

Statisztikai folyamatszabályozás (*statistical process control, SPC*)

Solymosi Norbert

2021. május 18.

- 1 Telepi adatok
- 2 Shewhart-chart
 - Ellenőrzési határok
 - Minta mérete, mintavétel gyakorisága, csoportok
 - Mintázat elemzése
 - I. és II. fázis
- 3 Cause-and-effect diagram
- 4 Méréses ellenőrző diagramok
 - \bar{x} -diagram
 - R -diagram
 - \bar{x} - és R -diagram
 - s -diagram
- 5 Minősítéses ellenőrző diagramok
 - p -diagram
 - np -diagram
 - c -diagram
 - u -diagram
- 6 Kisebb eltérések vizsgálata
 - CUSUM
 - EWMA
- 7 Mikor, melyiket?
- 8 ARL
 - \bar{x} -diagram
 - p -diagram

- telepirányító rendszerek
- folyamatos adatgyűjtés, de csak pont/periódus-statisztika
- PLF (pl. vezetőképesség)
- limitált beépített elemzési lehetőségek
- rengeteg adat képződik, alig feldolgozva
- környezeti (pl. meteorológia) adatok illesztését nem teszik lehetővé a telepi szoftverek
- állomány-egészségügyi vizsgálatok
- állatorvosi feladatok átalakulása
- adatfeldolgozási, elemzési képességek új lehetőségeket jelenthetnek
- mivel arra nem lehet várni, hogy minden elemzésre alkalmas telepirányító szoftverek legyenek, olyan eszközt érdemes használni, amelynek segítségével mindenféle adatot kezelhetünk, elemezhetünk



BosTracto

Alapadatok

Pontozás

Tárgya:

ID: 12345

Pont: 1

Almozás
Bendő
Bélsár A
Bélsár B
Csülök
Kondíció
Padozat
Sántaság
Tőgybimbó
Tőgyhigiéna

Save

Mintavétel

SusTracto

Légtér: Hizlalda 1 l2

Start

Vizsgálat tartama: 09:04

✓ Hang mentése

Köhögés:

0 + -

Tüsszögés:

0 + -

Save Cancel

Mihez viszonyítsuk a telepi mutatókat?

- a viszonyítási alap mennyire felel meg az állománynak?
- fajta, takarmányozás, tartás



PigCHAMP
Benchmarking


USA 2016- 2nd Quarter Summary

Number of Farms = 344

Variable	Mean	SD	Median	Upper 10th Percentile	Lower 10th Percentile
Total number of services	1262.35	1113.040	865.00	3026.00	295.00
Number repeat services	79.97	92.448	46.00	212.00	6.00
Percent repeat services	6.51	5.187	5.56	13.17	1.19
Number of sows farrowed	1097.39	994.380	777.00	2678.00	246.00
Total pigs born	15521.32	14126.790	10735.50	35800.00	3297.00
Average total pigs per litter	13.95	1.080	14.01	15.20	12.60
Total pigs born alive	14002.68	12779.200	9642.50	32968.00	2883.00
Average pigs born alive/litter	12.60	0.944	12.68	13.76	11.52
Liveborn/female/yr	27.76	4.510	28.02	32.75	22.75
Total stillborn pigs	1068.51	1020.770	768.50	2280.00	173.00
Average stillborn pigs	0.99	0.374	0.94	1.48	0.56
Total mummified pigs born	450.14	732.814	242.50	1026.00	20.00
Average mummies per litter	0.35	0.365	0.30	0.61	0.07
Farrowing rate	84.03	7.405	84.97	91.45	76.07
Pre-weaning mortality (n=343)	14.44	6.889	13.10	21.00	8.22
Sows farrowed and weaned	1114.77	996.678	760.50	2625.00	240.00
Average age at weaning	20.30	2.105	20.11	22.44	18.31
Total pigs weaned	11694.24	10334.080	8124.00	27189.00	2582.00
Average litter weaning weight (n=73)	140.48	27.806	143.00	168.00	115.03
Pigs weaned per litter weaned	11.10	0.903	11.17	12.17	10.02
Pigs wmd / mated female / yr	24.93	4.641	25.25	30.30	19.28
Pigs wmd / female / year	23.79	4.542	24.11	28.95	18.30
Females entered	0.38	2.996	0.00	0.00	0.00
Sow and gilt deaths	51.50	57.378	28.00	119.00	8.00
Death rate	9.97	6.941	8.79	15.34	4.71
Sows and gilts culled	213.97	210.251	134.50	480.00	42.00
Culling rate (n=340)	44.31	18.386	42.47	69.26	20.72
Total sows	1885.68	1686.490	1294.00	4575.00	452.00
Ending boar inventory	7.34	30.348	3.00	14.00	0.00

TblCowLast30DayMilk_2014-04-06.xlsx - LibreOffice Calc

FájlSzerkesztésNézetBeszúrásFormátumEszközökAdatokAblakSúgó



MS Sans Serif 10

A1:A1048576 $f(x)$ Σ = HerdNo

	B	C	D	E	F	G	H	I		
1	CowNo	SessionDate	Time	Session	MilkDate	Time	Yield	YieldIsLow	PeakFlow	AverageFlow
2	8852	2014.03.17	2	2014.03.17	17,44	0	7,4	4,5		
3	6790	2014.03.12	2	2014.03.12	16,65		5,2	2,5		
4	8732	2014.03.12	2	2014.03.12	15,77		2,8	1,6		
5	9983	2014.03.17	2	2014.03.17	10,03		5	2,2		
6	105	2014.03.17	2	2014.03.17	16,54		5,4	3,3		
7	9162	2014.03.20	1	2014.03.20	8,96		4,3	2,1		
8	8540	2014.03.12	2	2014.03.12	13,76		4,1	2,7		
9	5413	2014.03.12	2	2014.03.12	16,1		4,1	2,2		
10	8831	2014.03.07	2	2014.03.07	17,18		6,7	3,6		
11	9867	2014.03.17	2	2014.03.17	14,05		4,3	1,9		
12	8746	2014.03.12	2	2014.03.12	15,11		2,9	1,5		
13	3469	2014.03.12	2	2014.03.12	8,16	1	3,1	1,4		
14	173	2014.03.17	2	2014.03.17	11,25	0	6,8	3,1		
15	35	2014.03.17	2	2014.03.17	12,31		3	1,5		
16	6919	2014.03.12	2	2014.03.12	15,62		4,7	2,8		
17	7441	2014.03.12	2	2014.03.12	17,75		6,7	2,9		
18	5413	2014.03.07	2	2014.03.07	14,8		4,9	2,3		
19	9648	2014.03.17	2	2014.03.17	11,56		3,6	2,2		
20	7133	2014.03.12	2	2014.03.12	17,22		3,2	1,9		

Folyamatos adatgyűjtés, értékelés, táblázatok

Information

Farm overview

Compartment overview

Cough counts

Activity

Feed detect

Weight detect

Environment

Online logbook

Feed reports

Slaughter transports

Operation

Operation journal

Service level agreements

New fattening period

Feed delivery

Transport registration

Administration

Companies

Farming periods

Change password

Log out

Cough Counting Data

Company

Dunahy

Farm

Dunahy

Compartment

Compartment 1

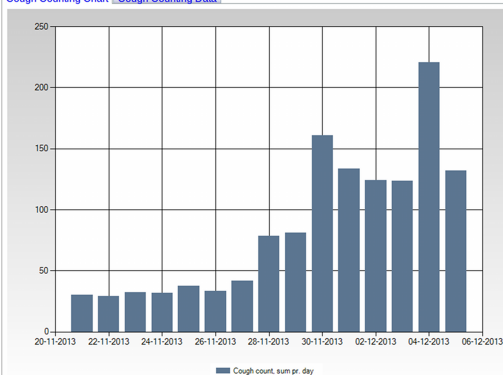
Date

21-11-2013

Count >=

Search

Cough Counting Chart | Cough Counting Data



Folyamatos adatgyűjtés, értékelés, ábrák

Comment on “A Bacterium That Can Grow by Using Arsenic Instead of Phosphorus”

István Csabai^{1,2*} and Eörs Szathmáry^{3,4}

Wolfe-Simon *et al.* (Research Articles, 3 June 2011, p. 1163; published online 2 December 2010) reported that bacterial strain GFAJ-1 can substitute arsenic for phosphorus in its biomolecules, including nucleic acids and proteins. Unfortunately, their study lacks crucial experimental evidence to support this claim and suffers from inadequate data and poor presentation and analysis.

The basic principles of life rest on the organization of different autocatalytic reaction systems (1), the exact chemical realization of which depends on the laws of chemistry, environmental conditions, and evolution. Finding that one element can be replaced by a similar one under certain conditions does not seem implausible, although finding such a system would be a major discovery. Wolfe-Simon *et al.* (2) reported that the bacterial strain GFAJ-1 can use arsenic (As) instead of phosphorus (P) in its basic biomolecules, but their conclusion is not well supported by the presented data.

To show the arsenate-dependent growth of strain GFAJ-1, Wolfe-Simon *et al.* measured the percentage of As and P in the dry weight of the bulk intracellular material of their samples. In table 1 in (2), they report that the mean intracellular As in +As/-P cells was $0.19 \pm 0.25\%$ by dry weight, but only 0.001 ± 0.0005 for cells

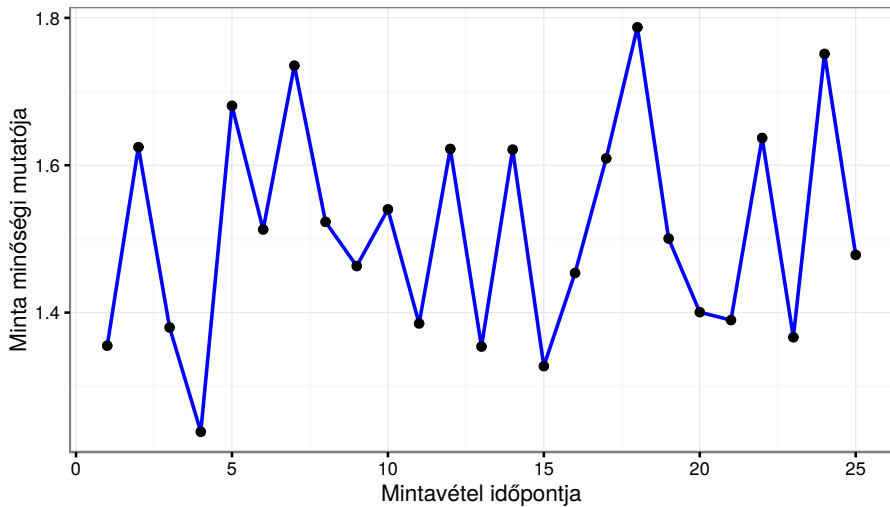
based on these values, we arrive at 5.05. From this value, we can calculate the P:As ratio that is the reciprocal of 5.05, namely 0.198, and conclude that on average, there is much more As than P in our samples. If we instead calculate the average P:As ratio first, it would be 5.05 and the As:P ratio 0.198, suggesting much more P than As. Wolfe-Simon *et al.* report a As:P ratio of 7.3 [shown in table 1 in (2)], but if calculated in reverse, this value would be a much less impressive 1.59.

Another concern with the data reported in table 1 (2) involves the calculated errors. The error for the As percent by dry weight ($\pm 0.25\%$) is larger than the value itself (0.19%), so the null hypothesis that the +As/-P sample contains no As, cannot be excluded. Although the P concentration seems to be considerably smaller for the +As/-P grown sample (0.019%) than for the -As/+P (0.54%), the authors do not discuss

This difference in As amount seems well beyond a reasonable standard fluctuation. Averaging together such bimodal data renders it difficult to draw meaningful conclusions from them.

Wolfe-Simon *et al.* used high-resolution secondary ion mass spectrometry to identify As in extracted, gel-purified DNA. Table S2 in (2) reports elemental concentrations and ion ratios for one +As/-P experiment, two -As/+P experiments, and one baseline measurement made on a blank agarose gel. The baseline value for P was 820 ± 143 parts per billion (ppb), yet the P concentration in the +As/-P DNA sample was only 299 ± 36 ppb. These two values alone indicate large fluctuations in the measurements and raise concerns about their accuracy. One possible solution is to use the reported baseline value in further calculations, because there is no reason a sample would contain less P than the blank gel itself. If we do so, we see that the amount of P is only slightly higher in the -As/+P samples. For the measurement of As concentration, the baseline value with estimated error is 15 ± 3 ppb. However, this error estimate appears to be too optimistic: The As concentrations reported for the two -As/+P experiments are 14 ± 3 ppb and 5 ± 1 ppb, meaning that even in three measurements, the value could fluctuate from 15 ppb down to 5 ppb, which suggests that the correct error should be around 10 ppb. If we take this value, we see that the As concentration reported for +As/-P sample (27 ± 10 ppb) is not significantly larger than the As concentration for the -As/+P sample or the baseline (15 ± 10 ppb). In several other places in their Research Article, Wolfe-Simon *et al.* do not provide specific error estimates, but rather use the ad hoc value of 10% as standard deviation.

Idősorok



A statisztikai folyamatszabályozás „nagyszerű hetese” (magnificent seven):

- ① Histogram (stem-and-leaf plot)
- ② Check sheet
- ③ Pareto chart
- ④ *Cause-and-effect diagram*
- ⑤ Defect concentration diagram
- ⑥ Scatter diagram
- ⑦ **Control chart**



Walter Andrew Shewhart
(1891-1967)

Shewhart: „A folyamatot akkor nevezzük stabilnak vagy statisztikailag kézbentartottnak (**in statistical control**), ha az ingadozás **véletlenszerű, időben állandó, nincsenek jól felismerhető és megnevezhető okai**, a jellemző jövőbeli értékei statisztikai módszerekkel megadható határok között vannak.”¹

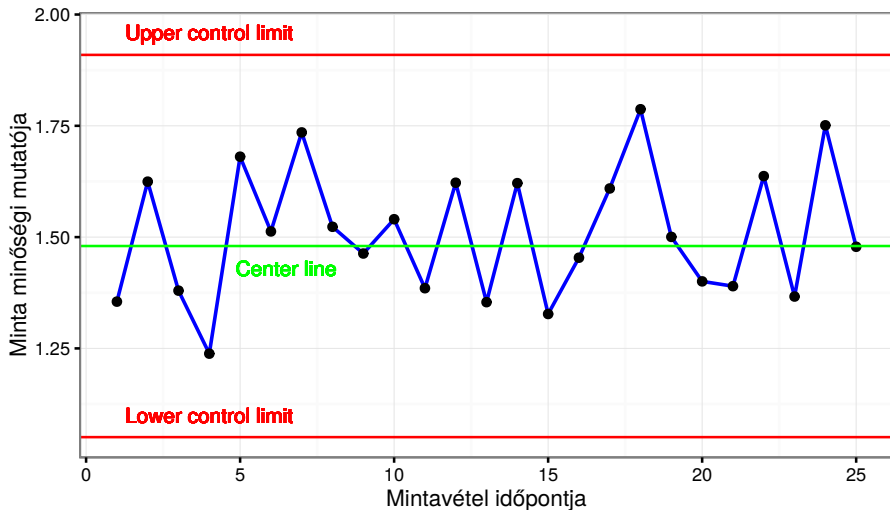
A megfigyelt variabilitás oka lehet:

- **véletlen** (chance causes of variation)
- **azonosítható** (assignable causes of variation)

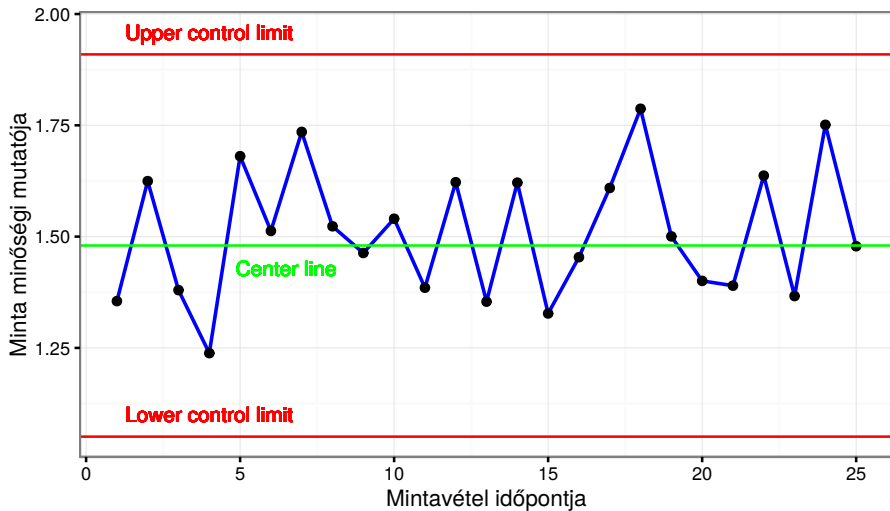
A variabilitás azonosítható okainak jelenléte esetén a folyamat kikerült az ellenőrzés alól (**out-of-control**)

A cél, hogy csökkentsük annak esélyét, hogy a folyamat kikerüljön az ellenőrzésünk alól. Illetve, ha kikerült, akkor gyorsan tudjuk detektálni. Ebben nyújthatnak segítséget az **ellenőrző diagramok**.

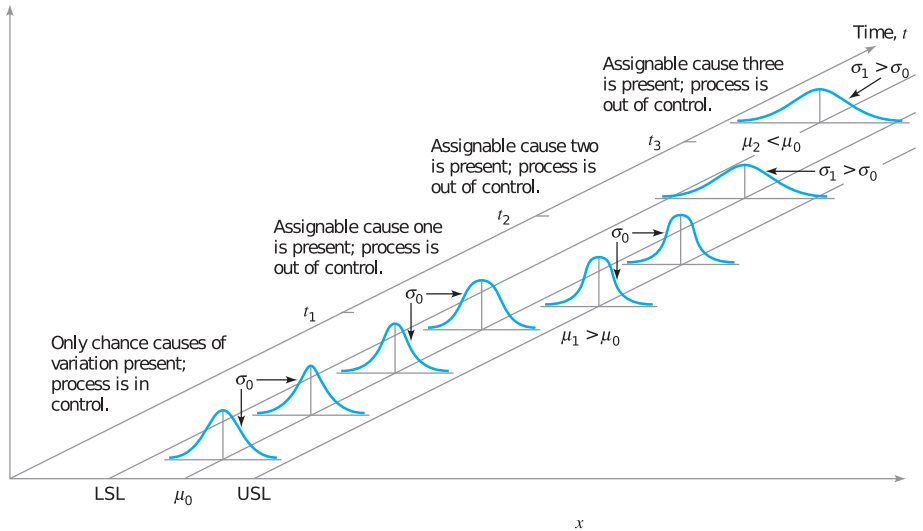
¹?



- A center line a folyamat stabil állapotának megfelelő minőségi mutató átlagát jelzi



- Az alsó és felső ellenőrzési határok a stabil folyamat határait jelzik



?

Egy sertéstelepen Shewhart-diagram használunk a malacok születési súlyának ellenőrzésére.

? által bemutatott táblázatból tudjuk a malacok születési súlyának átlagát és szórását.

Litter Size	≤ 9	10 – 11	12 – 13	14 – 15	≥ 16
% of litters	12.2	9.6	17.3	24.7	36.3
Parity	2.6	2.3	2.5	2.6	3.5
Litter size : total born	7.1	10.6	12.6	14.5	17.7
Litter size : live born	6.9	10.2	12.0	13.7	16.1
Litter size : stillborn	0.3	0.4	0.6	0.8	1.5
Mean birth weight (kg)	1.88	1.67	1.57	1.48	1.38
Within-litter SD (kg)	0.28	0.29	0.32	0.32	0.33
Coefficient of variation (%)	15	18	21	22	24
% small piglets	6	9	12	13	16

A 14-15 malacot fialó kocák malacainak átlagos súlya $\bar{x} = 1.48$ kg, a szórás pedig $\sigma = 0.32$ kg.

Az ellenőrző diagramunk rajzolásához minden nap lemérünk 5-5 újszülött malacot. Ezek súlyának átlagát pontokkal jelöljük az ábrán.

Az ellenőrző diagram határértékeit a standard hiba segítségével számítjuk:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.32}{\sqrt{5}} = 0.1431084$$

Ha $n \uparrow$, akkor $\sigma_{\bar{x}} \downarrow$!!!

Ha feltételezzük, hogy a kézbentartott „malactermelési” folyamat születési súly minták \bar{x} átlaga közelítőleg normális eloszlású, akkor várhatóan az \bar{x} mintaátlagok $100(1 - \alpha)\%$ -a fog a $1.48 - Z_{\alpha/2} \times 0.1431084$ és $1.48 + Z_{\alpha/2} \times 0.1431084$ közötti tartományba esni.

A $Z_{\alpha/2}$ önkényes konstans értéke 3^1 így a határértékek:

$$UCL = 1.48 + 3 \times 0.1431084 = 1.909325$$

és

$$LCL = 1.48 - 3 \times 0.1431084 = 1.050675$$

¹three-sigma control limits

Ha az ellenőrzési tartomány határait távolítjuk a centrális értéktől, akkor csökkentjük az I. típusú hiba valószínűségét, de növeljük a II. típusúét. Az ellenőrzési tartomány szűkítésével az ellenkező hatást érjük el.

A 3-sigma esetén $\alpha = 0.0027$, vagyis 99.73% a valószínűsége, hogy \bar{x} ebbe a tartományba esik. Vagyis 10000-ből 27 esetben tévesen jelezheti, hogy a folyamat kikerül az ellenőrzésünk alól. Ez mindkét irányban való határtúllépést jelent, az egyoldali ennek a fele, vagyis 0.00135.

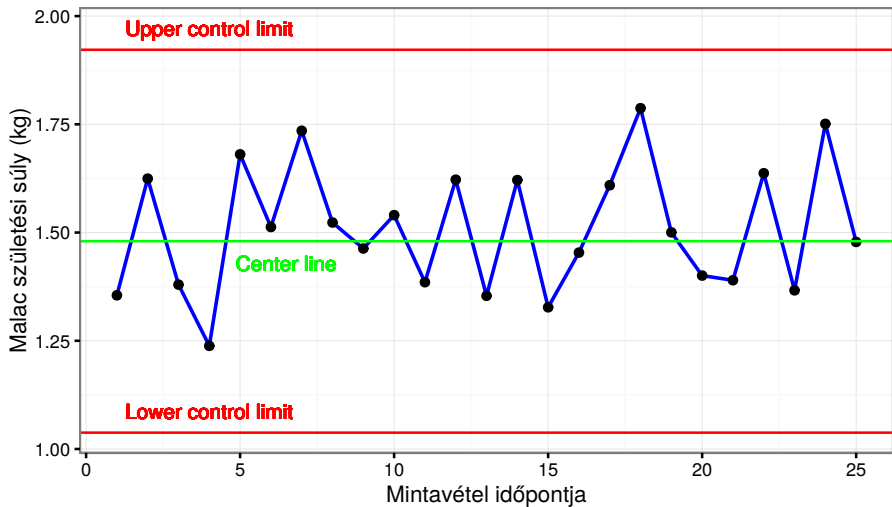
Esetenként nem szigmával adják meg az ellenőrzési tartományt, hanem az I. típusú hiba valószínűségének meghatározásával, így pl. ha ennek értéke 0.001 (egyik irányban), akkor az előző formulákban 3.09-es szorzót használunk:

$$UCL = 1.48 + 3.09 \times 0.1431084 = 1.922205$$

és

$$LCL = 1.48 - 3.09 \times 0.1431084 = 1.037795$$

Ezeket gyakran hívják 0.001 valószínűségi határoknak, azonban ez nem helyes, mert a két irányban 0.002 valószínűségi határnak kellene nevezni.



Vannak, akik nem csak egy határértéket használnak, hanem kettőt vagy hármat.

Ha két határértéket használunk az ellenőrzési tartomány meghatározására, akkor a legkülsőt cselekvési határnak (*action limits*), a belsőt figyelmeztetési határnak (*warning limits*) nevezzük. Ez utóbbi számszerűen:

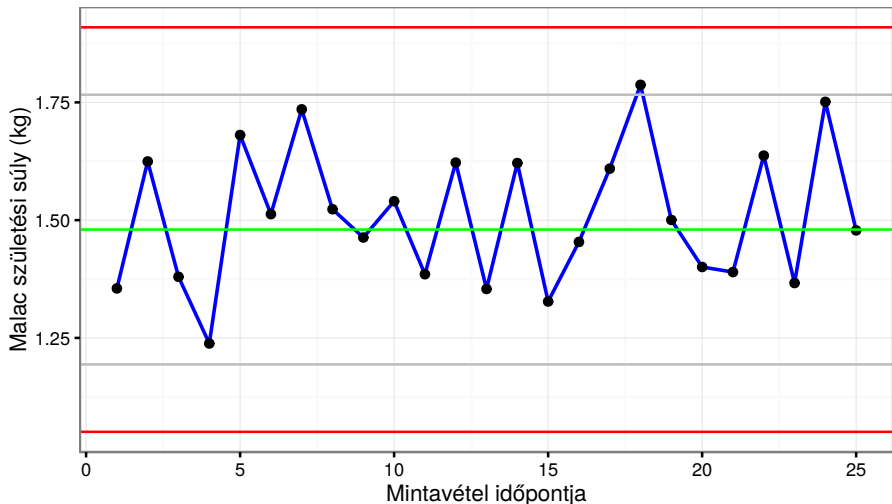
$$UCL = 1.48 + 2 \times 0.1431084 = 1.766217$$

és

$$LCL = 1.48 - 2 \times 0.1431084 = 1.193783$$

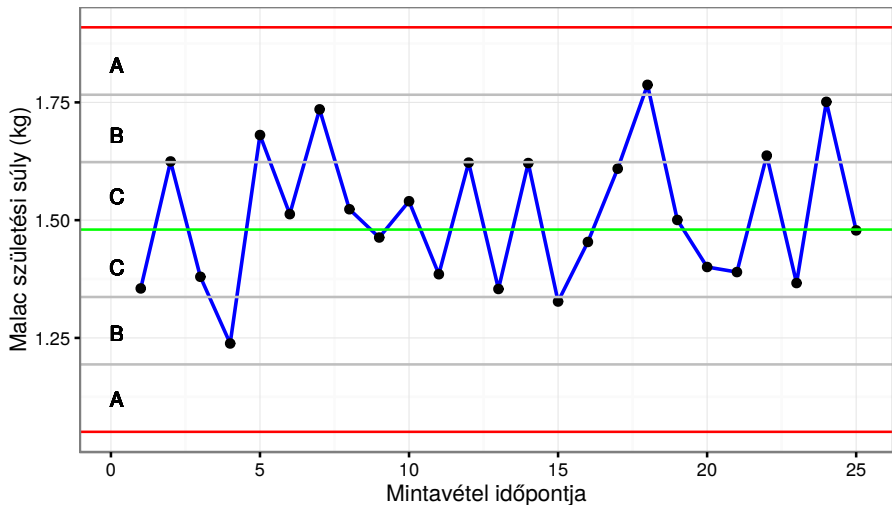
Ha valószínűségi határokat használunk, akkor a cselekvési limit 0.001, a figyelmeztetési pedig 0.025.

A figyelmeztetési határ használata növeli a diagram szenzitivitását, így gyorsabb reakciót tesz lehetővé. Azonban több téves riasztást is ad.



Ha egy vagy több pont a figyelmeztetési és cselekvési határ közé vagy a figyelmeztetési határhoz közel esik, akkor feltételezhetjük, hogy a folyamat nem működik tökéletesen.

Előfordul, hogy egy-szigmányira is meghúznak egy határt, amit alsó (a két-szigmányit felső) figyelmeztetési határnak nevezzük. A három határérték által létrehozott tartományok: A, B, C.



- a legjobb lenne, ha a lehető leggyakrabban, a lehető legtöbb mintát vehetnénk
- nagyobb mintamérettel kisebb változások könnyebben észlelhetők
- ésszerű mintavételi csoportok:
 - a csoportok közötti, az azonosítható okoknak tulajdonítható eltérések maximalizálhatók lehessenek
 - a csoporton belüli, az azonosítható okoknak tulajdonítható eltérések minimalizálhatók lehessenek
- ❶ azonos vagy nagyon közeli időpontban veszünk mintát (snapshot approach)
- ❷ az előző mintavétel óta létrejött egységeket reprezentáló minta (random sample approach)

- ha egy vagy több pont túl van az ellenőrzési határokon, vagy a pontok valamilyen nem véletlen mintázatot mutatnak, feltételezhető, hogy a folyamat kikerült az ellenőrzésünk alól
- az ábrán az 5-8. pontok mind az átlag felett helyezkednek el
- a 15. ponttól kezdve 4 pont folyamatos növekedést mutat (*run*, sorozat), ebben az esetben felfelé irányuló (*run up*)
- utána egy *run down* látható
- annak a valószínűsége, hogy 20-25 megfigyelt érték közül egy run 8 hosszúságú legyen, nagyon kicsi, így valamilyen rendellenességre utalhat
- ugyanígy, ha 8 egymást követő pont az átlag egyik oldalán helyezkedik el, akkor is gyanúnk lehet arra, hogy a folyamat „kicsúszott” a kezünkből
- ciklikus mintázatok utalhatnak arra, hogy a nem véletlen lefutás valamilyen szisztematikus hatás miatt lehet, még akkor is, ha a pontok egyébként a határok között vannak

Western Electric Statistical Quality Control Handbook (1956) által meghatározott döntési szabályok a nem véletlen mintázat detektálására:

- 1 egy pont kívül esik a három-sigma határon
- 2 három egymást követő pont közül kettő kívül esik a két-sigma figyelmeztetési határon
- 3 öt egymást követő pont közül négy egy-sigma távolságban vagy azon túl helyezkedik el
- 4 nyolc egymást követő pont a középvonal egyik oldalán helyeződik el

A szabályok a középvonal egyik oldalára vonatkoznak. A szenzitivitást növeli a 2-4. pont az 1. ponthoz viszonyítva.

További érzékenyítő szabályok:

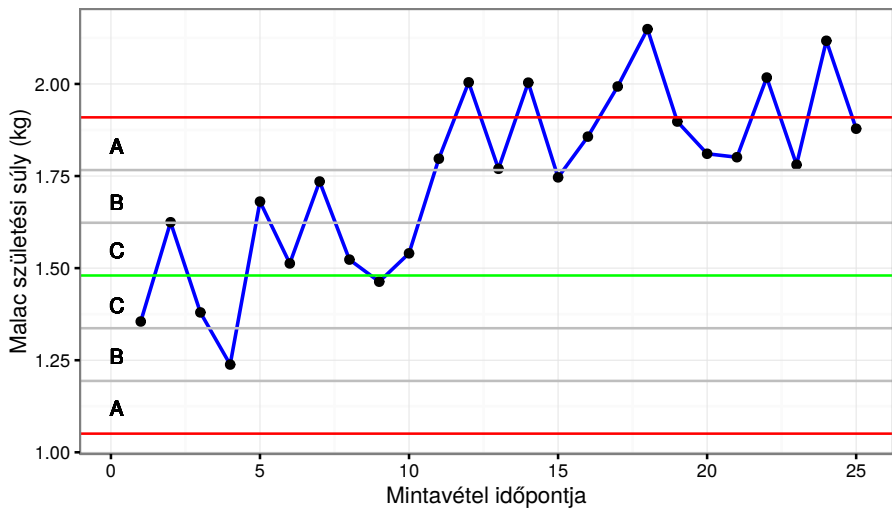
- 5 hat egymást követő pont folyamatosan növekvő vagy csökkenő értékekkel
- 6 tizenöt egymás utáni pont a C-zónában (a középvonal alatti és feletti)
- 7 tizennégy egymás utáni pont váltakozva le és fel
- 8 nyolc pont, ami egyik oldali C-zónába sem esik
- 9 szokatlan vagy nem véletlen mintázat az adatokban
- 10 egy vagy több pont közel a figyelmeztetési vagy cselekvési határhoz

- I. fázis:

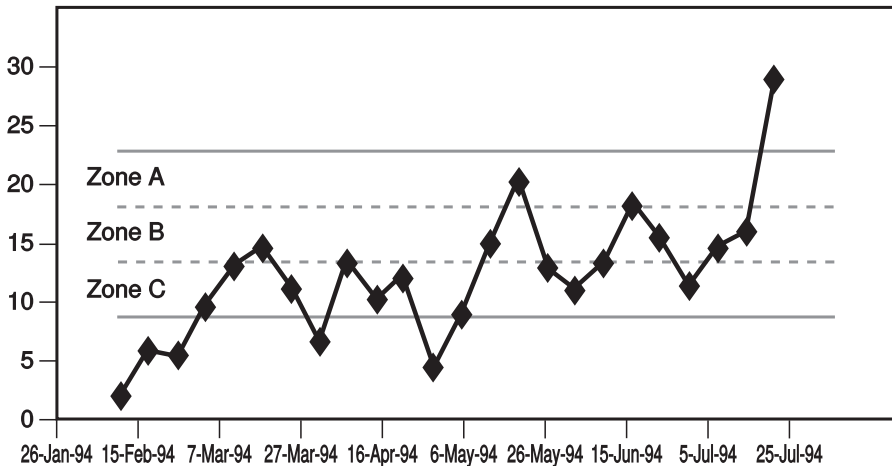
- retrospektív adatok vizsgálata
- stabil időszak, 20-25 időpontból, mintavételből származó minták
- célérték és határértékek számítása az alapnak tekintett időszak mintáiból
- határértéket meghaladó értékek ellenőrzése
- tisztáztuk, hogy mi lehet az oka
- ha egyértelműen tisztázódik (azonosítható okot találunk), kihagyhatjuk és újra számolhatjuk a diagram értékeit

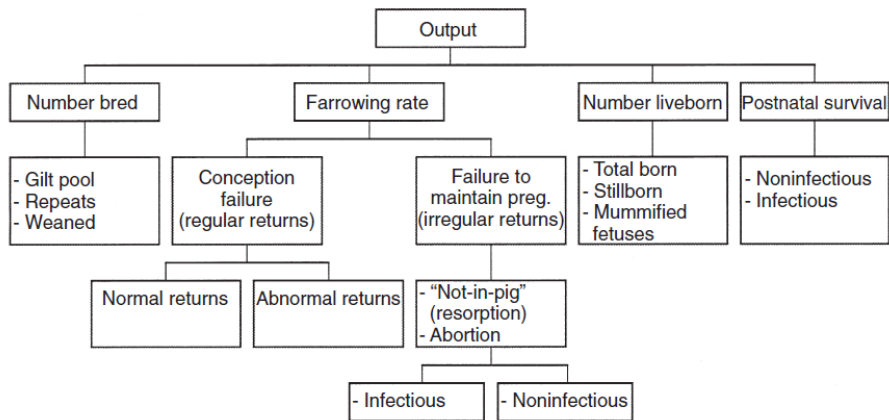
- II. fázis:

- monitoring
- folyamatos értékelés, minden új minta esetén
- ha szükséges, akkor a diagram értékeit módosítjuk



Kocák „visszaivarzásának” részaránya (%)





- gyakori, hogy a termelés valamely jellemzőjét számszerűleg mérjük (*numerical measurement*)
- pl.: testsúlyok, tejmennyiség, takarmány-, vízfogyasztás
- **változó**nak nevezzük a minőséget jellemző egyszerű mértéket
- ha a minőségi jellemző egy ilyen változó, akkor mind az **átlag**át, mind a **variabilitás**át be kell vonnunk az ellenőrzési folyamatba
- az **átlag**ot általában az ún \bar{x} -**diagram**mal ellenőrizzük
- a **variabilitást** a szórás felhasználásával, ekkor s - vagy az értékek terjedelmével, **R-diagram**mal ellenőrizzük
- a diagram értékeit a termelés stabil időszakában mért adatok alapján becsüljük
- a populáció valódi átlagának (μ) és varianciájának (σ^2) pontbecslése a mintából számított \bar{x} és s^2
- a σ^2 becslése
 - $n \leq 6$ esetén pontosabb az úgynevezett **range** módszerrel
 - $n \geq 10$ esetén a **szórás** alapján javasolható

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \cdots + \bar{x}_m}{m}$$

$$R = x_{max} - x_{min}$$

$$\bar{R} = \frac{R_1 + R_2 + \cdots + R_m}{m}$$

A relatív terjedelem (relative range):

$$W = \frac{R}{\sigma}$$

W az n mintaméret függvénye, W átlaga d_2 . Ebből következik σ becslése:

$$\hat{\sigma} = \frac{R}{d_2}$$

n	d_2
2	1.128
3	1.693
4	2.059
5	2.326
6	2.534
7	2.704
8	2.847
9	2.970
10	3.078
11	3.173
12	3.258
13	3.336
14	3.407
15	3.472
16	3.532
17	3.588
18	3.640
19	3.689
20	3.735
21	3.778
22	3.819
23	3.858
24	3.895
25	3.931

A diagram paramétereit a mintából becsülve:

$$CL = \bar{\bar{x}}$$

$$LCL = \bar{\bar{x}} - \frac{3}{d_2\sqrt{n}}\bar{R}$$

$$UCL = \bar{\bar{x}} + \frac{3}{d_2\sqrt{n}}\bar{R}$$

$$A_2 = \frac{3}{d_2\sqrt{n}}$$

$$LCL = \bar{\bar{x}} - A_2\bar{R}$$

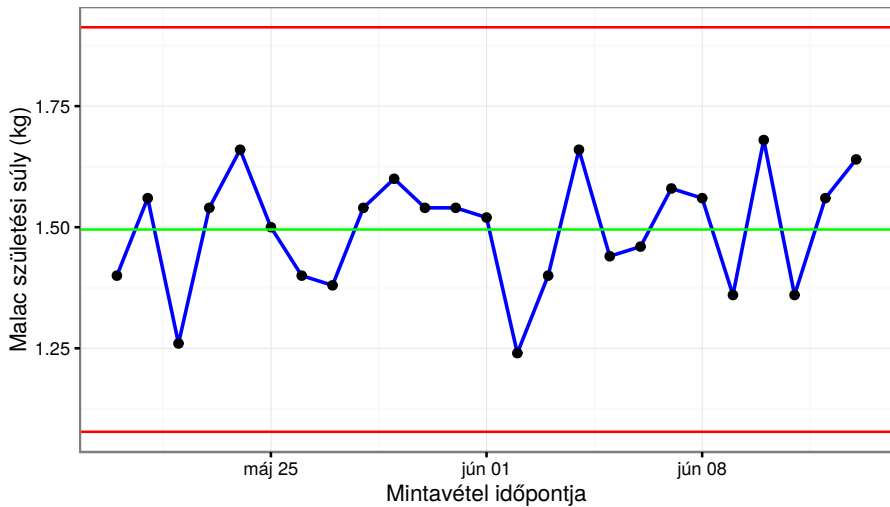
$$UCL = \bar{\bar{x}} + A_2\bar{R}$$

Ha $n > 25$:

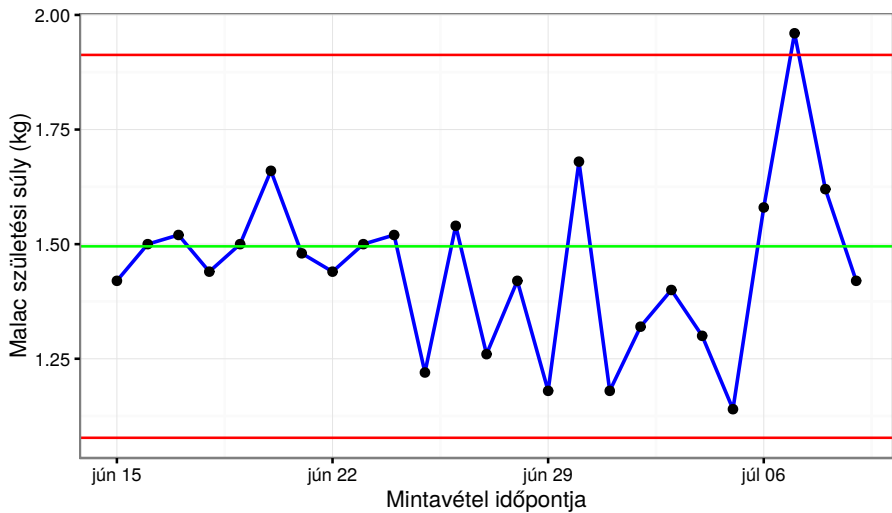
$$A = \frac{3}{\sqrt{n}} \quad A_3 = \frac{3}{c_4\sqrt{n}} \quad c_4 \cong \frac{4(n-1)}{4n-3}$$

n	A	A_2	A_3	c_4	$1/c_4$
2	2.121	1.880	2.659	0.7979	1.2533
3	1.732	1.023	1.954	0.8862	1.1284
4	1.500	0.729	1.628	0.9213	1.0854
5	1.342	0.577	1.427	0.9400	1.0638
6	1.225	0.483	1.287	0.9515	1.0510
7	1.134	0.419	1.182	0.9594	1.0423
8	1.061	0.373	1.099	0.9650	1.0363
9	1.000	0.337	1.032	0.9693	1.0317
10	0.949	0.308	0.975	0.9727	1.0281
11	0.905	0.285	0.927	0.9754	1.0252
12	0.866	0.266	0.886	0.9776	1.0229
13	0.832	0.249	0.850	0.9794	1.0210
14	0.802	0.235	0.817	0.9810	1.0194
15	0.775	0.223	0.789	0.9823	1.0180
16	0.750	0.212	0.763	0.9835	1.0168
17	0.728	0.203	0.739	0.9845	1.0157
18	0.707	0.194	0.718	0.9854	1.0148
19	0.688	0.187	0.698	0.9862	1.0140
20	0.671	0.180	0.680	0.9869	1.0133
21	0.655	0.173	0.663	0.9876	1.0126
22	0.640	0.167	0.647	0.9882	1.0119
23	0.626	0.162	0.633	0.9887	1.0114
24	0.612	0.157	0.619	0.9892	1.0109
25	0.600	0.153	0.606	0.9896	1.0105

I. fázis



II. fázis



A korábbiak alapján:

$$R = W \times \sigma$$

, illetve az R hibája:

$$\sigma_R = d_3 \times \sigma$$

mivel σ nem ismert így becsülhetjük:

$$\hat{\sigma}_R = d_3 \frac{\bar{R}}{d_2}$$

így az R -diagram paraméterei:

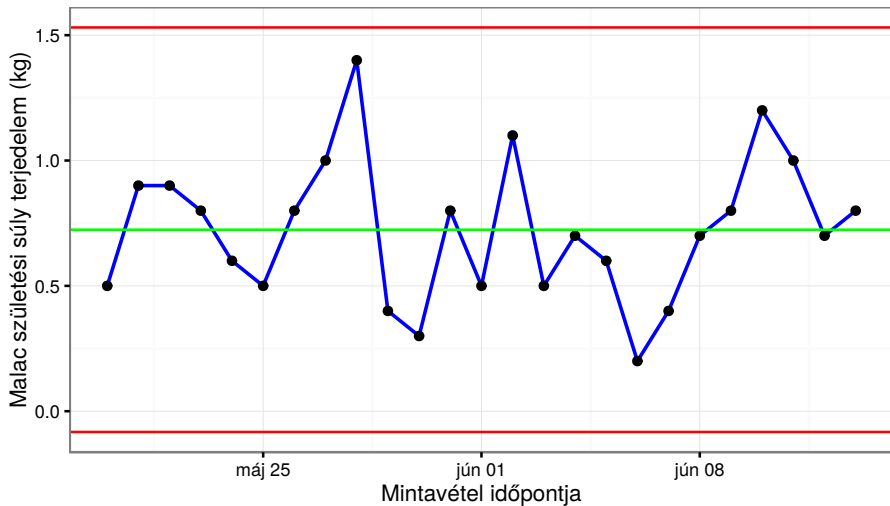
$$CL = \bar{R}$$

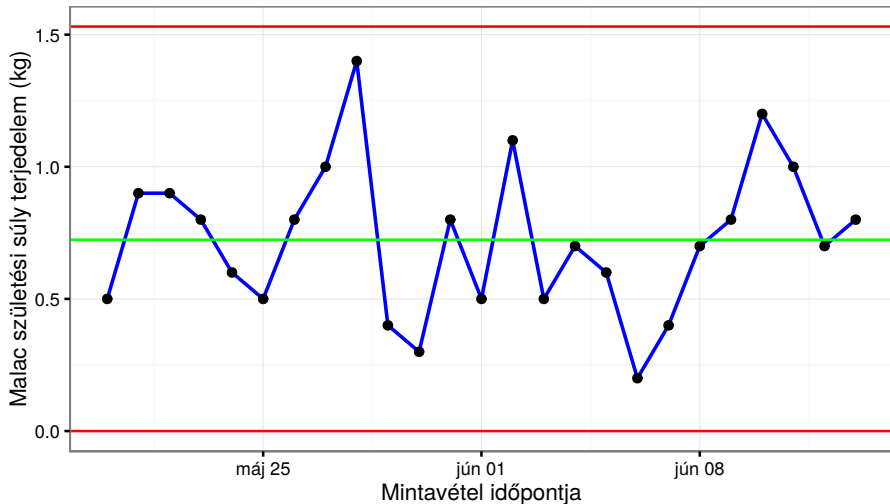
$$LCL = \bar{R} - 3\hat{\sigma}_R = \bar{R} - 3d_3 \frac{\bar{R}}{d_2} = D_3 \bar{R}$$

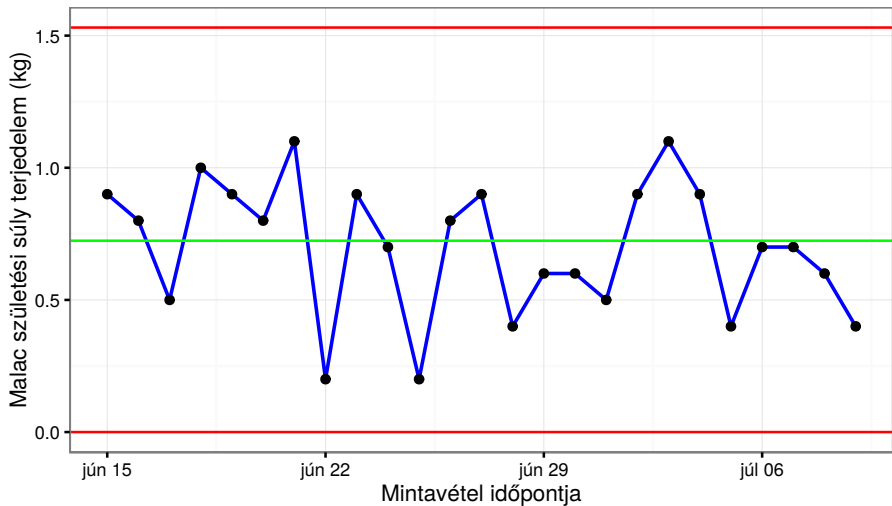
$$UCL = \bar{R} + 3\hat{\sigma}_R = \bar{R} + 3d_3 \frac{\bar{R}}{d_2} = D_4 \bar{R}$$

$$D_3 = 1 - 3 \frac{d_3}{d_2} \text{ és } D_4 = 1 + 3 \frac{d_3}{d_2}$$

n	d_3	D_1	D_2	D_3	D_4
2	0.853	0	3.686	0	3.267
3	0.888	0	4.358	0	2.574
4	0.880	0	4.698	0	2.282
5	0.864	0	4.918	0	2.114
6	0.848	0	5.078	0	2.004
7	0.833	0.204	5.204	0.076	1.924
8	0.820	0.388	5.306	0.136	1.864
9	0.808	0.547	5.393	0.184	1.816
10	0.797	0.687	5.469	0.223	1.777
11	0.787	0.811	5.535	0.256	1.744
12	0.778	0.922	5.594	0.283	1.717
13	0.770	1.025	5.647	0.307	1.693
14	0.763	1.118	5.696	0.328	1.672
15	0.756	1.203	5.741	0.347	1.653
16	0.750	1.282	5.782	0.363	1.637
17	0.744	1.356	5.820	0.378	1.622
18	0.739	1.424	5.856	0.391	1.608
19	0.734	1.487	5.891	0.403	1.597
20	0.729	1.549	5.921	0.415	1.585
21	0.724	1.605	5.951	0.425	1.575
22	0.720	1.659	5.979	0.434	1.566
23	0.716	1.710	6.006	0.443	1.557
24	0.712	1.759	6.031	0.451	1.548
25	0.708	1.806	6.056	0.459	1.541

I. fázis $(\bar{R} \pm 3d_3 \frac{\bar{R}}{d_2})$ 

I. fázis ($D_3\bar{R}$ és $D_4\bar{R}$)

II. fázis ($D_3\bar{R}$ és $D_4\bar{R}$)

- az \bar{x} -diagram a **minták közötti**, az R -diagram a **mintán belüli** variabilitást mutatja be
- az \bar{x} -diagram a közepes vagy nagy változások vizsgálatát teszi lehetővé az $n = 4, 5$ mintaméret, kisebb változások nagyobb $n = 15, \dots, 25$ mintákkal vizsgálhatók
- ennek alternatívája az érzékenyítő szabályok vagy CUSUM-, EWMA-diagramok alkalmazása
- a diagramok paramétereit **időről időre** felül kell vizsgálni, hogy az aktuális termelési folyamathoz igazodjon
- ezt meg lehet tenni hetente, havonta, vagy 25, 50, 100 mintánként
- az ellenőrző diagramokat kiegészíthetjük boxplotokkal (**tier chart**, tolerance chart) annak vizsgálatára, hogy bizonyos változásokat csak egy-egy mérési pont okoz-e vagy általános jelenségről van-e szó
- előfordul, hogy szakmai tapasztalatok alapján adnak meg felső és alsó határértékeket, ezeket ekkor nem control limitnek, hanem specification limitnek nevezik (nincs matematikai/statisztikai kapcsolat köztük)

- előfordul, hogy az n nem állandó, hanem változik mintáról mintára
- ilyen esetben nem szokták használni az \bar{x} - és R -diagramot
- olyan eset is előfordul, hogy a mintaméret tartósan megváltozik, ilyenkor az alábbi formulák segíthetnek:

$$CL = \bar{\bar{x}}$$

$$CL = \bar{R}_{new} = \left[\frac{d_{2new}}{d_{2old}} \right] \bar{R}_{old}$$

$$LCL = \bar{\bar{x}} - A_{2new} \left[\frac{d_{2new}}{d_{2old}} \right] \bar{R}_{old}$$

$$LCL = \max \left\{ 0, D_{3new} \left[\frac{d_{2new}}{d_{2old}} \right] \bar{R}_{old} \right\}$$

$$UCL = \bar{\bar{x}} + A_{2new} \left[\frac{d_{2new}}{d_{2old}} \right] \bar{R}_{old}$$

$$UCL = D_{4new} \left[\frac{d_{2new}}{d_{2old}} \right] \bar{R}_{old}$$

- habár az \bar{x} - és R -diagramokat széleskörben használják, előfordul, hogy az R helyett közvetlenül a folyamat szórásával vizsgálják a variabilitást
- az \bar{x} - és s -diagramot az alábbi esetekben érdemesebb az \bar{x} - és R -diagram helyett használni:
 - 1 ha a minta n nagysága nagyobb, mondjuk $n \geq 10$ vagy 12
 - 2 ha az n mintaméret változó
- ha az σ^2 variancia nem ismert, akkor a mintából becsülhetjük:

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

- érdemes megjegyezni, hogy s nem torzítatlan becslése σ -nak
- ha a vizsgált változó normális eloszlású, akkor s a $c_4\sigma$ becslése
- c_4 állandó, ami n -től függ
- ugyanakkor s hibája $\sigma\sqrt{1 - c_4^2}$

$$s = \sqrt{s^2}$$

$$CL = \bar{s} = \frac{1}{m} \sum_{i=1}^m s_i$$

$$B_3 = 1 - \frac{3}{c_4} \sqrt{1 - c_4^2}$$

$$B_4 = 1 + \frac{3}{c_4} \sqrt{1 - c_4^2}$$

$$LCL = \bar{s} - 3 \frac{\bar{s}}{c_4} \sqrt{1 - c_4^2} = B_3 \bar{s}$$

$$UCL = \bar{s} + 3 \frac{\bar{s}}{c_4} \sqrt{1 - c_4^2} = B_4 \bar{s}$$

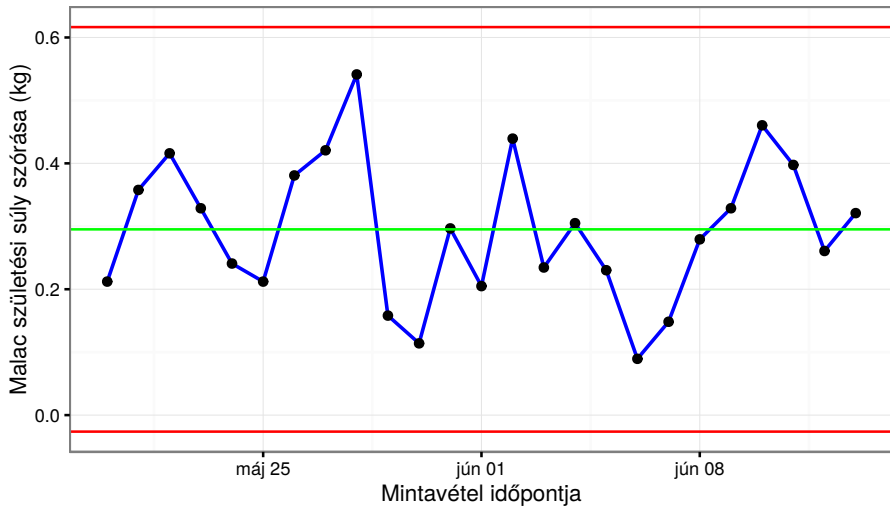
$$CL = \bar{\bar{x}}$$

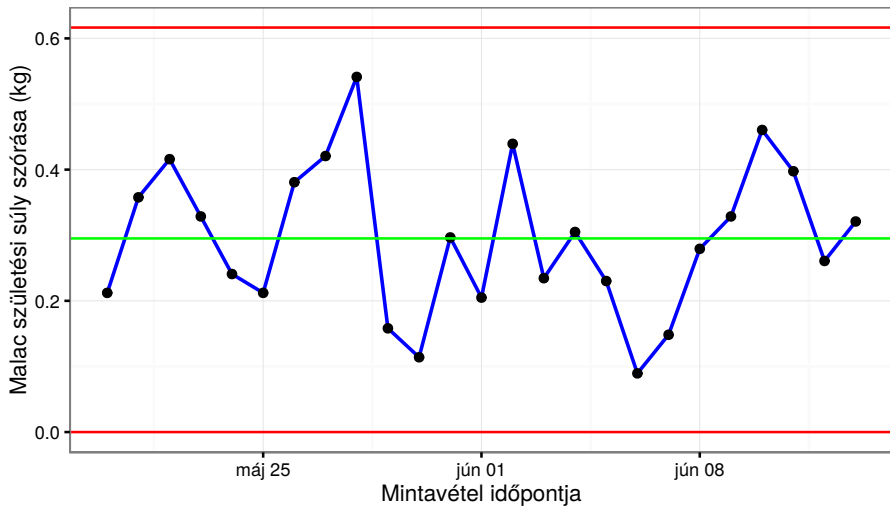
$$LCL = \bar{\bar{x}} - \frac{3\bar{s}}{c_4\sqrt{n}} = \bar{\bar{x}} - A_3\bar{s}$$

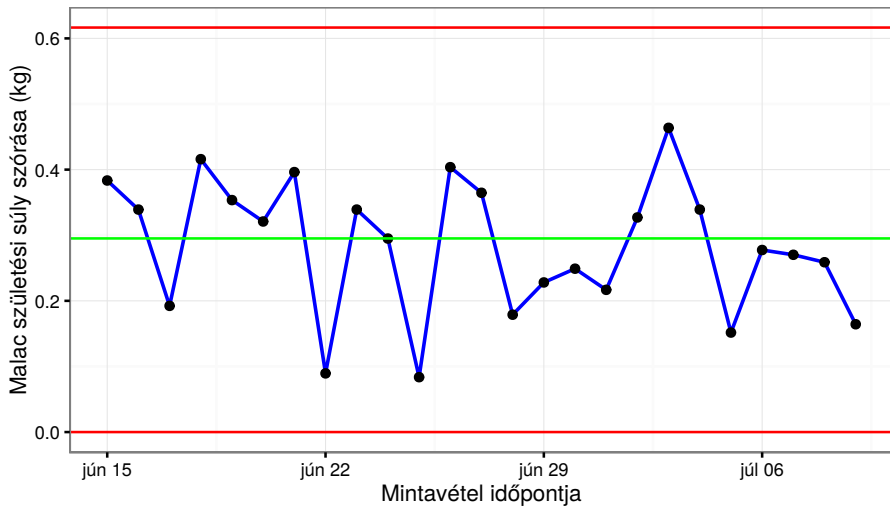
$$UCL = \bar{\bar{x}} + \frac{3\bar{s}}{c_4\sqrt{n}} = \bar{\bar{x}} + A_3\bar{s}$$

n	B_3	B_4	B_5	B_6
2	0	3.267	0	2.606
3	0	2.568	0	2.276
4	0	2.266	0	2.088
5	0	2.089	0	1.964
6	0.030	1.970	0.029	1.874
7	0.118	1.882	0.113	1.806
8	0.185	1.815	0.179	1.751
9	0.239	1.761	0.232	1.707
10	0.284	1.716	0.276	1.669
11	0.321	1.679	0.313	1.637
12	0.354	1.646	0.346	1.610
13	0.382	1.618	0.374	1.585
14	0.406	1.594	0.399	1.563
15	0.428	1.572	0.421	1.544
16	0.448	1.552	0.440	1.526
17	0.466	1.534	0.458	1.511
18	0.482	1.518	0.475	1.496
19	0.497	1.503	0.490	1.483
20	0.510	1.490	0.504	1.470
21	0.523	1.477	0.516	1.459
22	0.534	1.466	0.528	1.448
23	0.545	1.455	0.539	1.438
24	0.555	1.445	0.549	1.429
25	0.565	1.435	0.559	1.420

I. fázis $(\bar{s} \pm 3 \frac{\bar{s}}{c_4} \sqrt{1 - c_4^2})$



I. fázis ($B_3\bar{s}$ és $B_4\bar{s}$)

II. fázis ($B_3\bar{s}$ és $B_4\bar{s}$)

- változó n mintaméret esetén a súlyozott átlag módszerrel becsüljük $\bar{\bar{x}}$ és \bar{s} értékeit

$$\bar{\bar{x}} = \frac{\sum_{i=1}^m n_i \bar{x}_i}{\sum_{i=1}^m n_i}; \quad \bar{s} = \left[\frac{\sum_{i=1}^m (n_i - 1) s_i^2}{\sum_{i=1}^m n_i - m} \right]^{1/2}$$

- a középő egyenes ezekkel megrajzolható
- azonban a határértékek számításához a minták méretétől függően különböző A_3 , B_3 és B_4 konstans értékeket kell használnunk
- a változó szélességű ellenőrzési tartomány helyett használhatjuk az \bar{n} átlagos mintaméretet az ellenőrzési határok számítására
- ha a mintaméreten nem nagyon különbözőek, akkor ez jó közelítés lehet
- azonban mivel \bar{n} nem egész szám lesz, érdekesebb a mintaszámok móduszát használni

- vannak a termelési folyamatnak olyan mutatói, amelyek nem mérhetők numerikusan abban az értelemben, mint a fent említett mutatók esetén
- ezeket minősítéses adatokkal (*attributes data*) fejezzük ki
- ilyen pl. az, amikor a termelési folyamat eredményét jelentő elem lehet megfelelő (*conforming*) és nem megfelelő (*nonconforming*)
- természetesen ennek eldöntésében szerepelhetnek ún. méréses adatok, azonban a minősítésüket tekintve csak két kimeneti érték egyike rendelhető hozzájuk
- az ilyen adatokat eredményező termelési folyamatok statisztikai felügyeletére további ellenőrző diagramokat használnak:
 - *p*: nem megfelelő elemek **résaránya** (*control chart for fraction nonconforming*)
 - *c*: nem megfelelő elemek **száma** (*control chart for nonconformities*)
 - *u*: nem megfelelő elemek **egységenkénti száma** (*control chart for nonconformities per units*)

n méretű mintából D számú nem megfelelő elem esetén

$$\hat{p} = \frac{D}{n}$$

$$\mu_{\hat{p}} = p$$

$$\sigma_{\hat{p}}^2 = \frac{p \times (1 - p)}{n}$$

ha a termelési folyamatra vonatkozó p **ismert** vagy **standard** érték:

$$CL = p$$

$$LCL = p - 3 \times \sqrt{\frac{p \times (1 - p)}{n}}$$

$$UCL = p + 3 \times \sqrt{\frac{p \times (1 - p)}{n}}$$

ha nem ismert vagy standard érték, akkor becsüljük:

$$\hat{p}_i = \frac{D_i}{n} \quad i = 1, 2, \dots, m$$

m legalább 20-25 legyen
a \hat{p}_i értékek átlaga:

$$\bar{p} = \frac{\sum_{i=1}^m \hat{p}_i}{m} = \frac{\sum_{i=1}^m D_i}{m \times n}$$

$$CL = \bar{p}$$

$$LCL = \bar{p} - 3 \times \sqrt{\frac{\bar{p} \times (1 - \bar{p})}{n}}$$

$$UCL = \bar{p} + 3 \times \sqrt{\frac{\bar{p} \times (1 - \bar{p})}{n}}$$

- ha standard p -t használunk, akkor nincsen szükség a populációból való becslésére
- azonban óvatosan kell bánni a standard értékekkel, mivel nagyon ritka eset az, hogy valóban jól ismerhető a p
- a standard értéket inkább úgy kell értelmezni, mint egy célértéket
- pl. ha a standard (pl. $p = 0.01$) alapján meghatározott határértékeket meghaladják a II. fázis adatai, akkor nem eldönthető, hogy valóban kibillent-e a termelési folyamat, vagy csak más p jellemző rá (pl. $p = 0.05$)
- a p -diagram létrehozásához meg kell határozni a mintanagyságot, a mintavételi gyakoriságot, az ellenőrzési határokat
- előfordul, hogy a folyamat során, bevett periódusban képződő **összes** eredményt jelentő elemét megvizsgálják
- azonban általában az a cél, hogy csak n **mintát** használjunk

- ha p nagyon kicsi, akkor megfelelően nagy n szükséges
- ha n kicsi, akkor kisebb valószínűséggel azonosíthatunk „nem megfelelő elemet” a mintában
- pl. ha $p = 0.01$ és $n = 8$, akkor a felső ellenőrzési határérték:

$$UCL = p + 3 \times \sqrt{\frac{p \times (1 - p)}{n}} = 0.01 + 3 \times \sqrt{\frac{0.01 \times (1 - 0.01)}{8}} = 0.1155$$

- ha a mintában van egy nem megfelelő elem, akkor $\hat{p} = 1/8 = 0.125$, vagyis azt gondolhatjuk, hogy a folyamat kikerült a kezünk közül
- mivel valamely $p > 0$ esetén van valamekkora valószínűsége annak, hogy a termelési folyamatban „tökéletlen termék” jön létre, ezért nem szükségeszerű egyetlen ilyen „nem megfelelő elem” észlelése során arra következtetnünk, hogy a folyamat kikerült a kezünk közül
- mekkora n -t kell használnunk

1. megközelítés:

- tegyük fel, hogy $p = 0.01$
- mekkora n szükséges ahhoz, hogy a mintában lévő „nem megfelelő elemet” legalább 95%-os valószínűséggel megtaláljunk?
- ha D a „nem megfelelő elemek” száma a mintában, akkor keressük azt az n -t, aminél $P\{D \geq 1\} \geq 0.95$, vagy máshogy megfogalmazva $P\{D = 0\} = 0.05$:

$$P\{D = x\} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$$P\{D = 0\} = \frac{n!}{0!(n-0)!} 0.01^0 (1-0.01)^{n-0}$$

$$0.05 = (1-0.01)^n$$

$$n = \frac{\log(0.05)}{\log(1-0.01)} \simeq 298$$

Mean percent repeat services: 6.51%¹

$$n = \frac{\log(0.05)}{\log(1 - 0.0651)} \simeq 45$$

$$CL = p = 0.0651$$

$$LCL = 0.0651 - 3 \times \sqrt{\frac{0.0651 \times (1 - 0.0651)}{45}} = -0.0452$$

NB: $LCL < 0!!!$

$$UCL = 0.0651 + 3 \times \sqrt{\frac{0.0651 \times (1 - 0.0651)}{45}} = 0.1754$$

¹USA 2016- 2nd Quarter Summary

2. megközelítés

- ha alacsony a p , akkor könnyen lehet, hogy az LCL értéke kisebb lesz, mint 0, ami valószínűségek esetén nem értelmezhető
- ha azt szeretnénk, hogy az LCL is pozitív érték legyen, akkor az n meghatározása a következők szerint történhet:

$$LCL = p - L \times \sqrt{\frac{p \times (1 - p)}{n}} > 0$$

, amiből levezethető, hogy

$$n > \frac{1 - p}{p} \times L^2$$

, ha a $p = 0.0651$

$$n > \frac{1 - 0.0651}{0.0651} \times 3^2 \cong 129$$

Így ha $n \geq 130$, akkor az LCL pozitív értékű lesz.

3. megközelítés:

- a mintaszámnak elég nagynek kell lennie ahhoz, hogy kb. 50%-os valószínűséggel észlelhető legyen a folyamat egy meghatározott mértékű eltolódása
- pl. tegyük fel, hogy $p = 0.01$ és szeretnénk tudni azt az n mintaméretet, ami a $p = 0.05$ szintre való eltolódását 50%-os valószínűséggel lehetővé teszi
- feltételezve, hogy a binomiális eloszlás normális eloszlással való közelítése helytálló, és ha folyamatban detektálandó eltolódás δ :

$$\delta = L \times \sqrt{\frac{p \times (1 - p)}{n}}$$

, amiből

$$n = \left(\frac{L}{\delta}\right)^2 \times p \times (1 - p) = \left(\frac{3}{0.05 - 0.01}\right)^2 \times 0.01 \times 0.99 \cong 56$$

- *Mean percent repeat services: 6.51%*¹
- *Median percent repeat services: 5.56%*¹
- *Upper 10th percentile percent repeat services: 13.17%*¹
- mekkora n mintaméret szükséges, hogy a 13.17%-ra való eltolódást 50%-os valószínűséggel detektálhassuk?

$$n = \left(\frac{3}{0.1317 - 0.0651} \right)^2 \times 0.0651 \times (1 - 0.0651) \cong 123$$

$$n = \left(\frac{3}{0.1317 - 0.0556} \right)^2 \times 0.0556 \times (1 - 0.0556) \cong 82$$

¹USA 2016- 2nd Quarter Summary

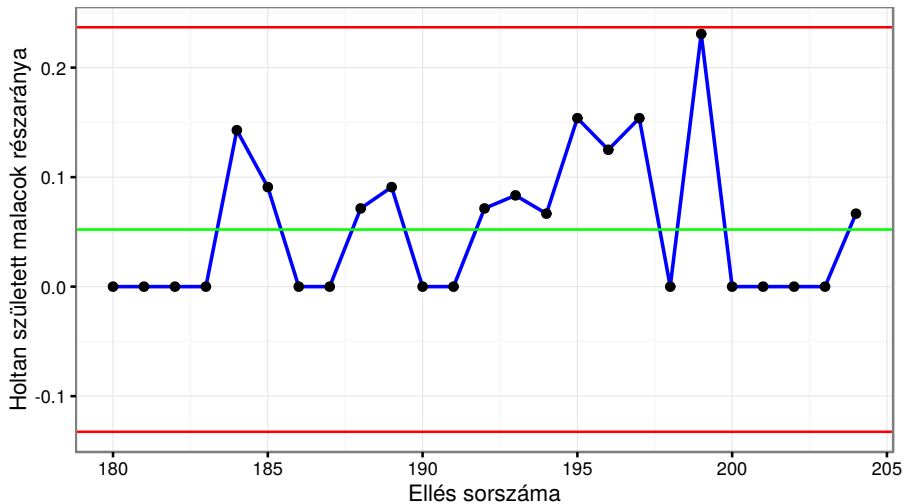
Nem állandó n mintaméret

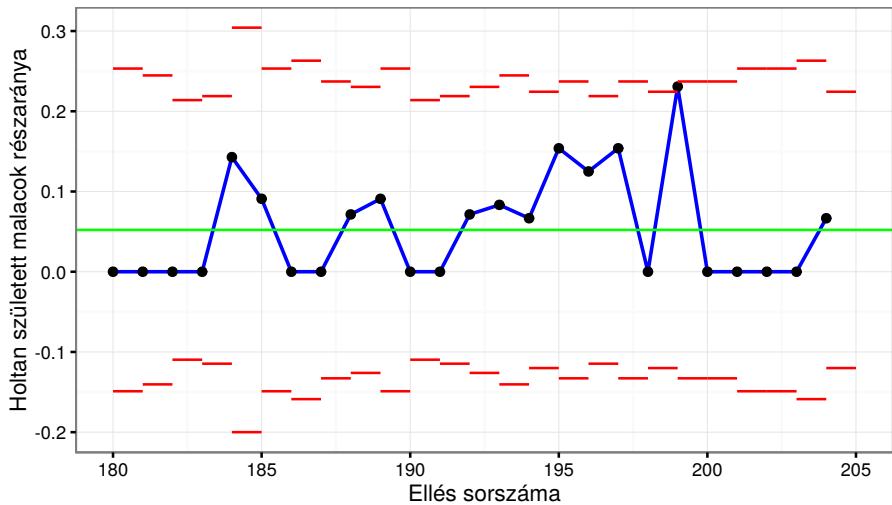
- gyakori, hogy nem mintát vesznek, hanem egy időben, időszakban képződő **összes** elemet bevonnak a vizsgálatba
- a haszonállattartásban ez a leginkább jellemző (pl. alomméret)
- a korábbiak alapján látható, hogy a változó mintaméret (n) befolyásolja az **ellenőrzési tartományokat**:

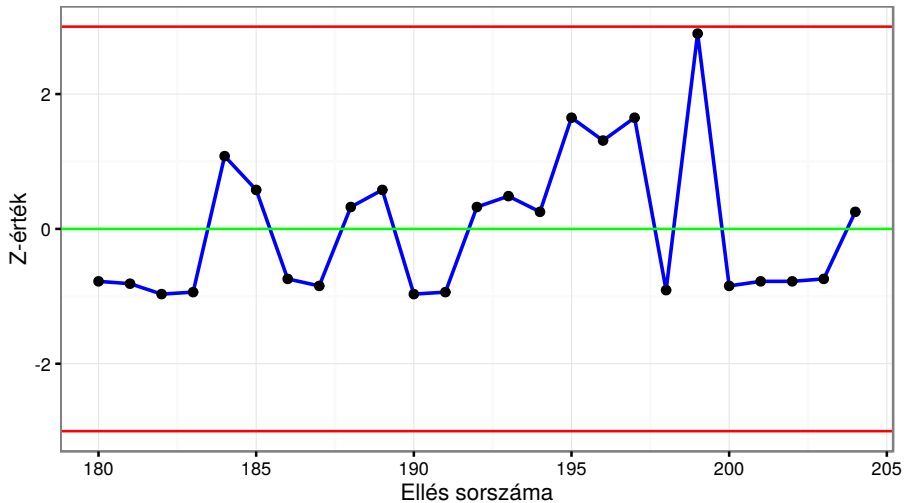
$$\bar{p} \pm 3 \times \sqrt{\frac{\bar{p} \times (1 - \bar{p})}{n}}$$

- a probléma kezelésére gyakoribb megoldások:
 - 1 átlagos mintaméret (\bar{n})
 - 2 mindegyik mintaméretre külön ellenőrzési határok (n_i)
 - 3 standardizált határok:

$$Z_i = \frac{\hat{p}_i - p}{\sigma_{\hat{p}}} = \frac{\hat{p}_i - p}{\sqrt{\frac{p \times (1 - p)}{n_i}}}$$

\bar{n} 

n_i 

Z_i 

- arra is lehetőség van, hogy a nem megfelelő adatokat ne részarányként fejezzük ki, hanem azok előfordulási gyakoriságával
- erre használják az np -diagramot

$$CL = n \times p$$

$$LCL = n \times p - 3 \times \sqrt{n \times p \times (1 - p)}$$

$$UCL = n \times p + 3 \times \sqrt{n \times p \times (1 - p)}$$

- ha p -re vonatkozóan nem elérhető standard érték, akkor becsülhetjük a mintából (\bar{p})
- az np -diagram sokszor jobban értelmezhető a statisztikában nem jártas felhasználók számára

- a c -diagram esetén a „nem megfelelő elemek” számát használjuk az ellenőrzési ábra létrehozásához
- pl. lehetne valamely betegség incidenciája, **azonos méretű** állatcsoportokban (időben, térben)

$$p(x) = \frac{e^{-c} c^x}{x!}$$

, ahol x a „nem megfelelő elemek” száma, $c > 0$ a Poisson-eloszlás paramétere

$$CL = c$$

$$LCL = c - 3 \times \sqrt{c}$$

$$UCL = c + 3 \times \sqrt{c}$$

- ha az $LCL < 0$, akkor nullára állítjuk be
- ha nem ismert a standard c , akkor \bar{c} -t becsülhetjük a mintából

- ha x a megfigyelt összes „nem megfelelő elem” száma, a megvizsgált n egységben
- pl. $n = 6$ kutrica, amelyekben azonos számú sertés van

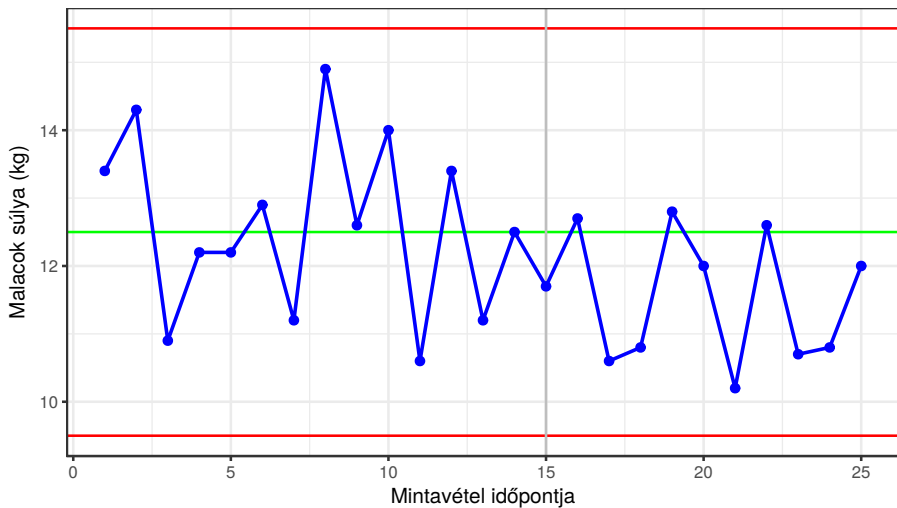
$$u = \frac{x}{n}$$

- x Poisson-eloszlású változó

$$CL = \bar{u}$$

$$LCL = \bar{u} - 3 \times \sqrt{\frac{\bar{u}}{n}}$$

$$UCL = \bar{u} + 3 \times \sqrt{\frac{\bar{u}}{n}}$$



Hathetes malacok súlyának átlaga 12.5 kg, $\sigma = 1$.

A Shewhart-diagramok megfelelő érzékenységűek, ha a folyamat-eltérés $1.5\text{-}2\sigma$. Ennél kisebb eltolódást csak az érzékenyítő szabályokkal detektálhatunk, amelyek viszont növelik a téves riasztások számát.

A célértéktől való eltérés **halmozódó összegének** (*cumulative sum*, CUSUM) számítása, ha $n \geq 1$:

$$C_i = \sum_{j=1}^i (\bar{x}_j - \mu_0)$$

Egyedi x -ek ($n = 1$) esetén:

$$\begin{aligned} C_i &= \sum_{j=1}^i (x_j - \mu_0) \\ &= (x_i - \mu_0) + \sum_{j=1}^{i-1} (x_j - \mu_0) \\ &= (x_i - \mu_0) + C_{i-1} \end{aligned}$$

i	x_i	C_i
1	13.4	
2	14.3	
3	10.9	
4	12.2	
5	12.2	
6	12.9	
7	11.2	
8	14.9	
9	12.6	
10	14.0	
11	10.6	
12	13.4	
\vdots	\vdots	\vdots

Mivel $C_{1-1} = C_0 = 0$

$$\begin{aligned}C_1 &= (x_1 - \mu_0) + C_0 \\&= (13.4 - 12.5) + 0 \\&= 0.9\end{aligned}$$

i	x_i	C_i
1	13.4	0.9
2	14.3	
3	10.9	
4	12.2	
5	12.2	
6	12.9	
7	11.2	
8	14.9	
9	12.6	
10	14.0	
11	10.6	
12	13.4	
\vdots	\vdots	\vdots

Mivel $C_{2-1} = C_1 = 0.9$

$$\begin{aligned} C_2 &= (x_2 - \mu_0) + C_1 \\ &= (14.3 - 12.5) + 0.9 \\ &= 2.7 \end{aligned}$$

i	x_i	C_i
1	13.4	0.9
2	14.3	2.7
3	10.9	
4	12.2	
5	12.2	
6	12.9	
7	11.2	
8	14.9	
9	12.6	
10	14.0	
11	10.6	
12	13.4	
\vdots	\vdots	\vdots

Mivel $C_{3-1} = C_2 = 2.7$

$$\begin{aligned}C_3 &= (x_3 - \mu_0) + C_2 \\&= (10.9 - 12.5) + 2.7 \\&= 1.1\end{aligned}$$

i	x_i	C_i
1	13.4	0.9
2	14.3	2.7
3	10.9	1.1
4	12.2	
5	12.2	
6	12.9	
7	11.2	
8	14.9	
9	12.6	
10	14.0	
11	10.6	
12	13.4	
\vdots	\vdots	\vdots

Mivel $C_{4-1} = C_3 = 1.1$

$$\begin{aligned}
 C_4 &= (x_4 - \mu_0) + C_3 \\
 &= (12.2 - 12.5) + 1.1 \\
 &= 0.8
 \end{aligned}$$

i	x_i	C_i
1	13.4	0.9
2	14.3	2.7
3	10.9	1.1
4	12.2	0.8
5	12.2	
6	12.9	
7	11.2	
8	14.9	
9	12.6	
10	14.0	
11	10.6	
12	13.4	
\vdots	\vdots	\vdots

Mivel $C_{5-1} = C_4 = 0.8$

$$\begin{aligned}C_5 &= (x_5 - \mu_0) + C_4 \\&= (12.2 - 12.5) + 0.8 \\&= 0.5\end{aligned}$$

i	x_i	C_i
1	13.4	0.9
2	14.3	2.7
3	10.9	1.1
4	12.2	0.8
5	12.2	0.5
6	12.9	
7	11.2	
8	14.9	
9	12.6	
10	14.0	
11	10.6	
12	13.4	
\vdots	\vdots	\vdots

Mivel $C_{6-1} = C_5 = 0.5$

$$\begin{aligned}
 C_6 &= (x_6 - \mu_0) + C_5 \\
 &= (12.9 - 12.5) + 0.5 \\
 &= 0.9
 \end{aligned}$$

i	x_i	C_i
1	13.4	0.9
2	14.3	2.7
3	10.9	1.1
4	12.2	0.8
5	12.2	0.5
6	12.9	0.9
7	11.2	
8	14.9	
9	12.6	
10	14.0	
11	10.6	
12	13.4	
\vdots	\vdots	\vdots

Mivel $C_{7-1} = C_6 = 0.9$

$$\begin{aligned}C_7 &= (x_7 - \mu_0) + C_6 \\&= (11.2 - 12.5) + 0.9 \\&= -0.4\end{aligned}$$

i	x_i	C_i
1	13.4	0.9
2	14.3	2.7
3	10.9	1.1
4	12.2	0.8
5	12.2	0.5
6	12.9	0.9
7	11.2	-0.4
8	14.9	
9	12.6	
10	14.0	
11	10.6	
12	13.4	
\vdots	\vdots	\vdots

Mivel $C_{8-1} = C_7 = -0.4$

$$\begin{aligned}C_8 &= (x_8 - \mu_0) + C_7 \\&= (14.9 - 12.5) - 0.4 \\&= 2.0\end{aligned}$$

i	x_i	C_i
1	13.4	0.9
2	14.3	2.7
3	10.9	1.1
4	12.2	0.8
5	12.2	0.5
6	12.9	0.9
7	11.2	-0.4
8	14.9	2.0
9	12.6	
10	14.0	
11	10.6	
12	13.4	
\vdots	\vdots	\vdots

Mivel $C_{9-1} = C_8 = 2.0$

$$\begin{aligned}C_9 &= (x_9 - \mu_0) + C_8 \\&= (12.6 - 12.5) + 2.0 \\&= 2.1\end{aligned}$$

i	x_i	C_i
1	13.4	0.9
2	14.3	2.7
3	10.9	1.1
4	12.2	0.8
5	12.2	0.5
6	12.9	0.9
7	11.2	-0.4
8	14.9	2.0
9	12.6	2.1
10	14.0	
11	10.6	
12	13.4	
\vdots	\vdots	\vdots

Mivel $C_{10-1} = C_9 = 2.1$

$$\begin{aligned}
 C_{10} &= (x_{10} - \mu_0) + C_9 \\
 &= (14.0 - 12.5) + 2.1 \\
 &= 3.6
 \end{aligned}$$

i	x_i	C_i
1	13.4	0.9
2	14.3	2.7
3	10.9	1.1
4	12.2	0.8
5	12.2	0.5
6	12.9	0.9
7	11.2	-0.4
8	14.9	2.0
9	12.6	2.1
10	14.0	3.6
11	10.6	
12	13.4	
\vdots	\vdots	\vdots

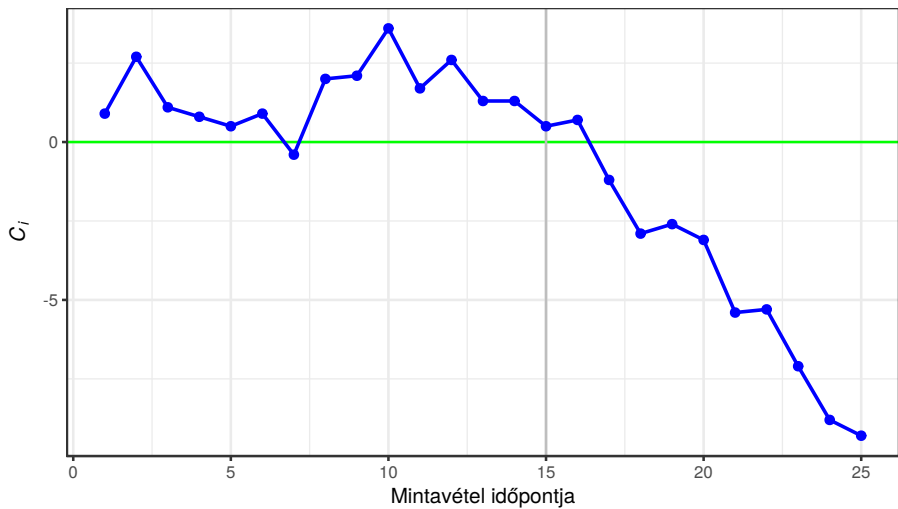
Mivel $C_{11-1} = C_{10} = 3.6$

$$\begin{aligned}
 C_{11} &= (x_{11} - \mu_0) + C_{10} \\
 &= (10.6 - 12.5) + 3.6 \\
 &= 1.7
 \end{aligned}$$

i	x_i	C_i
1	13.4	0.9
2	14.3	2.7
3	10.9	1.1
4	12.2	0.8
5	12.2	0.5
6	12.9	0.9
7	11.2	-0.4
8	14.9	2.0
9	12.6	2.1
10	14.0	3.6
11	10.6	1.7
12	13.4	
\vdots	\vdots	\vdots

Mivel $C_{12-1} = C_{11} = 1.7$

$$\begin{aligned}
 C_{12} &= (x_{12} - \mu_0) + C_{11} \\
 &= (13.4 - 12.5) + 1.7 \\
 &= 2.6
 \end{aligned}$$



Az egyoldali felső és alsó CUSUM számítása:

$$C_i^+ = \max \left[0, x_i - (\mu_0 + K) + C_{i-1}^+ \right]$$

$$C_i^- = \max \left[0, (\mu_0 - K) - x_i + C_{i-1}^- \right]$$

A kezdő értékek $C_0^+ = C_0^- = 0$.

A K referencia-érték, mely a μ_0 és az ellenőrzés alól kikerült folyamat μ_1 közti különbség fele.

$$K = \frac{\delta}{2} \sigma = \frac{|\mu_1 - \mu_0|}{2}$$

Mivel a példánkban a $\sigma = 1$ és egy σ eltérést szeretnénk detektálni ($\delta = 1$), így $K = 0.5$.

A CUSUM módszernél az ellenőrzési tartományokat ún. H -értékkel állítjuk be, aminek értéke első közelítésben a σ **ötszöröse**.

i	x_i	C_i	C_i^+	C_i^-
1	13.4	0.9		
2	14.3	2.7		
3	10.9	1.1		
4	12.2	0.8		
5	12.2	0.5		
6	12.9	0.9		
7	11.2	-0.4		
8	14.9	2.0		
9	12.6	2.1		
10	14.0	3.6		
11	10.6	1.7		
12	13.4	2.6		
\vdots	\vdots	\vdots		

Mivel $C_{1-1}^+ = C_0^+ = 0$

$$\begin{aligned}
 C_1^+ &= \max [0, x_1 - (\mu_0 + K) + C_0^+] \\
 &= \max [0, 13.4 - (12.5 + 0.5) + 0] \\
 &= \max [0, 0.4] \\
 &= 0.4
 \end{aligned}$$

Mivel $C_{1-1}^- = C_0^- = 0$

$$\begin{aligned}
 C_1^- &= \max [0, (\mu_0 + K) - x_1 + C_0^-] \\
 &= \max [0, (12.5 - 0.5) - 13.4 + 0] \\
 &= \max [0, -1.4] \\
 &= 0.0
 \end{aligned}$$

i	x_i	C_i	C_i^+	C_i^-
1	13.4	0.9	0.4	0.0
2	14.3	2.7		
3	10.9	1.1		
4	12.2	0.8		
5	12.2	0.5		
6	12.9	0.9		
7	11.2	-0.4		
8	14.9	2.0		
9	12.6	2.1		
10	14.0	3.6		
11	10.6	1.7		
12	13.4	2.6		
\vdots	\vdots	\vdots		

Mivel $C_{2-1}^+ = C_1^+ = 0.4$

$$\begin{aligned}
 C_2^+ &= \max [0, x_2 - (\mu_0 + K) + C_1^+] \\
 &= \max [0, 14.3 - (12.5 + 0.5) + 0.4] \\
 &= \max [0, 1.7] \\
 &= 1.7
 \end{aligned}$$

Mivel $C_{2-1}^- = C_1^- = 0.0$

$$\begin{aligned}
 C_2^- &= \max [0, (\mu_0 + K) - x_2 + C_1^-] \\
 &= \max [0, (12.5 - 0.5) - 14.3 + 0.0] \\
 &= \max [0, -2.3] \\
 &= 0.0
 \end{aligned}$$

i	x_i	C_i	C_i^+	C_i^-
1	13.4	0.9	0.4	0.0
2	14.3	2.7	1.7	0.0
3	10.9	1.1		
4	12.2	0.8		
5	12.2	0.5		
6	12.9	0.9		
7	11.2	-0.4		
8	14.9	2.0		
9	12.6	2.1		
10	14.0	3.6		
11	10.6	1.7		
12	13.4	2.6		
\vdots	\vdots	\vdots		

Mivel $C_{3-1}^+ = C_2^+ = 1.7$

$$\begin{aligned}
 C_3^+ &= \max [0, x_3 - (\mu_0 + K) + C_2^+] \\
 &= \max [0, 10.9 - (12.5 + 0.5) + 1.7] \\
 &= \max [0, -0.4] \\
 &= 0.0
 \end{aligned}$$

Mivel $C_{3-1}^- = C_2^- = 0.0$

$$\begin{aligned}
 C_3^- &= \max [0, (\mu_0 + K) - x_3 + C_2^-] \\
 &= \max [0, (12.5 - 0.5) - 10.9 + 0.0] \\
 &= \max [0, 1.1] \\
 &= 1.1
 \end{aligned}$$

i	x_i	C_i	C_i^+	C_i^-
1	13.4	0.9	0.4	0.0
2	14.3	2.7	1.7	0.0
3	10.9	1.1	0.0	1.1
4	12.2	0.8		
5	12.2	0.5		
6	12.9	0.9		
7	11.2	-0.4		
8	14.9	2.0		
9	12.6	2.1		
10	14.0	3.6		
11	10.6	1.7		
12	13.4	2.6		
\vdots	\vdots	\vdots		

Mivel $C_{4-1}^+ = C_3^+ = 0.0$

$$\begin{aligned}
 C_4^+ &= \max [0, x_4 - (\mu_0 + K) + C_3^+] \\
 &= \max [0, 12.2 - (12.5 + 0.5) + 0.0] \\
 &= \max [0, -0.8] \\
 &= 0.0
 \end{aligned}$$

Mivel $C_{4-1}^- = C_3^- = 1.1$

$$\begin{aligned}
 C_4^- &= \max [0, (\mu_0 + K) - x_4 + C_3^-] \\
 &= \max [0, (12.5 - 0.5) - 12.2 + 1.1] \\
 &= \max [0, 0.9] \\
 &= 0.9
 \end{aligned}$$

i	x_i	C_i	C_i^+	C_i^-
1	13.4	0.9	0.4	0.0
2	14.3	2.7	1.7	0.0
3	10.9	1.1	0.0	1.1
4	12.2	0.8	0.0	0.9
5	12.2	0.5		
6	12.9	0.9		
7	11.2	-0.4		
8	14.9	2.0		
9	12.6	2.1		
10	14.0	3.6		
11	10.6	1.7		
12	13.4	2.6		
\vdots	\vdots	\vdots		

Mivel $C_{5-1}^+ = C_4^+ = 0.0$

$$\begin{aligned}
 C_5^+ &= \max [0, x_5 - (\mu_0 + K) + C_4^+] \\
 &= \max [0, 12.2 - (12.5 + 0.5) + 0.0] \\
 &= \max [0, -0.8] \\
 &= 0.0
 \end{aligned}$$

Mivel $C_{5-1}^- = C_4^- = 0.9$

$$\begin{aligned}
 C_5^- &= \max [0, (\mu_0 + K) - x_5 + C_4^-] \\
 &= \max [0, (12.5 - 0.5) - 12.2 + 0.9] \\
 &= \max [0, 0.7] \\
 &= 0.7
 \end{aligned}$$

i	x_i	C_i	C_i^+	C_i^-
1	13.4	0.9	0.4	0.0
2	14.3	2.7	1.7	0.0
3	10.9	1.1	0.0	1.1
4	12.2	0.8	0.0	0.9
5	12.2	0.5	0.0	0.7
6	12.9	0.9		
7	11.2	-0.4		
8	14.9	2.0		
9	12.6	2.1		
10	14.0	3.6		
11	10.6	1.7		
12	13.4	2.6		
\vdots	\vdots	\vdots		

Mivel $C_{6-1}^+ = C_5^+ = 0.0$

$$\begin{aligned}
 C_6^+ &= \max [0, x_6 - (\mu_0 + K) + C_5^+] \\
 &= \max [0, 12.9 - (12.5 + 0.5) + 0.0] \\
 &= \max [0, -0.1] \\
 &= 0.0
 \end{aligned}$$

Mivel $C_{6-1}^- = C_5^- = 0.7$

$$\begin{aligned}
 C_6^- &= \max [0, (\mu_0 + K) - x_6 + C_5^-] \\
 &= \max [0, (12.5 - 0.5) - 12.9 + 0.7] \\
 &= \max [0, -0.2] \\
 &= 0.0
 \end{aligned}$$

i	x_i	C_i	C_i^+	C_i^-
1	13.4	0.9	0.4	0.0
2	14.3	2.7	1.7	0.0
3	10.9	1.1	0.0	1.1
4	12.2	0.8	0.0	0.9
5	12.2	0.5	0.0	0.7
6	12.9	0.9	0.0	0.0
7	11.2	-0.4		
8	14.9	2.0		
9	12.6	2.1		
10	14.0	3.6		
11	10.6	1.7		
12	13.4	2.6		
\vdots	\vdots	\vdots		

Mivel $C_{7-1}^+ = C_6^+ = 0.0$

$$\begin{aligned}
 C_7^+ &= \max [0, x_7 - (\mu_0 + K) + C_6^+] \\
 &= \max [0, 11.2 - (12.5 + 0.5) + 0.0] \\
 &= \max [0, -1.8] \\
 &= 0.0
 \end{aligned}$$

Mivel $C_{7-1}^- = C_6^- = 0.0$

$$\begin{aligned}
 C_7^- &= \max [0, (\mu_0 + K) - x_7 + C_6^-] \\
 &= \max [0, (12.5 - 0.5) - 11.2 + 0.0] \\
 &= \max [0, 0.8] \\
 &= 0.8
 \end{aligned}$$

i	x_i	C_i	C_i^+	C_i^-
1	13.4	0.9	0.4	0.0
2	14.3	2.7	1.7	0.0
3	10.9	1.1	0.0	1.1
4	12.2	0.8	0.0	0.9
5	12.2	0.5	0.0	0.7
6	12.9	0.9	0.0	0.0
7	11.2	-0.4	0.0	0.8
8	14.9	2.0		
9	12.6	2.1		
10	14.0	3.6		
11	10.6	1.7		
12	13.4	2.6		
\vdots	\vdots	\vdots		

Mivel $C_{8-1}^+ = C_7^+ = 0.0$

$$\begin{aligned}
 C_8^+ &= \max [0, x_8 - (\mu_0 + K) + C_7^+] \\
 &= \max [0, 14.9 - (12.5 + 0.5) + 0.0] \\
 &= \max [0, 1.9] \\
 &= 1.9
 \end{aligned}$$

Mivel $C_{8-1}^- = C_7^- = 0.8$

$$\begin{aligned}
 C_8^- &= \max [0, (\mu_0 + K) - x_8 + C_7^-] \\
 &= \max [0, (12.5 - 0.5) - 14.9 + 0.8] \\
 &= \max [0, -2.1] \\
 &= 0.0
 \end{aligned}$$

i	x_i	C_i	C_i^+	C_i^-
1	13.4	0.9	0.4	0.0
2	14.3	2.7	1.7	0.0
3	10.9	1.1	0.0	1.1
4	12.2	0.8	0.0	0.9
5	12.2	0.5	0.0	0.7
6	12.9	0.9	0.0	0.0
7	11.2	-0.4	0.0	0.8
8	14.9	2.0	1.9	0.0
9	12.6	2.1		
10	14.0	3.6		
11	10.6	1.7		
12	13.4	2.6		
\vdots	\vdots	\vdots		

Mivel $C_{9-1}^+ = C_8^+ = 1.9$

$$\begin{aligned}
 C_9^+ &= \max [0, x_9 - (\mu_0 + K) + C_8^+] \\
 &= \max [0, 12.6 - (12.5 + 0.5) + 1.9] \\
 &= \max [0, 1.5] \\
 &= 1.5
 \end{aligned}$$

Mivel $C_{9-1}^- = C_8^- = 0.0$

$$\begin{aligned}
 C_9^- &= \max [0, (\mu_0 + K) - x_9 + C_8^-] \\
 &= \max [0, (12.5 - 0.5) - 12.6 + 0.0] \\
 &= \max [0, -0.6] \\
 &= 0.0
 \end{aligned}$$

i	x_i	C_i	C_i^+	C_i^-
1	13.4	0.9	0.4	0.0
2	14.3	2.7	1.7	0.0
3	10.9	1.1	0.0	1.1
4	12.2	0.8	0.0	0.9
5	12.2	0.5	0.0	0.7
6	12.9	0.9	0.0	0.0
7	11.2	-0.4	0.0	0.8
8	14.9	2.0	1.9	0.0
9	12.6	2.1	1.5	0.0
10	14.0	3.6		
11	10.6	1.7		
12	13.4	2.6		
\vdots	\vdots	\vdots		

Mivel $C_{10-1}^+ = C_9^+ = 1.5$

$$\begin{aligned}
 C_{10}^+ &= \max [0, x_{10} - (\mu_0 + K) + C_9^+] \\
 &= \max [0, 14.0 - (12.5 + 0.5) + 1.5] \\
 &= \max [0, 2.5] \\
 &= 2.5
 \end{aligned}$$

Mivel $C_{10-1}^- = C_9^- = 0.0$

$$\begin{aligned}
 C_{10}^- &= \max [0, (\mu_0 + K) - x_{10} + C_9^-] \\
 &= \max [0, (12.5 - 0.5) - 14.0 + 0.0] \\
 &= \max [0, -2.0] \\
 &= 0.0
 \end{aligned}$$

i	x_i	C_i	C_i^+	C_i^-
1	13.4	0.9	0.4	0.0
2	14.3	2.7	1.7	0.0
3	10.9	1.1	0.0	1.1
4	12.2	0.8	0.0	0.9
5	12.2	0.5	0.0	0.7
6	12.9	0.9	0.0	0.0
7	11.2	-0.4	0.0	0.8
8	14.9	2.0	1.9	0.0
9	12.6	2.1	1.5	0.0
10	14.0	3.6	2.5	0.0
11	10.6	1.7		
12	13.4	2.6		
\vdots	\vdots	\vdots		

Mivel $C_{11-1}^+ = C_{10}^+ = 2.5$

$$\begin{aligned}
 C_{11}^+ &= \max [0, x_{11} - (\mu_0 + K) + C_{10}^+] \\
 &= \max [0, 10.6 - (12.5 + 0.5) + 2.5] \\
 &= \max [0, 0.1] \\
 &= 0.1
 \end{aligned}$$

Mivel $C_{11-1}^- = C_{10}^- = 0.0$

$$\begin{aligned}
 C_{11}^- &= \max [0, (\mu_0 + K) - x_{11} + C_{10}^-] \\
 &= \max [0, (12.5 - 0.5) - 10.6 + 0.0] \\
 &= \max [0, 1.4] \\
 &= 1.4
 \end{aligned}$$

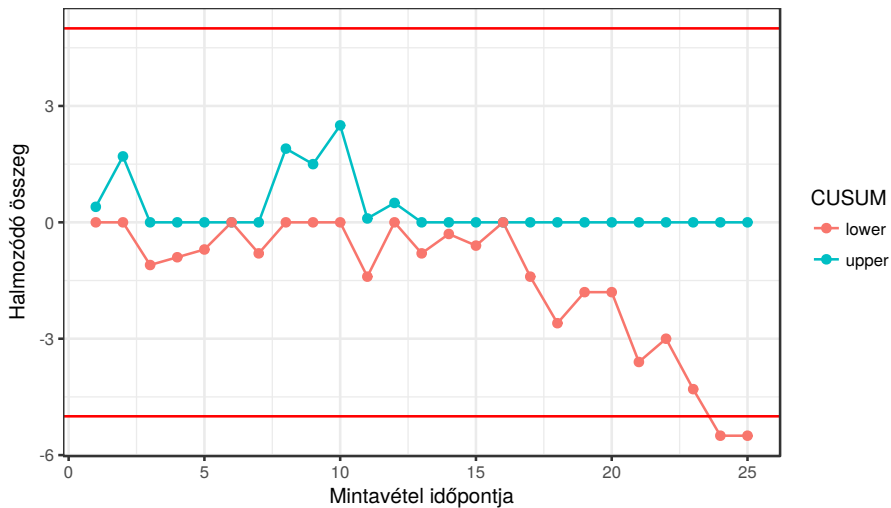
i	x_i	C_i	C_i^+	C_i^-
1	13.4	0.9	0.4	0.0
2	14.3	2.7	1.7	0.0
3	10.9	1.1	0.0	1.1
4	12.2	0.8	0.0	0.9
5	12.2	0.5	0.0	0.7
6	12.9	0.9	0.0	0.0
7	11.2	-0.4	0.0	0.8
8	14.9	2.0	1.9	0.0
9	12.6	2.1	1.5	0.0
10	14.0	3.6	2.5	0.0
11	10.6	1.7	0.1	1.4
12	13.4	2.6		
\vdots	\vdots	\vdots		

Mivel $C_{12-1}^+ = C_{11}^+ = 0.1$

$$\begin{aligned}
 C_{12}^+ &= \max [0, x_{12} - (\mu_0 + K) + C_{11}^+] \\
 &= \max [0, 13.4 - (12.5 + 0.5) + 0.1] \\
 &= \max [0, 0.5] \\
 &= 0.5
 \end{aligned}$$

Mivel $C_{12-1}^- = C_{11}^- = 1.4$

$$\begin{aligned}
 C_{12}^- &= \max [0, (\mu_0 + K) - x_{12} + C_{11}^-] \\
 &= \max [0, (12.5 - 0.5) - 13.4 + 1.4] \\
 &= \max [0, 0] \\
 &= 0
 \end{aligned}$$



Ha n állandó:

$$C_0^+ = H^+$$

$$C_0^- = H^-$$

$$C_i^+ = \max[0, C_{i-1}^+ + x_i - n \times k_u^+]$$

$$C_i^- = \min[0, C_{i-1}^- + x_i - n \times k_u^-]$$

Ha n változó, akkor ún. súlyozott binomiális CUSUM:

$$C_i^+ = \max[0, C_{i-1}^+ + x_i - n_i \times k_u^+]$$

$$C_i^- = \min[0, C_{i-1}^- + x_i - n_i \times k_u^-]$$

, ahol az emelkedő és csökkenő k^+ és k^- referenciaértékek egyaránt a $n \times k_u$ szorzatból adódnak, ahol a k_u az *egységenkénti* referenciaérték, ami csak a stabil folyamat π_0 és az ellenőrzés alól kikerült folyamat π_1 értékétől függ:

$$k_u = -\frac{\ln\left(\frac{1-\pi_1}{1-\pi_0}\right)}{\ln\left(\frac{\pi_1(1-\pi_0)}{\pi_0(1-\pi_1)}\right)}$$

Tegyük fel, hogy egy olyan folyamatot ellenőrzünk, amelyben a $\pi_0 = 0.01$ és ettől fölfelé a $\pi_1 = 0.03$ elcsúszást szeretnénk detektálni CUSUM-mal:

$$k_u^+ = -\frac{\ln\left(\frac{1-0.03}{1-0.01}\right)}{\ln\left(\frac{0.03 \times (1-0.01)}{0.01 \times (1-0.03)}\right)} = -\frac{\ln\left(\frac{1-0.03}{1-0.01}\right)}{\ln\left(\frac{0.03 \times 0.99}{0.01 \times 0.97}\right)} = -\frac{-0.02040885}{1.119021} = 0.018$$

Ugyanezen az ellenőrző diagramon a negatív irányba való eltérés vizsgálatára $\pi_1 = 0.005$ szeretnénk használni:

$$k_u^- = -\frac{\ln\left(\frac{1-0.005}{1-0.01}\right)}{\ln\left(\frac{0.005 \times (1-0.01)}{0.01 \times (1-0.005)}\right)} = -\frac{\ln\left(\frac{0.995}{0.99}\right)}{\ln\left(\frac{0.005 \times 0.99}{0.01 \times 0.97}\right)} = -\frac{0.005037794}{-0.698185} = 0.0072$$

$$C_i^+ = \max[0, C_{i-1}^+ + x_i - n \times 0.018]$$

$$C_i^- = \min[0, C_{i-1}^- + x_i - n \times 0.0072]$$

Exponenciálisan súlyozott mozgóátlag (*Exponentially Weighted Moving Average*, EWMA)

$$z_i = \lambda x_i + (1 - \lambda) \times z_{i-1}$$

, ahol $0 < \lambda \leq 1$, és a $z_0 = \mu_0$, illetve korábbi mintából becsülve $z_0 = \bar{x}$.

Az a tapasztalat, hogy a $0.05 \leq \lambda \leq 0.25$ jól használható; gyakran használt értékek a $\lambda = 0.05$, $\lambda = 0.10$ és $\lambda = 0.20$.

A kisebb λ -értékek kisebb eltérések detektálására jobban használhatók. Az ellenőrző diagramhoz szükséges formulák:

$$CL = \mu_0$$

$$LCL = \mu_0 - L \times \sigma \times \sqrt{\frac{\lambda}{2 - \lambda} \times [1 - (1 - \lambda)^{2i}]}$$

$$UCL = \mu_0 + L \times \sigma \times \sqrt{\frac{\lambda}{2 - \lambda} \times [1 - (1 - \lambda)^{2i}]}$$

, ahol L az ellenőrzési tartomány szélességét meghatározó érték. Nagyobb λ esetén az $L = 3$ megbízható érték. Azonban kisebb λ esetén, pl.

$\lambda \leq 0.1$, tanácsos csökkenteni az értékét (pl. $L = 2.7$).

i	x_i	z_i	LCL_i	UCL_i
1	13.4			
2	14.3			
3	10.9			
4	12.2			
5	12.2			
6	12.9			
\vdots	\vdots	\vdots	\vdots	\vdots

$$\begin{aligned}
 z_1 &= \lambda x_1 + (1 - \lambda)z_0 \\
 &= 0.1 * 13.4 + (1 - 0.1) * 12.5 \\
 &= 12.59
 \end{aligned}$$

$$\begin{aligned}
 h_1 &= L \times \sigma \times \sqrt{\frac{\lambda}{2 - \lambda} \times [1 - (1 - \lambda)^{2 \times 1}]} \\
 &= 2.7 \times 1 \times \sqrt{\frac{0.1}{2 - 0.1} \times [1 - (1 - 0.1)^{2 \times 1}]} \\
 &= 0.27
 \end{aligned}$$

$$LCL_1 = 12.5 - 0.27 = 12.23$$

$$UCL_1 = 12.5 + 0.27 = 12.77$$

$$\lambda = 0.1, L = 2.7, z_0 = \mu_0 = 12.5, \sigma = 1$$

i	x_i	z_i	LCL_i	UCL_i
1	13.4	12.59	12.23	12.77
2	14.3			
3	10.9			
4	12.2			
5	12.2			
6	12.9			
\vdots	\vdots	\vdots	\vdots	\vdots

$$\begin{aligned}
 z_2 &= \lambda x_2 + (1 - \lambda)z_1 \\
 &= 0.1 * 14.3 + (1 - 0.1) * 12.59 \\
 &= 12.76
 \end{aligned}$$

$$\begin{aligned}
 l_2 &= L \times \sigma \times \sqrt{\frac{\lambda}{2 - \lambda} \times [1 - (1 - \lambda)^{2 \times 2}]} \\
 &= 2.7 \times 1 \times \sqrt{\frac{0.1}{2 - 0.1} \times [1 - (1 - 0.1)^{2 \times 2}]} \\
 &= 0.36
 \end{aligned}$$

$$LCL_2 = 12.5 - 0.36 = 12.14$$

$$UCL_2 = 12.5 + 0.36 = 12.86$$

$$\lambda = 0.1, L = 2.7, z_0 = \mu_0 = 12.5, \sigma = 1$$

i	x_i	z_i	LCL_i	UCL_i
1	13.4	12.59	12.23	12.77
2	14.3	12.76	12.14	12.86
3	10.9			
4	12.2			
5	12.2			
6	12.9			
\vdots	\vdots	\vdots	\vdots	\vdots

$$\begin{aligned}
 z_3 &= \lambda x_3 + (1 - \lambda)z_2 \\
 &= 0.1 * 10.9 + (1 - 0.1) * 12.76 \\
 &= 12.57
 \end{aligned}$$

$$\begin{aligned}
 l_3 &= L \times \sigma \times \sqrt{\frac{\lambda}{2 - \lambda} \times [1 - (1 - \lambda)^{2 \times 3}]} \\
 &= 2.7 \times 1 \times \sqrt{\frac{0.1}{2 - 0.1} \times [1 - (1 - 0.1)^{2 \times 3}]} \\
 &= 0.42
 \end{aligned}$$

$$LCL_3 = 12.5 - 0.42 = 12.08$$

$$UCL_3 = 12.5 + 0.42 = 12.92$$

$$\lambda = 0.1, L = 2.7, z_0 = \mu_0 = 12.5, \sigma = 1$$

i	x_i	z_i	LCL_i	UCL_i
1	13.4	12.59	12.23	12.77
2	14.3	12.76	12.14	12.86
3	10.9	12.57	12.08	12.92
4	12.2			
5	12.2			
6	12.9			
\vdots	\vdots	\vdots	\vdots	\vdots

$$\begin{aligned}
 z_4 &= \lambda x_4 + (1 - \lambda)z_3 \\
 &= 0.1 * 12.2 + (1 - 0.1) * 12.57 \\
 &= 12.54
 \end{aligned}$$

$$\begin{aligned}
 l_4 &= L \times \sigma \times \sqrt{\frac{\lambda}{2 - \lambda} \times [1 - (1 - \lambda)^{2 \times 4}]} \\
 &= 2.7 \times 1 \times \sqrt{\frac{0.1}{2 - 0.1} \times [1 - (1 - 0.1)^{2 \times 4}]} \\
 &= 0.47
 \end{aligned}$$

$$LCL_4 = 12.5 - 0.47 = 12.03$$

$$UCL_4 = 12.5 + 0.47 = 12.97$$

$$\lambda = 0.1, L = 2.7, z_0 = \mu_0 = 12.5, \sigma = 1$$

i	x_i	z_i	LCL_i	UCL_i
1	13.4	12.59	12.23	12.77
2	14.3	12.76	12.14	12.86
3	10.9	12.57	12.08	12.92
4	12.2	12.54	12.03	12.97
5	12.2			
6	12.9			
\vdots	\vdots	\vdots	\vdots	\vdots

$$\begin{aligned}
 z_5 &= \lambda x_5 + (1 - \lambda)z_4 \\
 &= 0.1 * 12.2 + (1 - 0.1) * 12.54 \\
 &= 12.50
 \end{aligned}$$

$$\begin{aligned}
 l_5 &= L \times \sigma \times \sqrt{\frac{\lambda}{2 - \lambda} \times [1 - (1 - \lambda)^{2 \times 5}]} \\
 &= 2.7 \times 1 \times \sqrt{\frac{0.1}{2 - 0.1} \times [1 - (1 - 0.1)^{2 \times 5}]} \\
 &= 0.47
 \end{aligned}$$

$$LCL_5 = 12.5 - 0.50 = 12.00$$

$$UCL_5 = 12.5 + 0.50 = 13.00$$

$$\lambda = 0.1, L = 2.7, z_0 = \mu_0 = 12.5, \sigma = 1$$

i	x_i	z_i	LCL_i	UCL_i
1	13.4	12.59	12.23	12.77
2	14.3	12.76	12.14	12.86
3	10.9	12.57	12.08	12.92
4	12.2	12.54	12.03	12.97
5	12.2	12.50	12.00	13.00
6	12.9			
\vdots	\vdots	\vdots	\vdots	\vdots

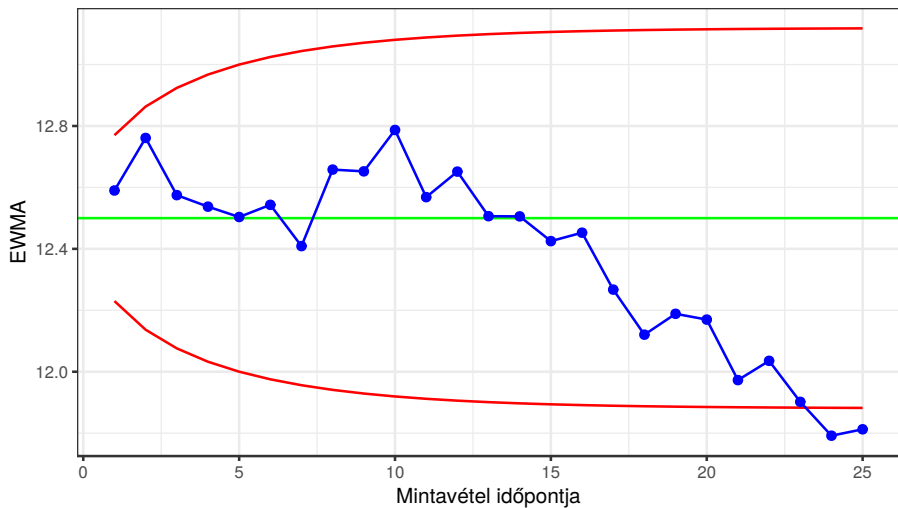
$$\begin{aligned}
 z_6 &= \lambda x_6 + (1 - \lambda)z_5 \\
 &= 0.1 * 12.2 + (1 - 0.1) * 12.54 \\
 &= 12.50
 \end{aligned}$$

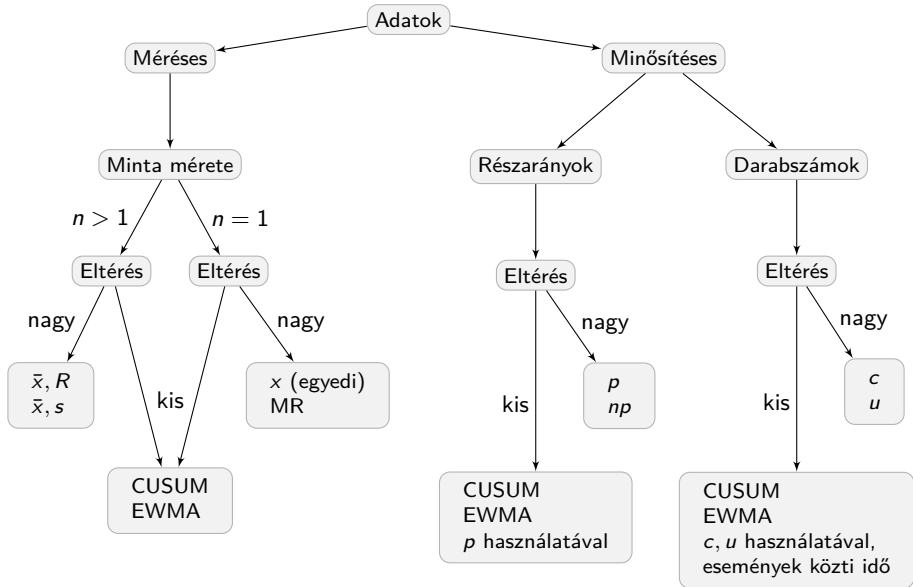
$$\begin{aligned}
 l_6 &= L \times \sigma \times \sqrt{\frac{\lambda}{2 - \lambda} \times [1 - (1 - \lambda)^{2 \times 6}]} \\
 &= 2.7 \times 1 \times \sqrt{\frac{0.1}{2 - 0.1} \times [1 - (1 - 0.1)^{2 \times 6}]} \\
 &= 0.47
 \end{aligned}$$

$$LCL_6 = 12.5 - 0.50 = 12.00$$

$$UCL_6 = 12.5 + 0.50 = 13.00$$

$$\lambda = 0.1, L = 2.7, z_0 = \mu_0 = 12.5, \sigma = 1$$





Átlagos sorozathossz (*average run length*, **ARL**): azoknak a pontoknak a száma, amelyek megelőzik azt a pontot, amikor a folyamat kikerül az ellenőrzés alól. Ha nincsen autokorreláció a folyamatban, akkor:

$$ARL = \frac{1}{p}$$

, ahol p annak a valószínűsége, hogy a folyamat kikerül az ellenőrzés alól. *Average time to signal* (**ATS**), ha a mintavételek között eltelt idő (h) mindig azonos: $ATS = ARL \times h$.

A három szigma \bar{x} -diagram esetén annak a valószínűsége, hogy egy pont kívülesik az ellenőrzési tartományon $p = 0.0027$, így az ARL_0 (a *kézben tartott folyamat ARL*-je):

$$ARL_0 = \frac{1}{\alpha} = \frac{1}{0.0027} = 370$$

Vagyis átlagosan minden 370. pont fog kívülesni az ellenőrzési tartományon, még abban az esetben is, ha a folyamat valójában nem tért el a céltól. Az az ideális, ha az ARL_0 nagy.

- II. típusú hiba: ha a H_0 -t nem utasítjuk el, pedig nem igaz. Ennek valószínűségét a β fejezi ki.
- Annak a valószínűsége, hogy NEM detektáljuk a folyamat eltérését az eltérés bekövetkezte utáni első mintában:

$$\beta = P\{LCL \leq \bar{x} \leq UCL | \mu = \mu_1 = \mu_0 + k\sigma\}$$

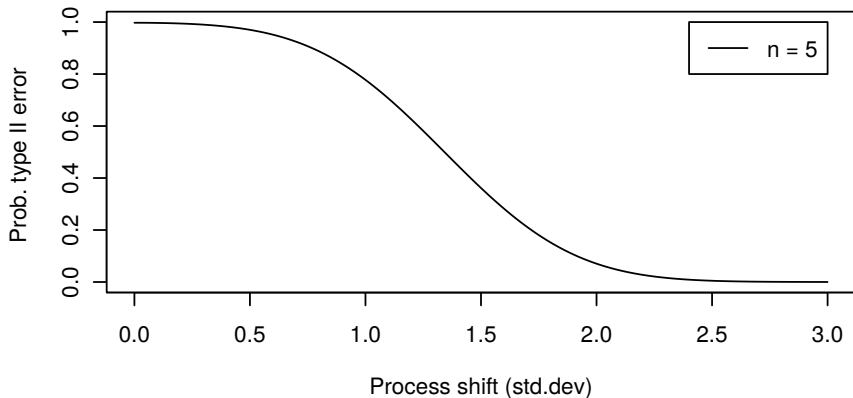
, ahol μ_0 a kézben tartott folyamat célértéke, $\mu_1 = \mu_0 + k\sigma$ az ellenőrzés alól kikerült folyamat átlaga.

- Ha a folyamatbeli elcsúszás megtörtént, annak a valószínűsége, hogy:
 - NEM detektáljuk az eltérést az *első* mintán: β
 - detektáljuk az eltérést az *első* mintán: $1 - \beta$
- Out-of-control ARL: amikor a folyamat kikerül az ellenőrzés alól, az eltérés detektálása előtti pontok átlagos száma:

$$ARL_1 = \frac{1}{1 - \beta}$$

Ideális esetben kicsi.

OC curves for xbar chart



Operating-characteristic curve: az y -tengelyen a β (annak a valószínűsége, hogy az ellenőrző diagram NEM jelez eltérést a folyamatban), az x -tengelyen a k , vagyis a stabil folyamat átlagától való eltérés mértéke σ szorzatként kifejezve.

Ha a malacok 1.495 kg-os átlag testsúlya, mondjuk 1.5 szigmányival alacsonyabb szintre ($1.495 - (1.5 * 0.31) = 1.03$ kg) süllyed, akkor β 0.36-ra csökken.

$$ARL_1 = \frac{1}{1 - \beta} = \frac{1}{1 - 0.36} = 1.56$$

Ami azt jelenti, hogy átlagosan 1.56 minta szükséges ahhoz, hogy a folyamatbeli ekkora eltérést detektálhassuk.

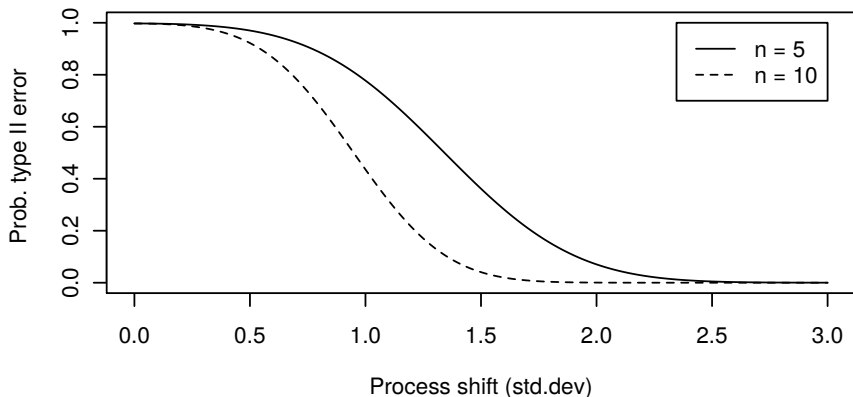
Ha mondjuk a malacsúly-mintavételünk hetente történik, vagyis hetente mérjük a malacokat, akkor:

$$ATS = ARL_1 \times h = 1.56 * 7 = 10.92$$

Azaz 11 nap szükséges ahhoz, hogy az eltérést észleljük. Túl késő:

- 1 Ha gyakrabban veszünk mintát, akkor csökkenthetjük a h -t és így hamarabb észlelhetjük az eltérést.

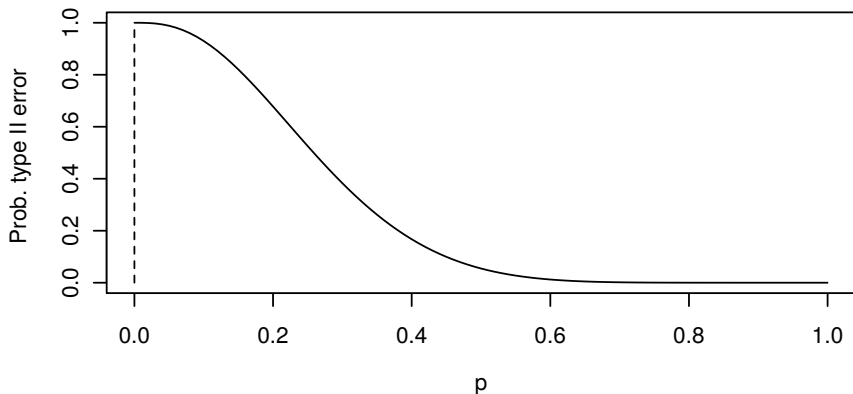
OC curves for xbar chart



- 2 Ha növeljük az n mintaméretet, akkor is csökkenthetjük a detektáláshoz szükséges időt, pl. ha $n = 10$:

$$ARL_1 = \frac{1}{1 - \beta} = \frac{1}{1 - 0.041} = 1.04$$

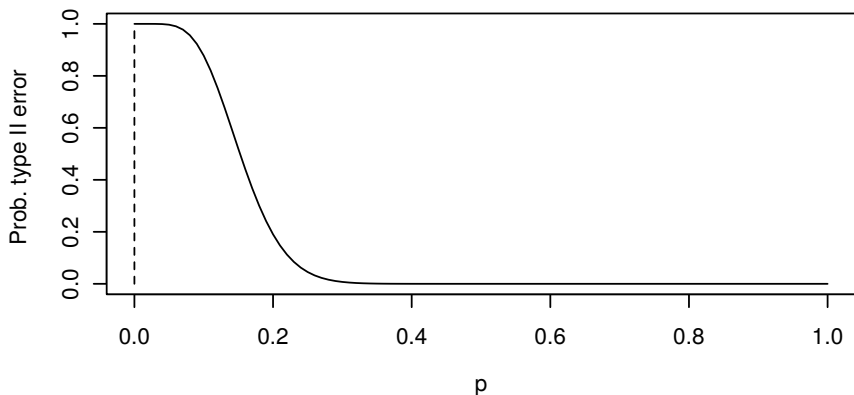
OC curves for p chart



A választásig történő elhullás valószínűsége a korábbi példából 5%. Ha $n = 10$, a 10%-os elhullás esetén a $\beta = 0.93$, így:

$$ARL_1 = \frac{1}{1 - \beta} = \frac{1}{1 - 0.93} = 14.29$$

OC curves for p chart



A választásig történő elhullás valószínűsége a korábbi példából 5%. Ha $n = 50$, a 10%-os elhullás esetén a $\beta = 0.88$, így:

$$ARL_1 = \frac{1}{1 - \beta} = \frac{1}{1 - 0.88} = 8.33$$

