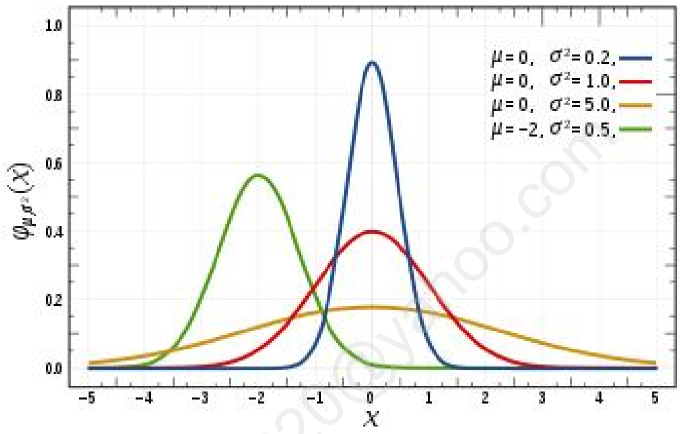
سمینار مجازی مقدمه ای بر استفاده از نرم افزار Excel و Matlab در آمار

27 بهمن 1399 دانشگاه و موسسه آموزش عالی زند شیراز امین صدیقی

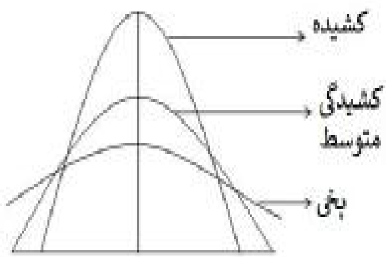
مقدمه کوتاه در آمار

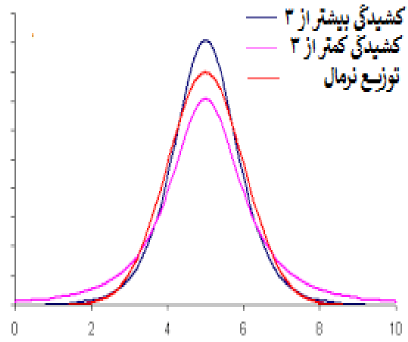
توزیع نرمال – نرمال استاندارد



منحنی توزیع نرمال

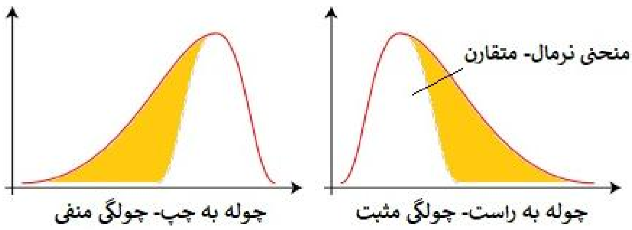
کشیدگی (Kurtosis)

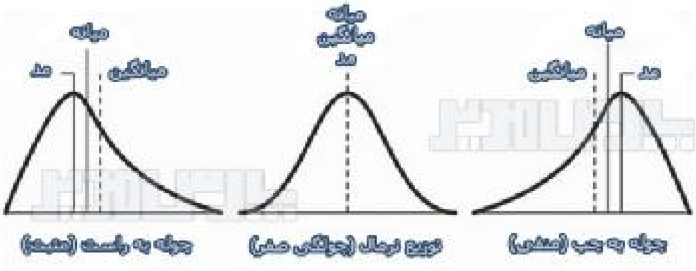




مقدار کشیدگی برای توزیع نرمال برابر ۳ می باشد
****

چولگی (Skewness)





چنانچه چولگی و کشیدگی در بازه (۲ ، ۲-) نباشند داده‌ها از توزیع نرمال برخوردار نیستند
****

سه–میانگین

این میانگین توسط آماردان برجسته آمریکایی «جان توکی» ابداع شده است. سه-میانگین با عبارت «Trimean» نشان داده می‌شود و شیوه محاسبه آن طبق رابطه زیر است:

Trimean=(Q1+2*Q2+Q3)/4

که در آن Q1 چارک اول، Q2 چارک دوم و Q3 چارک سوم هستند. این شیوه برای محاسبه میانگین، به نقاط مرکزی بیش از نقاط انتهایی اهمیت می‌دهد.

مثال

مقدار سه-میانگین برای مقدارهای 1،5،6،4،10 طبق مراحل گفته شده در بالا محاسبه می‌شود.

مرتب سازی: 1،4،5،6،10

میانگین حسابی برای این اعداد برابر با 5.2 خواهد بود

Q۱=۴ Q۲=۵ Q۳=۶

Trimean=(4+2×5+6)/4=20/4=5

درحالیکه میانگین حسابی برای این اعداد برابر با 5.2 خواهد بود. تاثیر مقدار ۱ و ۱۰ که نسبت به بقیه داده‌ها دور هستند، در میانگین حسابی دیده می‌شود.

ضریب چولگی گشتاوری پیرسون

اگر X یک متغیر تصادفی با میانگین μ و انحراف معیار σ
باشد ضریب چولگی گشتاوری به صورت زیر محاسبه می‌شود.

$$\gamma_1 = E[(\frac{X-\mu}{\sigma})^3] = \frac{\mu_3}{\sigma^3}$$

که در آن μ_3 گشتاور مرکزی سوم است.

بنابراین اگر لازم باشد براساس یک نمونه آماری، ضریب چولگی گشتاوری را بدست آورد باید واریانس و گشتاور مرکزی سوم نمونه‌ای را مبنا قرار داد. پس فرمول زیر را برای محاسبه ضریب چولگی گشتاوری نمونه‌ای پیرسون خواهیم داشت.

$$b_1 = \frac{m_3}{s^3} = \frac{\frac{1}{n}\sum(x_i-\bar{x})^3}{(\frac{1}{n-1}\sum(x_i-\bar{x})^2)^{\frac{3}{2}}}$$

ضریب چولگی اول پیرسون

ضریب چولگی اول پیرسون طبق رابطه زیر تعریف می‌شود که در آن نما مبنا در نظر گرفته شده است و انحراف میانگین از نما برحسب انحراف استاندارد محاسبه شده است.

$$\frac{Mean - Mode}{S}$$

البته گاهی به آن «چولگی نمای پیرسون» (Pearson Mode Skewness) نیز می‌گویند.

ضریب چولگی دوم پیرسون

ضریب چولگی دوم پیرسون اختلاف بین میانگین و میانه را مبنا قرار داده و نسبت آن را به انحراف استاندارد محاسبه می‌کند که گاهی آن را «چولگی میانه پیرسون» (Pearson Median Skewness) نیز می‌نامند.

$$\frac{3(Mean - Median)}{S}$$

نکته: اگر توزیع متقارن باشد میانگین و میانه برابر خواهند بود و ضریب گشتاوری دوم پیرسون برابر با صفر محاسبه خواهد شد.

چولگی برمبنای چارک‌ها

اگر از چندک‌ها برای محاسبه چولگی استفاده شود، شکل محاسباتی به صورت زیر خواهد بود. مشخص است که چارک دوم همان میانه است. این معیار توسط «آرتور بولی» (Arthur Bowley) دانشمند آماری در سال ۱۹۰۱ معرفی شده است.

$$B_1 = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

همانطور که دید می‌شود در مخرج کسر برای برآورد انحراف معیار از نصف فاصله بین چارک اول و سوم (دامنه میان چارکی) و در صورت کسر نیز میانگین چارک اول و سوم به عنوان برآورد میانگین محسوب شده است. زیرا با ساده کردن عبارت زیر به فرمول B1 خواهیم رسید.

$$B_1 = \frac{\frac{Q_3+Q_1}{2} - Q_2}{\frac{Q_3-Q_1}{2}}$$

چولگی G1

در محاسبه چولگی در بیشتر نرم‌افزارهای آماری نظیر SPSS، Minitab یا Excel از شیوه محاسبه زیر که توسط گیل (C. A. Gill) در سال 1998 ابداع شده، استفاده می‌شود.

$$G_1 = \frac{\sqrt{n(n-1)}}{n-2} \times \frac{\frac{1}{n}\sum(x_i-\bar{x})^3}{\left(\frac{1}{n-1}\sum(x_i-\bar{x})^2\right)^{\frac{3}{2}}}$$

به سادگی دیده می‌شود که بین G1 و b1 رابطه زیر برقرار است.

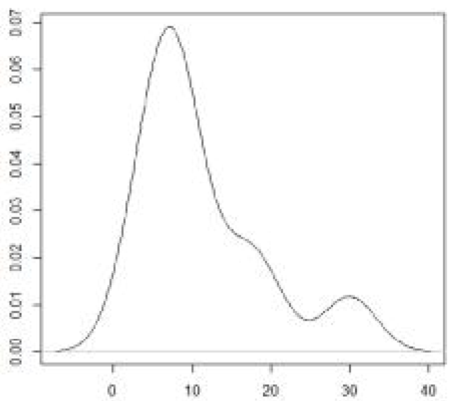
$$G_1 = \frac{n^2}{(n-1)(n-2)} b_1$$

ولی اگر تعداد مشاهدات زیاد باشد (n بزرگ باشد) مقدار G1 و b1 باهم برابر خواهند بود.

****

مثال

با توجه به داده‌های نمونه‌ای x=8,5,3,6,8,8,10,19,16,30 ضرایب چولگی پیرسون و چارکی طبق جدول زیر ارائه شده است. البته تصویر مربوط به منحنی توزیع احتمال این داده‌ها نیز در شکل دیده می‌شود.



روش محاسبه چولگی	چولگی گشتاوری	ضریب اول چولگی پیرسون	ضریب دوم چولگی پیرسون	چولگی برمبنای چارک
مقدار چولگی	$b_1 = 1.09995$ $G_1 = 1.52771$	0.4035	1.21038	0.59091

****

اگر داده ها نرمال نباشند

1) از آزمونهای پارامتری نمیتوان اســتفاده کرد زیرا آزمونهای پارامتری توان بیشــتری دارند خطای نوع دوم کمتر است کیفیت پایین میاید پس توان آزمون پایین میاید

2) نتایج آزمونهای پارامتری فاقد ارزش میشوند مثلا t

3) مثلا اگر در روش رگرسیون در حال اجرا باشیم باعث عدم اعتبار روش رگرسیون میشود باید یا خود داده یا باقیمانده داده نرمال باشند

****

بررسی پیشفرض نرمال بودن مانده‌ها

با ارائه مثالی مفهوم این که چرا باید باقی مانده های بدست آمده از مدل رگرسیون دارای توزیع نرمال باشند را توضیح داده میشود. فرض کنید نمرات ریاضی و آی‌کیو تعداد 10 نفر از دانش آموزان را به صورت فرضی داشته باشیم. این نمرات به شرح زیر هستند :

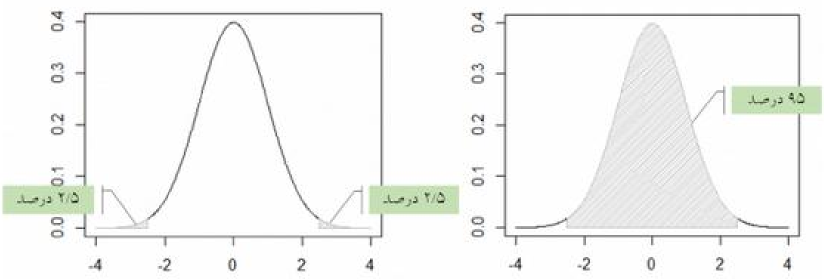
جدول (۱) داده های فرضی دانش آموزان

باقی‌مانده‌ها	مقادیر پیش بینی شده	نمرات ریاضی	نمرات آی‌کیو
-1.177	103.177	17.00	102.00
-0.528	107.528	18.00	107.00
-0.528	107.528	18.00	107.00
0.877	90.123	14.00	91.00
0.525	94.475	15.00	95.00
0.769	116.231	20.00	117.00
0.120	111.880	19.00	112.00
0.174	98.826	16.00	99.00
-1.177	103.177	17.00	102.00
0.945	114.055	19.50	115.00

به وسیله‌ی این داده‌ها مدل رگرسیونی برای پیش بینی نمرات آی کیو به توسط نمرات ریاضی دانش آموزان به وسیله مدل زیر برآورد می‌شود :

نمرات آی‌کیو = $29.206 + (4.351) *$ نمرات ریاضی

حال از طریق جای گذاری نمرات ریاضی دانش آموزان در فرمول فوق مقادیر پیش بینی شده نمرات آی کیو در جدول (1) بدست می‌آید. از تفریق مقادیر واقعی آی کیو از مقادیر پیش بینی شده، مانده ها یافت می شود (ستون آخر جدول (1)). حال مسئله اساسی این است که چرا این اعداد باید نرمال باشند ؟



شکل (۱) تابع چگالی توزیع نرمال

شکل (1) تابع چگالی توزیع نرمال را نشان می‌دهد. منحنی توزیع نرمال، زنگوله شکل است، نسبت به محور عمودی خود متقارن است و بیشتر داده‌ها را حول میانگین جای می‌دهد. با توجه به نمودار ملاحظه می‌شود میانگین برابر صفر است (وسط نمودار). همانطور که ملاحظه می‌کنید 95 داده‌ها حول صفر قرار دارند و فقط 5 درصد داده‌ها نسبت به میانگین اعدادی پرت هستند. حال اگر به مانده‌های جدول (1) نگاه کنیم می‌بینیم که همگی حول عدد صفر (میانگین) قرار دارند (اگر مانده‌ای عدد صفر اختیار کند بدین معناست که مقدار پیش‌بینی شده با مقدار واقعی برابر است.) بنابراین وقتی می‌گوییم یکی از پیشفرض‌های رگرسیون نرمال بودن مانده‌ها است بدین معنی است که اکثر مانده‌ها (95 درصد) نزدیک به صفر بوده و فقط اندکی از آن‌ها (5 درصد) از صفر دور باشند. به عبارت دیگر اکثر مقادیر پیش بینی شده نزدیک به مقادیر واقعی بوده و فقط اندکی از مقادیر پیش بینی شده با مقادیر واقعی تفاوت زیادی داشته باشد (بدین معنی که دقت پیش بینی بالا باشد).

****

نرم افزارها در آمار

SPSS	Minitab	Eviews	Statistica	SAS	R
Lisrel	Expert choice	NCSS	Microfit	STATA	
Excel	Mathlab				

در خصوص برخی نرم افزارها در زیر توضیحاتی ارائه میشود سپس نرم افزارها مقایسه میشوند

نرم افزار تحلیل آماری SAS

SAS مخفف Statistical Analysis System است. این نرم افزار آماری بیشتر روش -های رایج آماری را پوشش می- دهد و کاربر می -تواند فرمان- های خاص مورد نیازش را انتخاب نموده و به کار گیرد.

از مهم -ترین مزایای این نرم افزار می- توان به امکان برنامه نویسی برای تحلیل- های پیشرفته آماری بخصوص برنامه نویسی جهت تحلیل- های ماتریسی اشاره کرد.

نرم افزار تحلیل آماری Lisrel

لیزرل مخفف Linear Structural Relation به معنای ارتباطات ساختار خطی است، یک نرم افزار تحلیل آماری که برای مدل سازی معادلات ساختاری استفاده می- شود.

از لیزرل همچنین برای تست کردن روایی پرسشنامه استفاده می- شود. هدف لیزرل این است که ما را با انجام آزمونی به تحلیل مسیر، تحلیل عاملی و تحلیل خوشه- ای برساند.

نرم افزار تحلیل آماری Statistica

Statistica نیز یکی دیگر از نرم افزارهای معروف آماری است که قابلیت های منحصر به خود را دارد.

از آن-ها می -توان به امکان ارتباط بین عملیات در این نرم افزار با ویژوال بیسیک و انعطاف پذیری آن، قابلیت- های گرافیکی و نمودارهای زیبا و راهنمای کامل و کاربردی آن اشاره کرد.

نرم افزار تحلیل آماری Minitab

نرم افزار مینی تب از جمله نرم افزارهای مطرح آماری است که در برخی از زمینه- های علم آمار، از جمله کاربرد آمار در صنعت و در اقتصاد از قابلیت -های خاص و بالاتری برخوردار است.

رسم نمودارهای سه بعدی از مشاهدات، در این نرم افزار به راحتی امکان پذیر است. بسیاری از عملیات ها همانند تولید اعداد تصادفی که از توزیع- های آماری خاصی مانند توزیع نرمال، کی دو، گاما، اف و ... پیروی کند نیز به راحتی و سهولت در این نرم افزار امکان پذیر است.

مینی تب با داشتن خروجی -های بسیار دقیق و مناسب (در حد هزارم اعشار)، در زمینه- هایی از قبیل آمار توصیفی، رگرسیون، روش- های چند متغیره پیوسته و گسسته، طرح آزمایش -ها، کنترل کیفیت آماری و ... کاربرد دارد.

همچنین می -توان برای اجرای فرمان- هایی که در منو موجود نیست یک ماکروی کوچک نوشت و آن را اجرا کرد.

نرم افزار تحلیل آماری NCSS

نرم افزار NCSS یکی دیگر از نرم افزارهای آماری است که امکانات کامل و جامعی در اختیار شما قرار می- دهد تا داده- های خود را به بهترین شکل تحلیل کنید.

برای این کار می توانید از بیش از ۲۰۰ ابزار گرافیکی و آماری در این نرم افزار استفاده کنید.

از ویژگی- های نرم افزار NCSS می- توان به امکان رگرسیون خطی و غیرخطی، سازگاری با نسخه -های مختلف ویندوز، امکان رسم انواع نمودار و آمار توصیفی اشاره کرد. این نرم افزار به خاطر توانایی که در محاسبات تعیین حجم نمونه دارد بسیار مورد توجه تحلیلگران آماری است.

نرم افزار تحلیل آماری Eview

نرم افزار Eview یکی از نرم افزارهای تخصصی در زمینه بکارگیری روش- های اقتصاد سنجی، بخصوص جدیدترین روش- های آن است.

به کار گیری آسان روش -های VAR (نامحدود و ساختاری) و مدل -های ARIMA از مزیت- های این نرم افزار محسوب می- شود.

این نرم افزار اولین بار در سال ۱۹۹۴ به بازار عرضه شد و هدف آن انجام عملیات اقتصاد سنجی و تجزیه تحلیلی آماری داده- های اقتصادی بود.

نام این نرم افزار از دو کلمه- ی Econometric View گرفته شده است.

نرم افزار تحلیل آماری R

نرم افزار r یک نرم افزار آماری اوپن سورس است که بسیاری از آمار شناسان با استفاده از آن بسیاری از تکنیک -های آماری مدرن را اجرا می- کنند.

این نرم افزار یکی از قوی- ترین و به روز ترین نرم افزارهای آماری است که در زمینه- های مختلفی غیر از آمار نیز استفاده می- شود.

به دلیل رایگان بودن، بسیاری استفاده از این نرم افزار را به سایر نرم افزارها ترجیح می -دهند علاوه بر این -ها یکی دیگر از دلایل محبوبیت آن، قابل استفاده بود برای کاربران نرم افزار Splus است.

نکته ای که وجود دارد، این نرم افزار حتی ساده -ترین منوهای آماری را هم ندارد و تمامی امکانات آن از طریق دستورات و کدنویسی قابل استفاده است.

نرم افزار تحلیل آماری Expert Choice

نرم افزار Expert Choice اغلب برای انجام تحلیل سلسله مراتبی یا AHP و مقایسات زوجی استفاده می- شود و امروزه در اکثر علوم از جمله علم مدیریت به کار گرفته می- شود.

این نرم افزار تنها نرم افزاری است که به صورت اختصاصی در کشور ما برای انجام روش تحلیل AHP به کار می -رود.

بنابراین، هر فردی که نیاز به تحلیل AHP و تجزیه و تحلیل پرسشنامه -های مقایسات زوجی در پایان نامه یا مقاله خود دارد باید از این نرم افزار استفاده کنید.

نرم افزار تحلیل آماری Microfit

مایکروفیت یک نرم افزار ویژه اقتصاد سنجی- کاربرد علم آمار در اقتصاد یا Econometric است که قابلیت -های بسیاری در انجام روش- های رگرسیونی و تمامی آزمون- های مکمل آن در اقتصاد دارد و برای محققان رشته- های علوم اقتصادی می- تواند بسیار مفید واقع شود.

نرم افزارهای متعددی برای انجام تحلیل آماری مورد استفاده قرار می- گیرند که هر یک ویژگی و خصوصیت خاص خودشان را دارند، تسلط به تمامی این نرم افزارها امکان پذیر نیست بنابراین اگر تصمیم می- گیرید یک نرم افزار تحلیل آماری را انتخاب کنید باید بدانید به چه منظور است و برای چه کاری چنین تصمیمی دارید. همچنین اطلاعات تکمیلی درباره این نرم افزارها بخوانید تا بتوانید کاربردی -ترین آن- ها را برای رسیدن به هدفتان انتخاب کنید.

****

مقایسه نرم افزارهای آماری

۱- استفاده آسان:

در این بخش SPSS و Minitab بهترین هستند هر چند که SAS نیز در آخرین ویرایش خود سعی کرده منوهای بیشتری در اختیار کسانی قرار دهند که حوصله برنامه نویسی ندارند.

۲- یادگیری:

در این بخش هم مانند قبل SPSS و Minitab بهترین هستند. SAS و R سخت ترین نرم افزارها جهت یادگیری و STATA در مرز این دو گروه قرار دارد.

۳-عمق روش منویی:

این موضوع اشاره به طیف وسیعی از امکانات منو در نرم افزار دارد که SPSS از این لحاظ بهترین است.بعد از آن نیز Minitab. در این مورد R ضعیف ترین است.

۴- کیفیت و سهولت استفاده از روش آماری ارائه شده:

در این قسمت R و SAS بهترین هستند.

۵- اصلاح مشخصات خروجی:

این معیار به طیف وسیعی از گزینه های خروجی یک سیستم برای هر یک از روش های تحلیلی آن اشاره دارد . برای مثال ، SPSS طیف بسیار گسترده ای از گزینه های خروجی از آمار توصیفی، GLM آنالیز واریانس ، رگرسیون را ارائه می دهد . Minitab ارائه گزینه های به مراتب کمتر است .

۶- سهولت تبدیل خروجی به فرمت های دیگر مانند APA:

در این قسمت SPSS بهترین نرم فزار است.

۷- ارائه طیف وسیعی از خروجی گرافیکی:

در این قسمت R و SAS بهترین هستند.

۸- سرعت استفاده از داده هایی با حجم بالا:

نیازی به گفتن ندارد که SAS به قادر به مدیریت پایگاه داده های بزرگ است. این قابلیت SAS را در صف مقدم فن آوری نرم افزارها برای داده کاوی، قرار داده است.

۱۰- جامع بودن خروجی:

در این قسمت SPSS در مکان اول قرار دارد.

۱۱- جامعیت و ارزشمندی اسناد:

در این قسمت SAS بهترین و R بدترین است.

****

نرم افزار اکسل و استفاده آن در آمار

مقدمه :

نرم افزار اکسل یک ابزار محاسباتی قوی برای اکثر عملیات مهندسی و ریاضی و حسابداری و جداول است که قادر است عملیات سطر و ستون روی داده ها را انجام دهد.

توابع آمار در اکسل

مایکروسافت توابع آمار و احتمال زیادی (بیش از 80 تابع) در اکسل برای کاربران فراهم نموده است. بعضی در زیر معرفی میشوند

ABS	The absolute value of a number قدر مطلق عدد =abs(-34)
AVERAGE	The average or arithmetic mean for a group of numbers میانگین اعداد =average(a2:12)
COUNT	The number of cell locations in a range that contain a numeric character شمارش تعداد سلولهاي داراي عدد =count(a2:12)
COUNTA	The number of cell locations in a range that contain a text or numeric character شمارش تعداد سلولهاي غير خالي =count(a2:12)
MAX	The highest numeric value in a group of numbers ماگزیمم بین اعداد =max(a2:a12)
MEDIAN	The middle number in a group of numbers (half the numbers in the group are higher than the median and half the numbers in the group are lower than the median)میانه بین اعداد =median(a2:a12)
MIN	The lowest numeric value in a group of numbers مینیمم بین اعداد =min(a2:a12)
MODE	The number that appears most frequently in a group of numbers داده با بیشترین تکرار =mode(a2:a12)
PRODUCT	The result of multiplying all the values in a range of cell locationsضرب تمام اعداد =product(a2:a12)
SQRT	The positive square root of a number جذر عدد =sqrt(36)
STDEV.S	The standard deviation for a group of numbers based on a sample انحراف معیار نمونه =stdev.s(a2:a14)
SUM	The total of all numeric values in a groupجمع اعداد =sum(a2:a14)

توضیح بیشتر در مراجعه به نرمافزار اکسل و ارائه مثال از طریق توابع

منوی آمار در اکسل

غیر از توابع خود نرم افزار اکسل یک منو را جهت انجام عملیات آماری فراهم نموده است

معمولا هنگام اجرا Excel منوی بنام Data وجود دارد که درآیکنهای زیر آن افزونه ای بنام Data Analysis وجود ندارد که باید بشرح ذیل اضافه شود

File ⇨ options ⇨ Add ins ⇨ Manege(Go) ⇨ Analysis ToolPack☑

توضیح بیشتر در مراجعه به نرم‌افزار اکسل و ارائه مثال از طریق منوي آمار

نرم افزار مثلب و استفاده آن در آمار

مقدمه :

نرم افزار Matlab که می توان آنرا زبان ریاضیات مدرن نامید، ابزار قدرتمندی برای پردازش اطلاعات در ساختارهای ماتریسی است. این نرم افزار دارای توانمندی تحلیل عددی بسیار گسترده ای می باشد. ماتریسها، معادلات دیفرانسیل رشته های عددی اطلاعات، ترسیمات و گرافها، لوازم اصلی بکار رفته در ریاضیات و نیز در محیط نرم افزار Matlab هستند. این مجموعه امکانات، مت لب را به محیطی با راندمان بالا برای کاربردهای مهندسی و علوم تبدیل کرده است. نکته مهمتر آنکه وجود حالات مختلف محاوره ای با کاربر و عملکرد بلادرنگ این نرم افزار، آن را بسیار کاربردی تر ساخته است. نام نرم افزار Matlab، گرفته شده از علامت اختصاری Matrix Laboratory می باشد. وب سایت اینترنتی گروه نویسندگان و توسعه دهنده این نرم افزار به آدرس ذیل می باشد: www.mathworks.com

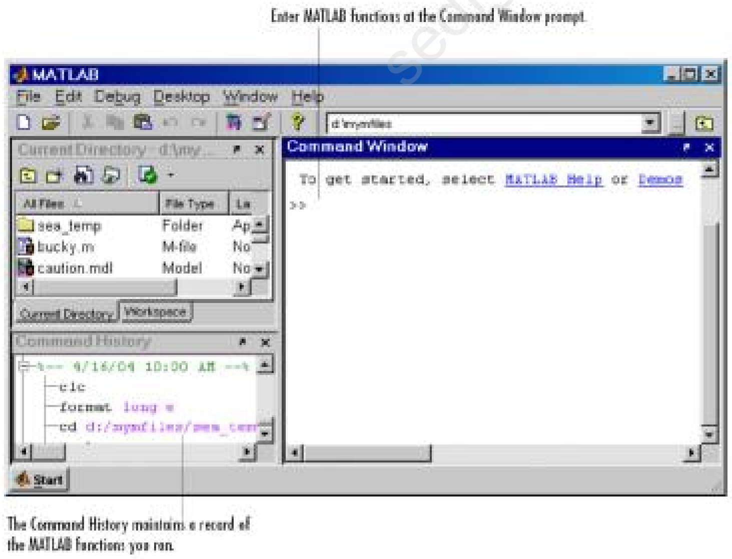
سیمولینک نیز برنامه ای محاوره ای و مرتبط با نرم افزار Matlab، جهت مدل کردن، شبیه سازی و تجزیه و تحلیل سیستمهای دینامیکی است که کاربردهای عمده آن در خصوص سیستمهای مختلف مهندسی از جمله مهندسی مکانیک، هوافضا، برق، دینامیکی، مخابرات، کنترل، برق قدرت، مهندسی پزشکی، و آمار و ...... میباشد

مثلا در سیمولینک ابزار SimPowerSystems، در سیستم های قدرت الکتریکی نوین، به کمک شبیه سازی آنها و ایجاد یک مرجع راهنما می باشد.

نصب نرم افزار Matlab

پس از قرار دادن دیسک Matlab که متناسب با نوع ویندوز باید باشد، این نرم افزار را شروع به نصب در محیط Windows می نماییم. این نرم افزار دارای یک کد شناسایی PLP می باشد که این کد را باید داشته باشید.

بعد از نصب و اجرا در محیط اجرایی (Command Window) دستورات اجرایی نرم افزار را تایپ می نماییم



روش های برنامه نویسی در Matlab:

اصولاً دو روش برنامه نویسی در نرم افزار Matlab وجود دارد:

1) روش برنامه نویسی در محیط اجرایی (Command Window) و یا فضای کاری (Workspace)

2) روش برنامه نویسی در فایل M-File

3) سیمولینک استفاده از ابزارهای گرافیکی در کتابخانه های سیمولینک

جزوه آموزش متلب از سایت www.aminsedighi.ir دانلود کنید

مثال1.) تابع $f(x) = x^3 + \sin(5\pi x)$ را رسم کنید

1- به x مقدار میدهیم x = linspace (-2 , 3) یعنی صد نقطه بین 3- تا 3+

2- مشخص نمودن تابع که در اینجا دوتابع داریم

تابع $f(x) = x^3 + 0x^2 + 0x^1 + 0$ که با آرایه ضرایب x مشخص میکنیم [1 0 0 0] و گرفتن جواب بازاء هر داده x که با استفاده از تابع polyval(p,x) حاصل میشود

تابع $f(x) = \sin(5\pi x)$ که بازاء هر داده x جواب با استفاده از خود تابع میتوان بدست آورد

4 - سپس مینویسم plot (x , y)

```
x=linspace(-3,3)
p=[1 0 0 0 ]
v=polyval(p,x)
u=sin(x*5*pi)
y=v+u
plot(x,y)
```

پاسخ متلب

```
% plot rasm nemodar
clc
clear all
x=linspace(-3,3)
p=[1 0 0 0 ]
v=polyval(p,x)
u=sin(x*5*pi)
y=v+u
plot(x,y)
x =
  Columns 1 through 7
   -3.0000   -2.9394   -2.8788   -2.8182   -2.7576   -2.6970   -2.6364
....................
  Columns 92 through 98
    2.5152    2.5758    2.6364    2.6970    2.7576    2.8182    2.8788
  Columns 99 through 100
    2.9394    3.0000

p =
     1     0     0     0

v =
  Columns 1 through 7
```

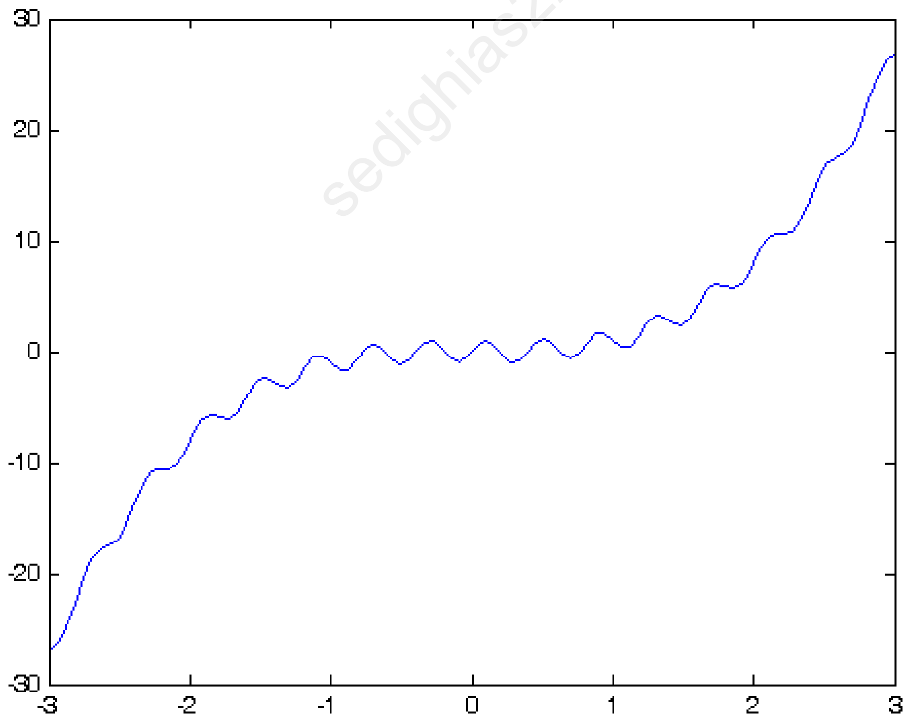
****

```
  -27.0000  -25.3965  -23.8577  -22.3824  -20.9692  -19.6168  -18.3238
  Columns 8 through 14
  -17.0889  -15.9108  -14.7881  -13.7195  -12.7037  -11.7393  -10.8250
..........................
  Columns 92 through 98
   15.9108   17.0889   18.3238   19.6168   20.9692   22.3824   23.8577

  Columns 99 through 100

   25.3965   27.0000

u =
  Columns 1 through 7
   -0.0000   -0.8146   -0.9450   -0.2817    0.6182    0.9989    0.5406
  Columns 8 through 14
   -0.3717   -0.9718   -0.7557    0.0951    0.8660    0.9096    0.1893
...............
  Columns 92 through 98
    0.9718    0.3717   -0.5406   -0.9989   -0.6182    0.2817    0.9450
  Columns 99 through 100
    0.8146    0.0000
y =
  Columns 1 through 7
  -27.0000  -26.2110  -24.8027  -22.6642  -20.3511  -18.6179  -17.7832
  Columns 8 through 14
  -17.4606  -16.8826  -15.5439  -13.6245  -11.8377  -10.8297  -10.6357
.....................
  Columns 92 through 98
   16.8826   17.4606   17.7832   18.6179   20.3511   22.6642   24.8027
  Columns 99 through 100
   26.2110   27.0000
```

توابع تولید عدد تصادفی

نام تابع	جواب مثلب	توضیح
b=rand	b = 0.6538	عدد تصادفی 0 تا 1
c=rand(3)	c = 0.4942 0.9037 0.6987 0.7791 0.8909 0.1978 0.7150 0.3342 0.0305	ماتریس 3 در 3 عدد تصادفی بین 0 تا 1
d=rand(2,3)	d = 0.7441 0.4799 0.6099 0.5000 0.9047 0.6177	ماتریس 2 در 3 عدد تصادفی بین 0 تا 1
d=10*rand(1,3)	d = 6.2797 2.9198 4.3165	
m=fix(d)	m = 6 2 4	
n=round(d)	n = 6 3 4	گرد کردن
a=randi(10,5,3)	a = 6 1 6 1 2 1 6 9 9 3 3 5 5 5 10	عدد صحیح تصادفی در 3 ستون بین صفر تا ده – در پنج ردیف و سه ستون
b=max(a)	b = 6 9 10	در مثال بالا ماگزیمم هر ستون بدست میاورد
a=randperm(10)	a = 10 7 2 1 8 3 4 5 6 9	بردار= ماتریس یک سطری برای عدد تصادفی 1 تا 10
a=randperm(10,3)	a = 5 1 9	سه عدد تصادفی بین 1 تا 10
b= randn(3)	b = 0.6164 1.8291 -0.5635 -0.5250 0.0853 -0.4736 1.0077 -0.0683 -1.7035	تولید اعداد تصادفی نرمال استاندارد(گوسی) در یک ماتریس 3 در 3 اگر اعداد خیلی زیاد باشند آنگاه نرمال استاندارد میشود (میانگین 0 و انحراف معیار 1)
b = randn(2 , 3)	b = -0.0537 1.2559 0.0549 -0.8813 0.1558 1.3986	تولید اعداد تصادفی نرمال استاندارد در یک ماتریس 2 در 3 اگر اعداد خیلی زیاد باشند (بجای 3 عدد 3000) آنگاه نرمال استاندارد میشود (میانگین 0 و انحراف معیار 1)
f=6+5*randn(1 , 3000) g=mean(f) k2=var(f) k3=sqrt(var(f))	f = 3.2275 -1.7150 ….. 2.1036 4.6005 g = 5.9902 k2 = 25.6394 k3 = 5.0635	تولید اعداد f تصادفی نرمال در یک ماتریس 1 در 3000 اگر اعداد خیلی زیاد باشند آنگاه نرمال میشود (میانگین 6 و انحراف معیار 5)

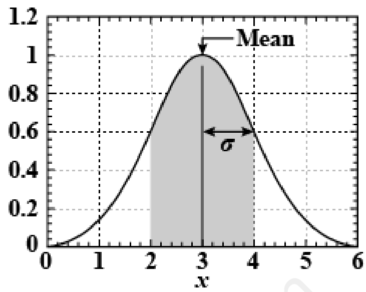
محاسبات آماری در مثلب :

میانگین در مثلب – دستور mean در مثلب – میانگین یا مقدار موردانتظار (Expected value)

از نخستین شاخص‌های است که برای متغیرهای تصادفی و مدل‌های آماری محاسبه می‌شود تا بتوان به کمک آن به طور کاملا تقریبی یک مقدار موردانتظار از متغیر تصادفی را در نظر گرفت. محاسبه میانگین براساس عملگر امید ریاضی و برای تعداد N نمونه برای متغیر تصادفی X بصورت زیر تعریف می‌شود:

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

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i$$



برای محاسبه میانگین در مثلب کافی است از دستور mean استفاده کنیم که برای یک ماتریس دلخواه این مقدار را محاسبه می‌کنیم:

مثال2.) میانگین

a= [4 6 8 10]
b=mean(a)
جواب
b =   7

مثال3.)

a=[1,4,7,10;15,16,17,18]
b=mean(a)
جواب = میانگین اعداد هر ستون
b =   8   10   12   14

c=mean(a)
جواب = همان میانگین اعداد هر ستون
c =   8   10   12   14

d=mean(a,2)
جواب = میانگین اعداد هر سطر
d = 5.5000
    16.5000

حالا اگر بخواهیم میانگین کل اعداد را حساب کنیم
f = mean (d)
جواب = میانگین کل اعداد

f =   11

همانطور که مشاهده می‌کنید اگر در ورودی دوم عدد 1 را قرار دهیم برای محاسبه میانگین بر روی سطرها حرکت می‌کند و میانگین اعداد روی یک ستون را در نظر می‌گیرد و اگر ورودی دوم را عدد 2 قرار دهیم میانگین اعداد روی هر سطر را محاسبه می‌کند و اصطلاحا روی ستون‌ها حرکت می‌کند .در حالتی که بدون ورودی دوم فراخوانی شود به طور پیش فرض حالت اول را در نظر می‌گیرد.

اگر بنویسم a=randn(10,5) و سپس بنویسم b=mean(a) میانگین روی ستون و اگر بنویسم b=mean(a,2) میانگین روی سطر انجام میشود

توجه: برای تولید اعداد تصادفی با توزیع نرمال استاندارد، از دستور randn استفاده میشود

در صورتی که در یک متغیر تصادفی اعدادی تعریف نشده (**nan**) وجود داشته باشد .می‌توان از دستور nanmean نیز استفاده کرد:

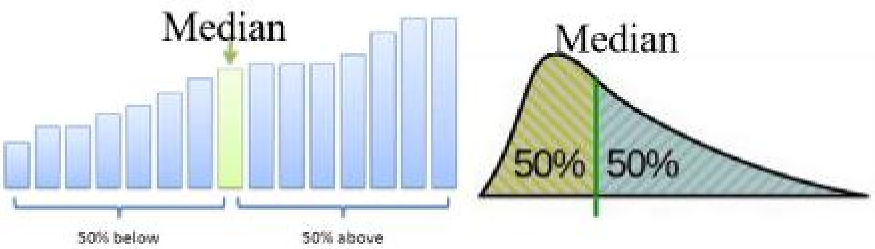
a= [1,4,7,10;15,16,17,nan]
b = nanmean (a)

جواب

b =   8   10   12   10

میانه در مثلب – دستور median در مثلب

همانطور که می‌دانید میانه داده‌ای است که پنجاه درصد داده‌ها از آن کوچکتر و یا بزرگتر هستند .در توزیع نرمال میانه و میانگین یکی هستند اما در حالت کلی باید بین میانگین و میانه تفاوت قائل شد. برای محاسبه میان در مثلب دستور medianقرار داده شده است.



اگر مانند محاسبه میانگین در مثلب داده‌هایی از جنس nan داشته باشیم، باید آن‌ها را حذف کنیم. برای بدست آوردن میانه در مثلب در این حالت نیز می‌توانیم از زیردستور omitna استفاده کنیم. همچنین برای محاسبه میانه در مثلب در این حالت می‌توانیم از دستور nanmedian نیز استفاده کنیم. به عنوان مثال می‌خواهیم برای 1000 عدد با توزیع استاندارد نرمال میانه را محاسبه کنیم:

```
B=randn(1,1000);
C = median(B)
D = mean(B)
```

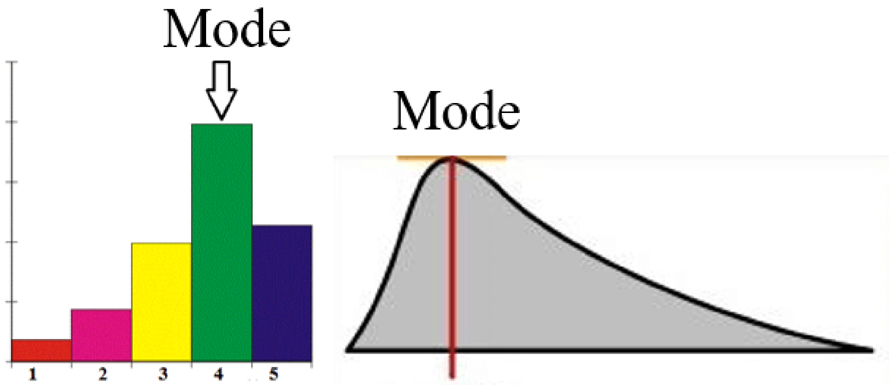
جواب

```
C  = -0.0304
D = -0.0326
```

همانطورکه مشاهده می‌شود چون توزیع نرمال می‌باشد میانه و میانگین تقریبا هر دو نزدیک صفر می‌باشند.

مد در مثلب – دستور mode در مثلب

در ادامه محاسبات آماری در مثلب شاخص مد را بررسی می‌کنیم. مد (mode) در تحلیل‌های آماری داده‌ای است که بیشترین فراوانی را دارد.



برای محاسبه مقدار مد در مثلب یا داده‌ای که بیشترین فراوانی را دارد، از دستور mode استفاده می‌کنیم .برای بدست آوردن مد در مثلب برای متغیرهای تصادفی روند کاملا مشابه با دستور mean برای محاسبه میانگین می‌باشد.

همچنین این دستور در حالت کلی دارای سه خروجی می‌باشد .به عنوان مثال فرض کنید که میانگین ریزش باران در هر ماه برحسب میلیمتر در یک شهر بصورت زیر باشد:

```
a=[2 8 3 8 3 4 9 3 2 3 8]
b=mode(a)
```

جواب

```
a =8    3    2    3    9    4    3    8    3    8    2
b =    3
```

مثال دیگر

```
a=[10 8 7 8 5 4 9 5 8 9 4 8 8 5]
[m,f,c]=mode(a)
```

پاسخ

```
a =
   10    8    7    8    5    4    9    5    8    9    4    8    8    5
m =
    8
f =
    5
c =
   [8]
```

یعنی mode سه عدد بر میگرداند m بیانگر آن عدد که بیشترین تعداد دارد و f بیانگر تعداد تکرار و c آرایه آن عدد

مثالی دیگر

```
a=randi(10,5,3)
[k,f,c]=mode(a)
```

پاسخ

```
a =
    3    2    6
```

```
    9     7     7
    4     4     9
    3     4     6
    9     9     4
k =
    3     4     6
f =
    2     2     2
c =
   [2x1 double]    [4]    [6]
```

اگر در ستون بیشترین تکرار نبود کوچکترین برمیگرداند

اگر بنویسیم [m,f,c]=mode(a,2) عدد 2 در ورودی دوم دستور mode مشابه با دستور mean برای این است که برای محاسبه مد در مثلب حرکت را بر روی سطرهای انجام دهد.

واریانس در مثلب – دستور var در مثلب

واریانس یک متغیر تصادفی براساس عملگر امید ریاضی بصورت لنگر دوم مرکزی تعریف می‌شود. همچنین برای تعداد N نمونه واریانس یک متغیر تصادفی بصورت زیر تعریف می‌شود:

$$M_2 = E\,[(x-\mu_x)^2]=\int_{-\infty}^{\infty}(x-\mu_x)f(x)$$

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

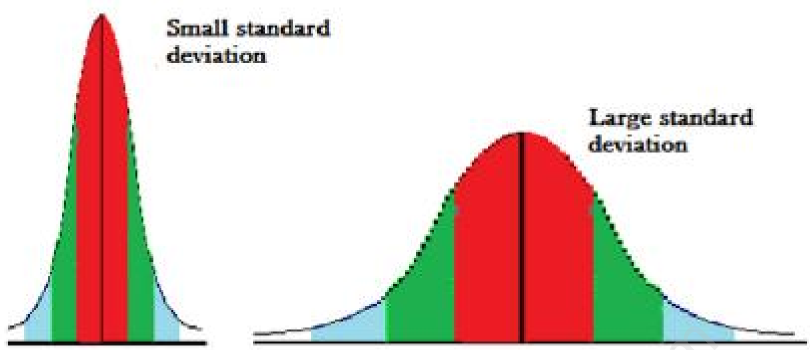
$$2)\frac{1}{N}\sum_{i}^{N}(x_i-\bar{x})^2$$

که رابطه اول اصطلاحا unbiased و رابطه دوم حالت biased می‌باشد. در محاسبه واریانس در مثلب رابطه اول (var(x,0 و رابطه دوم بصورت (var(x,1 ایجاد می‌شود .معمولا در بیشتر از موارد از رابطه اول استفاده می‌شود زیرا در صورتی که واریانس تعدادی نمونه بخواهد با واریانس جامعه برابر باشد ثابت خواهد شد که واریانس نمونه باید رابطه نخست را داشته باشد.

به طور کلی برای محاسبه واریانس در مثلب یا همان لنگر مرکزی دوم از دستور var استفاده می‌شود. برای حالتی که اعداد nan را بخواهیم حذف کنیم از دستور nanvar می‌توانیم استفاده کنیم.

انحراف معیار در مثلب – دستور std در مثلب

اما شاخصی که اهمیت بیشتری نسبت به واریانس در محاسبات آماری در مثلب و مدل‌های احتمالاتی دارد، انحراف معیار (Standard Deviation)یا همان جذر واریانس می‌باشد که پراکندگی داده‌ها نسبت به میانگین را نشان می‌دهد .انحراف معیار به دلیل اینکه هم بعد با متغیر تصادفی است می‌تواند در مقایسه چندین متغیر تصادفی که بعد یکسانی دارند مورد استفاده قرار گیرد.



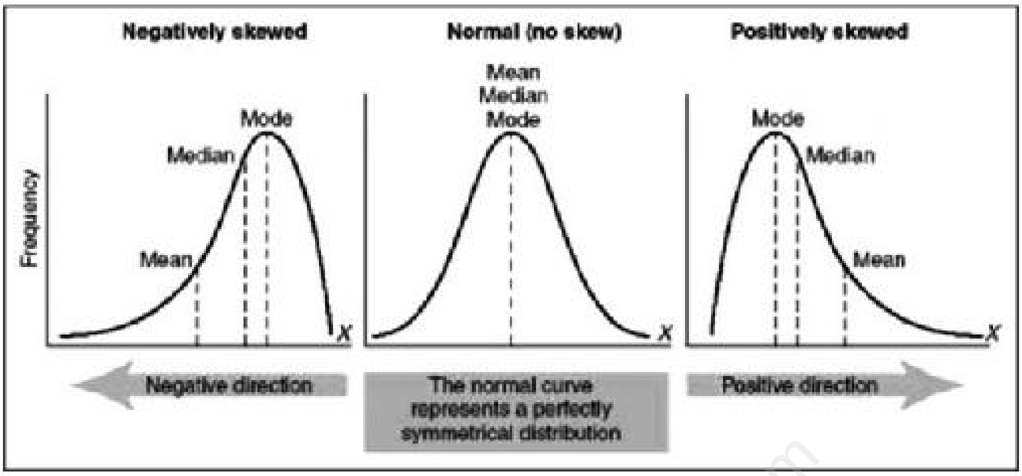
برای محاسبه انحراف معیار در مثلب از دستور std استفاده می‌شود و برای حذف اعداد تعریف نشده از دستور **nanstd** در حالت کلی استفاده می‌شود. به عنوان مثال برای 2000 عدد بصورت یکنواخت بین 10 تا 100 پارمترهای واریانس و انحراف معیار در مثلب را بدست می‌آوریم:

```
x=unifrnd (10,100,1,2000);
var(x,1) = 654.7821
var(x,0) = 655.1097
std(x) = 25.5951
```

همانطور که مشاهده می‌کنید در محاسبه واریانس در مثلب برای تعداد نمونه‌های زیاد هر دو رابطه پاسخ تقریبا یکسانی را نشان می‌دهند.

محاسبه ضریب چولگی در مثلب – دستور skewness در مثلب

در مدل‌های احتمالاتی و متغیرهای تصادفی ضریب چولگی (skewness) نشان دهنده میزان تقارن یک متغیر تصادفی حول میانگین می‌باشد .در شکل زیر ضریب چولگی در حالت‌های و تغییر کردن سایر پارامترها را مشاهده می‌کنید.



مطابق شکل فوق اگر ضریب چولگی مثبت باشد، شکل به سمت چپ اصطلاحا skewness دارد و اگر ضریب چولگی منفی باشد، شکل به سمت راست skewness خواهد داشت .در صورتی که این ضریب صفر باشد، این ضریب حول میانگین متقارن خواهد بود مانند توزیع نرمال.

این ضریب بصورت زیر محاسبه می‌شود:

$$\gamma = \frac{E[(x-\mu)^3]}{\sigma_x^3}$$

$$S_1 = \frac{\frac{1}{n}\sum_{i=1}^{N}(x_i-\bar{x})^3}{\left(\sqrt{\frac{1}{n}\sum_{i=1}^{N}(x_i-\bar{x})^2}\right)^3}$$

برای محاسبه ضریب چولگی در مثلب از دستور skewness استفاده می‌شود. به عنوان مثال برای یک میلیون داده از یک توزیع نرمال با میانگین 1 و انحراف معیار 2 می‌خواهیم این ضریب را محاسبه کنیم:

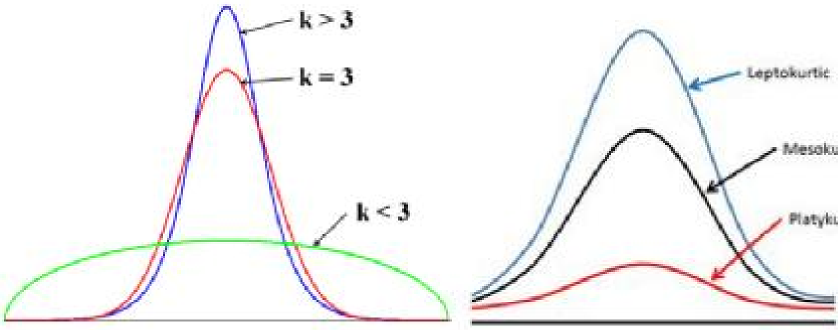
```
y=1+2*randn (1,1000000);
skewness(y) = -9.6266e-04
```

همانطور که مشاهده می‌کنید در محاسبه ضریب چولگی در مثلب به دلیل اینکه توزیع نرمال توزیع متقارن است این ضریب به عدد صفر بسیار نزدیک می‌باشد.

اگر در تعریف اعداد تصادفی در مثلب با توزیع نرمال مشکل دارید، حتما مقاله تولید عدد تصادفی را که در بالا معرفی شده است مطالعه نمایید.

محاسبه ضریب کورتوسیس در مثلب – دستور kurtosis در مثلب

در مدل‌سازی احتمالاتی ضریب کورتوسیس (kurtosis) معیاری از مسطح بودن تابع توزیع می‌باشد. هر چه مقدار ضریب کوتوسیس بیشتر باشد، تغییرات یک مقدار تصادفی بیشتر خواهد بود.



ضریب کورتوسیس برای توزیع یکنواخت عدد 1.8، برای توزیع نرمال عدد 3 و برای توزیع رایلی عدد 6 می‌باشد (که بیشترین مقدار این ضریب را در بین توزیع‌های رایج داراست).

این ضریب بصورت زیر محاسبه می‌شود:

$$K_x = \frac{E[(x-\mu_x)^4]}{\sigma_x^4}$$

$$K_1 = \frac{\frac{1}{n}\sum_{i=1}^{N}(x_i-\bar{x})^4}{\left(\sqrt{\frac{1}{n}\sum_{i=1}^{N}(x_i-\bar{x})^2}\right)^4}$$

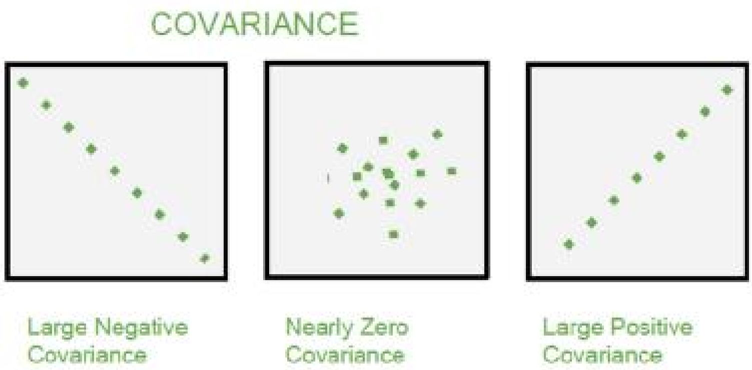
برای محاسبه ضریب کورتوسیس در مثلب از دستور kurtosis استفاده می‌شود .به عنوان مثال برای توزیع نرمال با میانگین 1 و انحراف معیار 2 خواهیم داشت:

```
y=1+2* randn (1,1000000);
kurtosis(y) = 3.0044
```

همانطور که مشاهده می‌کنید مطابق مطلب گفته شده این ضریب به عدد 3 نزدیک می‌باشد .در تعریفی دیگر از فرمول فوق عدد 3 را کم می‌کنند که در آن تعریف این ضریب به نسبت توزیع نرمال سنجیده می‌شود.

محاسبه کواریانس در مثلب – دستور cov در مثلب

مفهوم کواریانس برای دو متغیر تصادفی که برحسب هم ترسیم شده‌اند را در شکل زیر مشاهده می‌کنید. تغییرات دو متغیر را نسبت به هم مشاهده می‌کنید.



کواریانس برای دو متغیر تصادفی X و Y بصورت زیر تعریف می‌شود. همچنین رابطه سوم محاسبه ماتریس کواریانس در مثلب را نشان می‌دهد.

$$Covariance(X,Y) = E[(x-\mu_x)(y-\mu_y)]$$

$$cov(X,Y) = \frac{1}{N-1}\sum_{i=1}^{N}(X_i-\mu_X)^*(Y_i-\mu_Y)$$

$$C_{matrix} = \begin{pmatrix} cov(X,X) & cov(X,Y) \\ cov(Y,X) & cov(Y,Y) \end{pmatrix}$$

برای محاسبه ماتریس کواریانس در مثلب از دستور cov استفاده می‌شود .اگر دستور cov بصورت تک ورودی فراخوانی شود همان واریانس در عمل محاسبه خواهد شد.

در محاسبه ماتریس کواریانس به صورت فوق عمل خواهد شد. دقت شود که در محاسبه ماتریس کواریانس در مثلب حتما باید سایز هر دو بردار متغیر تصادفی با هم برابر باشد.


```
x=unifrnd (10,100,1,2000);
y=1+2*randn (1,2000);
cov (y, x) =
[ 4.0612 -0.7143 ; -0.7143 661.1731 ]
```

درایه‌های رو قطر اصلی ماتریس کواریانس در واقع همان واریانس هر متغیر خواهند بود.

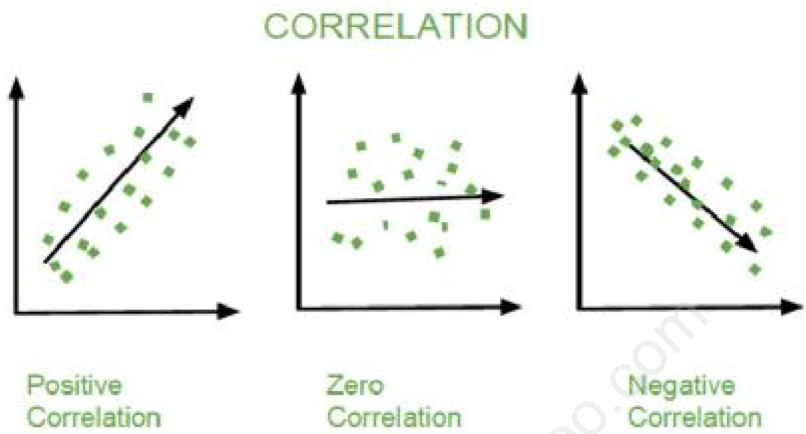
برای یک ماتریس که هر کدام از ستون‌های آن از یکسری مشاهدات از متغیر تصادفی است ماتریس کواریانس، کواریانس دوطرفه بین هر دو ترکیب ستون را محاسبه می‌کند .برای مثال زیر خواهیم داشت:

```
A = [1.77 -0.005 3.98; NaN -2.95 NaN; 2.54 0.19 1.01]
C = cov(A,'omitrows')
C =
0.2964 0.0751 -1.1435
0.0751 0.0190 -0.2896
-1.1435 -0.2896 4.4104
```

لازم به ذکر است همانطور که در محاسبه میانگین و انحراف معیار گفته شد، چون در ماتریس مشاهدات اعداد تعریف نشده داشتیم و می‌خواهیم آنها را حذف کنیم از زیردستور omitrowsاستفاده شده است.

محاسبه ضریب همبستگی در مثلب – دستور corrcoef در مثلب

ضریب همبستگی (correlation) بین دو متغیر تصادفی معیاری از وابستگی خطی است .این ضریب چون بدون بعد است می‌تواند روابط دو به دوی بین متغیرهای تصادفی با ابعاد مختلف را بیان کند. به عنوان مثال وابستگی زیادی بین مقاومت فشاری دو ستون در یک ساختمان وجود دارد که عملا همبستگی مثبت بین آنها وجود دارد و بین بارش برف و دمای هوای یک همبستگی منفی وجود دارد .مطابق شکل زیر:



ضریب همبستگی از طریق روابط زیر محاسبه می‌شود و ماتریس همبستگی مطابق رابطه سوم در محیط مثلب محاسبه می‌شود:

$$\rho_{XY} = \frac{cov[X,Y]}{\sigma_x.\sigma_y} = E[\frac{X-\mu_x}{\sigma_x}.\frac{Y-\mu_y}{\sigma_y}]$$

$$\rho_{XY} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{X_i-\mu_x}{\sigma_x}\right)\left(\frac{Y_i-\mu_y}{\sigma_y}\right)$$

$$R = \begin{pmatrix} \rho(X,X) & \rho(X,Y) \\ \rho(Y,X) & \rho(Y,Y) \end{pmatrix}$$

برای محاسبه ضریب همبستگی در مثلب از دستور corrcoef استفاده می‌شود .به عنوان مثال برای دو متغیر تصادفی X و Yاستاندارد نرمال بصورت زیر ضریب همبستگی را محاسبه می‌کنیم:

```
X=randn(1,1000);
Y=randn(1,1000);
corrcoef(X,Y) =
1.0000 -0.0017
-0.0017 1.0000
```

این دستور همچنین دارای خروجی‌های بیشتر و زیردستوراتی می‌باشد که به دلیل اهمیت کم آن‌ها از آوردن آن‌ها صرف نظر کرده‌ایم.

می‌توان با مراجعه به help نرم‌افزار مثلب یا سایت اصلی نرم‌افزار مثلب در صورت نیاز این تنظیمات را مشاهده نمود.

تابع polyfit بهترین منحنی گذرنده از زوج نقاط با درجات مختلف

برازش : تعدادی زوج نقطه داریم میخواهیم برازنده ترین منحنی که از این نقاط عبور میکند با درجات مختلف یک تابع بدست آوریم

مثال : نمودار رگرسیون خطی (معادله درجه یک) و درجات 2 و 3 را برای زوج نقاط زیر رسم کنید

x=[-2	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2]
y=[-8	-4.1	-1.7	-0.5	-0.1	0	0.1	0.5	1.7	4.1	8]

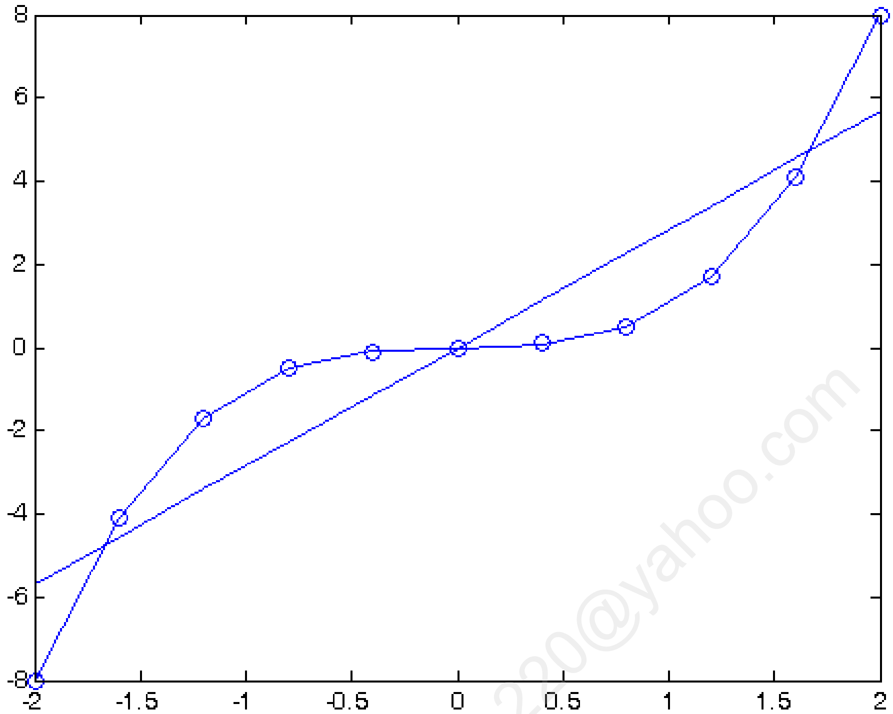
نوشتن script file در مثلب و سپس اجرا آن

```
clc
clear all
hold off
x=-2:0.4:2
y=[-8 , -4.1 , -1.7 , -0.5 , -0.1 , 0 , 0.1 , 0.5 , 1.7 , 4.1 , 8]
plot(x,y,'o')
hold on
pause
for n= 1:1:3
p1=polyfit(x,y,n)
y1=polyval(p1,x)
plot(x,y1)
pause
end
```

جواب متلب

```
% me poly fit
clc
clear all
x=-2:0.4:2
y=[-8 , -4.1 , -1.7 , -0.5 , -0.1 , 0 , 0.1 , 0.5 , 1.7 , 4.1 , 8]
plot(x,y,'o')
hold on
for n= 1:1:3
p1=polyfit(x,y,n)
y1=polyval(p1,x)
plot(x,y1)
end
x =
   -2.0000   -1.6000   -1.2000   -0.8000   -0.4000         0    0.4000
    0.8000    1.2000    1.6000    2.0000
y =
   -8.0000   -4.1000   -1.7000   -0.5000   -0.1000         0    0.1000
    0.5000    1.7000    4.1000    8.0000
p1 =
    2.8455         0
y1 =
   -5.6909   -4.5527   -3.4145   -2.2764   -1.1382         0    1.1382
    2.2764    3.4145    4.5527    5.6909
p1 =
    0.0000    2.8455   -0.0000
y1 =
```

```
   -5.6909   -4.5527   -3.4145   -2.2764   -1.1382   -0.0000    1.1382
    2.2764    3.4145    4.5527    5.6909
p1 =
    1.0022    0.0000   -0.0088   -0.0000
y1 =
   -8.0000   -4.0909   -1.7212   -0.5061   -0.0606   -0.0000    0.0606
    0.5061    1.7212    4.0909    8.0000
```



در دوران کرونا (کووید 19) سمینار مجازی مقدمه ای بر استفاده از نرم افزار Excel و Matlab در آمار 27 بهمن 1399 امین صدیقی

****
در تهیه این جزوه از تجربیات و اطلاعات شخصی خودم و و دیگر اساتید محترم و نوشته ها و نظرات دیگران در اینترنت استفاده نمودم . بدیهی است این جزوه دارای اشکالاتی است که با تذکرات شما و ارسال به آدرس زیر نقایص برطرف و تکمیل میشود. Sedighias220@yahoo.com

****
در هر حرفه ای که هستید نه اجازه دهید که به بدبینیهای بیحاصل آلوده شوید و نه بگذارید که بعضی لحظات تاسف بار که برای هر ملتی پیش می آید شما را به یاس و ناامیدی بکشاند. در آرامش حاکم بر آزمایشگاهها و کتابخانه هایتان زندگی کنید .

نخست از خود بپرسید : " برای یادگیری و خودآموزی چه کرده ام ؟ "

سپس همچنان که پیشتر میروید بپرسید : " من برای کشورم چه کرده ام ؟ "

و این پرسش را آنقدر ادامه دهید تا به این احساس شادیبخش و هیجان انگیز برسید که شاید سهم کوچکی در پیشرفت و اعتلای بشریت داشته اید.

اما هر پاداشی که زندگی به تلاشهایمان بدهد یا ندهد هنگامی که به پایان تلاشهایمان نزدیک میشویم هر کداممان باید حق آن را داشته باشیم که با صدای بلند بگوییم " من آنچه در توان داشته ام انجام دادهام "

لوئی پاستور 1895