

VILNIAUS UNIVERSITETAS

MATEMATIKOS IR INFORMATIKOS FAKULTETAS

**Tiesiniai modeliai**

Laboratorinis darbas

Atliko: 3 kurso 2 grupės studentai:

Matas Amšiejus

Salvija Račkauskaitė

Sandra Macijauskaitė

Darbo vadovė: doc. dr. Rūta Levulienė

Vilnius, 2021

**TURINYS**

ĮVADAS 4

1. TIRIAMASIS STRAIPSNIS 5

1.1.Tema 5

1.2.Duomenys 5

1.3.Atlikta analizė 6

2. ATLIKTA ANALIZĖ REMIANTIS STRAIPSNIU 6

2.1.Bendra analizės eiga 6

2.2Ūgiai 6

2.3.Masės 10

2.4.Sprintas 12

2.5.Vikrumas 13

2.6.Driblingas 16

2.7.Kamuolio kontrolė 17

2.8.Kamuolio smūgiavimas 19

2.9.Python programavimo kalba 20

IŠVADOS 21

ŠALTINIAI 22

ĮVADAS

Šiame laboratoriniame darbe siekiame ištirti jaunų futbolininkų sugebėjimų įtaką jų vėlesniai profesionaliai karjerai. Sieksime pritaikyti vienfaktorę dispersinę analizę. Laboratorinio darbo uždavinį įgyvendinti pasitelksime R, SAS ir Python programavimo kalbas. Tyrimo tikslas – išanalizuoti ir pateikti pasirinkto straipsnio spredimus.

1. TIRIAMASIS STRAIPSNIS

Straipsnis „The influence of speed abilities and technical skills in early adolescence on adult success in soccer: A long-term prospective analysis using ANOVA and SEM approaches“ pasirinktas iš „Official Journal of the American College of Sports Medicine“ žurnalo, 5 tomo. Straipsnio autoriai: Oliver Höner, Daniel Leyhr, Augustin Kelava. Išleistas 2017 metais rugpjūčio 14 dieną.

1.1.Tema

Straipsnyje analizuojama U12 futbolininkų sugebėjimai ir jų įtaka tolimesnei karjerai. Šiame tyrime buvo įvertinta ilgalaikė prognozuojama motorinių testų, žaidėjų greičio ir techninių įgūdžių reikšmė ankstyvoje paauglystėje.

1.2.Duomenys

Tyrimo imtį sudarė 14178 U12 žaidėjai iš Vokietijos talentų ugdymo programos. Buvo atlikti penki testai (sprinto, vikrumo, driblingo, kamuolio kontrolė, kamuolio smūgiavimas), o žaidėjų ūgis, masė ir santykinis amžius buvo įvertinti visoje šalyje atliekant diagnostiką 2004–2006 m. 2014–15 m. sezono metu žaidėjai buvo suskirstyti į profesionalus (n = 89), pusiau profesionalus (n = 913) arba neprofesionalus (n = 13176). Į pirmas tris geriausias lygas patekę futbolininkai buvo priskirti į APL 1 (suaugusiųjų futbolo lygis) grupę, į ketvirtą – penktą į APL 2, o visi kiti į APL 3. Motorinių testų bei ūgio ir svorio prognostinė svarba buvo nustatyta naudojant dispersinę analizę. Duomenys pateikti pagal 9 kriterijus:

1. Age – amžius, metai;
2. Spr – sprintas, sekundės;
3. Agi – vikrumas, sekundės;
4. Dri – driblingas, sekundės;
5. Bc – kamuolio valdymas, sekundės;
6. Sho – kamuolio smūgiavimas, taškai;
7. Hei – ūgis, centimetrai;
8. Wei – masė, kilogramai;
9. Apl – suaugusiojo veiklos lygis, 1-3 lygiai.

Prieš atliekant motorinius testus, buvo užregistruotas kiekvieno žaidėjo ūgis, svoris ir santykinis amžius (matuojamas pagal gimimo dieną per kalendorinius metus). Motorinių gebėjimų įvertinimus sudarė 5 testai, mažesnės vertės rezultatai rodė geresnius pasiekimus. Žaidėjai buvo išbandyti sprinto (20 m bėgimas), vikrumo ir driblingo (laikas slalomo trasoje atitinkamai be ir su kamuoliu), kamuolio valdymo (šešių perdavimų laikas prieš dvi priešingas sienas) ir kamuolio smūgiavimas ( 8 spyriai į tris skirtingus taikinius, kuriuos treneris įvertino pagal tikslumą ir greitį).

1.3.Atlikta analizė

Tyrime buvo atlikta vienfaktorė dispersinė analizė, kurią pakartojome ir gavome tuos pačius rezultatus, kuriuos aptarsime 2 skyriuje.

1. ATLIKTA ANALIZĖ REMIANTIS STRAIPSNIU

Atlikome vienfaktorę dispersinę analizę pagal straipsnį su SAS, R ir Python programavimo kalbomis.

2.1.Bendra analizės eiga

Pirmiausia nuskaitome duomenis iš S1File.txt. Tam, kad mūsų atliekamo tyrimo rezultatai sutaptų su straipsnio, mes perrenkame duomenis kaskart prieš tiriant naują kintamąjį. Atmetę nekorektiškas reikšmes, brėžiame vidurkių ir standartinių paklaidų grafikus (SAS) bei vidurkių ir standartinių nuokrypių grafikus (R). Tada atliekame dispersinę analizę (tolimesniame tyrime visur bus naudojamas reikšmingumo lygmuo ). Pirma tikriname nulinę hipotezę, kad dispersijos visose grupėse yra lygios, naudodami Levene testą. Jei nulinės hipotezės neatmetame (dispersijos lygios), tęsiame tyrimą. Jei ne, darbą tesiame naudodami Welch ANOVA. Tikriname nulinę hipotezę, kad vidurkiai visose grupėse yra lygūs. Jei taip, analizę baigiame. Jei ne, tikriname, kurie vidurkiai statistiškai reikšmingai skiriasi naudodami Tjukio testą (SAS) bei Bonferoni (R). Patikriname, kokią dalį skirtumų lėmė lyga, o kokią – individualūs žaidėjų skirtumai (R-square).

2.2Ūgiai

Pirma nusibraižome vidurkių ir standartinių paklaidų grafikus (SAS) bei vidurkių ir standartinių nuokrypių grafikus (R).

Chart, line chart

Description automatically generated

1 pav. Ūgio vidurkių ir standartinių paklaidų grafikas (SAS)

Chart, box and whisker chart

Description automatically generated

2 pav. Ūgio vidurkių ir standartinių nuokrypių grafikas (R)

Matome, kad ūgių vidurkiai tarp grupių mažėja. Taip pat mažėja ir standartinė paklaida, nes jos dydis smarkiai priklauso nuo imties dydžio (st. paklaida = st. nuokrypis / sqrt(imties dydis)). Tikriname nulinę hipotezę, kad dispersijos yra lygios.

Table

Description automatically generated

3 pav. Dispersijų lyginimo lentelė (SAS)

A picture containing text

Description automatically generated

4 pav. Dispersijų lyginimo lentelė (R)

Table

Description automatically generatedMatome, kad p reikšmė = 0,6778 didesnė už reikšmingumo lygmenį, todėl nulinės hipotezės atmesti negalime (SAS). Su R gauname kitokią p reikšmę, tačiau išvada padarome tą pačią. Toliau tikriname hipotezę apie vidurkių lygybę.

5 pav. Vidurkių lyginimo lentelė (SAS)



6 pav. Vidurkių lyginimo lentelė (R)

Gauname, kad p reikšmė yra mažiau už reikšmingumo lygmenį, todėl nulinę hipotezę atmetame, vidurkiai statistiškai reikšmingai skiriasi. Vadinasi ūgiai tarp grupių nėra vienodi. Tikriname, kurios grupės skiriasi.

Table

Description automatically generated

7 pav. Grupių tikrinimo lentelė (SAS)

A picture containing text

Description automatically generated

8 pav. Grupių tikrinimo lentelė (R)

Pagal SAS lentelę (Tjukio testą) matome, kad ūgiai statistiškai reikšmingai skiriasi tarp visų grupių (visur yra pažymėta \*\*\*), o pagal R (Bonferoni testą) pirmos ir antros grupės vidurkiai statistiškai reikšmingai nesiskiria su reikš mingumo lygmeniu 0,05. Patikrinkime R kvadrato reikšmes.

Calendar

Description automatically generated with medium confidence

9 pav. R kvadrato lentelė

Gauname, kad tik < 0,01 % ūgių lėmė skirtingi faktoriai (futbolo žaidimo lygiai), o > 99 % – individualūs žaidėjų skirtumai.

2.3.Masės

Chart, line chart

Description automatically generated

10 pav. Masės vidurkių ir standartinių paklaidų grafikas (SAS)

Chart, box and whisker chart

Description automatically generated

11 pav. Masės vidurkių ir standartinių nuokrypių grafikas (R)

Masių vidurkiai vėl mažėja. Tikriname nulinę hipotezę, kad dispersijos yra lygios.

Table

Description automatically generated

12 pav. Dispersijų lyginimo lentelė (SAS)

Text, letter

Description automatically generated

13 pav. Dispersijų lyginimo lentelė (R)

P reikšmė didesnė už reikšmingumo lygmenį, todėl nulinės hipotezės atmesti negalime. Tikriname hipotezę apie vidurkių lygybę.

Table

Description automatically generated

14 pav. Vidurkių lyginimo lentelė (SAS)

Text, letter

Description automatically generated

15 pav. Vidurkių lyginimo lentelė (R)

P reikšmė yra mažiau už reikšmingumo lygmenį, todėl nulinę hipotezę atmetame, vidurkiai statistiškai reikšmingai skiriasi. Vadinasi masės tarp grupių nėra vienodos. Tikriname, kurios grupės skiriasi.

Table

Description automatically generated

16 pav. Grupių lyginimo lentelė (SAS)

A picture containing text

Description automatically generated

17 pav. Grupių lyginimo lentelė (R)

Pagal lentelę matome, kad ūgiai statistiškai reikšmingai skiriasi tarp visų grupių. Patikrinę R kvadrato reikšmę gauname, kad tik < 0,01 % masių lėmė skirtingi faktoriai (futbolo žaidimo lygiai), o > 99 % – individualūs žaidėjų skirtumai.

2.4.Sprintas

Chart, line chart

Description automatically generated

18 pav. Sprinto vidurkių ir standartinių paklaidų grafikas (SAS)

Chart, box and whisker chart

Description automatically generated

19 pav. Sprinto vidurkių ir standartinių nuokrypių grafikas (R)

20 metrų sprinto laikų vidurkiai didėja. Tikriname nulinę hipotezę, kad dispersijos yra lygios. Gauname, kad p reikšmė = 0,28 yra daugiau nei reikšmingumo lygmuo, todėl nulinės hipotezės neatmetame. Tikriname hipotezę apie vidurkių lygybę. Gauname, kad p reikšmė < 0,001, tai ji mažesnė už reikšmingumo lygmenį, todėl nulinę hipotezę atmetame. Sprinto laikai nėra vienodi. Patikrinus, kurios grupės skiriasi, gauname, kad visų grupių laikai yra tarpusavyje skirtingi. Patikrinę R kvadrato reikšmę gauname, kad tik < 0,01 % sprinto greičių lėmė skirtingi faktoriai (futbolo žaidimo lygiai), o > 99 % – individualūs žaidėjų skirtumai.

2.5.Vikrumas

Chart, line chart

Description automatically generated

20 pav. Vikrumo vidurkių ir standartinių paklaidų grafikas (SAS)

Chart, box and whisker chart

Description automatically generated

21 pav. Vikrumo vidurkių ir standartinių nuokrypių grafikas (R)

Tikriname nulinę hipotezę, kad dispersijos yra lygios.

Table

Description automatically generated

22 pav. Dispersijų lyginimo lentelė (SAS)

Text, letter

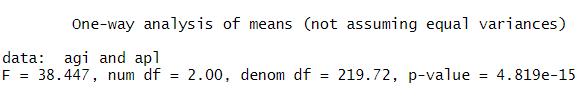
Description automatically generated

23 pav. Dispersijų lyginimo lentelė (R)

Table

Description automatically generatedGauname, kad p reikšmė yra mažiau nei reikšmingumo lygmuo, todėl nulinę hipotezę atmetame, dispersijos tarp grupių nėra lygios. Todėl vidurkių palyginimui naudosime Welch ANOVA. Tikriname nulinę hipotezę, kad vidurkiai tarp grupių yra vienodi.

24 pav. Vidurkių lyginimo lentelė (SAS)



25 pav. Vidurkių lyginimo lentelė (R)

Gauname, kad p reikšmė mažesnė už reikšmingumo lygmenį, todėl nulinę hipotezę atmetame. Vikrumo testo laikai nėra vienodi. Tikriname vidurkių skirtumus tarp kiekvieno faktoriaus.

Table

Description automatically generated

26 pav. Vidurkio skirtumo tarp faktorių lyginimo lentelė (SAS)

Text, letter

Description automatically generated

27 pav. Vidurkio skirtumo tarp faktorių lyginimo lentelė (R)

Gauname, kad pirmos (APL1) ir antros (APL2) grupės laikai statistiškai reikšmingai nesiskiria (nėra \*\*\*). O pirmos ir trečios (APL3) bei antros ir trečios laikai statistiškai reikšmingai skiriasi. Patikrinę R kvadrato reikšmę gauname, kad tik < 0,01 % vikrumo greičių lėmė skirtingi faktoriai (futbolo žaidimo lygiai), o > 99 % – individualūs žaidėjų skirtumai.

2.6.Driblingas

Chart, line chart

Description automatically generated

28 pav. Driblingo vidurkių ir standartinių paklaidų grafikas (SAS)

Chart

Description automatically generated

29 pav. Driblingo vidurkių ir standartinių nuokrypių grafikas (R)

Tikriname nulinę hipotezę, kad dispersijos yra lygios. Gauname, kad p reikšmė < 0,001, o tai yra mažiau nei reikšmingumo lygmuo, todėl nulinę hipotezę atmetame, dispersijos tarp grupių nėra lygios. Todėl vidurkių palyginimui naudosime Welch ANOVA. Tikriname nulinę hipotezę, kad vidurkiai tarp grupių yra vienodi. Gauname, kad p reikšmė < 0,001, o tai yra mažiau už reikšmingumo lygmenį, todėl nulinę hipotezę atmetame. Driblingo testo laikai nėra vienodi. Tikriname vidurkių skirtumus tarp kiekvieno faktoriaus.

Table

Description automatically generated

30 pav. Vidurkio skirtumo tarp faktorių lyginimo lentelė (SAS)

A picture containing text

Description automatically generated

31 pav. Vidurkio skirtumo tarp faktorių lyginimo lentelė (R)

Gauname, kad APL1 ir APL2 grupių laikai statistiškai reikšmingai nesiskiria. O APL1 ir APL3 bei APL2 ir APL3 laikai statistiškai reikšmingai skiriasi. Patikrinę R kvadrato reikšmę gauname, kad tik 2 % driblingo greičių lėmė skirtingi faktoriai (futbolo žaidimo lygiai), o 98 % – individualūs žaidėjų skirtumai.

2.7.Kamuolio kontrolė

Chart, line chart

Description automatically generated

32 pav. Kamuolio kontrolės vidurkių ir standartinių paklaidų grafikas (SAS)

Chart, box and whisker chart

Description automatically generated

33 pav. Kamuolio kontrolės vidurkių ir standartinių nuokrypių grafikas (R)

Tikriname nulinę hipotezę, kad komuolio kontrolės laiko dispersijos yra lygios. Gauname, kad p reikšmė = 0,16 yra daugiau nei reikšmingumo lygmuo, todėl nulinės hipotezės neatmetame. Tikriname hipotezę apie vidurkių lygybę. Gauname, kad p reikšmė < 0,001, tai ji mažesnė už reikšmingumo lygmenį, todėl nulinę hipotezę atmetame. Kamuolio kontrolės laikai nėra vienodi. Tikriname, kurių grupių laikai skiriasi tarpusavyje.

Table

Description automatically generated

34 pav. Vidurkio skirtumo tarp faktorių lyginimo lentelė (SAS)

Text

Description automatically generated with low confidence

35 pav. Vidurkio skirtumo tarp faktorių lyginimo lentelė (R)

Gauname, kad APL1 ir APL2 grupių laikai statistiškai reikšmingai nesiskiria. O APL1 ir APL3 bei APL2 ir APL3 laikai statistiškai reikšmingai skiriasi. Patikrinę R kvadrato reikšmę gauname, kad tik < 0,01 % kamuolio kontrolės greičio lėmė skirtingi faktoriai (futbolo žaidimo lygiai), o > 99 % – individualūs žaidėjų skirtumai.

2.8.Kamuolio smūgiavimas

Chart, line chart

Description automatically generated

36 pav. Kamuolio smūgiavimo vidurkių ir standartinių paklaidų grafikas (SAS)

Chart, box and whisker chart

Description automatically generated

37 pav. Kamuolio smūgiavimo vidurkių ir standartinių nuokrypių grafikas (R)

Tikriname nulinę hipotezę, kad kamuolio smūgiavimo taškų dispersijos yra lygios. Gauname, kad p reikšmė < 0,001, o tai yra mažiau nei reikšmingumo lygmuo, todėl nulinę hipotezę atmetame, dispersijos tarp grupių nėra lygios. Todėl vidurkių palyginimui naudosime Welch ANOVA. Tikriname nulinę hipotezę, kad vidurkiai tarp grupių yra vienodi. Gauname, kad p reikšmė < 0,001, o tai yra mažiau už reikšmingumo lygmenį, todėl nulinę hipotezę atmetame. Kamuolio smūgiavimo surinkti taškai nėra vienodi. Patikrinus, kurios grupės skiriasi tarpusavyje gauname, kad visų grupių laikai yra skirtingi. Patikrinę R kvadrato reikšmę gauname, kad tik < 0,01 % kamuolio smūgiavimo taškus lėmė skirtingi faktoriai (futbolo žaidimo lygiai), o > 99 % – individualūs žaidėjų skirtumai

2.9.Python programavimo kalba

Taip pat atlikome papildomą užduotį su Python kalba, tačiau išvados identiškos, dėl to apraše rezultatai nepateikiami.

IŠVADOS

Mūsų atlikto tyrimo ir tyrinėto straipsnio rezultatai sutapo identiškai. Gavome, kad tiek kūno sudėjimo, tiek žaidėjų įgūdžių parametrai statistiškai reikšmingai skyrėsi tarp žaidėjų profesionalumo lygio. Mažiausiai panašumų buvo tarp pirmosios ir trečiosios bei antrosios ir trečiosios grupės futbolininkų, kas leidžia teigti, kad futbolininkai, žaidžiantys top 5 Vokietijos lygose turi įgūdžių pranašumų prieš likusius neprofesionalius žaidėjus. Tokie požymiai kaip vikrumas, driblingas, kamuolio kontrolė statistiškai reikšmingai nesiskyrė tarp APL1 ir APL2 grupių, tad galime daryti išvadą, kad nebūtinai visi futbolininkai, žaidžiantys top 3 aukščiausiose lygose, pralenkia 4 ir 5 lygos žaidėjus.

ŠALTINIAI

1. „PLOS ONE“ tinklapis. Prieiga per internetą: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0182211>