

6-2. táblázat

k	$p(0.05)$	$F(0.05)$	$F(0.1)$	$F(0.025)$	$F(0.06)$	$F(0.04)$
0	.00000	.00000	.00000	.00004	.00000	.00000
1	.00000	.00000	.00000	.00045	.00000	.00000
2	.00000	.00000	.00000	.00255	.00000	.00001
3	.00000	.00000	.00000	.00968	.00000	.00007
4	.00001	.00001	.00000	.02784	.00000	.00033
5	.00004	.00005	.00000	.06472	.00000	.00118
6	.00014	.00020	.00000	.12697	.00001	.00350
7	.00042	.00062	.00000	.21680	.00003	.00895
8	.00110	.00172	.00000	.32996	.00011	.02009
9	.00251	.00423	.00000	.45634	.00031	.04032
10	.00517	.00940	.00000	.58305	.00083	.07328
11	.00965	.01905	.00000	.69823	.00200	.12197
12	.01646	.03551	.00000	.79397	.00442	.18773
13	.02586	.06136	.00000	.86724	.00902	.26952
14	.03762	.09898	.00000	.91917	.01715	.36371
15	.05095	.14993	.00000	.95344	.03050	.46471
16	.06452	.21445	.00001	.97458	.05101	.56598
17	.07671	.29116	.00002	.98683	.08057	.66128
18	.08591	.37707	.00004	.99351	.12072	.74578
19	.09090	.46797	.00009	.99695	.17225	.81656
20	.09114	.55911	.00021	.99864	.23490	.87275
21	.08680	.64591	.00044	.99942	.30727	.91511
22	.07870	.72462	.00088	.99976	.38685	.94551
23	.06808	.79269	.00168	.99991	.47033	.96634
24	.05628	.84898	.00308	.99996	.55403	.97997
25	.04455	.89353	.00542	.99999	.63438	.98851
26	.03382	.92735	.00917	1.00000	.70835	.99364
27	.02466	.95201	.01494	1.00000	.77376	.99660
28	.01729	.96930	.02348	1.00000	.82937	.99825
29	.01167	.98097	.03566	1.00000	.87491	.99912
30	.00760	.98856	.05239	1.00000	.91085	.99958
31	.00477	.99334	.07458	1.00000	.93824	.99980
32	.00290	.99623	.10301	1.00000	.95839	.99991
33	.00170	.99793	.13823	1.00000	.97274	.99996
34	.00097	.99890	.18048	1.00000	.98262	.99998
35	.00053	.99943	.22956	1.00000	.98922	.99999
36	.00028	.99971	.28486	1.00000	.99349	1.00000
37	.00015	.99986	.34530	1.00000	.99617	1.00000
38	.00007	.99993	.40946	1.00000	.99781	1.00000
39	.00004	.99997	.47563	1.00000	.99878	1.00000

40	.00002	.99999	.54198	1.00000	.99933	1.00000
----	--------	--------	--------	---------	--------	---------

Látjuk a táblázatból, hogy annak valószínűsége, hogy a selejtszám egy 400 elemű mintában 7 alatt (6 vagy kevesebb) legyen, ha $p=0.05$ (vagyis hogy a nullhipotézist elutasítva elsőfajú hibát kövessünk el negatív irányban) 0.0002. Annak valószínűsége, hogy a selejtszám 33 vagy kisebb legyen, 0.99793, azé, hogy 33-nál nagyobb (legalább 34) legyen, $1-0.99793=0.00207$.

A $\pm 3\sigma$ konvenció szerinti deklarált valószínűség az elsőfajú hibára 0.0027, egyik oldalra ennek fele, vagyis 0.00135. A pozitív irányú tévedésre az egyezés elfogadható (0.00135 helyett 0.00207), a negatív irányúra egy nagyságrendnyi eltérés van (0.00135 helyett 0.0002). Ez azt jelenti, hogy a valóságban 1000 (400 elemű) minta közül nem a deklarált ≈ 3 esetben, hanem csak 10000 közül két esetben hisszük azt, hogy lecsökkent a selejtarány, amikor nem változott.

Az átlagos sorozathosszakra nézve a pozitív irányú tévedésre ez a különbség azt jelenti, hogy a normális eloszlással közelítve $ARL_0=741$, a binomiális eloszlás pontos képleteivel számolva ez $ARL_0=483$. (Vagyis a közelítéssel számolva 741, egyenként 400 elemű minta vétele után várható a selejtarány növekedése miatti téves riasztás, a pontos számítással kapott 483 helyett.) Tehát valamivel rosszabb a valóságos helyzet, mint a deklarált. Ami a negatív irányú tévedést illeti, a selejtarány csökkenését nem 741 minta után, hanem csak 5000 minta után fedezzük föl.

Vizsgáljuk meg a közelítés következményét a másodfajú hibára.

$$\beta = F(UCL_{np} | p_1) - F(LCL_{np} | p_1) = F(33 | p_1) - F(6 | p_1),$$

ugyanis akkor fogadjuk el a nullhipotézist, ha a selejtszám ≤ 33 és > 6 . Számítsuk ki a másodfajú hiba valószínűségét és az átlagos sorozathosszt arra az esetre, ha $p_1=0.1$ ill. $p_1=0.025$, vagyis annak valószínűségét, hogy ne vegyük észre, ha a selejtarány kétszeresére nőtt ill. felére csökkent!

Helyettesítsük először a $p_1=0.1$ ellenhipotézist:

$$\beta = F(33|0.1) - F(6|0.1) = 0.13823 - 0 = 0.13823.$$

Ez azt jelenti, hogy a normális eloszlással való közelítés adta 0.138-del megegyezik a másodfajú hiba valószínűsége, vagyis annak kockázata, hogy nem vesszük észre a selejtarány növekedését.

Az átlagos sorozathossz, amely ahhoz szükséges, hogy a p paraméter kétszeresére való változását észrevegyük, $ARL_1 = \frac{1}{1-\beta} = \frac{1}{1-0.13823} = 1.16$ (a normális közelítéssel kapott értékkel azonos).

A $p_1=0.025$ ellenhipotézissel (vagyis amikor a tényleges selejtarány felére csökkent):

$$\beta = F(33|0.025) - F(6|0.025) = 1 - 0.12697 = 0.87303.$$

Ennek észleléséhez az átlagos sorozathossz: $ARL_1 = \frac{1}{1 - 0.87303} = 7.88$,

valamelyest eltér a normális eloszlással kapott 7.62 értéktől.

Számítsuk most ki a másodfajú hiba valószínűségét arra az esetre, ha $p_1=0.06$ ill. $p_1=0.04$, vagyis annak valószínűségét, hogy ne vegyük észre, ha a selejtarány 20%-kal nőtt ill. csökkent!

Helyettesítsük először a $p_1=0.6$ ellenhipotézist:

$$\beta = F(33|0.06) - F(60.06) = 0.97274 - 0.00001 = 0.97273.$$

Ez azt jelenti, hogy a normális eloszlással való közelítés adta 0.97714-nél kicsit kisebb a másodfajú hiba valószínűsége, vagyis csökken annak kockázata, hogy nem vesszük észre a selejtarány növekedését.

Az átlagos sorozathossz, amely ahhoz szükséges, hogy a p paraméter +20%-os változását észrevegyük, $ARL_1 = \frac{1}{1 - \beta} = \frac{1}{1 - 0.97273} = 36.7$ (a normális közelítéssel kapott 47.8-tól eltérő érték).

A $p_1=0.04$ ellenhipotézissel (vagyis amikor a tényleges selejtarány 20%-kal csökkent):

$$\beta = F(33|0.04) - F(6|0.04) = 0.99996 - 0.0035 = 0.99646.$$

Ennek észleléséhez az átlagos sorozathossz: $ARL_1 = \frac{1}{1 - 0.99646} = 282$, ez

nagyon eltér a normális közelítéssel kapott 130-tól, annak több, mint kétszerese. Tehát ha a binomiális eloszlás képletei helyett a normális eloszlással (a $\pm 3\sigma$ konvencióval) számolunk, a selejtarány csökkenése esetén ($p_1 < p_0$) alábecsüljük azt a mintaszámot, amely a változás észleléséhez szükséges. Másképpen fogalmazva alábecsüljük azt az időt, ami alatt észrevesszük a változást, ténylegesen hosszabb időre van szükség a változás fölfedezéséhez. A selejtarány megnövekedését ($p_1 > p_0$) azonban a valóságban a közelítéssel becsülnél rövidebb idő alatt észrevesszük.

A számítások eredményeit a 6-3. táblázatban tekinthetjük át.

6-3. táblázat

hipotézis	normális		binomiális	
	α	ARL_0	α	ARL_0
$p_0=0.05$	0.00135	741	0.00207	483
H_1	β	ARL_1	β	ARL_1
$p_1=0.04$	0.99232	130	0.99656	282
$p_1=0.06$	0.97714	47.8	0.97273	36.7

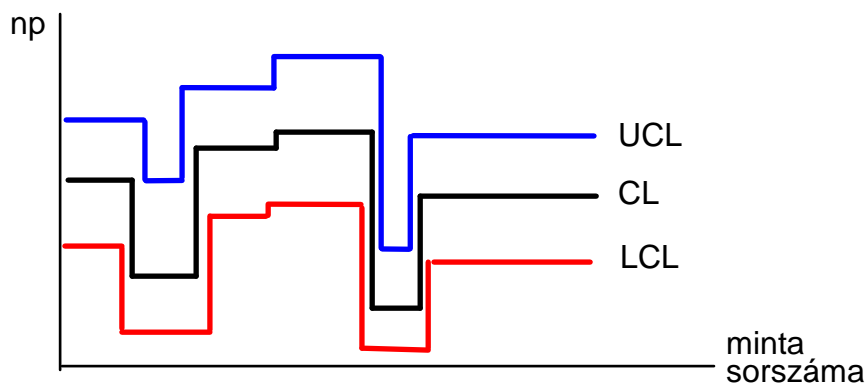
$p_1=0.025$	0.86883	7.62	0.87303	7.88
$p_1=0.1$	0.13933	1.16	0.13823	1.16

Az eltérés oka, hogy olyan esetben helyettesítettük a binomiális eloszlású valószínűségi változót normális eloszlásával, amelyben az nem volt teljesen jogos. A binomiális eloszlás p paraméterére ugyan teljesült a $\frac{1}{n+1} < p < \frac{n}{n+1}$ feltétel, de a várható érték

körüli $\pm 3\sigma$ [vagyis $np \pm 3\sqrt{np(1-p)}$] intervallum határán ill. azon kívül akartuk a valószínűségeket kiszámítani.

Az előzetes adatfelvételnél megállapított paramétereket, így a beavatkozási határokat is a halmozódó adatok alapján időszakonként fölül szokás vizsgálni. Grant és Leavenworth (1988) felhívja a figyelmet arra (idézi Banks, 1989, p. 149), hogy amennyiben a selejtarány csökkent, és ezt nem a felületesebb vizsgálat okozza, a beavatkozási határok szűkíthetők, tudomásul véve, hogy a folyamat megváltozott. Ha viszont a selejtarány nőtt, nem célszerű a beavatkozási határokat tágítani, hanem műszaki-szervezési intézkedéseket kell tenni a romlás okainak földerítésére és kiküszöbölésére.

Előfordul, hogy a minta elemszáma nem állandó. Például nehézkes lenne szigorúan adott számú kis méretű alkatrészt kivenni (pl. egy marék csavart vehetünk ki könnyen), vagy az aznap gyártott összes termék-egyedet egy mintának tekintjük, az pedig napról napra különböző lehet. Az np -kártya csak állandó mintaelemszám esetén használható. Mint a középvonal és a beavatkozási határok képletéből látható, e vonalak a minta n elemszámával változnának, tehát a 6-3. ábrán látható kártyát kapnánk, ami nehezen kezelhető. Ilyenkor a következő fejezetben ismertetendő p kártyát használjuk.



6-3. ábra. np -kártya változó mintaelemszám esetén

6.2. p -kártya

A $\hat{p} = \frac{D}{n}$ selejtarány a selejtszámhoz hasonlóan binomiális eloszlást követ, várható értéke és varianciája:

$$E(\hat{p}) = p, \quad \text{Var}(\hat{p}) = \frac{p(1-p)}{n}.$$

A selejtes darabok arányára vonatkozó Shewhart-kártya az ún. p -kártya, amely változó mintaelemszám esetén is használható. Középvonala p vagy annak becslése, a beavatkozási határokat a $\pm 3\sigma$ konvenció szerint szokás venni.

$$CL_p = \bar{p}, \quad UCL_p = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}, \quad LCL_p = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}.$$

A minták minimálisan szükséges elemszámára ugyanazon megfontolások érvényesek, mint az np -kártyánál.

6-8. példa

Készítsünk p -kártyát a 6-2. példa adataira!

A kártya középvonala az átlagos selejtarány:

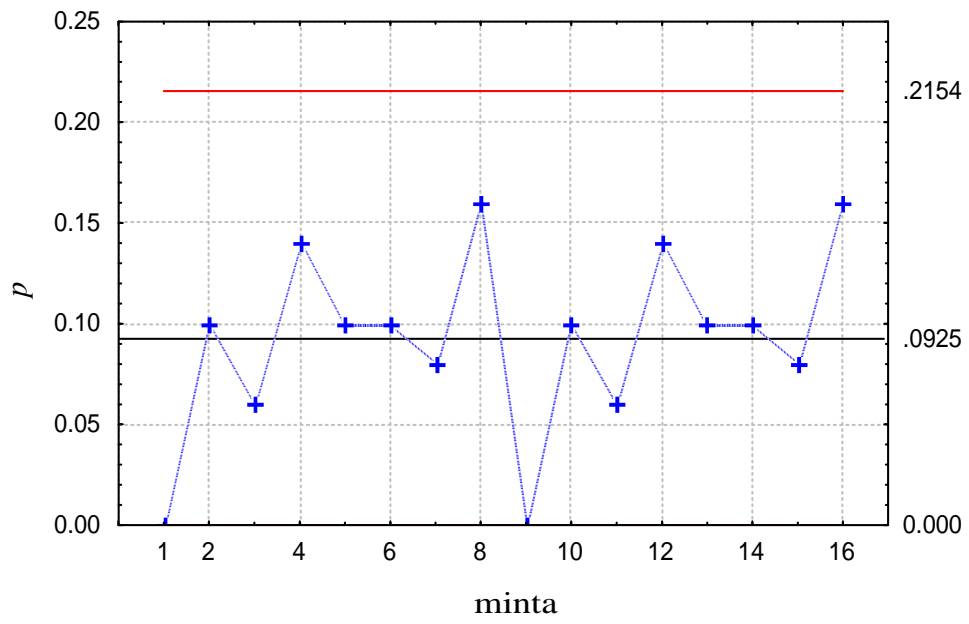
$$CL_p = \bar{p} = 0.0925.$$

A beavatkozási határok:

$$UCL_p = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} = 0.0925 + 3\sqrt{\frac{0.0925 \cdot (1-0.0925)}{50}} = 0.2154,$$

$$LCL_p = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} = 0.0925 - 3\sqrt{\frac{0.0925 \cdot (1-0.0925)}{50}} = -0.0304 \Rightarrow 0.$$

A kártya a 6-4. ábrán látható.



6-4. ábra. p -kártya a 6-8 példához, a STATISTICA programmal

Változó mintaelemszám

Célszerű a vizsgálatokat úgy szervezni, hogy a mintaelemszám állandó legyen. Előfordul azonban, hogy ez nem lehetséges vagy nem sikerül. Például ha az összes terméket minősítjük, és a termelt mennyiség ingadozik, n nem lesz azonos a különböző napokon. Más esetben, ha a minta egy része elvész, tönkremegy, elszennyeződik stb., a mintaelemszám ezért változhat. Analóg probléma, ha az előzetes adatfelvételnél a mintaelemszám más, mint az aktuális ellenőrzésnél.

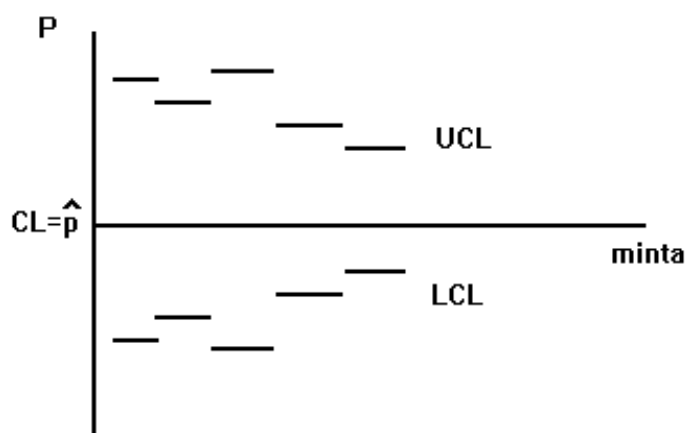
A következő lehetőségeink vannak:

- minden mintára a saját mintaelemszámához tartozó beavatkozási határt számítunk
- a beavatkozási határokat az átlagos mintaelemszámhoz számoljuk ki
- egyazon kártyán többféle határt használunk
- normalizáljuk (standardizáljuk) a \hat{p} valószínűségi változót, hogy u -eloszlású legyen

Nézzük részletesebben az egyes módszereket!

Minden mintára külön határokat számolunk

A határok kiszámítására szolgáló képletekben n mintánként különböző, tehát a beavatkozási határok különbözőek lesznek, a vonalak ugrálnak, ezt illusztrálja a 6-5. ábra. (A számításhoz p az előzetes adatfelvételnél kapott átlagos értékkel helyettesítendő.)



6-5. ábra. p -kártya mintánként különböző határokkal

Az üzemi helyszínen, ha nem számítógéppel dolgoznak, az ilyen, mintánként különböző határok kiszámítása nehéz, ez indokolja a következő egyszerűsítést.

A határokat az átlagos mintaelemszámhoz számoljuk ki

Az átlagos mintaelemszám:

$$\bar{n} = \frac{\sum_i^m n_i}{m},$$

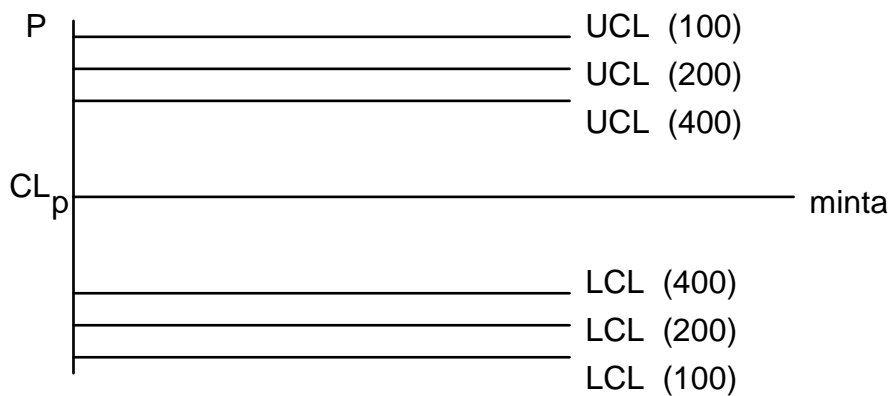
ez helyettesítendő a képletekben n helyére. Akkor használjuk ezt az eljárást, ha nem nagyon (pl. 15%-nál kevésbé) különböznek a mintaelemszámok. Ilyenkor a határokhoz közel eső pontok különös figyelmet érdemelnek, mert a határok csak közelítőek, tehát a szigorú határokkal más lehet a következtetés.

Ha $n_i < \bar{n}$, az átlagos mintaelemszámhoz számolt határok szűkebb tartományt adnak a szigorúnál, tehát ha az illető pont a közelítő határokon belülré esett, a szigorúakon is belül lenne. Ha ugyanekkor a pont a közelítő határokon kívülré esett, de a határhoz közel, a megnyugtató döntéshez ki kell számítani a szigorú határt is.

Amennyiben $n_i > \bar{n}$, a szigorú határok adnak szűkebb tartományt. Ha ilyenkor egy minta pontja a közelítő határokon kívül van, a szigorúak szerint sem lenne elfogadható. Ha a pont a közelítő határokon belülré esett, de a határhoz közel, a megnyugtató döntéshez pontosan kell számolni.

Többféle határ egy kártyán

Akkor járunk el így, ha néhány, de egymástól nagyon különböző mintaelemszám fordul elő, pl. 100, 200, 300. Ekkor mindegyiknek megfelelő határokat ábrázoljuk a kártyán, és mindegyik mintát a neki megfelelő határokhöz képest vizsgáljuk. Ilyen kártyát mutat a 6-6. ábra.



6-6. ábra. p -kártya többféle határral

Normalizált változó ábrázolása

A következő képlet szerint átalakított változó jó közelítéssel u -eloszlású:

$$u_i = \frac{\hat{p}_i - \bar{p}}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}}},$$

és mint ilyenre, a ± 3 határ közvetlenül alkalmazható.

6-9. példa

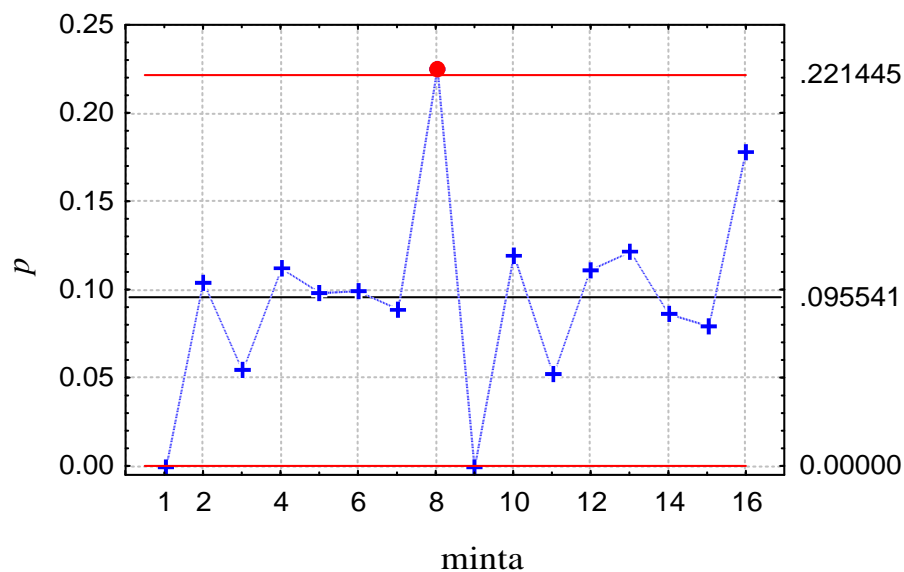
Egy gépen gyártott csapágygolyókból félóránként egy maréknyi mintát vesznek. A 6-4. táblázat mutatja a megvizsgált és a nem megfelelő darabok számát. Kezeljük az adatokat előzetes adatfelvétel eredményeiként, és készítsünk belőlük p -kártyát, a különböző mintaelemszám többféle módon való figyelembe vételével, a $\pm 3\sigma$ konvencióval!

6-4. táblázat

időpont	np	n	időpont	np	n
8:00	0	40	12:00	0	38
8:30	5	48	12:30	5	42
9:00	3	55	13:00	3	57
9:30	7	62	13:30	7	63
10:00	5	51	14:00	5	41
10:30	5	50	14:30	5	58
11:00	4	45	15:00	4	50
11:30	9	40	15:30	8	45
összeg	38	391	összeg	37	394

A 6-7. ábra mutatja az átlagos mintaelemszámmal számolt beavatkozási határokkal készült p -kártyát. A 8. pont eléri a felső beavatkozási határt, de csak éppen hogy, a gondos felhasználó ilyenkor pontos vizsgálatokat végez, kiszámítva a 8. pontra az egyedi határt.

$$\bar{p} = \frac{38 + 37}{391 + 394} = \frac{75}{785} = 0.0955$$



6-7. ábra. p -kártya a 6-9. példához átlagos beavatkozási határokkal, a STATISTICA programmal