

Biostatistiek

Biostatistiek

Richting Eysica

Jaar MFYS

Bespreking

- Het examen wordt via Blackboard afgenomen . Er zullen rond de 20 vragen worden gesteld, allemaal multiple choice. Hiervan zullen ongeveer een kwart theoretische vragen zijn en driekwart praktisch. Je zult vaak "select cases" moeten gebruiken!
- Er is geen klassieke giscorrectie meer, met puntenaftrek voor een fout antwoord, maar de totaalscore wordt wel gecorrigeerd. Een leeggelaten vraag wordt hetzelfde beoordeeld als een fout antwoord. Deze correctie is: het aantal juiste vragen - het aantal foute vragen gedeeld door 3, waarna dit omgeschaald wordt naar 20. Een excel-file met een voorbeeld van deze correctie is bijgevoegd bij het studiemateriaal ("voorbeeldCesuur"). Merk op dat je hierbij je punt nog kan bijtellen, dat je verdiend hebt met de practicum-testen.
- Het examen is geen open boek, maar je mag de beslissingsboom en de short manual van SPSS (die beide op Blackboard staan) printen, en meenemen naar het examen. Je mag op deze 2 papieren schrijven (aan de voorkant en de achterkant), en we gaan niet controleren wa er op staat. Andere hulpmiddelen (cursus nota's of enig elektronisch item) is niet toegestaan.
- Je zult geen verslag moeten schrijven zoals op het practicum, alleen de multiple choice vragen moeten beantwoord worden.
- Rekenmachines en kladpapier zijn niet toegestaan, maar je mag de ingebouwde calculator van Windows gebruiken (onder 'Accessories'), of MS Excel.

Puntenverdeling

Multiple choice (20 vragen)

- 1/4 theorie
- 3/4 oefeningen
- Cesuur giscorrectie
- (evt.) Bonuspunten van testjes tijdens het academiejaar

Examenvragen

Verander geen data, outliers of dergelijke. Indien je gegevens zou veranderen, dan komen de antwoorden niet meer overeen met de aangeboden keuzemogelijkheden.

Dus geen outliers verwijderen of veranderen!

- Als de SPSS output geeft $p = 0.000$, dubbel klik op de 0.000 en dan kan je het exacte cijfer zien.
- Alle data zijn reële data en er mogen geen gegevens aangepast of verwijderd worden in de database.
- Open de database BSGVD bis.sav
- Dit is een database over stemkarakteristieken zowel bij gezonde vrijwilligers als bij patiënten met stemproblemen. Verschillende centra namen deel aan deze studie, vandaar dat er zoveel gegevens inzitten.
- Van de meeste personen in de database werd opgevraagd of ze roken [smoking] (0 neen, 1 ja). Er werd ook hun leeftijd [age] nagegaan. Tevens werden de personen opgedeeld volgens het criterium of ze geld verdienen met hun stem, tzt in hoeverre ze in hun beroep afhankelijk zijn van hun stem. De variabele [professi] geeft dit weer: Hierin zijn er 4 niveaus:
 - Elite VP = Elite vocal performers zijn operazangers, radio of tv presentators enz.
 - Prof Voice User = Professional Voice user = advocaten, artsen, winkelbedienden, die gedeeltelijk financieel afhankelijk van de stem. Wanneer ze volledig geen stem hebben, kunnen ze nog een gedeelte van hun beroep uitoefenen.
 - Nonvocal Prof = Bedienden, secretaressen enz zijn slechts matig financieel afhankelijk van de stem.
 - Nonvocal Nonprof = mensen die totaal niet financieel afhankelijk van de stem zoals werkmans of werkvrouw, mensen die bandwerk doen, enz...

1. Welk type examen is dit

1. A
2. B
3. C
4. D

2. Selecteer de groep professi=2 (Professional Voice Users). Ga na of de rokers en niet rokers in deze groep gemiddeld even oud zijn. Welke p-waarde komt overeen met de analyse die dit probleem onderzoekt?

1. $p = 0.020$
2. $p = 0.861$
3. $p = 0.882$
4. geen van de vorige

3. Selecteer de Elite VP en ook de nonvocal Prof (select cases -> professi = 1 or professi = 3). Onderzoek het effect van roken en geslacht [sex] op de leeftijd. Het is een uitbreiding van de vorige vraag. Wat is dan juist wanneer men de gemiddelde leeftijd in deze groep onderzoekt?
 1. geslacht heeft een significant effect op de leeftijd, rekening houdende met de factor roken
 2. roken heeft een significant effect op de leeftijd rekening houdende met de factor geslacht
 3. het verschil in leeftijd tussen rokende mannen en niet-rokende mannen is significant verschillend van het verschil in leeftijd tussen rokende vrouwen en niet-rokende vrouwen"
 4. geen van de vorige is juist
4. Selecteer enkel de Elite VP (select cases -> professi = 1). Shimmer en Jitter zijn beide een indicatie voor de zuiverheid van de stem. Hoe groter deze zijn, hoe meer ruis er op de stem zit, en dus hoe ruwer en dus ook meer aangetast. Normale waarden zijn ongeveer rond 1%. Hoe hoger jitter en shimmer, hoe slechter de kwaliteit van de stem, en hoe meer kans op stembandaandoeningen. Het is logisch dat deze twee metingen een zekere gelijkenis vertonen. Welke waarde geeft het best het verband weer tussen [jitter] en [shimmer]?
 1. $p = 0.000002$
 2. $r = 0.647$
 3. $p = 0.001$
 4. $r = 0.497$
5. De variabele [Diagnose] geeft alle mogelijke ziektebeelden die werden gevonden bij de personen in de database. Selecteer alle diagnoses tot en met vocal nodules (stembandknobbels). Dus select cases -> diagnose < 5. Hebben vrouwen en mannen [sex] gemiddeld beschouwd dezelfde ziektebeelden? Ga dit na. Welke p-waarde komt overeen met het juiste antwoord op deze vraagstelling?
 1. $p = 0.025$
 2. $p = 0.063$
 3. $p = 0.022$
 4. geen van vorige
6. Selecteer de vrouwen van UZA (centre = 4). Select cases -> sex = 0 and centre =4. Ga in 1 enkele test na of roken [smoking] en leeftijd [age] een invloed hebben op de maximale fonatietijd [MPT]. De maximale fonatietijd is de maximale tijd (beste van 3 pogingen) tijdens dewelke men /aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa/ kan zeggen. Dit is typisch rond 20 seconden, behalve indien men getraind is. Welke p-waarde komt overeen met de significantietest die nagaat of roken de MPT beïnvloedt, rekening houdend met leeftijd.
 1. $p = 0.027$
 2. $p = 0.143$
 3. $p = 0.126$
 4. geen van vorige

7. Includeer enkel de vrouwen van UZA met diagnose = LaryngChronic (chronische laryngitis) en diagnose = Exces Musc Tens (sex=0 AND (diagnose=7 OR diagnose=1) AND centre=4). De variabele [jitter] is een maat voor ruis, of onregelmatigheid van de stem. Er moet een minimum aan ruis aanwezig zijn en typisch heeft men een jitter van de orde van 1%, maar hogere waarden geven aan dat er afwijkingen van de stembanden zijn. Welke p-waarde bekom je wanneer je test of er een verschil is qua jitter voor de diagnoses Exces Musc Tens versus LaryngChronic?
1. $p = 0.546$
 2. $p = 0.436$
 3. $p = 0.650$
 4. $p = 0.396$
8. Hoe interpreteer je het vorige resultaat?
1. Er zijn geen argumenten om de nulhypothese te verwerpen.
 2. de nulhypothese dient verworpen te worden.
 3. De kans voor het vinden van de huidige resultaten of extremere resultaten, puur op basis van toeval, is zodanig groot dat we de nulhypothese kunnen verwerpen.
 4. geen van de vorige is juist.
9. Includeer alle diagnoses, maar enkel de vrouwen van UZA/(sex=0 AND centre=4). De maximale fonatietijd is de maximale tijd (beste van 3 pogingen) tijdens dewelke men /aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa/ kan zeggen. Dit is typisch rond 20 seconden, behalve indien men getraind is. Vergelijk de maximale fonatietijd [MPT] tussen de verschillende professionele categorieën [professi]. Welke is de juist p-waarde die test of er een verschil in MPT-waarde is al naargelang de professionele categorie?
1. $p=0.000475$
 2. $p=0.834$
 3. $p=0.000067$
 4. geen van vorige
10. Voortgaande op de resultaten van de vorige test, willen we nagaan welke professionele categorieën qua MPT significant verschillend zijn van de EliteVP groep. Hou bij deze test rekening met het aantal uitgevoerde testen. Enkel de vergelijking met EliteVP is van belang. Welke categorieën hebben een MPT die sig. verschilt van de EliteVP groep?
1. Prof Voice User
 2. Nonvocal Prof
 3. Nonvocal Nonprof
 4. ze zijn allemaal verschillend van de Elite VP

11. Includeer alle observaties. Een aantal factoren bepalen mogelijk de maximale fonatietijd [MPT]. Bepaal uit de volgende lijst van variabelen welke een invloed hebben op MPT:
 1. Roken [smoking],
 2. leeftijd [age]
 3. Fundamentele frequentie [Fo]
 4. Geslacht [sex]
 5. stemmisbruik [abusus].
12. Welke van deze variabelen hebben een significante bijdrage tot het voorspellen van de MPT, wanneer ze tegelijk in 1 model zitten?
 1. Roken, leeftijd, fundam.freq, geslacht, stemmisbruik
 2. Roken, leeftijd, fundam.freq
 3. Fundam.freq
 4. Roken, Fundam.freq
13. Includeer enkel de vrouwen van UZA met diagnose = Exces Musc Tens (sex=0 AND diagnose=1 AND centre=4). Bij alle patienten wordt de fundamentele frequentie bij de hoogste [highest] en laagste [lowest] gemeten. Ga na of de fundamentele freq bij de hoogste verschilt van de fundamentele freq bij de laagste. Welke p-waarde vind je?
 1. $P < 0.000$
 2. $P = 0.005034$
 3. $P = 0.588$
 4. Geen van de voorgaanden
14. Includeer enkel de vrouwen van UZA/(sex=0 AND centre=4). Analyseer of de fundamentele frequentie [fo] beïnvloed wordt door de jitter. Welke p-waarde bekom je?
 1. $P = 0.031$
 2. $P = 0.619$
 3. $P = 0.610$
 4. Geen van voorgaande

Academiejaar 2010-2011 1^{ste} zit

- Als de SPSS output geeft $p = 0.000$, dubbel klik op de 0.000 en dan kan je het exacte cijfer zien.
- Open de database fokkerij.sav
- Dit is een database over anatomische gegevens van renpaarden, gefokt in verschillende stallen/fokkerijen in Frankrijk. De prestaties van dieren zijn opgemeten gedurende 3 opeenvolgende jaren evenals een aantal anatomische kenmerken (lengte van bepaalde segmenten van de benen etc).
- Alle data zijn reeel en er mogen geen gegevens aangepast of verwijderd worden in de database.
- Er zijn 4 jarige paarden en 5 jarige paarden [age=4 ofwel age=5].

- De paarden behoren toe aan verschillende stallen, verspreid over Frankrijk, aangegeven door de variabele [Site].
- Sommige paarden presteren middelmatig (medium) tijdens wedstrijden en sommige presteren goed (high). [Performance = 1 (medium) of = 4 (high)).

1. Duidt aan tot welke examentype uw opgave behoort. Dit staat bovenaan in het midden op deze bladzijde

1. Type A
2. Type B
3. Type C
4. Type D

2. Selecteer alle stallen behalve site =2, (dus select cases Site~2). De variabele ISOCAT geeft de prestaties van de paarden weer in 3 groepen: medium, good en excellent. Onderzoek de relatie tussen de prestaties en de verschillende stallen. Dan kan men de volgende uitspraak maken:

1. Men heeft geen argumenten om aan te tonen dat de prestaties van de paarden anders zijn tussen de verschillende stallen
2. Er is een significant verschil tussen de stallen wat betreft de prestaties van de paarden
3. Er is een associatie tussen stallen en prestaties.
4. geen van de vorige antwoorden is correct.

3. De variabele [Performance] geeft weer of de paarden middelmatig (medium) of zeer goed presteren (high). De variabele [age] geeft weer of het 4 of 5 jarige paarden zijn. Zijn er qua prestaties verschillen tussen de paarden van 4 of 5 jaar? [variabelen Performance en age].

Voer de analyse uit die bij deze vraagstelling past en geef aan welke p-waarde overeenkomt met het antwoord op de juiste vraagstelling.

1. $p = 0.203$
2. $p = 0.204$
3. $p = 0.201$
4. geen van de overige antwoorden is correct

4. De variabele [ISO2004] is een continue maat voor de prestaties van de paarden in het jaar 2004. Hoe hoger, hoe beter. De variabele [site] geeft de stallen of fokkerijen weer, verspreid over geheel Frankrijk waar de paarden toe behoren. Presteren de paarden van de verschillende stallen even goed? (variabelen ISO2004 en site) Voer de analyse uit die bij deze vraagstelling past en geef aan welke p-waarde overeenkomt met het antwoord op de juiste vraagstelling.

1. $p = 0.721$
2. $p = 0.561$
3. $p = 0.790$
4. geen van de overige antwoorden is correct

5. De variabele [ISO2004] is een continue maat voor de prestaties van de paarden in het jaar 2004. Hoe hoger, hoe beter. Is er een verschil tussen de prestaties van jonge (4 jaar) ten opzichte van oudere paarden (5 jaar), aangegeven met de variabele [age]?

Voer de analyse uit die bij deze vraagstelling past en geef aan welke p-waarde overeenkomt met het antwoord op de juiste vraagstelling.

1. $p = 0.0088$
2. $p = 0.0086$
3. $p = 0.702$
4. geen van de overige antwoorden is correct.

6. Aanvullend bij de vorige vraag kan men

1. concluderen dat de 4 jarige paarden significant beter presteren dan de 5 jarige paarden.
2. concluderen dat de 5 jarige paarden significant beter presteren dan de 4 jarige paarden.
3. concluderen dat er geen significant verschil is tussen de prestaties van de 4 of 5 jarige paarden
4. geen van de overige antwoorden is correct.

7. In de verschillende stallen [variabele site] worden er paarden gefokt. Selecteer de stallen 3,4,6,7 en 8 (select cases: site > 2 and site \sim 5). Een van de anatomische kenmerken van de paarden is de hoogte in het frontaal vlak [frontaal_H = frontale hoogte]. Is er een verschil qua hoogte van de paarden tussen de verschillende stallen? Voer de analyse uit die bij deze vraagstelling past en geef aan welke p-waarde overeenkomt met het antwoord op de juiste vraagstelling.

1. $p = 2.04 \cdot 10^{-31}$
2. $p = 1.95 \cdot 10^{-8}$
3. $p = 0.069$
4. geen van de overige antwoorden is correct.

8. Selecteer nog steeds de stallen 3,4,6,7 en 8 (select cases: site > 2 and site \sim 5).

Aansluitend bij vorige vraag wil men weten waar de verschillen qua frontale hoogte [frontaal_H] van de paarden uit de verschillende stallen zich situeren [site]. Indien men een analyse hiervoor uitvoert, dan blijkt o.a.

1. dat de frontale hoogte van de paarden van de stallen uit Compiègne en St Lo 2 niet significant verschillend is
2. dat de frontale hoogte van de paarden van de stallen uit Compiègne en St Lo 2 significant verschillend is
3. dat de voorwaarden om deze analyse uit te voeren niet vervuld zijn
4. geen van de vorige antwoorden is correct.

9. Selecteer alle paarden. Elk paard werd gedurende opeenvolgende jaren opgevolgd qua prestaties. Bekijk de prestaties van 2002 en 2003 [ISO2002 en ISO2003]. Is er een significante evolutie van de prestaties tussen beide momenten? Voer de analyse uit die bij deze vraagstelling past en geef aan welke p-waarde overeenkomt met het antwoord op de juiste vraagstelling.
 1. $p = 6.23 \cdot 10^{-13}$
 2. $p = 5.0 \cdot 10^{-49}$
 3. $p = 0.053$
 4. geen van de overige antwoorden is correct.
10. Bestudeer de evolutie van de prestaties tussen 2002 en 2003 [ISO2002 en ISO2003] rekening houdend met de leeftijd [age] van de paarden. Selecteer alle paarden. Welke conclusie is de juiste?
 1. Zowel jonge als oude paarden presteren beter in 2003 tov.2002
 2. Er is geen verschil in prestatie tussen 2002 en 2003, ongeacht de leeftijd
 3. De evolutie van de prestaties tussen 2002 en 2003 hangt af van de leeftijd van de paarden
 4. Geen van de vorige conclusies is correct.
11. De anatomische kenmerken van het paard werden o.a. opgemeten in de "been lengte segment" [beenlengte] en in frontale hoogte [Frontaal_h]. Zijn deze anatomische kenmerken aan elkaar gekoppeld of zijn het onafhankelijke observaties? Selecteer alle paarden. Voer de analyse uit die bij deze vraagstelling past en geef aan welke p-waarde overeenkomt met het antwoord op de juiste vraagstelling.
 1. $p = 0.164$
 2. $p = -0.051$
 3. $p = 3.50 \cdot 10^{-18}$
 4. geen van de overige antwoorden is correct.
12. Selecteer paarden van stal 1 (Select cases: site =1) Beenlengte segment en frontale hoogte zijn twee variabelen op basis van de welke we hun prestatie in 2003 [ISO2003] willen voorspellen. Welke van de onderstaande uitdrukkingen voorspelt optimaal hoe je ISO2003 kan voorspellen op basis van één of meer van deze variabelen?.
 1. $ISO2003 = 12.119 + 0.721 \cdot \text{frontale hoogte} - 1.146 \cdot \text{beenlengte}$
 2. $ISO2003 = 153,72 - 1.159 \cdot \text{beenlengte}$
 3. $ISO2003 = 23.759 + 0.578 \cdot \text{beenlengte}$
 4. Geen van de overige antwoorden is correct.
13. Welke van de volgende uitspraken is correct?
 1. De p-waarde geeft de kans weer op het vinden van de resultaten uit het experiment puur op basis van toeval
 2. De p-waarde geeft de kans weer op het vinden van de resultaten uit het experiment, of minder extreme resultaten, puur op basis van toeval
 3. De p-waarde geeft de kans weer dat de nulhypothese waar is.
 4. Geen van de overige uitspraken is juist.

Categorieën:

- Eysica
- MFYS