بسم الله الرحمن الرحيم

دانشکده ریاضی و رایانه

بخش آمار

پایان نامه تحصیلی برای دریافت درجه کارشناسی ارشد رشته آمار ریاضی

کاربرد مدلهای رگرسیون خطی چندگانه در کنترل کیفیت آماری

مؤلف:

فاطمه اقدامی اسفقن سری

استاد راهنما :

دکتر علیرضا عربپور

استاد مشاور:

دکتر وحید امیرزاده

پاییز ۱۳۹۱

این پایان نامه به عنوان یکی از شرایط درجه کارشناسی ارشد

به

بخش آمار

دانشکده ریاضی و رایانه

دانشگاه شهید باهنر کرمان

تسلیم شده است و هیچگونه مدرکی به عنوان فراغت از تحصیل دوره مزبور شناخته نمی شود.

دانشجو : فاطمه اقدامی اسفقن سری

استاد راهنما : دکتر علیرضا عربپور

استاد مشاور : دکتر وحید امیرزاده

داور ۱ : دکتر ماشاالله ماشینچی

داور۲ : دکتر محسن مددی

نماینده تحصیلات تکمیلی دانشکده : دکتر مرجان کوچکی رفسنجانی

معاون آموزشی و پژوهشی دانشکده: دکتر محمدعلی ولی

حق چاپ محفوظ و مخصوص به دانشگاه شهید باهنر کرمان است.

تقدیم به:

همه کسانی که لحظه ای بعد انسانی و وجدانی خود را فراموش نمی کنند و بر آستان گران سنگ انسانیت سر فرود می آورند و انسان را با همه تفاوت هایش ارج می نهند...

تشکر و قدردانی:

سپاس وستایش خدای را که رحمت بی‌پایانش روشنی بخش دل ودیده ام گردید ومرا یاری نمود تا کارتدوین این پایان نامه را به انجام برسانم .

دراینجا مراتب سپاس وقدردانی خود را ازجناب آقای دکترعرب پور که راهنمایی این پایان نامه رابرعهده گرفتند و زحمات زیادی برای آن کشیدند ابراز می دارم. همچنین از آقای دکتر امیرزاده که در سمت استاد مشاور مرا در تدوین این پایان نامه یاری نمودند کمال تشکر را دارم. از آقای دکتر ماشین چی و دکتر مددی که داوری این پایان نامه را تقبل نمودند بینهایت سپاسگذارم. از دوست بسیار عزیزم خانم سمیه زرنگار که صادقانه در این راه همراه و در کنارم بودند، قدردانی می‌نمائیم. در انتها نیز از خانواده ی دلسوزم، همراهان همیشه مهربانم،که در هر قدمی حضورشان الهام بخش بود و وجودشان دلگرمی برای ادامه ی راه، صمیمانه سپاسگزاری و قدر دانی می نمایم.

چکیده

در برخی از کاربردهای کنترل کیفیت، کیفیت یک محصول یا فرایند با رابطه بین دو یا چند متغیر مشخص می‌شود که این رابطه "پروفایل" نامیده می‌شود. این روابط در دو فاز I و II نشان داده می‌شود. فاز I تحلیل داده ، هنگامیکه کیفیت یک فرایند یا محصول با یک مدل رگرسیون خطی چندگانه نشان داده شود، بررسی می‌گردد و در کنترل کیفیت آماری به عنوان تحلیل پروفایل خطی ارجاع داده می‌شود. این پایان نامه شامل چندین روش برای تحلیل پروفایلهای رگرسیون خطی ساده می‌باشد. در فاز II ، ۳ طرح نمودار کنترل برای نشان دادن پروفایلهای خطی ساده چند متغیره به کار برده می‌شود و عملکرد آماری روشهای پیشنهاد شده با استفاده از متوسط طول دنباله اندازه گیری می‌شود و بعلاوه قابلیت اجرایی روشها با استفاده از یک مثال کاربردی نشان داده می‌شود.

کلید واژه : نمودارهای کنترل T^2 چند متغیره؛ رگرسیون خطی چند متغیره؛ کنترل فرایند آماری؛ میانگین طول دنباله؛ پروفایلهای خطی چند متغیره؛ میانگین متحرک موزون به طور نمائی چند متغیره.

فهرست مطالب:

فصل اول: مفاهیم و تعاریف مقدماتی

۱-۱ مقدمه.....۲

۱-۲ تاریخچه ای از کنترل کیفیت.....۳

۱-۲-۱ اهمیّت روشهای آماری کنترل کیفیت.....۵

۱-۲-۲ آزمون فرضیه برای پارامترهای فرایند.....۵

۱-۲-۳ نقش انحرافات تصادفی با دلیل در تغییر پذیری کیفیت.....٦

۱-۲-٤ اصول آماری از نمودارهای کنترل.....۷

۱-۲-۵ انتخاب حدود کنترل.....۹

۱-۳ محاسبات مربوط به طول دنباله.....۱۰

۱-٤ رگرسیون.....۱۱

۱-٤-۱مقدمه.....۱۱

۱-٤-۲ تاریخچه.....۱۱

۱-٤-۳ رگرسیون خطی ساده.....۱۱

۱-٤-٤ رگرسیون خطی چندگانه.....۱۲

۱-٤-۵ برآورد کننده های حداقل مربعات.....۱۳

۱-٤-٦ برآوردهای کمترین مربعات از مدل رگرسیون خطی چندگانه و خاصیت آنها.....۱٤

۱-٤-۷ آزمون F جزئی.....۱۵

۱-۵ نمودار میانگین متحرک موزون نمایی(EWMA).....۱۷

۱-۵-۱ مثال.....۱۹

۱-٦ توصیف داده های چند متغیره.....۲۲

۱-٦-۱ توزیع نرمال چند متغیره.....۲۲

۱-٦-۲ بردار میانگین و ماتریس کوواریانس نمونه ای.....۲۲

۱-٦-۳ نمودار کنترلی T^2 هتلینگ.....۲۳

۱-٦-٤ برآورد μ , Σ.....۲۳

۱-۳ مثال.....۲۵

فصل دوم: فاز I تحلیل پروفایلهای رگرسیون خطی چندگانه

۲-۱ مقدمه..... ۲۹

۲-۲ پایش پروفایلهای خطی در فاز اول..... ۲۹

۲ -۲-۱ فاز اول در مقابل فاز دوم..... ۳۰

۲-۳ تعیین حدود کنترل..... ۳۱

۲-٤ مدل پروفایل رگرسیون خطی چندگانه فاز۱ ..... ۳۳

۲-۵ روشهای استفاده شده در پروفایل رگرسیون خطی فاز ۱..... ۳٤

۲-۵-۱ (روشA) نمودار کنترل T^2 کانگ و آلبین..... ۳۵

۲-۵-۲ (روش B) آزمون F کلی پیشنهادی محمود و وودال..... ۳۸

۲-۵-۳ (روش C) نمودار T^2 بر اساس اختلاف پی در پی..... ٤۱

۲-۵-٤ (روشD) روش جایگزین..... ٤۲

۲-٦ مقایسه عملکردها..... ٤٤

۲-۷ مثال عددی..... ۵۸

۲-۷-۱ مثال با داده شبیه سازی شده..... ٦۱

فصل سوم: فاز II تحلیل پروفایلهای رگرسیون خطی چندگانه

۳-۱ مقدمه..... ٦۵

۳-۲ مدلسازی پروفایل خطی ساده چند متغیره..... ٦٦

۳-۳ فاز II طرح های نمودارهای کنترل..... ۷۰

۳-۳-۱ روش۱ : نمودار کنترل میانگین متحرک موزون نمایی چند متغیره منفرد..... ۷۰

۳-۳-۲ روش۲: ترکیب نمودار کای دو و نمودار کنترل میانگین متحرک موزون نمایی چند متغیره منفرد..... ۷۱

۳-۳-۳ روش۳ : نمودار کنترل نمودار کنترل میانگین متحرک موزون نمایی چند متغیره منفرد سه تایی..... ۷۲

۳-۱ قضیه..... ۷۳

۳-٤ مقایسه متوسط طول دنباله ها..... ۷٦

۳-۱ مثال..... ۸۷

نتیجه گیری.....۹۲
پیشنهاد.....۹۳

پیوست

پیوست الف : توزیع T^2 هتلینگ.....۹۵
پیوست ب : برنامه های R به کار گرفته شده برای رسم نمودارها ی فصل دوم.....۹۷
پیوست پ : برنامه های R به کار گرفته شده برای جداول فصل سوم.....۱۲۵
منابع و مراجع.....۱۳۰
چکیده انگلیسی.....۱۳۲

فصل اوّل

مفاهیم و تعاریف

مقدماتی

۱-۱ مقدمه

از زمانی که انسان توانایی ساخت محصولی را پیدا کرد جهت کنترل کیفیت آن نیز به تلاش پرداخت. ظرافت، دقت و هنرمندی خاصی که در آثار باستانی و محصولات متعلق به زمانه ای نه چندان دور مشاهده می‌شود حکایت از این تلاش دارد. با شروع انقلاب صنعتی در اواسط قرن هیجده، ماشین های تولیدی در روندی تدریجی جایگزین ابزار و مهارت افراد هنرمند و صنعتگر گردید. با پیدایش روشهای جدید و پیچیده تولید، اشتیاق به تولید بیشتر افزایش یافت و بدین ترتیب نیاز به کنترل کیفیت محصول نهایی نیز ابعاد تازه ای پیدا نمود.

روش‌های اولیه کنترل کیفیت، هر چند به نوبه خود موثر بودند امّا با روش‌های آماری امروز کنترل کیفیت، تنها شباهت اندکی دارند. پایه و اساس کنترل کیفیت آماری به مفهوم امروزی آن تنها در سال ۱۹۲۰ گذاشته شد. برای فهم بهتر واژه کنترل کیفیت، تعریف جداگانه کلمات کیفیت و کنترل لازم است. کیفیت یعنی شایستگی جهت استفاده به خصوص و میزانی است که محصول انتظارات مصرف کننده خود را برآورده می‌سازد و کنترل به معنی اعمال ضوابط و راهنمایی‌ها در مورد کسی یا چیزی جهت اطمینان از کسب نتایج مورد نظر می‌باشد. معنی کنترل کیفیت تعاریف جداگانه دو واژه فوق را در بر می گیرد.

در صنعت برداشت‌های مختلفی از کنترل کیفیت وجود دارد. بعضی از شرکت‌ها که تنها کار آن‌ها جهت اطمینان از کیفیت محصول و انجام عملیات بازرسی است، به این عملیات کنترل کیفیت می‌گویند. واژه کنترل کیفیت علاوه بر عملیات بازرسی به سایر عملیات از قبیل برنامه ریزی کیفیت، کنترل کیفیت در حین تولید، کنترل موارد ورودی، تجزیه و تحلیل و اقدام راهگشا در رابطه با نقص‌های تولید و تهیه گزارشات مربوط به مسایل کیفی نیز گفته می‌شود. کنترل کیفیت نباید به کاربرد محدود واژه آن منحصر گردد بلکه باید تمامی فعالیت‌های لازم جهت بدست آوردن سطح مطلوبی از درستی و بی عیبی محصول را شامل شود. کنترل کیفیت سیستمی است جهت رسیدن به سطح مطلوبی از کیفیت یک محصول یا یک فرایند تولید و نگهداری آن با برنامه ریزی دقیق، استفاده از ماشین آلات مناسب، بازرسی مستمر و عمل اصلاح کننده، هرگاه که لازم باشد. کیفیت یک محصول معمولاً در رابطه با سه عامل زیر تعیین می‌شود :

- کیفیت طرح
- کیفیت انطباق
- کیفیت عملکرد

عوامل فوق لزوماً با هزینه و اقتصاد کیفیّت ارتباط نزدیکی دارند که در ادامه به توضیح این عوامل می‌پردازیم:

–کیفیت طرح

دو محصول ممکن است برای کار مشابهی استفاده شوند ولی در طراحی آن‌ها اختلاف زیاد وجود داشته باشد. یک محصول ممکن است از پلاستیک معمولی و از قطعاتی ناجور ساخته شده باشد که مسلماً در این صورت عمر استفاده از آن هم کوتاه می‌باشد. محصول دیگر ممکن است به خوبی طراحی شده باشد و درآن مواد خوب ومقاوم به کار رفته و قطعات آن با تلورانس های دقیق ساخته شده باشد و تمام علائم استفاده طولانی را داشته باشد و به وضوح مشخص است که کیفیت طراحی محصول دوم به مراتب برتر از کیفیت طراحی محصول اول می‌باشد.

–کیفیت انطباق

کیفیت انطباق درجه هم سویی محصول با مشخصات، استانداردها و معیارهای تعیین شده برای ساخت آن محصول است. محصولی که طبق مشخصات و مطابق با حدود کنترل فرایند تولید ساخته می‌شود، چنانچه مشخصات آن به خوبی بیانگر نیازهای مصرف کننده باشد از کیفیت خوبی برخوردار بوده و رضایت مشتری را جلب می‌کند.

–کیفیت عملکرد

کیفیت عملکرد یک محصول تابعی از کیفیت طرح و کیفیت انطباق آن محصول است. حفظ سطح بالایی از هر دو کیفیت مذکور، به سطح بالایی از کیفیت عملکرد منجر می‌شود. اگر کیفیت طرح ضعیف باشد یا تطابق با مشخصات مناسب وجود نداشته باشد، کیفیّت عملکرد محصول تنزل می‌یابد. مسلماً داشتن بالاترین سطح تطابق با مشخصات محصول چنانچه طراحی محصول ضعیف باشد، فایده ای نخواهد داشت [۱].

۱-۲ تاریخچه ای از کنترل کیفیت

کیفیت همانطورکه از نامش پیداست شایستگی جهت استفاده تعریف می‌گردد و شایستگی جهت استفاده، یعنی توانایی فرآورده کردن ماده اولیه با حداقل قیمت ، دورریزی و دوباره کاری.کاربرد روشهای آماری در تولید و تضمین کیفیت از تاریخچه نسبتاً قدیمی برخوردار است . درسال ۱۹۲٤ والتر شوهارت از آزمایشگاههای تلفن بل اساس نمودار کنترل را پایه گذاری کرد. به طورکلی این مقطع، نقطه آغاز کنترل فرایندآماری شناخته می‌شود. اواخر دهه ۱۹۲۰ هارولداف داج[۱] و

[۱] *Haroldaf Dodge*

هارولدجی[1] رومیگ که هر دو در آزمایشگاههای تلفن بل مشغول به کار بودند روشهای نمونه گیری جهت پذیرش آماری را به جای بازرسی صد در صد معرفی نمودند. در اواسط دهه ۱۹۳۰ فنون کنترل فرایند آماری به طور نسبتا وسیعی در شرکت وسترن الکتریک که شاخه تولیدی سیستم بل بود به کار گرفته شد و این در حالی بود که ارزش و اهمیّت کنترل کیفیت آماری هنوز برای صنایع مشخص نگردیده بود که بعدها در زمان جنگ جهانی دوم بحث کنترل کیفیت آماری در صنایع به وجود آمد.

مجموعه کارهای شوهارت، داج و رومیگ اساس علمی را تشکیل می‌دهد که امروزه کنترل کیفیت آماری خوانده می‌شود. این افراد در دهه ۱۹۳۰ با همکاری جامعه ی آمریکایی برای آزمایش مواد، انجمن استانداردهای آمریکا وجامعه مهندسین مکانیک آمریکا کوشش های خود را برای معرفی روش های جدید آماری شروع کردند. صنایع آمریکا در اوائل علی رغم تبلیغات وسیعی که در مورد روشهای جدید صورت می‌گرفت به سختی حاضر به پذیرش آن بود. پروفسور فریمن که در انیستیتو تکنولوژی ماساچوست $(M.I.T)$ جهت ترویج روشهای آماری کنترل کیفیت تلاش می‌کرد، این عدم استقبال را به علل زیر نسبت داده است :

الف) اعتقاد راسخ مهندسین تولید مبنی بر اینکه وظیفه اصلی آنها تکمیل روشهایی فنی تا حدی است که هیچ گونه تغییر مهّمی در کیفیت محصولات تولیدی به وجود نیاید، و دیگر آنکه نظریه احتمالات و تغییرات تصادفی جایگاه مناسبی در روشهای تولید نمی‌توانند داشته باشند.

ب) مشکل پیدا کردن کارشناسان آمار صنعتی است که در این زمینه تقریباً پیچیده ، آموزشهای لازم را دیده باشند. تا سال ۱۹۳۷ تعداد مراکز صنعتی در آمریکا که روشهای جدید را پذیرفته بودند احتمالاً از ۱۲ عدد تجاوز نمی‌کرد. برخورد سرد صنایع آمریکا نسبت به روشهای آماری کنترل کیفیّت به سرعت در دوران جنگ جهانی دوم سپری شد. شروع جنگ در سال ۱۹۳۹ آمریکا را به فکر افزایش تجهیزات و نیروهای مسلح خود انداخت و به زودی نیروهای مسلح بصورت بزرگترین مشتری صنایع این کشور در آمدند و نیاز روز افزونی بر استانداردهای کیفیّت پیدا کردند. نقش ارتش در پذیرش کنترل کیفیت آماری دو جنبه داشت؛ اول اینکه نیروهای مسلح خود، روشهای علمی بازرسی نمونه ای را پذیرفته بودند و اولین قدم در این زمینه بلافاصله بعد از وارد شدن آمریکا در جنگ برداشته شد. به دعوت دولت گروهی از مهندسین برجسته آزمایشگاههای تلفن بل جهت تدوین یک برنامه بازرسی نمونه گیری برای اداره تدارکات ارتش مشغول به کار شدند. جداول بازرسی نمونه گیری اداره تدارکات ارتش و نیروهای مسلح که در

---

[1] *Horoldgi roming*

سال۱۹۴۳ منتشر گردید نتیجه کار این افراد بود. همین گروه کوشش وسیعی برای آموزش کارکنان دولت در استفاده ازجداول و روشهای جدید به راه انداختند.
دومین جنبه نفوذ ارتش آموزشی وسیع برای علاقه مندان وکارکنان صنایع بود. از اوائل سال ۱۹٤۰ ، انجمن استانداردهای آمریکا به دعوت وزارت جنگ پروژه ای را شروع کرد که سرانجام منجر به تدوین استانداردها و راهنمای کنترل کیفیّت و روشهای نمودارهای کنترل برای تجزیه و تحلیل داده ها و روش نمودارهای کنترل برای کنترل کیفیت در حین تولید گردید. مطالب فوق به عنوان مواد درسی دوره های آموزشی به کار گرفته شد[۱].

۱-۲-۱ اهمیّت روشهای آماری کنترل کیفیت

مؤثرترین راهی که تا کنون برای کنترل کیفیت محصولات پیدا شده است روشهای آماری می-باشد. با روشهای آماری می‌توان تصویری از وضعیت کل یک تولید بدست آورد. با توجه به اینکه تغییرپذیری، یک پدیده دائمی و جزء لاینفک همه محصولات است و مشخصه کیفی هر محصول تغییر می‌کند، روشهای آماری مؤثرترین وسیله بررسی و کنترل این تغییرات است. مادامی که از مواد، افراد، روشها و ماشین ها برای تولید استفاده می‌شود، مشکل تغییر کیفیت وجود خواهد داشت و مادامی که این مشکل وجود داشته باشد، روش های آماری کنترل کیفیت نیزلازم می-شود.
به طور کلی روشهای آماری کنترل کیفیت بیشتر در کنترل کیفیت محصولاتی که به صورت انبوه تولید می‌شوند مؤثر می‌باشند. هرگاه تولید تکراری باشد، چه بصورت جریان پیوسته و ساخت پی در پی محصولات ، روش های آماری کنترل کیفیت بیشترین استفاده را پیدا می‌کنند.
اگر قرار باشد که یک محصول مشخصات مورد نظر مشتری را دارا باشد آنگاه این محصول باید به وسیله یک فرایند پایدار یا تکرار پذیر تولید گردد. به عبارت دیگر، فرایند باید از تغییر پذیری کمی در حول مقدار هدف یا ابعاد اسمی مشخصات کیفی محصول برخوردار باشد. کنترل فرایند آماری (SPC) مجموعه ای قدرتمند وتوانا از ابزار حل مشکل است که در ایجاد ثبات در فرایند و بهبود کارایی آن از طریق کاهش تغییر پذیری مفید واقع می‌گردد[۱].

۱-۲-۲ آزمون فرضیه برای پارامترهای فرایند

یک فرضیه آماری جمله یا بیانی درباره مقادیر پارامترهای یک توزیع احتمال است. به منظور انجام یک آزمون فرضیه، یک نمونه تصادفی از جامعه مورد نظر تهیه، آماره آزمون مناسبی محاسبه و در نهایت نتیجه گیری می‌شود که آیا فرضیه مورد نظر رد می‌شود یا نه.

اساساً نمودار کنترل یک آزمون فرضیه است. یک نقطه که بین حدود کنترل رسم شود مانند این است که فرضیه "فرایند تحت کنترل آماری است"و از طرف دیگر رسم یک نقطه خارج از حدود کنترل بیانگر رد چنین فرضیه‌ای است. می‌توان فرضیه‌ها را به صورت زیر نشان داد:

$$\begin{cases} H_0: \text{فرایند تحت کنترل به سر می‌برد} \\ H_1: \text{فرایند تحت کنترل به سر نمی‌برد} \end{cases}$$

دو نوع خطا در زمان آزمون فرضیه‌ها ممکن است رخ دهد. اگر فرضیه خنثی رد شود، وقتی که صحیح است خطای نوع اول رخ داده است. اگر فرضیه خنثی رد نشود، وقتی که اشتباه است خطای نوع دوم رخ داده است. احتمال هر یک از این خطاها را می‌توان به صورت زیر تعریف کرد:

$$\alpha = P\{\text{خطای نوع اول}\} = P\{H_0 \text{ رد} \mid H_0 \text{ صحیح باشد}\}$$

$$\beta = P\{\text{خطای نوع دوم}\} = P\{H_0 \text{ عدم رد} \mid H_0 \text{ اشتباه باشد}\}$$

۱-۲-۳ نقش انحرافات تصادفی و با دلیل در تغییرپذیری کیفیت

در هر فرایند تولید، همیشه صرف‌نظر از طراحی خوب یا نگهداری مناسب از آن، مقدار خاصی از تغییرپذیری به طور ذاتی وجود دارد. این تغییرپذیری ذاتی یا اختلال، در اثر انباشته شدن مجموعه زیادی از انحراف ها کوچک و غیر قابل اجتناب به وجود می‌آید. اگر اختلالات موجود در یک فرایند کوچک باشد آنگاه عملکرد فرایند از لحاظ تغییرپذیری قابل قبول خواهد بود. در ساختار کنترل کیفیت آماری، این تغییرپذیری ذاتی را معمولاً به عنوان یک سیستم پایدار «انحرافات تصادفی» می‌شناسیم. فرایندی که فقط در حضور انحرافات تصادفی عمل کند را فرایند تحت کنترل آماری می‌نامند. به عبارت دیگر، انحرافات تصادفی بخش جدایی ناپذیر فرآیند محسوب می‌گردند.

گونه‌های دیگر از تغییرپذیری ممکن است گاه‌گاهی در خروجی یک فرایند مشاهده گردد. تغییرپذیری در مشخصات کیفی کلیدی معمولاً از سه منبع سرچشمه می‌گیرد: تنظیم نادرست دستگاه، خطاهای اپراتور و یا مواد اولیهٔ معیوب.

به طور کلی، یک چنین تغییرپذیری در مقایسه با اختلالات ذاتی موجود در فرآیند بزرگ‌تر است و معمولاً بیانگر سطح غیر قابل قبولی برای عملکرد فرآیند می‌باشد. این منابع ایجاد تغییرپذیری که بخشی از انحرافات تصادفی محسوب نمی‌گردند را انحرافات با دلیل نامند. فرآیندی که در حضور انحرافات با دلیل عمل می‌کند را فرآیند خارج از کنترل می‌نامند.

اغلب فرایندهای تولید معمولاً در حالت تحت کنترل به سر می‌برند و این باعث می‌گردد تا برای مدت طولانی محصولات قابل قبول تولید شود. با این حال در بعضی مواقع انحرافات با دلیل به وجود می‌آیند و باعث می‌گردند تا فرایند به حالت خارج از کنترل تغییر پیدا کند.

یکی از اهداف اصلی کنترل فرایند آماری پی بردن سریع به وجود انحرافات با دلیل یا تغییرات در فرایند است تا قبل از اینکه تعداد زیادی از محصولات معیوب تولید شود علل ایجاد چنین انحرافاتی بررسی و اقدامات اصلاحی انجام گیرد. نمودار کنترل یکی از روش‌های کنترل فرایند در حین تولید است که برای این منظور مناسب خواهند بود. نمودارهای کنترل را نیز می‌توان برای تخمین پارامترهای یک فرایند تولید استفاده نمود و اطلاعات حاصل را به منظور تعیین کارایی فرایند بکار گرفت. نمودار کنترل را همچنین می‌توان به منظور تهیه اطلاعات مفید جهت بهبود فرایند استفاده کرد.

هدف اصلی کنترل فرایند آماری حذف تغییرپذیری فرایند است. این ممکن نیست که بتوان کل تغییرپذیری فرایند را حذف نمود، اما نمودار کنترل را می‌توان به عنوان ابزار مؤثری جهت کاهش تغییرپذیری فرآیند استفاده نمود[۱] .

۱-۲-٤ اصول آماری از نمودار کنترل

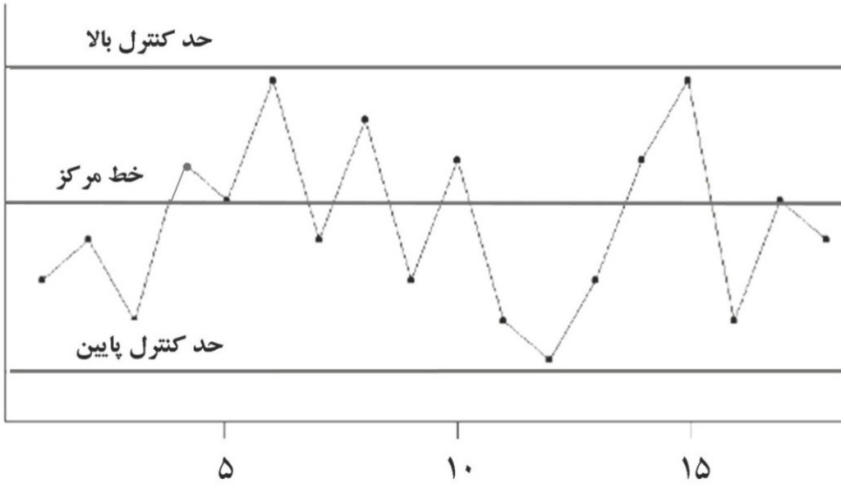
یک نمونه از نمودار کنترل در شکل ۱-۱ نشان داده شده است. این نمودار روش ارائه یک مشخصه کیفی که بر اساس اطلاعات نمونه، اندازه گیری یا محاسبه شده است را بر حسب نمونه یا زمان نشان می‌دهد.

نمودار شامل یک خط مرکز[۱] است که مقدار متوسط مشخصه کیفی را در حالت تحت کنترل نشان می‌دهد و یا به عبارت دیگر مرحله ای از فرایند را نشان می‌دهد که فقط خطاهای تصادفی

---

[۱] Center Line

حضور دارند. دو خط افقی دیگر که حد کنترل بالا (UCL)[1] و حد کنترل پایین (LCL)[2] نامیده می‌شوند در این نمودار نشان داده شده‌اند. این حدود کنترل به گونه ای انتخاب شده‌اند که اگر فرایند تحت کنترل باشد آنگاه تقریباً کلیهٔ نقاطی که بر اساس اطلاعات نمونه محاسبه شده‌اند بین این حدود واقع می‌شوند. تا زمانی که نقاط بین حدود کنترل قرار می‌گیرند، فرض می‌شود که فرایند تحت کنترل است و نیازی به فعالیت‌های اصلاحی نیست. اگر نقطه ای خارج از حدود کنترل رسم شود، نتیجه گیری می‌شود که فرایند در شرایط خارج از کنترل به سر می‌برد و اقدامات اصلاحی نیاز است تا منبع ایجاد انحراف یا انحراف با دلیل تعیین و حذف گردد. حتی اگر کلیهٔ نقاط در داخل حدود کنترل به گونه ای رسم شوند که از یک حالت سیستماتیک یا غیر تصادفی برخوردار باشند آنگاه فرایند خارج از حدود کنترل محسوب می‌گردد.



شکل۱-۱. یک نمودار کنترل

اگر فرایند تحت کنترل باشد باید روند نقاط بر روی نمودار کنترل به صورت تصادفی باشد. روش‌های بررسی روندهای غیر تصادفی را می‌توان در مورد نمودارهای کنترل جهت پی بردن به حالت خارج از کنترل استفاده نمود. معمولاً دلیلی برای پدیدار گشتن یک روند خاص غیر تصادفی بر روی نمودار کنترل وجود دارد و اگر بتوان آن را شناسایی و حذف نمود آنگاه عملکرد فرایند بهبود می‌یابد.

رابطهٔ بسیار نزدیکی بین نمودارهای کنترل و آزمون فرضیه وجود دارد. اساساً نمودار کنترل یک آزمون فرضیه است که به منظور ارزیابی شرایط تحت کنترل بودن فرایند از لحاظ آماری، استفاده

[1] Upper Control Limit
[2] Lower Control Limit

می‌گردد. یک نقطه بین حدود کنترل رسم می‌شود مانند این است که فرضیه مربوط به حالت تحت کنترل آماری رد نگردد. از طرف دیگر، رسم یک نقطه خارج از حدود کنترل بیانگر رد چنین فرضیه است. مشابه آزمون فرضیه، می‌توان احتمال خطای نوع اول (بیانگر حالت خارج از کنترل است هنگامی که فرایند تحت کنترل است) و احتمال خطای نوع دوم (بیانگر حالت تحت کنترل است در حالی که واقعاً فرایند خارج از حدود کنترل به سر می‌برد) برای نمودار کنترل تعریف کرد.

می‌توان یک مدل عمومی برای نمودار کنترل ارائه کرد. فرض کنید مشخصه کیفی مورد نظر به وسیله آماره ω اندازه گیری می شود و همچنین فرض کنید میانگین و انحراف معیار ω به ترتیب برابر با μ_ω و σ_ω است.

بنابراین خط مرکز، حد کنترل بالا و حد کنترل پایین برابر خواهد بود با:

$$\begin{cases} UCL = \mu_\omega + k\,\sigma_\omega \\ CL = \mu_\omega \\ LCL = \mu_\omega + k\,\sigma_\omega \end{cases}$$

در رابطه فوق، k فاصله حدود کنترل از خط مرکز را بر حسب واحد انحراف معیار نشان می‌دهد.

این تئوری عمومی نمودارهای کنترل بار اول توسط دکتر والتر شوهارت ارائه گردید و نمودارهای کنترلی که از چنین قانونی پیروی کنند را نمودارهای کنترل شوهارت می‌نامند

۱-۲-۵ انتخاب حدود کنترل

یکی از تصمیمات مهم در طراحی یک نمودار کنترل تعیین حدود کنترل است. با دور کردن حدود کنترل از خط مرکز احتمال خطای نوع اول و یا به عبارت دیگر ریسک رسم یک نقطه در پشت حدود کنترل در حالی که هیچ‌گونه انحراف با دلیلی وجود ندارد کاهش می‌یابد. از طرف دیگر، افزایش فاصله بین حدود کنترل سبب می‌شود تا احتمال خطای نوع دوم و یا به عبارت دیگر رسم یک نقطه بین حدود کنترل زمانی که فرایند در حالت خارج از کنترل به سر می‌برد افزایش یابد. اگر حدود کنترل به خط مرکز نزدیک‌تر شوند، اثر معکوسی به وجود می‌آید: خطای نوع اول افزایش ولی احتمال خطای نوع دوم کاهش می‌یابد.

معمولاً حدود کنترل، بدون توجه به توزیع مشخصه کیفی، به صورت ضریبی از انحراف معیار آماره‌ای که بر روی نمودار رسم می‌شود محاسبه می‌گردد. ضریبی که در این رابطه استفاده می‌شود معمولاً ۳ است و بنابراین حدود ۳ انحراف معیار را برای نمودارهای کنترل بدون توجه به نوع آن استفاده می‌کنند. در انگلیس و بخشی از کشورهای اروپایی حدود احتمال استاندارد ۰/۰۰۱ کاربرد دارد.

دلیل استفاده از حدود سه انحراف معیار، بازده خوبی است که در عمل از آن‌ها مشاهده شده است. علاوه بر این، در اغلب مواقع توزیع واقعی مشخصه کیفی به خوبی معلوم نیست تا بتوان حدود احتمال دقیقی محاسبه نمود[۱].

۳-۱ محاسبات مربوط به متوسط طول دنباله

در برقراری سیستمهای کنترل مسأله کارایی و حساسیت سیستم جهت کشف تغییرات در میانگین فرایند تولیدی از اهمیت بسزایی برخوردار است. معیاری که معمولاً برای ارزیابی کارآیی این سیستم ها مورد استفاده قرار می‌گیرد، متوسط طول دنباله نام دارد. متوسط طول دنباله عبارت است از متوسط تعداد نمونه ای که باید جهت کشف تغییر ایجاد شده در فرایند، از خط تولید برداشته شود. کشف تغییر عبارت است از قرار گرفتن نقطه ای در خارج از حدود عمل بر روی نمودار کنترل.

بنابراین متوسط طول دنباله ARL، متوسط تعداد نقاطی است که باید روی نمودار کنترل رسم گردند تا یک نقطه خارج از حدود کنترل مشاهده شود. مقدار ARL برای هر نمودار کنترل شوهارت می‌تواند به سادگی از طریق رابطه زیر محاسبه شود:

$$\text{متوسط طول دنباله} = \frac{۱}{P}$$

در رابطه فوق P احتمال یک نقطه خارج از حدود کنترل را نشان می‌دهد.

در صورتی که فرآیند تحت کنترل باشد ARL از رابطه زیر محاسبه می‌شود:

$$ARL = \frac{۱}{\alpha}$$

و اگر فرآیند خارج از کنترل باشد آنگاه:

$$ARL = \frac{۱}{۱ - \beta}$$

برای هر نمودار کنترل با حدود کنترل سه انحراف معیار مقدار P برابر ۰/۰۰۲۷ خواهد بود. بنابراین متوسط طول دنباله برای نمودار کنترل زمانی که فرآیند تحت کنترل است برابر می‌شود با :

$$\text{متوسط طول دنباله} = \frac{۱}{P} = \frac{۱}{۰/۰۰۲۷} = ۳۷۰/۳۷$$

به عبارت دیگر، حتی اگر فرآیند تحت کنترل باشد به طور متوسط یک هشدار تقریباً بعد از هر ۳۷۰ نمونه مشاهده خواهد شد[۱] .

۱-٤ رگرسیون

۱-٤-۱ مقدمه

بسیاری از پیشامدها و پدیده هایی که در جهان اتفاق می‌افتند، تابع برخی از پیشامدهای دیگر هستند. یکی از هدف های مهم در تحقیقات علمی، کشف وجود یا عدم وجود رابطه ی بین پدیده هاست. یکی از روش های آماری که بطور گسترده برای این منظور به کار می‌رود، روش تحلیل رگرسیون است.

۱-٤-۲ تاریخچه

واژه رگرسیون در لغت به معنی بازگشت است و اغلب جهت رساندن مفهوم «بازگشت به یک مقدار متوسط یا میانگین » به کار می‌رود. بدین معنی که برخی پدیده ها به مرور زمان از نظر کمّی به طرف یک مقدار متوسط میل می کنند.

در سال ۱۸۸۵ فرانسیس گالتون[۱] در مقاله ای که منتشر کرد اظهار داشت که متوسط قد پسران دارای پدران قد بلند، کمتر از قد پدرانشان می‌باشد. به نحو مشابه متوسط قد پسران دارای پدران کوتاه قد نیز بیشتر از قد پدرانشان گزارش شده است. به این ترتیب گالتون پدیده بازگشت به میانگین را در داده هایش مورد تأکید قرار داد.

برای گالتون رگرسیون مفهومی زیست شناختی داشت اما کارهای او توسط کارل پیرسون[۲] برای مفاهیم آماری توسعه داده شد. گرچه گالتون برای تأکید بر پدیده ی « بازگشت به سمت مقدار متوسط» از تحلیل رگرسیون استفاده کرد، اما امروزه واژه ی تحلیل رگرسیون جهت اشاره به مطالعات مربوط به روابط بین متغیرها به کار برده می‌شود[۲].

۱-٤-۳ رگرسیون خطی ساده

در حقیقت تحلیل رگرسیون فن و تکنیکی آماری برای بررسی و مدل سازی ارتباط بین متغیرهاست. رگرسیون تقریباً در هر زمینه ای از جمله مهندسی، مدیریت، بیولوژی برای برآورد و پیش بینی مورد نیاز است؛ وتحلیل رگرسیون پرکاربردترین تکنیک آماری است.

در رگرسیون خطی ساده ارتباط خطی به صورت زیر نشان داده می‌شود:

$$y = a + bx$$

که در آن a عرض از مبدأ و b شیب خط است.

شرایط محیط، تفاوتهای طبیعی و ..... سبب به وجود آمدن خطا در اندازه گیری می‌شود، بنابراین مقدار اولیه را به صورت زیر اصلاح می کنیم:

---

۱ Ferancies Galton

۲ K*arl* Pearson