



اصول کلی کنترل کیفیت در آزمایشگاه - (بخش دوم)

آمار در آزمایشگاه ۲

توجه داشت که رایانه فقط یک وسیله است و اگر از آن به خوبی استفاده نشود، مشکلات بیشتری به وجود خواهد آورد و مسوولیت آن به عهده مدیر آزمایشگاه است که نه تنها کنترل کیفیت اندازه گیری آزمایش ها بلکه کنترل کل آزمایشگاه را باید زیر نظر داشته باشد.

در قسمت قبل به بررسی شاخص های مرکزی از جمله میانه، میانگن و مد یا نما و کاربرد آنها پرداختیم همچنین درخصوص نمودارهای توزیع از جمله توزیع طبیعی یا گوسین، توزیع انحراف یافته، توزیع خطی و توزیع لگاریتمی نیز سخن به میان آوردیم و بعد از آن، شاخص های پراکندگی (دامنه تغییرات، میانگین انحرافات، واریانس، انحراف معیار، ضریب تغییرات، SEM یا خطای استاندارد از میانگین) را مورد بررسی قرار دادیم. در ادامه به بررسی دو موضوع همبستگی و آنالیز رگرسیون خطی می پردازیم.

همبستگی

حوادث متعددی در طبیعت اتفاق می افتد که بین آنها همبستگی یا رابطه وجود دارد. به عنوان مثال وضعیت خورشید و درجه حرارت زمین. چنانکه مشخص است ارتباط بین دو یا چند متغیر را همبستگی می گویند. هنگامی که افزایش در یک متغیر همراه با افزایش در متغیر دیگر است، همبستگی بین دو متغیر مستقیم و مثبت است. مثل همبستگی میان میزان هموگلوبین و شمارش گلبول های قرمز. چنانچه افزایش در یک متغیر همراه با کاهش در متغیر دیگری باشد، همبستگی از نوع منفی و معکوس است. وجود همبستگی بین متغیرها به این معنا نیست که یک متغیر، علت متغیر دیگریست، بلکه همبستگی رابطه بین دو متغیر را در یک جامعه توصیف می کند. به منظور سهولت در محاسبات آماری و ریاضی، متغیرها را با نمادهای X و Y نام گذاری می کنند.

یکی از روش هایی که توسط آن می توان همبستگی بین دو متغیر را نشان داد نمودار یا دیاگرام های پراکندگی است. این نمودار یک نمایش ترسیمی است که از طریق آن ارزش های دو متغیر نشان داده می شود. به طوریکه داده های مربوط به یک متغیر در محور افقی و داده های مربوط به متغیر دیگر در

امروزه نقش و جایگاه آزمایشگاه نه تنها به عنوان مرکزی جهت تشخیص بیماری ها بلکه محلی جهت پایش سلامت جامعه و به عنوان یکی از حساس ترین نقاط درمانی که نقش بسزایی را در تشخیص بیماری و درمان آن به عهده دارد و به پزشکان محترم کمک می کند که با تفسیر نتایج حاصل از آزمایشات بیمار، به تشخیص نهایی و درمان قطعی نزدیک و نزدیک تر شوند بر همگان واضح است و امید است که با وجود نوپا بودن موضوع و فرهنگ کنترل کیفی در کشور عزیزمان ایران، با همت و یاری عزیزان و مسئولان محترم امر، به هرچه نهادینه کردن، تثبیت و جا انداختن این مهم کمک شود.

کنترل کیفیت محصولات، قرن ها پیش به صنایع دستی و سپس به کارخانجات وارد شده و سازندگان مجبور بودند هر فرآورده ای را جداگانه مورد کنترل و بازرسی قرار دهند تا محصول خود را با کیفیتی قابل قبول به مشتری تحویل دهند. امروزه با پیشرفت سریع فناوری و پیدایش دستگاه های پیچیده سعی می شود که با آزمایش تعدادی از محصولات، کیفیت و دستگاه ها را کنترل نمایند. برای این کار از آمار استفاده نموده و به همین جهت SQC یا Statistical Quality Control وارد کارخانجات سازنده لوازم شد. تعریفی که از SQC شده است به قرار زیر است:

SQC پروسه ایست با تمرکز بر افشای هرگونه انحرافات از استانداردهای تعریف شده.

جمع آوری و تنظیم نتایج آزمایش ها و محاسبه آنها و سرانجام نتیجه گیری از داده های جزء، برای رسیدن به نتیجه کل، اساس آمار است. امروزه با به خدمت گرفتن رایانه در آزمایشگاه ها مسائل زیادی حل شده و سرعت امور آزمایشگاهی افزایش یافته است و به عنوان مثال، مشکلات یادداشت کردن اطلاعات و مشخصات بیمار اعم از نام بیمار، نام پزشک معالج، انواع آزمایش های مورد نیاز و .. و همچنین انتقال این اطلاعات به رایانه، دستگاه های اندازه گیری آزمایش ها و برگرداندن نتایج آزمایش ها به رایانه، تهیه گزارش ها، نگهداری اطلاعات مربوط به کالیبره کردن دستگاه ها و رسم نمودارهای مربوط به کنترل کیفیت و غیره نیز بسیار آسان شده است. اما باید به این نکته نیز

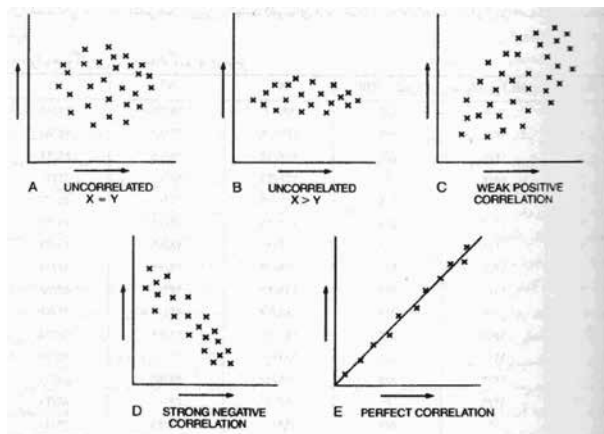
محور عمودی واقع می شود و بدین ترتیب وضعیت همبستگی بین دو متغیر مشخص می شود که در شکل ۱ پنج مورد از آن ها نشان داده شده اند.

A و B: گراف هایی که پراکندگی آنها شبیه به این دو نمودار است، عدم همبستگی مجموعه داده ها را نشان می دهد. (داده های ناهمبسته یا مستقل که در آنها مقادیر X و Y برابر است و داده های ناهمبسته ای که مقادیر X بزرگتر از Y دارد)

C: گراف هایی که پراکندگی آنها به شکل بیضی مایل است که به سمت بالا امتداد دارد، همبستگی مثبت و غیرخطی داده ها را نشان می دهد (همبستگی مثبت ضعیف - الگوی پراکندگی بیضی وار زاویه دار).

D: گراف هایی که پراکندگی آنها به شکل بیضی مایل و با امتداد رو به پایین است، همبستگی منفی و غیرخطی داده ها را نشان می دهد (همبستگی منفی قوی).

E: گراف هایی که پراکندگی آنها به شکل یک خط راست است، همبستگی خطی یا کامل داده ها را نشان می دهد (همبستگی کامل یا خطی داده ها که در آن تمامی داده ها به شکل کاملاً مناسبی در امتداد خطی با زاویه ۴۵ درجه یا همان Line of best fit قرار دارند).



شکل ۱. مثال هایی از نمودارهای پراکندگی

ضریب همبستگی

جهت محاسبه میزان همبستگی دو متغیر مورد مطالعه از ضریب همبستگی با نماد r استفاده می شود. این ضریب نشان می دهد که بین دو متغیر (به عنوان مثال X و Y) چه مقدار همبستگی وجود دارد.

$$r = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{[(n \sum X^2) - (\sum X)^2][(n \sum Y^2) - (\sum Y)^2]}}$$

این ضریب از رابطه زیر محاسبه می شود:

ضریب همبستگی صفر ($r=0$)، و اگر این دو صفت همبستگی کامل داشته باشند مقدار $r=1$ یا $r=-1$ است. بنابراین اگر پراکندگی داده ها به شکل یک خط راست و رو به بالا باشد (همبستگی مستقیم و کامل) میزان $r=1$ و زمانی که پراکندگی داده ها به صورت یک خط راست و به سمت پایین باشد (همبستگی معکوس و کامل) میزان $r=-1$ خواهد بود. ضریب همبستگی بین $+1$ و -1 همبستگی ناقص خواهد بود (نمودار C و D در شکل ۱).

لازم به ذکر است که می توان از رابطه $[100r^2]$ درصد همبستگی را محاسبه کرد.

از همبستگی و نمودارهای پراکندگی می توان جهت مقایسه دو روش و یا همبستگی دو پارامتر استفاده کرد. برای مثال جهت مقایسه اندازه گیری هموگلوبین به روش مرجع و روش SLS (Sodium lauryl sulphate).

توجه به این نکته الزامی است که محاسبه r در شرایطی که تفاوت بین نتایج به دست آمده خیلی نزدیک نباشد، نتیجه لازم را نمی دهد. برای درک این نکته مهم، مثالی می آوریم.

در مثال زیر اندازه گیری سدیم سرم، با دو روش مختلف A و B مورد مقایسه قرار گرفته است.

برای نمونه، برای نتایج ۱۰ آزمایش ثانویه، r عبارت است از:

شماره	نتایج ۸ آزمایش اولیه					نتایج ۱۰ آزمایش ثانویه				
	روش اندازه گیری سدیم					روش اندازه گیری سدیم				
	A(X)	B(Y)	X ²	Y ²	XY	A(X)	B(Y)	X ²	Y ²	XY
۱	-	-	-	-	-	۱۰۵	۱۰۸	۱۱۰۲۵	۱۱۶۶۴	۱۱۳۴۰
۲	۱۲۴	۱۲۶	۱۵۳۷۶	۱۵۸۷۶	۱۵۶۲۴	۱۲۴	۱۲۶	۱۵۳۷۶	۱۵۸۷۶	۱۵۶۲۴
۳	۱۲۶	۱۲۰	۱۵۸۷۶	۱۴۴۰۰	۱۵۱۲۰	۱۲۶	۱۲۰	۱۵۸۷۶	۱۴۴۰۰	۱۵۱۲۰
۴	۱۳۰	۱۳۵	۱۶۹۰۰	۱۸۲۲۵	۱۷۵۵۰	۱۳۰	۱۳۵	۱۶۹۰۰	۱۸۲۲۵	۱۷۵۵۰
۵	۱۳۲	۱۳۰	۱۷۴۲۴	۱۶۹۰۰	۱۷۱۶۰	۱۳۲	۱۳۰	۱۷۴۲۴	۱۶۹۰۰	۱۷۱۶۰
۶	۱۳۸	۱۳۵	۱۹۰۴۴	۱۸۲۲۵	۱۸۶۳۰	۱۳۸	۱۳۵	۱۹۰۴۴	۱۸۲۲۵	۱۸۶۳۰
۷	۱۴۰	۱۴۵	۱۹۶۰۰	۲۱۰۲۵	۲۰۳۰۰	۱۴۰	۱۴۵	۱۹۶۰۰	۲۱۰۲۵	۲۰۳۰۰
۸	۱۴۱	۱۳۵	۱۹۸۸۱	۱۸۲۲۵	۱۹۰۳۵	۱۴۱	۱۳۵	۱۹۸۸۱	۱۸۲۲۵	۱۹۰۳۵
۹	۱۴۳	۱۴۸	۲۰۴۴۹	۲۱۹۰۴	۲۱۱۶۴	۱۴۳	۱۴۸	۲۰۴۴۹	۲۱۹۰۴	۲۱۱۶۴
۱۰	-	-	-	-	-	۱۶۴	۱۶۰	۲۶۸۹۶	۲۵۶۰۰	۲۶۲۴۰
جمع	۱۰۷۴	۱۰۷۴	۱۴۴۵۵۰	۱۴۴۷۸۰	۱۴۴۵۸۳	۱۳۴۳	۱۳۴۳	۱۸۲۴۷۱	۱۸۲۰۴۴	۱۸۲۱۶۳
میانگین	۱۳۴.۲۵	۱۳۴.۲۵	۱۸۰۶۸.۷۵	۱۸۰۹۷.۵	۱۸۰۷۲.۸	۱۳۴.۳	۱۳۴.۳	۱۸۲۴۷.۱	۱۸۲۰۴.۴	۱۸۲۱۶.۳
r	۰,۸۵					۰,۹۵				

$$r = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{[(n \sum X^2) - (\sum X)^2][(n \sum Y^2) - (\sum Y)^2]}}$$

$$= \frac{10(182163) - (1343)(1342)}{\sqrt{[(10 \times 182471) - (1343)^2] \times [(10 \times 182044) - (1342)^2]}} = \frac{19324}{20253.001} = 0.954$$

با توجه به نکته ای که به آن اشاره شد (اگر تفاوت بین نتایج به دست آمده خیلی نزدیک نباشد، محاسبه r، نتیجه لازم را ارائه نمی دهد)، مشخص است که در آزمایش دوم (نتایج ۱۰

دامنه ارزش های ضریب همبستگی از -1 تا $+1$ است یعنی . اگر دو صفت مستقل از هم باشند (نمودار A در شکل ۱)، مقدار

آزمایش ثانویه) وقتی که دو نتیجه اضافه شد، مقدار r ، از 0.85 به 0.95 رسید که به عدد یک بسیار نزدیک است. به طور کلی می توان گفت انتخاب محاسبه آماری ضریب همبستگی یا r برای این مسئله که تفاوت سدیم در روش A (در نتایج ۸ آزمایش اولیه) بین کمترین و بالاترین مقدار، 19 است ($124-143$)؛ از ابتدا صحیح نبوده است.

در اغلب تحقیقات علمی، برای یافتن رابطه بین دو متغیر (X و Y) از r یا ضریب همبستگی استفاده می کنند. اما توجه به این امر لازم و ضروری است که بدون در نظر گرفتن نکته بالا، مقدار ضریب همبستگی، ارزش واقعی نخواهد داشت. محاسبه r در مثال زیر با توجه به تفاوت کمی که میان بیشترین و کمترین مقدار، از هر دو متغیر همگلوبین (X) و تعداد گلبول های قرمز (Y) وجود دارد، منطقی و درست است.

نمونه	هموگلوبین (g/dl)	RBC (میلیون)	X^2	Y^2	XY
۱	۹۸	۴.۳	۹۶۰۴	۱۸.۴۹	۴۲۱.۴
۲	۱۲.۸	۵.۹	۱۶۳.۰۴	۳۴.۸۱	۷۵.۵۲
۳	۱۳.۲	۵.۳	۱۷۴.۲۴	۲۸.۰۹	۶۹.۹۶
۴	۱۵.۷	۶	۲۴۶.۴۹	۳۶	۹۴.۲۰
۵	۱۱.۳	۴.۹	۱۲۷.۶۹	۲۴.۰۱	۵۵.۳۷
۶	۱۴.۳	۵.۱	۲۰۴.۴۹	۲۶.۰۱	۷۲.۹۳
۷	۱۲.۹	۵.۳	۱۶۶.۴۱	۲۸.۰۹	۶۸.۲۷
۸	۱۳.۲	۶.۳	۱۷۴.۲۴	۳۹.۶۹	۸۳.۱۶
۹	۱۱.۲	۴.۹	۱۲۵.۴۴	۲۴.۰۱	۵۴.۸۸
۱۰	۱۲.۶	۵.۱	۱۵۸.۷۶	۲۶.۰۱	۶۴.۲۶
۱۱	۱۵.۵	۶.۳	۲۴۰.۲۵	۳۹.۶۹	۹۷.۶۵
۱۲	۱۳.۷	۴.۸	۱۸۷.۵۹	۲۳.۰۴	۶۵.۷۶
۱۳	۱۳.۴	۶.۵	۱۷۹.۵۶	۴۲.۲۵	۸۷.۱۰
۱۴	۱۳.۹	۶	۱۹۳.۲۱	۳۶	۸۳.۴۰
۱۵	۱۴	۵.۷	۱۹۶	۳۲.۴۹	۷۹.۸۰
Σ	۱۹۷.۵	۸۲.۴	۲۶۳۳.۴۵	۴۵۸.۶۸	۱۰۹۴.۵۰

با توجه به فرمول ضریب همبستگی داریم:

$$r = \frac{15(1094.5) - (197.5)(82.4)}{\sqrt{[(15 \times 2633.45) - (197.5)^2] \times [(15 \times 458.68) - (82.4)^2]}} = \frac{143.5}{211.69} = 0.67$$

در نتیجه بین پارامترهای هموگلوبین (متغیر X) و تعداد گلبول های قرمز خون (متغیر Y) همبستگی مستقیم و ناقص وجود دارد.

آنالیز رگرسیون خطی

Linear regression analysis

رگرسیون خطی یک ابزار حسابی است که به آنالیز کامل تر دیاگرام های پراکندگی کمک

می کند. در این آزمون آماری، پاسخ های مربوط به دو روش آزمایشگاهی بر روی یک نمودار با هم مقایسه می شود. نتایج مربوط به روش قدیمی بر روی محور X و نتایج مربوط به روش جدید بر روی محور Y ثبت می شود. در حقیقت رگرسیون خطی یک روش جهت تعیین بهترین همبستگی خطی بین متغیرهاست. در این آنالیز، خط مستقیمی (Best Fitted Line) بهترین شرایط همخوانی دو روش نسبت به یکدیگر را نشان می دهد. با توجه به آنالیز رگرسیون خطی، اطلاعات توصیفی جامعی درخصوص اختلاف دو روش با توجه به سبب و علت آن به دست می آید. رگرسیون خطی به صورت معادله زیر بیان می شود:

$$Y = mX + Y_0$$

که Y ، متغیر وابسته، X ، متغیر مستقل یا غیروابسته، m ، شیب خط و Y_0 مقدار Y در نقطه $X=0$ است.

اگر دو روش، همخوانی کاملی با یکدیگر داشته باشند در این صورت $m=1$ و $Y_0=0$ خواهد بود و در نتیجه معادله خط به صورت $X=Y$ در می آید (نمودار A در شکل ۲). در نمودار B از شکل ۲، ارزش های مربوط به روش جدید به صورت ثابتی بیشتر از روش قدیمی است. در این حالت با وجود اینکه شیب منحنی یک است اما $Y_0=5$ است. در نمودار C، ارزش های مربوط به روش جدید بیشتر از روش قدیمی است اما $Y_0=0$ است. در نمودار D با اینکه $Y_0=0$ است شیب خط کمتر از یک است ($m=0.9$) که حاکی از پایین بودن مقادیر روش جدید (مقادیر روی محور Y) نسبت به قدیم (مقادیر روی محور X) است. رگرسیون خطی برآورد صحیحی از شیب و Y_0 ارائه می کند. مناسب ترین خط در این نوع آنالیز، خطی است که جمع مربعات فواصل عمودی نقاط مورد مشاهده نسبت به آن را به حداقل برساند. برای محاسبه شیب رگرسیون از دو معادله زیر استفاده می شود:

$$m = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\Sigma(X-\bar{X})^2}$$

و یا

$$m = \frac{n \Sigma XY - (\Sigma X)(\Sigma Y)}{n \Sigma X^2 - (\Sigma X)^2}$$

همچنین مقدار Y_0 به طریق زیر محاسبه می شود:

$$Y_0 = \frac{\Sigma Y}{n} - m \left(\frac{\Sigma X}{n} \right)$$

و یا

$$Y_0 = \bar{Y} - m\bar{X}$$

رگرسیون خطی بر این اصل استوار است که هیچ گونه خطای اندازه گیری در روش قدیمی وجود ندارد و انحراف از معیار خط

گرایش به دلیل خطای اتفاقی مربوط به روش جدید است.

در ادامه با یک مثال عملی به توضیح موارد و نکات بیان شده می پردازیم: فرض کنید می خواهیم اندازه گیری هموگلوبین به دو روش SLS یا سدیم لوریل فسفات را با روش سیان مت هموگلوبین که روش مرجع است مقایسه کنیم. بدین منظور ۱۰ نمونه خون تازه را جمع آوری کرده و میزان هموگلوبین آنها را با هر دو روش گفته شده اندازه گیری می کنیم که مقادیر اندازه گیری شده در زیر آورده شده اند. در ادامه به رسم نمودار پراکنندگی، محاسبه ضریب همبستگی (r)، درصد همبستگی، رگرسیون خطی، انحراف معیار خط گرایش (خط رگرسیون) یا همان $S_{Y/X}$ و محدوده ۹۵ درصد می پردازیم.

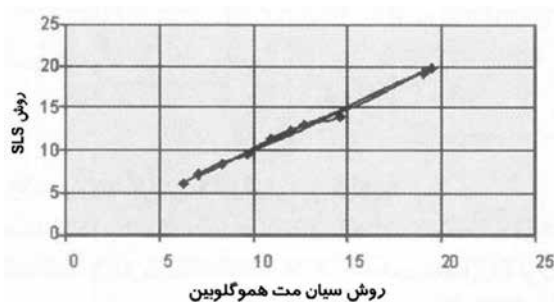
در جدول، متغیرهای X و Y به صورت زیر تعریف شده اند:

روش سیان مت هموگلوبین - X

روش SLS - نمودار پراکنندگی Y

نمونه	X	Y	X ²	Y ²	XY			
۱	۱۹.۱	۱۹.۲	۳۶۴.۸۱	۳۶۹.۶۴	۳۶۶.۷۲	۷.۰۷	۷.۰۰	۴۹.۹۸۴۹
۲	۹.۷	۹.۶	۹۴.۰۹	۹۲.۱۶	۹۳.۱۲	-۲.۳۳	-۲.۵۵	۵.۴۲۸۹
۳	۸.۳	۸.۵	۶۸.۸۹	۷۲.۲۵	۷۰.۵۵	-۳.۷۳	-۳.۶۵	۱۳.۹۱۲۹
۴	۷.۱	۷.۲	۵۰.۴۱	۵۱.۸۴	۵۱.۱۲	-۴.۹۳	-۴.۹۵	۲۴.۳۰۴۹
۵	۱۹.۵	۱۹.۹	۳۸۰.۲۵	۳۹۶.۰۱	۳۸۸.۰۵	۷.۴۷	۷.۷۵	۵۵.۸۰۰۹
۶	۶.۳	۶.۱	۳۹.۵۹	۳۷.۲۱	۳۸.۴۳	-۵.۷۳	-۶.۰۵	۳۲.۸۳۲۹
۷	۱۲.۷	۱۳.۱	۱۶۱.۲۹	۱۷۱.۶۱	۱۶۶.۲۷	-۰.۶۷	-۰.۹۵	۰.۴۴۸۹
۸	۱۱	۱۱.۴	۱۲۱	۱۲۹.۹۶	۱۲۵.۴۰	-۱.۰۳	-۰.۷۵	۱.۰۶۰۹
۹	۱۲	۱۲.۴	۱۴۴	۱۵۳.۷۶	۱۴۸.۸۰	-۰.۰۳	-۰.۲۵	۰.۰۰۰۹
۱۰	۱۴.۶	۱۴.۱	۲۱۳.۱۶	۱۹۸.۸۱	۲۰۵.۵۶	۲.۵۷	۱.۵۵	۶.۶۰۴۹
Σ	۱۲۰.۳	۱۲۱.۵	۱۶۳۷.۴۹	۱۶۷۷.۶۷	۱۶۵۴.۴۲	-	-	۱۹۰.۳۸۱
میانگین	۱۲.۰۳	۱۲.۱۵	-	-	-	-	-	-

نمودار پراکنندگی



در این نمودار پراکنندگی، اندازه گیری هموگلوبین به روش سیان مت هموگلوبین و SLS مقایسه شده است و چنانچه مشاهده می شود این دو روش از همبستگی خوبی برخوردارند.

ضریب همبستگی

$$r = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{[(n \sum X^2) - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}} = \frac{10(1654.42) - (120.3)(121.5)}{\sqrt{[(10 \times 1637.49) - (120.3)^2] \times [(10 \times 1672.25) - (121.5)^2]}} = \frac{1927.75}{1931.31} = 0.9981$$

پراکنندگی نقاط در اطراف خط گرایش، انحراف معیار خط گرایش (Standard deviation of the regression Line) نام دارد و آنرا با نماد $S_{Y/X}$ نشان می دهد. این پراکنندگی، خطای استاندارد برآورد نامیده می شود و از طریق رابطه زیر محاسبه می شود:

$$S_{Y/X} = \sqrt{\frac{\sum (y - \hat{y})^2}{n - 2}}$$

در نمودارهای E و F از شکل ۲، اثر افزایش پراکنندگی نقاط در اطراف خط گرایش نشان داده شده است. در این دو نمودار به ترتیب مقادیر ۲ و ۵ به صورت متناوب به ارزش های Y (نمودار A) اضافه و یا از آن کم شده اند. همان طور که ملاحظه می شود شیب (m) و Y تغییر نکرده اند و فقط $S_{Y/X}$ به ترتیب به ۲ و ۵ واحد افزایش یافته است. از آنجاییکه $S_{Y/X}$ برآورد انحراف معیار در اطراف خط گرایش است، محدوده های آماری برای هر نقطه ای روی این خط قابل محاسبه است. محدوده ۹۵ درصد (۹۵٪) برای هر ارزش Y خط گرایش عبارت است از:

$$Y \pm 1.96 S_{Y/X}$$

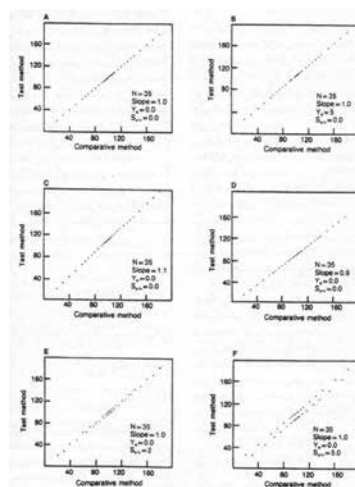
اگر $Y = mx + Y_0$ ، جایگزین Y شود، محدوده ۹۵٪ به صورت:

$$mx + Y_0 \pm 1.96 S_{Y/X}$$

تعریف می شود. لازم به ذکر است که محاسبه رگرسیون خطی تنها زمانی قابل استفاده و با ارزش است که متغیرها دارای همبستگی خطی باشند.

پیشتر به روش جدید و روش قدیم اشاره شد و بیان شد که نتایج مربوط روش قدیم (یا نخستین اندازه گیری یا اندازه گیری به روش اول) بر روی محور X و نتایج مربوط به روش جدید بر روی محور Y ثبت می شود. علت تفاوت روش ها از یکدیگر بروز خطا است. در مقایسه روش ها دو نوع خطا اندازه گیری می شود؛ خطای تصادفی و خطای سیستماتیک. خطای تصادفی در تمام اندازه گیری ها و به صورت شانسی مثبت و منفی بروز می کند. اندازه گیری پراکنندگی $S_{Y/X}$ برآوردی از خطای اتفاقی یا تصادفی به دست می دهد. خطای سیستماتیک باعث بروز خطا در یک جهت می شود. این خطا برخلاف خطای تصادفی، در یک روش نباید وجود داشته باشد. اندازه گیری Loca-

tion of slop یا موقعیت شیب و Y. در واقع اندازه گیری خطای سیستماتیک است.



شکل ۲. نمودارهای مقایسه روش ها

	X	y	Y	y-Y	(y-Y) ²
1	۱۹.۱	۱۹.۲	19.2907	-0.0907	0.00822
2	۹.۷	۹.۶	9.7967	-0.1967	0.03869
3	۸.۳	۸.۵	8.3827	0.1173	0.01375
4	۷.۱	۷.۲	7.1707	0.0293	0.00085
5	۱۹.۵	۱۹.۹	19.6947	0.2053	0.04214
6	۶.۳	۶.۱	6.3627	-0.2627	0.06901
7	۱۲.۷	۱۳.۱	12.8267	0.2733	0.07469
8	۱۱	۱۱.۴	11.1097	0.2903	0.08427
9	۱۲	۱۲.۴	12.1197	0.2803	0.07856
10	۱۴.۶	۱۴.۱	14.7457	-0.6457	0.41692
Σ					0.8271

درصد همبستگی

$$r^2 \times 100 = (0.9981)^2 \times 100 = 0.9962 \times 100 = \%99.62$$

محاسبه رگرسیون خطی

✓ محاسبه m با دو روش گفته شده

$$m = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sum(X-\bar{X})^2} = \frac{192.7935}{190.381} = 1.01$$

و

$$m = \frac{n \sum XY - (\sum X)(\sum Y)}{n \sum X^2 - (\sum X)^2} = \frac{10(1654.42) - (120.3)(121.5)}{10(1637.49) - (120.3)^2} = \frac{1927.75}{1902.81} = 1.01$$

✓ محاسبه Y₀ با دو روش

$$Y_0 = \frac{\sum Y}{n} - m \left(\frac{\sum X}{n} \right) = \frac{121.5}{10} - 1.01 \left(\frac{120.3}{10} \right) = -0.0003$$

و

$$Y_0 = \bar{Y} - m\bar{X} = 12.15 - 1.01(12.03) = -0.0003$$

رگرسیون خطی برابر است با:

$$Y = mX + Y_0$$

$$Y = -0.0003 + 1.01X$$

یافتن انحراف از معیار خط رگرسیون

برای محاسبه $S_{Y/X}$ به این صورت عمل می‌کنیم:
 ابتدا با توجه به معادله رگرسیون خطی که به دست آوردیم؛ تمام مقادیر Y را از این معادله به ازای قرار دادن X های جدول قبل به دست می‌آوریم.

برای مثال: برای نمونه ۱ از جدول قبل، داریم:

$$X = 19.1$$

$$Y = -0.0003 + 1.01(19.1) = 19.2907$$

سپس تمام مقادیر Y های جدول قبل (y) را از Y های به دست آمده از معادله رگرسیون خطی (Y) کم می‌کنیم (y-Y).

برای مثال: برای نمونه ۱ جدول قبل داشتیم:

$$y = 19.2$$

و Y به دست آمده از طریق معادله رگرسیون خطی برای همین نمونه برابر شد با:

$$Y = 19.2907$$

بنابراین

$$y - Y = 19.2 - 19.2907 = -0.0907$$

در نهایت پس از محاسبه، انحراف از معیار خط رگرسیون را از طریق رابطه زیر محاسبه می‌کنیم:

$$S_{Y/X} = \sqrt{\frac{\sum(y-Y)^2}{n-2}}$$

$$S_{Y/X} = \sqrt{\frac{\sum(y-Y)^2}{n-2}} = \sqrt{\frac{0.8271}{8}} = 0.32$$

با در دست داشتن $S_{Y/X}$ و معادله رگرسیون خطی می‌توان محدوده ی ۹۵ درصد را برای هر کدام از مقادیر X محاسبه کرد.

$$Y \pm 1.96 S_{Y/X}$$

$$mX + Y_0 \pm 1.96 S_{Y/X}$$

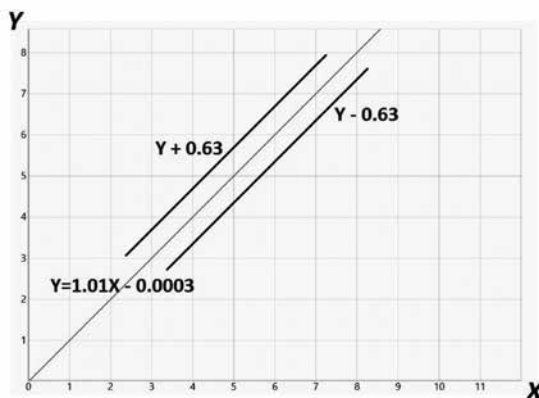
$$-0.0003 + 1.01X \pm 1.96 S_{Y/X}$$

برای بهتر مفهوم شدن موضوع به نمودار زیر توجه کنید:
 در این نمودار مشخص است که به ازای تمام مقادیر X ای که در معادله رگرسیون خطی قرار می‌گیرند، Y ها با ۹۵ درصد اطمینان در اطرف (بین) دو خط $Y + 0.63$ و $Y - 0.63$ قرار می‌گیرد.

$$Y \pm 1.96 S_{Y/X}$$

$$Y \pm 1.96 \times 0.32 = 0.63$$

$$Y \pm 0.63$$



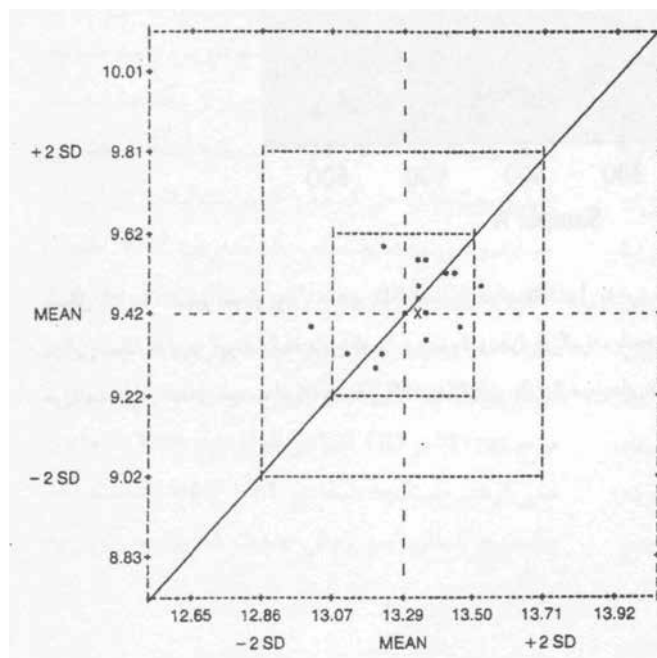
نمودارهای یودن (Youden Xy plots)

می‌دانیم برای تعیین میزان عدم دقت یک آزمایش می‌توان از دو سرم کنترل مختلف نرمال یا طبیعی و غیرنرمال استفاده

برخی از منابع

1. کتاب کنترل کیفیت در آزمایشگاه هماتولوژی، علی ملکی - دکتر سعید کاویانی
2. کتاب تجهیزات آزمایشگاهی، اصول فنی و نگهداری و روش های کنترل کیفی - مهندس سید بهزاد سیدعلیخانی - دانشگاه علوم پزشکی شهید بهشتی
3. کتاب مقدمات، آشنایی و اصول کلی آزمایشگاه تشخیص طبی - سعید طهماسبی
4. کتاب کنترل کیفی مواد و تجهیزات آزمایشگاهی - دکتر امیر سیدعلی مهبذ
5. دستورالعمل برخی استانداردهای ملی و بین المللی
6. کتاب مدیریت کنترل کیفیت متعادل برای آزمایشگاه تشخیص پزشکی - دکتر ملک پور و دکتر یوسفی و غیره ...

کرد. در نمودارهای یودن نیز از این مهم استفاده می شود. هنگامی که دو نمونه کنترلی در دو سطح مختلف (مثلا سطح طبیعی و سطح غیرطبیعی بالا) ران (Run) می شود، می توان چارت یودن را رسم کرد. این نمودار یک روش کارآمد در تفکیک خطاهای تصادفی از انحراف ثابت پاسخ ها یا همان خطاهای سیستماتیک است. با قرار دادن محدوده های $\pm 2SD$ میانگین \square از نمونه کنترل غیرطبیعی در محور عمودی و نیز محدوده های $\pm 2SD$ میانگین \square از نمونه کنترل طبیعی در محور افقی، نمودار طراحی می شود. سپس با متصل کردن نقاط $2SD$ به هم (چهار نقطه $\pm 2SD$) یک مربع حاصل می شود. خط موربی که نقطه $-2SD$ را به نقطه $+2SD$ متصل می کند، در حقیقت مسیر پراکندگی نرمال را مشخص می نماید (شکل ۳). در شرایط طبیعی داده های حاصل از کنترل ها بر روی این خط یا اطراف آن واقع خواهند شد و یا به صورت دسته ای بسیار نزدیک بهم در اطراف نقطه مرکزی تجمع می کنند. پراکندگی نابجای داده ها در گوشه های مربع مذکور نشان دهنده شیفت به بالا یا پایین در پاسخ هاست. نقاطی که خارج از این مربع قرار می گیرد، نشان می دهند که کنترل ها خارج از محدوده $\pm 2SD$ هستند (شکل ۳).



شکل ۳. نمودار یودن