

## IV.

### A matematikai statisztika alkalmazásának lehetőségei

A testnevelő tanár egy tanévre szóló tanmenet alapján dolgozik. A tanári munka egyik fő feladata a tanulók teljesítményének mérése. De nemcsak képességeikről, készség szintjükéről kell meggyőződni, hanem fokozni kell tervszerűen kidolgozott tanmenet segítségével a tanulók fizikai teljesítményét. Mérések révén különböző adatokhoz jutunk. Ezeket csoportosítjuk, a matematikai statisztika segítségével rendezzük, elemezzük.

Tervezésünk helyességéről, az alkalmazott módszereink megfelelő voltáról is csak mérések és számítások segítségével győződhetünk meg.

Mielőtt a testnevelésben használt konkrét adatok feldolgozására rátérnénk, ejtsünk néhány szót a matematikai statisztikáról.

#### 1. Statisztika

Statisztikán azt a tudományos kutatási módszert értjük, amelynek célja adatok gyűjtése, elemzése és értelmezése. A statisztikát majdnem minden tudományban, az élet szinte minden területén használják. A normatív statisztikai módszer a testnevelés egyik legfontosabb kutatási módszere. Kettős szerepe van, egyrészt tapasztalati adatok leírása, melynek segítségével óriási tényhalmazt egészében tekinthetünk át, másrészt olyan szabályok felállítása, amelyek lehetővé teszik, hogy korlátozott számú megfigyelésből egy olyan nagyobb sokaságra következtethessünk és általánosíthassunk, amelynek csak egy részét figyeltük meg.

*A statisztikai adatok értelmezéséhez néhány statisztikai alapfogalom ismerete szükséges.* Először is nézzünk meg két olyan fogalmat, amelyeket fontos tisztán látni. *A kutatással foglalkozó személy egy mintával, az elemek egy kisebb csoportjával dolgozik, amelyből következtetéseket von le egy populációra, sokaságra vonatkozóan.*

##### 1.1. Populáció

A közös tulajdonság alapján egymáshoz tartozó egyedek összessége. A testnevelők és edzők számára a populáció az általános és középiskolai tanulók, sportolók összességét, óriási halmazát jelenti. Ide tartoznak a valaha élt és a jövőben élni fogó hasonló korú tanulók és sportolók is.

##### 1.2. A minta

A populációnak az az általunk kiválasztott része, akik az elvégzendő vizsgálat vagy vizsgálatok alanyai. A minta a populáció egyes elemeinek véletlen kiválasztásával keletkezik. Amikor a testnevelő tanár egy vagy néhány osztály, vagy egy egész iskola, tehát egy kisebb létszámú populáció teljesítményét akarja mérni, és valamennyi egyén eredményét vizsgálja, *teljes körű eredménymérést* végez. A teljeskörű eredménymérésnek az a célja, hogy a mérésben résztvevőkről általános képet kapjon. Az ilyen felmérésből csak a mérésben résztvevők körére szabad következtetéseket levonni. Ezen túlmenő általánosítás hibás lehet.

Ha olyan nagy populációt kell felmérni, amely szinte lehetetlen feladatot jelent, szokás alkalmazni az úgynevezett *reprezentatív mérést*. Ennek lényege, hogy bizonyos szabályok megtartása esetén viszonylag kis számú résztvevő eredménye alapján következtethetünk az egész nagy populációra. A kevés résztvevőt nevezük mintának. A nagyobb elemszámú minta megbízhatóbb információt nyújt, de nagyobb idő- és anyagi ráfordítást igényel, míg a kis minta esetén bizonyítani kell a levont következtetések megbízhatóságát.

A megbízhatóság azt jelenti, hogy megismételt mérésnél vagy kísérletnél ugyanazt az eredményt kapjuk.

*Ha a minta jellemző az alapsokaságra, akkor a mintát reprezentatívnak nevezük.*

### 1.3. Elemszám

A mintát alkotó egyének (elemek) száma. **Jelölése**  $n$  (minta),  $N$  (alap-sokaság). A minta elemszáma függ a populáció nagyságától, általában 30-nál nagyobbra választják, és lehetőleg páratlannak veszik.

### 1.4. Valószínűségi változó

Az azonos módon és azonos módszerekkel rögzített, a véletlentől függő számértékeket felvevő változót nevezzük valószínűségi változónak. Jelölése-  $x$ ;  $y$ ;  $z$ ; vagy  $x_1; x_2; \dots; x_n$

A testnevelési tantárgytesztekben az egyes próbákban teljesített eredmények adatai (pl. az ugrás távolsága, a futások időeredményei, labdapasszolások száma, stb.) jelentik ezeket a valószínűségi változókat.

### 1.5. Szélsőértékek

Az összes adat közül a legkisebb és legnagyobb érték (a leggyengébb és a legjobb eredmény).

### 1.6. Középértékek

Azok az értékek, amelyek az adatok központosultságát mutatják. Ezek **az átlag, a medián és a módus**.

#### 1.6.1. Átlag

$\bar{x} = \frac{\sum x}{N}$ , azaz az adatok összege osztva az adatok számával. Ez a középérték érvényes, ahol minden

adattal dolgoznak, ahol az  $N \leq 30$ . Olyan esetekben, amikor a vizsgált adatok száma túl nagy, nem gazdaságos a korábbi módszer, ott **osztályba soroljuk** az adatokat. Ezzel a módszerrel majd a későbbiekben fogunk foglalkozni. Az átlag, vagy számtani közép - csak absztrakt mutatója a középértékeknek, lehet hogy a valóságban ez az adat nem is szerepel a felsorolt értékek között.

#### o Medián

Egy másik fajta középérték - az adatok nagyság szerinti sorában pontosan középpont elhelyezkedő adat.

#### o Módus

A harmadik fajta középérték - az az érték az adatok közül, amelyik a leggyakrabban fordul elő az egyes értékek között.

A három középértéknek más - más a szerepe. Legkevésbé megbízható a módus, mert az értékek sorában több gyakorisági csomópont is lehet, ezért csak homogén értéksorban ad jó jellemzést. A számtani középértéket a szélsőséges értékek nagyon befolyásolják. Egy kiugró jó vagy rossz eredmény eltorzítja a számtani középértéket. (pl. a 2, 18, 20, 24, 26, sorban a számtani közép = 18, a medián=20; de a 18, 20, 24, 26, 54, 61, 72 sorban, ahol a medián=26 torz, az átlag=39,3 jobban jellemzi a sort).

A középértékek önmagukban nem mindig adnak hű képet az adatsorról, csak homogén adatsor esetén nyújtanak megbízható jellemzést. Az adatok nagyobb eltérései esetén az eredmények hűbb és teljesebb jellemzéséhez ismerni kell az értékek határeseteit, **szóródásait** is. A szóródások azt mutatják meg, hogy a különféle értékű eredmények hogyan helyezkednek el a számtani középérték körül.

Tegyük fel, hogy két osztályban a kislabdahajításban mért eredmények számtani középértéke azonos. Az azonos átlag ellenére a két osztály teljesítménye alapvetően különböző jellegű lehet.

Az egyik osztályban lehet, hogy a tanulók többsége a középérték körüli eredményt érte el. A másik osztályban viszont lehet, hogy ez az eredmény elsősorban a nagyon kiváló és nagyon gyenge hajítási távolságok átlagaként jött létre.

Ha csak a középértéket néznénk, nem kapnánk hű képet a tanulók teljesítményéről. Igen fontos ezért az eredmények szóródását is vizsgálni.

### 1.7. Szórás

Szórás = standard deviáció. Jele  $s$  vagy **SD**, vagy  $\sigma$ . Megadja az adatok átlag körüli tömörülését.

$$s = \sqrt{\frac{\sum D^2}{N}}, \text{ ahol } D = x - \bar{x}, \text{ azaz az egyes eredmények átlagtól való eltérései.}$$

A tantárgytesztekben történő mérések alkalmával azt tapasztaljuk, hogy a legtöbben átlagos eredményt érnek el, vannak jobb és rosszabb eredményt elérők is, esetleg van kevés kiváló vagy egészen gyenge. Ha feltételezzük a normál eloszlást, akkor az adatok 68%-a az  $\bar{x} - 1s \leq x \leq \bar{x} + 1s$  közé esik, ezekre az adatokra vonatkozik közvetlenül az átlag. Pontosabb %-os megoszlást adnak meg egyes szerzők azzal,

hogy N helyett N-1 el osztanak. *Minél kisebb az „s” értéke, annál jobban képviseli az átlag a csoport teljesítményét.*

### 1.8. Gyakoriság

Azt mutatja meg, hogy a mérési eredmények egyes értékei hányszor fordulnak elő a mintában. Ha a két szélsőérték különbsége nagy, vagy nagy az elemszám, akkor az adatokat több értéket magába foglaló osztályokba, intervallumokra bontjuk. Ebben az esetben a gyakoriság az adott osztályban lévő adatok számával egyezik meg. A gyakoriságok összege egyenlő a minta elemszámával.

### 1.9. Eloszlás

A teljesítmények vizsgálatánál azt tapasztaljuk, hogy a legtöbb eredmény a középpérték körül tömörül, a kiváló és gyenge eredmények száma a szélsőérték felé haladva fokozatosan csökken. Ezt az igen gyakori eloszlását a teljesítményeknek *normáleloszlásnak* nevezzük. Ha ezt az eloszlást grafikusán is ábrázoljuk, kapjuk a Gauss- féle haranggörbét.

Erre az eloszlásra az jellemző, hogy az adatok eloszlása szimmetrikus, azaz a gyakorisági görbe bal- és jobb oldali ágának lefutása azonos vagy legalább erősen hasonló.

Ilyen Gauss- görbét kapunk tömegvizsgálatoknál, azonos életkorú tanulók valamely teljesítményének mérésénél, abban az esetben, ha a feladat se nem nehéz, se nem könnyű.

Ha a feladat túl könnyű, akkor a görbe negatív irányba, ha túl nehéz, akkor a pozitív irány felé tolódik el. Negatív eltolódás esetén a medián (az adatok sorában pontosan középpüth elhelyezkedő adat) mindig nagyobb mint az átlag, míg pozitív eltolódás esetén az átlag nagyobb mint a medián.

A teljesen kisimult, szabályos haranggörbét csak végtelen nagy populáció adataiból "ideális " esetben kapjuk. Véges számú eset csak megközelíteni képes.

1. ábra

2. ábra

3. ábra

A gyakorlati testnevelési sportpróbák anyagát úgy kell meghatározni, hogy azok a tanulók motoros képességeinek fejlődéséről minél megbízhatóbb képet adjanak, azaz egyaránt legyenek benne nehéz és könnyű elemek, hogy az eloszlás minél inkább megközelítse az ideális normáleloszlást.

Normáleloszlás esetén az adatok úgy szóródnak, hogy az átlagtól egy szórásnyi távolságon belül lefelé és felfelé is az adatok 68%-a helyezkedik el, egy hatoda a szóráson kívül felül és egy hatoda a szóráson kívül alul, és körülbelül az adatok 2,3%-ának távolsága az átlagtól meghaladja a kétszeres szórástávolságot, valamint 0,1%-uk háromszoros szórásnál is nagyobb távolságra található az átlagtól.

## 2. Próbalap

Térjünk át konkrét, a testnevelésben használt adatok feldolgozására.

A következőkben egy az általános - és középiskolában egyaránt használható 8 teljesítménypróbát tartalmazó testnevelési próbalap- mintát mutatunk be.

A felméréseket minden tanévben ősszel és a következő év tavaszán végezzük el, így az ősszel kapott adatok fogják a kezdeti értékeket, a tavasziak pedig a végső értékeket jelenteni.

### 2.1. A próbák leírása

#### 2.1.1. 60 méteres síkfutás

A tesztelés az atlétika rövid távú futására vonatkozó szabályainak megfelelően történik. A tanulók térdelőrajttal induljanak!

Tornacipőben egy alkalommal fussanak! Az eredményeket 0,1 mp-es pontossággal rögzítsük!

#### 2.1.2-4. 600m, 800m vagy 1000m-es síkfutás

Állórajttal, tornacipőben fussanak! Egy kísérlet történjen, melynek eredményét 1mp-es pontossággal rögzítsük, majd számoljuk át perc és másodperc értékre!

#### 2.1.5. Súlypontemelkedés (Sargent- ugrás)

Szükséges eszköz egy falra szerelhető egymástól 1cm-re lévő elmozdítható lapocskákból álló mérőoszlop.

Végrehajtás: A tanuló oldalt áll a falhoz. Leolvassuk a fal felőli felnyújtott kar keze középső ujjának azt a legnagyobb magasságát, amelyet a talajon állva sarkának felemelése nélkül el tud érni. Ezt követően páros lábról karlendítéssel helyből felugrást végez, és a fal felőli kézzel elmozdítja minél magasabban a lapocskákat. Az ugrás magasságát a felugrással elért és a talajtól mért érintőleges magasság különbsége adja cm-ben.

#### 2.1.6. Távolugrás helyből vagy rohammal

Az atlétikai helyből ugrások, ill. a rohammal történő ugrások szabályainak megfelelően, de csak két kísérlettel végeztessük!

Az eredményeket 10 cm-es pontossággal cm-ben rögzítsük!

#### 2.1.5. Tömöttlabdadobás (3-4 kg-os labdával hátra)

A tanulók a dobás irányára háttal álljanak fel! A labdát tartsák mellső középtartásban! Előrehajlítással szerzett lendülettel felegyenesedve a fej fölött hátrafelé dobják a labdát! Az atlétikai dobóversenyek szabályainak megfelelően a dobóvonalról történjen a dobás. A vonalra rálépni vagy azon túllépni nem szabad. Minden próbázó 3 dobást hajt végre. Eredménye a legjobb szabályos dobás, amit 10cm-es pontossággal rögzítsünk!

#### 2.1.8. Kislabdahajítás

Az atlétikai dobóversenyek szabályainak megfelelően. Az eredményeket m-es pontossággal jegyezzük!

**2.1.9. Négyütemű szabadgyakorlat** - A rajzon lévő szabadgyakorlatot egy percig folyamatosan szabályos guggolótámaszok-kal és mellső fekvőtámaszokkal kell végrehajtani! Az eredményt az egy perc alatt végrehajtott szabályos gyakorlatok száma adja (db).

4. ábra

#### 2.1.10. Labdapasszolás (kosárlabdával)

A faltól 2m távolságra meghúzott vonal mögül, tetszőleges dobásformával kell a kosárlabdát a falra dobni és utána elkapni. A gyakorlatot 30 mp-ig kell folyamatosan végrehajtani, a labdát nem szabad elejteni. A labdapasszolások száma adja az eredményt (db).

### 3. Az adatok elrendezése, feldolgozása

#### 3.1. Osztályba nem sorolt adatok feldolgozása

##### 3.1.1. Közéértékek

A tantárgytesztekben végzett felmérések közül a *helyből távolugrás* eredményét dolgozzuk fel.

Létszám: 31 fő. A táblázat tartalmazza az őszi (kezdeti) és a tavaszi (végső) eredményeket is. Az ugrásokat 10 cm-es pontossággal cm-ben rögzítettük.

<u>Sorszám:</u>	<u>kezdeti:</u>	<u>végső:</u>
1.	160	170
1.	180	200
1.	200	220
1.	200	210
1.	200	220
1.	160	170
1.	170	160
1.	170	180
1.	180	190
1.	200	200
1.	170	170
1.	170	210
1.	180	180
1.	160	170
1.	180	190
1.	220	210
1.	210	230
1.	230	250
1.	200	230
1.	150	150
1.	180	180
1.	220	240
1.	150	160
1.	170	190
1.	200	220
1.	180	200
1.	210	230
1.	180	190
1.	200	240
1.	170	170
1.	160	170

A kezdeti adatokat vizsgálva:  $N=31$

1. A legkisebb érték:  $x_{\min} = 150\text{cm}$ , a legnagyobb érték:  $x_{\max} = 230\text{cm}$

A szélsőértékek különbsége  $230-150=80\text{cm}$ , az adatok e 80 cm-es sávban szóródnak.

2. Szeretnénk az osztály teljesítményét egyetlen adattal jellemezni  $\bar{x} = \frac{\sum x}{N} = \frac{5710}{31} = 184,19 \approx 184$

Az átlag önmagában képviselheti az egész osztályt, és lehetőséget nyújthat más osztályok átlagteljesítményeivel való összehasonlításra. (Ez csak akkor igaz, ha a szórás értéke kicsi.)

Az adatokat megfigyelve (mivel 10cm-es pontossággal mértünk) ezt az átlagos teljesítményt egy tanuló sem érte el, 19 -en rosszabb, 12-en pedig jobb eredményt nyújtottak.

3. Keressük meg az adatok nagyságrendi sorában pontosan középpüti álló értéket - a **mediánt!**

A 16. elemet kell vennünk.

150	II
160	IIII
170	IIII I
180	IIII II
190	-
200	IIII II
210	II
220	II
230	I

A vonalkák az egyes adatok gyakoriságát jelzik. A 16. helyen a **180cm** áll.

A **módus = 180cm** -a leggyakrabban előforduló adat- gyakorisága 7.

(A fenti példában a 200 cm-hez is ugyanaz a gyakorisági érték tartozik. Ebben az esetben a számítógép automatikusan a kisebb csomópontot jelöli modusként.)

A három középpérték nem egyezik meg, nem adnak helyes képet az osztály teljesítményéről. Az átlag esetleges torz információjára még egy jobban szembeütő eset:

Tételezzük fel a következő helyzetet: egy 20 főnyi osztályban a mellső fekvőtámaszban végzett karhajlítás és nyújtás teljesítményt mérjük. Egy tanuló 41-t, míg a többi 19 csak 1-1 karhajlítást és nyújtást végez.

Az adatok összege  $\sum x = 1 \cdot 41 + 19 \cdot 1 = 60$ , az átlag  $\bar{x} = \frac{60}{20} = 3$

Ha ezzel jellemeznénk az osztályt, akkor azt mondhatnánk, hogy átlagosan 3 karhajlítást és nyújtást végeznének, holott ilyen eredményt egy tanuló sem ért el. Ebben az esetben az átlag hamisan tájékoztat bennünket.

### 3.1.2. Szórás

Az átlag megbízhatóságát, azt, hogy valóban a mintára jellemző mérőszám-e, a szórás értéke tükrözi. Az **átlag annál jobban képviseli a mintát, minél kisebb a szórás értéke**. A szórás azt mutatja meg, hogy az adatok kb. 68%-a milyen nagy területen helyezkedik el az átlag, azaz a 184cm körül. A felmérés eredményeit (csak a kezdeti értékeket) a 3.1.2./1. sz. táblázat tartalmazza.

$$s = \sqrt{\frac{\sum D^2}{N}}, \quad \text{ahol a } D = x - \bar{x}$$

Ssz.	x	D = x - $\bar{x}$	D <sup>2</sup>
1.	160	-24	576
2.	180	-4	16
3.	200	16	256
4.	200	16	256
5.	200	16	256
6.	160	-24	576
7.	170	-14	196
8.	170	-14	196
9.	180	-4	16
10.	200	16	256
11.	170	-14	196
12.	170	-14	196
13.	180	-4	16
14.	160	-24	576
15.	180	-4	16
16.	220	36	1296
17.	210	26	676

3.1.2./1.sz.táblázat

18.	230	46	2116
19.	200	16	256
20.	150	-34	1156
21.	180	-4	16
22.	220	36	1296
23.	150	-34	1156
24.	170	-14	196
25.	200	16	256
26.	180	-4	16
27.	210	26	676
28.	180	-4	16
29.	200	16	256
30.	170	-14	196
31.	160	-24	576

$$\Sigma D^2 = 13816$$

$$s = \sqrt{\frac{13816}{31}} = 21,06\text{cm} \quad \bar{x} - s = 184 - 21 = 163 \quad \text{és} \quad \bar{x} + s = 184 + 21 = 205$$

Az eredmény azt mutatja meg, hogy a 163cm és a 205cm-es intervallumba esik az adatok kb. 68%-a.

### 3.1.3. Variációsázalék

Nem állapítható meg pl., hogy a távolugrás:

$\Sigma y^2 = 8400$ -es eredmény, vagy a 3000m -es síkfutás

11:40 ± 17sec -os időeredmény esetén képviseli -e jobban az átlag a mintát.

Az átlag és szórás érték nem teszi lehetővé a különböző mértékegységekben mért adatok összehasonlítását. A szórások összehasonlításának alapjául a **variációsázalék** - az úgynevezett Pearson- féle együttható szolgál.

$$v = \frac{s}{\bar{x}} \cdot 100\%$$

Ennek segítségével összehasonlítható egy csoport két különböző tesztben elért teljesítménye.

Az előző osztály átlagos **helyből távolugró** eredménye és szórása:

$$x_t = 184 \text{ cm} \quad \text{és} \quad s_t = 21.06 \text{ cm}$$

Egy másik tesztben, pl. a négyütemű szabadgyakorlat végrehajtásában ugyanez az osztály nyújtson

$$x_{sz} = 24\text{db} \quad \text{és} \quad s_{sz} = 3.79\text{db -os teljesítményt.}$$

A fenti képlet alapján adódik :

$$v_t = \frac{21,06}{184} \times 100\% = 11,44\% \quad \text{és} \quad v_{sz} = \frac{3,79}{24} \times 100\% = 15,79\% \quad \text{a két tesztben a variációsázalékok.}$$

Annak ellenére, hogy az  $s_t$  értéke majdnem 6-szorosa a  $v_{sz}$ -nek, nem azt jelenti, hogy a szórás is 6-szoros, mert az  $\bar{x}_t$  átlag is több mint 7-szerese az  $\bar{x}_{sz}$  - nak.

$$\frac{v_t}{v_{sz}} = \frac{11,44\%}{15,79\%}$$

Ezzel az arányszámmal megszüntettük a dimenziókat, és mértékegység nélküli, abszolút értékeket hasonlíthatunk össze. Azt nem tudjuk, hogy a 21,06cm vagy a 3,79db-e a több, de a 11,44% és a 15,79% közötti különbség már értékelhető.

### 3.2. Az adatok osztályba sorolása

Az eddigiek folyamán 2 átlagos osztály tanulóinak adatait dolgoztuk fel. E fiktív osztályok létszámai 31 - 31 fő volt. Egy testnevelő tanár azonban nem csupán egy - két osztályt tanít, hanem -a legtöbb esetben- az iskola valamennyi osztályában van órája, azaz több 100 tanítvánnyal is foglalkozik. Számára az is fontos lehet, hogy akár a párhuzamos osztályokban, akár az iskola valamennyi osztályában mért, esetleg más is-

kolákból kapott adatokat is rendszerezze, értékelje. Ilyenkor olyan mennyiségű adathalmaz áll rendelkezésre, amelyeket az előzőekben ismertetett módon már nem képes feldolgozni.

Az I. osztályosok eredményei a helyből távolugrásban (tavaszi felmérést 3.2./1. sz. táblázat tartalmazza)

3.2./1. sz. táblázat

Sorszám	x	Sorszám	x	sorszám	x
1.	150	31.	220	61.	220
1.	150	32.	160	62.	160
1.	170	33.	240	63.	180
1.	180	34.	200	64.	180
1.	180	35.	250	65.	190
1.	200	36.	210	66.	200
1.	200	37.	170	67.	170
1.	170	38.	170	68.	210
1.	180	39.	220	69.	230
1.	170	40.	200	70.	230
1.	200	41.	210	71.	230
1.	200	42.	230	72.	170
1.	170	43.	250	73.	220
1.	200	44.	240	74.	240
1.	180	45.	230	75.	170
1.	230	46.	200	76.	200
1.	220	47.	220	77.	160
1.	170	48.	230	78.	180
1.	160	49.	260	79.	190
1.	160	50.	250	80.	170
1.	150	51.	180	81.	210
1.	170	52.	180	82.	170
1.	180	53.	180	83.	170
1.	190	54.	160	84.	180
1.	200	55.	170	85.	180
1.	220	56.	210	86.	230
1.	180	57.	200	87.	170
1.	190	58.	210	88.	180
1.	190	59.	160	89.	210
1.	210	60.	200	90.	220

Az adatokat osztályba kell sorolni. Ennek érdekében számítsuk ki a szélsőértékek különbségét.

$$x_{\max} = 250\text{cm} \quad x_{\min} = 150\text{cm} \quad x_{\max} - x_{\min} = 260\text{cm} - 150\text{cm} = 110\text{cm}$$

Legyen az osztályok száma 10, így az osztálynagyságok  $\frac{110\text{cm}}{10} = 11\text{cm}$

1. osztály	150-161	IIII IIII
2. osztály	161-172	IIII IIII IIII I
3. osztály	172-183	IIII IIII IIII
4. osztály	183-194	IIII
5. osztály	194-205	IIII IIII III
6. osztály	205-216	IIII III
7. osztály	216-227	IIII III
8. osztály	227-238	IIII III
9. osztály	238-249	III
10. osztály	249-260	III

Az adatok osztályokba való beosztásánál minden osztály után, ahová az illető adatot besoroljuk, egy rovátkát húzunk., és ötösével csoportosítjuk, hogy könnyebb legyen összeadni őket. A rovátkákat összeadva kapjuk az egyes osztályok gyakoriságát, azaz azt, hogy az illető osztályban hány adat van. Ha egy adat pont az osztályhatárra esik, akkor mindig a magasabb osztályhoz vegyük. PL: ha lenne 194cm-es adat, akkor ennek az 5. és nem a 4. osztályba kellene tartoznia.



Az osztályba sorolás után az egyes adatok elvesztik egyéni értéküket, és az egyes adatokat az illető osztály közepének segítségével fejezzük ki. Osztályba sorolt adatok esetén az átlag

$$\bar{x} = \frac{\sum F \cdot M}{N}, \text{ ahol } F \text{ a gyakoriság, az osztályközép, és } N \text{ az adatok száma.}$$

Ezután azt az osztályt, amelyikbe az átlag tartozik,  $D=0$  rangúnak véve megállapítjuk a többi osztály ettől való eltérését. A rangokat, az eltéréseket + és - egészszel jelöljük egyenként. (pozitívnek tekintjük azt az osztályt, amelybe az átlagnál jobb értékek tartoznak).

A különbségeket hangsúlyosabbá téve négyzetre emeljük  $D^2$ . Ezt követően az osztályok eltérésének négyzeteit súlyozzuk, azaz a  $D^2$ -t megszorozzuk a gyakorisággal,

$$F \cdot D^2, \text{ és kiszámítjuk a } \sum F \cdot D^2 \text{-t.}$$

A szórás ( standard deviáció ) a következőképpen számítható:

$$s = \sqrt{\frac{\sum F \cdot D^2}{N}} \cdot \text{osztálynagyság}$$

Nézzük meg az előző példa alapján milyen szórással bír az első osztályosok távolugró adatsora (3.2./2.sz.táblázat tartalmazza)

3.2./2.sz.táblázat

osztály	M oszt. közép	F gyakoriság	F • M	D	D <sup>2</sup>	F • D <sup>2</sup>
150-161	155,5	10	1555,0	-4	16	160
161-172	166,5	16	2664,0	-3	9	144
172-183	177,5	15	2662,5	-2	4	60
183-194	188,5	5	942,5	-1	1	5
194-205	199,5	13	2593,5	0	0	0
205-216	210,5	8	1684	+1	1	8
216-227	221,5	8	1772,0	+2	4	32
227-238	232,5	8	1860,0	+3	9	72
238-249	243,5	3	730,5	+4	16	48
249-260	254,5	4	1018,0	+5	25	100
			$\sum F \cdot M = 17482$			$\sum F \cdot D^2 = 629$

$$\bar{x} = \frac{\sum F \cdot M}{N} = \frac{17482}{90} = 194,24 \approx 194 \quad s = \sqrt{\frac{F \cdot D^2}{N}} \cdot 11 = \sqrt{6,99} \cdot 11 = 29 \text{ cm}$$

#### 4.4. Statisztikai hipotézisek és ellenőrzésük

Statisztikai hipotézisen egy, az alapeloszlás paramétereire, vagy magára az egész alapeloszlásra vonatkozó feltevést értünk.

A statisztikai hipotézisek bizonyítására, ellenőrzésére úgynevezett statisztikai próbákat fejlesztettek ki, amelyeket általánosan **hipotézisvizsgálatoknak nevezünk**.

A hipotézisvizsgálat első feladata a nullhipotézis képzése, amelyben az állítás jut kifejezésre, hogy az alapeloszlás paramétere és annak feltételezett értéke, vagy a tényleges és a feltételezett alapeloszlás között nincs különbség. PL. egy kísérleti csoport és egy ellenőrzőcsoport (kontroll) alapképességeit vizsgálva a kísérletre nézve nincs különbség.

A statisztikai próbák arra adnak választ, hogy a minták vagy csoportok közötti eltérések mennyire a véletlen és mennyire a kísérlet, vagy új módszer javára írhatók. Így állapítjuk meg, hogy az eredmény szignifikáns -e vagy sem. **A szignifikancia a nullhipotézistől való tényleges eltérés fokát jelzi.**

**Szignifikanciát kapunk akkor, ha a kapott érték a nullhipotézissel össze nem egyeztethető.** A valószínűség, amelyet  $p$ -vel jelölünk, azt adja meg, hogy sok kísérletet elvégezve, milyen valószínűséggel következik be az esemény.

A szignifikancia fokát e valószínűség határozza meg:

$p > 5\%$	nem szignifikáns	
$p \leq 5\%$	szignifikáns	( * )
$p \leq 1\%$	erősen szignifikáns	( ** )
$p \leq 0,1\%$	igen erősen szignifikáns	( *** )

Más könyvekben a szignifikanciaszintek jelzésére a ( ) -ben lévő jeleket is használják.

A valószínűséget, ezáltal a szignifikanciát befolyásolja az elemszám is, pontosabban az az elemszám, amely megmutatja, hogy egy adathalmaz konkrét értékeinek megismerését hány szabadon választható információ közlésével határozhatjuk meg. Ezt **szabadságfok**-nak nevezzük.

Megkülönböztetünk paraméteres és nem paraméteres próbákat. A paraméteres próbák közül a " T " és az " F " próbát tárgyaljuk

#### 4.1. Egymintás „t” próba, önkontrollos vizsgálat

A tantárgytesztek felmérései egy tanéven belül ősszel és tavasszal történnek. A tanári munka egyik célja, hogy a tanítványok teljesítményében fejlődés mutakozzon.

Jegyzetünk statisztikai eredménye, adatsora két fiktív ( képzelt ) osztályt vesz alapul, azt a célt szolgálja csak, hogy segítségével a mérési lehetőségeket bemutassuk.

Az egymintás T próba vagy másképpen önkontrollos vizsgálat segítségével azt ellenőrizzük, történt-e változás a tanév során, (esetleg egyik tanévről a másikra) azaz önmagukhoz viszonyítva tanulóink fejlődtek-e, és ha igen, fejlődésük jelentősnek (szignifikánsnak) tekinthető-e?

A láb dinamikus erejének növelésére több ugrógyakorlatot végeztek az év során a tanulók, mint pl. helyben és mozgás közbeni felugrásokat, sorozatugrásokat, súlyzós ugrógyakorlatokat. Azt vizsgáljuk az önkontrollos vizsgálat, hogy hatékonyaknak bizonyulnak-e ezek a gyakorlatok a dinamikus láberő fejlesztésére.

Feltételezésünk szerint a nullhipotézisünk az, hogy **nem befolyásolja**

Tekintsük a már korábban megismert felmérési adatokat:

A következő táblázatban a **helyből távolugrás** esetén:

sorszám	kezdeti $x_1$	végso $x_2$	változások $x = x_2 - x_1$	változások négyzete $x^2 = (x_2 - x_1)^2$
1.	150	160	10	100
1.	150	150	0	0
1.	170	170	0	0
1.	180	180	0	0
1.	180	190	10	100
1.	200	200	0	0
1.	200	220	20	400
1.	170	180	10	100
1.	180	190	10	100
1.	170	190	20	400
1.	200	210	10	100
1.	200	220	20	400
1.	170	160	-10	100
1.	200	240	40	1600
1.	180	200	20	400
1.	230	250	20	400
1.	220	210	-10	100

1.	170	170	0	0
1.	160	170	10	100
1.	180	180	0	0
1.	180	190	10	100
1.	180	200	20	400
1.	160	170	10	100
1.	170	210	40	1600
1.	210	230	20	400
1.	200	230	30	900
1.	210	230	20	400
1.	160	170	10	100
1.	200	220	20	400
1.	220	240	20	400
1.	160	170	10	100
			$\Sigma x = 390$	$\Sigma x^2 = 9300$
			változások összege	vált. négyzetösszege

a változások átlaga :  $\bar{x} = \frac{\Sigma x}{N} = \frac{390}{31} = 12,58$  jelenti az egy tanulóra eső változás mértékét (cm) .

$$(\Sigma x)^2 = 390^2 = 152100$$

$$\text{és a szórás } s = \sqrt{\frac{\Sigma x^2 - \frac{(\Sigma x)^2}{N}}{N-1}} = \sqrt{\frac{9300 - \frac{152100}{31}}{30}} = \sqrt{146,45} = 12,1 \text{ cm}$$

$$t = \frac{\bar{x}}{s} \cdot \sqrt{N} = \frac{12,58}{12,10} \cdot \sqrt{31} = 5,79$$

Mindezt a következőképpen számítottuk ki:

1. a végső eredményekből kivontuk a kezdeti értékeket, azaz meghatároztuk az egyes változásokat. 23 tanulónál pozitív, 2-nél negatív, 6-nál nem hozott változást. Kiszámoltuk a változások összegét, ügyelve az előjelekre

$$\Sigma x = \Sigma (x_2 - x_1)$$

o Kiszámoltuk a változások átlagát, azaz az előbbi összeget elosztottuk az adatok számával

$$\bar{x}$$

3. A közös szórás értékének kiszámításához a változások minden egyes értékét külön- külön négyzetre emeltük, és ezeket is összeadtuk.

$$\Sigma x^2$$

4. Ugyanakkor a változások összegét is négyzetre emeltük és elosztottuk az elemek számával:

$$\frac{(\Sigma x)^2}{N}$$

5. Kiszámítottuk az

$$s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N-1}} \quad \text{értékét.}$$

6. A "t" értéket úgy kaptuk meg, hogy a változások átlagát elosztottuk a szórás s értékével és megszoroztuk a felmérésben részt vett tanulók számának négyzetgyökével

$$t = \frac{\bar{x}}{s} \cdot \sqrt{N}$$

Az s képletben szereplő N-1 a szabadságfok.

A kapott eredményt a "t" próba értékeit tartalmazó táblázatból magyarázzuk tovább.

A táblázat felső sorában a p valószínűséget jelzik, a függőleges oszlopban az egyes szabadságfokokat, amelyek az adott elemszámnál mindig eggyel kevesebbek. E táblázatból állapítjuk meg, hogy a kezdeti és végső értékeket figyelembe véve a különbség szignifikáns-e vagy sem. Szignifikanciát állapítunk meg, ha a próba elvégzése után kapott érték a nullhipotézissel statisztikailag reálisan össze nem egyeztethetőnek látszik.

A táblázat N-1=30 sorában 3,646 található. Mivel a mi "t" értékünk 5,79 nagyobb a 3,646-nál, ebbe az oszlopba tartozik, mert amennyiben a kapott t érték eléri vagy meghaladja az 5%-os (0,05) 1%-os (0,01) vagy az 1 ezrelékes (0,001) oszlop és a szabadságfoknak megfelelő sor metszéspontjában lévő értéket, akkor azt mondjuk, hogy a  $p \leq 5\%$ -nál szignifikáns, a  $p \leq 1\%$ -nál erősen szignifikáns és hogy a  $p \leq 0,1\%$ -nál igen erősen szignifikáns.

Amennyiben a  $p > 5\%$ -nál, a fejlődés nem szignifikáns. A helyből távolugrás esetén a  $t=5,79$  és az  $N-1=30$  mellett igen erősen szignifikáns, tehát nem egyeztethető össze a nullhipotézissel, vagyis az év során alkalmazott gyakorlatok igen erősen befolyásolták a távolugrás eredmények javulását.

#### 4.2. Kétmintás "t" próba, azaz a kontrollcsoportos vizsgálat

Ezt a próbát akkor használjuk, ha egy új módszer hatékonyságát akarjuk ellenőrizni. Két különböző csoportunk (osztályunk) van, azaz egy *kísérleti* és egy *kontroll* (ellenőrző) csoport.

Mind a kísérleti, mind pedig a kontroll csoport :

- közel azonos létszámú,
- azonos nemű, életkorú
- azonos előképzettségű

Vizsgáljuk meg ismét a *helyből távolugrás* eredményeit:

sorsz.	KISÉRLETI CSOPORT				KONTROLL CSOPORT			
	kezdeti $x_1$	végső $x_2$	$x = x_2 - x_1$	$x^2$	kezdeti $y_1$	végső $y_2$	$y = y_2 - y_1$	$y^2$
1.	150	160	10	100	150	160	10	100
1.	150	150	0	0	150	150	0	0
1.	170	170	0	0	220	230	10	100
1.	180	180	0	0	180	180	0	0
1.	180	190	10	100	200	200	0	0
1.	200	200	0	0	200	200	0	0
1.	200	220	20	400	200	220	20	400
1.	170	180	10	100	210	220	10	100
1.	180	190	10	100	180	190	10	100
1.	170	190	20	400	170	190	20	400
1.	200	210	10	100	200	210	10	100

1.	200	220	20	400	200	220	20	400
1.	170	160	-10	100	230	230	0	0
1.	200	240	40	1600	200	240	40	1600
1.	180	200	20	400	180	200	20	400
1.	230	250	20	400	230	250	20	400
1.	220	210	-10	100	220	210	-10	100
1.	170	170	0	0	250	260	10	100
1.	160	170	10	100	240	250	10	100
1.	180	180	0	0	160	170	10	100
1.	180	190	10	100	180	180	0	0
1.	180	200	20	400	190	200	10	100
1.	160	170	10	100	170	180	10	100
1.	170	210	40	1600	170	210	40	1600
1.	210	230	20	400	210	230	20	400
1.	200	230	30	900	210	230	20	400
1.	210	230	20	400	210	230	20	400
1.	160	170	10	100	170	170	0	0
1.	200	220	20	400	200	220	20	400
1.	220	240	20	400	220	230	10	100
1.	160	170	10	100	180	200	20	400
$\Sigma x = 390$				$\Sigma y = 380$				
$\Sigma x^2 = 9300$				$\Sigma y^2 = 8400$				

$$\bar{x} = \frac{\Sigma x}{N_1} = \frac{390}{31} = 12,58 \qquad \bar{y} = \frac{\Sigma y}{N_2} = \frac{380}{31} = 12,26$$

$$\Sigma x^2 - \frac{(\Sigma x)^2}{N_1} = 9300 - \frac{152100}{31} = 4393,55 \qquad \Sigma y^2 - \frac{(\Sigma y)^2}{N_2} = 8400 - \frac{144400}{31} = 3741,93$$

$$s = \sqrt{\frac{\Sigma x^2 - \frac{(\Sigma x)^2}{N_1} + \Sigma y^2 - \frac{(\Sigma y)^2}{N_2}}{N_1 + N_2 - 2}} = \sqrt{\frac{4393,55 + 3741,93}{60}} = \sqrt{135,59} = 11,63$$

$$t = \frac{|\bar{x} - \bar{y}|}{s} \cdot \sqrt{\frac{N_1 N_2}{N_1 + N_2}} = \frac{12,58 - 12,26}{11,64} \cdot \sqrt{\frac{961}{62}} = 0,108$$

Kikeresve a " t " táblázatból: a 60 szabadságfok sorában a legkisebb értéket sem éri el a mi  $t = 0,108$  értékünk, ami azt jelenti, hogy  $p$  kisebb 90%, nem szignifikáns a különbség a két osztály teljesítményei között.

A kétmintás " t " próba alkalmazása csak akkor jogosult, ha az " F " próba nem ad szignifikáns eredményt.

#### 4.3. Mi is az az " F " próba?

Ebben az esetben a két adathalmaz szórásait hasonlítjuk össze. A két szórás  $s_1$  ill.  $s_2$ , a halmazok szabadságfokai

$N_1 - 1$  ill.  $N_2 - 1$ , akkor  $F = \frac{s_2^2}{s_1^2}$ , ahol az  $s^2$ -ek az egyes halmazok varianciái,

$N_1$  és  $N_2$  az első és második halmaz adatszám.

Pl.: már korábban kiszámoltuk az egymintás " t " próbával a távolugrás szórásértékét

$$s_1 = 12,10 \text{ cm}$$

Hasonló módon lehet egy másik adathalmaz (szintén 30 szabadságfokú) szórásértékét is kiszámolni, amit itt nem részletezünk. (lásd az adatokat a kétmintás " t " próba kontroll csoport )

$$s_2 = 11,17 \text{ cm}$$

$$F = \frac{12,10^2}{11,17^2} = \frac{146,41}{127,77} = 1,17 \text{ a számlálóba mindig a nagyobb szórás kerüljön!}$$

Megvizsgálva az F táblázatot: a 30-as szabadságfokú számláló (sor) és a 30-as szabadságfokú (oszlop) metszésében  $F=2,07$  áll, ez nagyobb mint az általunk kiszámított  $F = 1,15$  érték, tehát nem szignifikáns, így jogosult volt a kétmintás "T" próba alkalmazása.

## 5. KORRELÁCIÓSZÁMÍTÁS

Amikor két vagy több tulajdonság közötti összefüggést vizsgálunk, korrelációs számításhoz kell folyamodnunk. A korreláció összefüggést, függőséget jelent. A függőség azt jelenti, hogy valamely változó meghatározott értékéhez egy másik változó értéke egyértelműen hozzárendelhető.

A testsúly és testmagasság között nem mindig áll fenn összefüggés, ugyanazon magasságú egyének különböző súlyúak lehetnek. Hasonlóképpen egy meghatározott testsúlyhoz többfajta magasság tartozhat. A korrelációs számítással azt vizsgáljuk meg, hogy a tantárgytesztek közül bármely kettő között milyen az összefüggés szorossága. A korrelációs együttható " r " az alábbi tulajdonságokkal rendelkezik:

- +1 és -1 között helyezkedik el
- r nagyobb 0 esetén a kapcsolat egyirányú, vagyis minél nagyobb az egyik eredmény, annál nagyobb a másik is, r kisebb 0 esetén a kapcsolat fordított, minél nagyobb az egyik eredmény, annál kisebb a másik
- r =1, amikor a két mérési sorozat között közvetlen függvénykapcsolat áll fenn, azaz mindkét tesztben a tanulók azonos sorrendbeli teljesítményt nyújtottak
- r =0 , a változók között nincs összefüggés
- r = -1 , aki az egyikben a legjobb, az a másikban a leggyengébb, és így tovább.

A korrelációs számítás előnyei:

- Pozitív és szignifikáns összefüggés esetén következtethetünk a két mért tulajdonság közös tényezőjére. E tényező fejlesztése mindkét próbában teljesítménycsökkenést idéz elő.
- Az  $r \rightarrow 0$  arra utal, hogy a két mért tulajdonság független egymástól, így fejlesztésük eltérő módon kell hogy történjen.

### 5.1. Mértékkorreláció számítás

Vizsgáljuk meg próbatessztjeinkből azt, hogy milyen összefüggést mutatnak a medicinlabdadobás és a kislabdahajítás eredményei:

x	$x - \bar{x}$	$(x - \bar{x})^2$	y	$y - \bar{y}$	$(y - \bar{y})^2$	$(x - \bar{x}) \cdot (y - \bar{y})$
700	-50	2500	30	-8	64	400
690	-60	3600	29	-9	81	540
900	150	22500	42	4	16	600
810	60	3600	53	15	225	900
950	200	40000	52	14	196	2800
900	150	22500	45	7	49	1050
820	70	4900	49	11	21	770
720	-30	900	42	4	16	-120
710	-40	1600	38	0	0	0
500	-250	62500	22	-16	256	4000
560	-190	36100	25	-13	169	2470
700	-50	2500	30	-8	64	400
850	100	10000	42	4	16	400
600	-150	22500	35	-3	9	450
630	-120	14400	27	-11	121	1320
790	40	1600	37	-1	1	-40
590	-160	25600	24	-14	256	2240
920	170	28900	50	12	144	2040
630	-120	14400	28	-10	100	1200
1050	300	90000	60	22	484	6600
520	-230	52900	20	-18	324	4140
710	-40	1600	30	-8	64	320
900	150	22500	50	12	144	1800
870	120	14400	40	2	4	240
690	-60	3600	47	9	81	-540
640	-110	12100	30	-8	64	880
830	80	6400	34	-4	16	-320
860	110	12100	48	10	100	1100
780	30	900	50	12	144	360
840	90	8100	36	-2	4	-180
570	-180	32400	25	-13	169	2340
$\bar{x} = 750$			$\bar{y} = 38$			$\Sigma(x - \bar{x}) \cdot (y - \bar{y}) = 38160$
		$\Sigma(x - \bar{x})^2 = 577600$			$\Sigma(y - \bar{y})^2 = 3502$	

$$r = \frac{\Sigma(x - \bar{x}) \cdot (y - \bar{y})}{\sqrt{\Sigma(x - \bar{x})^2 \cdot \Sigma(y - \bar{y})^2}} = \frac{38160}{\sqrt{577600 \cdot 3502}} = 0,8484$$

A számításaink a következő lépcsőben történtek:

1. Kiszámítottuk a már ismert módon a medicinlabda-dobás (x), és a kislabdahajítás (y) értékeinek átlagait  $\bar{x}$  és  $\bar{y}$
2. Ezután mind az x mind az y oszlop minden egyes eredményéből kivontuk az átlagot, azaz kiszámítottuk minden sorban az  $x - \bar{x}$ -t, ill  $y - \bar{y}$ -t.
3. Az így kapott értékeket rendre négyzetre emeltük, és a négyzeteket összeadtuk. Kaptuk a  $\Sigma(x - \bar{x})^2 - t$ , és a  $\Sigma(y - \bar{y})^2 - t$ .
4. A két x, y oszlop átlagtól való eltéréseit összeszoroztuk, azaz kiszámítottuk soronként az

$(x - \bar{x})(y - \bar{y})$  értékeket, és ezeket is összeadtuk:  $\sum (x - \bar{x})(y - \bar{y})$

5. Kiszámítottuk a korrelációs együtthatót.

A korreláció szignifikanciáját is a Fischer-Yates táblázat segítségével határozhatjuk meg. A szignifikanciaszint u. a. , mint az egymintás "T" próba esetén:  $p \leq 5\%$ ,  $p \leq 1\%$ , ill.  $p \leq 0,1\%$ , valamint a szabadságfok is azonos  $n-1=30$ .

A táblázatban az  $n=30$  sorban a 0,5541 áll, mivel a számított  $r=0,85$  értékünk ennél nagyobb, ezért ebbe az oszlopba tartozik, amelynek valószínűsége 0,1%, azaz igen erősen szignifikáns.

### 5.2. Interkorrelációs mátrix

A többi próba esetén is kiszámíthatók a mértékkorrelációs együtthatók, velük kitölthető az úgynevezett **interkorrelációs mátrix**. A mátrix átlójában  $r = 1$  értékek állnak, mert bármely próba önmagával való összefüggése 1.

$$\text{A korrelációk száma} = \frac{\text{tesztek száma} \cdot \text{tesztek száma} - 1}{2}$$

ssz.	1	1	1	1	1	1	1	1	1	
1	1									
1		1								
1			1							
1				1						
1					1					
1						1				
1							1	0,8484		
1							0,8484	1		
1									1	
1										1

### 5.3. Rangkorreláció számítás

A mértékkorreláció számításán kívül , ahol is abszolút skálán mért mennyiségek összehasonlítását végeztük el, ismert az úgynevezett **rangkorreláció** számítás is .

Itt a korrelációt két **becsült** tulajdonság rangszámai között számítjuk ki. A rangkorreláció számítás munkamegtakarítást jelent, ha N nagyobb, mint 30. Ha N kisebb, mint 15, és a számadatok szóródása nagy, a rangkorreláció pontatlanul utal az összefüggésre.

Tegyük fel, hogy az egyik osztályban a tanulók tudásszintjét vizsgáljuk torna sportágban.

Megfigyeljük a kézállás a fejállás végrehajtásában nyújtott teljesítményüket.

Vizsgáljuk meg, milyen összefüggés van a két teljesítményben nyújtott rangsorok között:

sorszám	KÉZÁLLÁS rangszám x	FEJÁLLÁS rangszám y	$D =  x - y $	$D^2$
1.	1	1	0	0
1.	2	4	2	4
1.	4	3	1	1
1.	4	2	2	4
1.	4	7	3	9
1.	6	6	0	0
1.	8	5	3	9
1.	8	8	0	0
1.	8	9	1	1
1.	10	10	0	0



1.	12	11	1	1
1.	12	12	0	0
1.	12	14	2	4
1.	16	14	2	4
1.	16	14	2	4
1.	16	17	1	1
1.	16	17	1	1
1.	16	17	1	1
1.	20	19	1	1
1.	20	20.5	0.5	0.25
1.	20	20.5	0.5	0.25
1.	22	22	0	0
1.	23	27	4	16
1.	24	25	1	1
1.	25	26	1	1
1.	26	23	3	9
1.	27	24	3	9
1.	29	28	1	1
1.	29	31	2	4
1.	29	30	1	1
1.	31	29	2	4

$$\sum D^2 = 91,5$$

$$r = 1 - \frac{6 \cdot \sum D^2}{N^3 - N} = 1 - \frac{6 \cdot 91,5}{31^3 - 31} = 1 - \frac{549}{29760} = 0,982$$

A Spearman-féle rangkorrelációs együttható táblázatában az N-1. sorban 0,463 áll, ennél jóval nagyobb a kiszámított  $r = 0,982$ -s érték, tehát ebbe az oszlopba tartozik. A  $p = 1\%$ , azaz  $p = 0,01$ , tehát azt állapíthatjuk meg, hogy tanulóink kézállásban és fejjállásban nyújtott teljesítményei rangsora **erős szignifikanciájú** összefüggést mutat.

## 6. Grafikus ábrázolás

A statisztikai adatok szemléltetésének, népszerűsítésének közismert eszköze a grafikus ábrázolás.

A grafikus ábrák szerkesztése azonban nemcsak a szemléltetés és népszerűsítés céljait szolgálja, hanem az elemzés, az összefüggések feltárásának céljait is.

Sok esetben a grafikus ábra segít bennünket ahhoz, hogy a vizsgált jelenség természetéről teljes képet alkothassunk magunknak. A grafikus ábrázolás fontos tulajdonsága, hogy nem abszolút nagyságot, hanem arányokat érzékeltet. A grafikus ábrák statisztikai adatokat, ill. azok egymáshoz való arányát távolságokkal, területekkel vagy néha körtartalommal érzékeltetik. A grafikus ábrázolás legelterjedtebb formája a Descartes - féle koordináta-rendszer használata. Ma már e rendszerben rengeteg ábrázolási módot találtak ki. Így lehet pontthalmazt, ezeket összekötve vonal diagramot, és oszlopdiaagramot is ábrázolni. Sőt ezek kombinációi, változatai is ismertek.

A **pontdiagram** -ban a statisztikai adatokat a koordináta-rendszerben elhelyezett pontok ábrázolják. Ha ezeket a pontokat összekötjük, **vonaldiagram** -ot, vagy **gyakorisági poligon**-t kapunk.

Ha az x tengely beosztásának megfelelő szélességű téglalapokat rajzolunk, akkor **oszlopdiaagram**-ot, vagy **hisztogram**-ot kapunk. Az **oszlopdiaagram** x és y tengelyét felcserélve, az oszlopokat vízszintesen elhelyezve **szalagdiagram**-ot láthatunk.

A testnevelési kutatásmetodikában másik elterjedt szemléltetési mód a **kördiagram**. Ez egy sokaság, vagy minta valamilyen szempont szerinti összetételére ad szemléltetést. Az összetevőket a körcikkek területei érzékeltetik.

Mindig a legmegfelelőbb szemléltetési módot kell az elemzést végző egyénnek kiválasztania.

Lássunk néhány példát a különböző fajtájú grafikonokra:

### 6.1. Vonaldiagram (gyakoriság - teljesítmény)

5. ábra

a helyből távolugrás polygonja  
(31 adat esetén)

Az osztályba sorolt adatok esetén a vízszintes tengelyre az egyes számát írjuk. Itt az osztályközepek képviselik az egyes osztályok adatait, ezeket jelöljük be az X tengelyen. 90 adat esetén (6. ábra):

6. ábra

### 6.2. Hisztogram vagy oszlopdiaqram

Szerkesztése úgy történik, mint a gyakorisági grafikonoké, azzal a fontos különbséggel, hogy az egyes adatokat nem az osztályközepek képviselik, hanem azt feltételezzük, hogy az adatok az egész osztályban egyenlően vannak elosztva. Az egyes adatokat tehát olyan oszlopokkal ábrázoljuk, amelynek alapja az osztály nagysága, magassága pedig az illető osztály gyakorisága. (7. ábra)

7. ábra

### 6.3. Ogive vagy halmozó grafikon

Az adatok accumulatív (halmozó) összeadásán alapszik. Most a vízszintes tengelyre a gyakoriságot, az Y tengelyre pedig az osztályokat.

Ennek az ábrázolási módnak előnye, hogy százalékos értékeket is könnyen kiszámíthatunk vele: A grafikon legmagasabb pontjából merőlegest húzunk az X tengelyre. Ezt 10 egyenlő részre osztjuk, ezek adják a 10, 20, ... , 100 %-os értékpontok helyét.

8. ábra

A vastag vonal adja a halmozógörbét.

## 7. A számítástechnika lehetőségei a kutatómetodika területén

Mai korunk egyik legdinamikusabban fejlődő tudományágazata a számítástechnika. Nincs olyan nap, hogy ne hallanánk a számítógépről. Ugyanakkor a testnevelő tanár és edző számára is objektív és lényeges adatokat árulnak el a különböző statisztikák. A testnevelésben a matematikai statisztika játszik fontos szerepet. A sportban elsődleges cél az eredmény. Az eredményt pedig igen sok olyan tényező befolyásolja, amelyet mérni lehet. Ezek feldolgozására a legalkalmasabb módszer a számítógép használata. Ma már a testnevelő tanároknak nem szabad idegenkedniük a számítógéppel történő adatfeldolgozástól. Minden iskolában van már számítógép, ki kell használni az általa nyújtott lehetőségeket.

Ebben a fejezetben konkrét példákon keresztül szeretnénk segítséget nyújtani a leendő testnevelők számára az adatok számítógéppel történő felvitelére, elrendezésére és feldolgozására.

Előbb azonban néhány számítástechnikai fogalommal kell megismerkedni. Az egymással összefüggő adatok összessége az állomány vagy *file*. A lemezen a file-kat 8 karakterből álló név jelenti. Pl. Mat stat /a mi programunk esetén/.

Többféle file van, amelyek közül a mi számunkra a legfontosabbak az adatfile-ok. Egy ilyen file-ban tároljuk mi is a felmérésünk adatait.

A file-ok általában fix hosszúságú részekből állnak, - ezek a *rekordok*. Egy file-on belül az egy információegységhez tartozó információk összességét nevezzük rekordnak.

Pl. az egy sportolóhoz tartozó személyi adatok és sporteredményei egy rekordot alkotnak.

Végül a rekord mezőkre - *field*-ekre oszlik, melyek az információk alapegységeit képezik. Ilyen például a név, a 100m-es síkfutás időeredménye, a kislabdahajítás távolsága, stb.

A könnyű kezelés és áttekinthetőség érdekében a file-kat a számítógépben bizonyos rendszerbe célszerű osztályozni. Erre használjuk a *könyvtárakat*. A programokat nem ömlesztve tároljuk, hanem rendszerezve helyezük el a floppy-n vagy a merevlemez tárolón /winchester/.

Különböző, valamilyen szempontból összetartozó file-okat helyezünk azonos könyvtárba.

A file-ok kezelésére a DOS operációs rendszer parancsai szolgálnak. Első parancs a CD. E két betű segítségével tudunk közlekedni a könyvtárba be- és ki. A CD és az általunk adott könyvtárnév jelenti a be, a CD.. jelenti a könyvtárból való kilépést.

Minden programot saját nevének beírásával lehet elindítani. Ezek azok a parancsok, amelyeket szükséges ismerni ahhoz, hogy számítógéppel kezdjünk dolgozni. Ezek segítségével már el tudjuk indítani a programunkat.

Az általunk elkészített program IBM-PC gépen működik.

### 7.1. Adatok felvitele a számítógépre

Kapcsoljuk be a gépet és tegyük be a lemezt, fordítsuk rá a kart! A testnevelési adatok értékelését végző program indítása nevének (Mat stat) begépelésével és az Enter billentyű megnyomásával történik. Ehhez abban a könyvtárban kell lennünk, amely ezt a programot tartalmazza. A könyvtárak közötti váltás történhet vagy a fenti **DOS CD** paranccsal, vagy a **DOS** utasításokat helyettesítő **Norton Commander** használata esetén nyilakkal történő kurzormozgatások segítségével. Kezdő számítógép-használók kérjenek rövid eligazítást az adott számítógép kezeléséről!

Indítás után a képernyőn megjelenik :

#### Testnevelés kiértékelés:

<b>meglévő adatok használata</b>	<b>M</b>
<b>új adatok használata</b>	<b>U</b>
<b>válassz</b>	

Az adatok bevitele a gépbe a következő módon történt: új adatok felvétele **U** billentyűvel, beírtuk a **94RAD1EF** file nevet, amit az enter billentyű megnyomásával zártunk. /A név egyes részletei célszerűen utalnak az adatok forrására. Az általunk alkalmazott kódolás a következő: A rendelkezésre álló 8 karakterből az első kettő a felmérés évét, a következő három az iskola kódjelét jelenti, az ezután következő kettő az osztályt (pl. 1B vagy 1E), az utolsó betű a felmérésben résztvevő tanulók nemét jelenti, (F fiú, L lány). Ez a kódolás természetesen nem kötelező, de a későbbi visszakeresésnél segíti a tájékozódást.

A program kérdéseire válaszolva beírtuk egyenként az osztály tanulóinak nevét, életkorát, testmagasságát cm-ben, testtömegét kg-ban. Ezután következtek a teszteredmények felvételei a program kérdéseinek megfelelő sorrendben, először az őszi (kezdeti) adatok.

A program az adatok hihetőségét úgy vizsgálja, hogy az adott határokon belüli értékeket fogadja csak el, hibás adatok esetén módosításra szólít fel. Ha nincs rendelkezésre álló adat, akkor az Enter megnyomásával a következő mezőre ugorhatunk.

<b>1</b>	<b>60 m</b>	<b>időhatár 6,0 mp – 15 mp</b>
<b>2-4.</b>	<b>600 m, 800 m és 1000 m időértékei percben és másodpercben</b>	<b>időhatár 1 perc 30 mp - 6 perc 00 mp</b>
<b>2</b>	<b>Súlypontemelkedés cm-ben</b>	<b>határ : 20- 100 cm</b>
<b>3</b>	<b>Távolugrás / helyből / 10 cm-es pontossággal</b>	<b>határ : 150- 650 cm</b> (a felső határ azért ilyen magas, hogy az esetleges későbbi matematikai statisztikai számításoknál a rohammal történő ugrás eredményei is feldolgozhatók legyenek.)
<b>4</b>	<b>Medicinlabdadobás 10 cm-es pontossággal</b>	<b>határ : 300 - 1500 cm</b>
<b>5</b>	<b>Kislabdahajítás m- ben</b>	<b>határ : 10 - 80 m</b>
<b>6</b>	<b>Négyütemű szabadgyakorlat db</b>	<b>(60 mp alatt)</b>
<b>7</b>	<b>Labdapassz db</b>	<b>(30 mp alatt)</b>

Ha már korábban adtunk be adatokat (ehhez előkészítettük a **94RAD1EF**, a **94TIS1BF**, és a **94IOSZT** file-okat), akkor a program megismeréséhez a már korábban általunk bevitt adatokkal dolgozunk, ezért nyomjuk meg az **"M"** billentyűt! A program kilistázza a megfelelő file neveket, amelyek közül az egyik újragépelése jelenti a választást. Az egyszerűsítés érdekében előbb a **Caps Lock** billentyűvel a gépet folyamatosan "Nagy" betű írására állítsuk át!

Nyomjuk meg az Enter gombot! A gép kiírja ebben a file-ban szereplő neveket, Enter újbóli leütésével a monitoron megjelenik:

**További feldolgozás :**

<b>Adatbővítés / kezdeti /</b>	<b>B</b>
<b>Adatkiegészítés / végső /</b>	<b>V</b>
<b>Adattörlés</b>	<b>T</b>
<b>Adatmódosítás / kezdeti /</b>	<b>M</b>
<b>Adatmódosítás / végső /</b>	<b>G</b>
<b>Adatlistázás</b>	<b>L</b>
<b>Értékelés</b>	<b>É</b>
<b>Kilépés</b>	<b>K menü.</b>

A "B" billentyű megnyomásával újabb személy felmérési adatait lehet a gépbe vinni. A "V" billentyűvel lehet az őszi (végső) eredményeket a gépbe vinni. A "T" billentyűvel egyenként törölhetők egy tanuló adatai a felmérésből, ha pl. már nem az osztály tanulója az illető.

Az "M" billentyűvel a kezdeti, vagyis az őszi felméréskor kapott adatok bevitele esetén elkövetett hibát lehet módosítani. A keresés név szerint soronként történik.

A módosítható adat elütő / sárga/ színnel jelenik meg.

A helyesbített adat bevitelkor a korábbi téves adat törlődik, és az új kerül a helyére. Ugyanígy módosítható a "G" betűvel a végső (tavaszi) eredmények beírásakor elkövetett hiba is.

Az "L" gomb segítségével listázható egyszerre 6 tanuló minden tesztadata. Enterrel lapozható a kép a monitoron.

Az "E" billentyű leütésekor megjelenik a képernyőn egy újabb menü, amely a matematikai statisztikai kiértékelések között ad választási lehetőséget.

A sorok végén megjelenő betűk az egyes leütendő betűket jelentik.

## 7.2. Kezdeti értékek feldolgozása

Üssük le az "E" betűt, majd a képernyőn leolvasható tesztek utáni szám valamelyikének megfelelőjét a billentyűzet számsorán pl: nézzük a *Távolugrás helyből* (6. próbát):

A képernyőn olvasható: a teszt megnevezése, a felmérés ideje - kezdeti (mert korábban már a kezdeti adatok kiértékelését választottuk a "K" leütésével), az érvényes eredmények száma, a legkisebb és legnagyobb érték, a *medián*, a *modus* és az *átlag* (számtani közép), valamint a *szórás* ill. a *variáció százalék* értéke.

Az Enter leütése után a képernyőn megjelenik a *gyakoriság - teljesítmény grafikon*. Az ábra csak a két szélső érték közötti teljesítmények függvényében adja meg az egyes gyakoriságot. A legrövidebb vonal jelent egy egységet, 2x-es gyakoriságnak 2x-es, stb. magas vonal felel meg.

Az ábráról az is leolvasható, hogy pl. melyik teljesítményérték hiányzik az adatok közül.

Az Enter újbóli leütésével visszakapjuk a matematikai statisztikai kiértékelés korábban már megismert menüjét.

Bármely más teszt kezdeti értékeinek feldolgozása - a korábban felsorolt kifejezések konkrét értékei hasonló módon kaphatók, csak más - más számot kell leütni.

A végső értékek feldolgozása is ugyanígy - a "V" billentyű, majd a megfelelő szám leütésével - kapható.

A program tartalmazza az adatok *osztálybasorolását* mind a kezdeti, mind a végső adatok tekintetében. Ennek céljából vagy az "S" vagy az "Y" billentyűt kell leütni. A monitoron ismét a tesztek menüjéből történő választás után megjelenik: a teszt megnevezése, az érvényes eredmények száma, a legkisebb és a legnagyobb értékek, a kiszámolt átlag, ill. egy táblázat, amelyben az első oszlop az egyes osztályokat jelenti a "tól-ig" határok megjelölésével, a második oszlopban az egyes osztályok *közepi* jelennek meg, a harmadikban pedig a *gyakoriság*, azaz hogy hány adat esik az így kijelölt osztályokba. A következő oszlop a *gyakoriságok és az osztályközépek szorzatát* adja.

A "rang"-gal jelölt oszlopban van a "0" érték, ebbe az osztályba tartozik az adatok átlaga. Az utolsó oszlop a gyakoriság rang\*rang-ját adja.

Alul pedig megjelenik az

$$s = \sqrt{\frac{F \cdot D^2}{N}} \cdot \text{oszt.nagyság}$$

képlet alapján kiszámított szórásérték. Enterrel visszajuthatunk az előbbi menühöz.

### 7.3. Hipotézisvizsgálatok számítógéppel

#### 7.3.1. Egymintás „T” próba

Válasszuk most az összefüggés-vizsgálatok közül az **Egymintás T próbát**, azaz az önkontrollos vizsgálatot! Ennek céljából üssük le az "O" billentyűt. A próbatesztek közül válasszuk a 6-t, azaz a korábban már megismert helyből távolugrás értékeinek vizsgálatát!

A képernyőn először egy táblázatrész jelenik meg, a tanulók névsor szerinti adatsorából az első 15. A táblázat első oszlopa a sorszámot, a második a kezdő, majd a következő a végső távolugró értékeket tartalmazza. A " változás " oszlopban előjelesen veszi figyelembe a végső és kezdeti értékek különbségét. Az utolsó oszlop e különbségek négyzetét adja. Enter után folytatódik a táblázat a következő 15 adattal, ill. még egyszer leütve / Enter / a 31. adat utáni vonal alatt kapjuk a változások összegét,

$$\sum x = 390 \text{ ill. } \sum x^2 = 9300$$

Az Enter újbóli leütése után megjelenik a képernyőn a távolugrás adatainak értékelése:

<b>átlag</b>	$\bar{x} = 12,58$
<b>szórás</b>	<b>s = 12,10</b>
<b>" t " érték</b>	<b>t = 5,79</b>

A program megnevezése után másodperceken belül kaptuk ezeket az adatokat, korábban pedig mennyit kellett számolni a cél érdekében. Az Enterrel ismét visszakapjuk a menütáblát.

#### 7.3.2. Kétmintás „T” próba

Ha most arra vagyunk kíváncsiak, hogy két csoport kezdeti és végső adatai közötti eltéréseket összehasonlítva milyen eredményt kapunk, azaz szeretnénk elvégezni a **kontrollcsoportos vizsgálatot** azaz a Kétmintás "t" próbát, akkor üssük le a "C" billentyűt!

Ebben az esetben is a helyből távolugrás eredményeit vizsgáljuk, tehát a " G "-t üssük le! A képernyőn először megjelenik a **94RAD1E**osztály táblázatsora (Enterrel lapozható).

Az Enter billentyű 3x-i leütése után kapjuk az osztály átlagát

$$\bar{x} = 12,58$$

Az Enter újbóli leütése után be kell gépelni a másik osztály kódját, a **94TIS1BF**-et. Az Enter 2x-i leütése után megjelenik ennek az osztálynak a táblázatsora. 3x-i Enter után leolvasható az átlag

$$\bar{y} = 12,26$$

Még egyszeri Enter után megjelenik a képernyőn a

<b>közös szórás</b>	<b>s = 11,64</b>
<b>" t " érték</b>	<b>t = 0,11</b>

#### 7.3.3. „F” próba

Hogy egyáltalán jogosult-e a kétmintás " t " próba elvégzése, azt a képernyőn megjelenő "F" próba adatai adják.

Az első osztály (94RAD1EF), mint kísérleti csoport szórása az

$$s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N-1}} \text{ képlet alapján } s_2 = 12,10, \text{ hasonlóan a második (94TIS1BF) osztály, mint kontrollcsoport szórása } s_1 = 11,17$$

Valamint az  $F = \frac{s_2^2}{s_1^2}$  képlet alapján, ahol a számlálóban a nagyobb szórásérték szerepel.  $(s_2 > s_1)$  **F=1.15**

Az "F" táblázatból 30 szabadságfokú számláló és nevező metszéspontjában **2,07** található. A kiszámított érték  $F = 1,15$  ennél kisebb, tehát nem szignifikáns, jogosult volt a kétmintás T próba alkalmazása.

### 7.4. Korrelációs számítás számítógépes program segítségével

A program a mértékkorrelációt tartalmazza. Vegyük most is pl. a kislabdahajítás és a medicinlabda-dobás közötti korrelációs együttható értékének kiszámítását !

A szokásos értékek menüjéből kell kiindulni. A kezdeti értékek közötti korrelációhoz üssük le az " M " billentyűt, majd medicinlabda / 7 / és kislabda / 8 / számjelöléseinek beütése után a következő táblázat jelenik meg a monitoron :

**Mértékkorreláció vizsgálat  
medicinlabda – kislabda**

Az első oszlop a sorszámot, a második a medicinlabda-dobás teszt adatait (x), a harmadik ezek átlagtól való eltéréseit  $(x - \bar{x})$ , majd ezek soronkénti négyzeteit  $(x - \bar{x})^2$  adja. Hasonlóan az ezt követő oszlopok a kislabdahajítás eredményeit (y), az átlagtól való eltérésüket  $(y - \bar{y})$ , azok négyzeteit  $(y - \bar{y})^2$  t, majd a legutolsó oszlop a két eltérés szorzatát adja soronként  $(x - \bar{x}) \cdot (y - \bar{y})$

A 31. adatsor leolvasása után a vonal alatt megjelennek az egyes oszlopok adatainak összegei:

$$\sum(x - \bar{x})^2 = 577587 \qquad \sum(y - \bar{y})^2 = 3439,94 \qquad \sum(x - \bar{x})(y - \bar{y}) = 38154$$

$$r = \frac{\sum(x - \bar{x})\sum(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2}\sqrt{\sum(y - \bar{y})^2}} = \frac{38154,84}{\sqrt{577587,1}\sqrt{3439,94}} = 0,856$$

Bármely két teszt adatai közötti mértékkorrelációt hasonlóképpen lehet a programból megtudni.

A program mind a kezdeti mind a végső eredmények esetén tartalmazza az interkorrelációs mátrixot. Ennek megjelenítéséhez az "E" értékelés menüjéből az "I" ill. "J" billentyűk leütésével jutunk.

Irodalomjegyzék:

- |                       |  |
|-----------------------|--|
| Hajtman Béla:         | Bevezetés a matematikai statisztikába<br>Akadémia Budapest, 1968.                |
| Nagy-Báthori-Makszin: | Mérési és számítási módszerek a testnevelésben.<br>Tankönyvkiadó Budapest, 1983. |