

Név:

Hallgatói azonosító:

A 3 feladatsor kitöltésére 20-15-20 perc, összesen 55 perc áll rendelkezésre.  
 Összesen 76 pont érhető el. Az elégségeshez 20 pont (26%) elérése szükséges.  
 A vizsga során segédeszköz nem használható.  
 Jó munkát!

**Figyelem! A tesztfeladatok kitöltése során a rossz válasz -1 pontot ér.**

### 1. feladatsor (30 pont) , rendelkezésre álló idő 20 perc

#### 1. feladat (14 pont)

Eloszlástípusok:

- |                  |             |            |
|------------------|-------------|------------|
| a) chinégyzet    | c) gamma    | e) Poisson |
| b) exponenciális | d) normális | f) Wishart |

Egyértelműen jelölje, hogy a tanultak alapján tipikusan melyik eloszlástípus lép fel az alább adott esetekben?

- |  |          |          |          |          |          |          |
|--|----------|----------|----------|----------|----------|----------|
| • Esemény első bekövetkezésének ideje:                           | a        | <b>b</b> | c        | d        | e        | f        |
| • Esemény negyedik bekövetkezésének ideje:                       | a        | b        | <b>c</b> | d        | e        | f        |
| • Hányszor következett be egy adott esemény:                     | a        | b        | c        | d        | <b>e</b> | f        |
| • Sok, független, elemi véletlen hatás összegződéseként áll elő: | a        | b        | c        | <b>d</b> | e        | f        |
| • Véletlen standard normális vektor hossznégyzete:               | <b>a</b> | b        | c        | d        | e        | f        |
| • Normális vektor kovariancia mátrixa becslésének eloszlása:     | a        | b        | c        | d        | e        | <b>f</b> |
| • Hotelling féle $T^2$ eloszlás határeloszlása:                  | a        | b        | c        | d        | e        | <b>f</b> |

#### 2. feladat (6 pont)

Próbák:

- |                        |                     |                     |
|------------------------|---------------------|---------------------|
| a) Kolmogorov-Szmirnov | b) Cramér-von Mises | c) Anderson-Darling |
|------------------------|---------------------|---------------------|

Egyértelműen jelölje, hogy a **tanultak** alapján melyik próba teszteli az eloszlás

- |  |          |          |          |
|--|----------|----------|----------|
| • extrém értékeinek illeszkedését?           | a        | b        | <b>c</b> |
| • szokásos/gyakori értékeinek illeszkedését? | <b>a</b> | b        | c        |
| • értékeinek tendenciózus eltérését?         | a        | <b>b</b> | c        |

#### 3. feladat (10 pont)

Válassza ki a helyes választ! (helyes válasz: 2 pont, hiányzó válasz: 0 pont, rossz válasz: -1 pont)

Mi a tesztek helyes erősorrendje kollineáris alternatíva mellett a MANOVA esetén?

- |   |
|---|
| a) Wilks-Lambda $\leq$ Pillai $\leq$ Lawley-Hotelling $\leq$ Roy.   |
| <b>b) Pillai <math>\leq</math> Wilks-Lambda <math>\leq</math> Lawley-Hotelling <math>\leq</math> Roy.</b> |
| c) Roy $\leq$ Lawley-Hotelling $\leq$ Wilks-Lambda $\leq$ Pillai.   |
| d) Pillai $\leq$ Lawley-Hotelling $\leq$ Wilks-Lambda $\leq$ Roy.   |

Melyik a likelihood hányados teszt a MANOVA esetén

- a) Roy
- b) Pillai
- c) Wilks-Lambda
- d) Lawley-Hotelling

A MANOVA esetén melyik teszt eloszlása áll elő, mint nem azonos, de független béta eloszlású valváltozók szorzatának eloszlása.

- a) Roy
- b) Pillai
- c) Wilks-Lambda
- d) Lawley-Hotelling

Az egydimenziós szórásanalízis ANOVA esetén milyen próbával döntünk a nullhipotézisről

- a) t-próba
- b) F-próba
- c) chi-négyzet próba
- d) U-próba

Hogyan kapja meg a főkomponenseket?

- a) Az adatmátrix spektrálfelbontásából.
- b) Az adatok variancia-kovariancia mátrixának normált sajátvektoraiként.
- c) Az adatmátrix normált sajátvektoraiként.
- d) Az adatok variancia-kovariancia mátrixának QR felbontásában szereplő Q ortogonális mátrix oszlopvektoraiként (Gram-Schmidt ortogonalizáció).

Név:

Hallgatói azonosító:

**2. feladatsor (24 pont), rendelkezésre álló idő 15 perc**

**4. feladat. (14 pont)**

*Jelölje be, hogy Igaz, vagy Hamis-e az állítás (helyes válasz: 2 pont, hiányzó válasz: 0 pont, rossz válasz, vagy egyszerre 2 válasz: -1 pont)*

a) A valószínűségi változó négyzetének feltételes várható értéke nagyobb, mint a feltételes várható érték négyzete.

**Igaz** Hamis

b) A lineáris regresszió a magyarázó változók négyzete vagy koszinusza és a válasz közti zajjal terhelt lineáris kapcsolatot nem képes feltárni.

Igaz **Hamis**

c) A regresszióban a legkisebb négyzetes becslés a predikciós hibát minimalizálja.

**Igaz** Hamis

d) A regresszióban legkisebb négyzetes becslés a reziduálisok négyzetösszegét minimalizálja.

**Igaz** Hamis

e) A tolerancia azt méri, hogy az egyes magyarázó változóink tartalmaznak-e a többiektől eltérő információt.

**Igaz** Hamis

f) Cross validation-nal felmérhetjük, hogy a magyarázó változók információtartalma mennyire releváns a válaszcra nézve.

Igaz **Hamis**

g) A logisztikus regresszió jóságát a ROC görbe alatti terület nagyságával tudjuk jellemezni; minél nagyobb, annál lényegesebb a feltárt kapcsolat.

**Igaz** Hamis

**5. feladat (10 pont)**

Melyik teszttel vagy eljárással döntene a következő problémákról?

*Adja meg a helyes választ! (Egyetlen teszt vagy eljárás megnevezését fogadom csak el, minden más 0 pont. Helyes válasz: 2 pont)*

a)  
100 tesztalany közül 50 férfi, 50 nő, a férfiak és a nők közül is 25-25 dohányzik. Megadják, közülük kinek van magas vérnyomása. Eltérő módon hat-e a nők, ill. a férfiak magas vérnyomás kockázatára a dohányzás?

Válasz: **Khi négyzet próba**

**b)**

Egy kísérletben 100-100-100-100 növénynek külön-külön bórt, magnéziumot, cinket, vasat adnak (egy növénycsoport csak egyféle kiegészítő tápanyagot kap), illetve egy ugyancsak 100-as kontrol csoport nem kap kiegészítőt, és egyébként azonos körülmények közt nevelik őket. Feljegyzik mennyit nőttek egy hét alatt. Hogyan állapítaná meg, hat-e egyáltalán valamelyik anyag?

Válasz: **T-próba (kétmintás)**

**c)**

Autógumi gyártmányfejlesztése során a gumi fékutjának hosszát mérik adott sebességen, két komponens különböző arányának függvényében. Egy-egy rögzített arány mellett 100 gumi fékhosszát jegyzik fel. Az optimális komponensarányt szeretnék meghatározni, a fékút, mint a komponensarány egy polinomiális függvénye minimumhelyeként.

Hogyan illesszék az adott fokszámú polinomot az adatokra.

Válasz: **F-próba?**

**d)**

Az előző példában milyen eljárással választaná ki a legjobb fokszámot.

???? **Főkomponens**

**e)**

A városban 200 helyen mintát vesznek a csapvízből és elemzik azt nitritre, nitrátra, keménységre, vezetőképességre, vas- és kloridtartalomra. Milyen eljárással próbálná besorolni a mintavételi helyeket a vízjellemzők hasonlósága szerint (hogyan pl. ezzel szennyeződést lokalizáljon)?

Válasz: **Wilks-Lambda**

Név:

Hallgatói azonosító:

**3. feladatsor (22 pont) , rendelkezésre álló idő 20 perc**

**6. feladat (8 pont)**

*Válassza ki a helyes választ! (helyes válasz: 2 pont, hiányzó válasz: 0 pont, rossz válasz: -1 pont)*

Hol lép fel a Wishart eloszlás?

- a) Normális vektor hosszára vonatkozó próbában
- b) Normális vektor várható értékére vonatkozó próbában
- c) Normális vektor likelihood becslésének eloszlásaként
- d) A főkomponensek meghatározásában

Honnan következik, hogy ha normális együttes eloszlású valváltozók korrelálatlanok, akkor függetlenek is?

- a) Abból, hogy normális vektor lineáris leképezése normális vektor marad
- b) A normális vektor sűrűségfüggvényének formulájából
- c) Abból, hogy minden normális vektor független koordinátájába forgatható
- d) A normális vektor feltételes várható értékének linearitásából

Honnan következik, hogy szórásElemzésben (az alapfeltevések igaz volta mellett) jogos az F próba, mert a két szórásbecslés független „mintából” történik?

- a) A Fisher-Cochran tételből
- b) A Wishart eloszlás tulajdonságából
- c) Normális vektor sűrűségfüggvényének formulájából
- d) A variancia kovariancia mátrix spektrálfelbontásából

Honnan következik, hogy együttesen normálisak feltételes várható értéke a feltétel lineáris függvénye

- a) Abból, hogy normális vektor lineáris leképezése normális vektor marad
- b) A normális vektor sűrűségfüggvényének formulájából
- c) Abból, hogy minden normális vektor független koordinátájába forgatható
- d) Abból, hogy ha normális együttes eloszlású valváltozók korrelálatlanok, akkor függetlenek is

**Röviden válaszolja meg az alábbi kérdéseket.**

Csak rövid, néhány mondatos, vagy egyszerű formulás válaszokat kérek/fogadok el.

**7. feladat (4 pont)**

Definiálja a Wishart eloszlást. (A pontos feltételekre figyeljen.)

**8. feladat (4 pont)**

Mi alapján választja meg a faktorok/ megtartott főkomponensek számát?

**9. feladat (6 pont)**

Többváltozós problémákban mi a koordinátaként végzett tesztelés 3 fő problémája