6. Minősítéses ellenőrző kártyák

Sokszor előfordul, hogy a termék-egyedek minőségét nem tudjuk mérhető mennyiségekkel jellemezni, csak megfelelő/nem megfelelő kategóriákba sorolhatjuk őket, és a hibás darabokat, vagy az előforduló hibákat számláljuk le. Ezt nevezzük minősítéses ellenőrzésnek. Egyes esetekben a termék jellege olyan, hogy csak így kategorizálhatunk, pl. van-e az alkatrészen sorja, van-e szín-elcsúszás a nyomtatásnál. Más esetekben lehetne ugyan mérni a jellemzőt, de lényegesen olcsóbb csak minősíteni egy idomszerrel (nagyobb-e a geometriai mérete az előírtnál vagy nem). Az is lehetséges, hogy egy termék többféle, egyenként mérhető tulajdonságát szintetizáljuk minősítéssé.

Ilyenkor legegyszerűbb esetben azt veszik föl adatként, hogy a vizsgált n darab közül hány volt nem-megfelelő, ill. a selejt arányát. Ebből készül az np-kártya és a p-kártya. Az is lehet, hogy a vizsgált egységbe tartozó (pl. 5 darab, vagy egy doboz, vagy egy nap alatt gyártott) valamennyi termék-egyeden előforduló hibák száma az adat (c-kártya), vagy a termék valamilyen egységén (pl. egységnyi felületen) előforduló hibák száma (u-kártya). A minősítéses ellenőrző kártyákat nevezik selejt- ill. hiba-kártyáknak is.

A *gyártásközi ellenőrzés* célja itt is az, hogy megállapítsuk, az eloszlás paramétere (binomiális eloszlás esetén p, Poisson-eloszlás esetén λ) nem változott-e meg az elôzetes adatfelvételhez képest.

Az *előzetes adatfelvétel* célja az, hogy az eloszlás típusát meghatározzuk, és paramétereit becsüljük.

A minősítéses ellenőrző kártyák készítését és alkalmazását az MSZ 246/3-57 magyar szabvány írja elő. E könyvben csak a Shewhart-típusú kártyákkal foglalkozunk, de megjegyezzük, hogy a minősítéses jellemzőkre is lehet CUSUM és más ellenőrző kártyákat is készíteni (Montgomery, 1991).

6.1. *np*-kártya

Itt a vizsgált jellemző a p selejtarány, a hibás darabok száma az egész sokaság elemeinek számához viszonyítva. Ha nem az összes termék-példányt vizsgáljuk meg, csak mintát veszünk belőlük, akkor p becslése a mintabeli selejtarány, vagyis a talált selejtes darabok D számának és a minta n elemszámának hányadosa:

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

Azt, hogy az n elemű mintában véletlen kiválasztással hány selejteset találunk, a binomiális eloszlás írja le (ld. az 1. fejezetben). A selejtes darabok D számának várható értéke és varianciája:

$$E(D) = np\,,$$

$Var(D) = np(1-p)$.

Minthogy általában nem visszatevéses mintavételt végeznek, a binomiális eloszlás csak akkor alkalmazható, ha a minta elemszáma lényegesen kisebb a sokaság elemszámánál ($n < N/10$).

A selejtes darabok számára vonatkozó Shewhart-kártya az ún. *np*-kártya. Középvonala *np* vagy annak becslése, a beavatkozási határokat a $\pm 3\sigma$ konvenció szerint szokás venni:

$$CL_{np} = n\overline{p},$$

$$UCL_{np} = n\overline{p} + 3\sqrt{n\overline{p}(1-\overline{p})},$$

$$LCL_{np} = n\overline{p} - 3\sqrt{n\overline{p}(1-\overline{p})},$$

ahol $\overline{p} = \frac{\sum_i^m \hat{p}_i}{m} = \frac{\sum_i^m D_i}{mn}$ (ha minden minta *n* elemű).

Ha az alsó beavatkozási határra negatív érték adódik a képletből, helyette zérust veszünk.

Ha a binomiális eloszlású valószínűségi változót (a *k* selejtszámot) normális eloszlásúval közelítjük,

$$u = \frac{k - np}{\sqrt{np(1-p)}},$$

és a $\pm 3\sigma$ konvencióval kiszámított beavatkozási határokhoz az elsőfajú hiba valószínűsége 0.0027.

Amennyire a binomiális eloszlás eltér a normális eloszlástól, annyira tér el az elsőfajú hiba tényleges α valószínűsége a normális eloszlás esetén a $\pm 3\sigma$ határokhoz tartozó 0.0027-tôl.

Az *előzetes adatfelvétel* célja itt is a folyamat paraméterének (*p*-nek) a becslése, egyúttal a folyamat stabilitásának vizsgálata. Akkor fogadjuk el a becsült paramétert (és az ennek alapján számolt beavatkozási határokat), ha a folyamat kézbentartottnak (stabilnak) mutatkozik. Ezt ugyanúgy, mint a méréses esetben, ellenőrző kártya felvételével ellenőrizzük, mert ekkor a minták egymásutániságában rejlő információt is hasznosítjuk. A kártyán az egyes mintákhoz tartozó *np* értékeket, vagyis a selejtes darabok számát ábrázoljuk a minta sorszáma (vagy a mintavétel időpontja) függvényében. A pontok menetének vizsgálatára itt is használhatunk run teszteket (3.4. pont), de a méréses esetben megismertek közül csak azokat, amelyek nem igénylik normális eloszlás feltételezését (vagyis csak a nem-paraméteres próbákat), ugyanis nem átlag-értékeket ábrázolunk, tehát a centrális határeloszlási tétel sem teszi adatainkat közel normális eloszlásúvá.

Ha kiugró érték (vagy más rendellenesség) van, az ok azonosítása után a hozzá tartozó pontokat elhagyjuk, és a paramétereket (CL_{np}, UCL_{np}, LCL_{np}) újra kiszámoljuk. Ha nem találunk okot, kétféleképpen dönthetünk:

- Kihagyjuk így is a pontokat, ekkor azt kockáztatjuk, hogy az indokoltnál szűkebbek lesznek a beavatkozási határok.
- Nem hagyjuk ki a kérdéses pontokat, ez azzal járhat, hogy a helyesnél szélesebbek lesznek a beavatkozási határok; ha elég sok pontra alapozzuk a becslést, az eltérés nem lesz túlságosan nagy.

A folyamat tanulmányozása során szükség lehet arra, hogy az adatfelvétel két szakaszát összehasonlítsuk, vagyis azt a nullhipotézist vizsgáljuk, hogy a selejtszám (selejtarány) a két szakaszban azonos. Ha egy-egy minta összehasonlítása a feladat, a 2.3.5.3. pontban leírtak szerint járunk el. Ott a binomiális eloszlású valószínűségi változót normális eloszlásúval közelítettük, a helyettesítő normális eloszlás varianciáját az ismeretlen p paraméter mintabeli becsléséből számoltuk.
Ha a két szakasz több mintából áll, a több mintából kiszámolható átlagos selejtarány a centrális határeloszlási tétel értelmében akkor is jó közelítéssel normális eloszlású, ha az egy minta selejtaránya még nem lenne elég közel a normális eloszláshoz (p kicsi vagy nagy, n nem elég nagy). Ráadásul itt a varianciát sem kell az ismeretlen p paraméter mintabeli becsléséből számolnunk, hanem több ismétlés lévén, tapasztalati szórásnégyzetet használhatunk. Ekkor viszont nem u-, hanem t-próbát végzünk.

Jelölje az első szakasz átlagos selejtarányát p_I, a másodikét p_{II}, ekkor a nullhipotézis és ellenhipotézis:

$$H_0\colon E\left(\overline{p}_I\right)=E\left(\overline{p}_{II}\right);\ \ H_1\colon E\left(\overline{p}_I\right)\neq E\left(\overline{p}_{II}\right).$$

A próbastatisztika:

$$t_0=\frac{\overline{p}_I-\overline{p}_{II}}{\sqrt{\frac{s_1^2}{n_I}+\frac{s_2^2}{n_{II}}}}.$$

6-1. példa

Legyen a minták elemszáma 100, az első szakaszban vett 10 minta átlagos selejtaránya 0.1, a selejtarány korrigált tapasztalati szórásnégyzete 0.12; a második szakaszban vett 8 minta átlagos selejtaránya 0.08, a selejtarány korrigált tapasztalati szórásnégyzete 0.085. Döntsük el 0.05-os szinten, hogy a selejtarány a két szakaszban azonos-e!

$$H_0\colon E\left(\overline{p}_I\right)=E\left(\overline{p}_{II}\right);\ \ H_1\colon E\left(\overline{p}_I\right)\neq E\left(\overline{p}_{II}\right)$$

$$t_0=\frac{\overline{p}_I-\overline{p}_{II}}{\sqrt{\frac{s_1^2}{n_I}+\frac{s_2^2}{n_{II}}}}=\frac{0.1-0.08}{\sqrt{\frac{0.12}{10}+\frac{0.085}{8}}}=0.133$$

A $10+8-2=16$ szabadsági fokszámhoz és α=0.5 szignifikanciaszinthez tartozó kritikus érték a függelék III. táblázatából 2.12. Mivel a próbastatisztika talált értéke ez alatt van, elfogadjuk a nullhipotézist, mely szerint a két szakaszban a selejtarány csak a véletlen ingadozás miatt különbözik.

6-2. példa

Egy gépen gyártott csapágygolyókból félóránként 50 elemű mintákat veszünk. A 6-1. táblázat mutatja a selejtes (nem megfelelő méretű) darabok számát (*np*):

6-1. táblázat

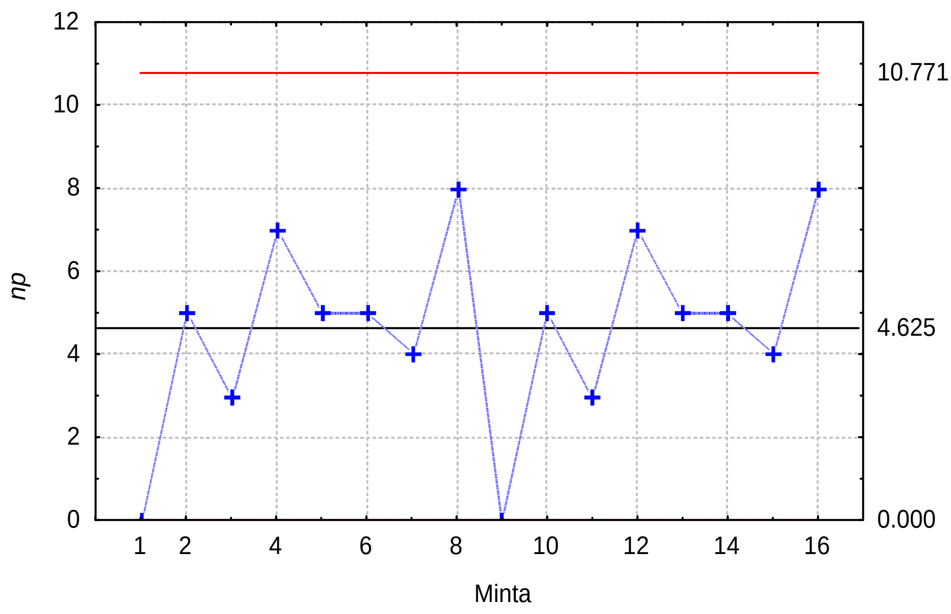
időpont	8:00	8:30	9:00	9:30	10:00	10:30	11:00	11:30
np	0	5	3	7	5	5	4	8

időpont	12:00	12:30	13:00	13:30	14:00	14:30	15:00	15:30
np	0	5	3	7	5	5	4	8

Készítsünk *np*-kártyát az adatokból, feltételezve, hogy előzetes adatfelvételnél kaptuk őket! A kártya *np* paraméterét az adatokból kell becsülni. Ez az átlagos selejtszám 4.625-nek adódik, ez lesz a középvonal helye. Az átlagos selejtarány: $\bar{p}=0.0925$. A fölső beavatkozási határ

$$UCL_{np}=n\bar{p}+3\sqrt{n\bar{p}\left(1-\bar{p}\right)}=4.625+3\sqrt{4.625\cdot\left(1-0.0925\right)}=10.771.$$

Az alsó beavatkozási határra -1.521 adódnék, helyette zérust veszünk. A kártya a 6-1. ábrán látható.



6-1. ábra. *np*-kártya a 6-2. példához, a STATISTICA programmal

A *gyártásközi ellenőrzésnél* az előzetes adatfelvételnél kapott paraméterekkel (CL_{np}, UCL_{np}, LCL_{np}) építjük föl az *np*-kártyát.

Ha *p előírt értéke* áll rendelkezésre, a föntebbi képletekbe azt helyettesítjük. Ez, mint a méréses ellenőrző kártyáknál, azért kockázatos, mert esetleg akkor is instabilnak minősítjük a folyamatot, ha stabil, csak nem az előírt *p* a paraméter, hanem annál nagyobb vagy kisebb a tényleges selejtarány.

A minták elemszámát célszerűen nagyra szokták választani, különben nagyon bizonytalanok lennének a statisztikai következtetések. Pyzdek szerint ökölszabály, hogy np>5 legyen, így ha pl. p=0.03, n>5/0.03=166.7, vagyis legalább 167 elemű mintát kell venni.

Arra is törekszenek, hogy a nem-megfelelő darabok előfordulási valószínűsége jelentős legyen, mert csak ekkor kapunk értékelhető információt a hibás darabok arányáról. Pl. ha p=0.03, 40 elemű mintából $\binom{40}{0}0.03^0 0.97^{40} = 0.2957$ annak valószínűsége, hogy nem fordul elő hibás darab, ugyanez 80 elemű mintánál $\binom{80}{0}0.03^0 0.97^{80} = 0.0874$. Vagyis 40 elemű mintáknál 100 esetből átlagosan harmincszor fordulna elő, hogy egyetlen hibás sincs, 80 elem esetén 100 mintából csak kb. 9 esetben fordul elő, hogy nincs benne hibás.

6-3. példa

Minimálisan hány elemű mintákat kell vennünk, ha azt akarjuk, hogy 99% valószínűséggel találjunk legalább 1 hibás darabot, vagyis $P(D>0) \geq 0.99$, amennyiben p=0.03?

A binomiális eloszlás összefüggéseivel számolunk.

$$P(D>0) = 1 - P(D=0) = 1 - \binom{n}{0} \cdot 0.03^0 \cdot 0.97^n = 1 - 0.97^n = 0.99$$

$$n = \frac{\ln 0.01}{\ln 0.97} = 151.2$$

Tehát a szükséges mintaelemszám 152.

Amennyiben a *p* sokaságbeli selejtarány kicsi, a normális eloszlás közelítő összefüggésével számolva az alsó beavatkozási határra negatív szám adódik, ezt zérusra igazítjuk. Ha azt akarjuk, hogy az alsó beavatkozási határ zérus fölött legyen, a következő feltételnek kell teljesülnie:

$$LCL_{np} = np - 3\sqrt{np(1-p)} > 0 .$$

Ebből a szükséges mintaelemszám:

$$n > \frac{9(1-p)}{p}.$$

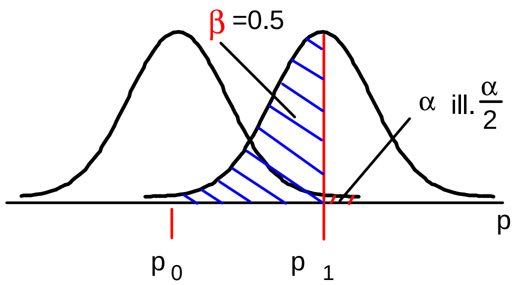
6-4. példa

Mennyi az ahhoz szükséges minimálisan szükséges mintaelemszám az előbbi p=0.03 esetén, hogy az alsó beavatkozási határ zérus fölött legyen?

$$n > \frac{9 \cdot 0.97}{0.03} = 291.$$

További szempont a minta nagyságának megválasztására az adott selejtarány-különbséghez tartozó másodfajú hiba nagysága. Szokás például vizsgálni [Duncan (1974), idézi Montgomery, 1991, p.161], mekkora minták kellenek, ha azt akarjuk, hogy 50% biztonsággal észrevegyük a zavar fellépése utáni első mintavételnél a $\Delta = p_1 - p_0$ nagyságú változást (növekedést) p-ben (vagyis az ekkora eltéréshez tartozó másodfajú hiba elkövetésének valószínűsége 0.5 legyen). A 6-2. ábra szerint a fölső beavatkozási határ ekkor meg kell, hogy egyezzék az ellenhipotézis szerinti p_1 értékkel. (Vigyázat, ez az előírás már meghatározza az elsőfajú hiba megengedett valószínűségét!)

$$\Delta = 3\sqrt{\frac{p_0(1-p_0)}{n}}, \text{ ahonnan } n = \frac{9p_0(1-p_0)}{\Delta^2}.$$



6-2. ábra. A $\beta = 0.5$ valószínűségű másodfajú hiba szemléltetése

6-5. példa

Hány elemű mintákat kell ahhoz vennünk, hogy a változást követő első mintavételnél 50%-os valószínűséggel észrevegyük, hogy a selejtarány 5%-ról 10%-ra nőtt (p_0=0.05 és p_1=0.1)?

$\Delta = 0.10 - 0.05 = 0.05$

$n = \dfrac{9 \cdot 0.05 \cdot 0.95}{0.05^2} = 171$, tehát legalább 171 elemű mintákra van szükség.

6-6. példa

Számítsuk ki az *np*-kártya paramétereit 400 elemű mintákra és p_0=0.05 selejtarányra a $\pm 3\sigma$ konvenció szerint (vagyis a binomiális eloszlás normális eloszlással való közelítésével)!

A kártya középvonala:

$CL_{np} = np_0 = 400 \cdot 0.05 = 20$,

$$UCL_{np} = np_0 + 3\sqrt{np_0(1-p_0)} = 20 + 3\sqrt{400 \cdot 0.05 \cdot 0.95} = 20 + 13.08 = 33.08 \to 34$$

$$LCL_{np} = np_0 - 3\sqrt{np_0(1-p_0)} = 20 - 3\sqrt{400 \cdot 0.05 \cdot 0.95} = 20 - 13.08 = 6.92 \to 6.$$

Ha egész értékre kerekítünk, ezt a középvonaltól elfelé kell tenni. UCL_{np} = 34 azt jelenti, hogy beavatkozásra van szükség, ha $x \ge 34$, de $x = 33$-nál még nem. LCL_{np} = 6 azt jelenti, hogy a folyamat stabilitására vonatkozó nullhipotézist akkor utasítjuk el, ha $x \le 6$ és elfogadjuk, ha $x > 6$.

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!

$\beta = P(6 < x < 34)$

$$\beta = F(u_{UCL} | p_1) - F(u_{LCL} | p_1) = \Phi\left(\frac{UCL - np_1}{\sqrt{np_1(1-p_1)}}\right) - \Phi\left(\frac{LCL - np_1}{\sqrt{np_1(1-p_1)}}\right)$$

A Φ itt most az *u*-eloszlás (standardizált normális eloszlás) eloszlásfüggvényét jelöli. Minthogy diszkrét eloszlást (a binomiális eloszlást) helyettesítünk folytonossal, az ún. folytonossági korrekciót is alkalmazni kell.

A p_1=0.1 ellenhipotézisre:

$$\beta = \Phi\left(\frac{34 - 0.5 - 400 \cdot 0.1}{\sqrt{400 \cdot 0.1 \cdot (1 - 0.1)}}\right) - \Phi\left(\frac{6 + 0.5 - 400 \cdot 0.1}{\sqrt{400 \cdot 0.1 \cdot (1 - 0.1)}}\right) =$$

$$= \Phi(-1.0833) - \Phi(-5.583) = 0.13933 - 0 = 0.13933.$$

A p_1=0.025 ellenhipotézisre:

$$\beta = \Phi\left(\frac{34 - 0.5 - 400 \cdot 0.025}{\sqrt{400 \cdot 0.025 \cdot (1 - 0.025)}}\right) - \Phi\left(\frac{6 + 0.5 - 400 \cdot 0.025}{\sqrt{400 \cdot 0.025 \cdot (1 - 0.025)}}\right) =$$

$$= \Phi(7.526) - \Phi(-1.1209) = 1 - 0.13117 = 0.86883.$$

Vagyis annak valószínűsége, hogy ne vegyük észre, ha a selejtarány kétszeresére nőtt, csak 14%, de annak valószínűsége, hogy ne vegyük észre, ha a selejtarány felére csökkent, már 87%! Az átlagos sorozathossz, amely p megváltozásának észleléséhez szükséges,

p_1=0.1 esetén

$$ARL_1 = \frac{1}{1 - \beta} = \frac{1}{1 - 0.13933} = 1.16,$$

p=0.025-hez

$$ARL_1 = \frac{1}{1 - 0.86883} = 7.62\,.$$

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!

$$\beta = F\left(u_{UCL} \middle| p_1\right) - F\left(u_{LCL} \middle| p_1\right) = \Phi\left(\frac{UCL - np_1}{\sqrt{np_1(1 - p_1)}}\right) - \Phi\left(\frac{LCL - np_1}{\sqrt{np_1(1 - p_1)}}\right)$$

A p_1=0.06 ellenhipotézisre:

$$\beta = \Phi\left(\frac{34 - 0.5 - 400 \cdot 0.06}{\sqrt{400 \cdot 0.06 \cdot (1 - 0.06)}}\right) - \Phi\left(\frac{6 + 0.5 - 400 \cdot 0.06}{\sqrt{400 \cdot 0.06 \cdot (1 - 0.06)}}\right) =$$

$$= \Phi(2.000) - \Phi(-3.684) = 0.97726 - 0.00012 = 0.97714.$$

A p_1=0.04 ellenhipotézisre:

$$\beta = \Phi\left(\frac{34 - 0.5 - 400 \cdot 0.04}{\sqrt{400 \cdot 0.04 \cdot (1 - 0.04)}}\right) - \Phi\left(\frac{6 + 0.5 - 400 \cdot 0.04}{\sqrt{400 \cdot 0.04 \cdot (1 - 0.04)}}\right) =$$

$$= \Phi(4.465) - \Phi(-2.424) = 1 - 0.00768 = 0.99232.$$

Vagyis annak valószínűsége, hogy ne vegyük észre, ha a selejtarány 20%-kal nôtt, 97.7%, annak valószínűsége, hogy ne vegyük észre, ha a selejtarány 20%-kal csökkent, 99.2%! Az átlagos sorozathossz, amely p megváltozásának észleléséhez szükséges,

p=0.06 esetén

$$ARL_1 = \frac{1}{1-\beta} = \frac{1}{1-0.97714} = 43.8,$$

p=0.04-höz

$$ARL_1 = \frac{1}{1-0.99232} = 130\,.$$

6-7. példa

Számítsuk most ki az első- és másodfajú hiba valószínűségét a 6-6. példabeli beavatkozási határokra, de a binomiális eloszlás pontos képleteivel!

A 6-2. táblázat (amelyet a STATISTICA programmal állítottunk elő) tartalmazza a 400 elemű mintára az egyes k selejtszámokhoz tartozó eloszlásfüggvény-értékeket vagy kumulált valószínűségeket (vagyis annak valószínűségét, hogy a selejtszám k-nál kisebb, vagy azzal egyenértékű). A táblázat előállításának megértéséhez a p=0.05 paraméter-értékhez feltüntettük az egyes k selejtarányok előfordulási valószínűségét, vagyis a sűrűségfüggvény értékét.

Emlékeztetőül:

$$F_k(p) = P(D \le k)\,.$$

Például:

$$p^4(0.05) = \binom{400}{4} \cdot 0.05^4 \cdot 0.95^{396} = 0.00001;\ \text{és mivel}$$

$$p^0(0.05) = p^1(0.05) = p^2(0.05) = p^3(0.05) = 0\,,$$

$$F_4(0.05) = \sum_{k=0}^{4} p^k(0.05) = 0.00001\,,$$

$$p^5(0.05) = \binom{400}{5} \cdot 0.05^5 \cdot 0.95^{395} = 0.00004;$$

$$F_5(0.05) = \sum_{k=0}^{5} p^k(0.05) = 0.00001 + 0.00004 = 0.00005\,.$$