فصل



مروری بر احتمال و آمار

❖ اهداف آموزشی

- ❖ آشنایی با مباحث احتمالات و آمار
- ❖ بررسی مراحل استنباط آماری
- ❖ آشنایی با ابزارهای مهم آمار توصیفی
- ❖ آشنایی با شاخص‌ها و ابزارهای آمار توصیفی و مفاهیم آن
- ❖ امید ریاضی و واریانس
- ❖ بررسی توزیع نرمال
- ❖ آشنایی با تابع‌های آماری مهم در نرم‌افزار Excell

⊡ مقدمه

آمار چیست و چگونه ما را در حل مسائل به ویژه مسائل مربوط به کنترل کیفیت یاری می‌کند؟ این اولین سؤالی است که در ذهن ما نقش می‌بندد. تکنیک‌های آماری یکی از مهمترین ابزارهای کیفیت می‌باشد و همانند روش‌های کنترل هزینه و کنترل مواد اهمیت دارد. تکنولوژی مدرن کامپیوتری با امکان‌پذیر نمودن جمع‌آوری و تحلیل فوری و پیوسته داده‌ها در مراکز QC، کاربرد تکنیک‌های آماری در هر نوع را آسان نموده است.

در ادامه با استفاده از تکنیک‌های آماری در فصل‌های آتی به این پرسش پاسخ خواهیم داد:

مثال ۱: محصولات فرایندی که تحت کنترل آماری نیست ممکن است مطابق با مشخصات باشند؟

مثال ۲: کشف انحرافات با دلیل در فرایند تولید چگونه انجام می‌شود؟

مثال ۳: کنترل فرایند تولید با روش‌های آماری چگونه انجام می‌شود؟

⊡ تعریف آمار و انواع آن

آمار مجموعه‌ای از روش‌ها را برای جمع‌آوری و خلاصه کردن داده‌ها، طبقه‌بندی آنها و روش‌های تحلیلی برای پیش‌بینی و برآورد[1]، تصمیم‌گیری در شرایط مختلف ارائه می‌دهد.

انواع آمار به دو دسته آمار توصیفی و استنباطی تقسیم می‌گردد. آمار توصیفی شامل روش‌هایی است که برای خلاصه کردن و رده‌بندی داده‌های موجود در مجموعه‌ای از داده‌ها، محاسبه مشخصات عددی این مجموعه و نمایش داده‌ها در قالب نمودار و شکل‌های مختلف به کار می‌رود.

مثال: فرض کنید بخواهیم اطلاعاتی در مورد قطعه نمدی خودرو، در خردادماه سال جاری به مدیریت ارائه دهیم. این اطلاعات می‌تواند شامل کل نمدهای پژو و میانگین مشخصات نمد، تعداد

---

1. Estimation

نمدهای ضایعاتی و میانگین ضایعات در شیفت‌های A و B بر اساس هر ایستگاه و تعداد کل قطعات سالم در آخرین مرحله کنترل کیفیت (SQFE) باشد. برای این منظور کافی است از روش‌های آمار توصیفی استفاده کرده، با تهیه جداول و نمودارهای آماری و محاسبه میانگین معدل‌ها، اطلاعات مورد نیاز را که به این ترتیب خلاصه شده‌اند را به مدیریت کارخانه ارائه کنیم.

با توجه به مثال گفته شده چون بررسی تمام عناصر جامعه ضرورت ندارد و از طرفی به دلیل کمبود امکانات، نمی‌توانیم داده‌های موجود در کل جامعه را بررسی کنیم، در این مرحله به سراغ آمار استنباطی می‌رویم.

آمار استنباطی شامل روش‌هایی است که با استفاده از آنها، اطلاعات موجود در نمونه را به کل جامعه تعمیم می‌دهیم. یکی از مهمترین هدف آمار، انجام استنباط‌هایی در مورد مشخصات جامعه، با توسعه به اطلاعات موجود در نمونه است. این دقیقاً همان مصداق ضرب‌المثل ایرانی مشت نمونه خروار را در ذهن ایجاد می‌کند.

▣ مراحل انجام استنباط آماری

معدودی از مسائل آماری در مرحله آمار توصیفی پایان می‌پذیرند، ولی بیشتر مسائل آماری، شامل استنباط دربارهٔ مشخصه یک جامعه، با استفاده از اطلاعات قابل دسترس در یک نمونه است. استنباط آماری را طی پنج مرحله می‌توان به انجام رسانید.

- اولین و مهمترین بخش یک مسأله آماری، شناسایی دقیق و روشن مسأله مورد نظر و جامعهٔ وابسته به این مسأله است.
- دومین مرحله، جمع‌آوری داده‌هاست. چون جمع‌آوری داده‌ها هزینه و وقت زیادی می‌خواهد، این مرحله اهمیت دارد. اگر اندازهٔ نمونه خیلی زیاد باشد هزینه بالا می‌رود و اگر اندازهٔ نمونه خیلی کم باشد، اطلاعات حاصله مفید نخواهند بود. بنابراین اندازهٔ نمونه باید به گونه‌ای انتخاب شود که داده‌های مورد نیاز با حداقل هزینه جمع‌آوری گردد.
- این مرحله تجزیه و تحلیل اطلاعات نمونه است که شامل رده‌بندی و خلاصه کردن مجموعهٔ داده‌ها و رسم نمودارها و شکل‌هاست.
- مرحلهٔ چهارم استفاده از اطلاعات موجود در نمونه و به دست آوردن نتایجی در مورد جامعه است که این نتیجه‌گیری‌ها یا استنباط‌ها می‌توانند شامل برآورد یا تصمیم‌گیری در مورد مشخصات جامعه باشند.

- مرحلهٔ نهایی ارائه پاسخی برای این سؤال است که «استنباط انجام شده چقدر مورد اطمینان است؟»

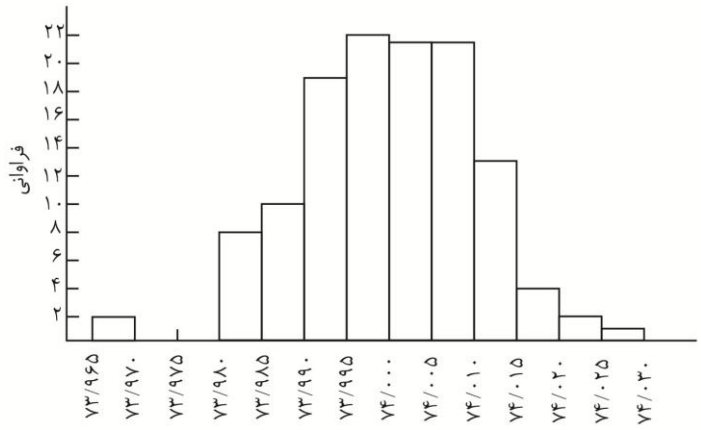
در این مرحله میزان اطمینان به استنباط انجام شده، در قالب معیاری عددی بیان می‌گردد.

با توجه به تفاوت بین آمار توصیفی و آمار استنباطی، جزئیات در ادامه بررسی خواهد شد.

⊡ آمار توصیفی و ابزارهای آن

هدف آمار استفاده از اطلاعات موجود در یک نمونه و استخراج نتایج و پیش‌بینی‌هایی در مورد جامعه است. آمار توصیفی شامل دو قسمت است. قسمت اول روش‌های توصیفی نموداری است و با استفاده از نمودار، شکل و جدول، **توصیفی ظاهری** از داده‌ها را ارائه می‌دهد. قسمت دوم شامل روش‌**های توصیفی عدد** داده‌هاست که برای خلاصه کردن و به دست آوردن مشخصات عددی داده‌ها از قبیل میانگین و انحراف معیار به کار می‌روند.

مثال:



شکل ۲- هیستوگرام برای داده‌های قطر رینگ پیستون

برای مثال، توزیع فراوانی داده‌های مربوط به رینگ پیستون در شکل (۲) هیستوگرام آنها را نشان می‌دهد. در این شکل، ملاحظه می‌شود که یک رینگ، قطری بین ۷۳/۹۶۵ ملی‌متر و ۷۳/۹۷۰ میلی‌متر دارد؛ هشت رینگ دارای قطرهای بین ۷۳/۹۸ میلی‌متر و ۷۳/۹۸۵ میلی‌متر هستند و بقیه نیز به همین ترتیب گروه‌بندی شده‌اند. ارتفاع هر ستون در شکل (۲) نمایانگر فروانی قطر رینگ در آن است.

در داده‌های مربوط به قطر رینگ پیستون، توزیع فراوانی قطر رینگ حدوداً متقارن بوده و تمایل مرکزی آن بسیار نزدیک به ۷۴ میلی‌متر است. به نظر می‌رسد تغییرپذیری موجود در قطر رینگ،

زیاد باشد. این امر بدین دلیل است که بعضی از قطرها به کوچکی ۷۳/۹۶۷ میلی‌متر و بعضی از آنها به بزرگی ۷۴/۰۳۰ میلی‌متر هستند. بنابراین، از هیستوگرام، اطلاعاتی به دست می‌آید که با بررسی داده‌های خام حاصل نمی‌شود. در ادامه وارد این ابزارها و تکنیک‌ها خواهیم شد.

■ **مراحل ساخت جدول فراوانی برای داده‌های طبقه‌بندی شده**

- ✧ کلیه اعداد به ترتیب از بزرگ به کوچک یا کوچک به بزرگ مرتب شوند.
- ✧ محاسبه برد (R) مجموعه ارقام جمع‌آوری شده

(بزرگترین رقم) – (کوچکترین رقم)

- ✧ محاسبه اندازه طبقه ارقام با استفاده از فرمول زیر:

$$\text{اندازه هر طبقه} = \frac{\text{برد}}{(\text{Log n})(۳/۳۲۲)+۱}$$

جدول فراوانی یکی از ساده‌ترین و متداول‌ترین جداول آماری است. برای محاسبه تعداد طبقات از دستور تقریبی $K=\sqrt{n}$ نیز می‌توان استفاده نمود.

فراوانی بر دو حالت فراوانی مطلق (F) و فراوانی نسبی (f) تقسیم می‌شود و داریم:

که n تعداد داده می‌باشد. $f=\frac{F}{n}$

لذا اگر داده‌های ما در K طبقه دسته‌بندی شده باشد ما فراوانی تجمعی (F_{C_j}) و فراوانی تجمعی نسبی (f_{C_j}) و رابطه آن را داریم:

$$F_{Cj}=F_۱+...+F_j \ , \quad f_{C_j}=f_۱+...+f_j$$

نکات مربوط به فراوانی بسیار بوده و به عنوان نمونه داریم، همواره جمع فراوانی‌های مطلق برابر با کل تعداد داده‌ها می‌باشد یعنی: $\sum_{i=۱}^{k} F_i = n$

نکته قابل توجه استفاده از جدول فراوانی در بررسی ضایعات خط تولید در کنترل کیفیت آماری و محاسبه شاخص مهم PPM می‌تواند باشد. این شاخص در صنایع خودروسازی به‌خصوص شرکت‌هایی که تأمین‌کننده برای خودروسازی کشور از جمله ایران خودرو و سایپا، کاربرد دارد و می‌توان توسط نرم‌افزار ماکروسافت Excell نیز رسم نمود. در واقع یکی از گزارشات مهم مدیریتی به حساب می‌آید که از سوی واحد کنترل کیفیت Qc تهیه می‌گردد. سایر ابزارهای آمار توصیفی در جدول (۱) خلاصه گردیده است:

جدول ۱- شاخص‌ها و روابط

عنوان	نام شاخص	فرمول	توضیحات
شاخص مرکزی	میانگین حسابی	$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$	میانگین نمونه را با $\bar{X}$ و جامعه را با μ نمایش می‌دهند.
	میانگین هندسی	$\mu_G = \sqrt[N]{x_1^{F_1}. \dots .x_N^{F_N}}$	در مواردی که داده‌ها به صورت دو عدد یا نرخ رشد یا شاخص باشد استفاده می‌شود.
	میانگین هارمونیک	$\mu_H = \frac{\sum F_i}{\frac{F_1}{x_1}+\dots+\frac{F_k}{x_k}}$	مقیاس داده‌ها ترکیبی باشد مانند متر به ثانیه از این میانگین استفاده می‌شود.
	میانه (Md) یا (Me)	زوج: $M_d = \frac{x_{\frac{N}{2}} + x_{\frac{N}{2}+1}}{2}$ فرد $N = Md = x(\frac{N+1}{2})$	در داده‌های طبقه‌بندی شده فرمول تغییر می‌کند.
	مد (Mo)	نما یک مجموعه، عددی است که در آن مجموعه از بقیه بیشتر تکرار شده باشد و منحصر به فرد نیست.	در داده‌های طبقه‌بندی شده فرمول تغییر می‌کند.
شاخص پراکندگی	دامنه تغییرات	$R = Max\, x_i - \min x_i$	در داده‌های پیوسته فرمول تغییر می‌کند.
	میانگین قدر مطلق انحرافات MAD	$MAD = \frac{\sum_{i=1}^{n} F_i \mid x_i - N \mid}{N}$	
	واریانس نمونه σ^2 یا S^2 یا Var (X)	$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} F_i (x_i - \bar{x})^2$	فرمول واریانس جامعه تغییر می‌کند.
شاخص نسبی پراکندگی	ضریب تغییرات (ضریب پراکندگی)	$CV = \frac{\sigma}{\mu}$	
	گشتاورها	$M_n(a) = \frac{\sum F_i (x_i - a)^n}{N}$	

		ضریب چولگی (SK)	به کتاب‌های آمار مراجعه گردد.	برای نشان دادن عدم تمرکز داده‌ها از لحاظ تقارن و مقایسه با توزیع نرمال استفاده می‌شود
		ضریب کشیدگی (E)	به کتاب‌های آمار مراجعه گردد.	معیاری برای بلندی یا کوتاهی نمودار توزیع داده‌ها را نشان می‌دهد.
رسم نمودارهای گوناگون	نمودار کمی	نمودار هیستوگرام		داده‌های این نمودار می‌توانید پیوسته یا گسسته باشد ولی داده‌ها باید کمی باشند.
		نمودار چندضلعی		با اتصال وسط مستطیل‌های هیستوگرام به دست می‌آید.
	نمودار وصفی	نمودار ستونی		محور عمودی این نمودار معمولاً فراوانی می‌باشد.
		نمودار دایره‌ای		قطاع‌های دایره نشان‌دهنده درصد می‌باشد.
		نمودار پاراتو		برای تحلیل در کنترل کیفیت کاربرد فراوانی دارد.
	تحلیل اکتشافی	نمودار شاخه و برگ		می‌توان انحراف چارکی (SIQR) حساب نمود.
		نمودار جعبه‌ای		از روی این نمودار متقارن بودن و چولگی قابل احصا می‌باشد.

در علم آمار به شاخص‌های میانگین و واریانس و انحراف معیار، آماره گفته می‌شود. دلیل این امر مشخص بودن مقدار آنها می‌باشد. رابطه تجربی بین میانگین، میانه و مد بستگی به توزیع داده‌ها دارد.
به طور مثال در حالت متقارن بودن توزیع، میانگین، میانه و مد هر سه با هم برابر می‌باشند.
چندک‌ها نیز جزء شاخص‌های مرکزی می‌باشند و داده‌ها را در فواصل مورد نیاز تقسیم می‌کنند و بسته به داده‌های گسترده و پیوسته روش محاسبه فرق می‌کند. تحلیل اکتشافی داده‌ها مربوط به شناسایی انفرادی داده‌ها می‌باشد.
یک کاربرد عملی انحراف معیار، قضیه چبیشف می‌باشد. این قضیه نشان می‌دهد که حداقل نسبت $\left[1-(\frac{1}{k^2})\right]$ از مقادیر موجود در جامعه در فاصله ($\mu+K\sigma$ و $\mu-K\sigma$) قرار دارد. مهمترین نتایج نظری در نظریه احتمال، قضایای حدی هستند. از این میان مهمترین آنها عبارتند از قضایایی که با عنوان قانون اعداد بزرگ یا قضایای حد مرکزی طبقه‌بندی شده‌اند. نامساوی مارکف، نامساوی

چبیشف نیز جزء این قضایای حدی هستند. از طرفی قواعد تجربی مربوط به شکل توزیع‌های احتمال نیز به ما در کنترل کیفیت کمک خواهند کرد.

⊡ نظریه احتمال و توزیع‌های احتمالی

همچنان که قبلاً اشاره گردید هدف علم آمار استفاده از اطلاعات موجود در نمونه برای به دست آوردن نتایجی در مورد جامعه است و چون نمونه بخشی از اطلاعات موجود در جامعه است و نه تمامی آن، لذا نمی‌توان به درستی نتایج حاصل به طور قطع اطمینان داشت. بنابراین به روش‌هایی نیاز داریم که میزان درستی نتایج را بتوان با به کارگیری آنها محاسبه نمود. نظریه احتمال چنین روش‌هایی را در اختیار می‌گذارد. مفاهیم اصلی و مورد نیاز آشنا بودن با تعاریف، آزمایش تصادفی، فضای نمونه، پیشامد، پیشامد ساده و مرکب، نمودار درختی و نمودار ون و احتمال یک پیشامد می‌باشد. شرایط مدل احتمال برای پیشامد ساده E_i به صورت زیر است:

۱- برای هر پیشامد ساده E_i : $0 \le P(E_i) \le 1$

۲- $P(S) = \sum_{i=1}^{K} P(E_i) = 1$

پیشامدهای مرکب به اجتماع و اشتراک و متمم یک پیشامد اشاره دارد. احتمال شرطی نیز در نمودارهای کنترل بسیار کاربرد داشته و با نماد $P(A|B)$ احتمال وقوع A به شرط رخداد پیشامد B را نشان می‌دهد. استقلال پیشامدها و ناسازگاری آنها و قانون ضرب احتمالات و جمع احتمالات و قانون بیز و قواعد شمارش نیز جزء مقدمات مباحث احتمال می‌باشد.

از آنجایی که نتایج آزمایش‌ها می‌تواند کمی یا کیفی باشد، می‌توان این نتیجه را با تعریف متغیر تصادفی برای نمایش نتایج کیفی نشان داد.

متغیر تصادفی، تابعی است که مقدار عددی‌اش به وسیله برآمد یک آزمایش تصادفی تعیین می‌شود. در علم آمار متغیر تصادفی را با حروف بزرگ X، Z و Y و مقادیر خاصی را که این متغیر می‌تواند بپذیرد با حرف کوچک x، z و y نشان می‌دهند. متغیر تصادفی به دو دسته گسسته و پیوسته تعلق دارد. از آنجایی که وقوع مقادیر متغیر تصادفی بسته به نوع آن دارد، لذا ما با دو دسته توزیع احتمالی مواجه هستیم. این توزیع‌های احتمالی به دو دسته توزیع‌های احتمالی گسسته و توزیع‌های احتمالی پیوسته با حفظ شرایط مدل احتمال دسته‌بندی می‌گردد.

تعاریف این متغیرهای به صورت زیر می‌باشد:

■ متغیرهای تصادفی گسسته

متغیر تصادفی X را گسسته گوییم اگر مجموعه مقادیری که این متغیر تصادفی اختیار می‌کند، یک مجموعهٔ قابل شمارش باشد.

برای متغیر تصادفی گسستهٔ X، تابع چگالی احتمال را به صورت زیر تعریف می‌کنیم.

$$A=\{x_1,x_2,...,x_n\} \quad P_X(x_i)=P(A=x_i), x_i\in A \qquad (۲-۴)$$

A: مجموعه مقادیری که متغیر تصادفی گسستهٔ X می‌تواند اختیار کند.

■ متغیرهای تصادفی پیوسته

تعریف: متغیر تصادفی X، یک متغیر تصادفی پیوسته است، اگر یک تابع حقیقی نامنفی وجود داشته باشد، به طوری که:

$$\forall B \quad B\subseteq R \quad p(X\in B)=\int_B f(x)\,dx \text{ و } D_f=R$$

f(x) را تابع چگالی احتمال متغیر تصادفی X می‌نامیم.

جدول (۲) برخی از توزیع‌های احتمالی پیوسته به همراه فرمول‌های آنها را نشان می‌دهد.

جدول ۲- خلاصه انواع توزیع‌های احتمال پیوسته

نام توزیع	تابع احتمال	میانگین	واریانس
توزیع یکنواخت پیوسته	$\begin{cases}\frac{1}{b-a}; a<x<b\\ 0 \quad ;ow\end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
توزیع گاما	$\begin{cases}\frac{1}{\beta^\alpha \Gamma(\alpha)}x^{\alpha-1}e^{\frac{-x}{\beta}}; x\ge\beta\\ 0 \quad ; x<\beta\end{cases}$	$\alpha\beta$	$\alpha\beta^2$
توزیع نمایی	$\begin{cases}\frac{1}{\theta}e^{\frac{-x}{\theta}}; x\ge 0\\ 0 \quad ; x<0\end{cases}$	θ	θ^2
توزیع نرمال	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-1}{2}(\frac{x-\mu}{\sigma})^2}$	μ	σ^2
توزیع نرمال استاندارد	$\frac{1}{\sqrt{2\pi}}e^{\frac{-1}{2}z^2}$	0	1

با استفاده از این توزیع‌ها می‌توان میانگین (امید ریاضی)، واریانس و انحراف معیار و سایر اندازه‌های توصیفی متغیر تصادفی را به دست آورد.

■ امید ریاضی و واریانس

امید ریاضی: امید ریاضی متغیر تصادفی X به صورت زیر تعریف می‌شود:

$$E(X)=\begin{cases}\sum x\, p_x(x) & ; \quad \text{اگر X گسسته باشد}\\ \int_{-\infty}^{+\infty} x\, f_x(x)\,dx & ; \quad \text{اگر X پیوسته باشد}\end{cases} \qquad (۱-۵)$$

در واقع E(X) متوسط وزنی مقادیر ممکن متغیر تصادفی X می‌باشد.

واریانس: اگر X یک متغیر تصادفی با میانگین μ باشد، آنگاه واریانس X که آن را با Var (X) نشان می‌دهیم به صورت زیر تعریف می‌شود:

$$Var(X)=E\left[(X-\mu)^{۲}\right]=E(X^{۲})-\left[E(X)\right]^{۲} \qquad (۲-۵)$$

نکته: اگر X متغیر تصادفی و a و b مقادیر ثابت باشند، داریم:

$$E(aX+b)=aE(X)+b$$

$$Var(aX+b)=a^{۲}Var(X)$$

امید ریاضی و واریانس متغیرهای تصادفی مهم:

۵-۱) امید ریاضی و واریانس متغیرهای تصادفی مهم

۱- اگر X یک متغیر تصادقی برنولی با پارامتر p باشد:

$$E(X)=p$$

$$Var(X)=p(۱-p)$$

۲- اگر X یک متغیر تصادفی دوجمله‌ای با پارامترهای n و p باشد:

$$E(X)=np$$

$$Var(X)=np(۱-p)$$

۳- اگر X یک متغیر تصادفی هندسی با پارامتر p باشد:

$$E(X)=\frac{۱}{p}$$

$$Var(X)=\frac{۱-p}{p^{۲}}$$

۴- اگر X یک متغیر تصادفی دو جمله‌ای منفی با پارامترهای r و p باشد:

$$E(X)=\frac{r}{p}$$

$$Var(X)=\frac{r(۱-p)}{p^{۲}}$$

۵- اگر X یک متغیر تصادفی فوق هندسی با پارامترهای N (تعداد کل توپ‌ها)، n (تعداد نمونه) و m (تعداد توپ‌های سفید) باشد و داشته باشیم $p=\frac{m}{N}$:

$$E(X)=np=\frac{nm}{N}$$

$$Var(X)=\frac{N-n}{N-۱}=np(۱-p)=\frac{N-n}{N-۱}.\frac{nm}{N}(۱-\frac{m}{N})$$

۶- اگر X یک متغیر تصادفی پواسون با پارامتر $\lambda(\lambda>۰)$ باشد:

$$E(X)=Var(X)=\lambda$$

۷- اگر X یک متغیر تصادفی پیوستهٔ یکنواخت روی فاصله (a و b) باشد:

$$E(X)=\frac{a+b}{۲}$$

$$Var(X)=\frac{(b-a)^{۲}}{۱۲}$$

۸- اگر X یک متغیر تصادفی نمایی با پارامتر λ باشد:

$$E(X)=\frac{۱}{\lambda}$$

$$Var(X)=\frac{۱}{\lambda^{۲}}$$

۹- اگر X یک متغیر تصادفی نرمال با پارامترهای μ و $\sigma^{۲}$ باشد:

$$E(X)=\mu$$

$$Var(X)=\sigma^{۲}$$

۱۰- اگر X یک متغیر تصادفی گاما با پارامترهای α و β باشد:

$$E(X)=\alpha\beta$$

$$Var(X)=\alpha\beta^{۲}$$

آشنایی با اصول آماری و احتمالات به ما کمک خواهد کرد تا بتوان به طور مثال ادعاهای تأمین‌کنندگان مواد اولیه در کارخانجات اعم از خورویی و غیر خودرویی را در قالب آزمون فرض‌های آماری و فاصله‌های اطمینان را اثبات کنیم. حتی می‌توان با کمک روش‌های رگرسیونی و طراحی آزمایش‌های (DOE)، بر روی وابستگی داده‌هایی که در محاسبات کنترل کیفیت و رد یا قبول خروجی فرایند استفاده شایانی نمود. به کارگیری این اصول آماری باعث هر چه توانمندتر عمل نمودن مهندسین کیفیت در کارخانجات و یا حتی کیفیت خدمات سازمانی خواهد شد.

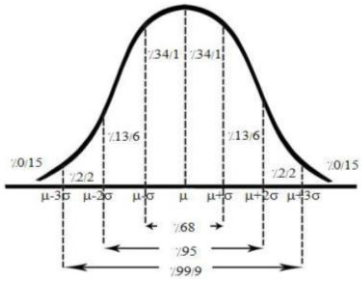
لازم به ذکر است که تعریف کیفیت صرفاً مخصوص کالا نمی‌باشد و می‌تواند شامل خدمت، مفهوم و... نیز باشد.

در چنین مواردی افق‌های پیش روی مهندسین کیفیت گسترده‌تر بوده و لذا نیازمند آگاهی از سایر تکنیک‌های آماری می‌باشد. لذا در این فصل خلاصه‌ای از این اصول آماری مطرح گردید و مابقی اصول آماری و جزئیات فرمول‌ها نیازمند مراجعه به کتب آماری می‌باشد.

در بخش بعدی پاره‌ای از مواد در نرم‌افزار آماری Minitab آموزش داده خواهد شد.

⊡ آشنایی با توزیع نرمال[1] و رفتار فرایندها طبیعی

نمودار منحنی توزیع احتمال نرمال به شکل زنگوله در شکل (۲) آمده است.



شکل ۲- منحنی توزیع احتمال نرمال

بسیاری از پدیده‌های طبیعی ساختار منحنی نرمال دارد مانند توانایی انسان در طول زمان، توزیع نمره‌های بسیاری از آزمون‌های روان‌شناختی نظیر روان‌شناختی نظیر آزمون هوش و پیشرفت تحصیلی و ...

1. Normal distribution

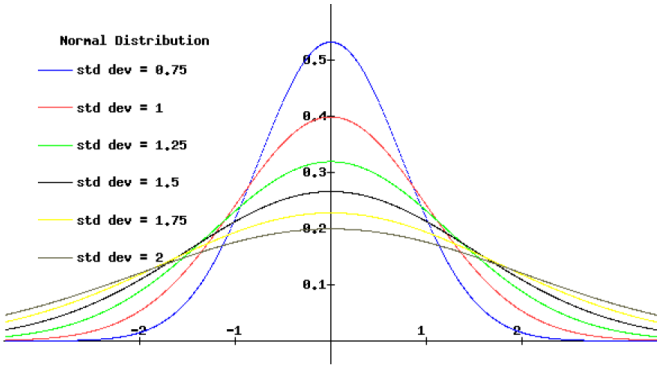
شکل توزیع نرمال به μ و σ بستگی دارد (μ را میانگین و σ را انحراف معیار توزیع نامند) توزیع نرمال را با نماد $N(\mu,\sigma^2)$ نشان می‌دهند. هر گاه بخواهیم بگوییم متغیری مانند X دارای توزیع نرمال است می‌نویسیم:

$$X \sim N(\mu,\sigma^2)$$

مثال: نماد $IQ \sim N(100,225)$ یعنی بهره هوشی گروهی از افراد جامعه دارای توزیع نرمال با میانگین ۱۰۰ و انحراف معیار ۱۵ است.

❐ منحنی نرمال استاندارد

اگر $X \sim N(\mu,\sigma^2)$ آنگاه $Z=\frac{X-\mu}{\sigma} \sim N(0,1)$ است. منحنی توزیع $N(0,1)$ را منحنی نرمال استاندارد گویند. این منحنی موارد استفاده زیادی در آمار دارد. زیرا پژوهشگر به جای کار کردن با بی‌نهایت منحنی که هر یک میانگین‌ها و انحراف معیارهای مختلف دارند و با یک منحنی که همیشه میانگین صفر و انحراف معیار یک دارد، روبرو است. (شکل ۴)



شکل ۳- منحنی نرمال استاندارد

❐ جدول توزیع نرمال استاندارد

این جدول شامل یک سطر و یک ستون است. رقم قبل از اعشار و بعد از اعشار در ستون و دو رقم بعد از اعشار در سطر قرار دارند مثلاً برای عدد ۲/۷۸ عدد ۲/۷ را در ستون و ۰/۰۸ را در سطر پیدا می‌کنیم. متناظر با هر عدد در فاصله (۰/۰۰ , ۳/۵۹) یک عدد در فاصله (۰/۵۰۰۰ , ۰/۹۹۹۰) وجود دارد. مثلاً برای عدد، عدد ۰/۹۹۷۳ پیدا می‌شود که تعبیر آن مساحت زیر منحنی از ابتدای

توزیع تا عدد ۲/۷۸ است (سطح هاشور خورده). همچنین این عدد احتمال $P(z<۲/۷۸)$ نیز است. برای محاسبات از جدول (۳) استفاده می‌گردد.

مثال: احتمالات زیر را با استفاده از جدول توزیع نرمال استاندارد بیابید.

$P(z<۱/۶۵)\ ,\ P(z<۰/۶۸)\ ,\ P(z<۲/۹۹)$

حل) جواب‌ها از چپ به راست عبارتند از: ۰/۹۵۰۵ , ۰/۷۵۱۷ , ۰/۹۹۸۶

■ **قواعد احتمال**

$$p(z>a)=۱-p(z<a)$$

$$p(z>۱/۸۵)=۱-p(z<۱/۸۵)=۱-۰/۹۶۷۸=۰/۰۳۲۲$$

$$p(z<-a)=۱-p(z<a)$$

$$p(z<-۱/۸۵)=۱-p(z<۱/۸۵)=۱-۰/۹۶۷۸=۰/۰۳۲۲$$

$$p(z>-a)=p(z<a)$$

$$p(z>-۱/۸۵)=p(z<۱/۸۵)=۰/۹۶۷۸$$

$$p(z<-a)+p(z>a)=۲(۱-p(z<a))$$

$$p(z<-۱/۸۵)=p(z<۱/۸۵)$$

$$=۲(۱-p(z<۱/۸۵))=۲(۱-۰/۹۶۷۸)=۰/۰۶۴۴$$

$$p(-a<z<a)=۲p(z<a)-۱$$

$$p(-۱/۸۵<z<۱/۸۵)=۲p(z<۱/۸۵)-۱=۲(۰/۹۶۷۸)-۱=۰/۹۳۵۶$$

$$p(a<z<b)=p(z<b)-p(z<a)$$

$$p(۰/۸۵<z<۲/۸۹)=p(z<۲/۸۹)-p(z<۱/۸۵)$$

$$=۰/۹۹۸۱-۰/۹۶۷۸=۰/۰۳۰۳$$

$$p(-۱/۸۵<z<۲/۸۹)=p(z<۲/۸۹)-p(z<-۱/۸۵)$$

$$=۰/۹۹۸۱-۰/۰۳۲۲=۰/۹۶۵۹$$

$$p(z>b \quad or \quad z<a)=p(z>b)+p(z<a)$$

$$p(z<-۱/۸۵ \ \ or \ \ z>۲/۸۹)=p(z<-۱/۸۵)+p(z>۲/۸۹)$$

$$=(۱-۰/۹۶۷۸)+(۱+۰/۹۹۸۱)=۰/۰۳۴۱$$

عمل دیگری که می‌توان با استفاده از جدول نرمال استاندارد انجام داد یافتن Z_α یا $Z_{\frac{\alpha}{۲}}$ است. α را سطح خطا نامند و مساحت زیر منحنی از نقطه Z_α تا انتهای منحنی توزیع نرمال استاندارد است. همچنین $۱-\alpha$ را سطح اطمینان نامند. برای یافتن Z_α، مقدار $۱-\alpha$ را یافته (در صورت عدم

وجود این عدد در جدول نزدیک‌ترین عدد به $1-\alpha$ را پیدا می‌کنیم) و در جدول عدد نظیر آن را در سطر و ستون پیدا می‌کنیم.

مثال: Z_α را برای $\alpha=۰/۰۵۹۴$ بیابید.

حل) $1-\alpha=۰/۹۴۰۶$ است. عدد متناظر با آن در جدول ۱/۵۶ است.

مثال: $1-\alpha=۰/۹۷$ است. این عدد در جدول وجود ندارد اما نزدیک‌ترین عدد به آن ۰/۹۶۹۹ است که متناظر با ۱/۸۸ است.

مثال: Z_α را برای $\alpha=۰/۰۳$ بیابید.

حل) $1-\alpha=۰/۹۷$ است. این عدد در جدول وجود ندارد اما نزدیک‌ترین عدد به آن ۰/۹۶۹۹ است که متناظر با ۱/۸۸ است.

مثال: Z_α را برای $\alpha=۰/۰۵$ بیابید.

حل) $1-\alpha=۰/۹۵$ است. این عدد در جدول وجود ندارد اما دو عدد ۰/۹۴۹۵ و ۰/۹۵۰۵ هر دو به یک اندازه به ۰/۹۵۰۰ فاصله دارند. بنابراین اعداد متناظر با این دو عدد را یافته و میانگین آنها Z_α است. پس $Z_\alpha=\frac{۱/۶۴+۱/۶۵}{۲}=۱/۶۴۵$

مثال: Z_α را برای سطح اطمینان ۹۹٪ بیابید.

حل) $1-\alpha=۰/۹۹۰۰$ و عدد متناظر آن ۲/۳۳ است.

مثال: $Z_{\frac{\alpha}{۲}}$ را برای $\alpha=۰/۰۵$ بیابید.

حل) $\frac{\alpha}{۲}=۰/۰۲۵$ و $1-\frac{\alpha}{۲}=۰/۹۷۵۰$ و عدد متناظر آن ۱/۹۶ است.

مثال: $Z_{\frac{\alpha}{۲}}$ را برای سطح اطمینان ۸۸٪ بیابید.

حل)

$$1-\alpha=۰/۸۸\Rightarrow\alpha=۰/۱۲\Rightarrow\frac{\alpha}{۲}=۰/۰۶\Rightarrow1-\frac{\alpha}{۲}=۰/۹۴$$

$$\Rightarrow Z_{\frac{\alpha}{۲}}=\frac{۱/۵۵+۱/۵۶}{۲}=۱/۵۵۵$$

بعضی از موارد کاربردی در جدول (۳) آمده است:

جدول ۳- جدول توزیع نرمال

$1-\alpha$	۰/۸	۰/۸۵	۰/۹	۰/۹۵	۰/۹۹
$Z_{\frac{\alpha}{2}}$	۱/۲۸	۱/۴۴	۱/۶۴	۱/۹۶	۲/۵۸

⊡ قضیه حد مرکزی

اگر از یک جامعه آماری با یک توزیع نامشخص، n نمونه تصادقی تهیه و میانگین هر نمونه را بیابیم بافت‌نگار فراوانی نسبی میانگین‌ها، وقتی n بزرگ می‌شود $(n \geq ۳۰)$ به منحنی نرمال نزدیک می‌شود.

جدول ۴- جدول توزیع نرمال



Probability Content from -oo to Z

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

⊡ تابع‌های آماری مهم در اکسل

در این قسمت توابع آماری نرم‌افزار اکسل مورد بررسی قرار گرفته است. سعی شده است از توضیحات اضافی اجتناب شود. در زیر به تعدادی از این توابع که مورد استفاده در آمار و کنترل کیفیت می‌باشد به همراه فرمول تابع و مثال آن در اکسل آورده شده است:

برای استفاده از توابع کافی است فرمول‌های مربوط به تابع را در قسمت نوار ابزار تابع اکسل وارد کرده و داده‌ها مورد نظر تابع را به شرح زیر وارد نموده تا بتوان مقدار تابع محاسبه شده را به دست آورد. در ادامه با ساده‌ترین تابع که تابع میانگین می‌باشد شروع می‌کنیم:

❐ Average: محاسبه میانگین

■ مثال در اکسل

	A
1	Data
2	10
	7
4	9
5	27
6	2

F rmula	Description (Result)
=AVERA E(A2:A6)	Average of the numbers above (11)
=AV RAGE(A2:A6 ،5)	Average of the numbers above and 5 (10)

فرمول تابع

نتیجه

❐ COVAR: محاسبه کوواریانس

فرمول آن به صورت زیر می‌باشد:

$$Cov(X,Y)=\frac{\sum(x-\bar{x})(y-\bar{y})}{n}$$

	A	B
1	Data1	Data2
2	3	9
3	2	7
4	4	12
6	5	15
		7

Formula	Description (Result)
COVAR(A2:A6 ، B2:B6)	Covariance ،the average of the products of deviations for each data point pair above (5. 2)

❐ Geomean: تعیین میانگین هندسی داده‌ها

فرمول آن به صورت روبه‌رو می‌باشد:

$$GM_{\bar{y}} = \sqrt[n]{y_1 y_2 y_3 \ldots y_n}$$

	A
1	Data
2	4
3	
4	
5	8
6	7
7	1
8	4
	3

Formula	Description esult)
=GEOMEAN(A2:A8)	Geometric mean of the data set above (5. 476987

❐ Harmean: محاسبه میانگین هارمونیک

A Da
4
5
8
7
1
4
3

Formula	Description (R sult)
=HARMEAN(A2:A8)	Harmonic mean of the data set above (5. 028376)

❐ Mode: تعیین مد داده‌ها

A Data
5. 6
4
3
2
4

Formula	Description (Result)
=MODE(A2:A7)	Mode ،or most frequently occurring number above (4)

❐ STDEV. P: محاسبه انحراف معیار جامعه

فرمول آن به صورت زیر می‌باشد:

$$\sqrt{\frac{\sum(x-\bar{x})^{2}}{n}}$$