

Nemparametrikus tesztek

A csonka félévre való tekintettel

A tananyag kényszerű átszerkesztése miatt idén nem lesz szó a legismertebb parametrikus próbáról, a varianciaanalízisről. A mai anyag megértéséhez ennyit fontos tudni róla:

- ▶ Kettőnél több mintát hasonlítunk össze. A teszt végrehajthatóságának feltétele, hogy a minták varianciái homogének legyenek, mert a teszt eredménye ezen a feltételezésen alapszik.
- ▶ Egytényezős varianciaanalízis: mint a kétmintás t -próba, csak kettőnél több független minta van. Ezek egyetlen paraméterben különböznek egymástól, pl. szöveg nyelve, csak nem két nyelv van, hanem több. Ezeknek a teszteknek tehát egyetlen független változója van, kettőnél több faktorial (= szinttel), pl. nyelv: angol, portugál, szuahéli.

Folytatás

- ▶ Többtenyezős varianciaanalízis: két vagy több független változója van, két vagy több szinttel. Például nyelv: angol, portugál, szövegtípus: eredeti, fordítás. Ilyen a `trans.RData`-ban található adatsor is.
- ▶ Ismételt méréses varianciaanalízis: a páros t -próba általánosítása. Egy elemről kettőnél több adatunk van, pl. vérnyomás a gyógyszerzedés 0., 6. és 12. hetében. Lehet egy- vagy többtenyezős.

Nemparametrikus tesztek

Alkalmazásuk feltétele: az eloszlás nem normális, de szimmetrikus és unimodális.

- ▶ ordinális adatok esetén,
- ▶ metrikus adatok esetén (intervallum és arányskála), ha nem normális eloszlásúak, vagy ha többmintás dizájn esetén a varianciahomogenitás feltétele nem teljesül,
- ▶ nominális adatok (gyakoriságok) esetén.

Háttérolvasmány Reiczigel et al., 7.6. és 7.3. rész.

A Friedman-próbáról Dalgaard könyvében lehet olvasni.

Rangpróbák (nemparametrikus próbák)

Alapgondolat: a próbastatisztikát nem a megfigyelt értékekből, hanem azok sorrendbeli számoljuk ki, hasonlóan a Spearman-féle ρ és Kendall-féle τ számításához.

Példa: 5 nő és 5 férfi testtömege. Igaz-e, hogy a két csoport tömegének mediánja különböző?

Nők tömege: $nok = c(50, 55, 60, 70, 80)$

Férfiak tömege: $ferfiak = c(60, 75, 80, 90, 100)$

Eljárás: sorbaállítjuk mind a tíz elemet növekvő sorrendben, nőket és férfiakat vegyesen. A sorszámok összege 55 ($1+2+3 \dots +10$). Ezután összeadjuk külön a nők és külön a férfiak sorszámait.

Számítás az R-ben

A teszt alapja a t -próbához hasonlóan egy előre definiált eloszlás, ami függ a minta elemszámától.

R-ben: `wilcox.test(nok,ferfiak)`

Figyelmeztetés: “cannot compute exact p-value with ties” – azért kapjuk, mert a 60 kg kétszer fordul elő a mintában.

Milyen következtetést vonunk le? Más eredményt kapnánk, ha a hipotézisünk nem kétoldalú lenne (a tömegek különböznek), hanem egyoldalú (a férfiak tömege nagyobb)?

Próbák típusai

Próbák:

- ▶ Mann-Whitney-próba, U-próba: a független mintás t-próba megfelelője: két ordinális vagy nem normális eloszlású független minta.
- ▶ Wilcoxon-próba: a páros t-próba megfelelője: két ordinális vagy nem normális eloszlású páros minta. Alternatív eset: egy minta hasonlítása egy középértékhez, mint az egymintás t-próbánál.
- ▶ Kruskal-Wallis-próba, H-próba: a független mintás egytényezős varianciaanalízis megfelelője: kettőnél több ordinális vagy nem normális eloszlású független minta.
- ▶ Friedman-próba: az ismételt méréses egytényezős varianciaanalízis megfelelője.

R-függvények:

Mann-Whitney és Wilcoxon-próba: `wilcox.test(paired=F vagy paired=T)`,
Kruskal-Wallis-próba: `kruskal.test()`,
Friedman-próba: `friedman.test()`.

Mann-Whitney-próba

Hatékony-e egy tesztelt vaskészítmény a vérszegénység ellen? Az adatok a kezelés (szer és placebo) utáni hemoglobinszintet mutatják.

```
kezelt = c(9.1, 10.3, 11.0, 11.5, 11.9, 9.5, 10.6,  
9.3, 11.0, 9.8)
```

```
kontroll = c(8.1, 8.4, 9.2, 9.4, 8.8, 9.8, 8.2, 10.3,  
9.5)
```

```
wilcox.test(kezelt, kontroll)
```

A nullhipotézist elvetjük. Miért?

Wilcoxon-próba

Mennyire elfogadható a *hotelban*, ill. *hotelben* alak? Egy 1-től 5-ig terjedő skálán kell értékelni, 1: egyáltalán nem elfogadható, 5: teljesen elfogadható. Tíz megkérdezett:

```
hatsom = c(5,5,5,5,4,5,5,5,4,5)
```

```
elsom = c(1,3,5,4,2,3,2,4,5,2).
```

Itt a tíz megkérdezett mindkét alakot értékelte, ezért a páros Wilcoxon-próbát alkalmazzuk:

```
wilcox.test(hatsom,elsom,paired=T)
```

A minták rangsorának azonosságát elvetjük. Miért?

Kruskal-Wallis-próba

A kettőnél több független mintás varianciaanalízis nemparametrikus megfelelője. Feltétele a varianciák azonossága és az eloszlások hasonlósága. Igaz-e, hogy a három magánhangzó átlagos tartama eltérő?

A `longvow.RData` mátrix magyar á, ó és ú magánhangzók tartamát tartalmazza (clara.nytud.hu/~mady/courses/statistics/materials). Alternatíva közvetlen letöltésre:

```
load(url("http://clara.nytud.hu/~mady/courses/statistics/materials/longvow.RData"))
```

Ellenőrizzük, hogy a tartamok a három magánhangzócsoportban normális eloszlást mutatnak-e. Ezt megtehetjük egyetlen függvénnyel is:

```
tapply(longvow$dur, longvow$vowel, shapiro.test)
```

p /u:/-ra és /a:/-ra $p < 0.05$, tehát nem teljesül a normális eloszlás feltétele.

Helyette:

```
kruskal.test(longvow$dur~longvow$vowel)
```

A különbség szignifikáns. Miért?

Friedman-próba

Az egytényezős, ismételt méréses ANOVA nemparametrikus megfelelője.

Az `accdur.RData` hangsúlyos szótagok tartamának megvalósulását tartalmazza. Azt feltételezzük, hogy a hangsúly típusának erősödésével egyre hosszabb lesz a szótag hossza. A leggyengébb a tág vagy mondatfókusz (`broad`), a legerősebb a kontraszthangsúly, `contrastive`.

Friedman-próba:

```
friedman.test(dur ~ focus | subj, accdur.mean)
```

A `|subj` azt jelenti, hogy a fókusztypusokon mért értékeket egyazon beszélőn belüli összehasonlításban vizsgáljuk.

Milyen eredményt kapunk?

Helyes eljárás

A mátrix beszélőnként és fókusz típusonként nem egy, hanem 12 adatpontot tartalmaz. A teszt bemeneteként viszont beszélőnként és cellánként egyetlen érték szerepelhet.

Cellaátlagok számítása:

```
accdur.mean =  
aggregate(accdur$dur, by=list(accdur$subj, accdur$focus),  
mean)
```

Az új mátrixban nem szerepelnek a korábbi oszlopnevek, az R automatikusan számozta be őket.

Oszlopnevek definiálása:

```
names(accdur.mean) = c("subj", "focus", "dur")
```

Így már lefut a Friedman-próba:

```
friedman.test(dur ~ focus | subj, accdur.mean)
```

Eredmény?

χ^2 -próba

Egy vagy két nominális skálájú minta eloszlásának illeszkedését teszteli a χ^2 -eloszlás mentén. Egymintás változatában az egyenletes eloszlás leggyakoribb tesztje.

Egy minta: khi-négyzet-próba eloszlásvizsgálatra. Megfigyelések gyakoriságát összehasonlítjuk a várt gyakorisággal, azaz n/k -val.

Például: ugyanannyi gyerek születik-e minden hónapban? 100 fős minta esetén megfigyelt gyakoriságok és várt gyakoriságok száma:

	jan	feb	már	ápr	máj	jún	júl	aug	szept	okt	nov	dec
megf.	8	9	10	4	14	7	9	10	6	9	8	6
várt	8,3	8,3	8,3	8,3	8,3	8,3	8,3	8,3	8,3	8,3	8,3	8,3

Példa

Illeszkednek-e a megfigyelt gyakoriságok a várt gyakorisághoz?
szulettes = c(8,9,10,4,14,7,9,10,6,9,8,6)

```
chisq.test(szulettes)
```

Eredmény? Az illeszkedés hipotézisét nem vetjük el. Akkor sem, ha májusban legalább négyel több gyerek született, mint bármely más hónapban.

Ha $p < 0.05$, a megfigyelt gyakoriságok nem illeszkednek a várt gyakoriságokhoz, azaz legalább egy érték kilóg (pl. májusban 20 gyerek született). Ellenőrzés: újabb teszt a kilógónak tűnő érték nélkül. Ha az nem szignifikáns, ez az érték okozza az egyenlőtlenséget.

χ^2 -próba két mintára

Khi-négyzet-próba függetlenségvizsgálatra: függetlenek-e a gyakoriságok a nominális skála szintjeitől?

Itt a megfigyelt gyakoriságokat nem a várt gyakorisággal, hanem a másik mintával hasonlítjuk össze.

Két paprikanövény fejlődését vizsgáljuk június elejétől.

Megszámoljuk, hány termést hoznak egy-egy héten. Az egyik növényt az erkélyre tesszük, a másikat a konyhaablakba. Jobban fejlődik-e az egyik vagy a másik növény?

	1. hét	2. hét	3. hét
erkély	11	17	14
konyha	31	25	28

H_0 : az eloszlások függetlenek a nominális változó szintjeitől, azaz a két növény egyformán jól fejlődik.

Példa

```
paprika = cbind(c(11,31),c(17,25),c(14,28))  
chisq.test(paprika)
```

Eredmény? Az eloszlás nominális változótól való függetlenségének hipotézisét nincs okunk elvetni. A gyakoriságok tehát függetlenek a növénytől.

A teszt szerint mindegy, hogy a paprikánkat az erkélyen vagy a konyhaablakban neveljük-e a fejlődésük szempontjából.

Fontos: a chí-négyzet próba csak nagy minták esetén megbízható. Elvárt cellagyakoriság legalább 5.

Fisher-próba

Fisher's exact test: kontingenciatáblázatokra alkalmazható. Mivel a konkrét gyakoriságokkal számol, nem egy eloszlásra illeszt, kisszámú adat esetén is megbízható.

Példa: idén nem paprikapalántát vettünk, hanem magról keltettük a paprikát. Ismét az erkélyen és a konyhaablakban próbálkozunk.

Az eredmény satnya. Kérdés: befolyásolja-e a hely a növény fejlődését?

```
paprika2 = cbind(c(1,8),c(2,4),c(3,7))  
fisher.test(paprika2)
```