

Név:

Hallgatói azonosító:

A 3 feladatsor kitöltésére háromszor 15, összesen 45 perc áll rendelkezésre.  
 Összesen XX pont érhető el. Az elégségeshez 30% (YY pont) 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)

## 1. feladat (14 pont)

Eloszlástípusok:

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

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

- |  |          |          |          |          |          |
|--|----------|----------|----------|----------|----------|
| • Örökifjú eloszlás:   | a        | <b>b</b> | c        | d        | e        |
| • Független ilyen eloszlásúak összege is ilyen eloszlású:    | a        | b        | <b>c</b> | <b>d</b> | e        |
| • Standardizált összeg határeloszlása:                       | <b>a</b> | b        | c        | d        | e        |
| • Véletlen standard normális vektor hossz négyzete:          | <b>a</b> | b        | c        | d        | e        |
| • Normális vektor kovariancia mátrixa becslésének eloszlása: | a        | b        | c        | d        | <b>e</b> |
| • Hotelling féle $T^2$ eloszlás határeloszlása:              | a        | b        | c        | d        | <b>e</b> |

## 2. feladat (6 pont)

Próbák:

- |                              |              |                |
|------------------------------|--------------|----------------|
| a) Brock-Dechert-Scheinkmann | b) Ljung-Box | c) Fordulópont |
|------------------------------|--------------|----------------|

Egyértelműen jelölje, hogy a **tanultak** alapján melyik próba tesztel

- |                            |          |          |          |
|----------------------------|----------|----------|----------|
| • függetlenséget?          | <b>a</b> | b        | <b>c</b> |
| • csak korrelálatlanságot? | a        | <b>b</b> | c        |

## 3. 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)

Egy bank ismeri ügyfelei életkorát, havi jövedelmét, lakásviszonyát, rezsiköltségeit, a nekik folyósított kölcsön nagyságát, futamidejét, havi törlesztőrészletét. Statisztikát vezet arról, hogy ügyfelei közül kik váltak nemfizetővé a folyósított banki kölcsön futamideje alatt. Új kölcsönigénylőjéről ugyancsak rendelkezésre állnak a fenti adatok. Ha 5%-nál nagyobb eséllyel lenne nemfizető a futamidő alatt, akkor elutasítja a kölcsönigényt. Milyen eljárással döntsön erről?

Válasz: Logisztikus regresszió

**b)**

Egy kísérletben 100-100-100 betegnek külön-külön háromféle gyógyszert adnak (egy betegcsoport csak egyféle gyógyszert kap), illetve egy ugyancsak 100-as kontrol csoport placebo-t kap. Feljegyzik mennyi idő alatt gyógyulnak meg. Hogyan állapítaná meg, hogy hat-e egyáltalán valamelyik gyógyszer? Vaktesztet végeznek, tehát NEM TUDJÁK melyik a kontrol csoport.

Válasz: Kétmintás T-próba

**c)**

Autógumi gyártmányfejlesztése során a gumi fékútjának hosszát mérik adott sebességről fékezve, 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.

Válasz:

**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:

**2. feladatsor (26 pont)**

**4. feladat. (16 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 feltételes várható értékének várható értéke nagyobb, mint a változó közönséges várható értéke.

*Igaz Hamis*

b) A lineáris regresszióban vizsgált függvénykapcsolat a magyarázó változók ismeretlen együttthatóiban lineáris.

*Igaz Hamis*

c) A determinációs együtttható a magyarázó változók által megmagyarázott variancia arányát adja meg. *Igaz Hamis*

d) A regresszióban a t-próbával azt vizsgáljuk, hogy a predikció várható értékben mennyire jól adja meg a választ.

*Igaz Hamis*

e) A Cook távolság a predikció távolságát méri a választól.

*Igaz Hamis*

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

*Igaz Hamis*

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

*Igaz Hamis*

h) A Cramér-von Mises próba teszteli az eloszlás extrém értékeinek illeszkedését.

*Igaz Hamis*

## 5. feladat (10 pont)

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

Az alábbiak közül melyik likelihood hányados teszt?

- a) Anderson-Darling
- b) U-próba
- c) F-próba
- d) Ljung-Box

Mi a tesztek helyes erősorrendje **diffúz** alternatíva mellett a MANOVA esetén?

- a)  $\text{Roy} \leq \text{Lawley-Hotelling} \leq \text{Pillai} \leq \text{Wilks-Lambda}$ .
- b)  $\text{Roy} \leq \text{Wilks-Lambda} \leq \text{Lawley-Hotelling} \leq \text{Pillai}$ .
- c)  $\text{Roy} \leq \text{Lawley-Hotelling} \leq \text{Wilks-Lambda} \leq \text{Pillai}$ .
- d)  $\text{Pillai} \leq \text{Wilks-Lambda} \leq \text{Lawley-Hotelling} \leq \text{Roy}$ .

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).

Melyik próba teszteli az eloszlás szokásos/gyakori értékeinek illeszkedését?

- a) Jarque-Bera
- b) Kolmogorov-Szmirnov
- c) Cramér-von Mises
- d) Anderson-Darling

Név:

Hallgatói azonosító:

**3. feladatsor (22 pont)**

**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 szórásbecslésben (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.

Legyen  $S = ZZ^T$  egy  $p \times p$  dimenziós pozitív definit szimmetrikus véletlen mátrix. Ekkor  $S$  Wishart eloszlású,  $n$  szabadsági fokkal.  
 $S \sim \mathcal{W}_p(\Sigma, n)$ .

**8. feladat (4 pont)**

Mi alapján választja meg a faktorok/ megtartott főkomponensek számát?

**9. feladat (6 pont)**

Többsváltozós problémákban mi a koordinátáinként végzett tesztelés 3 fő problémája