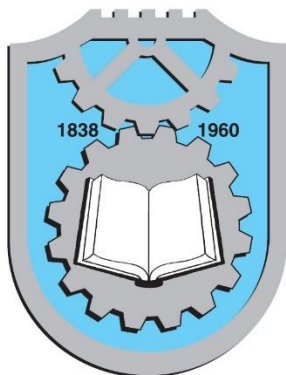


Факултет инжењерских наука Универзитета у Крагујевцу



Назив студијског програма: Аутомобилско инжењерство

Ниво студија: Основне академске студије

Предмет: Основи метрологије и контроле квалитета

Број индекса: 809/2013

Милош Р. Величковић

Статистичка контрола квалитета

Завршни рад

Комисија за преглед и одбрану:

1. доцент др Фатима Живић, дипл. инж. – ментор

2. _____

Датум одбране: _____

3. _____

Оцена: _____

У оквиру овог завршног рада, кандидат треба да:

- објасни методе статистичког управљања квалитетом,
- објасни процес статистичке анализе квалитета кроз контролне карте.

Препоручена литература:

[1] Crosby P. (1995). *Philip Crosby's Reflections on Quality*. McGraw-Hill

[2] Ishikawa K. (1985) *What is total quality control? The Japanese way*.

[3] Nelson, Lloyd, (1984) "The Shewhart Control Chart—Tests for Special Causes," *Technical Aids, Journal of Quality Technology*, No. 16, Oct., pp. 238-239.

[4] Myers & Mullet, (2011) Classification and competition analysis of air cargo logistics providers: *The case of Taiwan's high-technology industry*, *Journal of Air Transport Management*.

Крагујевац, август 2018.

Ментор:

доцент др Фатима Живић, дипл. инж.

Резиме

У овом раду представљен је основни приступ статистичких модела који се користе у сврху контроле и анализе квалитета. На првом месту то су контролне карте које пружају увид у комплетан процес који се одвија у производњи. Оне указују на то да ли је процес под контролом и да ли су аспекти квалитета задовољени.

Кључне речи: Статистичка анализа квалитета, Контролне карте, Нелсонова правила, Регресивна и Корелациона анализа.

Abstract

This paper presents the basic approach of statistical models used for the purpose of quality control and analysis. In the first place, these are control cards that provide insight into the whole process that takes place in production. They indicate whether the process is under control and whether the aspects of quality are satisfied.

Key words: Statistical Quality Analysis, Control Cards, Nelson Rules, Regressive and Correlation analysis.

Садржај

1. Увод	1
2. Појава статистичких метода у контроли квалитета	3
3. Нормална расподела	4
4. Значај контролних карата у статистичкој анализи квалитета.....	7
4.1. XR контролне карте (Xbar R-range)	10
4.1.1. Израчунавање контролних граница за XRкарту	12
4.2. Контролне карте $X\sigma$	16
4.3. Контролне X карте	18
4.4. Интерпретација контролних карата.....	19
4.4.1. Нелсонова правила.....	21
5. Регресиона и корелациона анализа.....	26
5.1. Регресиона анализа	26
5.2. Корелациона анализа	30
6. Вишеструка регресија и корелација.....	36
6.1. Мултиколинearност	37
6.2. Индекс детерминације	38
6.3. Значај вишеструке регресије	39
6.4. Тестирање статистичке значајности.....	39
7. Закључак	41
8. Литература.....	42

1. Увод

Са развојем индустријализације, односно масовне производње, појавила се потреба за контролом и ефективним управљањем квалитетом, који може бити дефинисан на различите начине.

Према међународном стандарду ISO 9000, квалитет је дефинисан као ниво до којег скуп одређених карактеристика (производа или услуге) испуњава захтеве. Са друге стране, гуруи квалитета, односно, најзначајнији аутори из ове области, који су у великој мери утицали на савремене методе управљања квалитетом, придодавали су значај посебним атрибутима које карактерише сам квалитет.

Тако нпр. према Кросбију квалитет је просто удовољавање захтевима потрошача (*Crosby 1995*), док је према Фајгенбауму квалитет оно што купац или корисник каже да жели. Колико год то звучало поједностављено, Фајгенбаум полази од тезе да се на турбулентним и високо конкурентним тржиштима промене у захтевима тржишта јако брзо мењају, у складу са захтевима крајњих корисника, па се у складу са тим, мора извршити конкретан приступ купцима и на тај начин установити шта конкретно они очекују од производа или услуге.

У односу на то, Кауори Ишикава под квалитетом подразумева производњу без грешке, односно, постизање жељеног квалитета већ у првој фази производње, па је у складу са тим развио низ корака који кроз учење и праксу доводе до тога. (*Ishikawa 1985*)

Постоји више разлога за овакав приступ проблему квалитета односно појаве некавалитета. Један од основних разлога је појава трошкова, везаних за квалитет који, у великој мери, могу смањити профит предузећа, или чак, довести до појаве додатних трошкова, уколико се квалитетом не управља на правилан начин. Трошкови квалитета су најчешће скривени у склопу других трошкова и могу се јавити на различитим трошковним местима, па их је због тога тешко уочити. Из тог разлога су развијене различите методе управљања квалитетом као и трошковима везаних за квалитет.

Све методе које данас срећемо теже ка успостављању уједначеног квалитета који се може мерити, исто као и трошкови који га прате. Једна од тих метода је и статистичка анализа квалитета којом се могу мерити одступања која се јављају у току производње, о чему ће у даљем тексту бити више речи.

Из свега наведеног, може се констатовати да квалитет подразумева следеће карактеристике:

- способност за употребу,
- способност за примену,
- задовољство корисника,
- усклађеност са захтевима.

На основу поменутих карактеристика, формиран су различити стандарди када се ради о квалитету као што су:

- **ISO**(*International Organization for Standardization* - Интернационална организација за стандардизацију),
- **CEN**(*European Committee for Standardization*-Европски комитет за стандардизацију),
- **HACCP**(*Hazard Analysis and Critical Control Point*-Анализа ризика и критичне контролне тачке, користи се у прехранбеној индустрији), итд.

2. Појава статистичких метода у контроли квалитета

Статистика, као наука, и њене методе нашле су велику примену у условима масовне производње. Први покушаји примене статистичке методе у контроли квалитета датирају почетком двадесетог века. Волтер Шухарт (*Walter A. Shewhart*) из *Bell Telephone Laboratories* је дао први нацрт контролне карте, на основу које су касније Харолд Доџ (*Harold F. Dodge*) и Хари Роминг (*Harry G. Roming*), отпочели обимне радове на разради метода узимања узорка у току серијске производње. (*Мајсторовић 2012*)

Током педесетих година прошлог века, у Јапану почиње озбиљнији приступ контроли квалитета оснивањем одбора за контролу квалитета у склопу Удружења Јапанских научника и инжењера. Отприлике у исто време јапанско друштво за стандардизацију организује семинаре о контроли квалитета са становишта статистичких метода, када се на њихов позив одазива Вилијем Деминг, амерички стручњак из области статистичког унапређења квалитета.

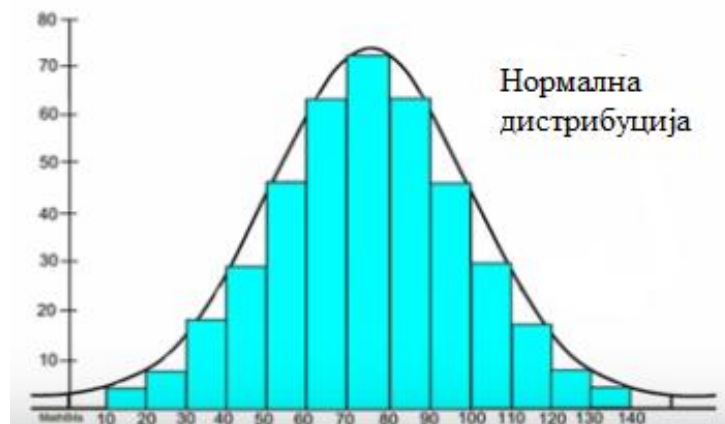
Статистичке методе управљања квалитетом нису фокусиране на појединачне факторе који утичу на квалитет појединог производа. На овакав начин врши се праћење укупног стања и кретања производа у целини, па се, сходно томе, може благовремено реаговати на посебне појаве које се могу уочити у анализи процеса и на тај начин спровести одређене корективне мере. (*Мајсторовић 2012*)

3. Нормална расподела

Приликом праћења неке од карактеристика квалитета, као статистичке величине, добијају се вредности које имају своју дистрибуцију односно, расподелу. Резултати, добијени на тај начин, у пракси најчешће гравитирају око средње вредности, а онда се са све мање података удаљавају од средине. Расподела учесталости података указују на Гасусову расподелу, која у великом броју случајева представља теоријску основу статистичке анализе квалитета. (Ивковић, 1992)

На пример:

У току серијске производње неминовно долази до различитих варијација у квалитету производа, које се могу мерити према различитим критеријумима. Ако су то, на пример, машински елементи који се израђују на стругу, пре или касније, ће доћи до варијација у димензијама израђених елемената у оквиру задатих толеранција. У овом случају, то су пречници елемената бројчано изражени у милиметрима и они варирају око неке средње вредности. Ако се ове мерене вредности представе као хистограм уочиће се расподела, односно нормална дистрибуција у току процеса производње, као што је то приказано на слици 3.1.



Слика 3.1. Нормална дистрибуција

Случајна променљива је дистрибуирана по закону нормалне или Гаусове расподеле ако је подручје њених вредности од $-\infty$ до $+\infty$, а функција вероватноће је представљена изразом (Поповић, 2008):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (3.1)$$

Из овога следи да је нормална расподела одређена параметрима аритметичке средине и стандардног одступања.

У стандардизованом облику, то је:

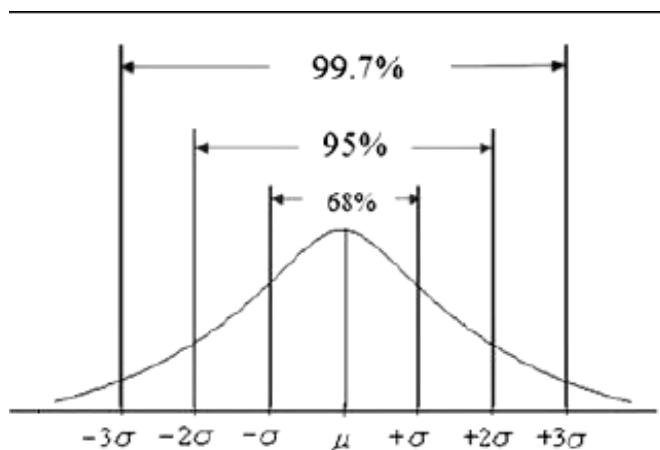
$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad (3.2)$$

Стандардизовањем променљиве (z), врши се стандардизовање сваке реализоване вредности, па се за сваку вредност (x), добија мера њеног одступања од стандардне вредности (μ) у односу на девијацију популације σ .

То конкретно значи да ће вредост $z=3$ променљиве x бити удаљена од средње вредности **3 σ** .

Како дата релација нема примитивну функцију вредност се може узети из таблица стандардне нормалне расподеле. (Букадиновић, 2008)

На тај начин долази се до појаве вредности x у одређеним интервалима, у односу на средњу вредност као што је приказано на слици 3.2.



Слика 3.2. Расподела вероватноће

У контексту статистичке анализе квалитета, мерењем узорака у различитим временским интервалима, према задатим атрибутима, може се уочити да 99.7% узорака варира у овом опсегу $\pm 3\sigma$. (Nelson, 2006)

Уколико подаци следе нормалну расподелу, способност процеса се може дефинисати као расипање и мери се толеранцијом $T=6\sigma$, при чему расипање садржи 99,73% производа. (Лазих, 2011)

Другим речима, расипање изван интервала ширине 6σ заузима јако мали број података, што говори о јако малој вероватноћи да се у том опсегу уопште и појаве подаци.

Из тог разлога, у пракси се границе од $\pm 3\sigma$ узимају као границе вероватноће којом се оцењује нека појава.

Са гледишта статистичке анализе квалитета и разматрањима битним за праћење процеса, од посебне важности су карактеристике средње вредности и карактеристике расипања.

Ако су толеранције квалитета неког производа, чији се процес прати, у границама $\pm 2\sigma$, онда ће свако одступање од тих граница указати на проблеме везане за квалитет производа и то у току самог процеса. Другим речима, указаће на неконтролисаност процеса што се може уочити из контролних карата.

4. Значај контролних карата у статистичкој анализи квалитета

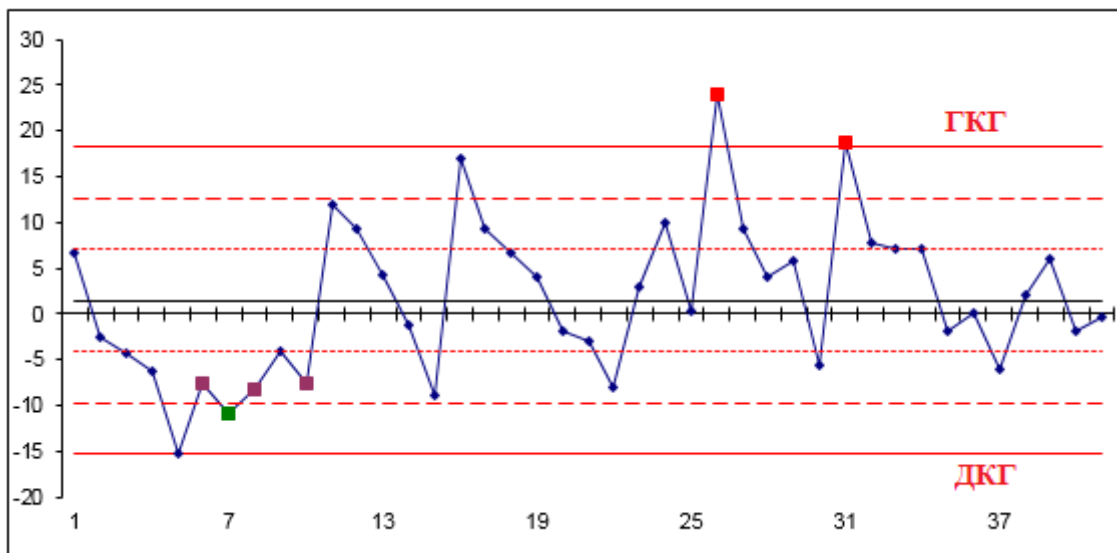
У току процеса реализације неког производа, појављује се велики број фактора који утичу на његов квалитет. Како би се ти фактори открили, и како би се на њих утицало, неопходно је да се процес стално прати. Од сваког процеса, неодвојиви су недостаци који се неминовно јављају и за последицу имају дораде, поправке, губитке, као и губитке у времену потребном за израду неког производа.

Све наведено доводи до нежељених повећања трошкова производње. Усаглашавањем процеса производње, тако да се постиже уједначен квалитет производа, смањиће се и време које је потребно за уклањање недостатака као и трошкови који их прате. *(Рајковић 2012)*

За сваки процес потребно је установити толеранције. Односно, границе у којима ће се одвијати процес контролисати, при чему ће се пратити карактеристике квалитета. Код новог производа, ове границе је потребно установити у само почетку производње. Често се дешава да и ако је квалитет неког производа јасно дефинисан толеранцијским пољима у складу са неком операцијом (као што је то операција на стругу), процес производње некада није у стању да задовољи циљеве квалитета у овом случају димензије. До тога долази првенствено због различитих промена у самом процесу које могу бити значајног или случајног карактера, као што су: дотрајали алат, промене режима обраде или чак измене које су настале у току одржавања машине. *(Бедовић, 2009.)*

Из тог разлога потребно је пратити карактеристике квалитета помоћу јединствене контролне карте. Оне указују на варијације просека, распона и стандардне девијације за неки контролни узорак, а графички се представља као дијаграм. На тај начин се може визуелно уочити да ли је неки процес изван контроле у одређеном временском интервалу, или је дошло до побољшања процеса. Употреба оваквих карата у праћењу процеса назива се статистичка контрола процеса (*SPC - Statistical Process Control*).

На слици 4.1. приказана је једна оваква карта.



Слика 4.1. Контролна карта

Као што је приказано на слици 4.1, једноставним графиком је представљена контролна карта, коју чине узорци узимани у различитим временским интервалима. Горња и доња црвена линија означава контролне границе у распону од $\pm 3\sigma$. У конкретном случају може се уочити да два узорка прелазе горњу контролну границу (ГКГ), према томе може се закључити да је дошло до грешке у процесу, па процес захтева регулисање.

Метода коришћења контролних карата у сврху анализе квалитета представља статистичко регулисање процеса, односно, статистичку контролу активности, које су у току ради интервенисања у случајевима када процес изађе ван прописаних контролних граница. (Carter, 2012)

За неки процес, који је под контролом, може се рећи да је варирање квалитета у границама интервала поверења за задати интервал значаја и да је процес стабилан. У случају нестабилности процеса појавиће се велике варијације у квалитету производа. Задатак лица која су задужена за управљање процесима квалитета је да:

- одрже процес у контролисаном стању (Carter, 2012),
- доведу процес у контролисано стање, ако се уочи да је процес неконтролисан,
- утврде да ли је успостављена контрола процеса.

Идеја контролних карата је веома једноставна, а заснива се на конструкцији критичне области теста значајности који полази од хипотезе о математичком очекивању. Посматрано обележје X које се посматра код сваког производа треба да има вредност X_0 (посматрано статистички).

Због различитих случајева у којима долази до флуктација у производном процесу X представља случајну променљиву, за коју се по централној теореме може сматрати да има нормалну расподелу.

У том случају вредност X_0 је остварена уколико је математичко очекивање променљиве X једнако нули.

Приликом проверавања процеса и његовог одвијања у задатим границама узима се периодично n узорака, након чега се мери њихово обележје X , па се на тај начин добија средња вредност.

Када се усвоји праг значајности α , потребно је тестирати хипотезу $H_0(X=X_0)$. Другим речима, проверава се да ли x припада критичној области. Уколико припада, то указује на то да је процес одступио од дозвољених граница, па је потребно уклонити узрок одступања, при чему се полази од варијансе среднине дате релацијом:

$$\frac{\sigma^2}{x} = \frac{\sigma^2}{n} \quad (4.1)$$

и стандардне грешке средине исказане релацијом:

$$\sigma \bar{x} = \frac{\sigma}{\sqrt{n}} \quad (4.2)$$

На овај начин приказана је у потпуности упрошћена процедура коришћења контролних карата у које се уносе средње вредности x (на слици 4.1, $X=0$) и подаци узетих узорака n и на којима је критична област одвојена паралелним правима које означавају горњу и доњу контролну границу.

Интервал оцене средње вредности исказан је релацијама:

- за доњу контролну границу:

$$\bar{x} = x - Z_{1-\alpha} * \frac{\sigma}{\sqrt{n}} \quad (4.3)$$

- за горњу контролну границу:

$$\bar{x} = x + Z_{1-\alpha} * \frac{\sigma}{\sqrt{n}} \quad (4.4)$$

Контролне карте пружају објективну и визуелно прегледну информацију о стању квалитета у току процеса кога прате лица задужена за контролу квалитета.

Контролне карте такође карактерише низ особина које поједностављују контролу процеса и саму статистичку анализу квалитета. За њих се још може рећи (Carter, 2012):

- да представљају прикладан начин за уношење података од стране контроле,
- да омогућавају јасан преглед стања и кретања процеса,
- да су врло осетљиве на промене у процесу, па се могу употребљавати и за евентуална предвиђања одређених појава у току одвијања процеса,
- да раздвајају случајеве значајних и случајних узрока промена у процесима, и као такве представљају континуирани тест хипотезе статистичке стабилности процеса,
- да се могу користити као средство помоћу кога ће се спречити прекомерно регулисан процес, када радник новим подешавањима реагује на случајна одступања, што доводи до мењања добро подешеног процеса,
- да су контролне карте погодно средство као начин комуникације везаних за конкретне проблеме унутар процеса између радника и контролора, као и између осталих сектора у организацији.

Контролне границе могу се израчунати на три начина:

- на основу прикупљених података технолошког процеса,
- на основу задатих толеранција праћеног квалитета,
- помоћу технолошког система **6σ**.

У пракси се за праћење квалитета помоћу мерних величина и атрибутивних оцењивања употребљавају различите контролне карте, док се за праћење средњих вредности најчешће употребљавају $\bar{X}R$ контролне карте, а за појединачне случајеве X карте.

4.1. $\bar{X}R$ контролне карте (Xbar R-range)

Контролне карте, помоћу којих се прате аритметичке средине $\bar{X}$ и распона R узорака, састављају се тако да користе предност оцена на основу математичке статистике.

Овакве карте се често користе у процесној индустрији, а нарочито су погодне у индустријама, у којима се производи на одређеној операцији појединачно посматрају.

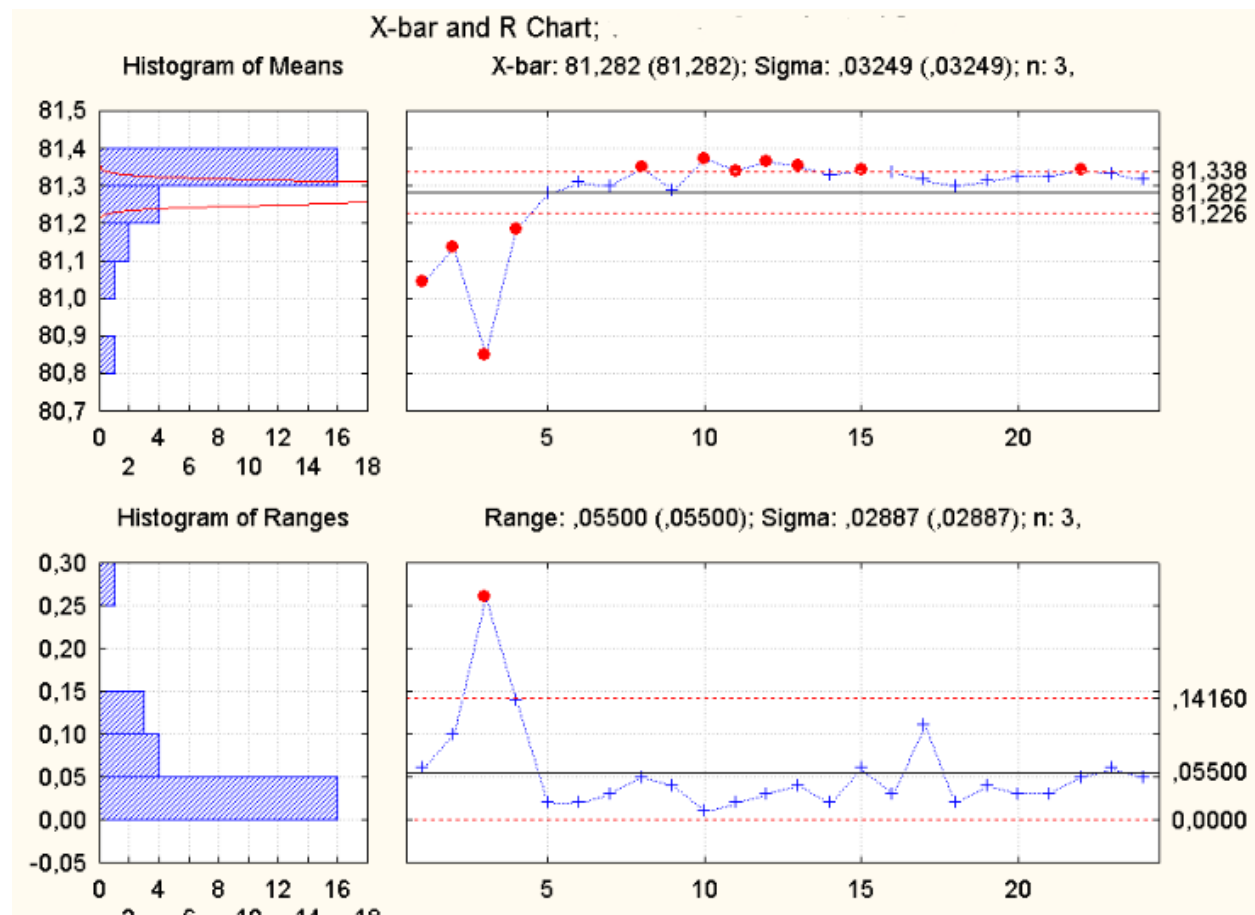
Контролисањем, односно, мерењем врши се дефинисање величина узорака као и учесталост контролисања и на тај начин се мери карактеристика квалитета, чије се варијације посматрају.

Том приликом се за свако посматрање израчунава просек и распон **R**. Другим речима, анализа варијација заснива се на понашању три основне компоненте: **тенденције, расипања и облика**. (Дреновац, А. 2015)

X – карта даје податке о тенденцији процеса зависно од временског интервала,

R – карта даје податке о расипању процеса

Заједничка $\bar{X}R$ карта приказује уз наведене податке и податке о својствима и абнормалностима облика варијације процеса који се контролише. На слици 4.2 приказан је пример једне овакве карте.



Слика 4.2. Заједничка $\bar{X}R$ карта

Израда $\bar{X}R$ контролне карте треба да прати следеће кораке:

- из процеса који се прати, узимају се мали узорци у различитим временским размацима што зависи од стабилности процеса,
- на узорку се мери карактеристика квалитета, при чему се добија n измерених вредности,
- након тога, за измерене вредности израчунавају се аритметичка средина и распон,
- израчунате вредности уписују се у $\bar{X}R$ карту,
- за аритметичке средине и распоне рачунају се контролне границе и уцртавају у контролну карту.

У табели 4.1 приказане су основне поставке процеса узроковања.

	Посматрање				
	1	2	3	...	k
Посматране вредности	X11	X12	X13	...	X1k
	.	.	.	...	.
	Xn1	Xn2	Xn3	...	Xnk
Средње вредности	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	...	$\bar{x}_k$
Распон	R1	R2	R3	...	Rk

Табела 4.1. Основне поставке процеса узроковања

4.1.1. Израчунавање контролних граница за $\bar{X}R$ карту

Контролне границе за $\bar{X}R$ карту добијају се на више начина и то, на основу прикупљених података технолошког процеса, задате толеранције карактеристике квалитета и на основу познавања технолошког процеса.

Уколико се разматрају подаци снимљени у току технолошког процеса, тада се за критичне границе задаје праг значаја $\alpha=0,0027$, што значи да ће се само 27 од 10000 примерака наћи ван задатих граница.

У том случају може се рећи да се 99,73% вредности налази у интервалу $\pm 3\sigma$ од средње вредности. За упозоравајуће границе најчешће се у пракси узима $\alpha=0,005$. Према томе, за $1-\alpha=0,95$ следи да је $\alpha=0,05$, $\alpha/2=0.025$ а $z_{\alpha/2}=z_{0,25}=1,96$.

Према томе контролне границе ће бити:

$$\text{ДКГ: } \bar{x} = x_0 - 3 * S_{\bar{x}} \quad (4.5)$$

$$\text{ГКГ: } \bar{x} = x_0 + 3 * S_{\bar{x}} \quad (4.6)$$

док ће упозоравајуће,

$$\text{ДУГ: } \bar{x} = x_0 - 1,96 * S_{\bar{x}} \quad (4.7)$$

$$\text{ГУГ: } \bar{x} = x_0 + 1,96 * S_{\bar{x}} \quad (4.8)$$

Где је $S_{\bar{x}}$ стандардна девијација која је једнака корену варијансе узрока $S_{\bar{x}} = \sqrt{S^2}$.

Из овога се може закључити да контролне границе зависе од x_0 и σ , као и од нивоа α и величине n . У пракси често параметри x_0 и σ^2 нису познати, па се непознате замењују њиховим оценама добијеним на нивоу узорка.

Како би се ова контролна карта могла подвргнути квалитетнијој статистичкој анализи потребно је да узорак, и ако иницијално мали, буде 5–10, док број посматрања минимум између 20–30.

За оцену x_0 узима се вредност:

$$x_0 = \bar{\bar{x}} = \frac{1}{k} \sum_{j=1}^k \bar{x}_j \quad (4.9)$$

а за оцену варијансе:

$$s^2 = \frac{1}{k} \sum_{j=1}^k s_j^2 \quad (4.10)$$

где је:

$$s_j = \frac{1}{n-1} \sum_{i=1}^n (x_{ij}^2 - \bar{x}_j^2) \quad (4.11)$$

Рачунање контролних граница се може додатно поједноставити помоћу табела фактора којима се израчунавају контролне границе.

n	A	A ₁	A ₂	B ₃	B ₄	D ₁	D ₂	D ₃	D ₄	d ₂
2	2,121	3,760	1,880	0,000	3,267	0,000	3,687	0,000	3,269	1,128
3	1,732	2,394	1,023	0,000	2,569	0,000	4,357	0,000	2,574	1,693
4	1,500	1,880	0,729	0,000	2,267	0,000	4,699	0,000	2,282	2,059
5	1,342	1,596	0,577	0,000	2,090	0,000	4,918	0,000	2,114	2,326
6	1,225	1,410	0,483	0,030	1,970	0,000	5,078	0,000	2,004	2,534
7	1,134	1,277	0,419	0,117	1,883	0,205	5,203	0,076	1,924	2,704

Табела 4.2. Табела фактора за израчунавање контролне границе

У табели 4.2 су приказани параметри исказани као пондери у израчунавању горњих и доњих контролних граница за различите контролне карте. Помоћу података датих у овој табели може се једноставно доћи до контролних граница.

Средњу вредност представља $\bar{x}_0$, која је добијена помоћу једначина (4.5) и (4.6).

Затим се за сваки узорак одређује распон **R**, израчунава аритметичка средина распона $\bar{R}$ и стандардна девијација узорка $S_{\bar{x}}$, односно:

$$\bar{R} = \frac{1}{k} \sum_{j=1}^k R_j \quad (4.12)$$

$$S_{\bar{x}} = \frac{R}{d_2 \cdot \sqrt{n}} \quad (4.13)$$

Контролне границе представљају кретање процеса као и његову способност и стабилност. Оне такође представљају могућности једног процеса као и варијације које се могу очекивати унутар њега. За вредности којима се располаже, добијене мерењем узорака, за израчунавање контролних граница користе се коефицијенти **A2**, **D3** и **D4**, који зависе само од величине узорка **n**. (Бакија, 1978)

$$ДКГ_{\bar{x}} = x_0 - A_2 \cdot \bar{R} \quad (4.14)$$

$$ГКГ_{\bar{x}} = x_0 + A_2 \cdot \bar{R}$$

$$ДКГ_{\bar{R}} = D_3 \cdot \bar{R} \quad (4.15)$$

$$ГКГ_{\bar{R}} = D_4 \cdot \bar{R}$$

Код израчунавања контролних граница $\bar{X}\bar{R}$ карте на основу задатих толеранција карактеристика квалитета, посматрана појава понаша се по закону нормалне расподеле. Како би се израчунале контролне границе потребно је пре свега одредити задату толеранцију T и меру средњег толерацијског поља x_0 , под претпоставком да је $T=6\sigma$.

Контролне границе средњих вредности у овом случају израчунавају се према релацији:

$$ДКГ_{\bar{x}} = x_0 - A \cdot T \quad (4.16)$$

$$ГКГ_{\bar{x}} = x_0 + A \cdot T$$

и распона:

$$ДКГ_R = D_1 \cdot T \quad (4.17)$$

$$ГКГ_R = D_2 \cdot T$$

Израчунавање граница на основу толеранција врши се у случајевима када из неког разлога није могуће снимити цео процес.

За разлику од израчунавања контролних граница на основу толеранција, израчунавање граница на основу снимљеног технолошког процеса захтева много више времена и подробнију анализу како снимљени подаци не би дали погрешну слику о стварном стању. Вредност контролних граница задржава се током праћења процеса и има знатне предности у односу на границе израчунате помоћу задатих толеранција, јер се узима у обзир способност технолошког процеса.

Контролне границе се израчунавају као:

$$ДКГ_{\bar{x}} = x_0 - A \cdot \sigma_0 \quad (4.18)$$

$$ГКГ_{\bar{x}} = x_0 + A \cdot \sigma_0$$

а за распон као:

$$ДКГ_R = D_1 \cdot \sigma_0 \quad (4.19)$$

$$ГКГ_R = D_2 \cdot \sigma_0$$

4.2. Контролне карте $\bar{X}\sigma$

Код праћења сложенијих технолошких процеса са већим бројем комада у узорку користе се $\bar{X}\sigma$ контролне карте. За њих је карактеристично праћење технолошког процеса преко аритметичке средине и стандардне девијације величина узорка. Користи се најчешће у процесу производње када је трајање операције дуге и код кога постоји веће расипање, односно, када су потребни поузданији подаци, при чему је већи узорак неопходан за детаљан увид у резултате.

У овом случају стандардна девијација се уместо распона појављује као мера расипања. У пракси се показало да се распон са довољном тачношћу може употребљавати уместо стандардне девијације.

Контролне границе се израчунавају за иста три случаја као и код претходно поменуте контролне карте односно: на основу снимљених података процеса, на основу задате толеранције и на основу способности технолошког процеса.

За израчунавање контролних граница на основу снимљених података процеса, пре свега је потребно одредити централну линију стандардне девијације и централну линију средњих вредности узорка. Линија средњих вредности добија се на исти начин као и код предходне контролне карте, док се линија стандардне девијације рачуна према релацији:

$$\bar{\sigma} = \frac{\sum_{i=1}^k \sigma_i}{k} \quad (4.20)$$

На основу добијених података, контролне границе стандардне девијације израчунавају се као:

$$ДКГ_{\sigma} = B_1 \cdot \bar{\sigma}$$

(4.21)

$$ГКГ_{\sigma} = B_2 \cdot \bar{\sigma}$$

а границе средње вредности као:

$$ДКГ_{\bar{x}} = \bar{x} - A_1 \cdot \sigma_0$$

(4.22)

$$ГКГ_{\bar{x}} = \bar{x} + A_1 \cdot \sigma_0$$

Ако је толеранција предходно задата контролне границе стандардне девијације израчунавају се као:

$$ДКГ_{\sigma} = B_1 \cdot T$$

(4.23)

$$ГКГ_{\sigma} = B_2 \cdot T$$

а контролне границе средњих вредности као:

$$ДКГ_{\bar{x}} = \bar{x} - A \cdot T$$

(4.24)

$$ГКГ_{\bar{x}} = \bar{x} + A \cdot T$$

Када су познате карактеристике технолошког процеса, контролне границе стандардне девијације израчунавају се као:

$$ДКГ_{\sigma} = B_1 \cdot \sigma_0$$

(4.25)

$$ГКГ_{\sigma} = B_2 \cdot \sigma_0$$

а контролне као:

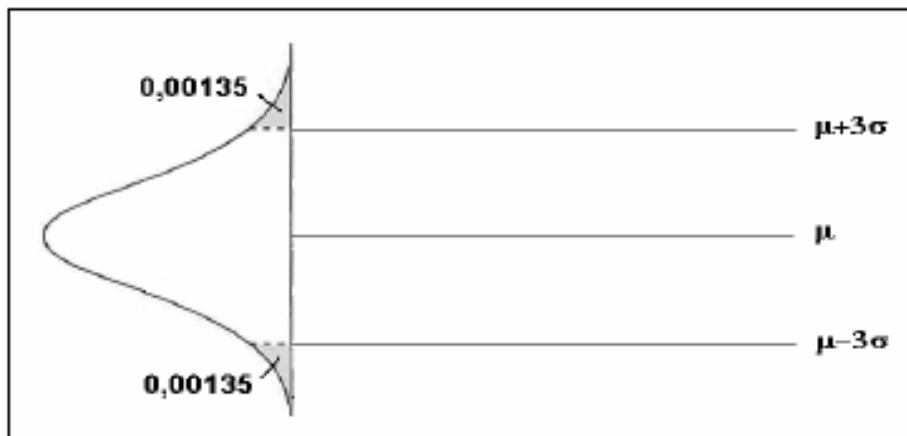
$$\text{ДКГ}_{\bar{x}} = \bar{x} - A \cdot \sigma_0 \quad (4.26)$$

$$\text{ГКГ}_{\bar{x}} = \bar{x} + A \cdot \sigma_0$$

4.3. Контролне X карте

Контролна X карта је најједноставнији облик контролне карте и користи се за праћење појединачних случајева и представљена је као график временског низа појединачних мерења карактеристике и резултата процеса. Као и код осталих контролних карата на њој је уцртана централна линија и контролне границе. На тај начин X карта испуњава основне захтеве постављене контролном картом, јер омогућава брз увид у стање и довољно детаљан приказ. Помоћу ње се лако може установити где се налазе мерене вредности у односу на толеранцијско поље, а такође пружа и увид у распон измерених величина.

Централном линијом је код ове карте представљена просечна вредност карактеристике квалитета када је процес у фази статистичке контроле. Контролне границе одређују се тако да, када је процес под контролом, постоје веома мала вероватноћа да појединачне вредности буду ван контролних граница. У пракси се обично оне постављају у опсегу од три стандардне девијације од централне линије. Када је процес под контролом, вероватноћа да појединачна мерења буду ван контролних граница је $\alpha=0,0027$ одн. $\alpha/2=0,00135$. На слици 4.3 приказане су вероватноће мерења посматраних процеса. (Merkle, М. Дреновац. А. 2015)



Слика 4.3. Вероватноће мерења посматраног процеса (Дреновац. А.2015).

Уколико је процес, који се разматра, унутар контролних граница, сматра се да је под контролом. Ако се нека од вредности нађе ван граница, то указује на неку ретку појаву у процесу, или да је процес ван контроле и у том случају треба предузети одговарајуће корективне мере.

У овом случају контролне границе функционишу као критичне вредности, па се према томе за сваку тачку, унету у график, врши двострани тест хипотезе. Заправо оне су функције природне варијабилности процеса, а уколико се користе 3σ границе, оне представљају функцију стандардне девијације процеса као што је приказано на слици 4.4.



Слика 4.4. Приказ контролне $\bar{X}$ карте (Дреновац, А. 2015)

Осим контролних граница, процес може бити одређен спецификационим границама које су независне од процеса. Овакве границе, релевантне за посматрање процеса, одређују руководиоци, технолози или купци што додатно олакшава дефинисање граница.

4.4. Интерпретација контролних карата

Праћење процеса у току статистичке анализе квалитета узимањем контролних узорака не указује само на то да ли се неки процес одвија под контролом или не, односно, да ли је квалитет у одговарајућим границама, већ указује и на потенцијалне проблеме који се касније могу јавити, па према томе, контролне карте, посматрано у ширем смислу, дају својеврсну прогнозу о томе како ће се процес даље одвијати. Управо у томе је и највећи значај анализе квалитета статистичким методама јер оне указују на вероватноћу грешке пре него што до ње дође.

У пракси се често дешава да се на контролној карти уочава да је један од узорака изашао изван области контролних граница, а да је остатак процес под контролом. Ово се првенствено дешава код $\bar{X}$ карти и у случајевима када се узима појединачни узорак.

И ако је вероватноћа за овакав случај јако мала, они се ипак дешавају, или као последица грешке при мерењу, или је дошло до промена у процесу које су случајног карактера. Узрок оваквих појава је најчешће људски фактор, или неки други случајни узрок што се може установити применом методе коренске анализе узрока (*Root cause analysis-RCA*), односно, помоћу алата као што је Ишикава дијаграм познат због свог карактеристичног игледа још и као „дијаграм рибља кост“.

Пример конструкције оваквог дијаграма приказан је на слици 4.5.

У приказаном примеру анализиран је процес серијске производње, чије димензије производа варирају у оквиру задатих толераниција. За узорке, који су се нашли изван контролних лимита, примењен је дијаграм рибље кости.

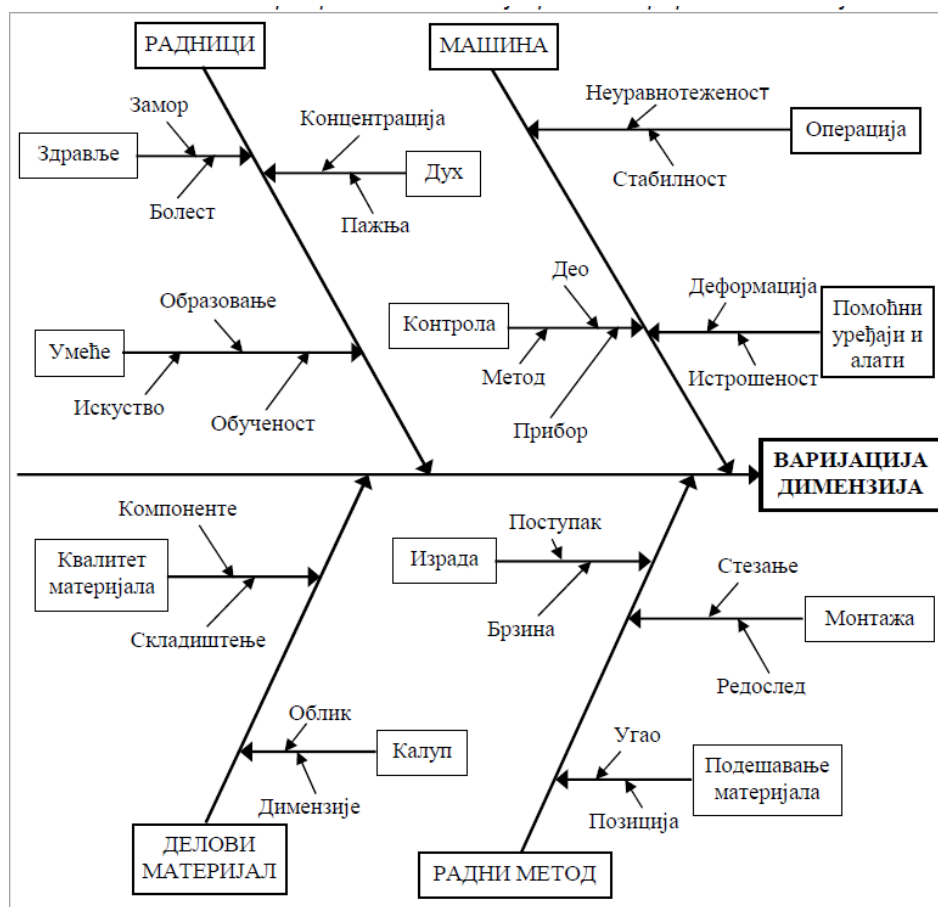
Први корак у овом процесу је дефинисање последице (у овом случају варијације у димензијама), након чега је последица рашчлањена на четири основна фактора (*Ishikawa K.1968*):

- *Machinery* – опрема,
- *Manpower*- радна снага,
- *Method*–метод,
- *Material* – материјал.

Ова четири фактора се обично узимају када се у процесу примете одступања случајног карактера, али се овим факторима у неким случајевима могу придодати још два, и то су (*Ishikawa K.1968*):

- *Maintenance* – одржавање,
- *Environment* – околина.

На тај начин се графички приказује однос између узрока и последице. Дијаграм се може додатно развијати рашчлањивањем на поткатегорије, све док се не пронађу узроци који су довели до настанка проблема.



Слика 4.5. Ишикава дијаграм за варијације у квалитету димензија

Овако конструисан дијаграм се може даље развијати, па тако постаје веома ефикасан алат којим се брзо и једноставно могу открити узроци који доводе до потенцијалних проблема чиме се већ у раној фази снимања процеса могу уочити фактори који доводе до нестабилности процеса.

4.4.1. Нелсонова правила

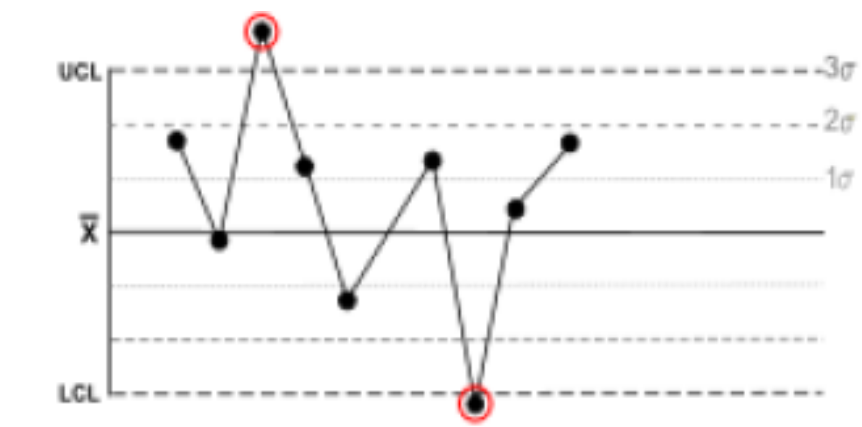
Ако се узме у обзир да контролне карте могу да садрже велики број података и да је њихова обрада сложен процес, па је у неким случајевима тешко одредити да ли је неки процес изван контроле, поготово, ако нема очигледних назнака да се неки од узорака нашао изван контролних граница, поставља се питање, које су то смернице које треба узети у обзир приликом анализе контролних карата и на какве појаве оне указују.

Током педесетих година прошлог века Валтер Шухарт је установио три стандардна правила која указују на нестабилност процеса.

Његова правила су прикупљена у низу тестова за значајне узроке од стране Лојда Нелсона (*Lloyd Nelson*), који је дошао до сета од осам правила која указују на нестабилност процеса. Ова правила се примењују на контролне карте на којима су величине неких варијабли исцртане у односу на време.

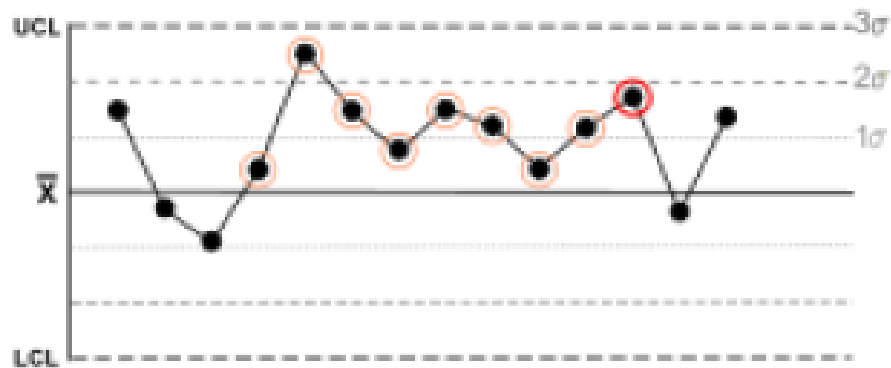
Ова правила су заснована на средњој вредности и стандардној девијацији узорака . (Nelson, 1984)

1. **Једна тачка је удаљена више од три стандардне девијације од средње вредности**— у овом случају један узорак се налази изван контролних граница. На слици 4.6 је приказан пример који илуструје ово правило где се проблематични узорци налазе са изван горење и доње контролне границе.



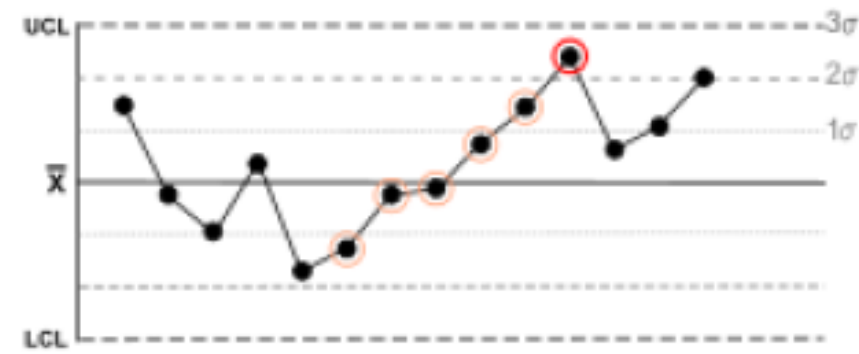
Слика 4.6. Прво правило

2. **Девет или више тачака се налазе са исте стране у односу на средњу вредност**—овај случај указује на склоност узорака ка гравитирању према једној од контролних граница. Другим речима, за овај случај је карактеристичан губитак варијабилности, па самим тим указује на нестабилност процеса у раној фази. На слици 4.7 је приказан такав процес код кога девет од четрнаест тачака теже ка горњој контролној граници.



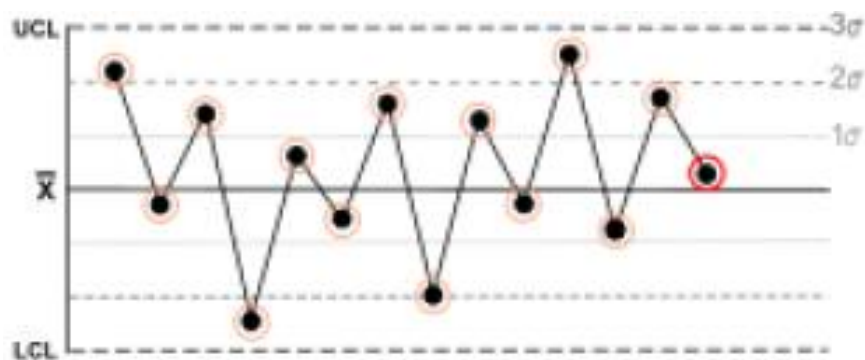
Слика 4.7. Друго правило

3. **Шест или више тачака континуирано расту или опадају**— овај случај указује на формирање тренда услед појаве значајних фактора, који се повећавају са сваким следећим узорком. Тренд може расти (ка горњој контролној граници), или опадати (ка доњој). На слици 4.8 је приказан случај појаве растућег тренда.



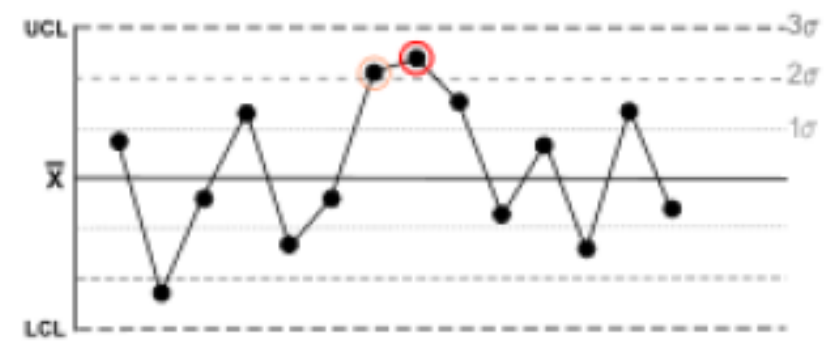
Слика 4.8. Треће правило

4. **Четрнаест (или више) тачака за редом се мењају у правцу, а затим се смањују**— у овом случају јављају се превелике осцилације које варирају у опсегу од три стандардне девијације, наизменично ка горњој и доњој контролној граници. Положај стандардне девијације не утиче на ово правило. На слици 4.9 је приказан случај оваквог процеса.



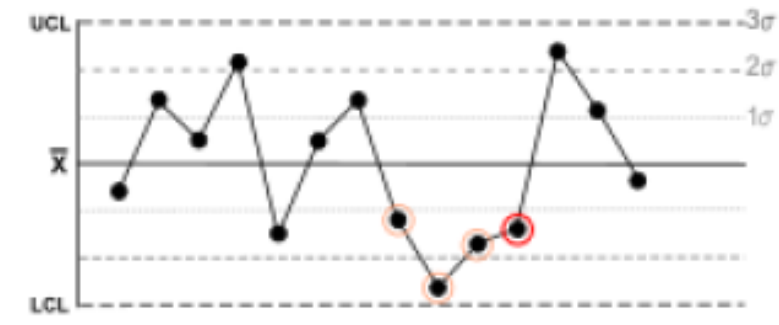
Слика 4.9. Четврто правило

5. **Две од три тачке су више од две стандардне девијације удаљене од средње вредности у истом правцу**—у овом случају постоји средња тенденција да узорци у средини процеса буду ван контроле и да теже ка некој од контролних граница. На слици 4.10 приказан је процес код кога је приказана оваква тенденција.



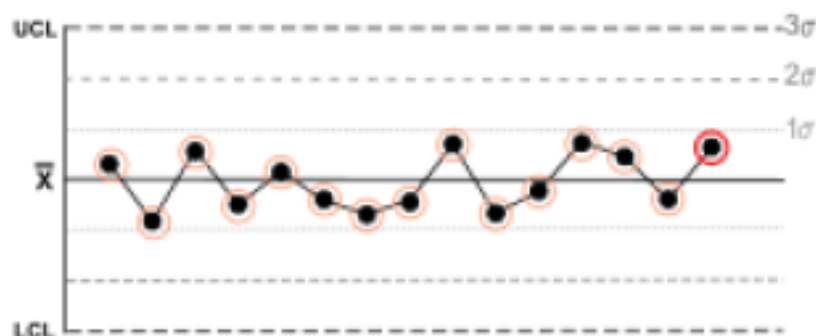
Слика 4.10. Пето правило

6. **Четири од пет или свих пет тачака су удаљене више од једне стандардне девијације у односу на средњу вредност и то у истом правцу**—ово указује на снажне тендеције процеса да се у једном тренутку нађе ван контроле. Као и у претходном случају, ово правило важи са обе стране средње вредности и примењује се без обзира да ли узорци имају тенденције према горњој или доњој контролној граници. На слици 4.11 је приказан процес који има тенденције ка доњој контролној граници.



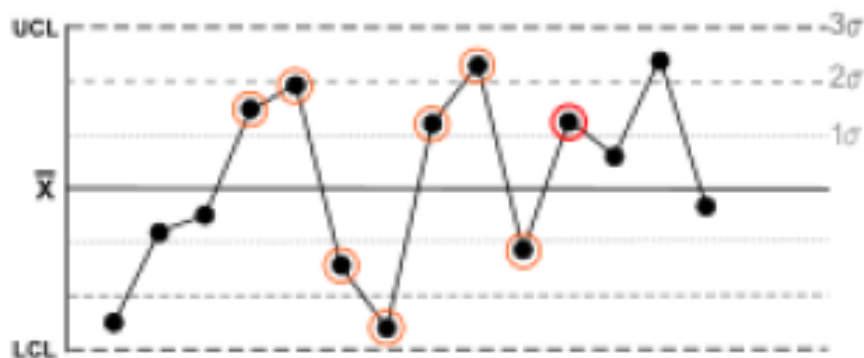
Слика 4.11. Шесто правило

7. **Петнаест тачака које се налазе у опсегу једне стандардне девијације и варирају са обе стране средишње линије**—вероватноћа да се узорак нађе у опсегу $\pm 1\sigma$ износи **68.27%**, али да се то понови петнаест пута за редом износи: $(0.6827)^{15} = 0.00326$, што је јако мала вероватноћа, али када до ње ипак дође она сигурно указује на то да нешто није у реду са разматраним процесом. Пример оваквог процеса приказан је на слици 4.12.



Слика 4.12. Седмо правило

8. **Осам тачака за редом изван једне стандардне девијација**—вероватноћа да се узорци нађу у изван једне стандардне девијације осам пута за редом је изузетном мала и ретко се дешава у пракси, али такође указује на нестабилност процеса. Пример оваквог процеса приказан је на слици 4.13.



Слика 4.13. Осмо правило

5. Регресиона и корелациона анализа

Регресиона анализа представља скуп статистичких метода којима се може открити да ли постоји веза између различитих појава и какве су оне по облику и смеру.

Коришћењем регресионих модела може се установити како се једна променљива мења под утицајем друге. У овом случају постоји међусобна веза између две променљиве и то на основу регресионе и корелационе анализе.

Промена одлика статистичког скупа обично утиче на промену других одлика због међусобне повезаности. Као што је већ речено, ова повезаност може се разликовати по смеру као и по једначини повезаности. Најужа или најснажнија веза између одлика назива се **функционалном или детерминистичком везом**. У том случају, веза између једне одлике тачно одговара вредности друге. Када се ради о мањој повезаности између одлика, која су подложна мањим или већим одступањима, таква веза назива се **корелативном или стохастичком везом**.

Обично се случајна променљива идентификује као независна x , а друга као зависно случајно променљива y .

Функционална веза се јавља када једној вредности независно променљиве x одговара једна вредност променљиве y .

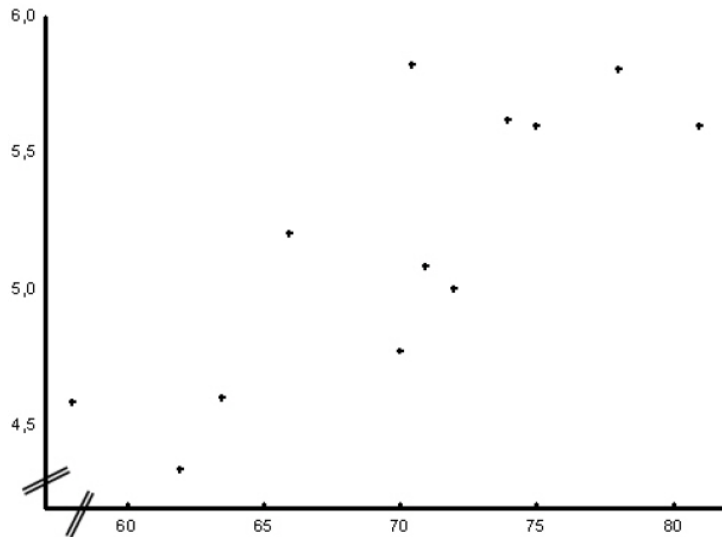
Стохастичка веза јавља се у случајевима када једној вредности независно променљиве x одговара више могућих вредности зависно променљиве y .

Испитивање зависности у статистичкој анализи има два правца (*Chiu.1991*):

- облик зависности који истражује регресиона анализа,
- јачину зависности коју одређује корелациона анализа.

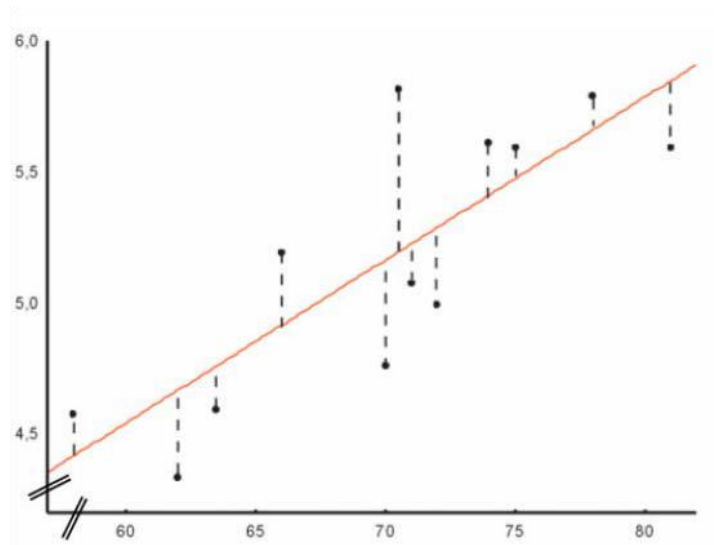
5.1. Регресиона анализа

Код регресионе анализе уочава се облик повезаности између две променљиве помоћу регресионе линије. Како је однос променљиве y према променљивој x често различит, први корак ка откривању повезаности је уцртавање дијаграма растурања, односно, дисперзије измђу узорака, као што је приказано на слици 5.1.



Слика 5.1. Дијаграм дисперзије

Како би се приближно одредила линеарна веза између те две величине, може се конструисати правац, заснован на аритметичкој средини вредности узорака, што је приказано на слици 5.2.



Слика 5.2. Аритметичка средина изражена правом

На слици 5.2 је приказано екзактно утврђен правац линеарне везе, који је одређен из услова да је збир квадрата вертикалних удаљености тачака од правца најмањи, односно, методом најмањих квадрата. Тако утврђен правац повезаности две варијабле одређује регресиона линија која је изражена једначином:

$$y = a + b \cdot x \quad (5.1)$$

при чему је:

y – зависно променљива ,

x – независно променљива,

a – регресиона константа,

b – коефицијент регресије.

У овом случају, зависно променљива y је непозната, а израчунава се на основу вредности променљиве x која је позната. Тада је регресиона константа и коефицијент регресије одређен методом најмањих квадрата, а имају формуле:

$$a = \bar{y} - b \cdot \bar{x} \quad (5.2)$$

$$b = \frac{n \cdot \sum xy - \sum x \cdot \sum y}{n \cdot \sum x^2 - (\sum x)^2} \quad (5.3)$$

У једначин n представља број парова који не би требао да буде мањи од 12 како би се добила одговарајућа регресиона права, и прави облик који указује на зависност између појава. Параметар a одређује ниво регресионе праве и представља вредност y за $x=0$, односно, представља тачку у којој се секу регресиона права и y оса. То је почетна вредност зависне y пре почетка деловања независне x .

Карактеристике параметара су:

- ($a = 0$) – регресиона права пролази кроз координатни почетак. То значи да не могу да имају негативне вредности јер полазе од нултог нивоа.
- ($a > 0$) – регресиона права сече ординату изнад координатног почетка.
- ($a < 0$) – регресиона права сече ординату испод координатног почетка.

Параметар b представља коефицијент регресије и чини тангенс угла кога регресиона права заклапа за x осом.

Карактеристике параметра b су:

- ($b = 0$) – регресиона права је паралелна са x осом. У том случају обележје y има исту вредност па је независно од x .
- ($b > 1$) – регресиона права се удаљава од x осе.
- ($b < 1$) – регресиона права је ближа x оси, а удаљава се од y .

Линијом регресије могуће је извршити интерполацију, односно, одредити вредност y за било коју вредност x .

Пример 1: За дванаест узорака материјала дати су параметри густине и тврдоће. Параметри су дати у табели 5.1 (X-густина, Y-тврдоћа)

N	x	y
1	4.21	108.4
2	4.3	112
3	3.6	87.3
4	4.41	99
5	3.8	93
6	3.7	92.3
7	3.8	90
8	3.8	94
9	3.81	95
10	3.7	92.3
11	2.9	77.96
12	3.9	80
Σ	45.93	1121.26

Табела 5.1. Разматрани параметри

Како би се израчунала једначина регресије, формира се радна табела (табела 5.2) која у овом примеру изгледа:

N	x	y	x ²	y ²	x*y
1	4.21	108.4	17.7241	11750.56	456.364
2	4.3	112	18.49	12544	481.6
3	3.6	87.3	12.96	7621.29	314.28
4	4.41	99	19.4481	9801	436.59
5	3.8	93	14.44	8649	353.4
6	3.7	92.3	13.69	8519.29	341.51
7	3.8	90	14.44	8100	342
8	3.8	94	14.44	8836	357.2
9	3.81	95	14.5161	9025	361.95
10	3.7	92.3	13.69	8519.29	341.51
11	2.9	77.96	8.41	6077.7616	226.084
12	3.9	80	15.21	6400	312
Σ	45.93	1121.26	177.4583	105843.192	4324.488

Табела 5.2. Радна табела

Одатле следи да је :

$$y_c = 17,71 + 19,783 \cdot x \quad (5.4)$$

Како би се конструисала регресиона линија потребно је одредити минимум две координатне тачке, при чему се узима најнижа и највећа вредност за променљиву x .

$$\text{за } x = 2,9 \rightarrow y_c = 17,71 + 19,78 \cdot 2,9 = 75,082$$

$$\text{за } x = 4,41 \rightarrow y_c = 17,71 + 19,78 \cdot 4,41 = 104,95$$

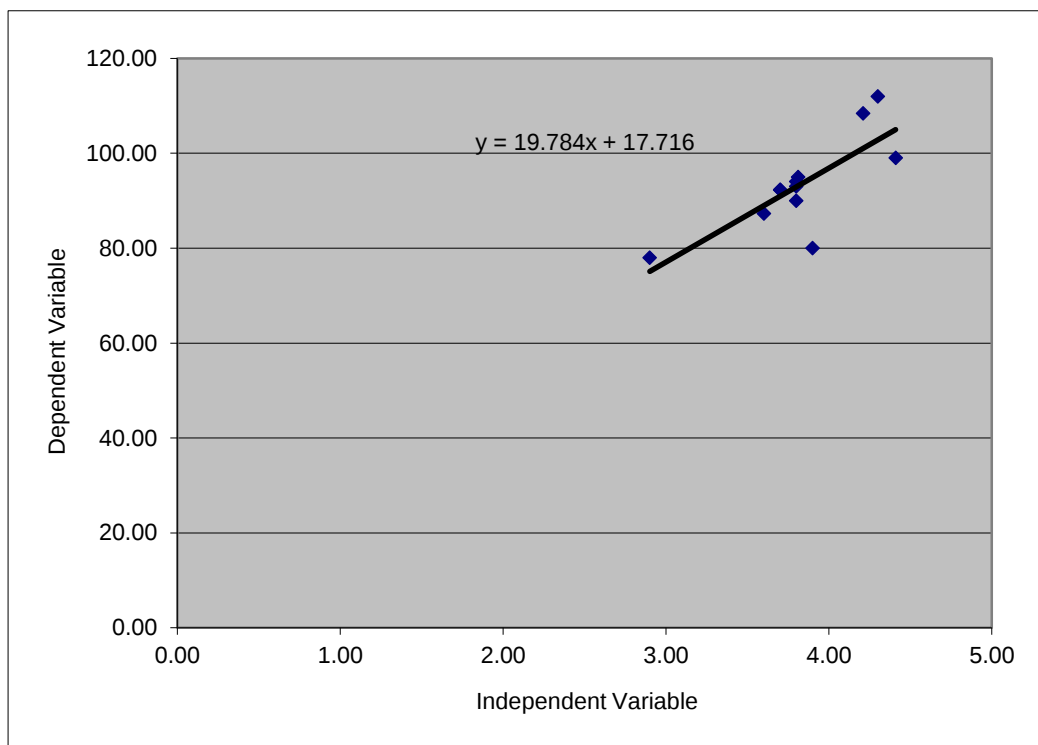
$$\bar{x} = \frac{\sum x}{n} = 3,8275$$

$$\bar{y} = \frac{\sum y}{n} = 93,4383$$

$$b = \frac{n \cdot \sum xy - \sum x \cdot \sum y}{n \cdot \sum x^2 - (\sum x)^2} = 19,783$$

$$a = \bar{y} - b \cdot \bar{x} = 17,71$$

y_c – оцена просечне вредности за вредност независно променљиве.



Слика 5.3. *Графикон регресије*

Као што је приказано на слици 5.3, из графика се види да је регресиона линија близу свих тачака, што зналчи да је збир квадрата одступања минималан.

5.2. Корелациона анализа

Корелациона анализа указује на степен повезаности између две променљиве, односно, корелацијом се мери интензитет и карактеристике већ утврђених вредности између две променљиве.

Степен интензитета повезаности између променљивих у линеарном односу мери се (Chiu.S1991):

- **коваријансом**—апсолутном мером интензитета корелације,
- **коэффицијентом просте линеарне корелације** — као релативном мером интензитета везе.

Коваријанса према томе представља заједничку меру варијабилности, па се математички може представити као збир обе варијабле:

$$C_{xy} = SD_x^2 + SD_y^2 \quad (5.5)$$

$$C_{xy} = \frac{\sum(x-\bar{x})^2}{n} + \frac{\sum(y-\bar{y})^2}{n} \quad (5.6)$$

Из чега следи да је формула за коваријансу:

$$C_{xy} = \frac{\sum xy}{n} - \bar{xy} \quad (5.7)$$

Где је **n** величина узрока.

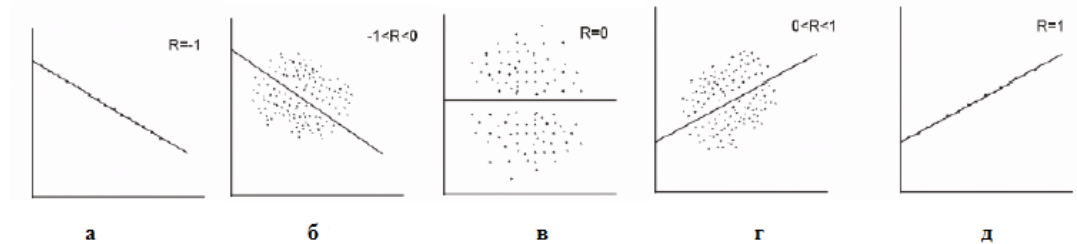
Упркос томе, коваријанса није погодна за процену, па се приступа израчунавању релативне мере израчунавањем коэффициента просте линеарне корелације.

Коефицијент просте линеарне корелације или **Пирсонов коефицијент** чини коваријасу, изражену јединицама стандардних девијација за обе варијабле, а израчунава се као количник коваријансе и производа стандардних девијација обе варијабле. Формула просте линеарне корелације је :

$$r_{xy} = \frac{C_{xy}}{SD * SD} = \frac{\text{коваријаса}}{\text{производ стандардних девијација } x \text{ и } y} \quad (5.8)$$

Коефицијент просте линеарне корелације указује на зависности између променљивих и одређује величину растурања података око регресионе линије.

Повећање линеарне повезаности, расипање, се смањује док у случајевима апсолутног слагања сви подаци леже на истој линији. На слици 5.4 су приказани графици за пет различитих случајева.



Слика 5.4. (а) идеална негативна корелација, (б) негативна, (в) без корелације, (г) позитивна, (д) идеална позитивна

Из овога следи да су варијабле уско повезане тако да када опадају вредности променљиве x опадају и вредности зависне y . У том случају је корелација позитивна, односно, $r > 0$, док, када већим вредностима независно променљиве x , одговарају мање вредности y , односно, долази до падања вредности за x долази до раста вредности y . Тада долази до негативне корелације $r < 0$.

Коефицијент корелације никад нема вредност 1 или -1, јер у том случају не би постојала статистичка веза, већ математичка. Вредност корелације се може оценити преко одговарајућих таблица за граничне вредности xy .

Оне показују колика мора да буде најмања вредност за xy , да би се одговарајући праг значајности могао сматрати статистички значајним. Број степена слободе израчунава се по формули $S.S. = n - 2$, где је n број парова.

Пример 2: Утврдити за претходни пример да ли постоји корелација између тврдоће и густине материјала.

Но: Између густине и тврдоће не постоји корелација.

На: Између густине и тврдоће постоји корелација.

Пре свега треба израчунати коефицијент просте линеарне корелације.

Формула је :

$$r_{xy} = \frac{C_{xy}}{SD_x \cdot SD_y} \quad (5.9)$$

Нако тога, је потребно израчунати C_{xy} , SD_x и SD_y .

$$C_{xy} = \frac{\sum xy}{n} - \bar{x}\bar{y} \quad (5.10)$$

Због тога што за дати пример већ постоји радна табела (приказана у склопу примера 1), може се одмах израчунати коваријанса.

$$C_{xy} = \frac{4226,56}{12} - 3,727 \cdot 93,575 = 3,46$$

$$SD_x = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2} = \sqrt{\frac{168,61}{12} - 3,727^2} = 0,4$$

$$SD_y = \sqrt{\frac{\sum y^2}{n} - \bar{y}^2} = \sqrt{\frac{106101,59}{12} - 93,575^2} = 9,25$$

Одатле следи да је:

$$r_{xy} = \frac{C_{xy}}{SD_x \cdot SD_y} = \frac{3,46}{0,4 \cdot 9,25} = \frac{3,46}{3,7} = 0,94$$

Из приложеног се види да је вредност коефицијента корелације **0,94**, што значи да између густине и тврдоће материјала постоји директна корелација.

У ово случају је $S.S. = n - 2 = 12 - 2 = 10$.

За степен слободе 10 и $p = 0.05$, гранична вредност за $g_{xy} = 0,576$

$R_{xy} = 0,94 > g_{xy}(10; 0,05) = 0,576$ и $p < 0,05$

Због тога што је добијена вредност $g_{xy} = 0,94$ већа од граничне вредности из табеле, одбацује се нулта и прихвата се алтернативна хипотеза са грешком $p < 0,05$ и $P > 95\%$. Тада се може тврдити да постоји јака позитивна корелација између густине и тврдоће материјала.

За степен слободе 10 и $p = 0,01$ гранична вредност $g_{xy} = 0,708$.

$R_{xy} = 0,94 > g_{xy}(10; 0,01) = 0,708$ и $p < 0,01$

И у овом случају се може тврдити да постоји јака позитивна корелација.

Коефицијент прости линеарне корелације се израчунава из узорка. Због тога се поставља питање да ли је узорак довољно значајан за доношење непристрасне оценое коефицијента основног скупа. Тада се тестира хипотеза, која треба да пружи одговор на питање да ли је израчунати прости коефицијент линеарне корелације из узорка (g_{xy}) у исто време и прецизна оцена простог коефицијента линеарне корелације основног скупа (R_{xy}).

Ако се тестом одбаци нулта хипотеза, прихвата се израчуната вредност коефицијента корелације из узорка као права оцена коефицијента у основном скупу.

Тестирање просте линеарне корелације заснива се на Студентовом распореду за $n-2$ степена слободе, који се најчешће употребљава као тест значаја нулте хипотезе, а добијена вредност се тумачи на исти начин као и код Студентовог t – теста.

Тест се математички дефинише формулом:

$$t = \frac{r_{xy}}{\sqrt{\frac{1-r_{xy}^2}{n-2}}} \quad (5.11)$$

односно,

$$t = r_{xy} \frac{\sqrt{n-2}}{\sqrt{1-r_{xy}^2}} \quad (5.12)$$

где је:

- r_{xy} – добијена вредност из узорка
- n – величина узорка

Број степена слободе се израчунава по обрасцу $S.S. = n - 2$.

У овом случају добијена t вредност се тумачи на исти начин као и код класичног Студентовог t теста, што значи да добијену вредност $r_{xy} = 0,94$ за дванаест случајева:

Но: $R_{xy}(\text{основног скупа}) = 0$

На: $R_{xy}(\text{основног скупа}) \neq 0$

$$t = \frac{0.94}{\sqrt{\frac{1-0.94^2}{12-2}}} = \frac{0.94}{0.1078} = 8,712$$

$$t = 8,712 > t(10 \text{ и } 0,05) = 2,23 \text{ и } p < 0,05$$

Због тога што је добијена вредност за t већа од 8,712, односно, већа од вредности из табеле $t = 2,23$, за степен слободе 10 и праг значаја $p = 0.05$, одбацује се нулта хипотеза и прихвата алтернативна са грешком $p < 0,05$ и сигурношћу $P > 95\%$, закључује се да између својства материјала постоји висок степен корелације, а добијена вредност $r_{xy} = 0,94$ представља стварну меру, па се закључак може применити на целокупну популацију узорака.

Пирсонов коефицијент корелације даје информације да ли је повезаност варијабли слаба, умерена, јака или веома јака. Међутим, он нам не даје и информацију колико је зависна променљива условљена вредностима независно променљиве, а колико другим факторима.

Овај проблем решава коефицијент детерминације, који се најлакше израчунава као други степен коефицијента прости линеарне корелације и он је мера за објашњени варијабилитет: $r_{xy}^2 = 0,94^2 = 0,8836$

Вредност коефицијента детерминације указује да су вредности тврдоће материјала 88,36% детерминисане густином истог.

6. Вишеструка регресија и корелација

Испитивње зависности једне појаве од две или више назива се вишеструка регресија. Циљ регресије је да открије што је могуће више фактора који утичу на зависну променљиву. Полази се од хипотезе да, што је више независних варијабли, све је мањи утицај стандардне грешке ε_i , $i = 1, 2, \dots, n$. Важно је пажљиво одабрати променљиве, које ће се укључити у модел.

Основни вишеструки регресиони модел изгледа на следећи начин:

$$\hat{x}_{i1,23\dots m} = a_{1,23\dots m} + b_{12.34\dots m}x_{i2} + b_{13.24\dots m}x_{i3} + \dots + b_{1m.23\dots m-1}x_{im} + \varepsilon_i, i = 1, 2, \dots, n \quad (6.1)$$

$a_{1,23\dots m}$ – слободан члан

$\hat{x}_{i1,23\dots m}$, $i = 1, 2, \dots, n$ – појединачне вредности регресије

$\hat{x}_{i2\dots}x_{ij}$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$ – вредности независних променљивих

$b_{12.34\dots m}$, $b_{13.24\dots m}$, $b_{1m.23\dots m-1}$ – регресиони коефицијент

ε_i , $i = 1, 2, \dots, n$ – случајна грешка

m – број независних променљивих

n – број јединица у узорку

Овакав модела даје најбоља предвиђања вредности зависне променљиве на основу независних променљивих, под условом да су претпоставке испуњене. На основу величине регресионих коефицијената можемо закључити колики је релативни утицај, или важност сваке независне променљиве, ако се ти коефицијенти конвертују у коефицијенте β . Ови коефицијенти се добију када се све вредности променљивих стандардизују.

Једна од претпоставки за употребу регресионе анализе јесте постојање линеарне зависности између варијабли. Она је неопходна, јер анализа започиње израчунавањем коефицијената просте корелације (биваријантних корелација) за све парове варијабли, а сва ова израчунавања захтевају линеаран однос између парова варијабли. Вишеструка регресија ће бити приказана на хипотетичком примеру са 7 независних варијабли. Приликом корелационе анализе, од посебног интереса је одређивање степена повезаности између варијабли. Корелациона анализа нам пружа следеће:

- релативну важност сваке варијабле у поређењу или утицају на зависну варијаблу,
- степен до којег све независне варијабле комбиновано објашњавају варијације зависне варијабле.

Одговоре на ова питања добијају се преко величине стандардизованог регресионог коефицијента β и коефицијента просте корелације r . У табели 6.1 је приказан пример са седам независних варијабли.

Варијабла	Коефицијент просте корелације r	Стандардизовани регресио коефицијент β	Регресиони коефицијент b
1	2	3	4
X1	0,63	0,55	2,89
X2	0,52	0,27	10,41
X3	0,40	0,15	6,62
X4	0,21	0,17	1,32
X4	0,11	-0,04	-5,32
X6	0,06	0,22	3,44
X7	0,03	0,01	4,45

Табела 6.1. *Седам независних варијабли*

У другој колони, коефицијенти просте корелације показују јачину везе између сваке независне променљиве посебно са зависном променљивом Y . Овај коефицијент се креће у интервалу од 0,03 до 0,63. Када се ове вредности подигну на квадрат добијају се коефицијенти детерминације који објашњавају колико дата независна варијабла има удела у објашњавању варијација независне променљиве. На пример, ако се први регресиони коефицијент подигне на квадрат, добија се да независна променљива **X1** објашњава 39,7% варијација зависне променљиве Y .

6.1. Мултиколинеарност

Пошто се коефицијенти корелације и β коефицијенти узимају као мере релативне важности сваке независне варијабле, вредности у другој и трећој колони табели 6.1 би требале да буду пропорционалне или бар да опадају исти редом. Међутим, види се да то није случај. Разлог лежи у мултиколинеарности или просто колинеарности. Мултиколинеарност показује колика је међузависност између независних варијабли. Што је већа мултиколинеарност, то се више одражава на β коефицијенте и они све мање могу да се употребе као показатељи релативног утицаја сваке независне варијабле.

Разлог лежи у томе што се регресиони коефицијенти, b и β , увек израчунавају тако да дају најбоље могуће предвиђање зависне варијабле Y , а не да покаже релативну важност сваке независне променљиве X . Када је мултиколинеарност мала и не постоји онда су регресиони коефицијенти приближно пропорционални коефицијентима просте корелације, па и једни и други дају сличну представу о релативној важности независних варијабли. Ако постоји значајна мултиколинеарност, онда ће најзначајнијој независној варијабли бити додељена права вредност β коефицијента, док ће код осталих независних β вредност бити много мања да би се избегла међузависности и међусобни утицај независних варијабли.

У табели 6.1, пошто величине β коефицијената нису пропорционалне са коефицијентима корелације, може се закључити да постоји значајна мултиколинеарност.

На пример, видимо да најзначајнија независна варијабла **X1** има висок коефицијент корелације и **β** коефицијент, али већ **X2** има нешто мањи коефицијент корелације али дупло мању вредност **β** коефицијента. Несразмера се понавља и код других варијабли у моделу. То је због тога што се преклапа утицај независних варијабли, па су због тога **β** коефицијенти свих варијабли осим **X1** пуно мањи. Овај проблем се може решити преко степњисе регресије.

Прихватљиви ниво мултиколинearности није лако одредити. Он зависи од броја независних варијабли у моделу, колико њих је корелисано и у ком обиму. Потребно је на почетку израдити таблицу простих коефицијената корелације између свих варијабли. Прости коефицијенти просте корелације до **0,5** између неколико независних варијабли обично не би требало да утичу на регресионе коефицијенте. Ако су поменути коефицијенти просте корелације већи од **0,7** онда је у питању озбиљан проблем.

Могућа решења су следећа (Myers & Mullet, 2003):

- израдити табелу са свим варијаблама и њиховим коефицијентима просте корелације. Ако код неког пара варијабли коефицијент прелази **0,7**, онда се једна од две варијабле елиминише, обично она која има мању корелацију са зависном варијаблом **Y**.
- Уколико три или више независних варијабли имају велику међусобну корелацију, изабере се она са највећом корелацијом са **Y** и онда се елиминишу све остале или се изради нова заједничка варијабла од свих међузависних варијабли (на основу ваганих вредности, или на основу пропорција у корелацији са **Y**).
- Изради се анализа главних компоненти за све независне варијабле. Ова техника тражи групу од две или више варијабли које су високо или осредње међусобно корелисане, али су истовремено неповезане са осталим варијаблама. За сваку од ових група израђују се вредности које се зову фактор скорови, што је врста ваганих просека. Пошто су ови фактор скорови некорелисани и садрже већину информација из оригиналних варијабли, они могу да се употребе као нови сет независних варијабли у вишеструком регресионом моделу. Ова опција је најбоља и препоручује се посебно ако је у питању велики број варијабли (преко 50). Ипак, овим се губи могућност да се посматра свака оригинална варијабла појединачно.

6.2. Индекс детерминације

Вишеструка регресија такође показује колико је јака међузависност зависне варијабле са свим независним варијаблама преко индекса корелације **R**.

Индекс детерминације **R²** показује колики је проценат варијабилитета зависне променљиве објашњен варијабилитетом независних променљивих. У примеру из табеле 6.1, индекс детерминације је **48%**, што је далеко од пожељне величине од **70%**. То значи да неке варијабле, које имају значајну повезаност са независном променљивом **Y** недостају у моделу, али није познато које су то варијабле.

Пошто се индекс корелације и индекс детерминације рачунају на основу података који су прикупљени, дакле пост – фестум, не може се ништа учинити на његовом побољшању. Ипак, у пракси се препоручује да се прво уради пилот истраживање, где се на мањем узорку тестира што већи број варијабли да би се идентификовале све оне које имају најзначајнији утицај, а затим се уради велико посматрање на комплетном узорку где се прикупљају подаци о тим варијаблама.

6.3. Значај вишеструке регресије

На основу претходно реченог, вишеструка регресија се користи за добијање одговора на следећа питања:

- Колико добро све независне варијабле комбиновано објашњавају, или им се може приписати разлог за варијације зависне варијабле (R^2) ?
- Колика је релативна важност сваке независне варијабле у објашњавању варијација зависне варијабле (β коефицијенти), под условом да не постоји значајна мултиколинearност?
- Која је најбоља предвиђена вредност зависне варијабле за било коју комбинацију независних варијабли ?
- Који се обим промене зависне варијабле може очекивати за сваку јединицу промене сваке независне варијабле (коефицијенти просте корелације) ?

Претпоставке на којима се заснива модел вишеструке регресије су сличне онима које важе за просту регресију и оне гласе:

- Облик зависности између свих варијабли је линеаран, односно, права линија. Ово је поготово важно за однос независних варијабли са зависном варијаблом,
- Све варијабле су континуалне,
- Све варијабле имају интервал варијације, дисперзију, односно, варијансу које имају смисла, односно, већина опсервација није једна вредност, или интервал.

У бази се налази барем три до пет пута више јединица посматрања него што је варијабли јер би у супротном регресиони коефицијенти били непоуздани.

6.4. Тестирање статистичке значајности

Пре објашњавања резултата потребно је тестирати њихову статистичку значајност. Ако **R**, **b** и β нису статистички значајни, закључује се да ниједна независна варијабла нема стварну повезаности са зависном варијаблом. То значи да добијени модел нема практичну вредност. Већина статистичких софтвера има опцију тестирања.

Уколико су сви регресиони коефицијенти **b** статистички значајни, онда ће и индекс корелације **R** бити сигурно значајан.

У обрнутом случају то не мора да се деси, јер је могуће да се због великог броја варијабли добије статистички значајно **R**, а да **b** коефицијенти нису значајни.

Униформно оцењивање. Још један проблем, који може да се јави, јесте када за неку јединицу посматрања не постоје варијације у прикупљеним вредностима варијабли.

На пример, испитаник је на сва или скоро сва питања одговорио истом оценом (на скали од 1 до 10 он је на сва питања заокружио оцену 5). Пошто у том случају не постоје варијације за дату јединицу посматрања, не долази до коваријације са осталим варијаблама и јединицама посматрања. Повећава се само величина узорка **n** или скупа **N**, али се не повећава коваријанса. На тај начин се вештачки снижава корелација. Ни овде не постоји идеално решење. Уколико су све вредности једнаке, боље је такву јединицу елиминисати из анализе. Уколико је присутан део вредности који се понавља за дату јединицу посматрања може се урадити следеће:

- елиминисати јединицу посматрања код које не постоји интервал у вредностима варијабли у довољној мери. На пример на мерној скали са 10 вредности интервал за ту јединицу посматрања су само три суседне вредности).
- Елиминисати јединицу посматрања код које постоји мали број варијација у односу на најчешћу вредност, на пример до 25% посматраних варијабли.
- Израчунати стандардну девијацију свих вредности варијабли за сваку јединицу посматрања и елиминисати оне јединице посматрања код којих је израчуната вредности близу нуле.

7. Закључак

Статистичка контрола квалитета се показала као поуздан алат у контроли процеса производње и праћења самог квалитета који из ње произилази. Осим тога, статистички подаци у знатном броју могу дати прогнозу и указати на то у којој мери су захтеви квалитета испоштовани, као и чему производни процес треба да тежи. Они такође представљају својеврстан показатељ којим се дефинише прихватање одређеног производа или услуге од стране тржишта, односно, указује на интервал поузданости за дату карактеристку квалитета.

Поред осталих статистичких метода којима се може на основу прикупљених података пратити и измерити квалитет неког производа према његовим атрибутима, најзначајније средство у праћењу процеса производње представљају контролне карте. Оне су као математичко средство веома погодан алат за остваривање циљева статистичке контроле и анализе квалитета. Њиховом применом омогућава се испуњење захтевних карактеристика квалитета које захтевају одређени стандарди.

Прилагођавањем проблемима у производњи статистичка контрола квалитета и контролне карте налазе широку примену у производној индустрији.

8. Литература

1. Crosby P. (1995). *Philip Crosby's Reflections on Quality*. McGraw-Hill
2. Ishikawa K. (1985) *What is total quality control? The Japanese way*.
3. Bakija, I., (1978), *Kontrola kvalitete*, Zagreb, Tehnička knjiga
4. Carter, M., *Control charts in quality control- Shewart charts*, доступно на: <http://printfu.org/read>., Приступљено августа 2018.
5. Đedović, B., (2009), *Vođenje i vrednovanje projekata*, Beograd, Fakultet za menadžment MSP.
6. Ivković, Z., (1992) *Matematička statistika*, Beograd, Naučna knjiga.
7. Ishikawa, K (1968). *Guide to Quality Control*. Tokyo: JUSE.
8. Nelson, Lloyd, (1984) "The Shewhart Control Chart—Tests for Special Causes," Technical Aids, Journal of Quality Technology, No. 16, Oct., pp. 238-239.
9. Myers & Mullet, (2011) Classification and competition analysis of air cargo logistics providers: *The case of Taiwan's high-technology industry*, Journal of Air Transport Management.