

# Razpon rok in učinkovitost igralcev lige NBA

Aljaž Švigelj

Mentor: prof. dr. Aleksandar Jurišić

18. marec 2018

## 1. Uvod

Liga NBA je najelitnejša košarkarska liga na svetu in ena izmed najprestižnejših svetovnih športnih lig na sploh. Večini profesionalnih igralcev so največje športne sanje uvrstiti se v NBA in igrati z najboljšimi. Ekipe v ligi porabijo ogromno denarja za iskanje mladih talentov. V ta namen imajo zaposlene ljudi, ki so jim pripravljeni plačati letno tudi več kot 500,000 ameriških dolarjev v upanju, da jim bodo poiskali najbolj nadarjene igralce s celega sveta in jim s tem omogočili prednost pred konkurenco. Igralec se lahko uvrsti v NBA preko izbora, ki poteka enkrat letno. Ekipe izbirajo igralce glede na njihovo dosedanje uspešnost, delovne navade, ocenjen potencial za nadaljnji razvoj, košarkarski IQ itd.. To so vse lastnosti, ki jih lahko košarkaši izboljšujejo tekom kariere, vendar pa niso edine, ki štejejo. Iskalci igralcev so pozorni tudi na vnaprej določene fizične značilnosti. Vsem je znano, da je telesna višina zelo pomemben faktor pri ocenjevanju kvalitete in lastnosti igralca, v zadnjem času pa se je veliko pozornosti preusmerilo tudi k malo manj znani telesni lastnosti, razponu rok.

Tako NBA strokovnjaki kot tudi zagrizeni navijači menijo, da je nadpovprečen razpon rok v primerjavi z igralčevo višino velika prednost. Igralec, ki je enako velik kot nek drugi, ima zaradi daljših rok večjo možnost ujeti žogo, pri metu iz skoka izpustiti žogo na višji točki, imeti daljši doseg za prestrežanje žog in lažje blokirati mete na koš. Tudi marsikateri uspešni igralci, kot so na primer Kevin Durant, Anthony Davis in Kawhi Leonard, imajo zelo velik razpon rok v primerjavi z njihovo višino, kar debato le še pogloblja.

Raziskovalec Matt Ignal je v svojem članku "Making Sense of NBA Wingspan Data" [3, rpubs.com] že raziskoval povezanost med razponom rok in kvalitetno obrambo. Ugotovil je, da igra razpon rok dokaj pomembno vendar ne odločljivo vlogo pri ocenjevanju igralčeve sposobnosti igranja obrambe. Igralec z zelo kratkim razponom rok v primerjavi z njegovo višino je še vedno lahko izjemen v obrambi (kot primer navede Joakim Noah-a in Andrew Bogut-a) vendar pa to ni pogost pojav. Ignal je za izračune uporabil Defensive Box Plus-Minus ali DBPM [8, basketball-reference.com], ki pa je prepoznan kot ne nujno natančen način za vrednotenje igralčevih obrambnih sposobnosti, kar nekoliko zmanjša relevantnost njegovih ugotovitev.

Med tem, ko je Ignal v svojem članku opazoval povezanost med razponom rok in obrambo, je v tej raziskavi v središče postavljena povezava med razponom rok in igralčevo učinkovitostjo, definirano po novem standardu lige NBA [6, stats.nba.com], ki je natančnejša kot tista, katere del je uporabil Ignal. Cilj naloge je ugotoviti, ali sta razpon rok in učinkovitost v resnici povezana in če sta, na kakšen način.

Naloga je sestavljena iz štirih delov. V prvem oddelku je opisan postopek pridobivanja podatkov, za tem pa sta v razdelku 3. opisani analizi odstotka razpona rok v primerjavi s telesno višino ter učinkovitosti. V 4. delu je izračunana linearna korelacija med slednjima spremenljivkama ne glede na pozicijo igralcev, v 5. razdelku pa se enake izračune naredi še na podskupinah igralcev razdeljenih na organizatorje igre, krila in centre. V zaključku so povzeta opažanja in možni razlogi, zaradi katerih smo prišli do navedenih rezultatov.

## 2. Pridobivanje podatkov

Vsi podatki o telesnih značilnostih igralcev so pridobljeni z uradne spletne strani lige NBA, in sicer za vse igralce, ki so prvič igrali v ligi med letoma 2000 in 2014.

Najpomembnejši podatki za podan raziskovalni problem so razpon rok, telesna višina [4, stats.nba.com] in igralčeva učinkovitost v vsaki sezoni, ki jo je odigral v ligi [5, stats.nba.com].

**Razpon rok** je obravnavan v povezavi z igralčevo višino po spodnji enačbi in nam pove, za koliko odstotkov je razpon rok večji oziroma manjši kot igralčeva višina.

$$\text{razmerje razpona rok in višine} = \left[ \frac{\text{razponRok}}{\text{telesnaVišina}} - 1 \right] * 100$$

**Učinkovitost igralca** je uradno definira s strani lige NBA po spodnji enačbi PIE (Player Impact Estimate) [6, stats.nba.com]:

$$PIE = \frac{PTS + FGM + FTM - FGA - FTA + DREB + (.5 * OREB) + AST + STL + (.5 * BLK) - PF - TO}{GmPTS + GmFGM + GmFTM - GmFGA - GmFTA + GmDREB + (.5 * GmOREB) + GmAST + GmSTL + (.5 * GmBLK) - GmPF - GmTO}$$

|       |                                                     |
|-------|-----------------------------------------------------|
| PTS   | Število točk                                        |
| FGM   | Število zadetih metov iz igre                       |
| FTM   | Število zadetih prostih metov                       |
| FGA   | Število metov iz igre                               |
| FTA   | Število prostih metov                               |
| aDREB | Število skokov v obrambi                            |
| OREB  | Število skokov v napadu                             |
| AST   | Število podaj                                       |
| STL   | Število ukradenih žog                               |
| BLK   | Število blokad                                      |
| PF    | Število prekrškov                                   |
| TO    | Število izgubljenih žog                             |
| Gm*   | Vsota lastnosti označene z * vseh igralcev na tekmi |

Tabela 1: Spremenljivke, vključene v enačbo PIE in njihovi pomeni za posameznega igralca

Standard PIE definira, kolikšen vpliv je imel igralec na tekmi, v sezoni itd. Če obravnavamo eno samo tekmo, je vsota PIE vseh igralcev, ki so stopili na igrišče, enaka 100. Če so igralci ene ekipe skupaj dosegli več kot 50 PIE točk, NBA definira, da je ta ekipa tekmo najverjetneje dobila, igralec, ki je na tekmi dosegel več kot 10 PIE točk, pa je bil najverjetneje boljši kot povprečen igralec na tekmi. Več točk kot je dobil, boljši je bil. Visok PIE ekipe močno korelira z zmago (NBA navaja korelacijski koeficient R-kvadrat 0.908) .

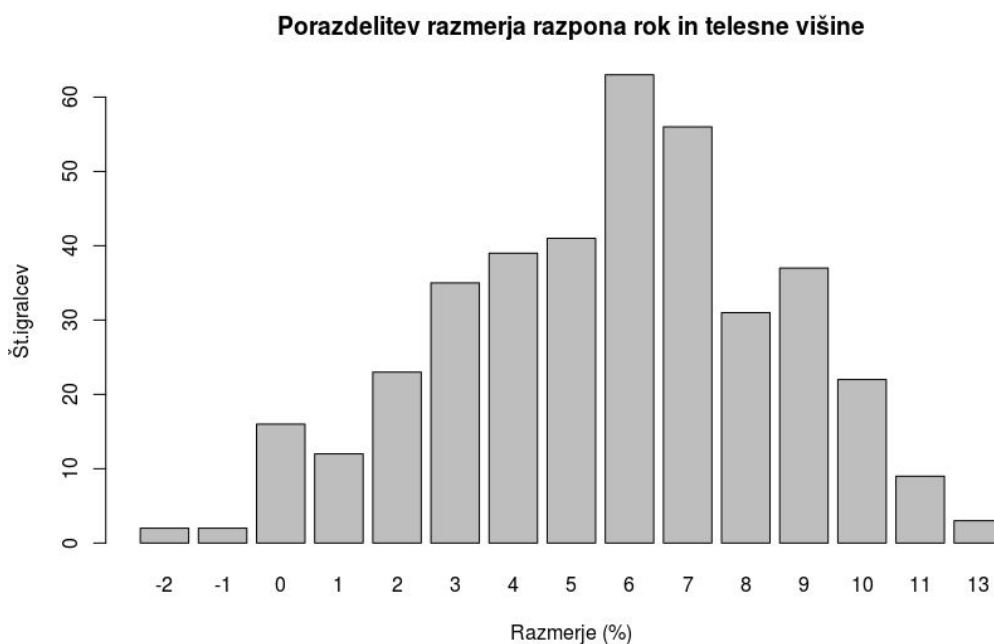
Za opazovanje povezave med razponom rok in učinkovitostjo potrebujemo za vsakega igralca v vzorcu reprezentativno PIE vrednost. Smiselno je, da se kot vrednost vzame povprečen PIE vseh igralčevih sezon, v katerih je odigral več kot 20 tekem (s tem izničimo možnost, da je igralec v neki sezoni odigral le nekaj tekem, ki so bile zelo dobre ali zelo slabe in bi imele prevelik vpliv na povprečen PIE) in upoštevamo le igralce, ki so v ligi odigrali vsaj 3 sezone, se s tem privadili tempu lige in tako bolje izkoristili svoj potencial in s tem tudi (nad/pod)povprečen razpon rok.

### 3. Osnovna analiza podatkov

Osnovne izračune in grafe sem izdelal na podlagi vzorca v velikosti 391 profesionalnih igralcev lige NBA, ki so v vsaj treh sezonah igrali 20 ali več tekem. Pri vseh izračunih je bil uporabljen programski jezik R.

#### 3.1. Odstotek razpona rok v primerjavi z višino

V spodnjem grafu si oglejmo, kako je porazdeljeno razmerje razpona rok in telesne višine med igralci.



Graf 1: Porazdelitev igralcev gledano za kolikšen odstotek je razpon rok večji ali manjši od višine. Najpogostejše pojavljajoča vrednost za razmerje razpona rok in višine je 6. Minimalna vrednost je -2, maksimalna pa 13.

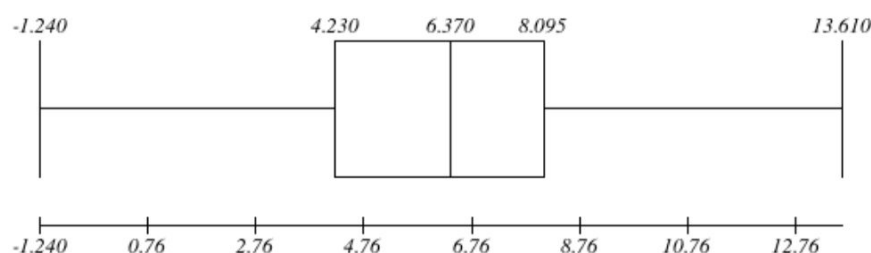
Izračunajmo povprečje razmerja razpona rok s telesno višino (1), standardni odklon (2) in varianco, ki je definirana kot kvadrat standardnega odklona.

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (1) \quad \sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}} \quad (2)$$

Dobimo sledeče vrednosti ( $\bar{x}$  - povprečna vrednost,  $\sigma$  - standardni odklon,  $\sigma^2$  - varianca):

$$\bar{x} = 6.178\%, \quad \sigma = 2.829\%, \quad \sigma^2 = 8.000\%.$$

Oglejmo si, kako so podatki predstavljeni s kvartilnim diagramom (graf 2). Razberemo, da je minimalna vrednost -1.24%, prvi kvartil se nahaja na 4.23%, mediana na 6.37%, tretji kvartil na 8.10% in maksimalna vrednost na 13.61%.



Graf 2: Škatla z brki za odstotek razpona rok

Nadalje izračunajmo interval zaupanja za povprečno vrednost, ki bo s 95% verjetnostjo določil območje, v katerem se nahaja povprečen razpon rok celotne populacije igralcev NBA. Če za stopnjo zaupanja  $\beta$  izberemo 95%, je interval zaupanja sledeč:

$$P(5.897\% \leq \bar{x} \leq 6.459\%) \geq 0.95$$

Za nekaj izmed trenutnih najbolj znanih in uspešnih košarkašev si za primer oglejmo njihove odstotke razmerja razpona rok in telesne višine v primerjavi s povprečno vrednostjo, ki je 6.178%.

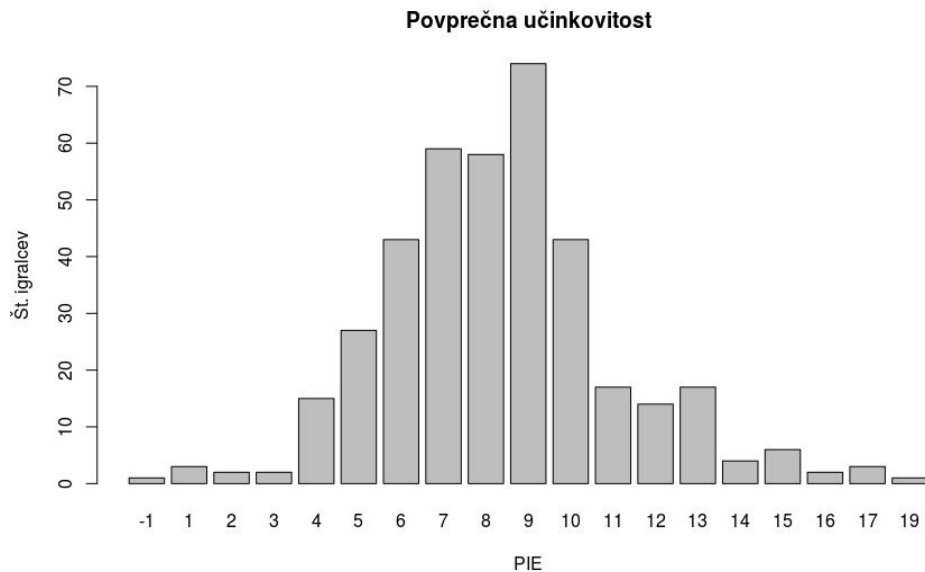
| Št. | Ime               | odstotek razpona rok (%) |
|-----|-------------------|--------------------------|
| 1.  | Kevin Durant      | 9.58                     |
| 2.  | Lebron James      | 6.31                     |
| 3.  | Russell Westbrook | 7.42                     |
| 4.  | Anthony Davis     | 10.13                    |
| 5.  | Kawhi Leonard     | 11.56                    |
| 6.  | Stephen Curry     | 2.02                     |

Tabela 3: Odstotek razpona rok pri nekaterih znanih igralcih (vrstni red je naključen)

Kot vidimo, imajo nadpovprečen odstotek razpona rok Kevin Durant, Russel Westbrook, Anthony Davis in Kawhil Leonard, med tem ko imata LeBron James povprečen odstotek, Stephen Curry pa izrazito podpovprečen. Povprečje razpona rok teh šestih igralcev je 7.88%, kar je za 27% več kot povprečje lige.

### 3.2. Učinkovitost igralcev

Spodaj je naveden graf, ki prikazuje porazdeljenost povprečne učinkovitosti igralcev (standard PIE) vseh odigranih sezon.

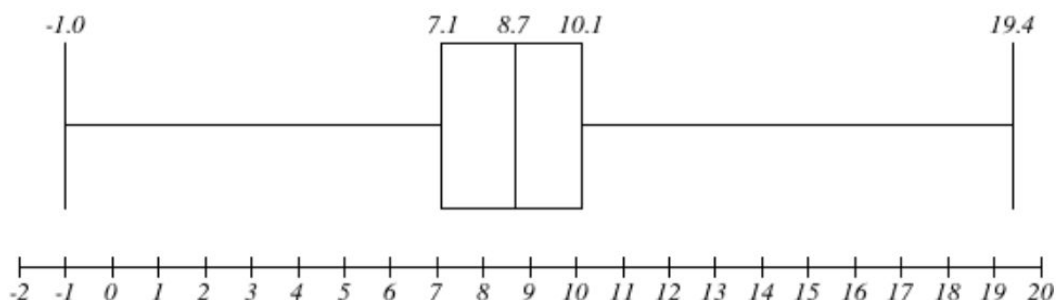


Graf 3: Porazdelitev učinkovitosti igralcev glede na oceno PIE. Največkrat zastopana vrednost PIE je 9, med tem ko levo in desno od te vrednosti št. zastopanj pada. Porazdelitev Je nekoliko desno asimetrična.

Tudi tokrat izračunajmo povprečno učinkovitost po Formuli 1, standardni odklon po Formuli 2 in varianco definirano kot kvadrat standardnega odklona. Dobimo sledeče vrednosti ( $\bar{x}$  - povprečna učinkovitost,  $\sigma$  - standardni odklon,  $\sigma^2$  - varianca):

$$\bar{x} = 8.785 \text{ PIE}, \quad \bar{\sigma} = 2.759 > \text{PIE}, \quad \overline{\sigma^2} = 7.579 \text{ PIE}.$$

Poglejmo si še podatke predstavljene s kvartilnim diagramom (graf 4). Razberemo, da je minimalna povprečna učinkovitost igralca v ligi -1.0 PIE, prvi kvartil se nahaja na 7.1 PIE, mediana na 8.7 PIE, tretji kvartil na 10.1 PIE in maksimalna vrednost na 19.4 PIE.



Graf 4: Škatla z brki učinkovitosti igralcev po standardu PIE

Če za stopnjo zaupanja  $1-\alpha$  izberemo 95%, je interval zaupanja za povprečno vrednost učinkovitosti enak:

$$P(8.5 \text{ PIE} \leq \bar{x} \leq 9.1 \text{ PIE}) \geq 0.95$$

Z zbranimi podatki si lahko ogledamo, kolikšen odstotek razpona rok ima nekaj igralcev z najvišjimi PIE učinkovitostmi v ligi.

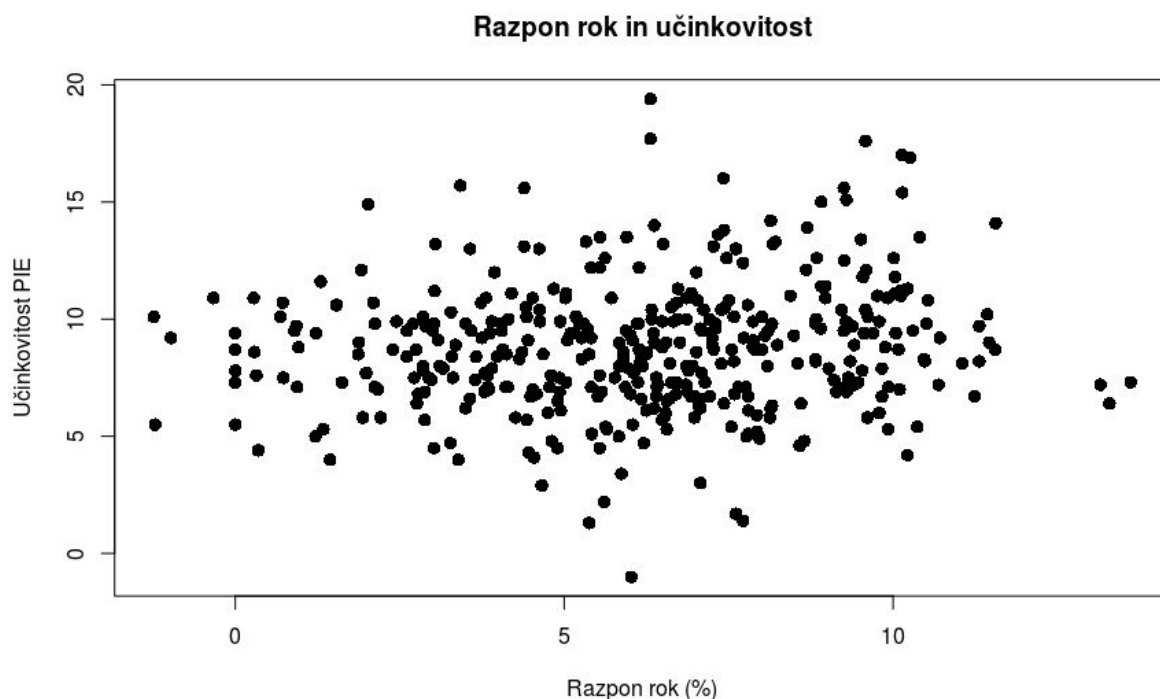
| Št. | Ime               | Učinkovitost (PIE) | odstotek razpona rok (%) |
|-----|-------------------|--------------------|--------------------------|
| 1.  | Lebron James      | 19.4               | 6.31                     |
| 2.  | Chris Paul        | 17.7               | 6.31                     |
| 3.  | Kevin Durant      | 17.6               | 9.58                     |
| 4.  | Anthony Davis     | 17.0               | 10.13                    |
| 5.  | Hassan Whiteside  | 16.9               | 10.26                    |
| 6.  | Russell Westbrook | 16.0               | 7.42                     |

Tabela 4: Najbolj učinkoviti igralci in njihov odstotek razpona rok

Opazimo, da imata dva najbolj učinkovita igralca lige popolnoma povprečen odstotek razpona rok, med tem ko imajo igralci med tretjim in šestim mestom nadpovprečen odstotek razpona. Kot lahko vidimo, med najboljšimi šestimi igralci v ligi ni nikogar s podpovprečnim odstotkom. Povprečen odstotek razpona rok navedenih igralcev je enak 8.34%, kar je za 35% več kot povprečje lige.

## 4. Korelacija med razponom rok in učinkovitostjo

V prejšnjem razdelku (tabela 6) smo izbrali 6 igralcev z največjo učinkovitostjo PIE v ligi. Ugotovili smo, da je njihov PIE višji od povprečja lige, kar bi lahko kazalo na to, da sta učinkovitost in razmerje razpona rok s telesno višino res v povezavi. V tem razdelku bom to poizkusil še statistično potrditi.



Graf 5: Graf igralcev, razvrščenih glede na učinkovitost in razmerje razpona rok in telesne višine.

V Grafu 5 vidimo razporeditev igralcev glede na njihovo učinkovitost ter na odstotek, za katerega je njihov razpon rok večji kot njihova višina. Kot vidimo, je graf precej razpršen, kljub temu pa opazimo, da ima zelo malo igralcev s podpovprečnim razponom učinkovitost več kot 10 PIE, med tem ko ima kar nekaj igralcev z nadpovprečnim razponom PIE vrednost več kot 10.

Poglejmo si raje, kakšna je izračunana korelacija med učinkovitostjo in razmerjem razpona rok in telesne višine. Uporabili smo Pearsonov koeficient korelacije, ki lahko zavzema vrednosti na intervalu  $[-1, 1]$  (formula 3) [2, 194]. Bolj kot je vrednost koeficienta blizu roba intervala, bolj je linearna odvisnost med spremenljivkama močna. Bliže kot je koeficient vrednosti 0, manjša je linearna povezanost med spremenljivkama. Formula za izračun Pearsonovega koeficienta korelacije je navedena kot

$$r_{XY} = \frac{k(X, Y)}{s_X * s_Y} = \frac{\sum_{i=1}^n (x_i - \bar{x}) * (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 * (y_i - \bar{y})^2}} \quad (3).$$

Če v enačbo vstavimo podatke, dobimo sledeč korelacijski koeficient:

$$r_{xy} = 0.128.$$

V našem primeru se izkaže, da je koeficient korelacije zgolj 0.128, kar nakazuje na izredno nizko, praktično zanemarljivo linearno povezavo med razmerjem razpona rok s telesno višino in učinkovitostjo.

Naredimo še test linearne odvisnosti med tema dvema spremenljivkama. Pri domnevi  $H_0$  (spremenljivki nista linearno povezani) velja sledeča enačba testne statistike, ki se porazdeljuje po Student( $n - 2$ ) in ima obliko

$$TS = \frac{r * \sqrt{n - 1}}{\sqrt{1 - r^2}} \quad (4).$$

Določimo stopnjo značilnosti  $\alpha = 0.05$ . Če v enačbo za testno statistiko vstavimo  $r = 0.128$  in  $n = 391$ , dobimo ( $TS$  - testna statistika,  $t$  - kritično območje testa,  $p$  -  $p$ -vrednost):

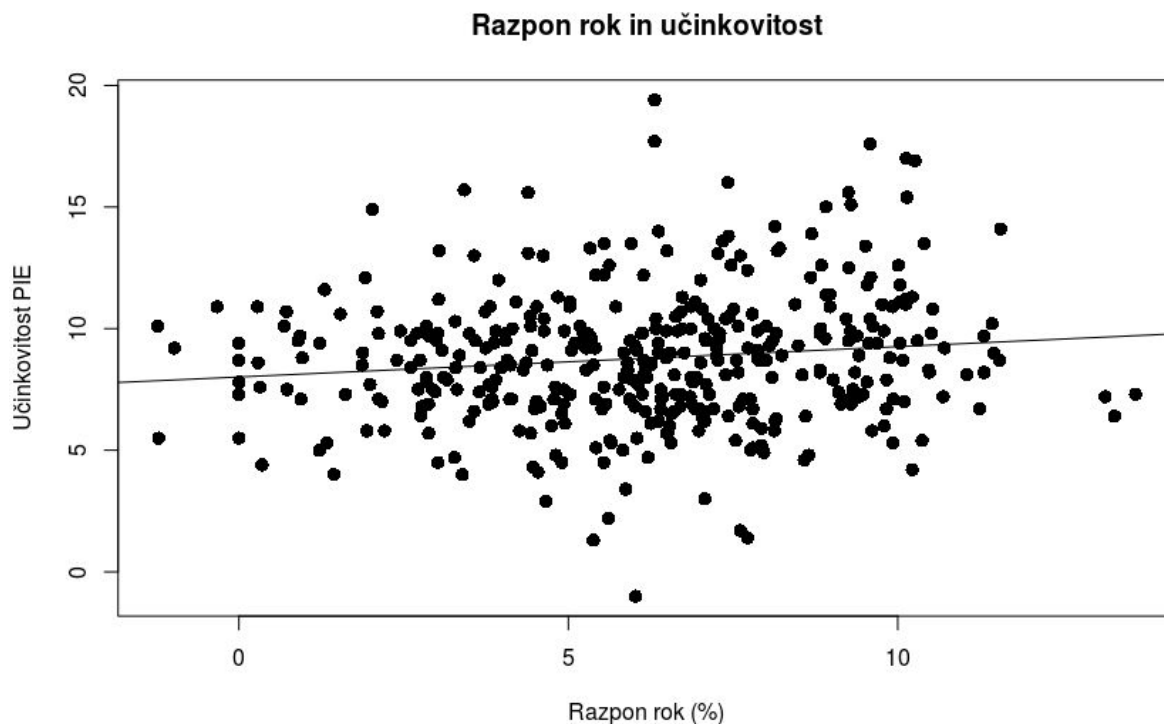
$$TS = 2.54, \quad t = \pm 1.66, \quad p = 0.01137.$$

Ker je  $p$ -vrednost (0.01137) za majhno stopnjo manjša kot stopnja značilnosti (0.05), lahko sprejmemo neničelno hipotezo, torej da je korelacija statistično pomembna, četudi je linearna povezanost med spremenljivkama nizka.

Kljub zelo majhni linearni korelaciji med spremenljivkama, pa vseeno izračunajmo še regresijsko premico. S pomočjo jezika R in funkcije `abline(A, B)`, kjer A predstavlja podatke o učinkovitostih in B podatke o odstotku razpona rok, dobimo spodaj navedeno enačbo. Spremenljivka  $y$  predstavlja učinkovitost PIE,  $x$  pa razpon rok.

$$y = 8.0143 + 0.1247x.$$

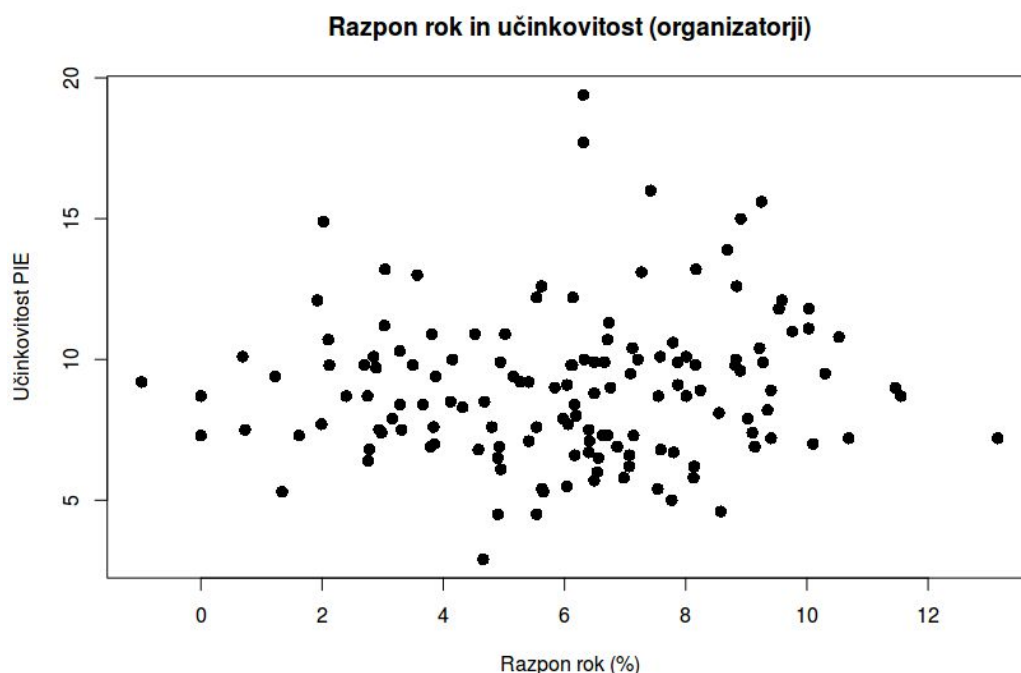
Regresijsko premico še izrišimo na grafu razpršenosti igralcev v odvisnosti odstotka razpona rok od učinkovitosti (Graf 6).



Graf 6: Graf igralcev, razvrščenih glede na učinkovitost in razpon rok z regresijsko premico

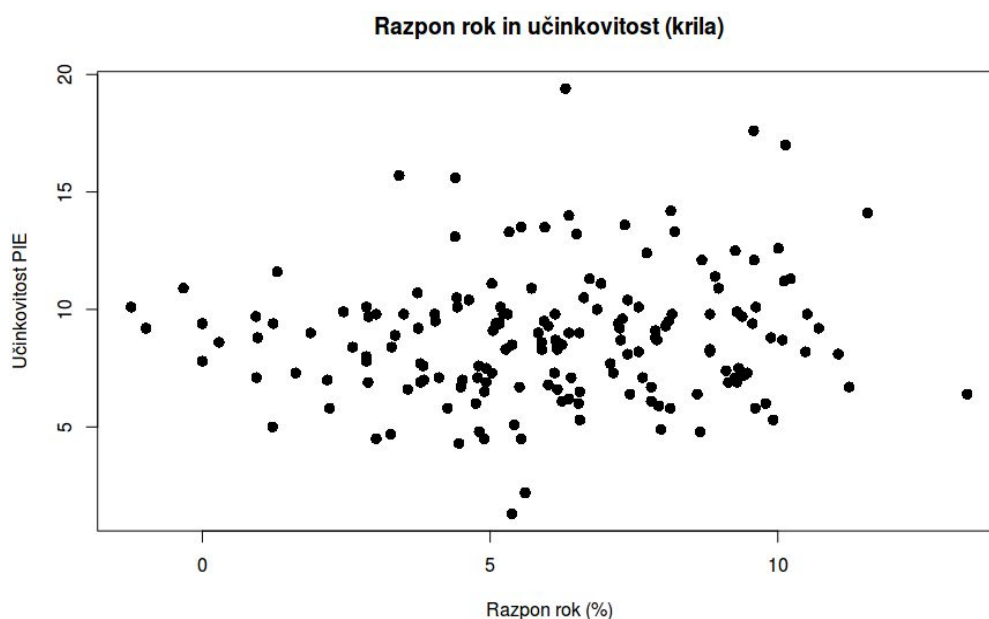
## 5. Obravnava posameznih igralških pozicij

Košarkaši na različnih pozicijah imajo na igrišču različne vloge. Organizatorji igre morajo poiskati najbolj prostega igralca in je za to dobro, da imajo odlično prostorsko predstavo. Naloga kril je prodirati v notranjost rakete in jim zato koristi atletičnost. Centri se borijo za prostor pod košem in so pogosto visoki in močni. Gledano iz tega vidika, je tudi lastnost kot je razpon rok različno koristna pri igralcih na različnih pozicijah. Predvidevamo lahko, da je najbolj koristna pri centrih, saj jim omogoči na primer ujeti žogo na višji točki, kot bi jo lahko, če bi imeli krajše roke, malo manj pri krilih, najmanj pa pri organizatorjih igre, ki se redkeje udeležujejo borbam za žogo. Spodaj so navedeni trije grafi, ki prikazujejo odvisnost razpona rok in učinkovitosti za posamezne pozicije.



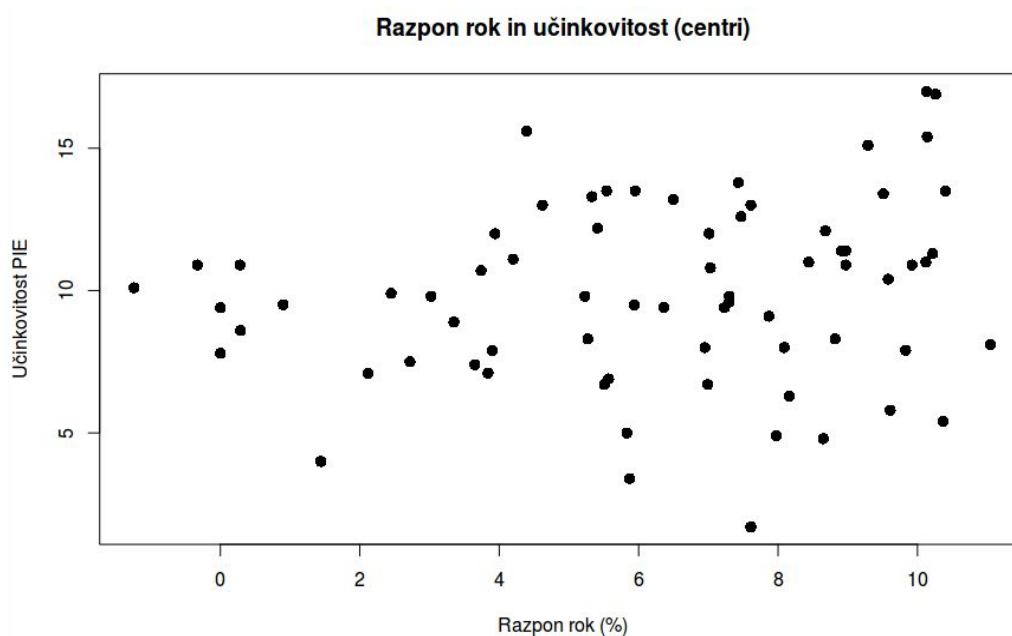
Graf 7: Graf organizatorjev igre, razvrščenih glede na učinkovitost in razpon rok

Kot vidimo, pri organizatorjih igre ni opazne velike povezanosti med obema spremenljivkama, kar odraža tudi izračunan linearni koeficient korelacije, ki znaša 0.079. Vrednost je občutno manjša kot vrednost linearnega koeficienta korelacije pri vseh igralcih, izračunane v prejšnjem razdelku (0.128).



Graf 8: Graf krilnih igralcev, razvrščenih glede na učinkovitost in razpon rok

Tudi pri krilnih igralcih se je izkazalo, da je linearni koeficient korelacije med razponom rok in učinkovitostjo le 0.099, kar je še vedno manj kot koeficient izračunan ne glede na pozicijo. Nizka linearna korelacija pri organizatorjih igre in krilnih igralcih v primerjavi s celoto pomeni, da bo pri centrih le ta nekoliko večja.



Graf 7: Graf centrov, razvrščenih glede na učinkovitost in razpon rok

Pri igralcih, ki imajo pozicijo centra, je izračunana linearna korelacija 0.183, kar je največ od vseh do sedaj opazovanih skupin. Iz tega lahko sklepamo, da je nadpovprečen razpon rok precej bolj pomemben pri visokih igralcih, katerih glavna naloga je braniti koš in dobivati borbe za žogo pod obročem. Precej manj na učinkovitost vpliva pri organizatorjih igre, kjer je linearna povezanost najnižja.

## 6. Zaključek

V seminarski nalogi sem pokazali, da sta razmerje razpona rok s telesno višino in učinkovitost igralcev lige NBA po standardu PIE povezana, četudi je povezava precej majhna. Povzava je najbolj očitna pri centrih, manj pri krilnih igralcih in najmanj pri organizatorjih igre. Ugotovljeno odstotnost povezanosti lahko pojasnimo na različne načine.

Učinkovitost igralcev sem meril po standardu PIE, ki je zasnovan tako, da daje precej večjo težo učinkovitosti v napadu kot v obrambi. Na veliko obrambnih elementov (npr. skoki, blokade in kvalitetna obramba) močno vplivata igralčeva višina in razpon rok, medtem ko je pri napadu takih elementov občutno manj. Standard PIE je tudi zgolj trenutni najboljši približek za merjenje učinkovitosti igralcev in številni strokovnjaki menijo, da ni najbolj optimalna mera uspešnosti. Čeprav naj bi bil boljši kot starejši standardi (PER, PM, RPM, ...), je še vedno razmeroma netočen.

V sezoni 2013-2014 so v NBA na vseh tekmah uvedli naprednejši način merjenja igralčevega vpliva na igrišču, imenovan Player tracking ("sledenje igralcem") [7, espn.com]. Sistem na vsaki tekmi uporablja 6 kamer, ki so zmožne slediti vsakemu igralcu na igrišču. Tako se med drugim lahko meri, za vsakega igralca posebej, njegova hitrost, pretečena razdalja, dotiki žoge na tekmi, podaje, oddaljenost od napadalca, ki naj bi ga igralec pokrival, kvaliteta pozicije pri skoku za žogo in drugo. S temi podatki, ter drugimi, ki jih bodo lahko strokovnjaki pridobili iz posnetkov kamer, bo veliko lažje določiti še dodatne parametre pri računanju učinkovitosti igralca ne le v celoti, temveč tudi posamično v napadu in v obrambi. Z izboljšanim sistemom bi prav gotovo dobili več kvalitetnih podskupin učinkovitosti, na katerih bi lahko analizirali vpliv razpona rok.

Verjetno je tudi, da je opazovanje povezave med razponom rok in učinkovitostjo preveč splošno in bi bilo veliko bolje opazovati povezavo med razponom rok in katero izmed bolj izoliranih skupin podatkov, kot so na primer skoki, blokade in ukradene žoge. Tako bi bili rezultati manj splošni in bolj relevantni.

Smiselno bi bilo, da nadpovprečen odstotek razpona rok vpliva na igralčevo učinkovitost, vendar pa te trditve žal nisem uspel popolnoma podpreti, opazna je bila zgolj majhna povezanost med spremenljivkama. Kljub temu, da ima veliko zvezdnikov visok odstotek razpona rok, pa igralci kot so Stephen Curry (razpon rok: 2%), Blake Griffin (3.4%), Kevin Love (4.3%) in Kyrie Irving (3%) dokazujejo, da si lahko med najboljšimi tudi brez te lastnosti.

## Literatura

- [1] W. Mendenhall, T. Sincich, Statistics for Engineering and the Sciences, 5. izdaja, Pearson, 2006
- [2] A. Jurišić. Verjetnostni račun in statistika, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, 2015.
- [3] rpubs.com (29.8.2016) (<https://rpubs.com/matthewignal/205006>) (18.3.2018)
- [4] stats.nba.com (8.1.2018) (<https://stats.nba.com/draft/combine-anthro/>) (8.1.2018)
- [5] stats.nba.com (8.1.2018) (<http://stats.nba.com/players/advanced>) (8.1.2018)
- [6] stats.nba.com (8.1.2018) (<https://stats.nba.com/help/faq/>) (8.1.2018)
- [7] espn.com (19.5.2012) ([http://www.espn.com/blog/playbook/tech/post/\\_/id/492/492](http://www.espn.com/blog/playbook/tech/post/_/id/492/492)) (10.2.2018)
- [8] basketball-reference.com (18.3.2018) (<https://www.basketball-reference.com/about/bpm.html>) (18.3.2018)