquad (۳-۲۶)$$

که با استفاده از این ماتریس دربررسی خطا ها وتجزیه وتحلیل اولیه (Pri_Analyses) مشاهدات می توانیم منحنی خطا ها وحدجا بجائی خطا را موردبررسی قرار دهیم ودرنتیجه بتوانیم از یک سری مشاهدات اولیه به یک سری اندازه های نهائی دست یابیم .

- If the elements of the multivariate are statistically independent (no correlation), then the variance-covariance matrix will be diagonal matrix.

$$C_L = \begin{bmatrix} \sigma_a^2 & 0 & 0 \\ 0 & \sigma_b^2 & 0 \\ 0 & 0 & \sigma_c^2 \end{bmatrix} = diag\left(\sigma_a^2\ \sigma_b^2\ \sigma_c^2\right)$$

- The following matrices cannot be variance-covariance matrices:

$$\mathbf{A} = \begin{bmatrix} 3 & 4 \\ 4 & -1 \end{bmatrix}$$

The variance cannot be a negative number

$$\mathbf{A} = \begin{bmatrix} 5 & 1 & -2 \\ 1 & 3 & 0 \\ 2 & 0 & 4 \end{bmatrix}$$

Matrix is not symmetric

$$\mathbf{A} = \begin{bmatrix} 6 & 6 \\ 6 & 6 \end{bmatrix}$$

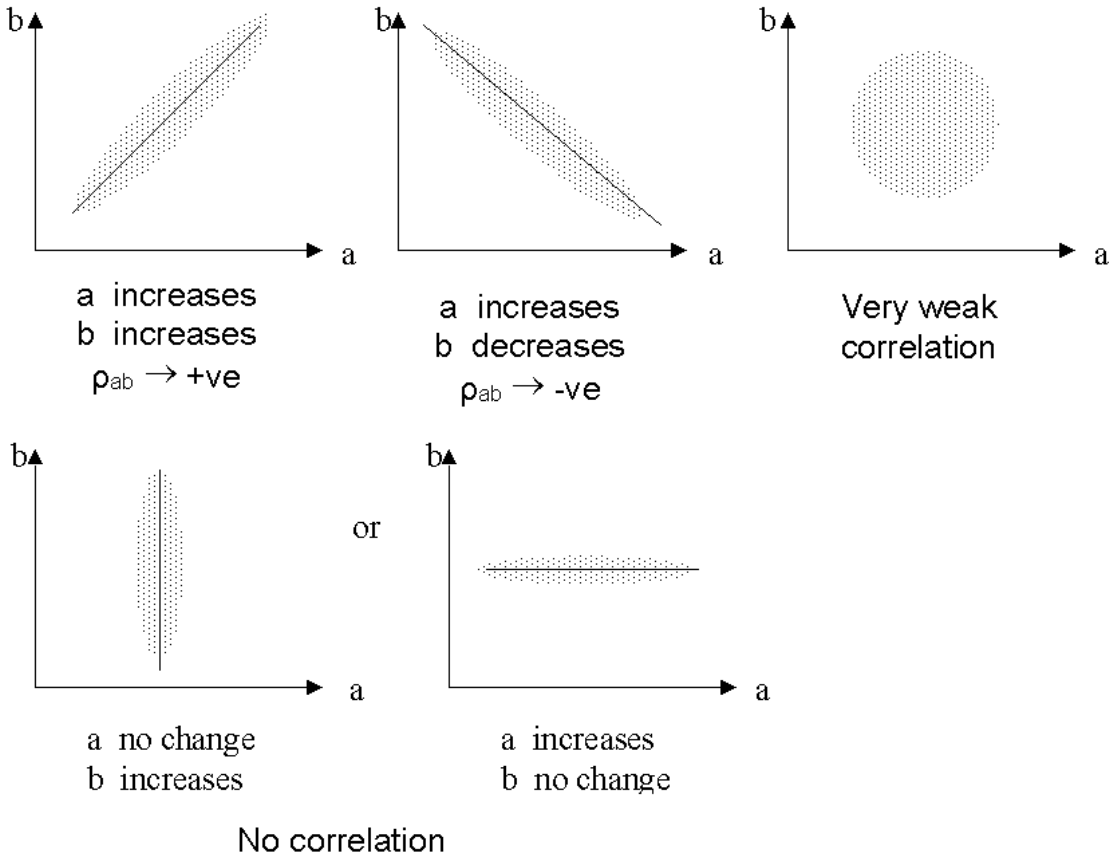
Matrix is not invertable – determinant is equal to zero (|**A**| = 0)

مقدار عددی کوواریانس دومتغیر تصادفی نمی‌تواند به‌روشنی بیان‌گر میزان وابستگی بین آن‌ها باشد. بر همین‌اساس امکان مقایسهٔ میزان وابستگی بین زوج متغیرهای تصادفی نیز وجودندارد. بنابراین یک‌کمیت جدید، به‌نام ضریب هم‌بستگی[1] بین دومتغیر تصادفی X و Y که دارای مفهوم نسبی است، معرفی و به‌جای کوواریانس استفاده‌می‌شود. رابطهٔ ضریب همبستگی به‌صورت زیر است و مقدار عددی آن در فاصلهٔ $[-1,1]$ قراردارد.

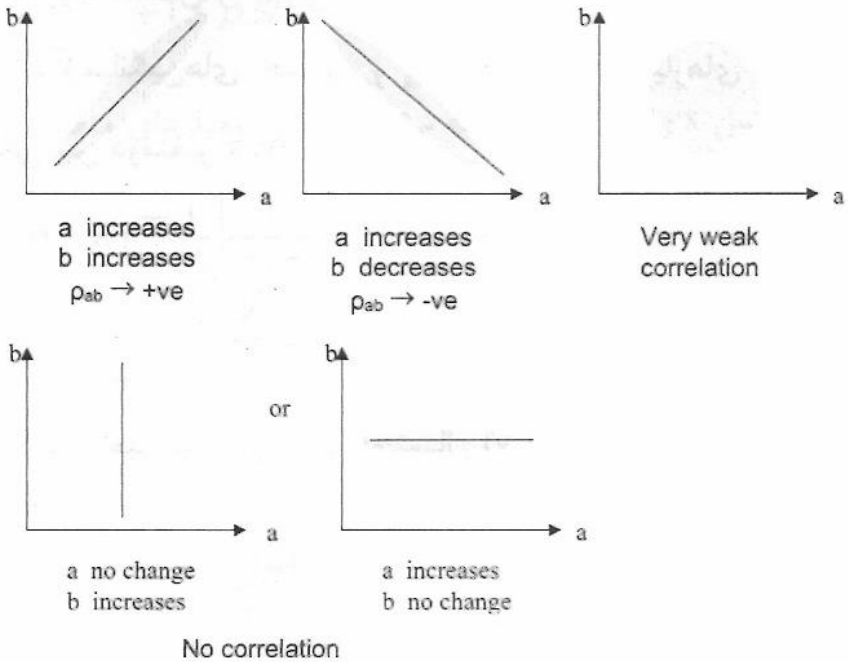
$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x . \sigma_y}$$

که در آن σ_{xy} کوواریانس بین دو متغیر تصادفی X و Y، σ_x انحراف‌معیار متغیر تصادفی X، و σ_y انحراف‌معیار متغیر تصادفی Y است.

به عنوان مثال مختصات x، y نقاط در یک پیمایش همیشه به یکدیگر ارتباط دارند، چون از زوایا و طولهای مشترکی محاسبه شده‌اند در نتیجه وابسته به هم‌اند. اما در اندازه‌گیری زوایا در ایستگاههای مختلف یک شبکه مثلث‌بندی در صورتیکه شامل هیچ گونه امتداد یا نشانه مشترکی نباشند، زوایا مستقل از هم می‌باشند.



اگر بین دو متغیر همبستگی وجود داشته باشد، این همبستگی به دو صورت مثبت و منفی بیان می‌شود.



شکل ۲-۴: انواع همبستگی

۲-۵-۱۱-۱-۱- همبستگی مثبت (مستقیم)

در این حالت، اندازه‌های دو متغیر در یک جهت تغییر می‌کنند یعنی x و y هر دو بزرگ و یا هر دو کوچک می‌شوند.

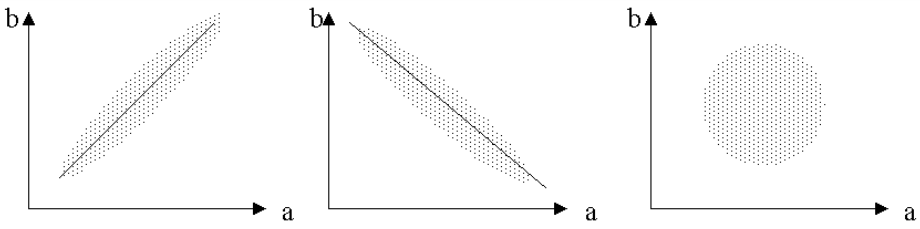
۲-۵-۱۱-۱-۲- همبستگی منفی (معکوس)

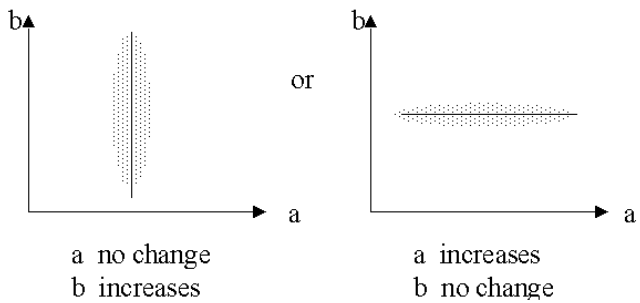
در این حالت، اندازه‌های دو متغیر در خلاف جهت هم تغییر می‌کنند. یعنی اگر x بزرگ شود y کوچک می‌شود و برعکس.

همانطوریکه گفته شد ضریب همبستگی خطی در فاصله $[-1,1]$ قرار دارد، لذا

- اگر $\rho = 0$ باشد، بین دو متغیر همبستگی خطی وجود ندارد. (عدم همبستگی)
- اگر $\rho = +1$ باشد، بین دو متغیر همبستگی خطی کامل و مستقیم (مثبت) وجود دارد.
- اگر $\rho = -1$ باشد، بین دو متغیر همبستگی خطی کامل معکوس (منفی) وجود دارد.

لازم به ذکر می‌باشد که در دو حالت اخیر، نقطه‌ها دقیقا روی یک خط راست (صعودی یا نزولی) قرار می‌گیرند.





No correlation

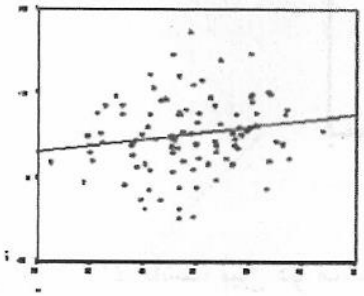
- اگر $0 < \rho < 1$ باشد بین دو متغیر همبستگی خطی ناقص مستقیم (مثبت) وجود دارد که هر قدر به 1 نزدیکتر می‌شود بر شدت این همبستگی افزوده خواهد شد.
- اگر $-1 < \rho < 0$ باشد بین دو متغیر همبستگی خطی ناقص معکوس (منفی) وجود دارد که هر قدر به -1 نزدیکتر می‌شود بر شدت این همبستگی افزوده خواهد شد.
- اگر $\rho = 0$ باشد فقط نشان دهنده آن است که بین x و y، رابطه خطی وجود ندارد و به معنی استقلال نیست، به عبارت دیگر ممکن است مثلا رابطه $y = x^2$ (درجه دوم، سوم و نمایی و لگاریتمی و غیره) برقرار باشد که نشان دهنده یک نوع همبستگی است ولی ما فقط بستگی خطی را مورد بررسی قرار می‌دهیم.

دقت کنید همانطور که می‌دانیم $-1 \le \rho \le 1$ حال اگر دو ضریب همبستگی مختلف $\rho_1 = 0.3, \rho_2 = 0.6$ در نظر گرفته شود می‌توان بیان نمود که این دو ضریب همبستگی، مثبت هستند و از طرفی نمی‌توان استنباط کرد که $\rho_2 = 0.6$ دارای دو برابر همبستگی خطی نسبت به $\rho_1 = 0.3$ است.

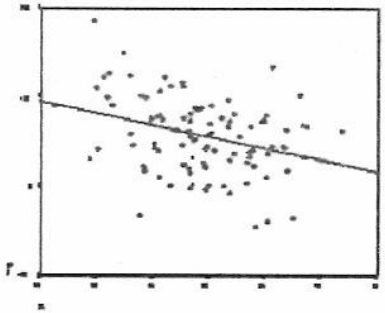
ضریب همبستگی

در زیر چند نمودار پراکنش رسم و نوع همبستگی نیز به وسیله ρ مشخص شده است.

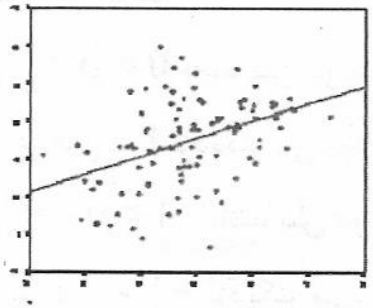
$\rho = +0.17$



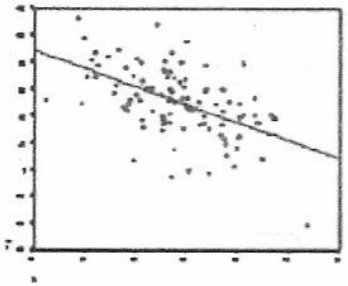
$\rho = -0.33$



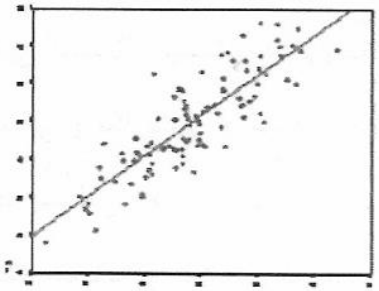
$\rho = +0.42$



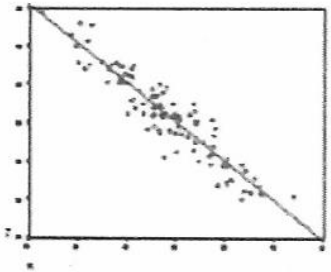
$\rho = -0.54$



$\rho = +0.85$



$\rho = -0.94$



ضریب همبستگی (مثال)

مثال : در چند نمونه روبرو میانگین، وریانس، کواریانس و ضریب همبستگی را محاسبه کنید.

$$D = \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} (1,2,3) \\ (3,2,1) \end{bmatrix} \qquad n_a = 3 \qquad n_b = 3$$

$$M = \begin{bmatrix} m_a \\ m_b \end{bmatrix} = \begin{bmatrix} \frac{1}{n_a}\sum_{i=1}^{n_a}(a_i) \\ \frac{1}{n_b}\sum_{i=1}^{n_b}(b_i) \end{bmatrix} = \begin{bmatrix} \frac{1}{3}(1+2+3) \\ \frac{1}{3}(3+2+1) \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$S^2 = \begin{bmatrix} s_a^2 \\ s_b^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{n_a}\sum_{i=1}^{n_a}(a_i - m_a)^2 \\ \frac{1}{n_b}\sum_{i=1}^{n_b}(b_i - m_b)^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3}((1-2)^2 + (2-2)^2 + (3-2)^2) \\ \frac{1}{3}((3-2)^2 + (2-2)^2 + (1-2)^2) \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

$$S_{a,b} = S_{b,a} = \frac{1}{n_{a\downarrow b}}\sum_{k=1}^{n_a}(a_i - m_a) \times (b_i - m_b)$$

$$= \frac{1}{3}[(1-2)(3-2) + (2-2)(2-2) + (3-2)(1-2)] = \frac{1}{3}[-1-1] = -\frac{2}{3}$$

$$\Sigma_D = \begin{bmatrix} S_a^2 & S_{ab} \\ S_{ba} & S_B^2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} \end{bmatrix}$$

$$\rho_{a,a} = \frac{S_{a,a}}{S_a \times S_a} = \frac{\frac{2}{3}}{\sqrt{\frac{2}{3}} \times \sqrt{\frac{2}{3}}} = 1$$

$$\rho_{b,b} = \frac{S_{b,b}}{S_b \times S_b} = \frac{\frac{2}{3}}{\sqrt{\frac{2}{3}} \times \sqrt{\frac{2}{3}}} = 1$$

$$\rho_{a,b} = \rho_{b,a} = \frac{S_{a,b}}{S_a \times S_b} = \frac{-\frac{2}{3}}{\sqrt{\frac{2}{3}} \times \sqrt{\frac{2}{3}}} = -1$$

مثال

مثال ۱۰-۳) سه نمونه بصورت زیر موجود است محاسبه کنید کوواریانس های موجود بین هردو-نمونه.

$$x^1 = (1, 2, 3, 4, 5)$$

$$x^2 = (6, 7, 8, 9, 10)$$

$$x^3 = (10, 11, 12, 13, 14)$$

حل :

$M_{x^1} = 3 \qquad M_{x^2} = 8 \qquad M_{x^3} = 12$

$$S_{x^2x^1} = S_{x^1x^2} = E(x^1 - M_x)(x^2 - M_{x^2}) = \frac{1}{n}\sum_{i=j=1}^{5}(x^1_i - M_{x^1})(x^2_j - M_{x^2})$$

$$= \frac{1}{5}\left[(1-3)(6-8) + (2-3)(7-8) + (3-3)(8-8) + (4-3)(9-8) + (5-3)(10-8)\right]$$

$$\Rightarrow S_{x^2x^1} = S_{x^1x^2} = \frac{1}{5}(4+1+1+4) = 2$$

$$S_{x^1x^3} = S_{x^3x^1} = \frac{1}{5}\left[(1-3)(10-12)+(2-3)(11-12)+(3-3)(12-12)+(4-3)(13-12)+(5-3)(14-12)\right]$$

$$= \frac{1}{5}(4+1+1+4) = 2$$

$$S_{x^3x^2} = S_{x^2x^3} = \frac{1}{5}\left[(6-8)(10-12)+(7-8)(11-12)+(8-8)(12-12)+(9-8)(13-12)+(10-8)(14-12)\right]$$

$$= \frac{1}{5}(4+1+1+4) = 2$$

مثال

$$\Sigma_{\underset{3\times 3}{x}} = \begin{bmatrix} S_1^2 & S_{12} & S_{13} \\ S_{21} & S_2^2 & S_{23} \\ S_{31} & S_{32} & S_3^2 \end{bmatrix}$$

واریانس ها به ترتیب عبارتند از :

$$S_1^2 = \frac{1}{5}\sum_{i=1}^{5}(x_i^1 - Mx^1)^2$$

$$\Rightarrow S^2 = \frac{1}{5}\left[(1-3)^2+(2-3)^2+(3-3)^2+(4-3)^2+(5-3)^2\right]$$

$$\Rightarrow S^2 = \frac{1}{5}(4+1+0+1+4) = 2$$

$$\Rightarrow S_2^2 = S_3^2 = S_1^2 = 2$$

وهمچنین داشتیم :

$$S_{12} = S_{21} = S_{13} = \cdots = S_{32} = 2$$

درنتیجه ماتریس واریانس کوواریانس خواهد شد :

$$\Sigma_x = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}$$

$$\rho_{12} = \rho_{21} = \frac{S_{12}}{\sqrt{S_1^2 \times S_2^2}} = \frac{S_{21}}{\sqrt{S_2^2 \times S_1^2}} = \frac{2}{|2|} = |1| = \rho_{13} = \rho_{31} = \cdots$$

سوال ۱) نشان دهید که تابع چگالی زیر، یک تابع چگالی احتمالی نیست؟

$$\begin{cases} 2x - x^2 & 0 \le x \le 2 \\ 0 & x > 2,\ x < 0 \end{cases}$$

سوال ۲) مقدار k را طوری تعیین کنید که تابع زیر تابع چگالی احتمال باشد؟

$$f(x) = \begin{cases} \frac{x}{6} + k & 0 \le x \le 3 \\ 0 & x < 0, x > 3 \end{cases}$$

سوال ۳) امید ریاضی کمیت $y = 2x_1 - 3x_2 + 4$ را محاسبه کنید؟

سوال ۴) اگر E[x]=1 و E[x(x-1)]=2 باشد. مطلوبست محاسبه واریانس x؟

سوال ۵) اگر $y = ax + b$ که a,b ثابت هستند و E[x] = m و $\mathrm{var}(x) = n^2$ باشد آنگاه مطلوبست E[xy] ؟

سوال ۶ : اگر متغیر تصادفی x نشان دهنده زمان کارکرد کامپیوتر قبل از خراب شدن باشد که دارای تابع چگالی احتمال به صورت $f(x) = \lambda e^{\frac{-x}{100}}$ بوده و $x \geq 0$ و λ هم معلوم است. حساب کنید احتمال اینکه:

الف) کامپیوتر قبل از خراب شدن (50,150) ساعت کار کند؟

ب) $x < 100$ ساعت کار کند؟

سوال ۷ : در چند نمونه زیر میانگین، وریانس، کواریانس، ماتریس وریانس-کواریانس و ضریب همبستگی را محاسبه کنید.

$$D = \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} (3,4,5,6) \\ (6,5,4,3) \end{bmatrix}$$

دانشجویان عزیز توجه داشته باشند که حل کامل و مرتب تمرینات این جلسه، تأثیر مستقیم در نمره پایان ترم شما خواهد داشت.

پایان جلسه سوم

درس تئوری خطاها

جلسه چهارم

برخی از توزیع های احتمال

فرید اسماعیلی
Farid_63@yahoo.com
www.faridesm.ir

تماس با استاد از طریق پست الکترونیکی
مشاهده اطلاعیه ها، نمرات، دریافت فایل ها در وب سایت

چرا تابع توزیع خاص؟

بعضی از متغیرهای تصادفی دارای یک الگوی خاص هستند و می‌توان برای آنها یک نوع توزیع احتمال را در نظر گرفت. برای مثال در انجام آزمایشات پرتاب یک تاس ۷ مرتبه و پرتاب یک نیزه به طرف هدف، فرض کنید که

X = تعداد مشاهده عدد ۴ در ۷ مرتبه پرتاب تاس

Y = تعداد برخورد به هدف نیزه در ۵ مرتبه پرتاب نیزه

همانطور که دیده می‌شود این دو متغیر تصادفی دارای یک شکل بخصوص هستند، در حقیقت هر دو تعداد موفقیتها در n آزمایش مستقل را بیان می‌کنند. حال اگر توزیع احتمال برای تعداد موفقیتها در n آزمایش مستقل را به دست آوریم آنگاه توزیع احتمال برای تعداد زیادی متغیر تصادفی مشابه با X و Y گفته شده در بالا را به دست آورده‌ایم.

مثال دو تاس را با هم پرتاب می‌کنیم و متغیر تصادفی X را برابر مجموع اعداد روی دو تاس مشاهده شده در نظر می‌گیریم.

الف- تابع احتمال X را به دست آورید.

ب- احتمال اینکه مجموع اعداد روی دو تاس حداکثر ۴ شود را بیابید.

ج- احتمال اینکه مجموع اعداد روی دو تاس بین ۶ و ۹ شود را بیابید.

حل الف- تکیه گاه X برابر $S_X = \{۲, ۳, ۴, ..., ۱۲\}$ است و برای مثال

$$P(X=۴) = P(\{(۱,۳),(۲,۲),(۳,۱)\}) = \frac{۳}{۳۶}$$

با محاسبه احتمالات مربوط به نقاط دیگر S_X، تابع احتمال X به صورت زیر به دست می‌آید.

x	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
$f_X(x) = P(X=x)$	$\frac{۱}{۳۶}$	$\frac{۲}{۳۶}$	$\frac{۳}{۳۶}$	$\frac{۴}{۳۶}$	$\frac{۵}{۳۶}$	$\frac{۶}{۳۶}$	$\frac{۵}{۳۶}$	$\frac{۴}{۳۶}$	$\frac{۳}{۳۶}$	$\frac{۲}{۳۶}$	$\frac{۱}{۳۶}$

جدول فوق یک تابع از S_X به اعداد حقیقی در فاصله [۰,۱] برقرار می‌کند که آن را با تابع $f_X(x)$ نمایش داده و به آن تابع احتمال متغیر تصادفی X گویند، بنابراین

$$f_X(x) = P(X=x)$$

این تابع یک تابع غیر منفی است و مجموع آن روی کلیه مقادیری که می‌تواند اختیار کند برابر ۱ است.

تابع توزیع احتمال گسسته (یاد آوری از جلسه قبل)

x	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
$f_X(x)=P(X=x)$	$\frac{۱}{۳۶}$	$\frac{۲}{۳۶}$	$\frac{۳}{۳۶}$	$\frac{۴}{۳۶}$	$\frac{۵}{۳۶}$	$\frac{۶}{۳۶}$	$\frac{۵}{۳۶}$	$\frac{۴}{۳۶}$	$\frac{۳}{۳۶}$	$\frac{۲}{۳۶}$	$\frac{۱}{۳۶}$

ب – $P(X\leq ۴)$ مورد سئوال است بنابراین از فرمول (۱.۳) داریم که

$$P(X\leq ۴)=f_X(۲)+f_X(۳)+f_X(۴)=\frac{۱}{۳۶}+\frac{۲}{۳۶}+\frac{۳}{۳۶}=\frac{۱}{۶}$$

ج – با استفاده از فرمول (۱.۳) داریم که

$$P(۶<X<۹)=f_X(۷)+f_X(۸)=\frac{۶}{۳۶}+\frac{۵}{۳۶}=\frac{۱۱}{۳۶}$$

تعریف تابع $f_X(x)=P(X=x)$ را تابع احتمال متغیر تصادفی X گویند هر گاه

۱ـ برای هر $x\in R$ داشته باشیم که $f_X(x)\geq ۰$

۲ـ $\sum_{x\in R} f_X(x)=۱$

تابع توزیع یکنواخت گسسته

ساده‌ترین تابع توزیع احتمال، تابع توزیع یک‌نواخت گسسته[1] است که تمام مقادیر متغیر تصادفی آن دارای احتمال مساوی است. بنابراین برای یک‌متغیر تصادفی X با n مقدار $(x_1, x_2, \ldots\ldots, x_n)$ تابع توزیع یک‌نواخت به‌صورت زیر معرفی‌می‌گردد.

$$f(x) = \frac{1}{n} \qquad (۲۸.۲)$$

تابع توزیع برنولی

آزمایشی را در نظر بگیرید که دارای دو نتیجه موفقیت با احتمال p و شکست با احتمال $q=۱-p$ باشد. چنین آزمایشی را آزمایش برنولی گویند. برای مثال پرتاب یک سکه یک آزمایش برنولی است که در آن موفقیت مشاهده شیر با $p=\frac{۱}{۲}$ و مشاهده خط شکست با $q=\frac{۱}{۲}$ می‌باشد. همچنین در پرتاب یک تاس که در آن موفقیت مشاهده عدد ۴ با $p=\frac{۱}{۶}$ و مشاهده عدد غیر از ۴ شکست با $q=\frac{۵}{۶}$ است. اگر متغیر تصادفی X را به صورت زیر تعریف کنیم

$$X=\begin{cases} ۱ & \text{اگر نتیجه آزمایش موفقیت باشد} \\ ۰ & \text{اگر نتیجه آزمایش شکست باشد} \end{cases}$$

در این صورت متغیر تصادفی X را متغیر تصادفی برنولی گوئیم و آن را با نماد $X\sim B(۱,p)$ نمایش داده و گوئیم X دارای توزیع برنولی با پارامتر p است (عدد ۱ نمایانگر انجام یک بار آزمایش است). تابع احتمال متغیر تصادفی برنولی به صورت زیر به دست می‌آید

x	۰	۱
$f_X(x)=P(X=x)$	$۱-p$	p

که آن را می‌توان در فرمول زیر خلاصه کرد

$$\boxed{f_X(x)=p^x(۱-p)^{۱-x}\ ,\ x=۰,۱ \quad \text{تابع احتمال توزیع برنولی}} \qquad (۱.۵)$$

در قضیه زیر امید ریاضی و واریانس این توزیع را به دست می‌آوریم.

قضیه ۱.۵ اگر $X \sim B(۱,p)$ آنگاه $E(X)=p$ و $Var(X)=pq$.

اثبات

$$E(X)=(۰)(۱-p)+(۱)(p)=p$$

$$E(X^۲)=(۰^۲)(۱-p)+(۱^۲)(p)=p$$

بنابراین

$$Var(X)=p-p^۲=p(۱-p)=pq$$

مثال ۱.۲.۵ یک تاس را یک مرتبه پرتاب می‌کنیم و قرار می‌دهیم:

$$X=\begin{cases} ۱ & \text{اگر عدد ۴ مشاهده شود} \\ ۰ & \text{در غیر اینصورت} \end{cases}$$

در این صورت $X \sim B(۱,\frac{۱}{۶})$ و همچنین

$$f_X(x)=(\frac{۱}{۶})^x(\frac{۵}{۶})^{۱-x} \quad , \quad x=۰,۱$$

$$E(X)=\frac{۱}{۶} \quad , \quad Var(X)=\frac{۱}{۶}(\frac{۵}{۶})=\frac{۵}{۳۶}$$

تابع توزیع دو جمله ای

آزمایش تصادفی که از انجام چند آزمایش مستقل برنولی بوجود می‌اید را ازمـایش دو جمله‌ای گویند. برای مثال انتخاب ۵ مهره با جایگذاری از ظرفی که دارای ۸ مهره سفید و ۶ مهره قرمز است و موفقیت برای ما مشاهده مهره قرمز است یک آزمایش دو جمله‌ای با احتمال موفقیت $\frac{۶}{۱۴}$ می‌باشد.

به طور کلی یک آزمایش دو جمله‌ای دارای خواص زیر است.

۱ـ آزمایش دو جمله‌ای از انجام n آزمایش مستقل برنولی بوجود آمده است.

۲ـ در آزمایشات برنولی احتمال موفقیت p (و شکست $q=۱-p$) می‌باشد که در تمام آزمایشات برنولی مقداری ثابت است.

مثال ۲.۳.۵ بسکتبالیستی ۶۰٪ از توپهایش گل می‌شود. اگر او ۵ پرتاب مستقل انجام دهـد و موفقیت برای او گل شدن توپ باشد، در این صورت یک آزمایش دو جمله‌ای با پارامترهای $n=۵$ و $p=۰/۶$ داریم.

تعریف ۱.۵ اگر متغیر تصادفی X به صورت زیر تعریف شود

X = تعداد موفقیتها در n آزمایش مستقل برنولی

آنگاه X را یک متغیر تصادفی دو جمله‌ای گویند. اگر احتمال موفقیت در هر آزمایش برنولی p باشد، آنگاه گوئیم که X دارای توزیع دو جمله‌ای با پارامترهای n و p است و آن را با نماد $X \sim B(n,p)$ نمایش می‌دهیم.

بنابراین تابع احتمال X برابر است با

$$f_X(x) = \binom{n}{x} p^x q^{n-x} \ , \ x = ۰, ۱, ۲, \dots, n \qquad \text{تابع احتمال توزیع دوجمله‌ای} \qquad (۲.۵)$$

قضیه ۲.۵ اگر $X \sim B(n,p)$ آنگاه $E(X) = np$ و $Var(X) = npq$.

مثال ۴.۳.۵ یک تاس را ۵ بار پرتاب می‌کنیم.

الف- احتمال اینکه عدد ۴ دقیقاً ۳ بار مشاهده شود را بیابید؟

$$f_X(x) = \binom{۵}{x} \left(\frac{۱}{۶}\right)^x \left(\frac{۵}{۶}\right)^{۵-x} \quad , \quad x = ۰, ۱, \ldots, ۵$$

بنابراین

الف- $P(X=۳) = f_X(۳) = \binom{۵}{۳} \left(\frac{۱}{۶}\right)^۳ \left(\frac{۵}{۶}\right)^{۵-۳} = ۰/۰۳۲$

یاد آوری :

$$\binom{n}{r} = \frac{n!}{r! \times (n-r)!}$$

تابع توزیع پواسون

آزمایشی که تعداد موفقیتها در یک فاصله زمانی یا ناحیه مشخص را به دست دهد یک آزمایش پواسون نامیده می‌شود. این فاصله زمانی می‌تواند هر فاصله زمانی مانند ثانیه، دقیقه و... باشد و ناحیه مشخص نیز می‌تواند یک فاصله خطی یا مساحت یا حجم یا... باشد.

تعریف ۳.۵ اگر متغیر تصادفی X به صورت زیر تعریف شود

X = تعداد موفقیتها در یک فاصله زمانی یا ناحیه مشخص

آنگاه X را یک متغیر تصادفی پواسون گویند.

برای مثال اگر X تعداد تلفن هایی باشد که در یک ساعت به رییس یک شرکت زده می‌شود و یا X تعداد غلطهای تایپی در هر صفحه یک کتاب باشد و یا X تعداد تاکسی‌هایی باشد که در یک ساعت معین از یک چهار راه عبور می‌کنند، آنگاه X یک متغیر تصادفی پواسون است.

تعریف ۴.۵ اگر متغیر تصادفی X برابر تعداد موفقیتها در یک فاصله زمانی یا ناحیه مشخص باشد آنگاه $X \sim P(\mu)$ و

تابع احتمال توزیع پواسون

$$f_X(x) = \frac{e^{-\mu}\mu^x}{x!} \quad , \quad x = ۰, ۱, ۲, \ldots \qquad (۴.۵)$$

قضیه ۴.۵ اگر $X \sim P(\mu)$ آنگاه $E(X) = \mu$ و $Var(X) = \mu$.

مثال ۱.۵.۵ یک دستگاه چاپگر کامپیوتری به طور متوسط در هر ماه ۲ بار سرویس می‌شود.

الف- احتمال اینکه در یک ماه کمتر از ۲ بار سرویس شود را بیابید.

حل الف- اگر متغیر تصادفی X برابر تعداد دفعاتی باشد که چاپگر در یک ماه سرویس می‌شود آنگاه $X \sim P(۲)$ و بنابراین

$$P(X < ۲) = f_X(۰) + f_X(۱) = \frac{e^{-۲}۲^۰}{۰!} + \frac{e^{-۲}۲^۱}{۱!} = ۰/۴۰۶$$

تابع توزیع احتمال پیوسته (یاد آوری از جلسه قبل)

متغیر تصادفی که مجموعه مقادیر آن یک فاصله عددی یا اجتماع چند فاصله عددی باشد را متغیر تصادفی پیوسته گویند. در مورد متغیر تصادفی پیوسته بایستی توجه کرد که احتمال اینکه یک متغیر تصادفی پیوسته بخواهد فقط یک مقدار بخصوص از مجموعه مقادیرش را بگیرد برابر صفر است. به مثال زیر توجه کنید.

بنابراین توزیع احتمال یک متغیر تصادفی پیوسته را نمی‌توان به صورت یک جدول نمایش داد. در این حالت توزیع احتمال متغیر تصادفی پیوسته X را به صورت یک تابع $f_X(x)$ نمایش داده و آن را **تابع چگالی احتمال** X می‌نامند.

تابع توزیع احتمال پیوسته (یاد آوری از جلسه قبل)

برای به دست آوردن تابع چگالی احتمال یک متغیر تصادفی پیوسته معمولاً ابتدا تابع توزیع آن را به دست می‌آورند، زیرا تابع توزیع احتمال را در فواصل محاسبه می‌کند و محاسبه این احتمالات در حالت پیوسته امکان‌پذیر است. تابع توزیع یک متغیر تصادفی پیوسته X به صورت $F_X(x)=P(X\leq x)$ تعریف می‌شود. با توجه به آنکه در حالت گسسته تابع توزیع به صورت $F_X(x)=P(X\leq x) = \sum_{t\leq x} f_X(t)$ محاسبه می‌شد، با قرار دادن تابع چگالی احتمال به جای تابع احتمال و تبدیل مجموع به انتگرال می‌توان تابع توزیع در حالت پیوسته را به طور مشابه و به صورت زیر محاسبه کرد:

$$\boxed{F_X(x)=P(X\leq x) = \int_{-\infty}^{x} f_X(t)dt} \qquad (۴.۳)$$

تابع $f_X(x)$ که در رابطه (۴.۳) صدق می‌کند را تابع چگالی احتمال و تابع $F_X(x)$ در این رابطه را تابع توزیع متغیر تصادفی پیوسته X گویند. همچنین با توجه به قضیه اساسی حساب از رابطه (۴.۳) نتیجه می‌شود که

$$\boxed{F_X'(x)=\frac{d}{dx}F_X(x)=f_X(x)} \qquad (۵.۳)$$

با توجه به روابط (۴.۳) و (۵.۳) با داشتن تابع توزیع X و مشتق گرفتن از آن به راحتی می‌توان تابع چگالی احتمال X را به دست آورد و برعکس با داشتن تابع چگالی احتمال X و انتگرال گرفتن از آن می‌توان تابع توزیع X را به دست آورد.

مثال نقطه‌ای را به تصادف از فاصله اعداد حقیقی [۰,۲] انتخاب می‌کنیم و متغیر تصادفی X را برابر نقطه انتخاب شده در فاصله [۰,۲] در نظر می‌گیریم.

تابع توزیع متغیر تصادفی پیوسته X را به دست آورده و از روی آن تابع چگالی احتمال X را محاسبه کنید و سپس این دو تابع را رسم کنید.

حل با توجه به نمودار روبرو و مفهوم تابع توزیع **تجمعی**

اگر $x < ۰$ آنگاه $F_X(x) = P(X \le x) = ۰$ و اگر $۰ \le x < ۲$ آنگاه

$$F_X(x) = P(X \le x) = \frac{\text{طول فاصله } [۰, x]}{\text{طول فاصله } [۰,۲]} = \frac{x}{۲}$$

و اگر $x \ge ۲$ آنگاه $F_X(x) = P(X \le x) = ۱$. بنابراین

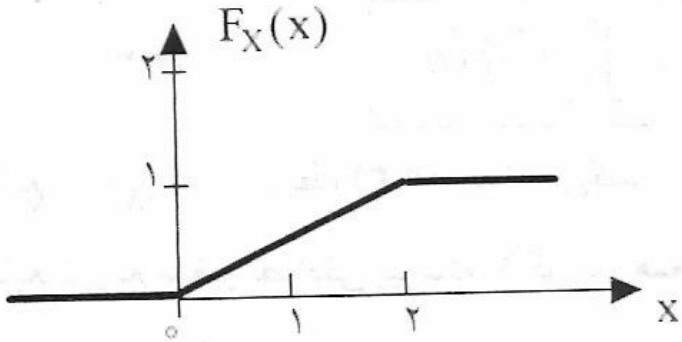
و اگر $x \geq ۲$ آنگاه $F_X(x) = P(X \leq x) = ۱$. بنابراین

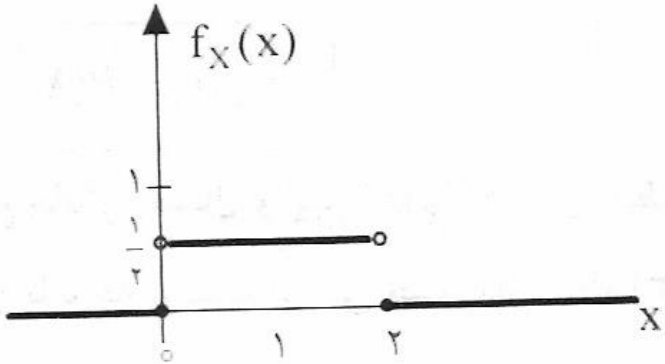
$$F_X(x) = P(X \leq x) = \begin{cases} ۰ & x < ۰ \\ \frac{x}{۲} & ۰ \leq x < ۲ \\ ۱ & ۲ \leq x \end{cases}$$

و در نتیجه

$$f_X(x) = F_X'(x) = \begin{cases} \frac{۱}{۲} & ۰ < x < ۲ \\ ۰ & \text{سایر نقاط} \end{cases}$$

نمودار این دو تابع در زیر رسم شده است. همان طور که دیده می‌شود نمودار تابع توزیع یک متغیر تصادفی پیوسته یک نمودار پیوسته است.

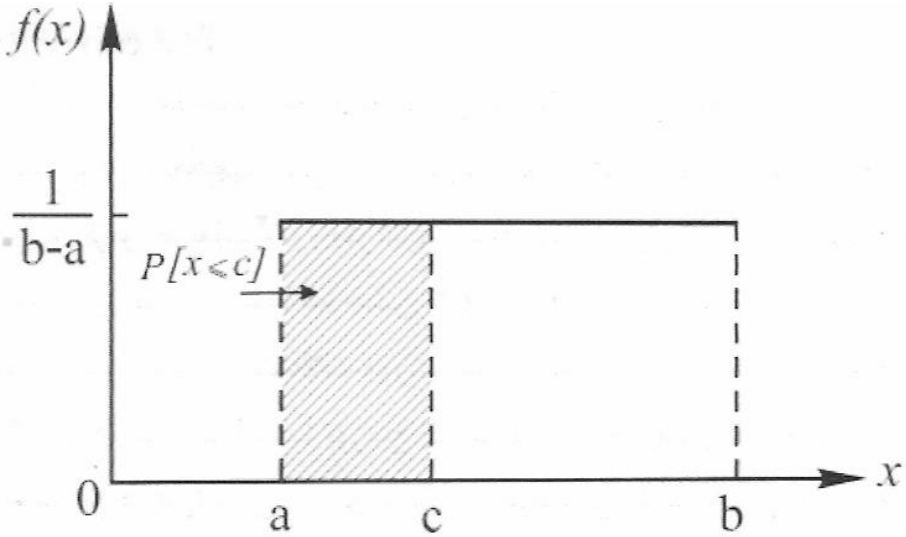




تابع توزیع یکنواخت پیوسته

مطابق نگارهٔ ۷.۲، متغیر تصادفی که توزیع آن‌در فاصلهٔ $[a,b]$ ثابت و در خارج از آن، صفر باشد دارای ساده‌ترین شکل تابع توزیع است که به‌آن تابع توزیع یک‌نواخت پیوسته[1] می‌گویند و آن‌را به‌صورت زیر معرفی‌می‌کنند:

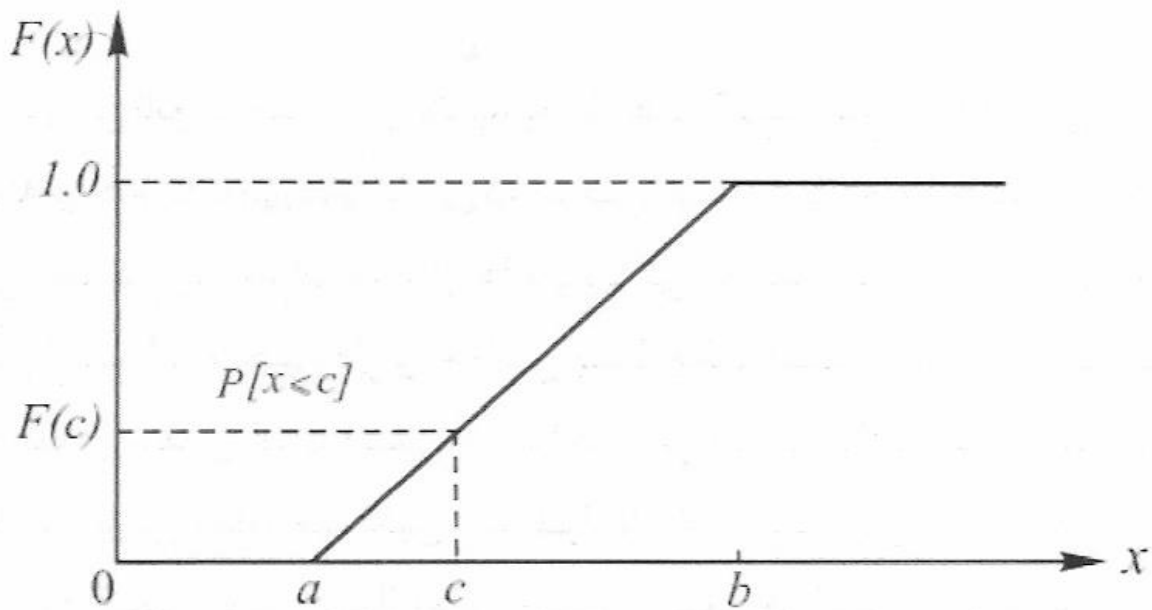
$$f(x)=\begin{cases}\dfrac{1}{b-a} & for \quad a<x<b\\ 0 & elsewhere\end{cases} \qquad (۳۳.۲)$$



نگارهٔ ۷.۲. نمایش هندسی تابع چگالی یک‌نواخت (G. Gracie & E.M. Mikhail, 1981)

بدیهی است با توجه به‌رابطهٔ بین تابع چگالی و تابع تجمعی، تابع توزیع تجمعی یک‌نواخت نیز به‌صورت زیر به‌دست‌می‌آید (نگارهٔ ۸.۲):

$$F(x)=\begin{cases}0 & for\ x\le a\\ \int_a^x \frac{1}{b-a}du=\frac{x-a}{b-a} & for\ a<x<b\\ 1 & for\ x\ge b\end{cases} \qquad (۳۴.۲)$$



نگارهٔ ۸.۲. نمایش هندسی تابع تجمعی یک‌نواخت (G. Gracie & E.M. Mikhail, 1981)

تابع توزیع نرمال

توزیع نرمال[1]، یکی از مهم‌ترین توابع توزیع در نظریهٔ احتمال است و کاربردهای بسیاری در علم فیزیک و در علوم مهندسی دارد. این‌توزیع را کارل فریدریش گاوس[2] در بارهٔ کاربرد روش کم‌ترین مربعات[3] در آمارگیری کشف‌شد. فرمول آن‌بر حسب، دو پارامتر میانگین و واریانس بیان‌می‌شود. هم‌چنین تابع توزیع نرمال یا گاوس از مهم‌ترین توابعی است که در مباحث آمار و احتمالات موردبررسی قرارمی‌گیرد، زیرا به‌تجربه ثابت شده‌است که در دنیای اطراف ما توزیع بسیاری از متغیرهای طبیعی از همین‌تابع پیروی‌می‌کنند. بر همین‌اساس، توزیع نرمال در مهندسی نقشه‌برداری نیز از جای‌گاهی ویژه برخوردار است و رفتار همهٔ اندازه‌ها و مشاهدات نقشه‌برداری به‌عنوان متغیرهای تصادفی نرمال از این‌تابع پیروی‌می‌کنند.

تعریف ۱۰.۵ گوئیم متغیر تصادفی X دارای توزیع نرمال با میانگین μ و واریانس σ^2 است و آن را با نماد $X \sim N(\mu, \sigma^2)$ نمایش می‌دهیم، هر گاه تابع چگالی احتمال آن به صورت زیر باشد

تابع چگالی احتمال توزیع نرمال

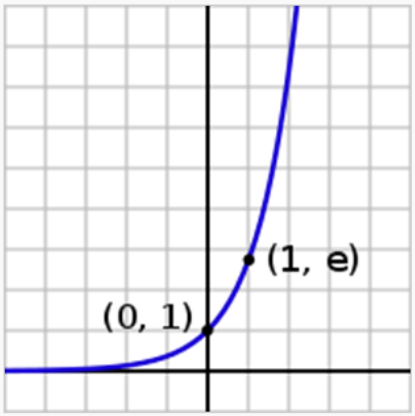
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \qquad -\infty < x < +\infty,\ \mu \in R,\ \sigma > 0$$

ویا :

$$N(\mu_x, \sigma_x) = f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left(\frac{-(x-\mu_x)^2}{2\sigma_x^2} \right)$$

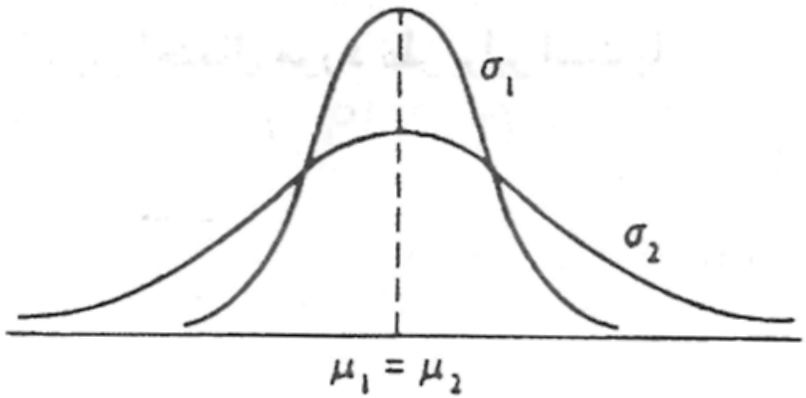
یاد آوری :

In mathematics, the **exponential** function is the function exp, where e is the number (approximately 2.718281828)



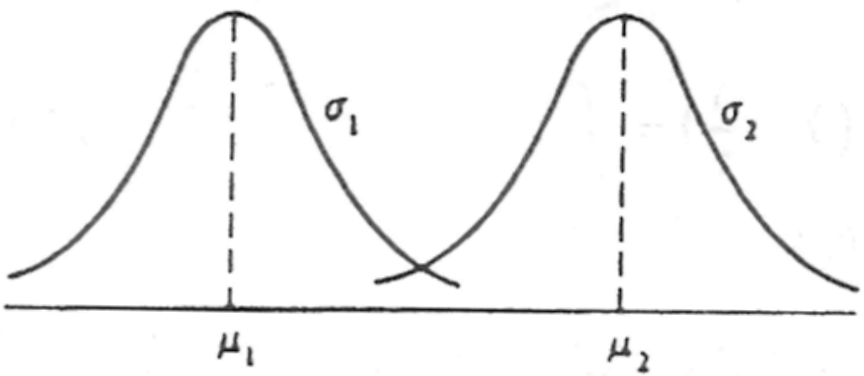
The natural exponential function
$y = e^x$

منحنی نرمال کاملاً بوسیله مقادیر μ و σ^2 مشخص می‌شود. با افزایش σ^2 پراکندگی توزیع افزایش می‌یابد و با افزایش μ منحنی به سمت راست انتقال پیدا می‌کند. در **زیر** منحنی نرمال در حالتهای مختلف μ و σ^2 رسم شده است.



شکل ۲.۵ – ب منحنی‌های نرمال با

$\mu_1 = \mu_2 , \sigma_1 < \sigma_2$

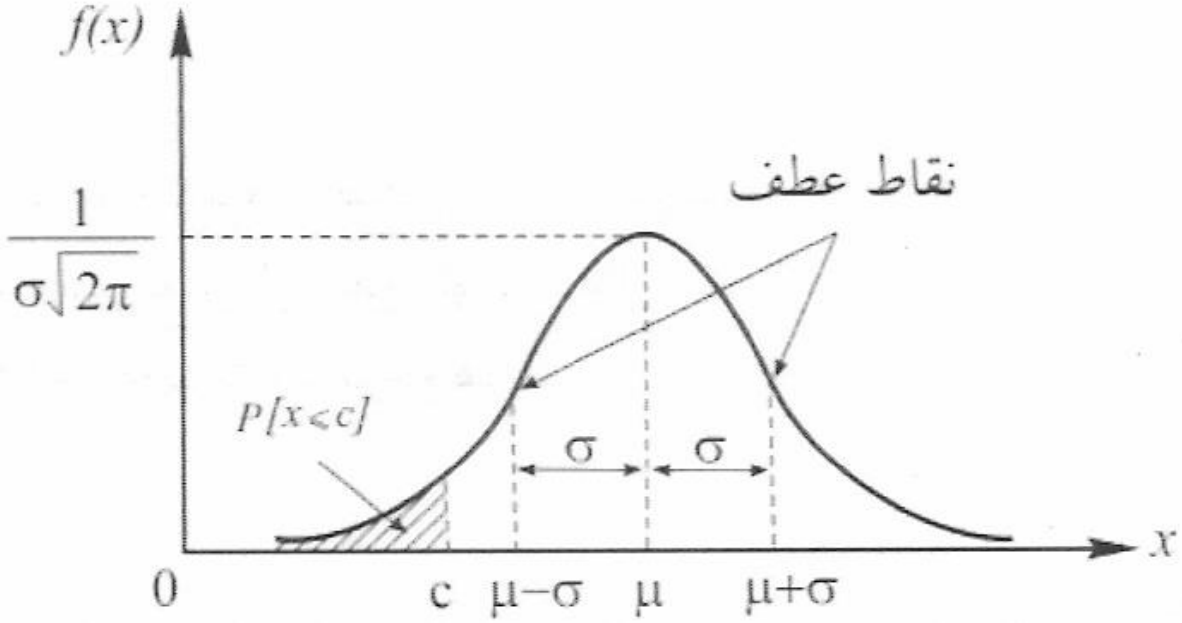


شکل ۲.۵ – الف منحنی‌های نرمال با

$\mu_1 < \mu_2 , \sigma_1 = \sigma_2$

منحنی رفتار این‌تابع تا حد زیادی شبیه به‌زنگ‌های کلیسا است و به‌همین دلیل به‌آن تابع زنگوله‌ای هم گفته‌می‌شود. با این‌که ممکن است ارتفاع و انحنای انواع مختلف این‌منحنی یک‌سان نباشد اما همهٔ آن‌ها یک‌ویژگی یک‌سان دارند و آن، مساحت واحد زیر منحنی است. این‌منحنی دارای خواصی بسیارجالب است از آن‌جمله این‌که نسبت به‌محور عمودی متقارن است (نیمی از مساحت زیر منحنی بالای مقدار میانگین μ_x و نیمهٔ دیگر در پایین مقدار میانگین μ_x قراردارد) و دیگر این‌که، هرچه از دوطرف به‌میانگین نزدیک‌می‌شویم احتمال وقوع بیشترمی‌شود.

فراگیری این تابع آن‌قدر زیاد است که دانشمندان اغلب برای مدل‌کردن متغیرهای تصادفی که با رفتار آن‌ها آشنایی ندارند، از آن‌استفاده‌می‌کنند. برای مثال، در یک‌امتحان درسی نمرات دانش‌آموزان اغلب اطراف میانگین بیشتر است و هر چه به‌سمت نمرات بالا یا پایین پیش برویم تعداد افرادی که این‌نمرات را گرفته‌اند کم‌تر می‌شود. این‌رفتار را به‌سهولت می‌توان با یک‌توزیع نرمال مدل‌کرد. در نگاره‌های **زیر** تفسیر هندسی توابع چگالی و تجمعی احتمال نرمال به‌طور شماتیک نمایش‌داده‌شده‌است.



خواص منحنی نرمال با استفاده از فرم تابع چگالی نرمال و همچنین نمودار آن خواص زیر را در مورد منحنی نرمال می‌توان بیان و اثبات کرد (اثبات به عنوان تمرین)

۱ـ منحنی تنها دارای یک نقطه ماکزیمم در نقطه $x=\mu$ است.

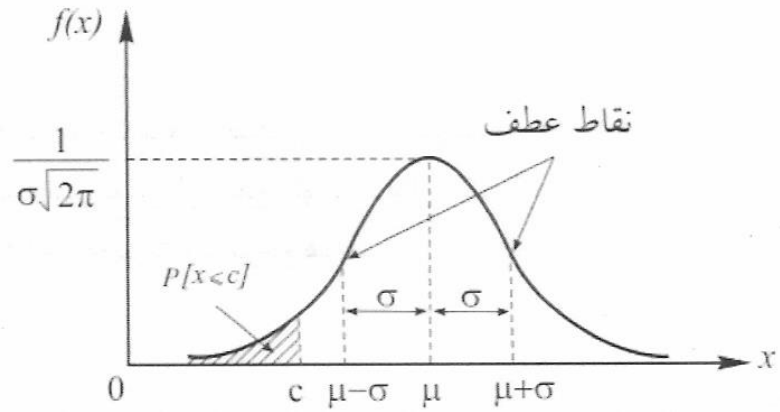
۲ـ منحنی دارای دو نقطه عطف در نقاط $x=\mu\pm\sigma$ است.

۳ـ منحنی نسبت به خط $x=\mu$ متقارن است، یعنی $f_X(\mu-a)=f_X(\mu+a)$

۴ـ در دو طرف حد متوسط μ منحنی به مجانب خود یعنی محور xها نزدیک می‌شود، یعنی

$$\lim_{x\to\pm\infty} f_X(x)=0$$

۵ـ سطح محصور بین منحنی و محور طولها برابر یک واحد است، یعنی $\int_{-\infty}^{+\infty} f_X(x)\,dx=1$



توزیع نرمال استاندارد

توزیع نرمالی که میانگین آن صفر و واریانس آن یک باشد را توزیع نـرمال اسـتاندارد می‌نامند و آن را با نماد $Z \sim N(۰,۱)$ نمایش می‌دهند.

چنان‌چه متغیر تصادفی x را به‌صورت زیر به‌حالت استاندارد تبدیل کنیم:

$$Z = \frac{x - \mu_x}{\sigma_x} \qquad (۳۷.۲)$$

متغیر جدید Z یک‌متغیر تصادفی نرمال با پارامترهـای $\mu_z = 0$ و $\sigma_z = 1$ اسـت. در این‌صورت به‌تابع توزیع آن، تابع توزیع نرمال استاندارد[1] گویند. در این‌حالت، تابع توزیـع چگالی $f(z)$ به‌صورت زیر درخواهدآمد:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} \qquad (۳۸.۲)$$

قضیه ۱۰.۵ اگر $Z \sim N(۰,۱)$ آنگاه $E(Z) = ۰$ و $Var(Z) = ۱$.

توزیع نرمال استاندارد

معمولا مقادیر مشخصی از تابع تجمعی نرمال استاندارد محاسبه‌می‌شوند و در انتهای اغلب کتب مرتبط به‌صورت جداول می‌آیند. چنان‌چه مقدار احتمال تجمعی برای هر متغیر نرمال استاندارد z_α موردنظر باشد، با توجه به‌روابط زیر، به‌راحتی از جداول آماری مرتبط قابل‌استخراج است (به‌بخش پیوست‌ها مراجعه‌شود).

$$F(z_\alpha) = \int_{-\infty}^{z_\alpha} f(z)dz = P(Z \le z_\alpha) = 1-\alpha \qquad (۴۰.۲)$$

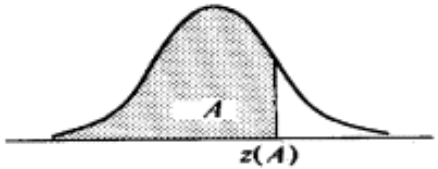
$$1-F(z_\alpha) = \int_{z_\alpha}^{+\infty} f(z)dz = P(Z \ge z_\alpha) = \alpha$$

که در آن α سطح زیرمنحنی از نقطهٔ z_α تا $+\infty$ است. بدیهی‌است چنان‌چه هدف محاسبهٔ احتمال بین دو مقدار $z_{\alpha 1}$ و $z_{\alpha 2}$ باشد، با استفاده‌از جداول مذکور به‌سادگی قابل‌محاسبه‌است.

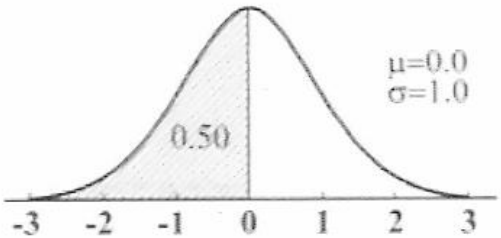
$$P(z_{\alpha 1} \le Z \le z_{\alpha 2}) = F(z_{\alpha 2}) - F(z_{\alpha 1}) \qquad (۴۱.۲)$$

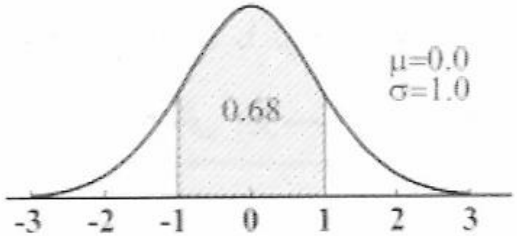
توابع توزیع پیوسته (توزیع نرمال)

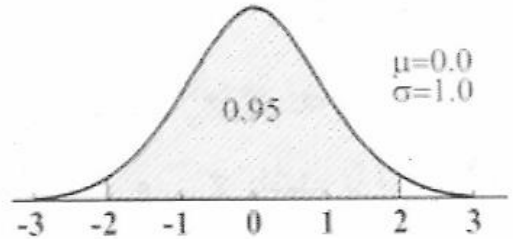
Entry is area A under the standard normal curve from $-\infty$ to $z(A)$



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998



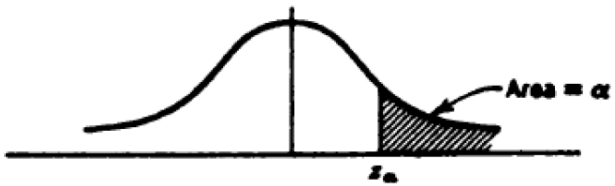




سطح زیر منحنی نرمال استاندارد برای سه‌حالت متفاوت

جدول آماری 1- سطح زیر منحنی تابع چگالی نرمال استاندارد

$$\alpha = \int_{z_\alpha}^{\infty} f(z)\,dz = P(z > z_\alpha) = 1 - \int_{-\infty}^{z_\alpha} f(z)\,dz$$



جدول آماری 1- سطح زیر منحنی تابع چگالی نرمال استاندارد

$$\alpha = \int_{z_\alpha}^{\infty} f(z)\,dz = P(z > z_\alpha) = 1 - \int_{-\infty}^{z_\alpha} f(z)\,dz$$

$z_\alpha \rightarrow$ ↓	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0076	0.0073	0.0071	0.0070	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0042	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0014	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002

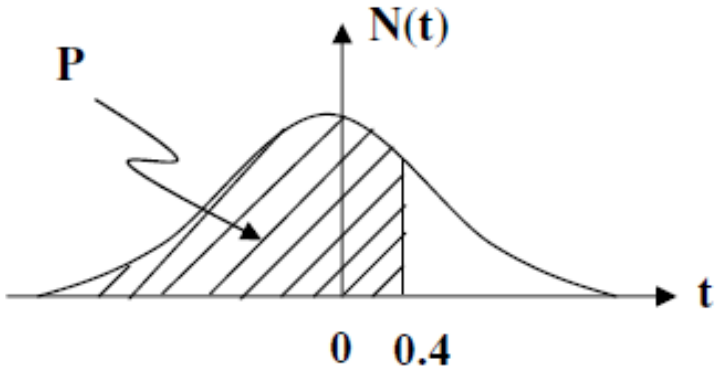
فرض می کنیم که قد دانشجو یک متغیر تصادفی با نام h و با تابع پخش نرمال بوده و دارای میانگین $\mu_h = 66$ و انحراف معیار $\sigma_h = 5$ اینچ باشد. تعداد دانشجویانی را که قد آنها در نامساوی های زیر صدق می کند به صورت عدد k در هزار پیدا کنید.

الف – $h \le 68$ (شکل 4-4a)

ب - $h \le 61$ (شکل 4-4b)

پ - $h \ge 74.6$ (شکل 4-4c)

ت - $[64.3 \le h \le 70]$ (شکل 4-4d)



(شکل 4-4a)

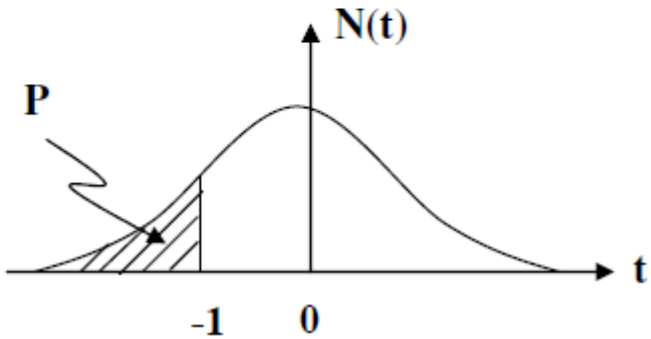
الف- در این قسمت از جدول II-B استفاده می کنیم

$$P(h \le 68) = P(t \le \frac{68-66}{5}) = P(t \le 0.4) = 0.6554$$

$$k_1 = (0.6554)(1000) = 655$$

ب-

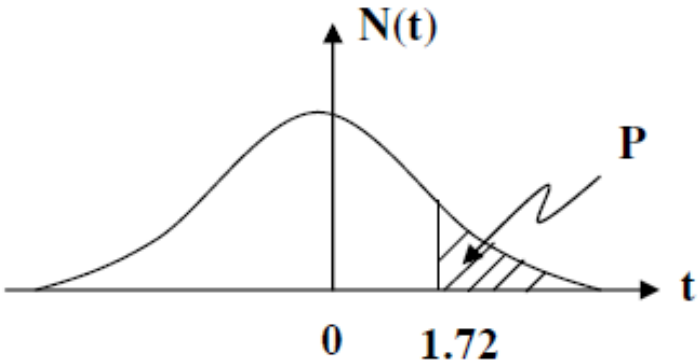
$$P(h \le 61) = P(t \le \frac{61-66}{5}) = P(t \le -1) = 1 - P(t \le 1) = 1 - 0.8413$$
$$= 0.1587, k_2 = (0.1587)(1000) = 159$$



(شکل 4-4b)

پ-

$$P(h \ge 74.6) = P(t \ge \frac{74.6-66}{5}) = P(t \ge 1.72) = 1 - P(t \le 1.72) = 1 - 0.9573 = 0.0427$$
$$k_3 = (0.0427)(1000) = 43$$



(شکل 4-4c)

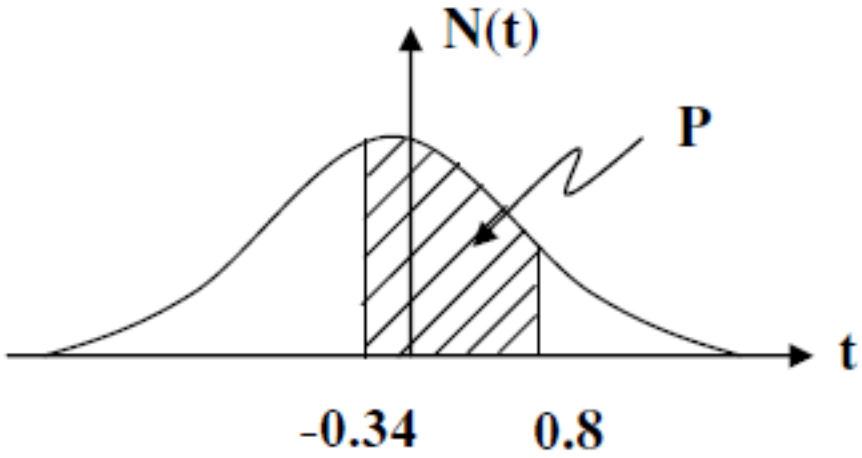
ت-

$$P(64.3 \le h \le 70) = P(\frac{64.3-66}{5} \le t \le \frac{70-66}{5}) = P(-0.34 \le t \le 0.8)$$

$$= P(t \le 0.8) - P(t \le -0.34) = P(t \le 0.8) - [1 - P(t \le 0.34)]$$

$$0.7881 - (1 - 0.6321) = 0.7881 - 0.3669 = 0.4212$$

$$k_4 = (0.4212)(1000) = 421$$



(شکل 4-4d)

مثال۳) فرض کنید مشاهدات طولی دارای یک توزیع نرمال با $\mu = 1200m$ و $\sigma = 0.2m$ باشد چه نسبتی از مشاهدات دارای طول مشاهده شده:

الف- بیشتر از 1200.4m می‌باشند؟

$$z_\circ = \frac{x - \mu}{\sigma} = \frac{1200.4 - 1200}{0.2} = 2$$

$$p(x > x_\circ) = p(z > z_\circ) = p(z > 2) \overset{table\ 2}{=} 0.0228 = \%2.28$$

ب- کمتر از 1200.4m می‌باشند؟

$$p(x < x_\circ) = p(z < z_\circ) = p(z < 2)$$

$$= 1 - p(z > 2) \overset{table\ 2}{=} 1 - 0.0228 = 0.9772 = \%97.72$$

به خاطر اینکه $p(x = x_\circ) = 0$ لذا داریم: $p(x = 2) = 0$

پ- کمتر از 1200.2m و بیشتر از 1199.7m می‌باشند؟

$$z_1 = \frac{1199.7 - 1200}{0.2} = -1.5$$

$$z_2 = \frac{1200.2 - 1200}{0.2} = +1$$

$$p(x_1 < x < x_2) = p(z_1 < z < z_2) = p(-1.5 < z < 1)$$

لازم به یادآوری می‌باشد که به علت متقارن بودن منحنی نرمال استاندارد همواره داریم:

$$p(z < -z_\circ) = p(z > z_\circ)$$

$$p(z < -1.5) = p(z > 1.5) \overset{table\ 2}{=} 0.0668$$

$$p(z > 1) \overset{table\ 2}{=} 0.1587$$

$$p(-1.5 < z < 1) = 1 - [p(z > 1) + p(z < -1.5)] = 1 - [p(z > 1) + p(z > 1.5)]$$

$$\overset{table\ 2}{=} 1 - (0.1587 + 0.0668) = 0.7745 = \%77.45$$

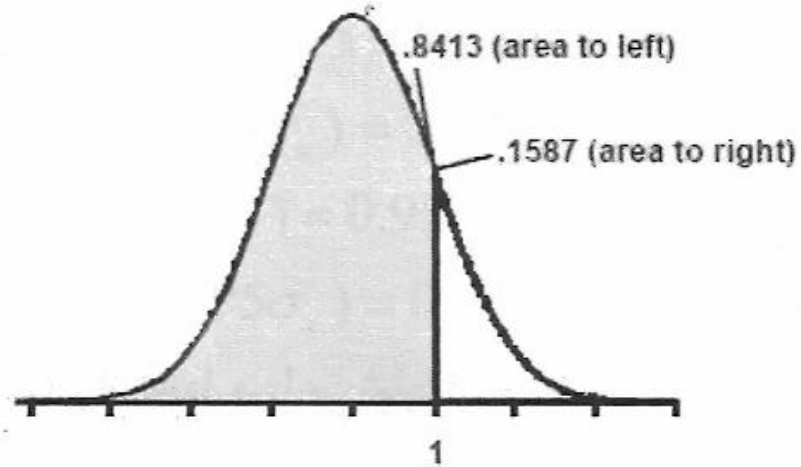
ت- حداقل مشاهدات جهت قرار گرفتن در بین 20.05% بالای منحنی (مشاهدات با طول بیشتر) چقدر است؟

$$p(z_\circ < z < +\infty) \overset{table\ 2}{=} 0.2005$$

از جدول شماره ۲ مقدار z_0 را محاسبه می‌کنیم:

$$z_\circ = 0.84$$

$$x_\circ = \mu + z \times \sigma = 1200 + (0.84 \times 0.2) = 1200.168m$$



پایان جلسه چهارم

درس تئوری خطاها

جلسه پنجم

فرید اسماعیلی
Farid_63@yahoo.com
www.faridesm.ir

تماس با استاد از طریق پست الکترونیکی
مشاهده اطلاعیه ها، نمرات، دریافت فایل ها در وب سایت

نمونه، مجموعه‌ای از مشاهدات داخل یک جامعه می‌باشد. توصیف کننده‌های عددی یک نمونه را آماره گویند. برای مثال میانگین $(\bar{x})$ و واریانس نمونه (s^2) آماره‌های یک نمونه هستند.

- یک مشخصه عددی از یک جامعه مثل میانگین یا انحراف معیار را یک **پارامتر جامعه**[1] یا به طور ساده یک پارامتر می‌نامند.
- کمیت های محاسبه شده از یک نمونه، مانند میانگین یا انحراف معیار آن را یک **آماره نمونه**[2] (برآورد نمونه) یا به طور ساده یک آماره یا برآورد می نامند.

درجه آزادی به صورت زیر تعریف می شود[2]:

$$df = n - u = \text{تعداد مشاهدات} - \text{تعداد پارامترهای برآورد شده}$$

n : تعداد مشاهدات

u: تعداد پارامترهای برآورد شده (تعداد مجهولات)

هرگاه $x_1, x_2, \ldots, x_n$ متغیرهای تصادفی مستقل از یک نمونه تصادفی به تعداد n از یک جامعه $N(\mu, \sigma^2)$ باشد در این صورت آماره t استیودنت با درجه آزادی m = n-1 به صورت زیر تعریف می‌شود:[6]

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

که در آن $\bar{x}$ (میانگین نمونه) و s (انحراف معیار نمونه) به صورت زیر تعریف می‌شوند:

$$\bar{x} = \frac{\sum x_i}{n}, \qquad s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

تابع چگالی احتمال t به صورت زیر می‌باشد که در آن Γ، تابع گاما می‌باشد.

$$f(t)=\frac{1}{\sqrt{(n-1)\pi}}\times\frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})}\times\frac{1}{(1+\frac{t^2}{n-1})^{\frac{n}{2}}}$$

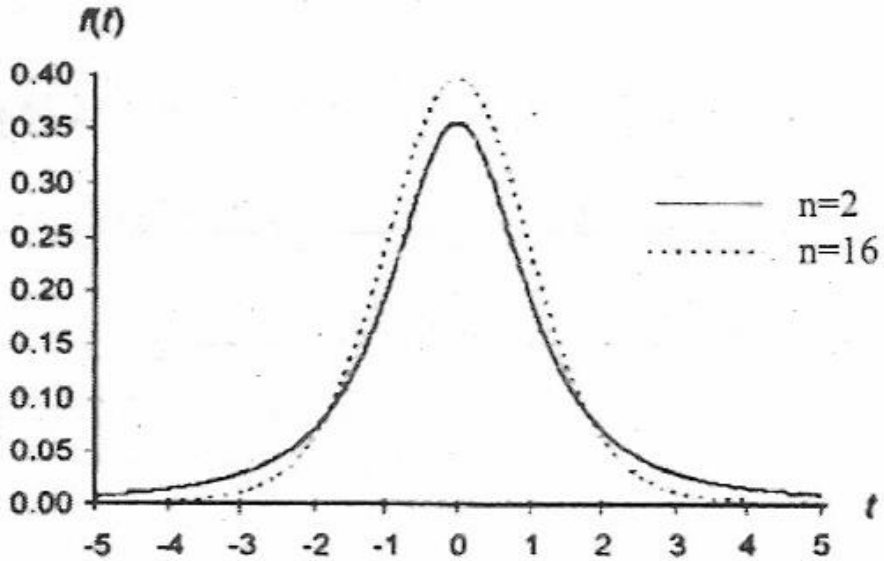
$$\Gamma(n)=\int_0^\infty e^{-t}t^{n-1}dt,\qquad n>0$$

$$\Gamma(n+1)=n\times\Gamma(n),\qquad \Gamma(\frac{1}{2})=\sqrt{\pi},\qquad \Gamma(1)=1$$

تابع تعریف شده دارای m = n-1 درجه آزادی بوده و نسبت به میانگین حالت متقارن داشته و وقتی درجه آزادی زیاد می‌شود ($n \geq 30$) توزیع t به توزیع نرمال با میانگین 0 و واریانس 1 (تابع توزیع نرمال استاندارد) نزدیک می‌شود.

خواص توزیع t استیودنت عبارتند از:

الف- شکل توزیع t استیودنت به درجه آزادی آن بستگی دارد.



شکل ۳-۱۹: تابع چگالی احتمال t با درجه آزادی ۲ و ۱۶

ب- شکل توزیع t استیودنت مشابه شکل توزیع نرمال است با این تفاوت که منحنی توزیع t با کاهش درجه آزادی در وسط پهن‌تر شده و به طرف انتهاها گسترده‌تر می‌شود.

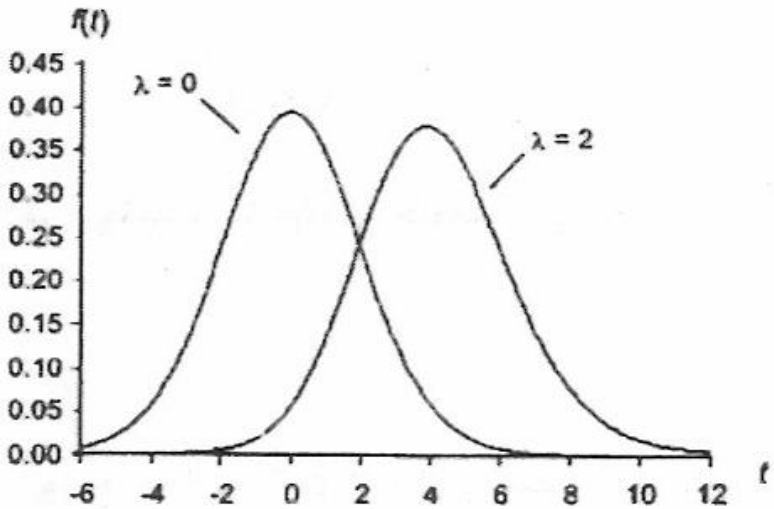
پ- انحراف معیار توزیع t استیودنت بزرگتر از یک است.

ت- وقتی درجه آزادی زیاد می‌شود $(n \geq 30)$ توزیع t استیودنت به توزیع نرمال با میانگین 0 و واریانس 1 نزدیک می‌شود.

ث- به طور خلاصه شرایط توزیع t استیودنت به شرح زیر می‌باشد:

- نمونه تصادفی کوچک است. $(n < 30)$
- واریانس جامعه مجهول است.
- جامعه اصلی، نرمال است.

توزیع t مرکزی و غیرمرکزی در شکل زیر ارائه شده است.



شکل ۳-۲۱: توزیع t مرکزی $(\lambda = 0)$ و غیرمرکزی $(\lambda = 2)$ با درجه آزادی ۲۰

کاربرد تابع t استیودنت عبارتند از:[6]

الف- آزمون t برای معنی‌دار بودن میانگین جامعه، وقتی واریانس معلوم نیست.

ب- آزمون t برای معنی‌دار بودن اختلاف بین دو نمونه وقتی که واریانسهای جامعه‌هایی که از آنها نمونه‌ها انتخاب شده‌اند مساوی ولی معلوم نیست.

پ- آزمون t برای معنی دار بودن ضریب همبستگی یک نمونه مشاهده شده

هرگاه $x_1, x_2, \ldots, x_n$ متغیرهای تصادفی مستقل دارای توزیع نرمال با میانگین 0 و واریانس 1 باشد آنگاه مجموع مربعات این متغیرهای تصادفی یک متغیر تصادفی دیگری را به صورت زیر تعریف می کند:

$$X^2 = x_1^2 + x_2^2 + \ldots + x_n^2$$

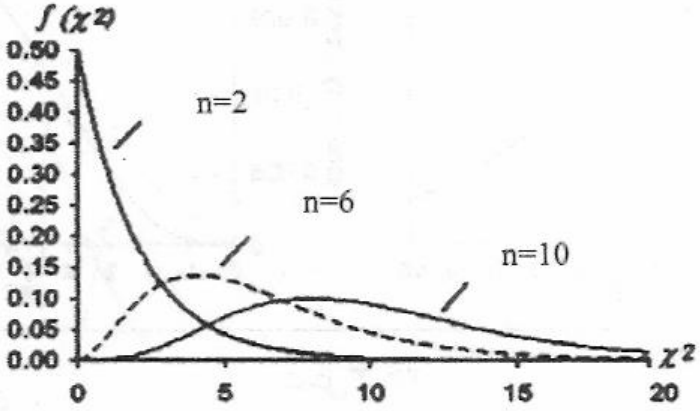
متغیر تصادفی نیز دارای n درجه آزادی است. تابع چگالی احتمال χ^2 به صورت زیر تعریف می‌شود:

$$f(x) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}} \times \Gamma(\frac{n}{2})} \times x^{\frac{n-2}{2}} \times e^{-\frac{x}{2}} & x > 0 \\ 0 & x \le 0 \end{cases}$$

وقتی درجه آزادی توزیع کای اسکور از 10 بیشتر می‌شود، آنگاه توزیع کای اسکور به توزیع نرمال نزدیک می‌شود و وقتی $n \to \infty$ رفتار آن همانند توزیع نرمال با میانگین n و واریانس $2n$ خواهد بود.

خواص توزیع کای اسکور عبارتند از:

الف- شکل توزیع کای اسکور به درجه آزادی آن بستگی دارد.



شکل ۳-۲۲: تابع چگالی متغیر کای اسکور با درجه آزادی ۲ و ۶ و ۱۰

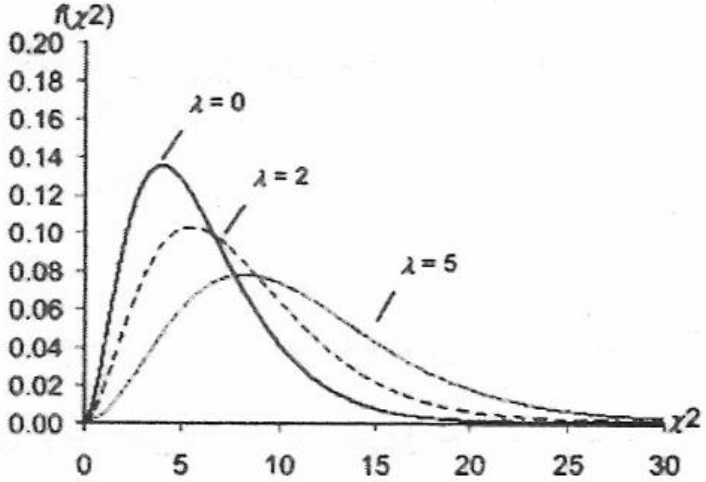
ب- این توزیع مثل توزیع نرمال متقارن نیست.

پ- مقادیر این تابع نمی‌تواند منفی باشد.

امید ریاضی متغیر کای اسکور غیر مرکزی برابر است با:

$$E[\chi^2] = n + \lambda$$

با مقایسه توزیع کای اسکور مرکزی و غیرمرکزی بدلیل وجود پارامتر غیرمرکزی λ، میانگین به طرف راست متمایل می‌شود.



شکل ۳-۲۳: تابع چگالی متغیر کای اسکور مرکزی $(\lambda = 0)$ و غیر مرکزی $(\lambda = 2,5)$ با درجه آزادی ۶

کاربردهای توزیع کای اسکور

کاربردهای توزیع کای اسکور عبارتند از:[6]

الف- آزمون کای اسکور برای خوبی برازش[1] (حسن تطابق)

ب- برای آزمون این که آیا جامعه دارای واریانس σ^2 است یا نه؟

هرگاه $x_1, x_2, \ldots, x_n$ و $y_1, y_2, \ldots, y_n$ دو دسته از متغیرهای مستقل با توزیع نرمال استاندارد یعنی دارای میانگین صفر و واریانس یک باشند. دو متغیر تصادفی X_n^2 و X_m^2 را به صورت مجموع مربعات متغیرها به صورت زیر تعریف می‌کنیم: (اگر X_m دارای توزیع کای اسکور با m درجه آزادی و اگر X_n دارای توزیع کای اسکور با n درجه آزادی باشد)

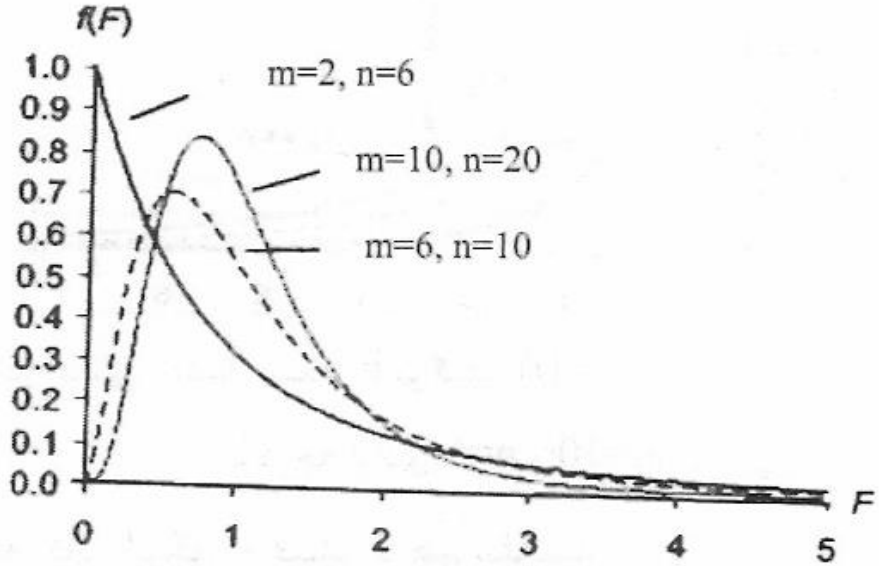
$$X_n^2 = x_1^2 + x_2^2 + \ldots + x_n^2$$

$$X_m^2 = y_1^2 + y_2^2 + \ldots + y_n^2$$

سپس متغیر تصادفی F را به صورت زیر تعریف می‌کنیم: (یعنی، این توزیع در رابطه با نمونه‌گیری از جامعه‌های نرمال می باشد. این توزیع به عنوان توزیع نمونه‌ای نسبت دو متغیر تصادفی مستقل با توان دوم کای اسکور که بر درجات آزادی خودشان تقسیم شده‌اند، تعریف می‌شود.)

$$F_{m,n} = \frac{\frac{X_m^2}{m}}{\frac{X_n^2}{n}}$$

توزیع F دارای n و m درجه آزادی است. لازم به ذکر است که شکل توزیع F به درجه آزادی آن بستگی دارد.



شکل ۳-۲۵: تابع چگالی احتمال متغیر F با درجه آزادی مختلف

تابع چگالی احتمال $f_{n,m}(u)$ به صورت زیر تعریف می‌شود:

$$f_{m,n}(u) = \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})} \times \left(\frac{m}{n}\right)^{\frac{m}{2}} \times \frac{u^{\frac{m-2}{2}}}{(1+\frac{m}{n}\times u)^{\frac{m+n}{2}}}, \quad u > 0$$

میانگین و واریانس توزیع فیشر با m و n درجه آزادی به صورت زیر می‌باشد:

$$E\left(f_{m,n}\right)=\frac{n}{n-2}, \qquad n>2$$

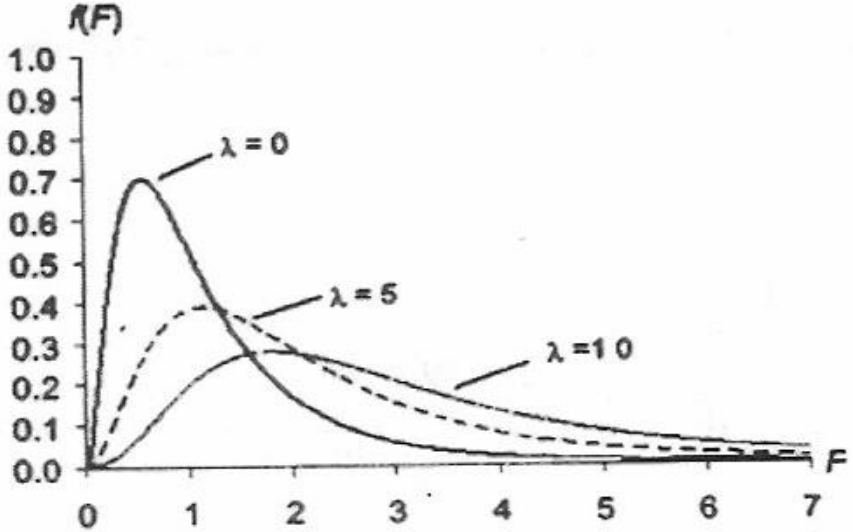
$$Var(f_{m,n})=\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}, \qquad n>4$$

تابع توزیع $f_{n,m}$ با m و n درجه آزادی وقتی $n\to\infty$، آنگاه توزیع فیشر مانند توزیع کای اسکور با m درجه آزادی است.

امید ریاضی تابع توزیع $f_{m,n}$ غیر مرکزی برابر است با:

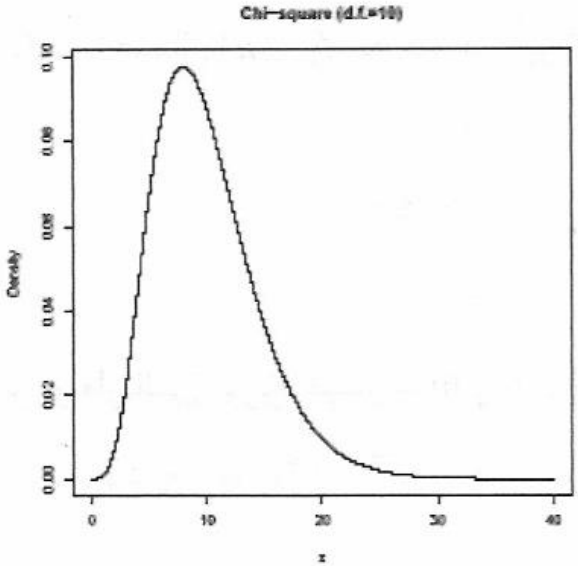
$$E[f_{m,n}]=\frac{n}{n-2}[1+\frac{2\lambda}{m}]$$

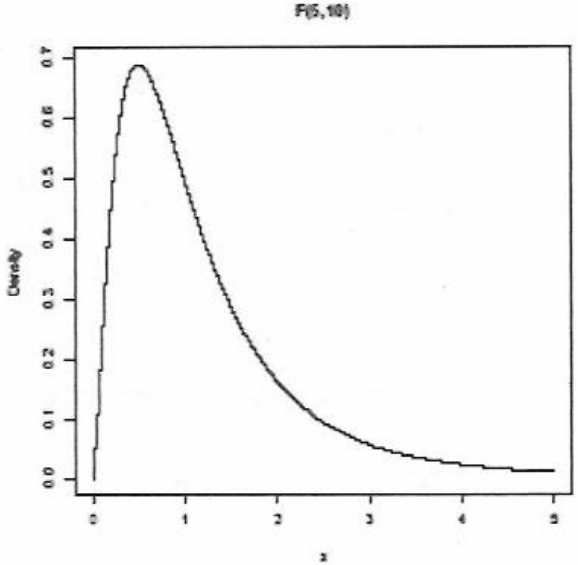
با مقایسه توزیع F مرکزی و غیرمرکزی بدلیل وجود پارامتر غیرمرکزی λ، میانگین بـه طـرف راست متمایل می‌شود.



شکل ۳-۲۶: تابع چگالی احتمال متغیر F مرکزی $(\lambda = 0)$ و غیر مرکزی $(\lambda = 5,10)$ با درجه آزادی n=6 و m=10

در دو شکل زیر دو تابع کای اسکور و فیشر با هم مقایسه شده‌اند.





کاربردهای توزیع فیشر

کاربردهای توزیع فیشر عبارتند از:[6]

الف- آزمون F برای مساوی بودن واریانس‌های دو جامعه

ب- آزمون F برای معنی‌دار بودن ضریب همبستگی چندگانه یک نمونه مشاهده شده

پ- آزمون F برای معنی‌دار بودن نسبت ضریب همبستگی یک نمونه مشاهده شده

ت- آزمون F برای معنی‌دار بودن رگرسیون خطی

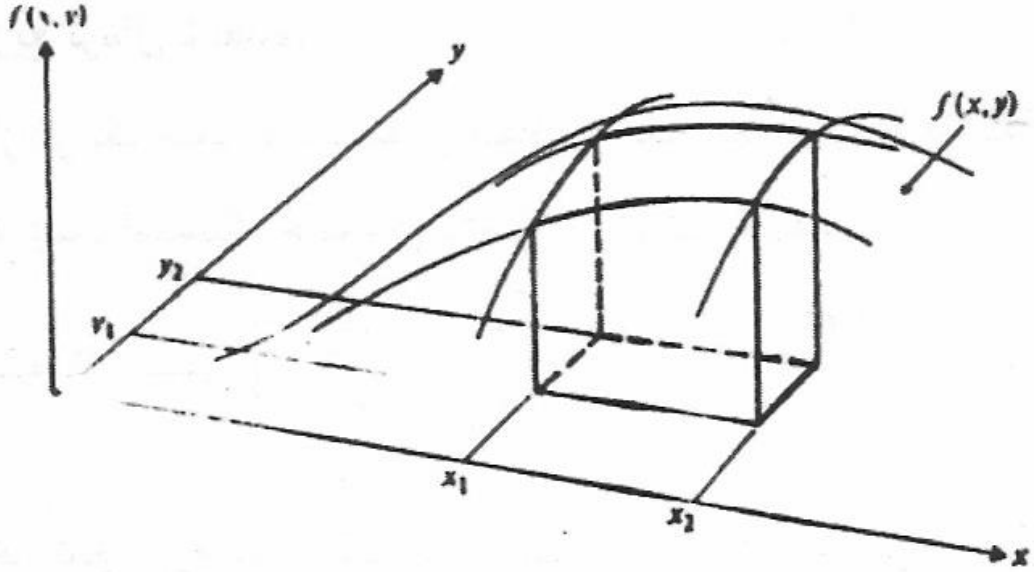
ث- آزمون F برای معنی‌دار بودن تساوی میانگین‌های چند جامعه

اگر برای هر اندازه از یک متغیر x یک مقدار متناظر از یک متغیر ثانی Y داشته باشیم، سری داده‌ها دوتائی به دست آمده یک جامعه دو متغیره[1] را تشکیل می‌دهد.

اگر داشته باشیم: $\sum_{xx} = \begin{bmatrix} \sigma_{x_1x_1} & \sigma_{x_1x_2} \\ \sigma_{x_2x_1} & \sigma_{x_2x_2} \end{bmatrix} = \begin{bmatrix} \sigma^2_{x_1} & \sigma_{x_1x_2} \\ \sigma_{x_2x_1} & \sigma^2_{x_2} \end{bmatrix}$ ، $\mu = \begin{bmatrix} \mu_{x_1} \\ \mu_{x_2} \end{bmatrix}$ و $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ رابطه زیر برای تابع چگالی بدست می‌آید:

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\sigma^2_{x_1}\sigma^2_{x_2} - \sigma^2_{x_{12}}}} e^{\left\{-\frac{\sigma^2_{x_1}\sigma^2_{x_2}}{2(\sigma^2_{x_1}\sigma^2_{x_2} - \sigma^2_{x_{12}})}\left[(\frac{x-\mu_{x_1}}{\sigma_{x_1}})^2 - 2\sigma_{12}\frac{(x-\mu_{x_1})(y-\mu_{x_2})}{\sigma^2_{x_1}\sigma^2_{x_2}} + (\frac{y-\mu_{x_2}}{\sigma_{x_2}})^2\right]\right\}}$$

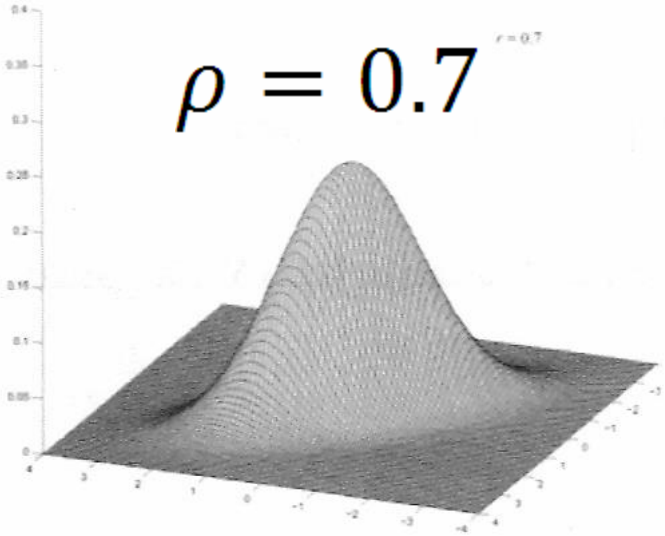
که در آن μ_{x_1} ، μ_{x_2} ، $\sigma^2_{x_1}$ ، $\sigma^2_{x_2}$ ، به ترتیب میانگین‌ها و واریانسهای دو متغیر تصادفی x_1 و x_2 و $\sigma_{x_1x_2}$ نیز کواریانس دو متغیر تصادفی x_1 و x_2 می‌باشد.

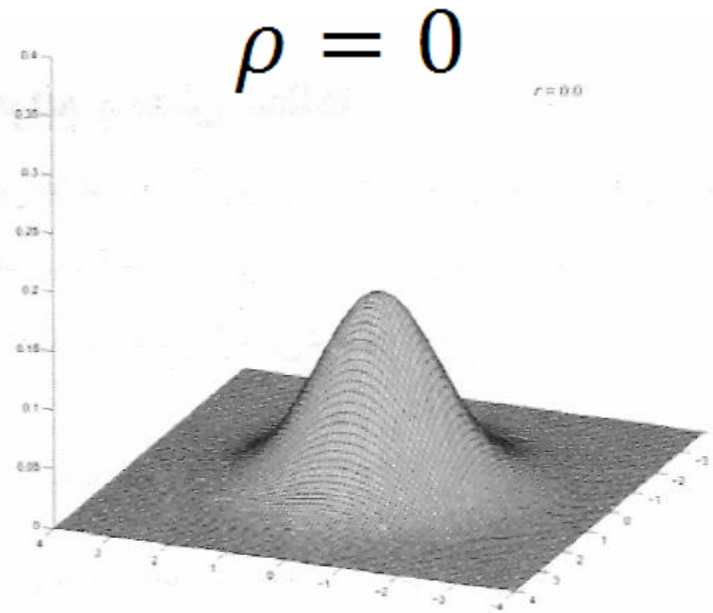


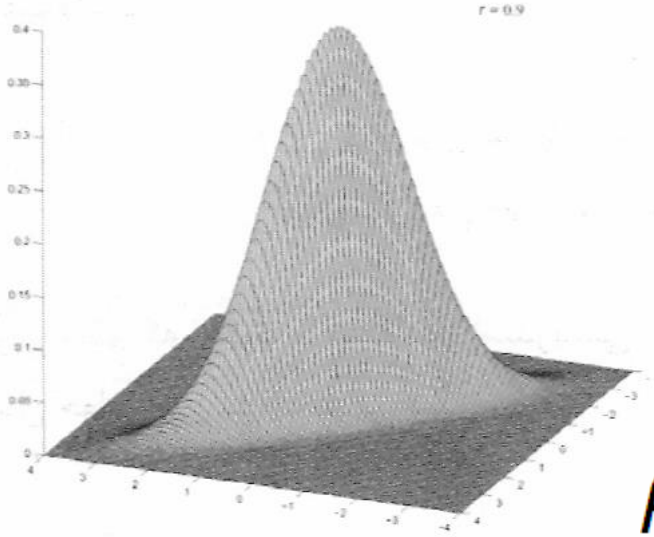
شکل ۳-۱۹: تابع توزیع نرمال دو متغیره

توزیع نرمال دو بعدی شبیه به زنگ می‌باشد. تقاطع حاصل این شکل به صفحات قائم یک تـابع چگالی نرمال یک بعدی می‌باشد.

همچنین هر صفحه موازی صفحه y، x شکل زنگوله‌ای تابع را در یک منحنی بیضی شکل قطع می‌کند.

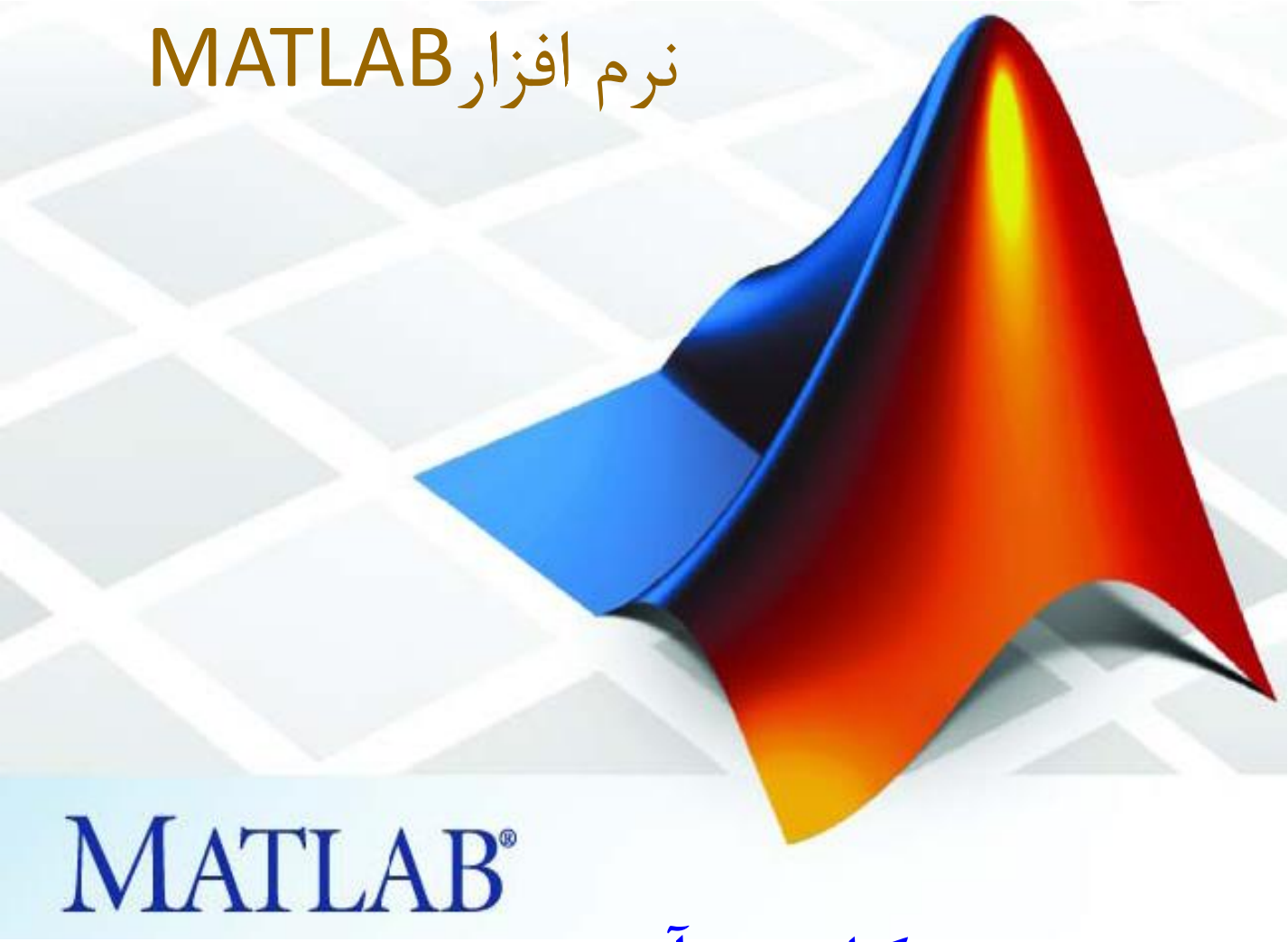






فرید | **نمایش هندسی تابع چگالی نرمال دو متغیره برای ضریب همبستگی ۰، ۰/۷، و ۰/۹**

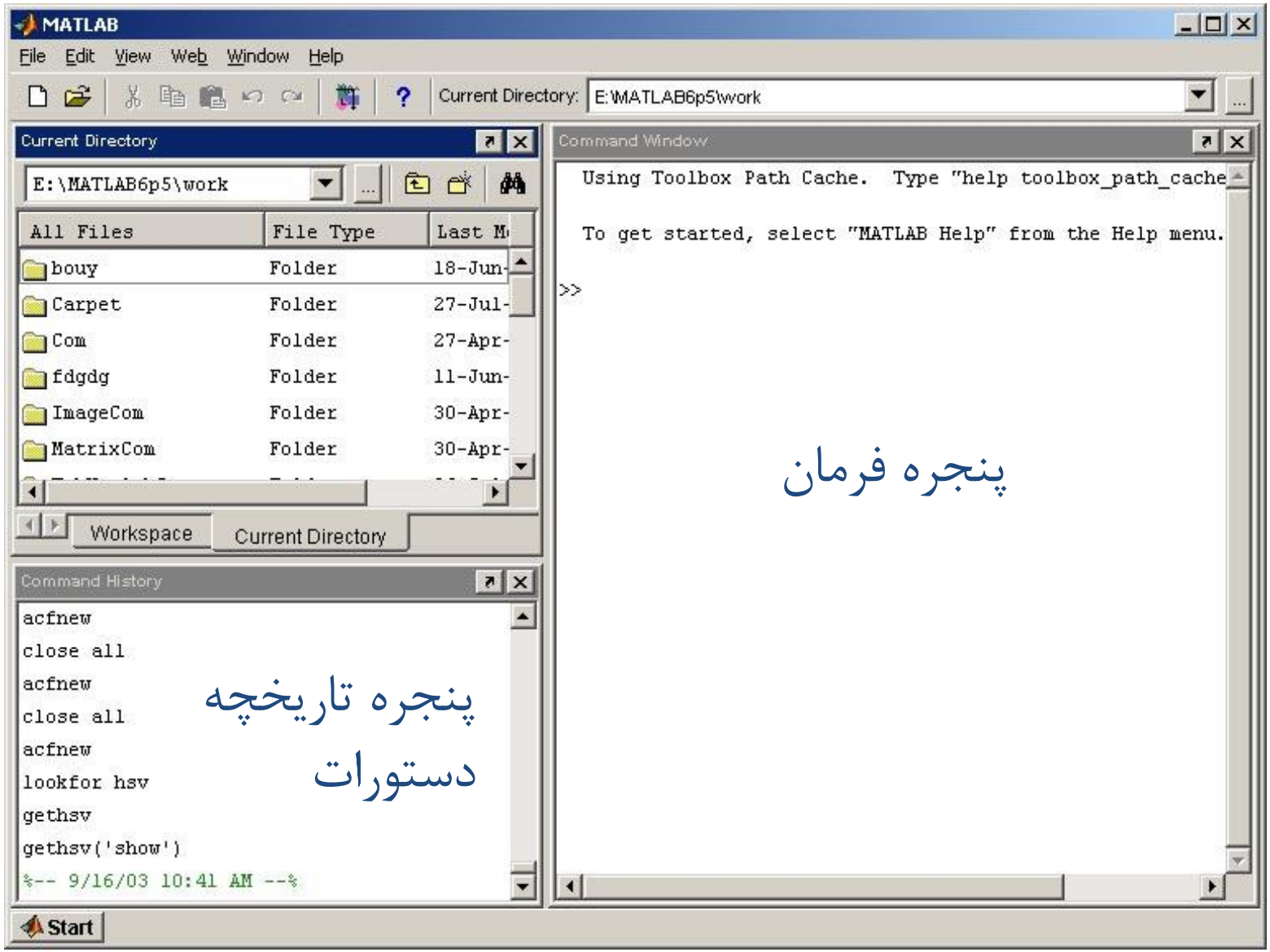
نرم افزار MATLAB



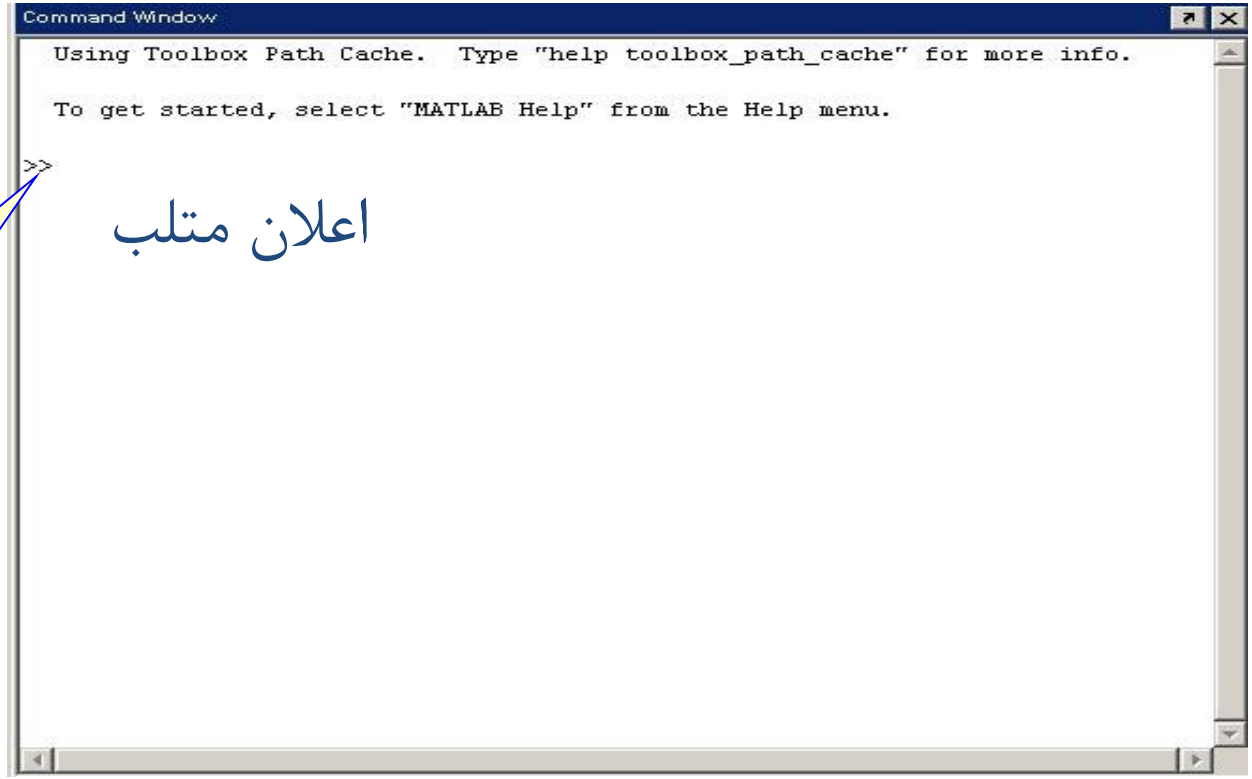
کلاس آموزشی

آشنایی با محیط متلب

- پنجره فرمان : Command window
- پنجره تاریخچه دستورات: Command History
- پنجره دایرکتوری جاری : Current Directory
- پنجره فضای کاری : Work Space
- دایرکتوری جاری
- منوی Start

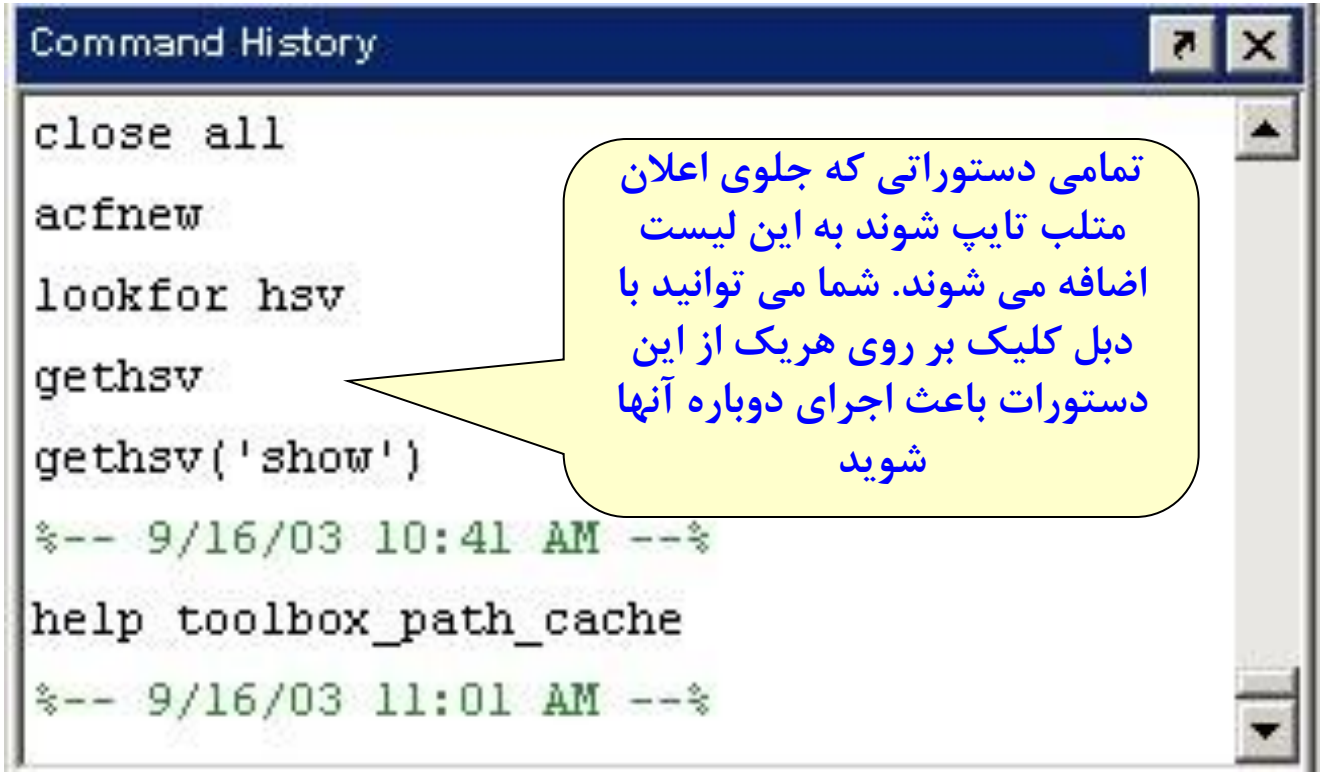

MATLAB
File Edit View Web Window Help
Current Directory: E:\MATLAB6p5\work
Current Directory
E:\MATLAB6p5\work
All Files File Type Last M
bouy Folder 18-Jun-
Carpet Folder 27-Jul-
Com Folder 27-Apr-
fdgdg Folder 11-Jun-
ImageCom Folder 30-Apr-
MatrixCom Folder 30-Apr-
Workspace
Current Directory
Command History
acfnew
close all
acfnew
close all
acfnew
lookfor hsv
gethsv
gethsv('show')
%-- 9/16/03 10:41 AM --%
Command Window
Using Toolbox Path Cache. Type "help toolbox_path_cache
To get started, select "MATLAB Help" from the Help menu.
>>
پنجره فرمان
پنجره تاریخچه
دستورات
Start

پنجره فرمان : Command Window



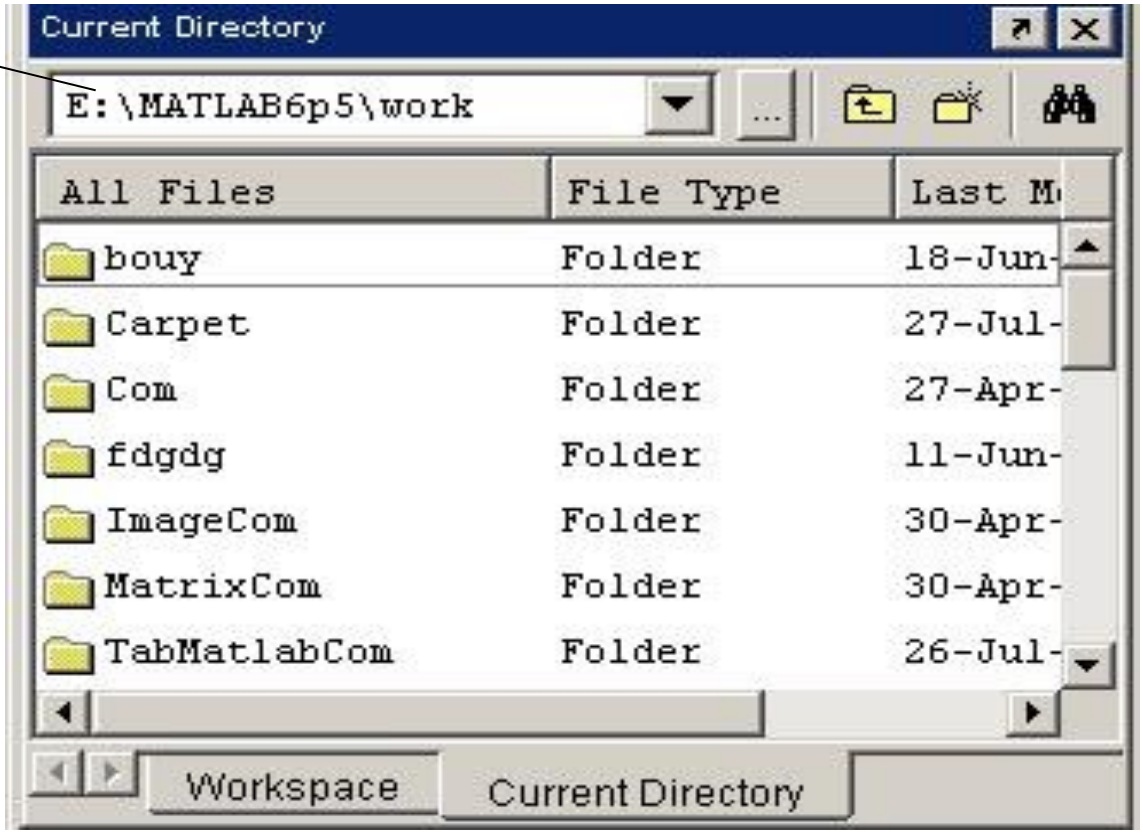
فرامین متلب را در
جلوی اعلان متلب
تایپ کنید

پنجره تاریخچه دستورات: Command History



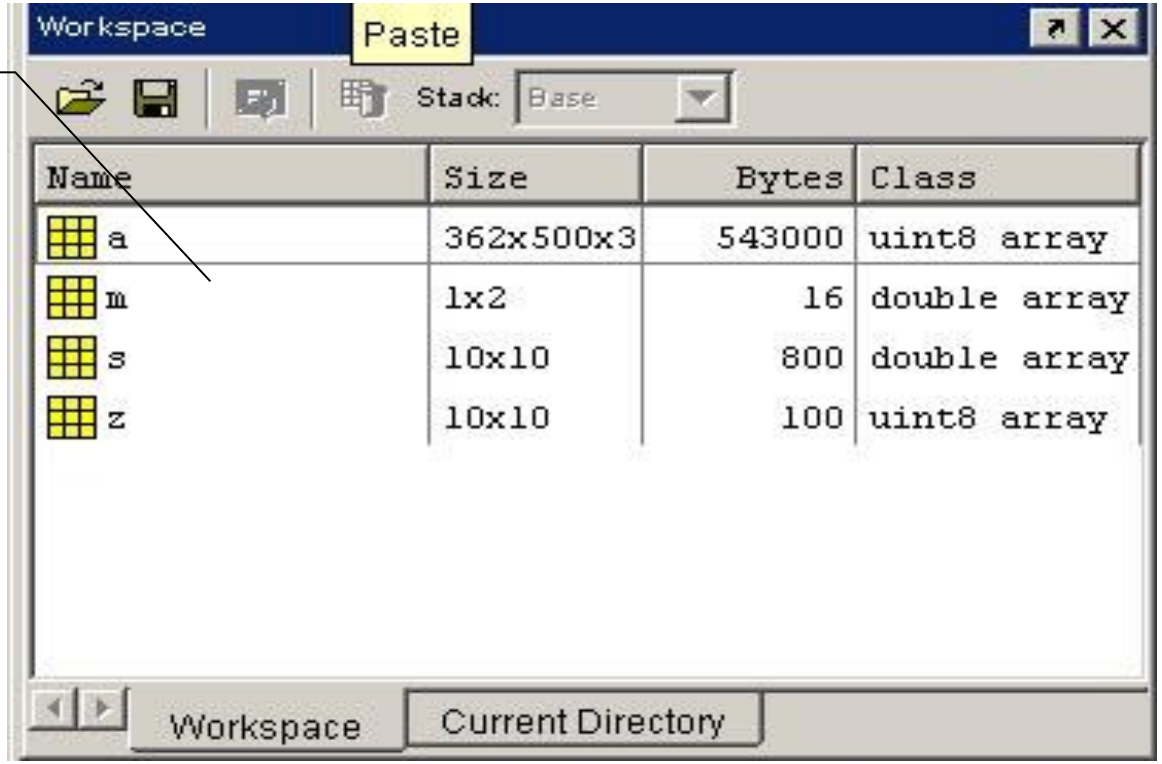
پنجره دایرکتوری جاری : Current Directory

در هر زمان تنها یک دایرکتوری یا پوشه به عنوان دایرکتوری جاری در متلب شناخته می شود. هر فایل متلب(برنامه نوشته شده توسط شما) که نام آن جلوی اعلان متلب تایپ شود تنها درصورتی اجرا می شود که در دایرکتوری جاری یا در مسیر متلب باشد



فضای کاری : Work Space

متغیرهایی که در حال حاضر در محیط کاری متلب وجود دارند و شما می توانید از مقادیر آنها استفاده کنید یا آنها را تغییر دهید



عملیات ریاضی ساده

مثال: محاسبه یک عبارت:

راه اول:

```
>> 4*25 + 6*22 + 2*99
   ans=
                430
```

عملیات ریاضی ساده

مثال: محاسبه یک عبارت:

راه دوم:

```
>>a=25;
>>b=22; c=99;
>>d=4*a+6*b+2*c
   d=
        430
>>
```

نکته ۱: علائم ;

نکته ۲: تعریف متغیرها

عملگرهای ریاضی متلب:

\ / , * , - , + , ^

مثال:

```
>>5^2
 ans=
        25
```

/ و \ هر دو عملگر تقسیم میباشند. / تقسیم از چپ و \ تقسیم از راست است. مثلا حاصل 56/8 و 8\56 یکسان است.

- ترتیب حق تقدم: - + < * \ / < ^

زمان اعتبار متغیرها

متغیرهایی که در فضای کاری تعریف می شوند تنها در دو حالت زیر از حافظه پاک خواهند شد:

- خروج متلب
- استفاده از دستور clear :

>> clear تمامی متغیرها از حافظه پاک می شوند

>> clear a b c تنها متغیرهای نامبرده شده از حافظه پاک می شوند

فایلهای متنی(Script) یا فایلهای m

بمنظور اجرای چند دستور بطور همزمان و بدون نیاز به تایپ مجدد، از فایلهای متنی استفاده می‌شود.

این فایلها باید دارای پسوند m باشند.

مراحل ایجاد فایلهای متنی

1. باز کردن یک فایل جدید در ویرایشگر متلب:

File>New>m-file

2. تایپ کردن دستورات متلب در فایل مذکور
3. ذخیره کردن فایل با نامی مشخص:

File>Save As…

روش اجرای یک فایل متنی

برای اجرای یک فایل متنی کافی است نام آنرا در جلوی اعلان متلب تایپ کرده کلید Enter را بزنیم.

مثال: برنامه sample1.m

```
% SAMPLE1: A Simple m-file
n=10;a=2;b=4;
c=n*a^3/b + 3*n*a^2/b^2+6*n*a/b^3
```

```
>> sample1
   c=
        29.3750
```

ایجاد ماتریس با استفاده از علائم ; ، , و [ ]

از علامت ; برای تعیین سطر جدید و از علامت , برای تعیین ستون جدید استفاده می‌شود.

مثال:

```
>> a=[1,2,3;4,5,6]
a=
    1 2 3
    4 5 6
>> b=[1,2,3,4,5,6]
b=
    1 2 3 4 5 6
```

ایجاد ماتریس با استفاده از علائم ; ، , و [ ]

نکته: بجای علامت ; از enter و بجای علامت , از فاصله خالی نیز می‌توان استفاده کرد

مثال:

```
>> c=[1 2,3
        4 5 6;7 8,9]
c=
    1   2   3
    4   5   6
    7   8   9
```

ماتریسهای ویژه

- [ ] : ماتریس تهی
- eye : یک ماتریس یکه با ابعاد داده شده ایجاد می‌کند
- ones : یک ماتریس که تمامی عناصر آن یک می‌باشند با ابعاد داده شده ایجاد می‌کند
- zeros : یک ماتریس صفر با ابعاد داده شده ایجاد می‌کند
- rand : یک ماتریس با عناصر راندوم با توزیع یکنواخت به ابعاد داده شده ایجاد می‌کند
- randn : یک ماتریس با عناصر راندوم با توزیع نرمال به ابعاد داده شده ایجاد می‌کند

ماتریسهای ویژه

مثال:

```
>>ones(2,3)
ans =

   1   1   1
   1   1   1
>>ones(2)
ans =

   1   1
   1   1
```

```
>>A=3
A=
3

>>A=3;
>>
```

```
>>[1 2]+[3 4]

Ans=
3  6

>>[1 2 3;4 5 6]+4


Ans=
5 6 7
8 9 10
```

```
>>A=[1 2 3];
>>B=[1;2;3];

>>A*B

Ans=
14
>>B*A

Ans=

1 2 3
2 4 6
3 6 9
```

```
>>A=[1 2 3 4]

>>A'

Ans=
1
2
3
4
```

```
>>A=[1 2 3 4]

>>a.*a

ans=
1  4  9  16
```

عدد پی Pi

ریشه دوم sqrt

توان power

```
>>power(2,3)

Ans=
8
```

تابع نمایی exp

log همان لگاریتم طبیعی یا بر مبنای e است .

log2 لگاریتم بر مبنای دو است .

log10 لگاریتم بر مبنای ده است .

```
>>A=input('please enter the number :')
Please enter the number
123

A=
123
```

پاک کردن صفحه نمایش clc

پاک کردن متغیر clear

format bank	دو رقم اعشار	3.14
format short	چهار رقم اعشار	3.1416
format short e	چهار رقم اعشار ممیز شنا	3.1416e+000
format short g	بهترین حالت short	3.1416
format long	پانزده رقم اعشار	3.14159265358979
format long e	پانزده رقم اعشار ممیز شناور	3.141592653589793 e+000
format long g	بهترین حالت long	3.14159265358979
format hex	بر اساس هگزا دسیمال	400921 fb54442d18
format rat	بصورت کسر منطقی	355 / 113
format +	بر اساس تابع علامت sign	+

گرد کردن

تابع	عملکرد تابع	عملکرد برای 2.4	عملکرد برای-2.4
fix	عدد را به سمت صفر گرد می کند	2	-2
floor	عدد را به سمت منفی بینهایت گرد می کند	2	-3
ceil	عدد را به سمت مثبت بینهایت گرد می کند	3	-2
round	عدد را به سمت نز دیکترین همسایگی گرد	2	-2

توابع sin cos tan cot مقادیر سینوس کسینوس تانژانت و کتانژانت مقدار را می دهد .

توابع asin acos atan acot مقادیر معکوس توابع بالا را می دهد .

توابع sinh cosh tanh coth مقادیر هیپربولیک توابع بالا را می دهد .

توابع asinh acosh atanh acoth مقادیر معکوس توابع هیپربولیک را میدهد .

توابع آرایه ای

```
>>a=[1 2 3;4 5 6];
>>numel(a)
```

Ans=
6

تعداد اعضای آرایه

```
>>a=[1 2 3];
>>length(a)
```

Ans=
3

طول بردار

```
>>s=[1 2 3;4 5 6];
>>size(s)
```

Ans=
2 3

اندازه گیری مرتبه ماتریس

```
>>eye(3)
```

ماتریس همانی

Ans=
1 0 0
0 1 0
0 0 1

```
>>zeros(2)
```

ماتریس صفر

ans=
0 0
0 0

```
>>ones(3)
```

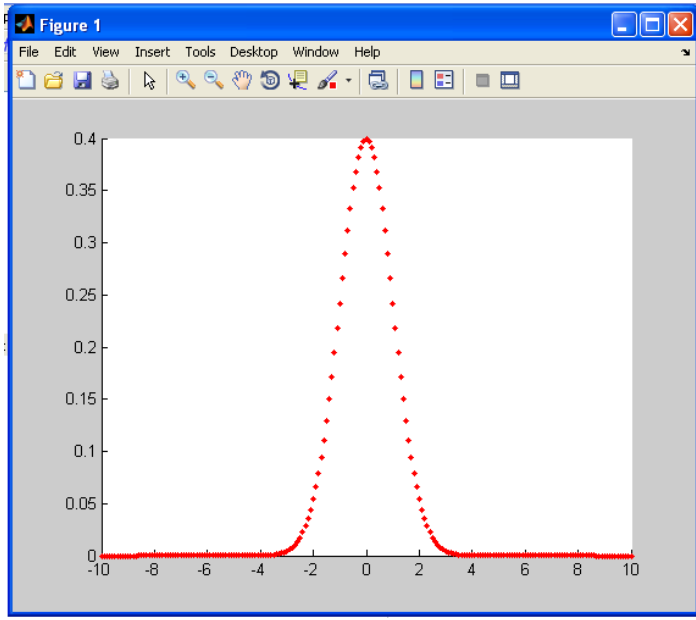
ماتریس یک

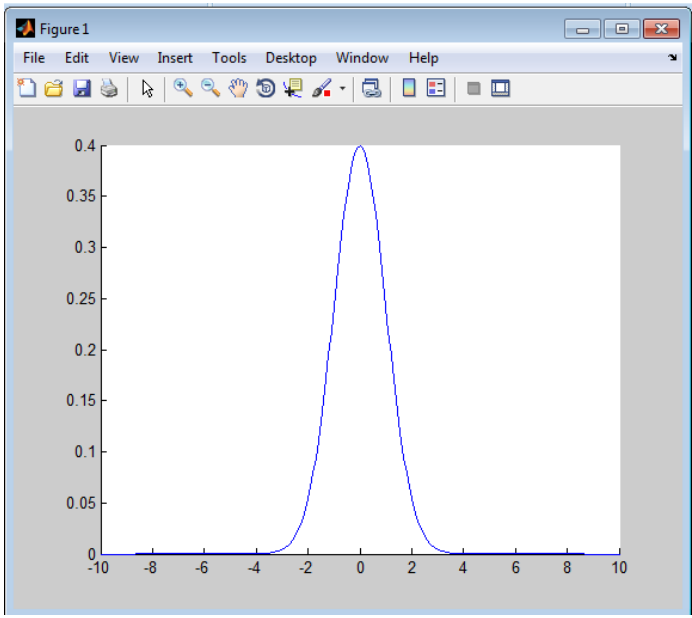
Ans=
1 1 1
1 1 1
1 1 1

```
>> a=[2 3 -4;0 -4 2;1 -1 5]
a =
    2    3   -4
    0   -4    2
    1   -1    5

>> inv(a)
ans =
    0.3913    0.2391    0.2174
   -0.0435   -0.3043    0.0870
   -0.0870   -0.1087    0.1739
```

برنامه ترسیم تابع چگالی نرمال

```
clc
clear all
close all

n=10;
x=[-n:0.1:n];
d=20*n+1;
s=1; %enheraf meyar
m=0; %miyangin

for i=1:d
  y(i)=(1/(s*((2*pi)^.5)))*(exp((-(x(i)-m)^2)/(2*s^2)));
 hold on
 plot(x(i),y(i), 'r.')
End
plot(x,y)
```





پایان جلسه پنجم

درس تئوری خطاها

جلسه ششم

فرید اسماعیلی
Farid_63@yahoo.com
www.faridesm.ir

تماس با استاد از طریق پست الکترونیکی
مشاهده اطلاعیه ها، نمرات، دریافت فایل ها در وب سایت

تابع plot

شکل کلی:

plot (x1,y1,’c1s1’,x2,y2,’c2s2,x3,y3,’c3s3’,…)

در این رابطه، sn می‌تواند هر یک از کاراکترهای زیر باشد:

. , o , x , + , - , * , -. , -- , penta , hexa

و cn نیز می‌تواند یکی از رنگهای زیر باشد:

y , m , c , r , g , b , w , k

که به ترتیب معرف رنگهای زرد، سرخابی، فیروزه‌ای، قرمز، سبز، آبی، سفید و سیاه می‌باشد

برچسب گذاری محورهای افقی و عمودی و عنوان

بمنظور برچسب‌گذاری محورها و ایجاد عنوان برای نمودار می‌توان از توابع xlabel, ylabel , title استفاده کرد.

```
>> xlabel('یک رشته متنی');
>> ylabel('یک رشته متنی');
>> title('یک رشته متنی');
```

این دستورات بر روی آخرین نمودار ترسیم شده اعمال میشوند بنابراین بعد از هر دستور plot یا دستور ترسیمی دیگر بلافاصله باید از این دستورات استفاده گردد.

رسم خطوط شبکه‌ای بر روی نمودار

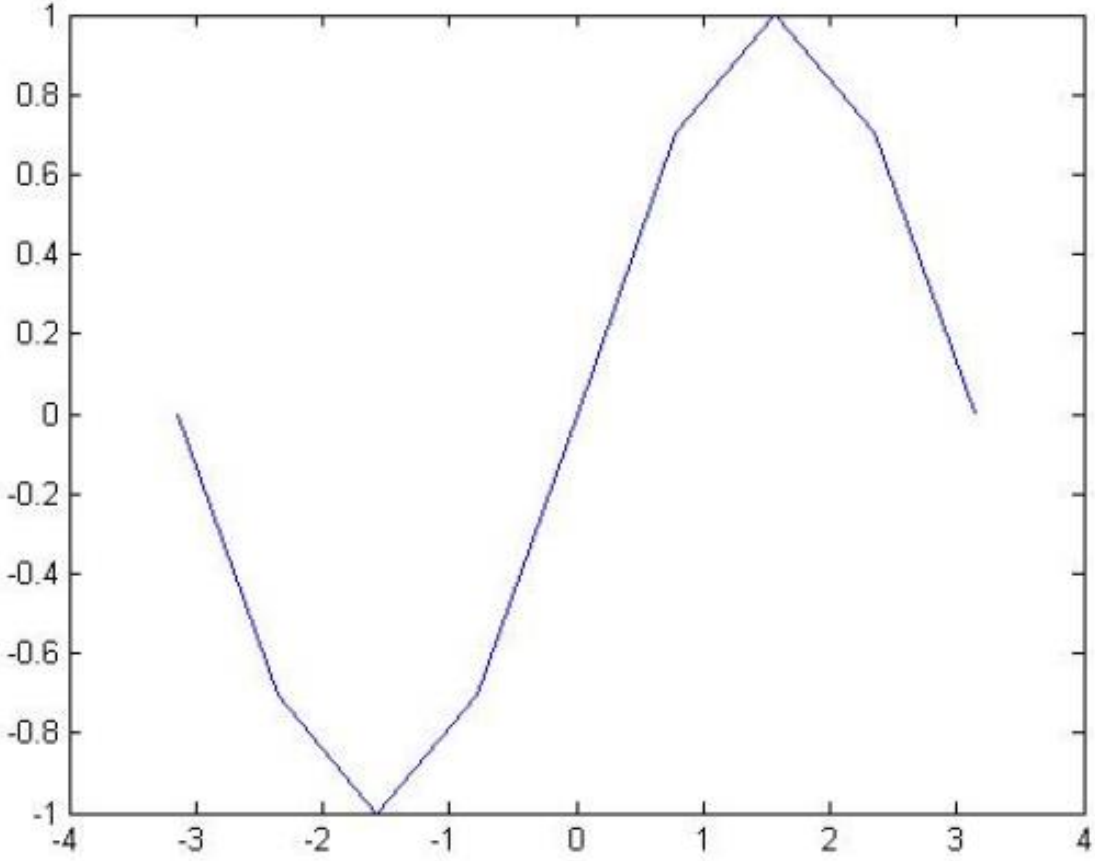
بمنظور ایجاد خطوط شبکه‌ای(چهارخانه‌های نقطه‌چین) بر روی یک نمودار، می‌توان از دستور grid استفاده کرد. شکل کلی استفاده از دستور grid بصورتهای زیر است:

```
>> grid on
```
حالت شبکه‌ای را فعال می‌کند

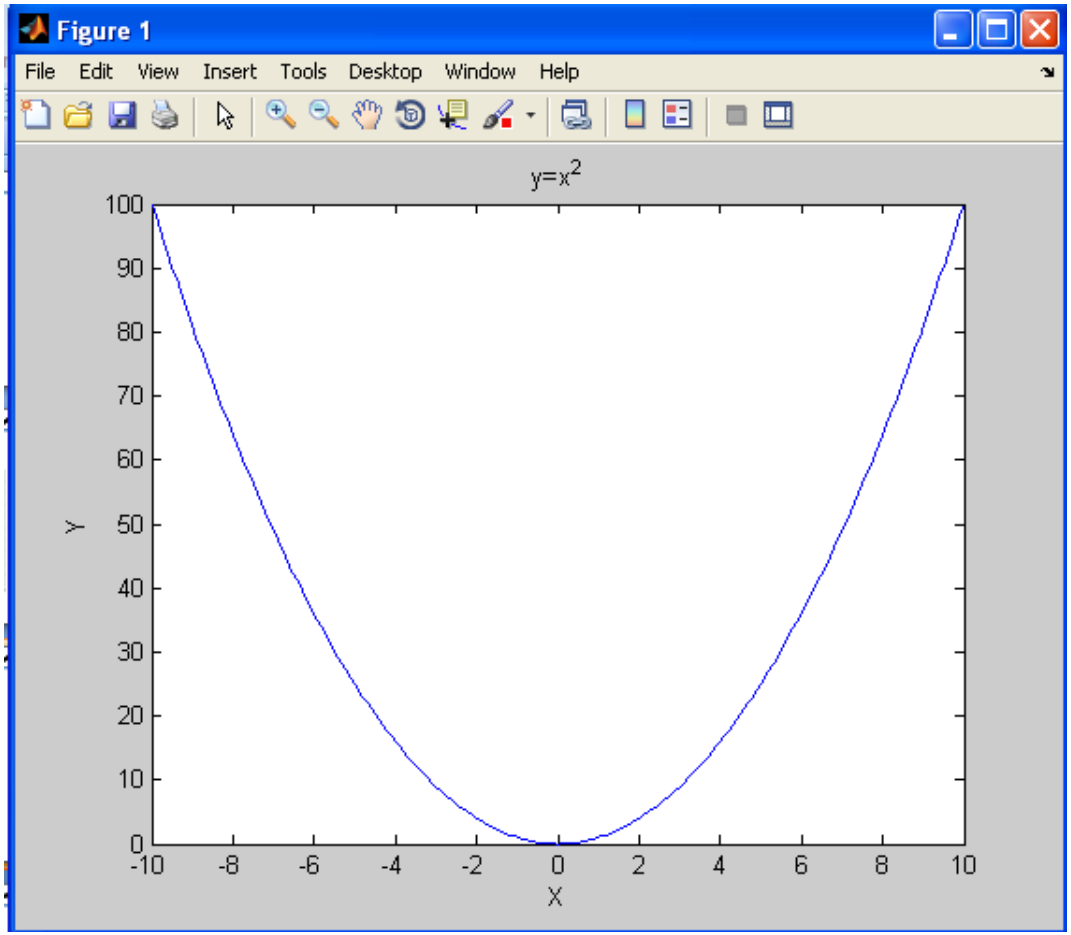
```
>> grid off
```
حالت شبکه‌ای را غیر فعال میکند

```
>> grid
```
حالت شبکه‌ای را از فعال به غیرفعال و از غیر فعال به فعال تغییر می‌دهد

```
>>x=-pi:pi/4:pi;
>>y=sin(x);
>>plot(x,y)
```

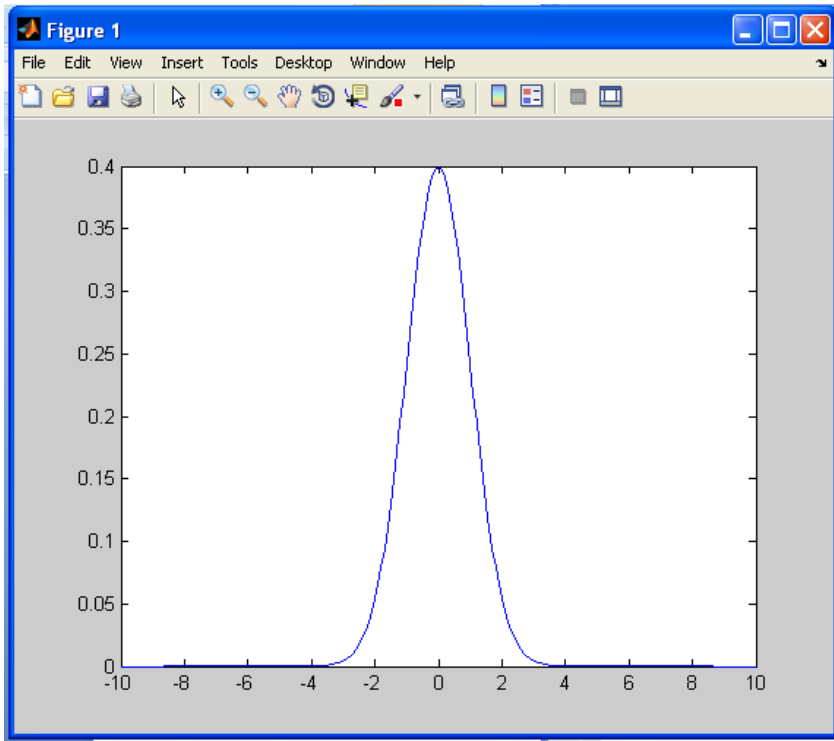


```
clc
clear all
close all
x=[-10:0.1:10];
y=x.^2;
plot(x,y);
title('y=x^2')
xlabel('X')
ylabel('Y')
```



ترسیم تابع توزیع نرمال گوس

```
clc
clear all
close all
n=10;
x=[-n:0.1:n];
d=20*n+1;
s=1; %enheraf meyar
m=0; %miyangin
y=(1/(s*((2*pi)^.5)))*(exp((-(x-m).^2)/(2*s^2)));
plot(x,y)
```



خمهای فضایی- تابع plot3

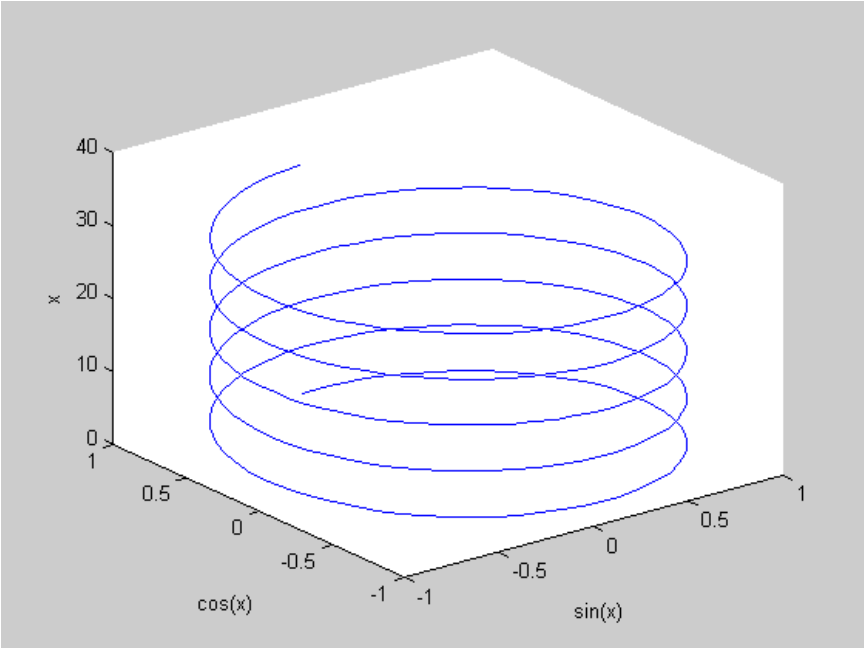
با استفاده از تابع plot3 در متلب می‌توان یک منحنی را در فضای سه‌بعدی ترسیم کرد. روش استفاده از این تابع بسیار شبیه تابع plot است. جز اینکه بازای هر منحنی به سه بردار هم طول نیاز است.

مثال: رسم یک فنر با شعاع برابر با یک:

$$\begin{cases} x=t \\ y=\sin(t) \\ z=\cos(t) \end{cases} \qquad t \in R$$

خمهای فضایی-ادامه

```
>>t=0: pi/50:10*pi;
>>plot3(sin(t) , cos(t) , t);
>> xlabel('sin(x)'); ylabel('cos(x)'); zlabel('x')
```

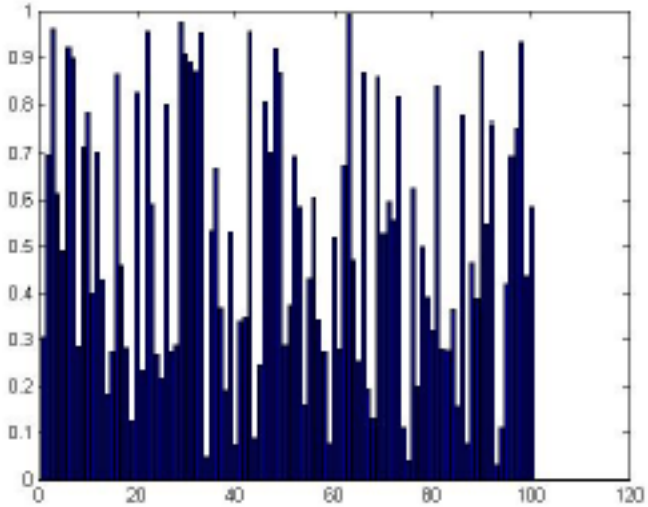


نمودار های آماری

Bar نمودار میله ای

این دستور نمودار میله ای یک مجموعه را رسم می کند .

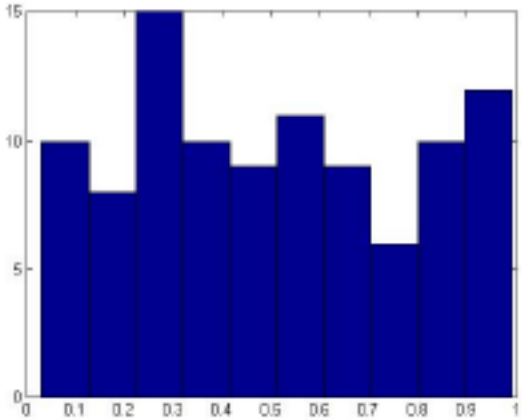
```
>>a=rand(1,100);
>>bar(a)
```



Hist نمودار فراوانی

نمودار هیستوگرام مربوط به مجموعه را رسم میکند .

```
>>a=rand(1,100);
>>hist(a)
```



مفهوم RMSE در قالب یک مثال (جزئیات بر روی تخته توضیح داده شده است)

RMSE.xlsx - Microsoft Excel

Computation

spread 5

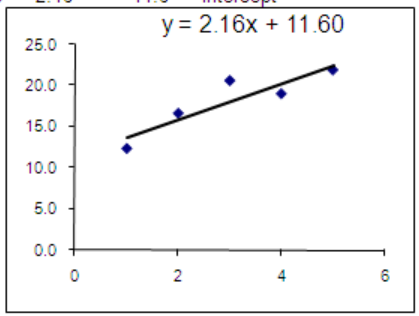
Hit F9 to change the data.
The graph and regression results change also.

1) The Data

x	y
1	12.3
2	16.6
3	20.6
4	19.0
5	21.9

2) Summarizing the Data with Regression

slope 2.16 11.6 intercept



3) Computing the RMSE

Predicted y	Residual
13.76	-1.460
15.92	0.680
18.08	2.520
20.24	-1.240
22.4	-0.500

1.46506 SD of the residuals is the RMSE

4) Using LINEST to get the RMSE

2.16	11.6
0.598108	1.9837
0.812992	1.891384 RMSE from cell 3,2 of LINEST output
13.04212	3
46.656	10.732

1.891384 RMSE from Excel function STEYX

Explaining the Difference

The difference in cells J12 and K18 (or K22) are due to adjustments if the data are a sample.
This is exactly like the sample SD (STDEV) and population SD (STDEVP) issue.
LINEST assumes you have a sample and is adjusting the RMSE with appropriate degrees of freedom.
The number of degrees of freedom is n-k, where n is the number of observations and k is the number of coefficient
Here's the =SQRT(5/3)*STDEVP(J6:J10)
1.891384
This is the standard deviation of the residuals adjusted for "degrees of freedom."

Intro | Computation | SATData | Accordion | PivotTable | RegResults | Pictures

دانلود فایل Excel که در قالب مثال های مختلف مفهوم RMSE را توضیح می دهد :

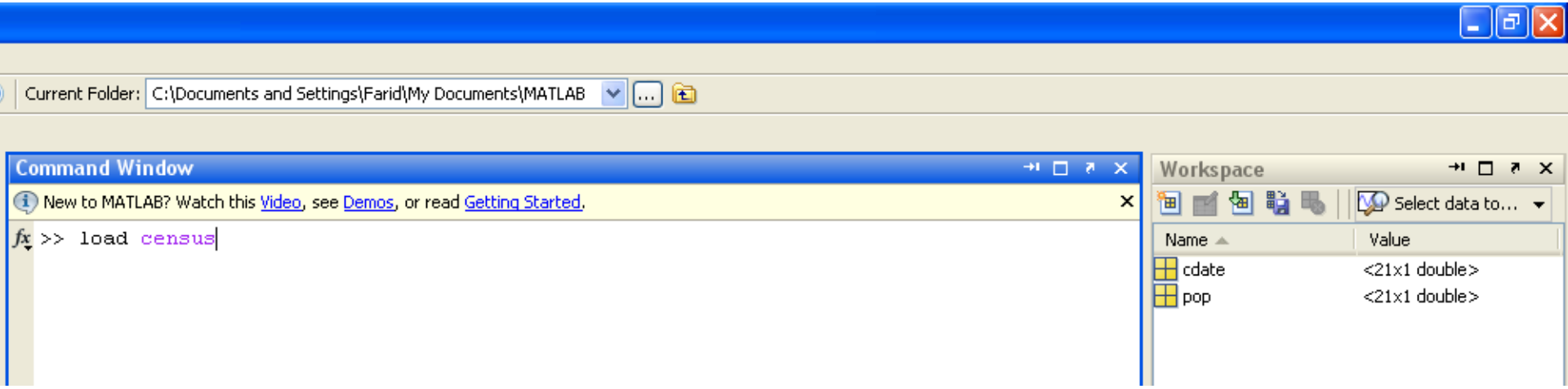
www.faridesm.ir/RMSE.xlsx

۱-۱۱ برازش منحنی به سری نقاط دو بعدی

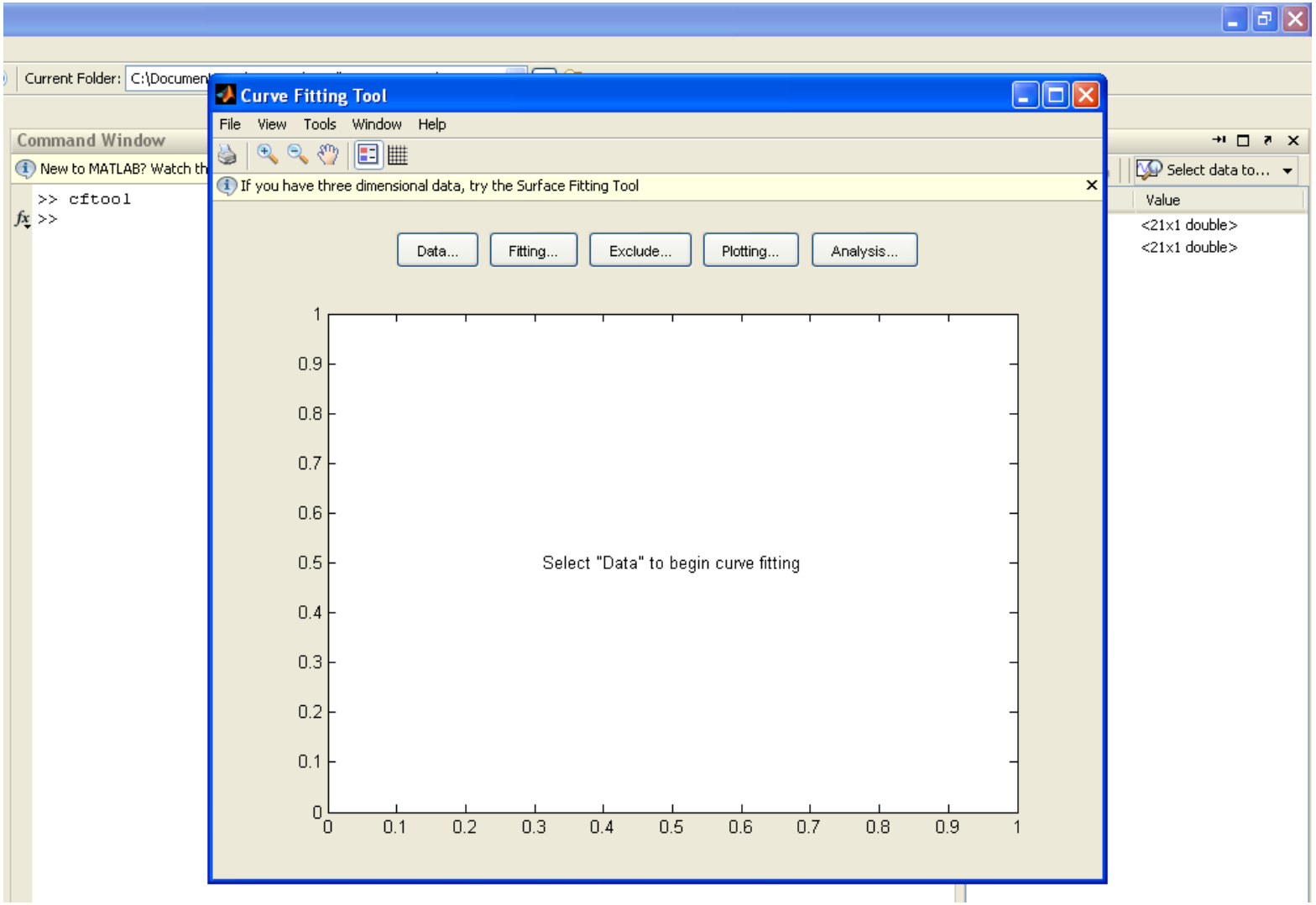
برای برازش منحنی دو بعدی به یک سری نقاط دارای مختصات x,y از دستور cftool استفاده میکنیم.

برای نمونه ابتدا فایل census را که بصورت پیش فرض در MATLAB موجود هست بارگذاری نمائید.

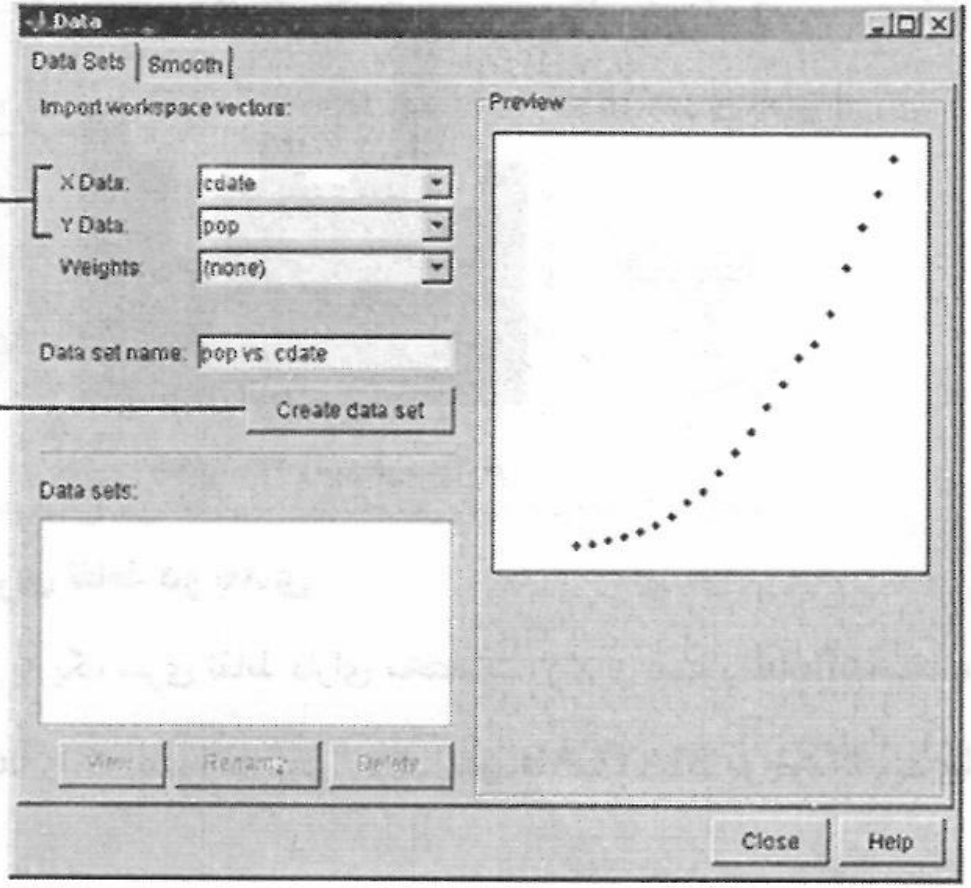
```
load census
cftool
```



برازش منحنی به نقاط ۲ بعدی و تعیین خطای RMSE

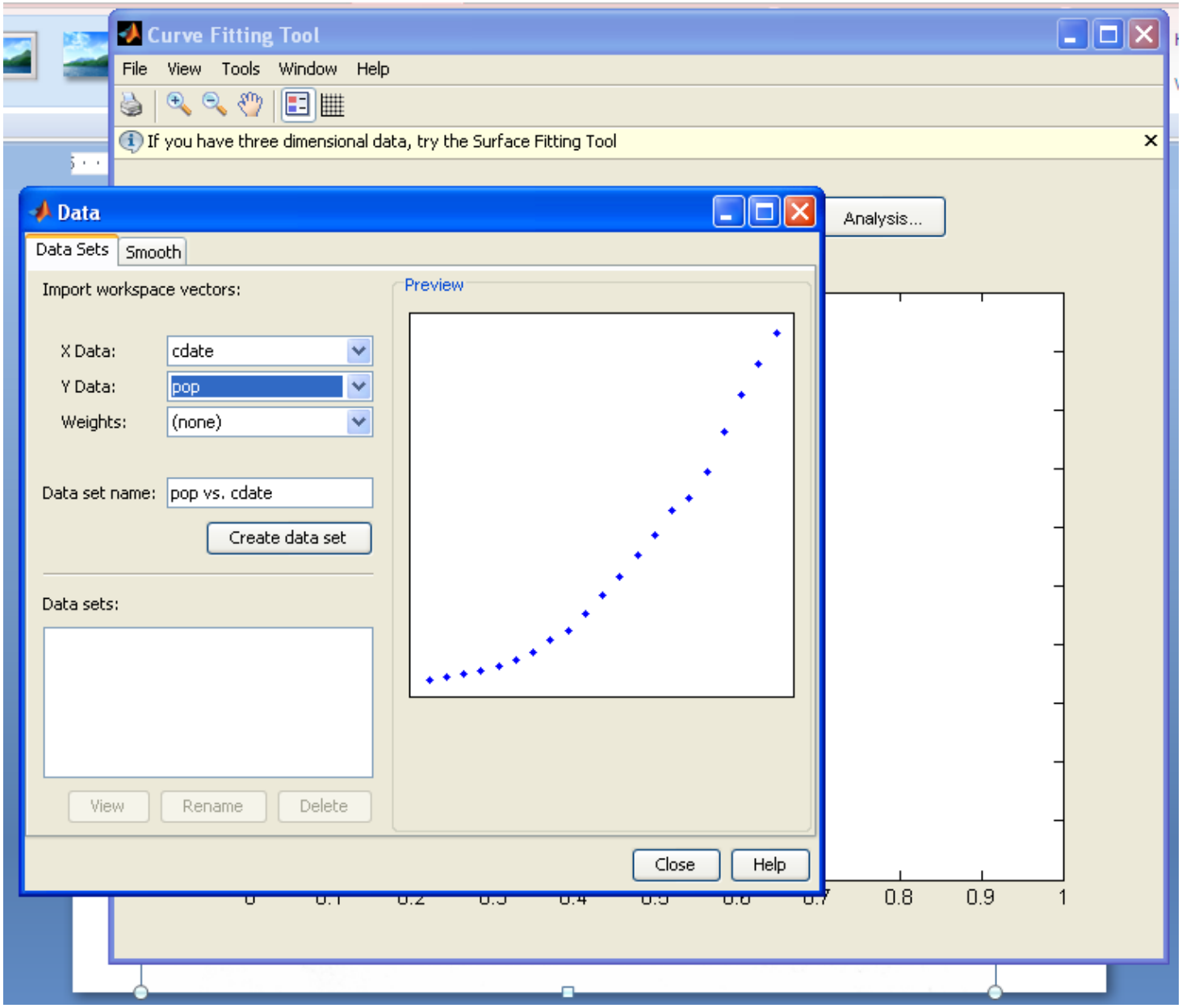


از منو file→ import data (یا دکمه **Data**) را دنبال کرده و y data, x data را بترتیب، **cdata** و **pop** انتخاب نموده و دکمه Create data set را فضار دهید. توزیع نقاط مطابق شکل ۱-۲۶ قابل روئیت میباشند.



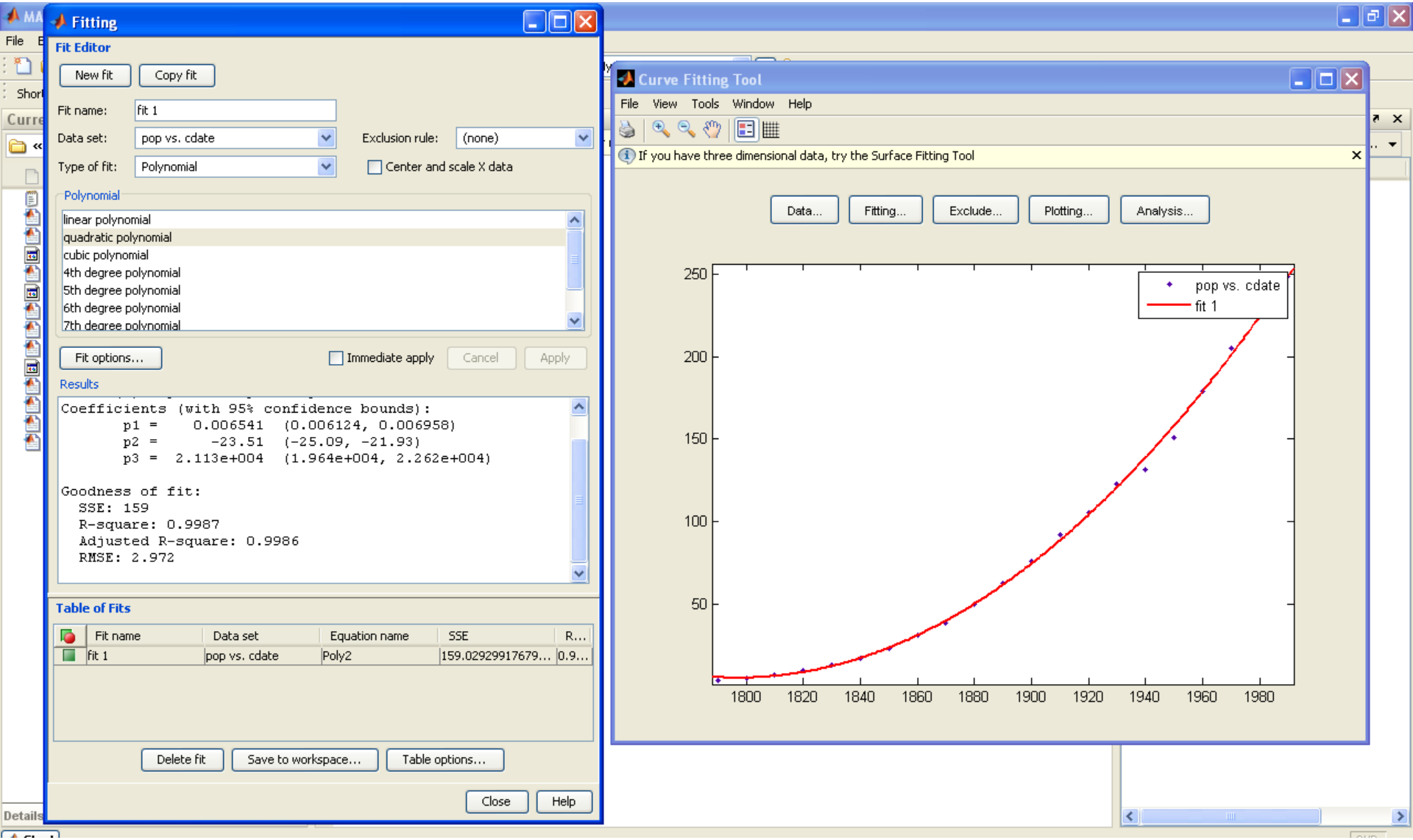
شکل ۱-۲۶

برازش منحنی به نقاط ۲ بعدی و تعیین خطای RMSE



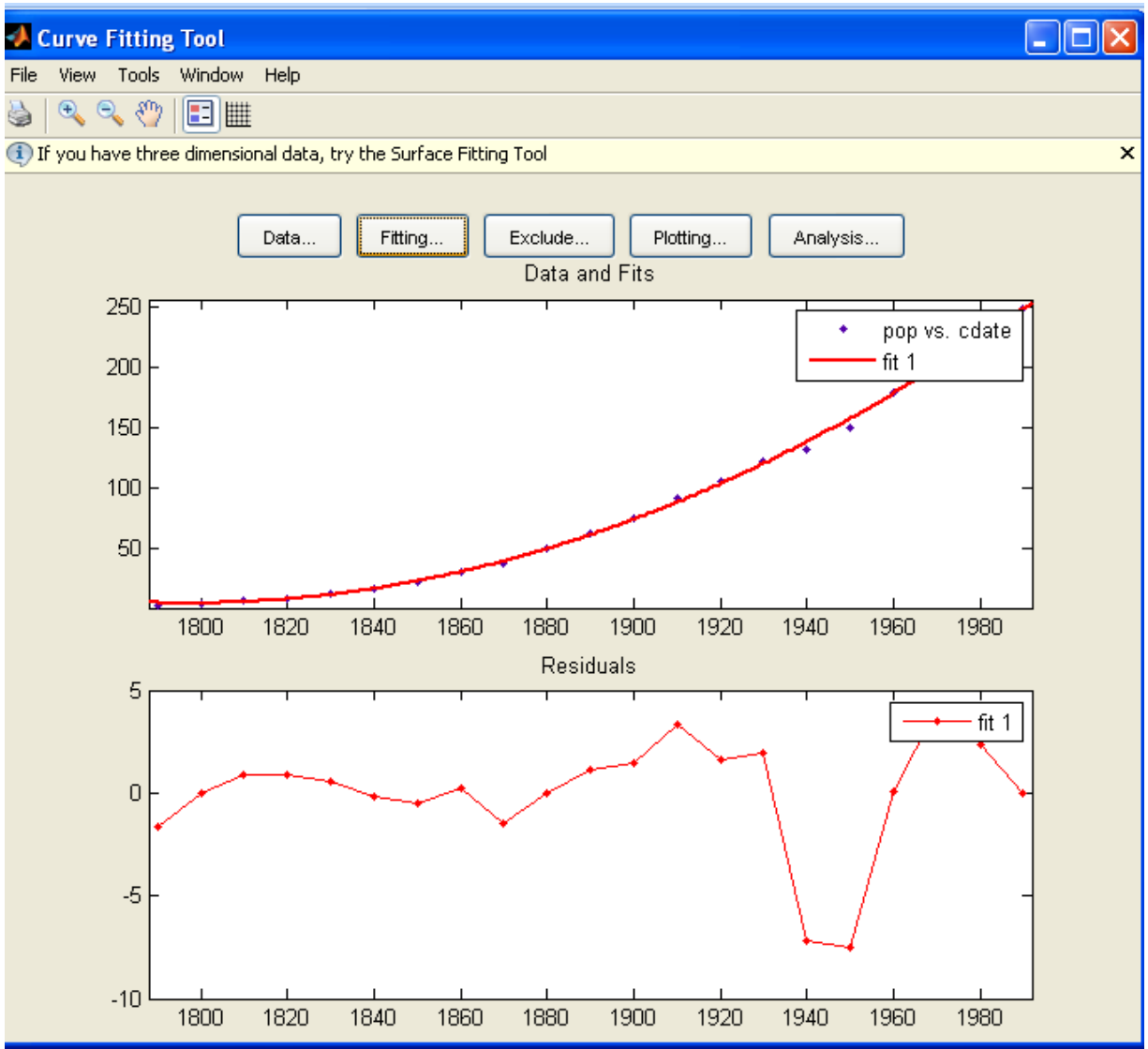
از طریق دکمه Fitting→ New fit. نوع تابع برازش را در بخش Type of Fit انتخاب فرمائید. این تابع میتواند Polynomials, Fourier,... باشد. در صورت انتخاب Polynomials میتوان نوع Polynomials را خطی، درجه ۲ . ۳ و... انتخاب نمود و دکمه **Apply** را برای برازش وارد کرد. برای نمونه برای نوع درجه ۲ نتایج ذیل حاصل میشود. مقدار ضرایب تابع p_i بوده و کیفیت نیکویی برازش با عدد (Root Mean Square Error) RMSE تعیین میگردد. عدد کوچکتر بیانگر فیت بهتر تابع انتخابی هست. توابع مختلف را آزمایش نمائید و تابعی که دارای RMSE کوچکتر هست را به عنوان تابع بهینه انتخاب فرمائید.

برازش منحنی به نقاط ۲ بعدی و تعیین خطای RMSE



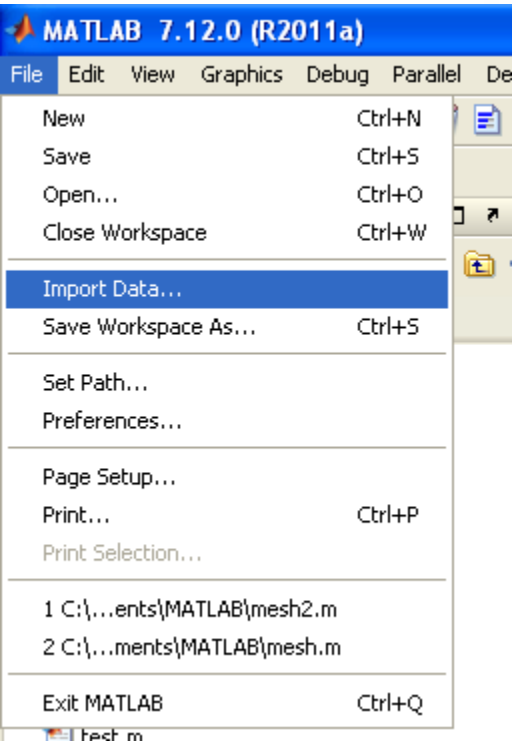
برازش منحنی به نقاط ۲ بعدی و تعیین خطای RMSE

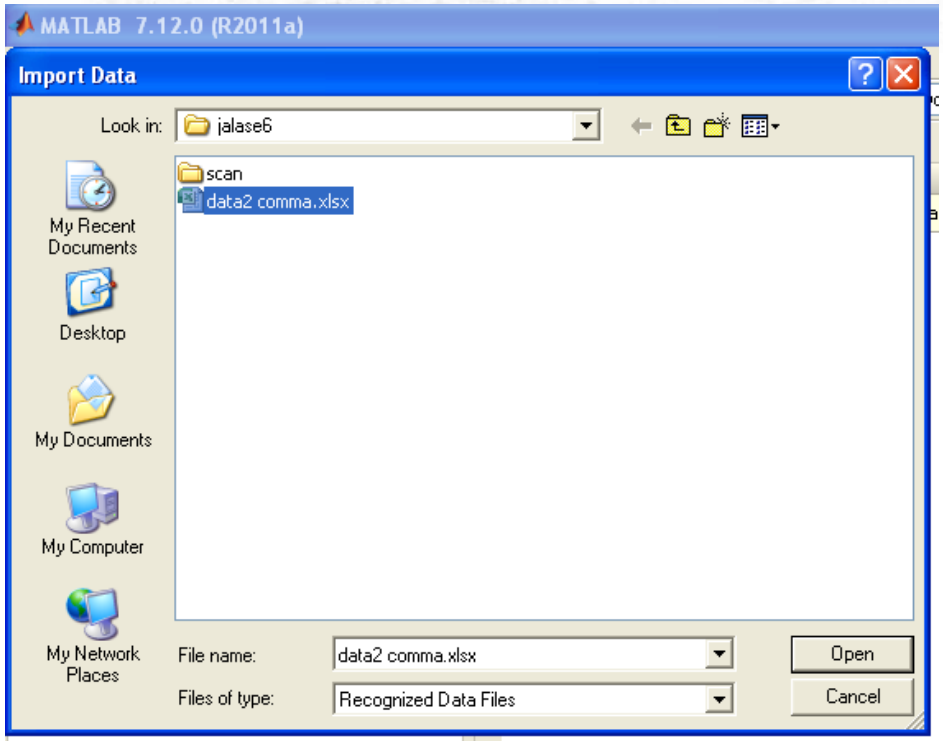
با انتخاب Residual ← view امکان بررسی اختلاف هر نقطه از منحنی برازش شده از نقاط اصلی میسر میگردد.



۱-۱۲ برازش رویه به سری نقاط سه بعدی

برای برازش رویه سه بعدی به سری نقاط دارای مختصات x, y, z از دستور sftool استفاده میکنیم. برای نمونه ابتدا فایل فایل حاوی x, y, z نقاط را در برنامه اکسل آفیس آماده می نماییم سپس فایل را فرخوانی میکنیم. مشابه برازش منحنی میتوان فایلی را که بصورت پیش فرض در MATLAB موجود هست بارگذاری نمائید. بدین منظور از منوی اصلی MATLAB مسیر file→ import data → next→ cerate vector from each column using column names را دنبال کرده و فایل اکسل را بارگذاری مینماییم. فایل اکسل و روش ورود مطابق اشکال ذیل خواهند بود.





Import Wizard

Preview of D:\znu\teory\jalase6\data2 comma.xlsx

Worksheets

data2 comma

data | textdata | colheaders

	1	2	3	4
1	1	144.7720	68.1280	100
2	2	85.1600	-945.2730	106.5580
3	3	170.3200	-945.2730	117.3400
4	4	255.4790	-945.2730	97.0180
5	5	340.6390	-945.2730	90.4270
6	6	425.7990	-945.2730	111.5100
7	7	510.9590	-945.2730	109.8120
8	8	596.1180	-945.2730	109.8120
9	9	681.2780	-945.2730	101.0980
10	10	766.4380	-945.2730	110.7380
11	11	851.5980	-945.2730	110.7380
12	12	936.7570	-860.1140	117.4930
13	13	851.5980	-860.1140	104.8470
14	14	740.8900	-851.5980	115.9640
15	15	689.7940	-851.5980	125.9880
16	16	604.6340	-851.5980	116.1170
17	17	545.0220	-851.5980	120.5400
18	18	493.9270	-851.5980	120.6920
19	19	434.3150	-851.5980	116.1170
20	20	349.1550	-851.5980	115.0450

Help | < Back | Next > | Finish | Generate MATLAB code | Cancel

Import Wizard

Select variables to import using checkboxes

- Create variables matching preview.
- Create vectors from each column using column names.
- Create vectors from each row using row names.

Variables in D:\znu\teory\jalase6\data2 comma.xlsx, Worksheet: data2 comma

Import	Name	Size	Bytes	Class
☑	no	251x1	2008	double
☑	x	251x1	2008	double
☑	y	251x1	2008	double
☑	z	251x1	2008	double

No variable selected for preview.

Help | < Back | Next > | Finish | Generate MATLAB code | Cancel

Workspace

Select data to...

Name	Value
no	<251x1 double>
x	<251x1 double>
y	<251x1 double>
z	<251x1 double>

Surface Fitting Tool

File Fit View Tools Desktop Window Help

untitled fit 1

Fit name: untitled fit 1

X input: x

Y input: y

Z output: z

Weights: (none)

Interpolant

Method: Cubic

Center and scale

Auto fit

Fit

Stop

Results

Duplicate x-y data points detected: using average of the z values.

Piecewise cubic interpolant:
f(x,y) = piecewise cubic surface computed from p
where x is normalized by mean 341.4 and std 341.9
and where y is normalized by mean -80.95 and std 583.9
Coefficients:
p = coefficient structure

Goodness of fit:
SSE: 2.464e-026
R-square: 1
Adjusted R-square: NaN
RMSE: NaN

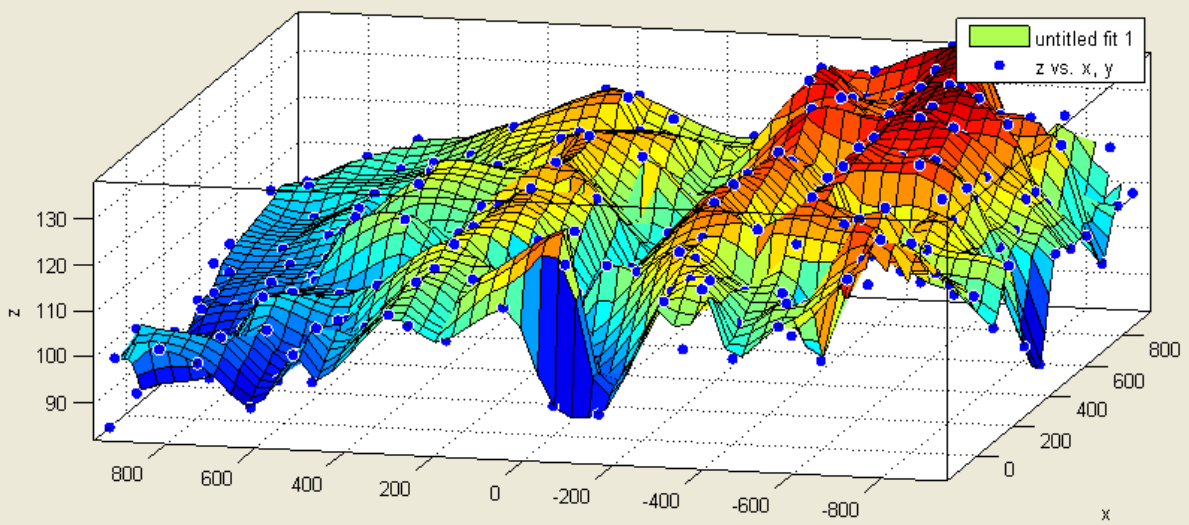


Table of Fits

Fit name	Data	Fit type	SSE	R-square	DFE	Adj R-sq	RMSE	# Coeff	Validation Data	Validation SSE	Validation RMSE
untitled fit 1	z vs. x, y	cubicinterp	2.4638e-26	1	0	NaN	NaN	248			

Surface Fitting Tool

File Fit View Tools Desktop Window Help

untitled fit 1

Fit name: untitled fit 1

X input: x

Y input: y

Z output: z

Weights: (none)

Interpolant

Method: Cubic

Center and scale

Auto fit

Fit

Stop

Results

Duplicate x-y data points detected: using average of the z values.

Piecewise cubic interpolant:
f(x,y) = piecewise cubic surface computed from p
where x is normalized by mean 341.4 and std 341.9
and where y is normalized by mean -80.95 and std 583.9
Coefficients:
p = coefficient structure

Goodness of fit:
SSE: 2.464e-026
R-square: 1
Adjusted R-square: NaN
RMSE: NaN

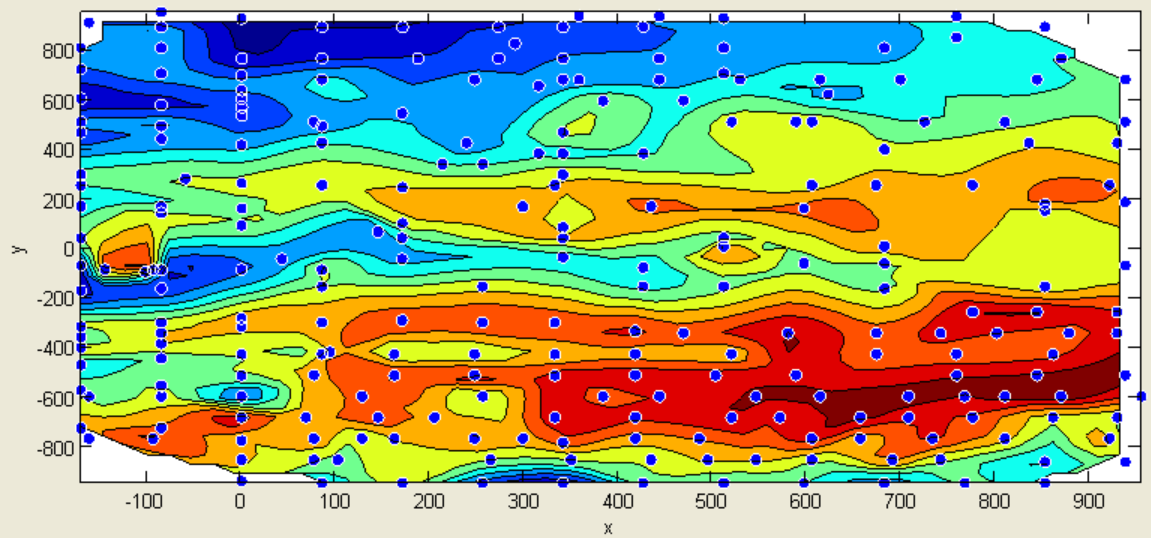


Table of Fits

Fit name	Data	Fit type	SSE	R-square	DFE	Adj R-sq	RMSE	# Coeff	Validation Data	Validation SSE	Validation RMSE
untitled fit 1	z vs. x, y	cubicinterp	2.4638e-26	1	0	NaN	NaN	248			

پایان جلسه ششم

درس تئوری خطاها

جلسه هفتم

فرید اسماعیلی
Farid_63@yahoo.com
www.faridesm.ir

تماس با استاد از طریق پست الکترونیکی
مشاهده اطلاعیه ها، نمرات، دریافت فایل ها در وب سایت

به مجموعه روشهایی که برای تعیین و ارزیابی کمیت هایی که به طور مستقیم و یا غیر مستقیم برای بیان شکل هندسی و یا میدان جاذبه زمین به کار گرفته می شوند ، **روش ژئودتیک** گفته می شود.

هر پروژه ای باید بر اساس کیفیت خواسته شده از کمیت های مورد نظر (کمیت های مجهول)، طرح ریزی گردد . این طرح ریزی شامل تعیین نوع و مقدار اطلاعات (یا مشاهدات) مورد نیاز برای جمع آوری و همچنین تعیین دقت آنها می باشد. اطلاعات جمع آوری شده مورد ارزیابی قرار می گیرند و دقت آنها کنترل می شود تا در محدوده دلخواه باشد. بنابراین مشاهدات کم دقت حذف می شوند . پس از بررسی دقیق اولیه ، این اطلاعات پردازش می شوند و مقادیر کمیت های مجهول بدست می آیند و در نهایت نتایج به دست آمده مورد ارزیابی قرار می گیرند .

مراحل مختلف روش ژئودتیک به شرح ذیل دسته بندی می گردد :

۱ – نخست کمیت هایی که مقادیر آنها باید تعیین شود (پارامترهای مجهول) مشخص می شوند بعد دقت دلخواه برای آنها در نظر گرفته می شود . البته برای اینکار روش خاصی موجود نیست باید بر اساس شناختی که از پروژه وجود دارد، انجام گیرد .

۲ – به علت اینکه نمی توان پارامترهای مجهول را مستقیماً اندازه گیری کرد ، لازم است روابط آنها را با کمیت هایی که مستقیماً مورد اندازه گیری هستند (مشاهدات) بدانیم. بنابراین مرحله دوم عبارت از فرموله کردن روابط فوق به صورت توابع ریاضی است . این توابع (روابط) را **مدل ریاضی** می نامیم که اساس تعیین پارامترهای مجهول است .

۳ – قبل از انجام اندازه گیری های لازم ، باید دقت آنها مشخص شده باشد . واضح است که دقت اندازه گیریها (دقت مشاهدات) مبتنی بر دقت خواسته شده برای پارامترهای مجهول و همچنین روابط موجود در مدل ریاضی می باشد. روش برآورد دقت مشاهدات را به طوری که دقت خواسته شده برای پارامترهای مجهول از طریق مدل ریاضی مربوطه تعیین گردد را **تحلیل اولیه** می نامند . در تحلیل اولیه روش اندازه گیری نیز مشخص می شود .

مراحل مختلف روش ژئودتیک به شرح ذیل دسته بندی می گردد (ادامه):

۴ – اندازه گیری های لازم انجام شده و صحت اندازه گیری مشاهدات (تا حد امکان) با استفاده از **روشهای کنترل مقدماتی و آزمونهای پیش از سرشکنی** تحقیق می شود. سپس مشاهدات بدست آمده مورد آزمایش قرار می گیرند تا معلوم شود که دارای دقت تعیین شده در مرحله تحلیل اولیه هستند یا خیر ؟ اگر مشاهدات دارای دقت لازم نباشند باید تکرار شوند .

۵ – مشاهدات ارزیابی شده به مدل معرفی می شوند از طریق حل مدل ریاضی پارامترهای مجهول و دقت آنها تعیین می شوند . **(سرشکنی).**

۶ – در این مرحله کامل بودن مدل ریاضی مورد ارزیابی قرار می گیرد و همچنین بعد از محاسبات سرشکنی صحت مشاهدات مورد آزمایش قرار می گیرد **(آزمون های پس از سرشکنی)**

۷ – مرحله آخر عبارت از **ارزیابی** پارامترهای محاسبه شده می باشد و دقت و صحت مقادیر بدست آمده برای کمیت های مجهول تعیین می گردد .

نظریه تصمیم گیری (روشهای رسیدن از نمونه به جامعه) به دو دسته کلی تقسیم بندی می شود.
یکی از آنها نظریه برآورد کردن (تئوری تخمین) می باشد و دیگری آزمون فرض ها است .
در نظریه برآورد سعی می شود که پارامترهای مجهول جامعه با برآورد گرهای نقطه ای مانند میانگین و وریانس تخمین زده شوند یا اینکه می توان یک برآورد فاصله ای برای پارامترهای جامعه در نظر گرفت. آزمون هائی را که در این قسمت بررسی می کنیم برای تشخیص وجود خطای سیستماتیک در یک مشاهده ، مستقل از مشاهدات دیگر (فارغ از مدل ریاضی) و همچنین تعیین دقت یک دستگاه اندازه گیری کاربرد دارند .

قبل از پرداختن به این آزمون ها چند تعریف آماری را مرور می کنیم :

آماره : به یک متغیر تصادفی خاص اتلاق می شود که خود تابعی از یک یا چند متغیر تصادفی دیگر است و بستگی به هیچیک از پارامترهای مجهول جامعه ندارد.

فرض آماری: فرض یا بیان یا حدس درباره توزیع جامعه یا پارامترهای جامعه را فرض آماری می نامیم . درست بودن فرض را باید از نتایج بدست آمده از نمونه گیری بررسی نمود .

اگر بنا بفرض پارامترهای جامعه (θ_1, θ_2) دارای مقادیر مشخصی نظیر (θ_1^0, θ_2^0) باشند ، آنرا فرض صفر(null hypothesis) نامیده و با H_0 نشان می دهند.
در برابر هر فرض صفر ممکن است تعداد زیادی مفروضات دیگر باشد که آنها را فرض مقابل (alternative hypothesis) نامیده و با H_1 نشان می دهند.

سنجش صحت هر فرض آماری نیازمند برقراری آزمون می باشد ، ماحصل آزمون آماری دو نتیجه است : قبول فرض H_0 و یا رد H_0 .

دو نوع فرض آماری وجود دارد :
H0 : فرض صفر – اظهار نظر در مورد ویژگی آماری یک جامعه
H1 : نظریه مقابل فرض صفر

نتیجه آزمون آماری دو چیز بیشتر نیست :
۱- قبول H0 و رد H1
۲- رد H0 (که الزاماً به معنی قبول H1 نیست ولی در برخی شرایط می تواند باشد)

خطاهای نوع اول و دوم :

رد فرض H0 وقتی که H0 در واقع درست باشد را خطای نوع اول می نامند.

احتمال ارتکاب خطای نوع اول را با α نمایش می دهند و به آن سطح معنی دار بودن (یا سطح اعتبار) آزمون می گویند .

مقدار α را معمولاً ۰.۰۵ یا ۰.۰۲۵ و یا حتی ۰.۰۰۱ انتخاب می کنند.

در جلسات قبل با این مفهوم در بحث تعیین فاصله اطمینان هم برخورد کردیم که در آنجا ، احتمال اینکه مقدار واقعی کمیت خارج از فاصله مذکور باشد را به همین شکل بیان نمودیم.

در مقابل خطای نوع اول ، خطای نوع دوم را داریم که به معنی قبول فرض H0 است هنگامی که H0 در واقع حقیقت نداشته باشد . احتمال ارتکاب خطای دوم را با β نمایش می دهیم و (۱-β) را توان آزمون می گویند .power of test.

تصمیم / واقعیت	قبول فرض H_0	رد فرض H_0
H_0 حقیقت دارد	تصمیم گیری صحیح $p = 1 - \alpha$	خطای نوع اول $P = \alpha$
H_0 حقیقت ندارد H_1 حقیقت دارد	خطای نوع دوم $P = \beta$	تصمیم گیری صحیح $P = 1 - \beta$

آزمون فرض H0 و خطاهای نوع اول و دوم

طبیعی است که در مقابل هر فرض صفر می توان بینهایت فرض مقابل در نظر گرفت که در میان آنها قویترین آزمون ، آزمونی است که احتمال خطای نوع دوم آن کمترین باشد و این فقط وقتی اتفاق می افتد که فرض مقابل در تضاد مطلق با فرض صفر باشد بگونه ای که بتوان با رد فرض صفر ، فرض مقابل را پذیرفت و یا با قبول فرض صفر ، فرض مقابل را رد کرد.

به عنوان مثال اگر فرض های H1,H0 به صورتهای زیر باشند:

$$\begin{cases} H_0: \mu = 57.326 \\ H_1: \mu < 57.326 \end{cases} \quad (۱)$$

در این حالت رد فرض H0 به معنای پذیرش H1 نیست زیرا که ممکن است μ >۵۷.۲۳۶ و در اینجا احتمال ارتکاب خطای نوع دوم زیاد است.

ولی اگر فرض های H1,H0 به صورتهای زیر باشند:

$$\begin{cases} H_0: \mu = 57.326 \\ H_1: \mu \neq 57.326 \end{cases} \quad (۲)$$

رد H0 لزوماً به معنای تائید H1 است و احتمال ارتکاب خطای نوع دوم به مراتب کمتر از حالت قبل است و بنابراین آزمون شماره (۲) قویتر است (توان آزمون بیشتر است) .

آزمون میانگین (در صورت نرمال بودن تابع توزیع مشاهدات) صحت ادعا در مورد یک مقدار معین برای میانگین جامعه را بررسی می کند.
به بیان دیگر در این آزمون یک کمیتی که مقدار آن از پیش معلوم است توسط یک دستگاه اندازه گیری می شود و سپس مقدار به دست آمده از اندازه گیری جدید با مقدار از پیش معلوم مقایسه می شود. اگر اختلاف بین این دو معنی دار تشخیص داده شد، وجود خطای سیستماتیک تأیید می شود. و در غیر این صورت وجود خطای سیستماتیک تکذیب می شود.
فرض صفر در اینجا میانگین را مساوی یک مقدار از قبل معلوم فرض می کند و فرض مقابل به طور کلی یکی از سه حالت ذکر شده ذیل است :

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \\ H_1: \mu > \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$$

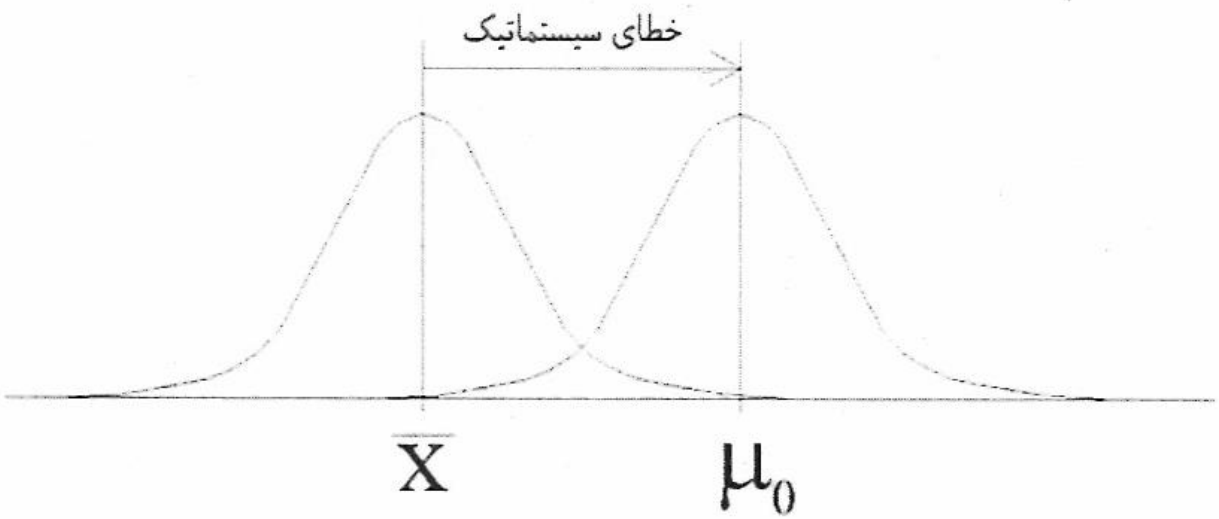
مسئله دیگر معلوم بودن یا مجهول بودن وریانس جامعه است که باعث بوجود آمدن دو حالت می شود : آماره مورد استفاده وقتی که σ^2 معلوم است $\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ و در حالتی که σ^2 مجهول است ، به صورت $\frac{\bar{x}-\mu_0}{S/\sqrt{n}}$ می باشد .

که برای آماره های ذکر شده n تعداد مشاهدات نمونه است . برای حالتی که σ^2 معلوم است فرض H_0 در موارد زیر رد می شود (ζ_α و $\zeta_{\alpha/2}$ مقادیر استخراج شده از جدول توزیع نرمال هستند)

$H_0: \mu = \mu_0 \ \ vs \ H_1: \mu \neq \mu_0$	$H_0: \mu = \mu_0 \ \ vs \ H_1: \mu > \mu_0$	$H_0: \mu = \mu_0 \ \ vs \ H_1: \mu < \mu_0$
H_0 رد می شود اگر : $\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} > \zeta_{\alpha/2}$ و یا $\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} < -\zeta_{\alpha/2}$	H_0 رد می شود اگر : $\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} > \zeta_\alpha$	H_0 رد می شود اگر : $\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} < -\zeta_\alpha$
سطح اطمینان : $P(-\zeta_{\frac{\alpha}{2}} < \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} < \zeta_{\frac{\alpha}{2}}) = 1-\alpha$	احتمال خطای نوع اول : $P(\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} > \zeta_\alpha) = \alpha$	احتمال خطای نوع اول : $P(\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} < -\zeta_\alpha) = \alpha$

آزمون میانگین (σ^2 معلوم)

در حالتی که σ^2 نامعلوم است از S به جای σ استفاده شده و به جای توزیع نرمال از توزیع t استیودنت با n-1 درجه آزادی استفاده می شود و بقیه موارد مانند جدول فوق رعایت می گردد . در نقشه برداری هدف از انجام آزمون میانگین بررسی دستگاه های اندازه گیری از نظر وجود خطاهای سیستماتیک می باشد . در صورت وجود خطای سیستماتیک μ_0 با $\bar{x}$ اختلاف معنی داری خواهد داشت



: اختلاف بین میانگین نمونه و مقدار واقعی یک کمیت در صورت وجود خطای سیستماتیک معنی دار است.

گفتیم که $\bar{x}$ برآوردگر خوبی برای μ می باشد و در صورت نرمال بودن تابع توزیع مشاهدات این ادعا کاملاً صحیح است . هنگامی که خطای سیستماتیک وجود دارد ، امید ریاضی جامعه آماری مشاهده انجام شده (μ) نسبت به واقعیت (μ_0) انحراف خواهد داشت؛ پس برآورد کننده آن $\bar{x}$ هم نسبت به μ_0 انحراف خواهد داشت . اگر این انحراف از یک سطح معنی دار معین (α) تجاوز کند فرض صفر رد می شود و وجود خطای سیستماتیک تأئید می گردد .

فرض می کنیم در مورد کالیبره یا کالیبره نبودن یک دستگاه طولیاب یا زاویه یاب تردید داریم (می دانیم که در صورت کالیبره نبودن ، خطای سیستماتیک در مشاهدات وارد می شود) برای این که در این مورد تحقیق کنیم دستگاه را بر روی یک باز(base line)یا زاویه از قبل معلوم قرار داده و با رعایت کلیه موارد لازم ، اندازه گیری های مکرری را به تعداد نسبتاً زیاد انجام می دهیم و آن گاه با برگزار نمودن آزمون میانگین وجود خطای سیستماتیک در نمونه جمع آوری شده را بررسی می کنیم و در این جا دو حالت به وجود می آید یکی این که از دقت دستگاه مطلع باشیم (σ^2 معلوم) یا از آن اطلاع نداشته باشیم که در این حالت مجبور به استفاده از وریانس نمونه خواهیم شد . در صورتی که آزمون میانگین رد شود ، دستگاه نیاز به بررسی و کالیبراسیون خواهد داشت .

مثال - جدول زیر اندازه گیری های انجام شده توسط یک دستگاه طولیاب از یک باز با طول معلوم $819.1227m$ را نشان می دهد . دقت ادعا شده از سوی کارخانه سازنده دستگاه برای آن فاصله $5.34mm$ می باشد . با فرض صحیح بودن دقت اسمی دستگاه ، مطلوبست بررسی دستگاه از نظر وجود خطای سیستماتیک در سطوح اطمینان ۹۹٪ و ۹۵٪ .

819.124	819.128	819.126	819.126
819.126	819.121	819.124	819.122
819.131	819.118	819.123	819.123
819.120	819.120	819.127	819.124
819.129	819.130	819.131	819.129

فواصل اندازه گیری شده برای بررسی صحت اندازه گیری با یک دستگاه طولیاب (واحد : متر)

گامهای حل مسأله :

۱- تشخیص نوع آزمون : آزمون میانگین

۲- تعیین وضعیت : σ معلوم و یا σ مجهول : در این مسأله σ معلوم است. پس آماره : $\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$

۳- محاسبه آماره :

$$\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{0.0024}{0.00534/\sqrt{20}} = 2.01 \,, \bar{x} = 819.1251$$

۴- اجرای قاعده تصمیم و استخراج مقادیر از جدول توزیع آماری مربوطه (در این مثال توزیع نرمال است)

α= 0.05 % -- > 1- α=0.95 %

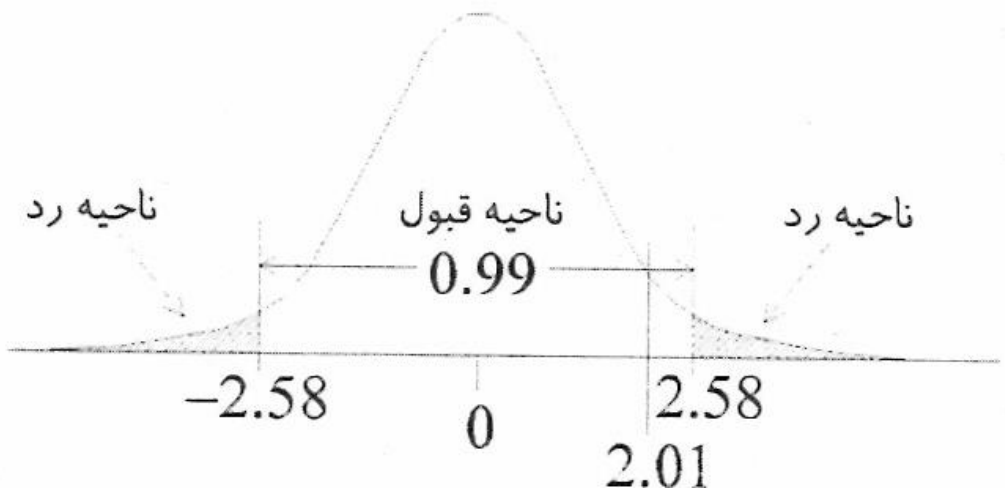
α= 0.01 % -- > 1- α=0.99 %

فرض های صفر و مقابل را تشکیل می دهیم :

$$\begin{cases} H_0: \mu = 819.1227m \\ H_1: \mu \neq 819.1227m \end{cases} \quad , \alpha = 0.01 \quad (\text{برای سطح اطمینان } 99\%)$$

۵- صدور حکم :

مقدار $\zeta_{\alpha/2}$ به ازای $\alpha = 0.01$ از جدول توزیع نرمال 2.58 استخراج می شود و چون $-2.58 < 2.01 < 2.58$ و بنابر جدول توضیح داده شده فرض صفر در سطح اطمینان ۹۹ ٪ پذیرفته می شود و خطای سیستماتیک وجود ندارد.



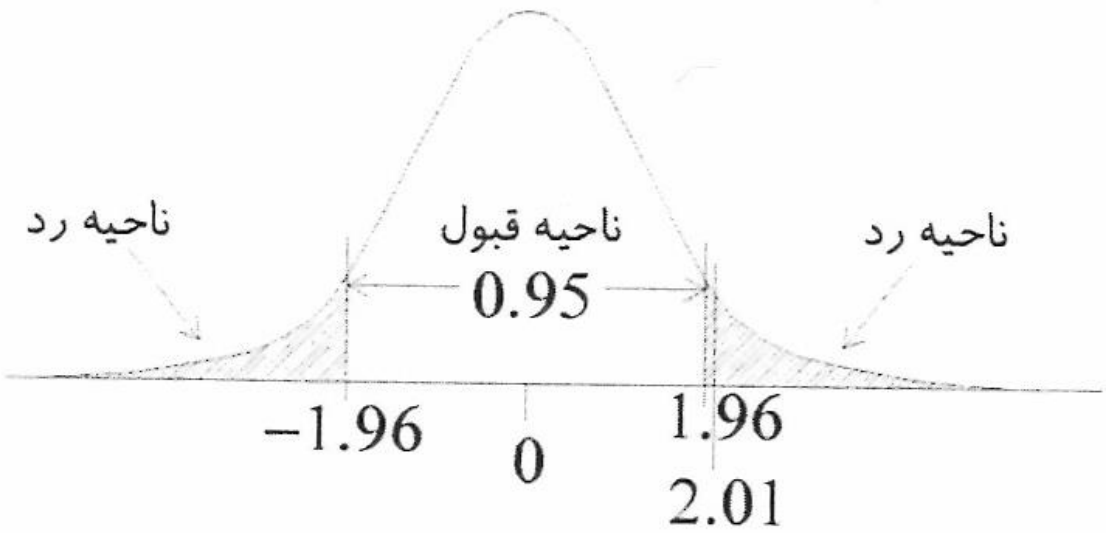
$H_0: \mu = \mu_0 \;\; vs \; H_1: \mu \neq \mu_0$
H_0 رد می شود اگر : $\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} > \zeta_{\alpha/2}$ و یا $\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} < -\zeta_{\alpha/2}$
سطح اطمینان : $P(-\zeta_{\frac{\alpha}{2}} < \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} < \zeta_{\frac{\alpha}{2}}) = 1 - \alpha$

آزمون میانگین در سطح اطمینان ۹۹ ٪

برای سطح اطمینان 95% داریم : $\alpha = 0.05$,

$$\begin{cases} H_0: \mu = 819.1227m \\ H_1: \mu \neq 819.1227m \end{cases}$$

مقدار $\zeta_{\alpha/2}$ به ازای $\alpha = 0.05$ از جدول توزیع نرمال 1.96 استخراج می شود و چون $1.96 < 2.01$ است و بنابر جدول توضیح داده شده فرض صفر در سطح اطمینان ۹۵ ٪ رد می شود و وجود خطای سیستماتیک تأئید می گردد .



آزمون میانگین در سطح اطمینان ۹۵ ٪

آزمون های آماری Decision theorem – مثال

جدول آماری 1– سطح زیر منحنی تابع چگالی نرمال استاندارد

$$\alpha = \int_{z_\alpha}^{\infty} f(z)\,dz = P(z > z_\alpha) = 1 - \int_{-\infty}^{z_\alpha} f(z)\,dz$$

$z_\alpha \rightarrow$ ↓	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0076	0.0073	0.0071	0.0070	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0042	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0014	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002

در این مثال ملاحظه شد که برای دو مقدار مختلف α ممکن است وجود یا عدم وجود خطای سیستماتیک مورد تردید قرار بگیرد و سؤالی که مطرح می شود این است که کدام یک را باید انتخاب کرد و معیار چیست ؟ در اصطلاح α (احتمال وقوع خطای نوع اول) را احتمال رد کردن یک عنصر مرغوب (مثلاً یک دستگاه سالم) و β (احتمال وقوع خطای نوع دوم) را احتمال قبول کردن یک دستگاه غیر سالم می نامند. یعنی اگر $\alpha = 0.05$ در نظر گرفته شود احتمال رد کردن دستگاه سالم را افزایش داده ایم و اگر $\alpha = 0.01$ در نظر گرفته شود احتمال رد کردن دستگاه سالم را کاهش داده ایم. در یک آزمون خوب هم α و هم β باید مینیمم باشند . اما عملاً در هر مسئله ای که مورد مطالعه قرار می گیرد مینیمم کردن آن ها میسر نیست و هر عملی که منجر به مینیمم کردن α گردد معمولاً منجر به افزایش β می شود . در اغلب تصمیم گیری ها قبول یک فرض غلط خیلی خطرناک تر از رد یک فرض صحیح است یعنی نتایج حاصل از خطای نوع دوم خیلی جدی تر از نتایج حاصل از خطای نوع اول است . برای مثال پذیرفتن دستگاهی که کالیبراسیون آن بهم خورده در مقابل نپذیرفتن دستگاهی که کالیبره می باشد زیرا که هزینه کالیبراسیون یک دستگاه معیوب ممکن است ۱۰۰ هزار تومان باشد در حالی که کار کردن با یک دستگاه معیوب می تواند ده ها میلیون تومان و یا بیشتر ضرر وارد کند . پس باید سعی شود که تا حد امکان از خطای نوع دوم پرهیز شود عملاً استفاده از سطح معنی دار $\alpha = 0.05$ برای آزمون کالیبراسیون دستگاه ها منطقی تر است.

آزمون واریانس در بررسی دقت یک دستگاه اندازه گیری به کار گرفته می شود و برای تحقیق در مورد این که آیا یک جامعه با توزیع نرمال دارای وریانس به خصوصی هست یا نه باید از این آزمون استفاده کرد . بر این اساس فرض صفر و فرض های مقابل بدین شرح می باشند :

$$\begin{cases} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 < \sigma_0^2 \\ H_1: \sigma^2 > \sigma_0^2 \\ H_1: \sigma^2 \neq \sigma_0^2 \end{cases}$$

در این جا نیز دو حالت به وجود می آید یکی وقتی که میانگین جامعه (μ) معلوم است که از آماره $\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2 = \frac{nS^2}{\sigma^2}$ استفاده می شود (در این حالت وریانس نمونه از فرمول $S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$ محاسبه می شود .) و وقتی μ مجهول است از آماره $\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\sigma}\right)^2 = \frac{(n-1)S^2}{\sigma^2}$ استفاده می شود . n تعداد مشاهدات نمونه است .

جدول زیر کیفیت انجام آزمون وریانس را نشان می دهد

وضعیت	$H_0: \sigma^2 = \sigma_0^2$ vs $H_1: \sigma^2 \neq \sigma_0^2$	$H_0: \sigma^2 = \sigma_0^2$ vs $H_1: \sigma^2 > \sigma_0^2$	$H_0: \sigma^2 = \sigma_0^2$ vs $H_1: \sigma^2 < \sigma_0^2$
μ معلوم	H_0 رد می شود اگر : $\frac{nS^2}{\sigma_0^2} > x^2_{\alpha/2,n}$ و یا $\frac{nS^2}{\sigma_0^2} < x^2_{1-\alpha/2,n}$	H_0 رد می شود اگر : $\frac{nS^2}{\sigma_0^2} > x^2_{\alpha,n}$	H_0 رد می شود اگر : $\frac{nS^2}{\sigma_0^2} < x^2_{1-\alpha,n}$
μ مجهول	$\frac{(n-1)S^2}{\sigma_0^2} > x^2_{\alpha/2,n-1}$ و یا $\frac{(n-1)S^2}{\sigma_0^2} < x^2_{1-\alpha/2,n-1}$	$\frac{(n-1)S^2}{\sigma_0^2} > x^2_{\alpha,n-1}$	$\frac{(n-1)S^2}{\sigma_0^2} < x^2_{1-\alpha,n-1}$

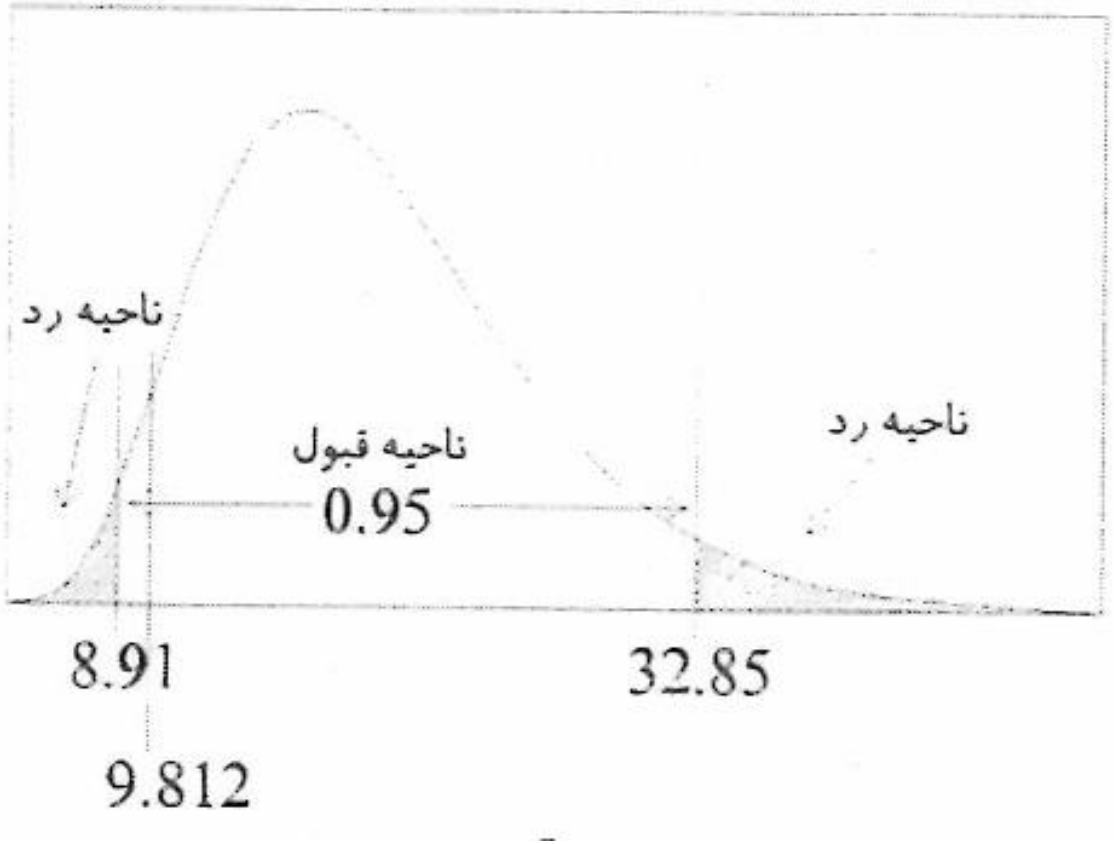
X توزیع کای اسکور می باشد

مثال : مطلوب است بررسی دقت ادعا شده برای دستگاه طولیاب ذکر شده در مثال قبلی در سطح معنی دار $\alpha=0.05$.

حل – چون در سطح معنی دار $\alpha = 0.05$ وجود خطای سیستماتیک برای دستگاه تأئید گردید پس $\mu \neq \mu_0$ و از میانگین جامعه نمی توانیم استفاده کنیم و به جای آن از میانگین نمونه $\bar{x} = 819.1227\text{m}$ استفاده می شود :

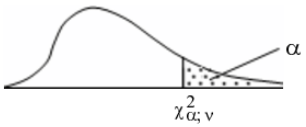
$$\begin{cases} H_0: \sigma^2 = (5.34)^2 mm^2 \\ H_1: \sigma^2 \neq (5.34)^2 mm^2 \end{cases} , \alpha = 0.05 \qquad \frac{(n-1)S^2}{\sigma_0^2} = \frac{(19)(14.7263)}{(5.34)^2} = 9.812$$

چون $8.91 < 9.812 < 32.85$ ، فرض صفر پذیرفته می شود ، دقت ادعا شده برای دستگاه صحیح است.


ناحیه رد
ناحیه قبول
0.95
ناحیه رد
8.91
32.85
9.812

جدول توزیع کای اسکور

Table of the Chi-square Distribution



$\alpha =$	0.995	0.99	0.98	0.975	0.95	0.90	0.80	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.001	$=\alpha$
$\nu = 1$	0.0000393	0.000157	0.000628	0.000982	0.00393	0.0158	0.0642	1.642	2.706	3.841	5.024	5.412	6.635	7.879	10.827	$\nu = 1$
2	0.0100	0.0201	0.0404	0.0506	0.103	0.211	0.446	3.219	4.605	5.991	7.378	7.824	9.210	10.597	13.815	2
3	0.0717	0.115	0.185	0.216	0.352	0.584	1.005	4.642	6.251	7.815	9.348	9.837	11.345	12.838	16.268	3
4	0.207	0.297	0.429	0.484	0.711	1.064	1.649	5.989	7.779	9.488	11.143	11.668	13.277	14.860	18.465	4
5	0.412	0.554	0.752	0.831	1.145	1.610	2.343	7.289	9.236	11.070	12.832	13.388	15.086	16.750	20.517	5
6	0.676	0.872	1.134	1.237	1.635	2.204	3.070	8.558	10.645	12.592	14.449	15.033	16.812	18.548	22.457	6
7	0.989	1.239	1.564	1.690	2.167	2.833	3.822	9.803	12.017	14.067	16.013	16.622	18.475	20.278	24.322	7
8	1.344	1.646	2.032	2.180	2.733	3.490	4.594	11.030	13.362	15.507	17.535	18.168	20.090	21.955	26.125	8
9	1.735	2.088	2.532	2.700	3.325	4.168	5.380	12.242	14.684	16.919	19.023	19.679	21.666	23.589	27.877	9
10	2.156	2.558	3.059	3.247	3.940	4.865	6.179	13.442	15.987	18.307	20.483	21.161	23.209	25.188	29.588	10
11	2.603	3.053	3.609	3.816	4.575	5.578	6.989	14.631	17.275	19.675	21.920	22.618	24.725	26.757	31.264	11
12	3.074	3.571	4.178	4.404	5.226	6.304	7.807	15.812	18.549	21.026	23.337	24.054	26.217	28.300	32.909	12
13	3.565	4.107	4.765	5.009	5.892	7.042	8.634	16.985	19.812	22.362	24.736	25.472	27.688	29.819	34.528	13
14	4.075	4.660	5.368	5.629	6.571	7.790	9.467	18.151	21.064	23.685	26.119	26.873	29.141	31.319	36.123	14
15	4.601	5.229	5.985	6.262	7.261	8.547	10.307	19.311	22.307	24.996	27.488	28.259	30.578	32.801	37.697	15
16	5.142	5.812	6.614	6.908	7.962	9.312	11.152	20.465	23.542	26.296	28.845	29.633	32.000	34.267	39.252	16
17	5.697	6.408	7.255	7.564	8.672	10.085	12.002	21.615	24.769	27.587	30.191	30.995	33.409	35.718	40.790	17
18	6.265	7.015	7.906	8.231	9.390	10.865	12.857	22.760	25.989	28.869	31.526	32.346	34.805	37.156	42.312	18
19	6.844	7.633	8.567	8.907	10.117	11.651	13.716	23.900	27.204	30.144	32.852	33.687	36.191	38.582	43.820	19
20	7.434	8.260	9.237	9.591	10.851	12.443	14.578	25.038	28.412	31.410	34.170	35.020	37.566	39.997	45.315	20
21	8.034	8.897	9.915	10.283	11.591	13.240	15.445	26.171	29.615	32.671	35.479	36.343	38.932	41.401	46.797	21
22	8.643	9.542	10.600	10.982	12.338	14.041	16.314	27.301	30.813	33.924	36.781	37.659	40.289	42.796	48.268	22
23	9.260	10.196	11.293	11.688	13.091	14.848	17.187	28.429	32.007	35.172	38.076	38.968	41.638	44.181	49.728	23
24	9.886	10.856	11.992	12.401	13.848	15.659	18.062	29.553	33.196	36.415	39.364	40.270	42.980	45.558	51.179	24
25	10.520	11.524	12.697	13.120	14.611	16.473	18.940	30.675	34.382	37.652	40.646	41.566	44.314	46.928	52.620	25
26	11.160	12.198	13.409	13.844	15.379	17.292	19.820	31.795	35.563	38.885	41.923	42.856	45.642	48.290	54.052	26
27	11.808	12.879	14.125	14.573	16.151	18.114	20.703	32.912	36.741	40.113	43.194	44.140	46.963	49.645	55.476	27
28	12.461	13.565	14.847	15.308	16.928	18.939	21.588	34.027	37.916	41.337	44.461	45.419	48.278	50.993	56.893	28
29	13.121	14.256	15.574	16.047	17.708	19.768	22.475	35.139	39.087	42.557	45.722	46.693	49.588	52.336	58.302	29
30	13.787	14.953	16.306	16.791	18.493	20.599	23.364	36.250	40.256	43.773	46.979	47.962	50.892	53.672	59.703	30
40	20.706	22.164	23.838	24.433	26.509	29.051	32.345	47.269	51.805	55.759	59.342	60.436	63.691	66.766	73.402	40
50	27.991	29.707	31.664	32.357	34.764	37.689	41.449	58.164	63.167	67.505	71.420	72.613	76.154	79.490	86.661	50
60	35.535	37.485	39.699	40.482	43.188	46.459	50.641	68.972	74.397	79.082	83.298	84.580	88.379	91.952	99.607	60
70	43.275	45.442	47.893	48.758	51.739	55.329	59.898	79.715	85.527	90.531	95.023	96.388	100.425	104.215	112.317	70
80	51.171	53.539	56.213	57.153	60.391	64.278	69.207	90.405	96.578	101.880	106.629	108.069	112.329	116.321	124.839	80
90	59.196	61.754	64.634	65.646	69.126	73.291	78.558	101.054	107.565	113.145	118.136	119.648	124.116	128.299	137.208	90
100	67.327	70.065	73.142	74.222	77.929	82.358	87.945	111.667	118.498	124.342	129.561	131.142	135.807	140.170	149.449	100

پایان جلسه هفتم

درس تئوری خطاها

جلسه هشتم

فرید اسماعیلی
Farid_63@yahoo.com
www.faridesm.ir

تماس با استاد از طریق پست الکترونیکی
مشاهده اطلاعیه ها، نمرات، دریافت فایل ها در وب سایت

همان‌طور که می‌دانیم مقدار بسیاری از کمیت‌ها با اندازه‌گیری کمیت‌های دیگر و با به‌کارگیری یک رابطهٔ ریاضی بین آن کمیت‌ها و کمیت مورد نظر به‌دست می‌آید. مثلاً اندازهٔ مساحت زمینی که به شکل مستطیل است با اندازه‌گیری a و b (طول وعرض) آن زمین و با استفاده از رابطهٔ ریاضی A=a.b حاصل می‌شود. هدف از این بحث این است که اگر در اندازه‌گیری کمیت‌های اولیه با خطاهایی مواجه باشیم تأثیر این خطاها در اندازهٔ کمیت مورد نظر چگونه است.

به طور خلاصه می‌توان گفت محاسبه خطای متوسط هندسی در اندازه‌گیریهای غیرمستقیم به صورت زیر می‌باشد، اگر فرض کنیم کمیت u از طریق اندازه‌گیری چند کمیت دیگر به عنوان مثال از x , y , z بدست بیاید. یعنی:

$$u = f(x, y, z)$$

در این صورت دقت اندازه‌گیری کمیت u از رابطه زیر بدست می‌آید:

$$\boxed{\delta_u = \pm\sqrt{(\frac{\partial f}{\partial x})^2 \delta_x^2 + (\frac{\partial f}{\partial y})^2 \delta_y^2 + (\frac{\partial f}{\partial z})^2 \delta_z^2}}$$

در صورتیکه دقت کمیتهای x , y , z به ترتیب برابر $\delta_z, \delta_y, \delta_x$ می‌باشند.

مثال۱) ابعاد یک زمین مستطیل شکلی به صورت زیر می‌باشد مطلوب است محاسبه دقت تعیین مساحت این مستطیل؟

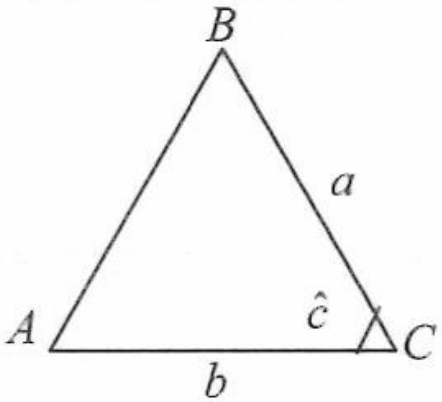
$$\begin{cases} a = 100m \pm 0.02m \\ b = 50m \pm 0.01m \end{cases}$$

$$s = a \times b$$

$$\delta_s = \pm\sqrt{b^2 \times \delta_a^2 + a^2 \times \delta_b^2} \Rightarrow \delta_s = \pm\sqrt{(50)^2 \times (\frac{2}{100})^2 + (100)^2 \times (\frac{1}{100})^2} = \pm\sqrt{2}m^2$$

مثال۲) دو ضلع و یک زاویه مثلثی به همراه خطاهای استاندارد آنها در زیر آمده است مطلوب است محاسبه مساحت مثلث و خطای استاندارد مساحت مثلث؟

$$a = 155.25m, \quad \delta_a = \pm 0.03\,m$$
$$b = 71.25m, \quad \delta_b = \pm 0.02m$$
$$C = 40^\circ 20', \quad \delta_C = \pm 20''$$



$$S = \frac{1}{2} ab \sin C = \frac{1}{2} \times 155.25 \times 71.25 \times \sin 40^\circ 20' = 3579.71 m^2$$

$$\frac{\partial S}{\partial a} = \frac{1}{2} b \sin C = \frac{1}{2} \times 71.25 \times \sin 40^\circ 20' = 23.06m$$

$$\frac{\partial S}{\partial b} = \frac{1}{2} a \sin C = \frac{1}{2} \times 155.25 \times \sin 40^\circ 20' = 50.24m$$

$$\frac{\partial S}{\partial C} = \frac{1}{2} ab \cos C = \frac{1}{2} \times 71.25 \times 155.25 \times \cos 40^\circ 20' = 4216.07m$$

$$\delta_C = 20'' = \frac{20}{206265} rad = 0.000097 rad$$

$$\delta_S = \pm\sqrt{\left(\frac{\partial S}{\partial a}\delta_a\right)^2 + \left(\frac{\partial S}{\partial b}\delta_b\right)^2 + \left(\frac{\partial S}{\partial C}\delta_C\right)^2}$$

$$\delta_S = \pm\sqrt{\left(\frac{1}{2}b\sin C \times \delta_a\right)^2 + \left(\frac{1}{2}a\sin C \times \delta_b\right)^2 + \left(\frac{1}{2}ab\cos C \times \delta_C\right)^2}$$

$$\Rightarrow \delta_S = \pm\sqrt{23.06^2 \times 0.03^2 + 50.24^2 \times 0.02^2 + 4216.07^2 \times 0.000097^2} \Rightarrow \delta_S = \pm 1.29m^2$$

حالت‌های خاص

۳-۱۰-۱ خطای مجموع دو یاچندکمیت

هرگاه کمیت S از مجموع چند کمیت دیگر با مقادیر x ، y و z و ... به دست آمده باشد. (مثل وقتی که بخواهیم مسافتی را که طولش نزدیک به ۱۰۰ متر است با استفاده از یک نوار ۳۰ متری اندازه بگیریم). در این حالت

$$s = x+y+z+\ldots$$

هم‌چنین :

$$\frac{\partial s}{\partial x} = ۱ \qquad \frac{\partial s}{\partial y} = ۱ \qquad \frac{\partial s}{\partial z} = ۱$$

و خطای معیارطبق رابطهٔ ۳-۲۰ به صورت زیر محاسبه می‌شود .

$$\delta A = \sqrt{(\frac{\partial F}{\partial x})^۲.(\delta x)^۲ + (\frac{\partial F}{\partial y}).^۲(\delta y)^۲ + (\frac{\partial F}{\partial z})^۲.(\delta z)^۲} \longrightarrow \delta = \sqrt{\delta_x^۲ + \delta_y^۲ + \delta_z^۲ + \ldots} \qquad (۲۱-۳)$$

در حالتی که $\delta_x = \delta_y = \delta_z = \delta$ یعنی خطای کمیت‌ها یکسان باشد خطای مجموع می‌شود.

$$\delta_S = \delta\sqrt{n} \qquad (۲۲-۳)$$

که درآن n تعداد کمیت‌ها و δ خطای مشترک آنها است.

مثال ۳-۵ مطلوب است محاسبهٔ خطای اندازه‌گیری یک مسافت ۴۵۰ متری که با یک نوار فلزی ۵۰ متری اندازه‌گیری شده و خطای معیار هر دهانهٔ اندازه‌گیری ۲ میلی‌متر باشد.

تعداد دهانه‌ها ۴۵۰÷۵۰ =۹

$$\delta = ۲ \times \sqrt{۹} = ۶mm$$

اگراین‌خطا، سیستماتیک باشد یعنی طول حقیقی نوار از طول اسمی آن ۲ میلی‌متر کوتاه‌تر یا بلندتر باشد اثر این خطا در کل مسافت به مقدار

$$n \times e = ۹ \times ۲ = ۱۸mm$$

ظاهر خواهد شد که اگر طول حقیقی نوار بیشتر از طول اسمی آن باشد عدد حاصل به مسافت قرائت شده اضافه می‌شود و اگر طول حقیقی از طول اسمی کمتر باشد مقدار مذکور از فاصلهٔ قرائت شده کم می‌شود.

به طور کلی می‌توان گفت که حاصل جمع خطاهای سیستماتیک به صورت $n \times e$ در نتیجهٔ کل ظاهر می‌شود درحالی که حاصل جمع خطاهای تصادفی $\pm e\sqrt{n}$ است.

۳-۱۰-۲ خطای تفاضل دوکمیت

فرض می‌کنیم:

$$d = x - y$$

در این صورت:

$$\frac{\partial d}{\partial y} = -۱ \qquad \frac{\partial d}{\partial x} = ۱$$

و خطای معیار طبق فرمول۳-۲۰ می‌شود :

$$\delta = \sqrt{(\delta x)^۲ + (\delta y)^۲} \qquad (۲۳-۳)$$

و اگر فرض کنیم

$$\delta_x = \delta_y = \delta$$

در این صورت

$$\delta_d = \sqrt{\delta_x^۲ + \delta_y^۲} = \delta\sqrt{۲} \qquad (۲۴-۳)$$

مثال ۳-۶ : مطلوب است محاسبهٔ خطای اندازه‌گیری یک خط به طول تقریبی ۲۵ متر که با یک نوار ۳۰ متری انجام شده و خطای قرائت ابتدا و انتهای نوار ۲ میلی‌متر باشد .

حل: چون طول این خط از تفاضل قرائت ابتدا و انتهای نوار به دست می آید لـذا اگـر قرائـت متر را در ابتدا و انتهای نوار به ترتیب x و y فرض کنیم، معادلهٔ اندازه‌گیری به صورت

$$d = y-x$$

است وخطای اندازه‌گیری طبق رابطهٔ ۳-۲۳ می‌شود:

$$۲ \times \sqrt{۲} = ۲/۸\,mm$$

۳-۱۰-۳ خطای حاصل‌ضرب دویا چند کمیت

هرگاه کمیت P از حاصل‌ضرب دو یا چند کمیت مستقیم x ، y و z به دست آمده باشد یعنی داشته باشیم ،

$$P = x.y.z$$

در این صورت خطای معیار کمیت طبق رابطهٔ ۳-۲۰ عبارت است از :

$$\delta_P = \sqrt{(y.z.\delta_x)^2 + (x.z.\delta_y)^2 + (x.y.\delta_z)^2} = P.\sqrt{(\frac{\delta_x}{x})^2 + (\frac{\delta_y}{y})^2 + (\frac{\delta_z}{z})^2}$$

(۳-۲۵)

و خطای نسبی کمینت می‌شود.

$$\frac{\delta_P}{P} = \sqrt{(\frac{\delta_x}{x})^2 + (\frac{\delta_y}{y})^2 + (\frac{\delta_z}{z})^2} \qquad (۳-۲۶)$$

مثال۳-۸ مطلوب است تعیین خطای مساحت زمین مستطیل شکلی که طول و عرض آن به ترتیب ۱۲۵٫۴۵ متر و ۸۲٫۴۸ متر و خطای معیار آنها به ترتیب ۳ سانتی‌متر و ۱ سانتی‌متر است.

حل : معادلهٔ اندازه‌گیری مساحت زمین:

$$A = x.y = ۱۰۳۴۷٫۱۲$$

است و

$$\delta_A = \sqrt{(y.\delta_x)^۲ + (x.\delta_y)^۲} = \sqrt{(۸۲٫۴۸ \times ۰٫۰۳)^۲ + (۱۲۵٫۴۵ \times ۰٫۰۱)^۲} = ۲٫۷۷\, m^۲$$

بنابراین نتیجهٔ اندازه‌گیری مساحت را می‌توان چنین نوشت .

$$A = (۱۰۳۴۷٫۱۲ \pm ۲٫۷۷)\, m^۲$$

مثال ۳-۹ برای تعیین مساحت زمینی که به شکل مثلث است اندازه‌گیری‌های زیرانجام شده است. مساحت زمین وخطای نسبی آن را حساب کنید.

a=۲۳۵٫۷۰ m ±۰٫۰۵ m	b=۱۵۷٫۹۰ m ±۰٫۰۲ m	$\hat{C}$ =۶۱° , ۳۰′ ±۲′

حل:

معادلهٔ اندازه‌گیری

$$A = \frac{۱}{۲} a.b.\sin\hat{C}$$

$$A = ۱۶۳۵۳٫۴۸\ m^۲$$

و هم‌چنین :

$$\delta_A = \sqrt{(\frac{\partial A}{\partial a}\delta_a)^۲ + (\frac{\partial A}{\partial b}\delta_b)^۲ + (\frac{\partial A}{\partial \hat{C}}\delta_{\hat{C}})^۲}$$

در اینجا

$$\delta_a = ۰/۰۵ \text{ m}$$

$$\delta_b = ۰/۰۲ \text{ m}$$

$$\delta\hat{C} = ۲ \times ۲/۹ \times ۱۰^{-۴} = ۵/۸ \times ۱۰^{-۴} \text{ radian}$$

$$\frac{\partial A}{\partial a} = \frac{۱}{۲} b.\sin\hat{C} = ۶۹/۳۸$$

$$\frac{\partial A}{\partial b} = \frac{۱}{۲} a.\sin\hat{C} = ۱۰۳/۵۷$$

$$\frac{\partial A}{\partial \hat{C}} = \frac{۱}{۲} a.b.\cos\hat{C} = ۸۸۷۹$$

$$\delta A = ٦/۵۵ \text{ m}^۲$$

و خطای نسبی

$$\frac{\delta A}{A} \cong \frac{۱}{۲۵۰۰}$$

۳-۱۰-۴ خطای خارج قسمت دو کمیت

فرض می‌کنیم :

$$Q = \frac{y}{x}$$

در این صورت :

$$\delta Q = \sqrt{\left(-\frac{y}{x^{۲}}\delta x\right)^{۲} + \left(\frac{۱}{x}\delta y\right)^{۲}} = Q\sqrt{\left(\frac{\delta x}{x}\right)^{۲} + \left(\frac{\delta y}{y}\right)^{۲}} \qquad ۳-۲۵$$

مثال ۳-۱۰ فاصلهٔ افقی و اختلاف ارتفاع بین دونقطهٔ A و B به ترتیب ۱۵۰ متر و ۱۲ متر است. اگر خطای معیار فاصلهٔ افقی ۲۰ سانتی‌متر و خطای معیار اختلاف ارتفاع ۵ سانتی‌متر فرض شود خطای شیب خط AB چقدر است؟

$$p = \frac{h}{d} = \frac{۱۲}{۱۵۰} = ۰/۰۸ = ٪۸ \qquad \text{شیب خط}$$

$$\delta p = p\sqrt{\left(\frac{\delta h}{h}\right)^{۲} + \left(\frac{\delta d}{d}\right)^{۲}} = ۰/۰۸ \times \sqrt{\left(\frac{۰/۰۵}{۱۲}\right)^{۲} + \left(\frac{۰/۲۰)}{۱۵۰}\right)^{۲}} = ۳/۵ \times ۱۰^{-۴} = ٪۰/۰۳۵$$

سوال ۱) برای اندازه گیری فاصله افقی و اختلاف ارتفاع بین دو نقطه a و b فاصله شیب دار بین a و b برابر 100m با دقت $\pm 5cm$ و زاویه شیب برابر 30° با دقت یک دقیقه اندازه گیری شده است. دقتهای فاصله افقی و اختلاف ارتفاع را تعیین کنید؟

سوال ۳) مساحت قطعه زمینی به شکل مستطیل با ابعاد a=120 m و b=50 m را می خواهیم با دقت 4 متر مربع به دست آوریم. طول ها باید با چه دقت نسبی اندازه گیری شوند؟

پایان جلسه هشتم

درس تئوری خطاها

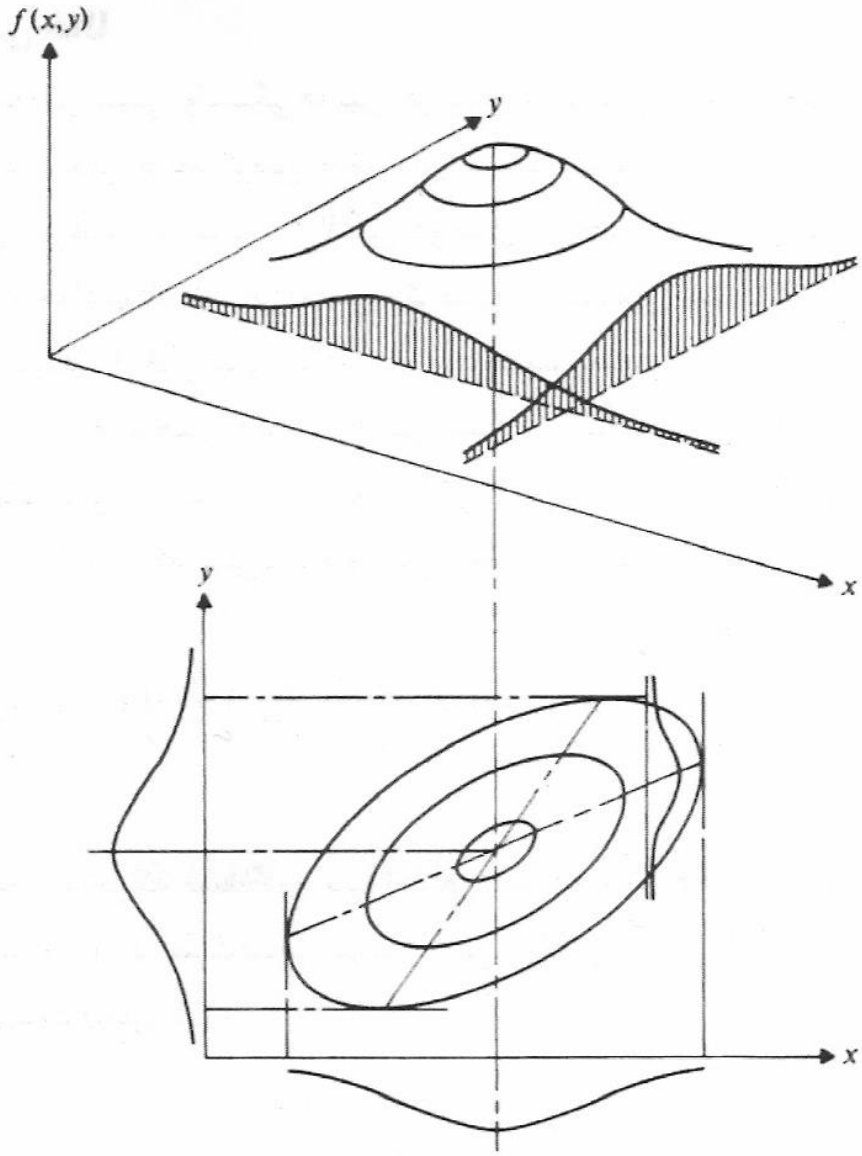
جلسه نهم

فرید اسماعیلی
Farid_63@yahoo.com
www.faridesm.ir

تماس با استاد از طریق پست الکترونیکی
مشاهده اطلاعیه ها، نمرات، دریافت فایل ها در وب سایت

برای بیان توام عدم‌قطعیت یا خطای چندمتغیر تصادفی از کمیت‌های واریانس و کوواریانس بین آن‌ها استفاده‌می‌شود. کمیت‌های بالا عددی‌اند و تصویری روشن، به‌ویژه هنگامی‌که با متغیرهای مکانی[1] سروکارداریم، برای ما فراهم‌نمی‌سازند. به‌همین‌منظور از یک‌نمایش گرافیکی[2] برای متغیرهای تصادفی توام استفاده‌می‌شود که چگونگی توزیع خطا اطراف نقطهٔ برخورد آن‌ها را نشان‌می‌دهد. معمولا این‌نمایش گرافیکی، در حالت دوبعدی، بابیضی خطای دوبعدی[3]، و در حالت سه‌بعدی با بیضوی[4] انجام‌می‌گیرد.

ازآن‌جاکه به‌طور سنتی وابستگی‌ها بین هردومتغیر تصادفی ارائه‌می‌شوند، لذا ساده‌ترین حالت ممکن، یعنی درنظرگرفتن متغیرهای تصادفی به‌صورت دو به‌دو، بررسی‌می‌شود. بنابراین در این‌**قسمت با** نمایش گرافیکی دوبعدی از پخش خطا برای دومتغیر تصادفی توام سروکارخواهیم‌داشت. این‌نمایش گرافیکی یک‌بیضی است که درواقع بیان‌گر یک‌سطح اطمینان از قرارگیری مقادیر واقعی دومتغیر تصادفی به‌صورت توام در آن‌است.



نگارهٔ ۱.۵. نمایش تابع چگالی نرمال دومتغیره و تصویر برش دومقطع عمودبرهم

(E.M. Mikhail & F. Ackermann, 1976)

بیضی خطای مطلق هر نقطه از طریق زیر ماتریسی از ماتریس وریانس–کوریانس مجهولات برآورد شده که دارای واریانس و کوریانس مختصات نقطه مورد نظر است، به دست می‌آید. توجیه این بیضی، بگونه‌ای است که قطر اطول و اقصر آن در جهت بیشترین و کمترین مقدار خطای مختصات و اندازه قطر اطول و اقصر آن نیز برابر بیشترین و کمترین مقدار وریانس موقعیت نقطه مورد نظر است. طول و جهت اقطار بیضی خطا را می‌توان از روابط زیر به دست آورد :

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

معادله مشخصه این ماتریس بصورت زیر می‌باشند:

$$\begin{cases} (-\lambda)^2 + trace(\Sigma)(-\lambda) + |\Sigma| = 0 \\ trace(\Sigma) = \sigma_1^2 + \sigma_2^2 \\ |\Sigma| = \sigma_1^2\sigma_2^2 - \sigma_{12}^2 \end{cases} \Rightarrow \lambda^2 - (\sigma_1^2 + \sigma_2^2)\lambda + (\sigma_1^2\sigma_2^2 - \sigma_{12}^2) = 0$$

$$\lambda_1 = \frac{\sigma_x{}^2 + \sigma_y{}^2}{2} + \frac{1}{2}\sqrt{(\sigma_x{}^2 - \sigma_y{}^2)^2 + 4\sigma_{xy}{}^2}$$

$$\lambda_2 = \frac{\sigma_x{}^2 + \sigma_y{}^2}{2} - \frac{1}{2}\sqrt{(\sigma_x{}^2 - \sigma_y{}^2)^2 + 4\sigma_{xy}{}^2}$$

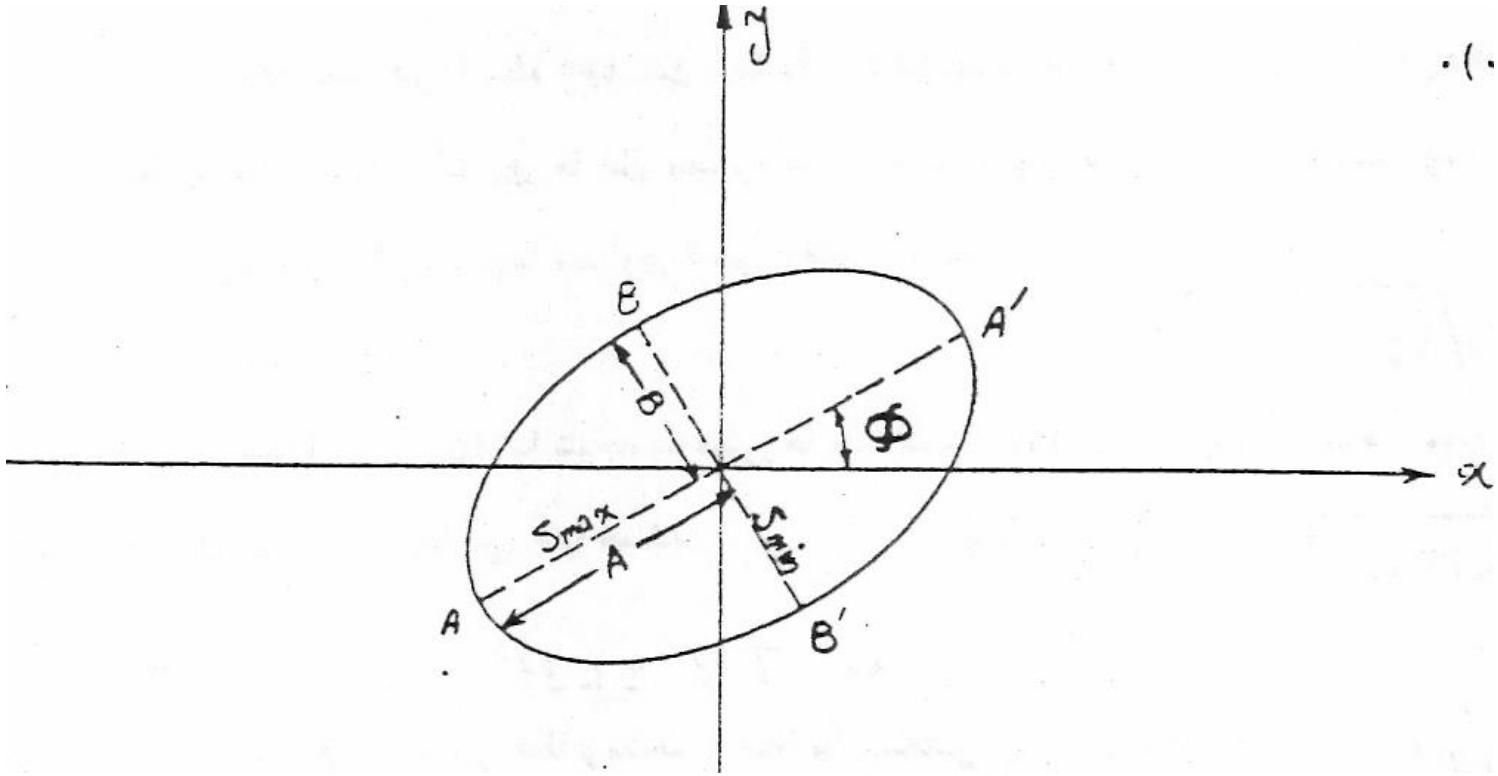
$$\mathrm{tg}\, 2\gamma = \frac{2\sigma_{xy}}{\sigma_x{}^2 - \sigma_y^2}$$

γ زاویه بین نیم قطر اطول بیضی و محور xها می‌باشد.

رابطه بین مقادیر ویژه ماتریس واریانس کواریانس و ابعاد بیضی خطا بصورت زیر می‌باشد:

$$a = \sqrt{\lambda_1}, \quad b = \sqrt{\lambda_2}$$

بیضی خطای حاصل را در یک فاصله اطمینان 39.4% به مرکز موقعیت واقعی نقطه مورد نظر است. اما عملا این بیضی را بخاطر عدم دسترسی به موقعیت واقعی، به مرکز موقعیت برآورد شده رسم می‌کنند. اگر چه بیضی خطای مطلق شمای کلی از دقت شبکه را به دست می‌دهد اما وابسته به سیستم مختصات است، زیرا مختصات، کمیتهای inestimable هستند.



سوال ۵) به منظور تعیین مختصات سه بعدی نقطه A، روشی اتخاذ شده است که در آن مولفه ارتفاعی مستقل از مولفه‌های مسطحاتی اندازه‌گیری می شود. در جدول زیر نتیجه اندازه‌گیری مختصات نقطه A که در آن 10 نوبت اندازه‌گیری انجام پذیرفته است مطلوبست:

الف) تشکیل ماتریس واریانس – کواریانس مختصات نقطه A.

ب) تعیین ابعاد بیضی خطای استاندارد موقعیت مسطحاتی نقطه A (نیم قطر اطول، نیم قطر اقطار و زاویه انحراف بیضی)

ج) تعیین ضریب گسترش (k) بیضوی خطای 90% نسبت به بیضوی خطای استاندارد.

NO	X(m)	NO	Y(M)	NO	Z(m)
1	10.017	1	50.016	1	100.110
2	10.015	2	50.011	2	100.120
3	10.010	3	50.007	3	100.115
4	10.017	4	50.018	4	100.130
5	10.020	5	50.017	5	100.113
6	10.016	6	50.011	6	100.125
7	10.021	7	50.009	7	100.133
8	10.018	8	50.008	8	100.118
9	10.029	9	50.015	9	100.130
10	10.006	10	50.010	10	100.120

الف) برای تشکیل ماتریس واریانس– کواریانس مختصات نقطه A به صورت زیر عمل می‌کنیم:

$$\bar{X} = \frac{\sum X_i}{10} = 10.0169$$

$$\bar{Y} = \frac{\sum Y_i}{10} = 50.0122$$

$$\bar{Z} = \frac{\sum Z_i}{10} = 100.1214$$

$$\sigma_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \rightarrow \sigma_X^2 = 0.00003832$$

$$\sigma_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 \rightarrow \sigma_Y^2 = 0.00001573$$

$$\sigma_Z^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Z_i - \bar{Z})^2 \rightarrow \sigma_Z^2 = 0.00006137$$

$$\sigma_{X,Y} = \frac{1}{n-1}\sum_{k=1}^{n}(X_i - \bar{X}) \times (Y_i - \bar{Y}) \rightarrow \sigma_{X,Y} = 0.00001057$$

$$\sigma_{X,Z} = \frac{1}{n-1}\sum_{k=1}^{n}(X_i - \bar{X}) \times (Z_i - \bar{Z}) \rightarrow \sigma_{X,Z} = 0$$

$$\sigma_{Y,Z} = \frac{1}{n-1}\sum_{k=1}^{n}(Y_i - \bar{Y}) \times (Z_i - \bar{Z}) \rightarrow \sigma_{Y,Z} = 0$$

$$\sum\nolimits_{\mathrm{A}} = \begin{bmatrix} 0.00003832 & 0.00001057 & 0 \\ 0.00001057 & 0.00001573 & 0 \\ 0 & 0 & 0.00006137 \end{bmatrix}$$

$$\begin{bmatrix} \sigma_X^2 & \sigma_{X,Y} \\ \sigma_{X,Y} & \sigma_Y^2 \end{bmatrix} = \begin{bmatrix} 0.00003832 & 0.00001057 \\ 0.00001057 & 0.00001573 \end{bmatrix}$$

$$\lambda_1 = a^2 = \frac{\sigma_X^2 + \sigma_Y^2}{2} + \frac{1}{2}\sqrt{(\sigma_X^2 - \sigma_Y^2)^2 + 4{\sigma_{X,Y}}^2}$$

$$\lambda_2 = b^2 = \frac{\sigma_X^2 + \sigma_Y^2}{2} - \frac{1}{2}\sqrt{(\sigma_X^2 - \sigma_Y^2)^2 + 4{\sigma_{X,Y}}^2}$$

$$\lambda_1 = a^2 = \frac{0.00003832 + 0.00001573}{2} + \frac{1}{2}\sqrt{(0.00003832 - 0.00001573)^2 + 4 \times (0.00001057)^2}$$
$$= 0.00004249$$

$$\lambda_2 = b^2 = \frac{0.00003832 + 0.00001573}{2} - \frac{1}{2}\sqrt{(0.00003832 - 0.00001573)^2 + 4 \times (0.00001057)^2}$$
$$= 0.00001155$$

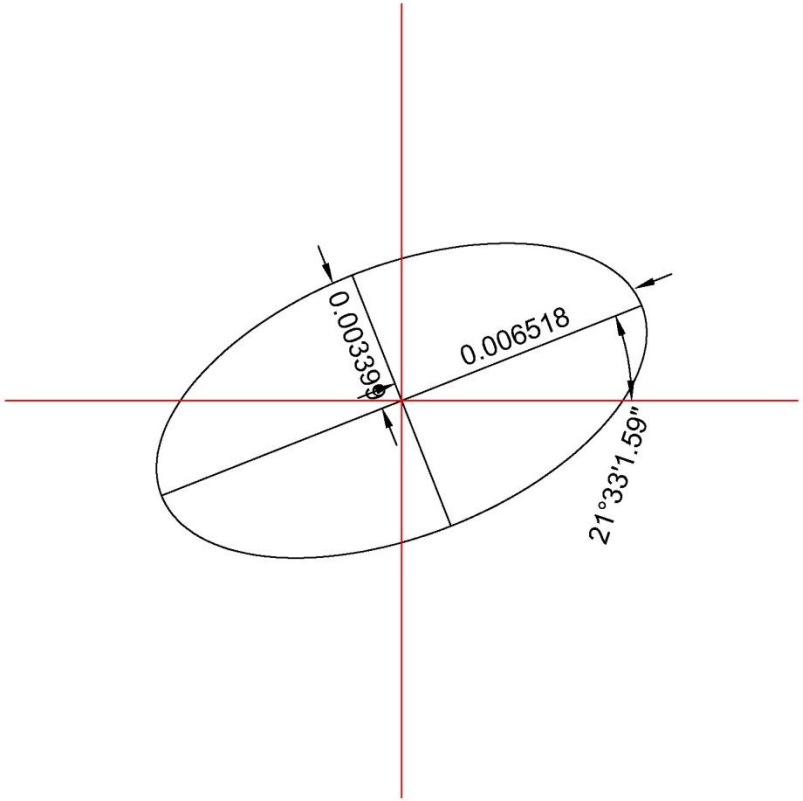
$$a = \sqrt{\lambda_1} = \sqrt{0.00004249} = 0.006518$$

$$b = \sqrt{\lambda_2} = \sqrt{0.00001155} = 0.003399$$

ابعاد بیضی خطای استاندارد
٪ ۳۹.۴

$$\tan(2\gamma) = \frac{2\sigma_{X,Y}}{\sigma_X^2 - \sigma_Y^2} \rightarrow \tan(2\gamma) = 0.93581230$$

$$\rightarrow \gamma = 21^\circ\, 33'1.59''$$



بیضی خطای استاندارد ٪ ۳۹.۴

1-α=0.90 => α=0.10 =>

α/2=0.05 =>

C=2.5002

از جدول توزیع کای اسکور

ابعاد بیضی خطای ۹۰ ٪

$$a\ (90\%) = c\sqrt{\lambda_1} = 2.5002 * \sqrt{0.00004249} = 0.01629$$

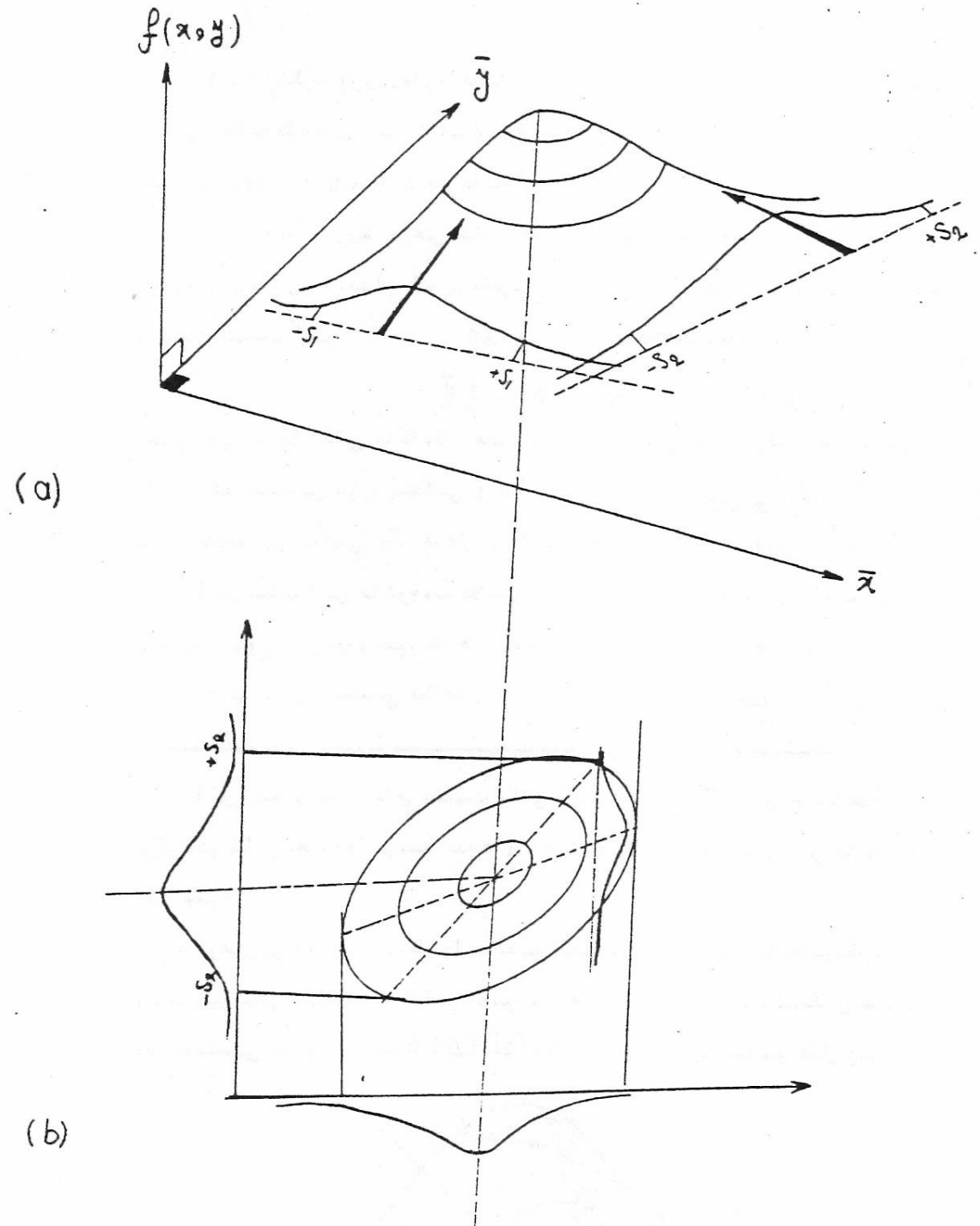
$$b(90\%) = c\sqrt{\lambda_2} = 2.5002 * \sqrt{0.00001155} = 0.008498$$

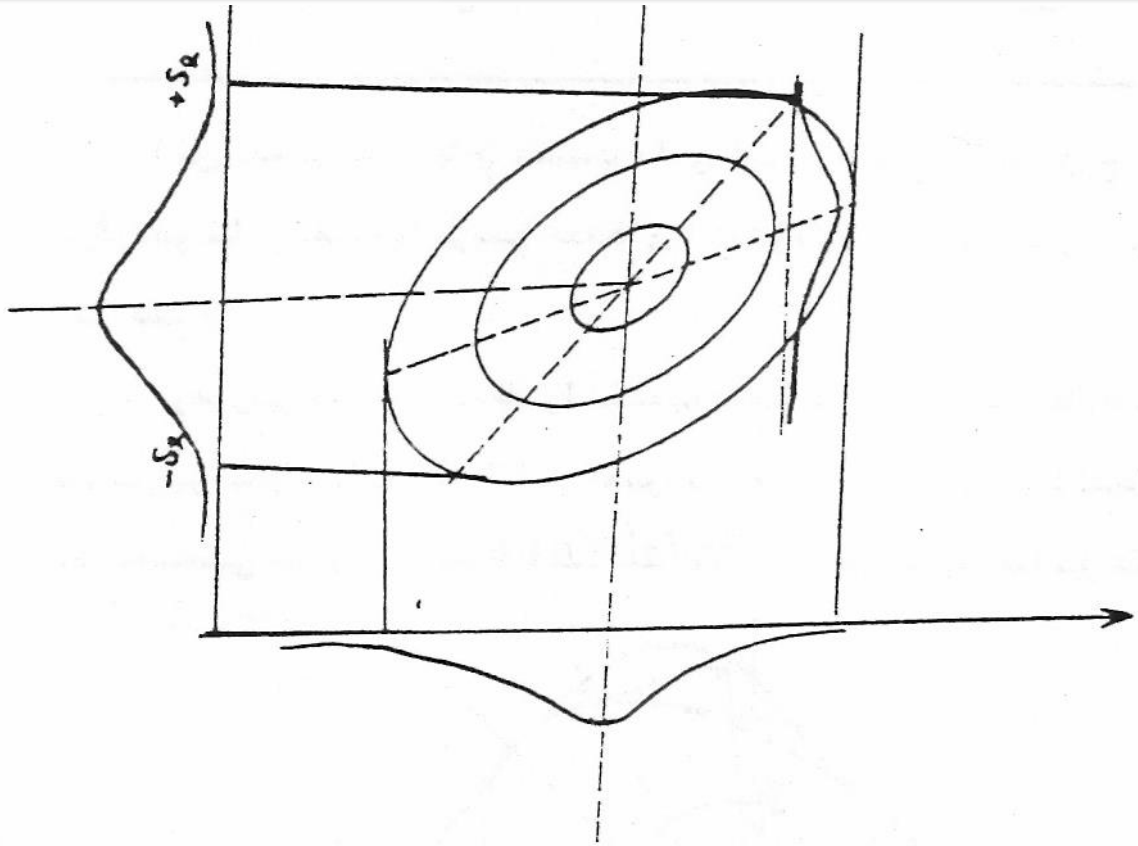
۲ –۸ –۵ ویژگیها ی بیضی خطا ها

۱– با توجه به مقا دیرقطربزرگ وقطرکوچک بیضی که درروا بط قبل دا ده شد درصورتیکه وریا نس ا ندا زه‌گیری یک زوج یکسا ن با شدا ولا S_1^2 مسا وی S_2^2 خوا هدبود ثا نیا "گوویلانس برا برصفرثا لثا " ا قطا ربیضی با هم‌برا برخوا هندشد . درنتیجه بیضی تبدیل به‌دا یره‌می گردد ودرچنین حا لتی‌مکا ن هندسی خطا مکا نی یکسا ن (Homogenecus)بطوریکه تغییرا ت خطا درحدودا حتما لوقوع – آ ن ا زروند معین وثا بتی به‌ا ندا زه‌شعا ع دا یره خطا پیروی خوا هدکرد .

بعبا رت دیگرجنا نچه S_1 برا بر S_2 نبا شد تلفیق شکل توا بع پخش ا ین دوخطا بصورت زیرخوا هد بودکه درا ین شکل فا صله از $+S_1$ تا $-S_1$ برا برطول قطربزرگ و $+S_2$ تا $-S_2$ برا برطول کوچک بیضی خوا هدبود :

خواص بیضی خطاهای مطلق





(شکل ۶ ـ ۵).

وچنانچه فاصله $-S_1$ تا $+S_1$ برابر $-S_2$ تا $+S_2$ باشد ($S_1 = S_2$)درا ینصورت شکل فوق (a) دارای مقاطع دایره‌ای درحالت (b) خواهدبود.

۲ ـ درحالتی که مقادیر A و B بهمدیگرمتفاوت باشند درا ینصورت قطربزرگ نشان دهنده خطای ماگزیمم وقطرکوچک نشان دهنده خطای‌می‌نیمم خواهدبود.

۳ ـ از نظر تئوری آمار و احتمالات و مباحث گذشته در زمینه محاسبه احتمال سطح زیر منحنی خطا فاصله‌های $-\sigma$ تا $+\sigma$ و -2σ تا $+2\sigma$ و تا $+2\sigma$ و -3σ تا $+3\sigma$ به ترتیب دارای مقادیر ۶۸٪ ، ۹۵٪ و ۹۹٪ می باشد.

۴ ـ مقدار زاویه چرخش بیضی خطا همانگونه که ملاحظه شد بستگی به مقادیر واریانس و ـ کوواریانس یک زوج اندازه‌گیری دارد بطوری که درحالت های مختلف ممکن است مقادیر خاصی را پیدا کند مثلاً اگر $\sigma_{12} = 0$ باشد نتیجه میشود:

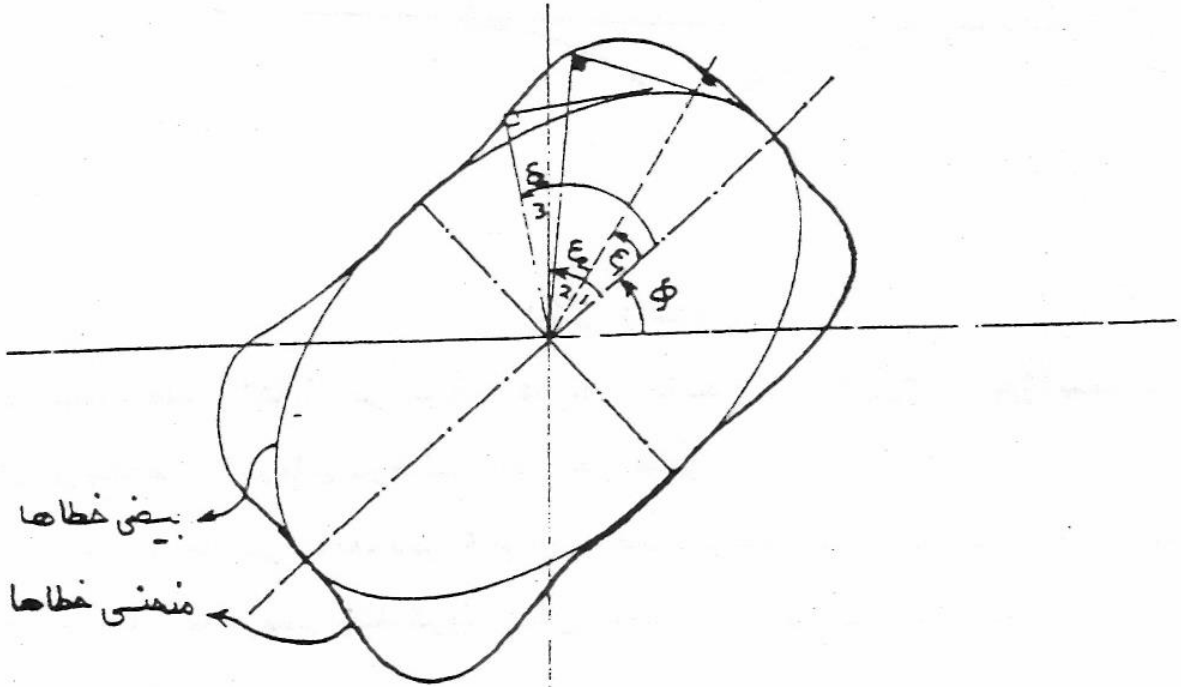
$$\tan 2\phi = 0 \Rightarrow 2\phi = 0 \text{ یا } \pi \Rightarrow \phi = 0 \text{ یا } \frac{\pi}{2}$$

یعنی بیضی کاملاً افقی یا کاملاً عمودی است. یعنی قطر بزرگ بیضی منطبق بر محورهای x یا y باشد. همچنین در رابطه کلی (۳۰ ـ ۵) در واقع $\tan 2\phi = \sin 2\phi / \cos 2\phi$ میباشد که درنتیجه صورت کسر σ_{12} هم‌علامت با $\sin 2\phi$ ومخرج آن یعنی $\sigma_2^2 - \sigma_1^2$ با $\cos 2\phi$ هم علامت است بنابراین با توجه به علامتهای صورت ومخرج کسر مکان زاویه ϕ درروی شکل مشخص میباشد (شروع زاویه از محور x ها درجهت عکس عقربه های ساعت).

۳ ـ ۸ ـ ۵ ـ منحنی خطا ها (Pedal Curve).

این منحنی برمبنای وضعیت خطای اندازه‌گیری دریک زوج مشاهده بروی بیضی خطا ترسیم می‌گرددونشان دهنده ماگزیمم حدخطای اندازه‌گیری های یکزوج می باشدونحوه‌ترسیم آن بصورت زیراست :

درصورتی که امتداد خطا را نقطه‌به‌نقطه درروی بیضی خطا ها درنظربگیریم ودرهرنقطه ـ مماسی بربیضی چنان رسم کنیم که‌برادامه این امتداد عمودباشدمکان هندسی پای عمودها تشکیل دهنده‌منحنی خطا یا Pedal Curve خواهدبودمطابق شکل زیر :



چنانچه ملاحظه شد مقادیر ما گزیمم و می نیمم خطای یک زوج اندازه‌گیری به ترتیب منطبق بر اقطار بزرگ وکوچک بیضی می‌باشد لیکن امکان دارد خطا مقداری بین این دوحدنیزدریک اندازه گیری داشته باشد بنابراین تمام سطح‌بیضی‌خطا می توانند محدوده‌نشان دهنده خطای اندازه گیریها باشد حال اگردریک سری اندازه‌گیری برای یک زوج مقادیرخطا ها به‌ترتیب e_1 با زاویه ε_1 نسبت به‌محوربزرگ بیضی و e_2 با زاویه ε_2 نسبت به‌همین محورو.....و e_n با زاویه ε_n باشدحدجا بجائی ما گزیمم خطای درروی منحنی خطا به‌روش ذکرشده بدست خواهد آمد. ملاحظه میگرددکه درحالت خاص نیزانتهای قطرهای کوچک وبزرگ بیضی روی منحنی خطا قرار دارد. مقدارخطا درهرحالت با توجه به زاویه آن نسبت به‌قطربزرگ ازرابطه زیرقابل‌محاسبه است.

$$S_{\varepsilon_i}^2 = a^2 \cos(\varepsilon_i^2) + b^2 \sin(\varepsilon_i^2) \qquad (۵-۳۱)$$

توجیه این رابطه‌بدین صورت است که اگرمقدارخطای روی‌قطارمعینی تصویرکنیم به‌ترتیب تصاویرآن $a\cos(\varepsilon_i)$ و $b\sin(\varepsilon_i)$ خواهدبودکه مولفه این دومقدارراابطه‌بالاست.

(a و b به‌ترتیب قطعاتی از A و B می‌باشند).

پایان جلسه نهم