8. A mérőeszközök képességvizsgálata[1]

A vizsgálat célja annak megállapítása, hogy a használt mérőeszköz elég kis hibával használható-e ahhoz, hogy vele a folyamatról információt szerezzünk.

Az AIAG (Automotive Industry Task Force) javasolta R&R (reproducibility and repeatability) vizsgálatnál azt elemezzük, hogy a mérőeszközzel való mérés ismételhetősége, a különböző személyek általi reprodukálhatósága hogyan viszonylik a mérendő objektumok közötti különbségekhez, egyúttal becslést is adunk az eltérés-összetevők varianciájára. A professzionális statisztikai programok ezt az elemzést grafikus szemléltetéssel is kiegészítik. Ezt a vizsgálatot írja elő az autóipari beszállítók QS-9000 előírás-csomagja is.

Az elemzéshez szükséges mérések a következők:

Kiválasztunk véletlenszerűen a mérendő darabok (pl. alkatrészek) közül valahányat (pl. tizet), mindegyiket több (mondjuk 5) kezelő többször (pl. háromszor) megméri.

A mérési eredmények a következő okokból különböznek:

- az alkatrészek különbözőek, remélhetően ez adja az eltérések döntő részét;
- a mérés ismétlésekor a mérőeszköz jellegétől függően minden kezelő véletlen mérési hibákat követ el (pl. parallaxis-hibát, a vastagságmérőt jobban vagy kevésbé szorítja rá az alkatrészre, a mérőszalagot jobban vagy kevésbé húzza meg, a vonalzót csak többé-kevésbé tudja a mérendő élhez illeszteni stb.), ez határozza meg a mérés *ismételhetőségét* (repeatability);
- a kezelők ügyessége, figyelmessége, munkájuk megbízhatósága is különböző lehet, ez határozza meg a mérés *reprodukálhatóságát* (reproducibility).
- lehetséges, hogy egyes kezelők a különböző alkatrészeket különböző hibával képesek mérni, vagyis *kölcsönhatás* lehet a kezelő és az alkatrész között.

Az eltérések fölbontását szemlélteti a 8-1. ábra.

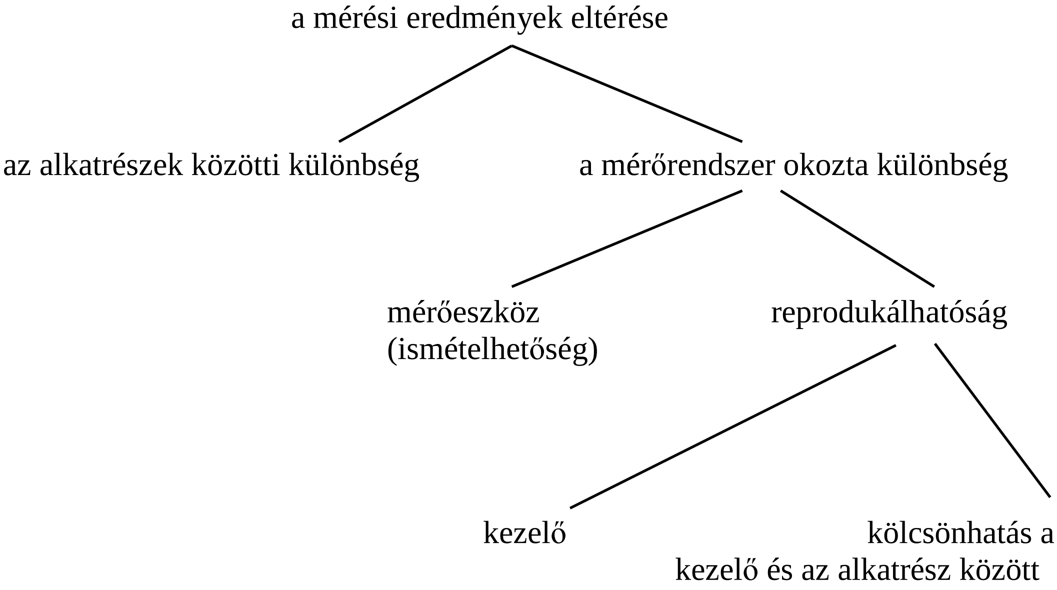
Ennek megfelelően a különböző alkatrészekre kapott mérési eredmények eltérését (ingadozását) jellemző variancia a következőképpen bontható föl:

$$\sigma^2_{\text{teljes}} = \sigma^2_{\text{alkatrész}} + \sigma^2_{\text{mérés}}\,.$$

Ha képet akarunk alkotni arról, hogy a teljes ingadozás mekkora, nemcsak az alkatrészek mérhető jellemzőinek ingadozását, hanem a mérés bizonytalanságát is ismernünk kell.

Két alapvető módszer használatos. Az egyik (R&R study) az alkatrészek jellemzőjének ingadozásával együtt méri a mérés bizonytalanságát, a másik külön csak a mérés bizonytalanságának megismerésére irányul.

[1] A fejezet feltételezi a varianciaanalízis ismeretét. A könyv terjedelme nem teszi lehetővé, hogy ezt kellően megalapozzuk. Utalunk ezért a szakirodalomra (Meszéna - Zierman, 1981; Vincze, 1975; Kemény - Deák, 1990).



8-1. ábra. Az ingadozás forrásainak felbontása

8.1. A reprodukálhatóság és ismételhetőség vizsgálata (AIAG R&R study)

A mérésnek tulajdonítható ingadozás is két részből áll:

$$\sigma^2_{\text{mérés}} = \sigma^2_{\text{reprod}} + \sigma^2_{\text{ism}}$$

Tovább bontva a reprodukálhatósági varianciát:

$$\sigma^2_{\text{reprod}} = \sigma^2_{\text{kezelõ}} + \sigma^2_{\text{alkatrész*kezelõ}}$$

Maga a kísérleti terv két faktor (alkatrész, kezelő) szerinti kereszt-osztályozás, többszöri ismétléssel, mindkét faktor véletlen jellegű.

A 8-1. példában szereplő adatokat fogjuk az elemzésekhez használni.

8-1. példa (Papp L., Róth P., Németh L., 1992)

> Édesipari üzemben MEO laboratóriumában csokoládé-massza zsírtartalmát a törésmutató mérésével (Abbe-refraktométerrel) ellenőrzik. Az előírás: 33.8±1.5%. A mérési adatokat a 8-1. táblázat tartalmazza.
>
> Az “alkatrész” helyett itt az elemzendő minta áll.

8-1. táblázat

operátor:	A: Kiss Rózsa			B: Nagy Judit			C: Kovács Edit		
minta	1.mérés	2.mérés	3.mérés	1.mérés	2.mérés	3.mérés	1.mérés	2.mérés	3.mérés
1	33.9	34.0	33.9	33.6	33.7	33.8	33.2	33.7	33.7
2	34.2	34.3	34.2	34.2	34.3	34.4	34.3	34.7	34.5
3	33.5	33.6	34.0	34.3	34.0	34.0	33.9	33.8	33.7
4	34.7	34.9	35.0	34.5	34.8	34.4	34.2	34.5	34.4
5	32.5	32.9	32.6	33.3	32.8	33.0	32.9	33.0	32.7
6	32.7	32.7	33.2	33.0	32.6	32.8	32.7	33.2	32.8
7	34.8	35.0	35.2	34.9	35.2	35.0	35.1	34.9	34.8
8	33.3	33.3	33.8	33.3	33.4	33.4	33.1	33.5	33.5
9	33.6	33.7	33.9	33.6	33.5	33.4	33.4	33.7	33.2
10	33.4	33.6	33.5	33.5	33.7	33.4	33.4	33.6	33.2

Az eredmények feldolgozására (a varianciák becslésére) két módszer ismeretes: a terjedelem fölbontásával és varianciaanalízissel, utóbbinak ismeretét most feltételezzük. A két módszer közötti alapvető különbség az egyes hatások varianciája becslésének módjában van (a terjedelemből ill. az eltérés-négyzetösszegekből), ennek következménye, hogy a terjedelem-módszernél kölcsönhatásokat nem vehetünk figyelembe.

8.1.1. Terjedelem-módszer

Itt a varianciákat az átlag-terjedelem ellenőrző kártyáknál megismert módon, a terjedelemből becsüljük. Nevezik átlag-terjedelem módszernek is (QS-9000).
Először az ismétlések varianciájának becslését végezzük el:

$$\hat{\sigma}_{\text{ism}} = \frac{\overline{R}_{\text{ism}}}{d_2}\ .$$

ahol $\overline{R}_{\text{ism}}$ az ismétlések átlagos terjedelme, vagyis az ismétlések terjedelmének átlaga; d_2 értéke az ismétlésszám függvényében a függelék terjedelem-kártyákkal kapcsolatban megismert V. táblázatából veendő.
A következő lépés a különböző személyek végezte mérések reprodukálhatóságára jellemző variancia becslése:

$$\sigma_{\text{reprod}} = \frac{R_{\text{reprod}}}{d_2}\ ,$$

ahol R_{reprod} a kezelőknek (operátoroknak) az összes alkatrészre kapott átlagos mérési eredményei átlagának terjedelme (a legnagyobb és legkisebb érték különbségének abszolut értéke); d_2 a kezelők számához veendő, tehát nem feltétlenül ugyanaz az érték, amelyet az ismétlések terjedelmével összefüggésben használtunk.

A pontosság kedvéért figyelembe szokták venni (nem mindenütt), hogy a reprodukálhatóság így becsült varianciája tartalmazza az ismétlések bizonytalanságát is, mert az átlagokat, amelyekből számítottuk, ez a bizonytalanság terheli. Ha az operátorok között semmiféle eltérés nem lenne, akkor is valamelyest eltérő eredményeket kapnának, a mérőeszköz bizonytalansága miatt, és ez utóbbit éppen az ismétlések varianciája méri. Megjegyezzük, hogy a reprodukálhatóság varianciája itt egyértelműen a kezelők közötti eltérésekből adódik, a másik, az ANOVA-módszerrel végzett feldolgozásnál nem.

A korrigált érték számítása:

$$\hat{\sigma}_{\text{reprod}} = \sqrt{\left(\frac{R_{\text{reprod}}}{d_2}\right)^2 - \frac{\hat{\sigma}^2_{\text{ism}}}{n_{\text{ism}} \cdot n_{\text{alkatrész}}}}\,.$$

Végül az alkatrészek közötti eltérés varianciájának becslése:

$$\hat{\sigma}_{\text{alkatrész}} = \frac{R_{\text{alkatrész}}}{d_2}\,,$$

ahol $R_{\text{alkatrész}}$ az egyes alkatrészekre az összes kezelők kapta eredmények átlagának terjedelme; d_2 az alkatrészek számához veendô.

Szokás azt vizsgálni, hogy az eltérések mekkora részéért felelősek az egyes források, ez azt jelenti, hogy az egyes varianciákat a teljes eltérések varianciájának százalékában fejezik ki.

Ugyancsak szokás a (C_P) folyamatképességi mutatók analógiájára a teljes ingadozást és az egyes összetevők hatását viszonyítani a tűrésmező szélességéhez, ehhez 99%-os (5.15σ szélességű) konfidenciaintervallumokat számolnak, és belőlük egy P/T (precision to tolerance) mutatót képeznek:

$$\frac{P}{T} = \frac{5.15 \cdot \hat{\sigma}_{\text{mérés}}}{USL - LSL}\,.$$

(Lehet az 5.15-os szorzó helyett 6-tal is számolni, ez a jól ismert $\pm 3\sigma$ határ, amelyhez nem 99%, hanem 99.73% valószínűség tartozik; a számítógépi programoknál ez általában megadható paraméter).

Azt is meg szokták adni, hogy az alkatrészek mért jellemzőjének tapasztalt ingadozási tartományában a mérési módszer hány kategóriát tud megkülönböztetni, ehhez az alkatrészekre kapott σ-becslést (ill. a konfidenciaintervallumot) viszonyítják a mérési módszer becsült σ-jához. A megkülönböztethető kategóriák száma:

$$\frac{\hat{\sigma}_{alkatrész}}{\hat{\sigma}_{mérés}}\sqrt{2}$$ (lefelé egészre kerekítve).

Ha pl. 10 alkatrészre ez az érték 4, ez azt jelenti, hogy a 10 alkatrész mért jellemzőjében nem különbözik egymástól eléggé ahhoz, hogy mindegyiket különbözőnek találjuk, csak 4 csoportba sorolhatjuk őket. Ha finomabb megkülönböztetést szeretnénk tenni, nagyobb felbontású (kisebb varianciájú) mérésre van szükségünk. Ha a megkülönböztethető csoportok száma 2 alá kerül, a mérőeszköz nem alkalmas, mert minden alkatrészt egyformának talál. Pontosabban ez úgy értendő, hogy amennyiben a vizsgált alkatrészek mind a tűréshatáron belül voltak, az ilyen mérőeszköz jó/nem jó döntéshez használható csak. Vagyis méréses ellenőrző kártyához ill. méréses átvételi ellenőrzéshez nem használható.

8-2. példa

Végezzük el a 8-1. táblázat adatainak elemzését az ismételhetőségi-reprodukálhatósági elemzést a terjedelem-módszerrel!

A számításokhoz szükséges átlag- és terjedelem-adatokat a 8-2. táblázatban készítettük elő.

8-2. táblázat

minta	kezelő						átlag
	Kiss Rózsa		Nagy Judit		Kovács Edit		
	átlag	terje-delem	átlag	terje-delem	átlag	terje-delem	
1	33.93	0.1	33.70	0.2	33.53	0.5	33.722
2	34.23	0.1	34.30	0.2	34.50	0.4	34.344
3	33.70	0.5	34.10	0.3	33.80	0.2	33.867
4	34.87	0.3	34.57	0.4	34.37	0.3	34.600
5	32.67	0.4	33.03	0.5	32.87	0.3	32.856
6	32.87	0.5	32.80	0.4	32.90	0.5	32.856
7	35.00	0.4	35.03	0.3	34.93	0.3	34.989
8	33.47	0.5	33.37	0.1	33.37	0.4	33.400
9	33.73	0.3	33.50	0.2	33.43	0.5	33.556
10	33.50	0.2	33.53	0.3	33.40	0.4	33.478
átlag	33.797	0.33	33.793	0.29	33.710	0.38	33.7667

Az ismétlések varianciájának becslése:

Az ismétlések terjedelmét a minták ("alkatrészek") és az operátorok szerint kell átlagolni.

$\overline{R}_{\text{ism}} = (0.33 + 0.29 + 0.38) / 3 = 0.333$, az ismétlések 3 mérésből álltak, tehát a függelék V. táblázatából d_2=1.693, ezzel

$$\hat{\sigma}_{\text{ism}} = \frac{\overline{R}_{\text{ism}}}{d_2} = \frac{0.333}{1.693} = 0.197\,.$$

A reprodukálhatóság varianciájának becslése:

A kezelők (operátorok) átlagos mérési eredményei a 8-2. táblázat utolsó sorában láthatók. A három operátor mérései átlagának terjedelme:
$R_{\text{reprod}} = 33.797 - 33.710 = 0.087$.

Itt is 3 eredmény terjedelméről van szó, tehát ugyanazt a d_2=1.693 értéket használhatjuk, mint az előbb az ismétléseknél, ezzel:

$$\hat{\sigma}_{\text{reprod}} = \frac{R_{\text{reprod}}}{d_2} = \frac{0.087}{1.693} = 0.0454 .$$

Az ismétlések varianciájával korrigált érték:

$$\hat{\sigma}_{\text{reprod}} = \sqrt{\left(\frac{R_{\text{reprod}}}{d_2}\right)^2 - \frac{\hat{\sigma}^2_{\text{ism}}}{n_{\text{ism}} \cdot n_{\text{alkatrész}}}} = \sqrt{\left(\frac{0.087}{1.693}\right)^2 - \frac{0.0389}{3 \cdot 10}} = 0.0277$$

Az "alkatrészek" közötti eltérés varianciájának becslése:

Az egyes alkatrészekre kapott átlagos eredmények terjedelme a 8-2. táblázat utolsó oszlopából:

$R_{\text{alkatrész}} = 34.989 - 32.856 = 2.133$.

Itt most 10 eredmény terjedelméről van szó, tehát a függelék V. táblázatából d_2=3.078, amellyel

$$\hat{\sigma}_{\text{alkatrész}} = \frac{R_{\text{alkatrész}}}{d_2} = \frac{2.133}{3.078} = 0.671 .$$

Vizsgáljuk meg, hogy az eltérések mekkora részéért felelősek az egyes összetevők!
Az eredményeket (a STATISTICA program számítási eredményeit) a 8-3. táblázat első oszlopában foglaltuk össze.

8-3. táblázat

az ingadozás forrása	$\hat{\sigma}$	$\hat{\sigma}^2$	% R&R-ben	% a teljes ingadozásban
(1) ismétlés	0.1972	0.0389	98.08	7.94
(2) reprodukálhatóság	0.0276	0.0008	1.92	0.16
(1+2) R & R		0.0397	100.00	8.10
(3) alkatrészek között	0.6709	0.4501		91.90
(1+2+3) teljes		0.4897		100.00

Az első számoszlop ($\hat{\sigma}$) első három sora tartalmazza az előbbiekben kiszámított $\hat{\sigma}$ értékeket, az ismétlésre, a reprodukálhatóságra és az "alkatrészekre". A második számoszlop ($\hat{\sigma}^2$) az első négyzete. Az ismétlés és a reprodukálhatóság varianciájának összege szerepel az R&R (a mérőrendszer okozta különbség) sorban. Ha ehhez még az “alkatrészek” közötti eltérés varianciáját is hozzáadjuk, kapjuk a teljes varianciát, amely a 8-1. ábra legfölső eleme.

A harmadik számoszlop (% R&R-ben) a mérés varianciájának összetevőkre bontását mutatja, ezért a mérőrendszer okozta különbség (R&R) 100%.

A negyedik oszlop (% a teljes ingadozásban) a teljes (tehát a különböző alkatrészek közötti különbségeket is tartalmazó) ingadozás varianciájának fölbontása. Ezen belül az első két sor összege a harmadik sorbeli adat (R&R), ezt az “alkatrészek” közötti eltérés varianciája egészíti ki 100%-ra.

Nemcsak a varianciákra szokás összehasonlításokat végezni, hanem a mérési adat mint valószínűségi változó 99%-os konfidenciaintervallumának szélességére, az ún. ingadozási tartományra is. A 99%-os konfidenciaintervallum u=2.575-hez tartozik, vagyis az intervallum szélessége $5.15\cdot\sigma$. A 8-4. táblázat mutatja ezeket az összehasonlításokat.

8-4. táblázat

az ingadozás forrása	$\hat{\sigma}$	az ingadozási tartomány szélessége ($5.15\cdot\hat{\sigma}$)	a teljes ingadozás %-ában	a tűrésmező %-ában
ismételhetőség (mérőeszköz)	0.1972	1.0158	28.18	33.86
reprodukálhatóság (operátor)	0.0276	0.1422	3.94	4.74
R & R	0.1992	1.0257	28.46	34.19
alkatrészek	0.6709	3.4549	95.86	115.16
teljes ingadozás	0.6998	3.6040	100.00	120.13
tűrésmező		3.0000		100.00

Itt az első számoszlop a 8-3. táblázat első oszlopa, kiegészítve két számított értékkel (R&R és teljes ingadozás). Az első számoszlop ($\hat{\sigma}$) harmadik és ötödik eleme a 8-2. táblázat második oszlopában kiszámított $\hat{\sigma}^2$ megfelelő elemének négyzetgyöke, tehát nem az első oszlopon belüli összegzéssel keletkezik (pl. 0.67086≠0.197239+0.027607). A második oszlopban az 5.15σ szélességű intervallumok láthatók, amelyek az első oszlop adataiból számolhatók (pl. 5.15·0.197239=1.015779). Az utolsó sorban van a tűrésmezőnek a feladatban megadott szélessége (itt 3 egység).

A harmadik számoszlopban a folyamat teljes ingadozásának tartományához (“alkatrész”+mérőrendszer), a $\hat{\sigma}_{teljes}$=0.699798-ból 5.15-tel való szorzással kapott 3.603961-hez viszonyítjuk az egyes összetevők 99%-os valószínűségű (

$5.15\cdot\hat{\sigma}$) ingadozási tartományát. Vegyük észre, hogy mivel az összegzés a varianciákra érvényes, az azok négyzetgyökével arányos konfidencia-intervallumokra nem, az összetevők intervallumainak összege nem egyezik meg a négyzetösszeg intervallumának négyzetgyökével, így az intervallumok összege 100% fölött van! Ugyanígy az ismételhetőség és a reprodukálhatóság konfidenciaintervallumának összege nem adja ki a mérés konfidencia-intervallumát.

A negyedik oszlopban a tűrésmező szélességéhez viszonyítjuk a konfidencia-intervallumokat, azt véve 100%-nak.

Az eredmények azt mutatják, hogy a példában az ismételhetőség (tehát amikor ugyanazon személy ugyanazt az alkatrészt többször méri) nem kielégítő, ez adja a mérési variancia 98%-át (8-3. táblázat, harmadik számoszlop, első sor), tehát a mérőeszközön kell javítani. Ezzel a mérési módszerrel a mérés bizonytalansága a tűrésmező szélességének 34.2%-a.
Az is látható, hogy az “alkatrészek” jellemzőjének ingadozása 15%-kal meghaladja a tűrésmezőt, tehát maga a folyamat sem felel meg az előírásoknak ($C_P < 1$).
A megkülönböztethető kategóriák száma:

$$\frac{\hat{\sigma}_{\text{alkatrész}}}{\hat{\sigma}_{\text{mérés}}}\sqrt{2} = \frac{0.671}{0.199}\sqrt{2} = 4.76 \Rightarrow 4\,.$$

8.1.2. ANOVA-módszer

Itt a varianciák becslését az ANOVA-táblázatból számolják ki, (ezzel a becslésre a momentumok módszerét alkalmazzák). A modell a két véletlen faktor szerinti kereszt-osztályozásra a következő (x a vizsgált minőségi jellemző):

$$x_{ijk} = \mu + A_i + O_j + AO_{ij} + \varepsilon_{k(ij)}$$

ahol A jelöli az alkatrészeket, O az operátorokat (kezelőket), ε az ismétlési hibát, melynek indexében a $k(ij)$ azt jelöli, hogy a k-adik ismétlés egy i-j (alkatrész-kezelő) kombinációhoz tartozik.

Legyen általánosan az A faktornak (alkatrészek) r szintje, az O faktornak (kezelők) q, és végezzünk p-szer ismétlést. A 8-5. táblázatban mutatjuk be az ANOVA szokásos összesítő táblázatát, ν jelöli a szabadsági fokok számát.

Ha a $\sigma_A^2 = 0$ nullhipotézist (az alkatrészek között nincs különbség) kívánjuk vizsgálni, az s_A^2 / s_{AO}^2 arányra végzünk F-próbát, mert a H_0 fennállása esetén mindkét szórásnégyzet $\left(p\sigma_{AO}^2 + \sigma^2\right)\chi^2/\nu$ eloszlású. Hasonlóan a $\sigma_B^2 = 0$ nullhipotézist (a kezelôk között nincs különbség) az s_O^2 / s_{AO}^2 aránnyal vizsgáljuk. Az AO kölcsönhatás nagyságának vizsgálatára az s_{AO}^2 / s_R^2 arány alkalmas. Az ismétlések varianciájának becslése a szokásos error szórásnégyzet:

$$\hat{\sigma}_{\mathrm{ism}}^2 = s_{\mathrm{error}}^2 = MS_{\mathrm{error}} = \frac{SS_{\mathrm{error}}}{\nu_{\mathrm{error}}} = \frac{\sum_{i}^{r}\sum_{j}^{q}\sum_{k}^{p}\left(x_{ijk} - \bar{x}_{ij}\right)^2}{rq\left(p-1\right)},$$