

**Akademia Górniczo-Hutnicza im. Stanisława
Staszica w Krakowie**



Paulina Korzeniowska

**Analiza czynników wpływających na przebieg
profesjonalnej rozgrywki w grze League of
Legends.**

**Wydział Zarządzania
Informatyka i Ekonometria
rok II semestr I
Grupa projektowa: 2**

Kraków, 2021 rok

Spis treści

| | |
|--|----|
| Wstęp..... | 3 |
| Cel pracy..... | 3 |
| Pytania badawcze i hipotezy..... | 3 |
| Wprowadzenie do badań | 5 |
| Część teoretyczna (metody) | 5 |
| Test istotności wartości oczekiwanej | 5 |
| Test istotności dwóch frakcji..... | 6 |
| Test zgodności z rozkładem normalnym | 7 |
| Analiza wariancji | 7 |
| Test istotności dwóch wariancji | 8 |
| Część praktyczna..... | 9 |
| Charakterystyka danych | 9 |
| Testy | 10 |
| Wstęp | 10 |
| Test istotności wartości oczekiwanej zmiennej czas gry | 15 |
| Test istotności wartości oczekiwanej zmiennej czas gry | 16 |
| Test istotności dwóch frakcji | 17 |
| Test istotności wariancji | 18 |
| Analiza wariancji | 19 |
| Wnioski | 20 |
| Bibliografia..... | 23 |
| Spis tabel i rysunków | 23 |

Wstęp

Zawody i inne formy rywalizacji stanowią od dawna część kultury gier komputerowych, ale przez dłuższy czas były traktowane jedynie rozrywkowo, rozgrywane przez amatorów. Dopiero wraz z początkiem lat 2000, dzięki możliwości strumieniowania rozgrywek na żywo, zaczęły nabierać obecnej formy wraz z gwałtownym wzrostem ich popularności. Pojawili się zawodowi gracze oraz komentatorzy, specjalizujący się w danej grze. Szybki wzrost zainteresowania i znaczenia e-sportu spowodował, że już w latach 2010-tych stanowił istotny czynnik w branży gier komputerowych, a twórcy aktywnie wspierali i współtworzyli go, przystosowując swoje gry i finansując turnieje. Dużą popularnością cieszy się gra League of Legends, która na całym świecie organizuje turnieje z nagrodami.

Cel pracy

Celem mojej pracy jest analiza czynników, wpływających na przebieg profesjonalnej rozgrywki w grze League of Legends. Swoje badania ograniczyłam do roku 2017 oraz do trzech regionów jakimi są: Europa, Ameryka i Korea. Zamierzam pokazać różne zależności pomiędzy danymi zebranymi na temat profesjonalnych rozgrywek i które z nich istotnie wpływają na przebieg rozgrywki.

Pytania badawcze i hipotezy

1. Czy można przyjąć że średnia czasu gry jest równa 36,1?

H0: $m = 36,1$ – średnia długość czasu gry jest równa 36,1 minut.

H1: $m \neq 36,1$ – średnia długość czasu gry nie jest równa 36,1 minut.

2. Czy można przyjąć że średnia czasu gry jest wyższa w Europie niż w Korei?
H0: $m_1 = m_2$ – średnia długość czasu gry w Europie jest równa średniej długości czasu gry w Korei
H1: $m_1 > m_2$ – średnia długość czasu gry w Europie nie jest równa średniej długości czasu gry w Korei
3. Czy można przyjąć że odsetek gier postacią Caitlyn jest taki sam jak odsetek gier postacią Xayah?
H0: $p_1 = p_2$ – odsetek gier zagranych postacią Caitlyn jest taki sam jak odsetek gier zagranych Xayah
H1: $p_1 \neq p_2$ – odsetek gier zagranych postacią Caitlyn nie jest taki sam jak odsetek gier zagranych Xayah
4. Czy można stwierdzić, że odsetek zagranych liczby gier przez Jankosa jest taki większy jak Perkza?
H0: $p_1 = p_2$ – odsetek gier zagranych przez Jankosa jest taki sam jak Perkza
H1: $p_1 > p_2$ - odsetek gier zagranych przez Jankosa jest większy niż Perkza
5. Czy istnieją statystyczne różnice średnich czasów rozgrywki w 3 badanych regionach?
H0: $m_1 = m_2 = m_3$ – wartości średnie w badanych regionach są równe
H1: $m_i = m_j$ – dla pewnych i, j – wartości średnie czasu rozgrywki w danych regionach nie są równe
6. Czy można przyjąć że wariancja złota w 15 minucie jest niższa dla ADC niż Supporta?
H0: wariancja zdobytego złota w 15 minucie gry przez ADC jest równa wariancji zdobytego złota w 15 minucie gry przez Supporta
H1: wariancja zdobytego złota w 15 minucie gry przez ADC jest większa niż wariancja zdobytego złota w 15 minucie gry przez Supporta

Wprowadzenie do badań

Badania zostały wykonane w środowisku języka R za pomocą programu Rstudio. We wszystkich moich badaniach przyjmuję poziom istotności $\alpha=5\%$.

Część teoretyczna (metody)

Test istotności wartości oczekiwanej

Test t-Studenta {t.test ()} zakłada normalny rozkład obserwacji w każdej populacji. Służy on do porównywania wartości średnich odpowiadających poszczególnym obserwacjom.¹ T.test pozwala nam na przeprowadzenie testu na jeden lub dwóch próbach przy założeniu, że rozkład danych jest w normalny. Formuła testu pozwala nam zweryfikować 3 rodzaje hipotez. Kiedy chcemy porównać wartości oczekiwane nie dodajemy żadnego argumentu, jeśli chcemy zbadać nierówność to w przypadku znaku większości używamy alternative="greater" a w przypadku znaku mniejszości alternative="less". Wynik testu przedstawia nam wartość p-value, które interpretujemy w zależności od poziomu istotności. Jeśli wartość p-value jest niższa lub równa poziomowi istotności 0,05, możemy odrzucić hipotezę zerową i zaakceptować hipotezę alternatywną.

Do testowania, czy wartość średniej (parametru położenia) w jednej próbie jest równa zadanej wartości μ . W tym przypadku hipoteza zerowa jest postaci:

$$H_0: \mu = \mu_0$$

gdzie μ to nieznany parametr średniej rozkładu próby x, a μ_0 to zadana stała.

¹ Biecek P – „przewodnik po pakiecie R” s. 206

Do testowania równości dwóch średnich (parametrów położenia) dla dwóch prób. W tym przypadku hipoteza zerowa jest postaci:

$$H_0: \mu_1 = \mu_2$$

gdzie μ_1 to nieznany parametr średniej dla rozkładu pierwszej próby, μ_2 to nieznany parametr średniej dla rozkładu drugiej próby.²

Test istotności dwóch frakcji

Używamy jej gdy, populacja generalna ma rozkład zero-jedynkowy z parametrem p . Parametr p jest prawdopodobieństwem, że badana cecha z populacji przyjmuje wyróżnioną wartość (jedną z dwóch możliwych). Można go interpretować jako frakcję elementów populacji mających tę wartość. Przy tej interpretacji parametr ten jest nazywany wskaźnikiem struktury w populacji.³

W języku R używamy funkcji `prop.test()`.

W przypadku testowania równości wskaźników struktury w różnych grupach hipoteza zerowa jest postaci:

$$H_0: p_1 = p_2$$

Gdzie p_i to prawdopodobieństwo sukcesu w grupie i . Innymi słowy hipoteza zerowa to przypuszczenie, że prawdopodobieństwa sukcesów we wszystkich grupach są równe.⁴

² „przewodnik po pakiecie R” Biecek P.

³ „Statystyka od podstaw” Janina Józwiak, Jarosław Podgórski

⁴ „przewodnik po pakiecie R” Biecek P.

Test zgodności z rozkładem normalnym

Normalna zmienna losowa to ciągła zmienna losowa, która pojawia się jako granica skokowej zmiennej dwumianowej, gdy liczba doświadczeń, n , wzrasta. Rozkład normalny nie jest tylko granicznym przypadkiem rozkładu dwumianowego. Jest on rozkładem wielu zmiennych losowych w sytuacjach, w których ujawniają się skutki różnych przypadkowych czynników pochodzących z różnych źródeł. Rozkład normalny był kiedyś nazywany prawem (rozkładem) błędów, ponieważ przypadkowe skutki o różnym znaczeniu i różnych źródłach pochodzenia były traktowane jako „błędy”. Rozkład normalny jest tak rozpowszechniony, że Francis Galton nazwał go „najwyższym prawem przypadkowości”⁵

Shapiro.test – Test Shapiro-Wilka jest to jeden z najbardziej popularnych i jednocześnie jeden z lepszych testów normalności. Liczba obserwacji powinna mieścić się w przedziale od 3 do 5000, ponieważ dla takich liczebności stabilizowane są wartości kwantyli statystyki testowej. Wywołanie tej funkcji z wektorem o większej lub mniejszej liczebności zakończy się komunikatem o błędzie.⁶

Analiza wariancji

Metoda statystyczna służąca do badania obserwacji, które zależą od jednego lub wielu działających równocześnie czynników. Metoda ta wyjaśnia, z jakim prawdopodobieństwem wyodrębnione czynniki mogą być powodem różnic między obserwowanymi średnimi grupowymi. Analiza wariancji została stworzona w latach dwudziestych przez Ronalda Fishera.

Modele analizy wariancji można podzielić na:

modele jednoczynnikowe – wpływ każdego czynnika jest rozpatrywany oddzielnie, tą klasą zagadnień zajmuje się jednoczynnikowa analiza wariancji

modele wieloczynnikowe – wpływ różnych czynników jest rozpatrywany łącznie, tą klasą zagadnień zajmuje się wieloczynnikowa analiza wariancji.

⁵ „Statystyka w zarządzaniu – pełny wykład” Amir D. Aczel

⁶ „przewodnik po pakiecie R” Biecek P.

Analizę wariancji scharakteryzujemy następująco. Obserwujemy n obiektów, które można pogrupować ze względu na pewną cechę jakościową (jedną lub więcej cech) X . Dla każdego z tych obiektów obserwujemy pewną interesującą nas cechę ilościową Y . Interesuje nas, czy średnia wartość cechy Y różni się w różnych grupach cechy X . Zakładamy, że dla ustalonej grupy zmienna Y ma rozkład normalny o wariancji σ^2 , równej dla każdej populacji. Nieparametrycznym odpowiednikiem analizy wariancji test Kruskala-Wallisa dla wielu prób. Jest uogólnieniem testu Wilcoxa na więcej niż dwie podpopulacje.⁷

Test istotności dwóch wariancji

Przy konstrukcji testu weryfikującego hipotezę o identyczności wariancji w dwóch populacjach generalnych korzystamy z rozkładu F-Snedecora. Niezbędne założenia to dwie populacje mają rozkład normalny.⁸

Do testowania równości dwóch wariancji (parametru położenia) dla dwóch prób. W tym przypadku hipoteza zerowa jest postaci:

$$H_0: S^2 = S_0^2$$

gdzie S^2 to nieznany parametr wariancji rozkładu próby x , a S_0^2 to to nieznany parametr wariancji dla rozkładu drugiej próby.⁹

`var.test()` jest to test F dla dwóch prób do weryfikacji hipotezy o jednorodności wariancji. Formuła testu pozwala nam zweryfikować 3 rodzaje hipotez. Kiedy chcemy porównać wariancje nie dodajemy żadnego argumentu, jeśli chcemy zbadać nierówność to w przypadku znaku większości używamy `alternative="greater"` a w przypadku znaku mniejszości `alternative="less"`. Wynik testu przedstawia nam wartość `p-value`, które interpretujemy w zależności od poziomu istotności. Jeśli wartość `p-value` jest niższa lub równa poziomowi istotności 0,05, możemy odrzucić hipotezę zerową i zaakceptować hipotezę alternatywną.

⁷ „przewodnik po pakiecie R” Biecek P.

⁸ „Statystyka od podstaw” Janina Józwiak, Jarosław Podgórski

⁹ „przewodnik po pakiecie R” Biecek P.

Część praktyczna

Charakterystyka danych

Dane zostały pobrane ze strony:

LINK: <https://www.kaggle.com/chuckephron/leagueoflegends?select=LeagueofLegen>

Do swoich badań użyłam danych zatytułowanych matchinfo.csv oraz gold.csv. Zaimportowałam bazy do środowiska R-Studio odpowiednio dobierając właściwości takie, żeby uwzględniono nazwy kolumn.

Przedstawiam opisane dane, których użyłam w ramach mojego projektu

DANE

| Nazwa zmiennej | Opis |
|----------------|---|
| Address | Link do strony z której zebrane zostały dane, używam ich jako klucza danych |
| League | Region, w którym została rozegrana gra |
| Year | Rok, w którym została rozegrana gra |
| gamelength | Długość czasu trwania gry |
| blueTeamTag | Nazwa drużyny, która grała po stronie blue ¹⁰ |
| redTeamTag | Nazwa drużyny, która grała po stronie red ¹¹ |
| bResult | Wynik gry dla drużyny po stronie blue |
| rResult | Wynik gry dla drużyny po stronie red |
| blueJungle | Nazwa zawodnika grającego Jungle ¹² po stronie blue |
| redJungle | Nazwa zawodnika grającego Jungle po stronie red |
| blueMiddle | Nazwa zawodnika grającego Middle po stronie blue |
| redMiddle | Nazwa zawodnika grającego Middle po stronie red |
| blueADCChamp | Nazwa postaci, która można grać na roli ADC po stronie blue |
| redADCChamp | Nazwa postaci, która można grać na roli ADC po stronie red |
| min_15 | Złoto zdobyte w 15 minucie gry |
| Type | Typ zmiennej złoto |
| Min_15.x | Złoto zdobyte przez ADC w 15 minucie gry po stronie blue |
| Min_15.y | Złoto zdobyte przez Supporta w 15 minucie gry po stronie blue |

¹⁰ Blue – niebieska strona-w grze występuje dwie strony: niebieska i czerwona.

¹¹ Red- czerwona strona - w grze występuje dwie strony: niebieska i czerwona.

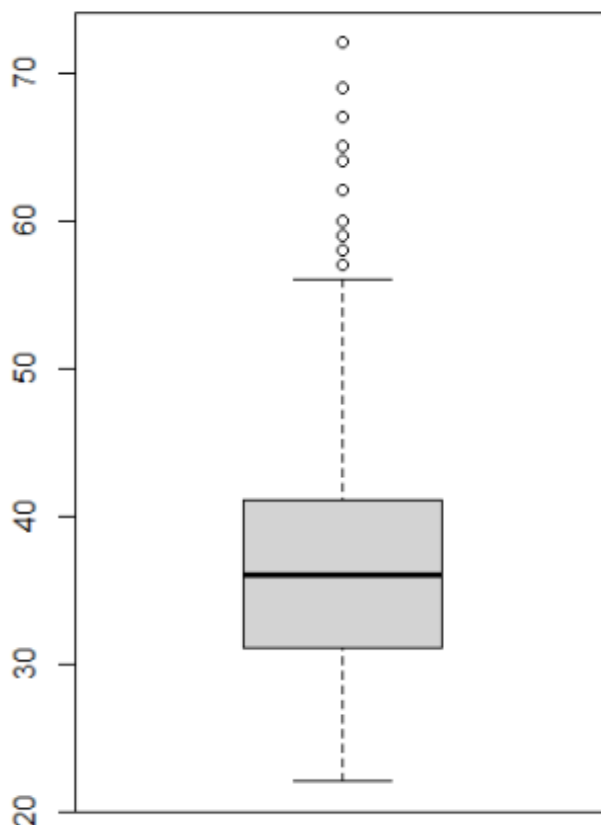
¹² Jedna z ról, którymi można grać w grze. Jest ich pięć (Top,Middle,Jungle,ADC,Support)

Testy

Wstęp

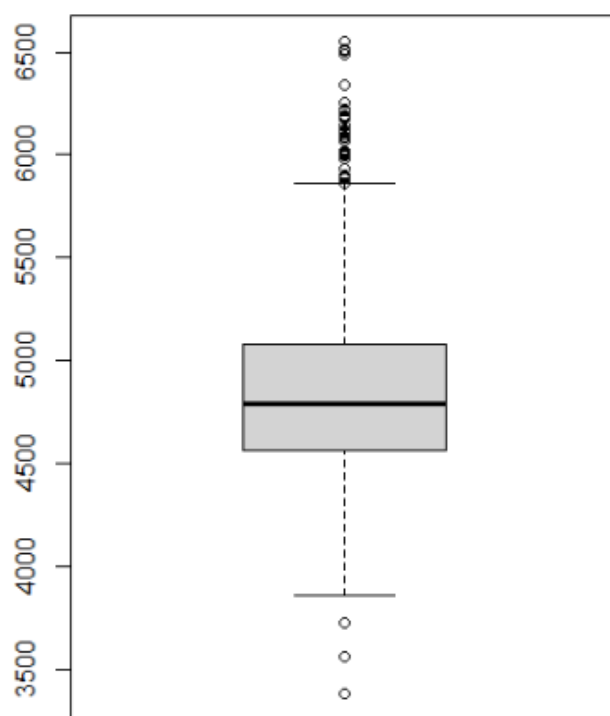
Przed rozpoczęciem badań przygotowałam odpowiednią próbkę.

Połączyłam te dwie bazy w jedną przy pomocy funkcji merge w R. Z bazy wybrałam interesujące mnie kolumny, dane dla roku 2017 i dla regionów takich jak Ameryka, Europa i Korea. Następnie utworzyłam losową próbkę 1000 elementów. Usunęłam puste dane i zajęłam się badaniem danych odstających.



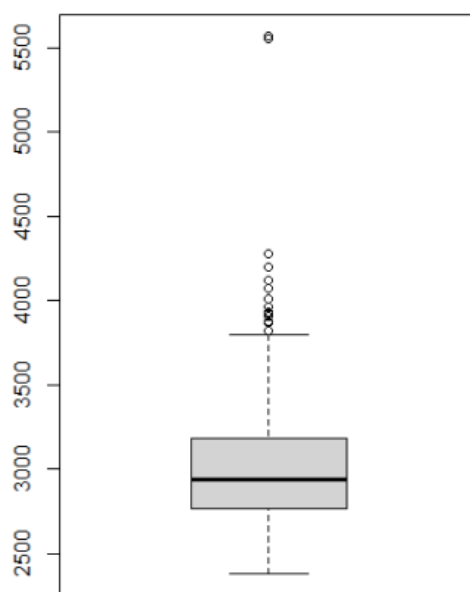
Rysunek – wykres pudełkowy zmiennej czas gry

Na powyższym rysunku przedstawiony został wykres pudełkowy zmiennej czas gry, Dostrzegłam odstające dane dlatego postanowiłam ograniczyć moją próbkę do danych, w których czas gry jest mniejszy niż 55 minut. Następnie wykonałam wykres pudełkowy dla zmiennej min_15.x



Rysunek – wykres pudełkowy zmiennej `min_15.x`

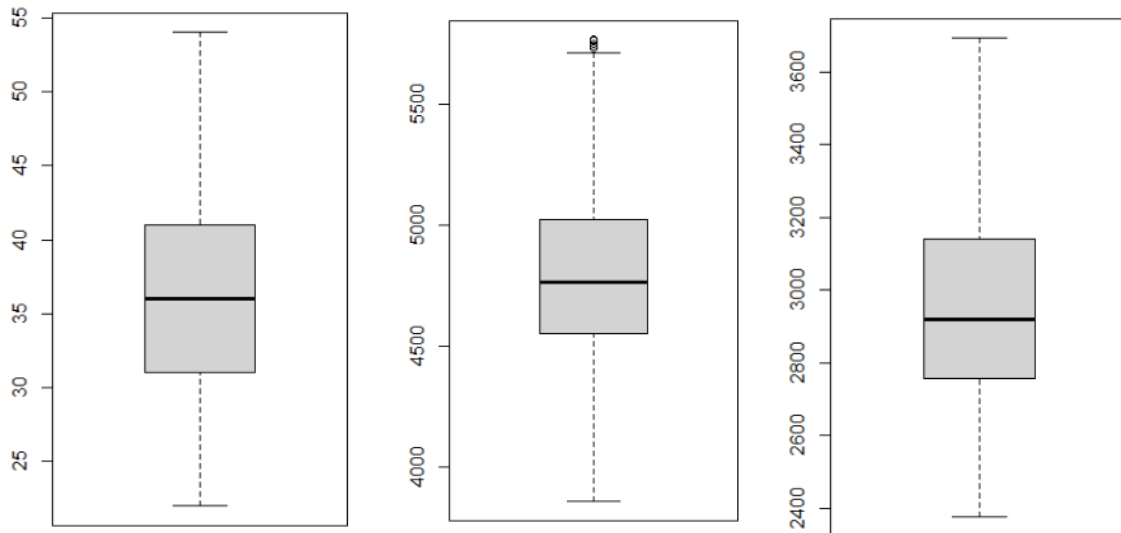
Wykazuje on dane odstające, które później ograniczyłam do danych gdzie złoto w 15 minucie dla ADC jest większe od 3800 ale mniejsze niż 5800.



Rysunek – wykres pudełkowy zmiennej `min_15.y`

Dane z zmiennej min_15.y ograniczyłam do danych gdzie złoto w 15 minucie dla Supporta jest mniejsze niż 3700.

Wykresy pudełkowe ostatecznie prezentują się w następujący sposób:



czas gry, złoto w 15 minucie dla ADC, złoto w 15 minucie dla Supporta

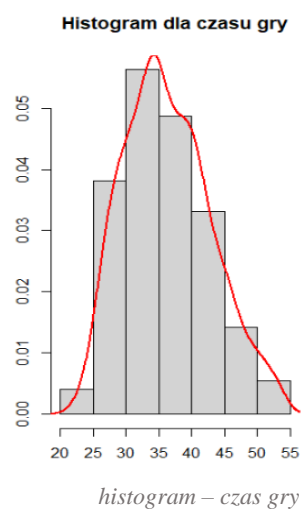
Dokonałam obliczeń wskaźników dla zmiennej czas gry

```
> v  
[1] 43.63608  
> sd  
[1] 6.605761  
> m  
[1] 36.38277
```

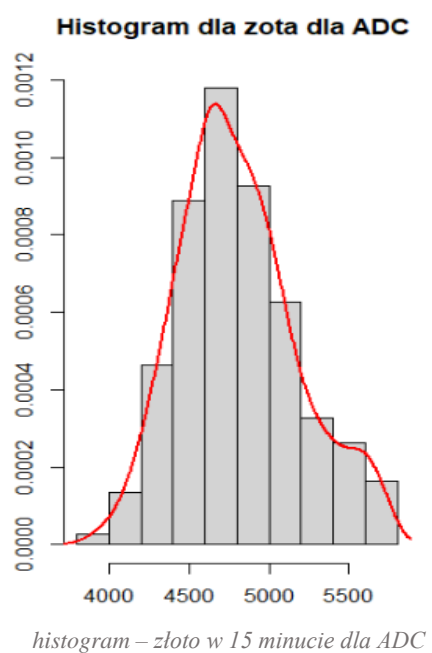
Wskaźniki zmiennej czas gry

Histogramy i krzywa gęstości:

- czas gry

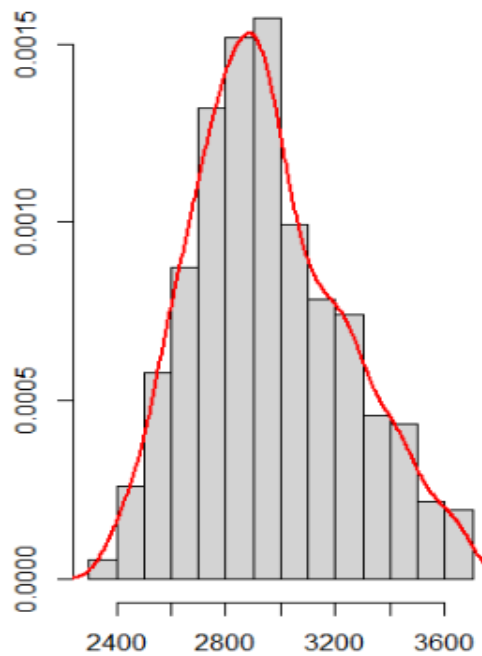


- złoto w 15 minucie dla ADC

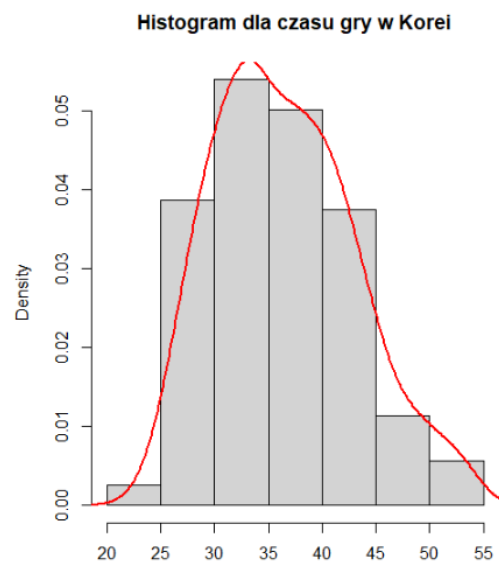
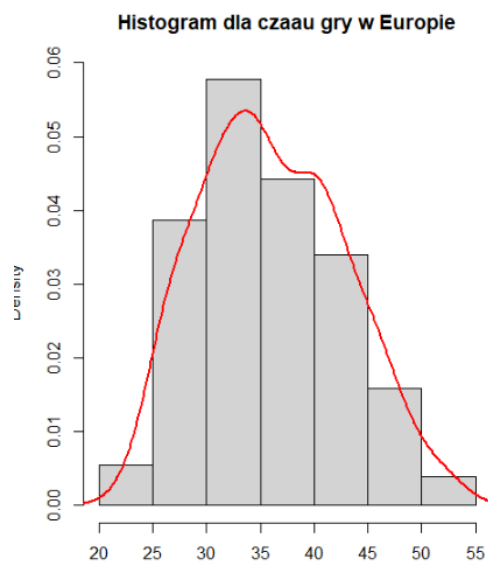


- złoto w 15 minucie dla Support

Histogram dla zota dla Supporta



Histogram- złoto w 15 minucie dla Supporta



Czas gry w Europie i Korei

Również wykonałam histogramy dla czasu gry w Europie i Korei

Wszystkie te histogramy wykazują rozkład danych, na potrzeby moich badań uznaję, że mają one w przybliżeniu rozkład normalny.

Przed wykonaniem testu dla dwóch frakcji musiałam stworzyć odpowiedni zestaw danych. Dla danych, które mnie interesowały napis zmieniałam na wartość 1 lub 0. Następnie połączyłam kolumny blue i red aby wykazać realną wartość wybieranych postaci lub rozegranych gier.

Test istotności wartości oczekiwanej zmiennej czas gry

H0: $m = 36,1$ – średnia długość czasu gry jest równa 36,1 minut.

H1: $m \neq 36,1$ – średnia długość czasu gry nie jest równa 36,1 minut.

```
One Sample t-test

data: probaf$gamelength
t = 1.676, df = 914, p-value = 0.09407
alternative hypothesis: true mean is not equal to 36.1
95 percent confidence interval:
 36.03750 36.89364
sample estimates:
mean of x
 36.46557
```

Rysunek 1. Wyniki testu wartości oczekiwanej w R

Przedział ufności

```
> hip1$conf.int
[1] 36.03750 36.89364
attr(,"conf.level")
[1] 0.95
> |
```

Rysunek 2. Przedział ufności wartości oczekiwanej

Wniosek: W wykonanym teście wartości oczekiwanej $p\text{-value} > 0,05$ zatem nie ma podstaw do odrzucenia hipotezy głównej. Na poziomie ufności 95% możemy stwierdzić, że średnia wartość czasu gry jest równa 36,1 minut.

Test istotności wartości oczekiwanej zmiennej czas gry

$H_0: m_1 = m_2$ – średnia długość czasu gry w Europie jest równa średniej długości czasu gry w Korei

$H_1: m_1 > m_2$ – średnia długość czasu gry w Europie nie jest równa średniej długości czasu gry w Korei

```
Welch Two Sample t-test

data: probaf_Europa$gamelength and probaf_Korea$gamelength
t = -0.62434, df = 583.04, p-value = 0.7337
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -1.248324      Inf
sample estimates:
mean of x mean of y
 36.15693  36.50000
```

Rysunek 3. Test dwóch wartości oczekiwanych

Wniosek: W wykonanym teście wartości oczekiwanej $p\text{-value} > 0,05$ zatem nie ma podstaw do odrzucenia hipotezy głównej. Na poziomie ufności 95% możemy stwierdzić, że średni czas gry w Europie jest taki sam jak w Korei.

Test istotności dwóch frakcji

H0: $p_1 = p_2$ – odsetek gier zagranych postacią Caitlyn jest taki sam jak odsetek gier zagranych Xayah

H1: $p_1 \neq p_2$ – odsetek gier zagranych postacią Caitlyn nie jest taki sam jak odsetek gier zagranych Xayah

```
2-sample test for equality of proportions with continuity correction

data:  c(xayah$suma, caitlyn$suma) out of c(xayah$nc, caitlyn$nc)
X-squared = 3.2331, df = 1, p-value = 0.07216
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.068394548  0.002820777
sample estimates:
   prop 1    prop 2 
0.1551913 0.1879781
```

Rysunek 4. Test istotności dwóch frakcji

Wniosek: W wykonanym teście wartości oczekiwanej $p\text{-value} > 0,05$ zatem nie ma podstaw do odrzucenia hipotezy głównej. Na poziomie ufności 95% możemy stwierdzić, że odsetek gier zagranych Caitlyn jest taki sam jak odsetek gier zagranych Xayah. Czyli gracze wybierają te dwie postacie równie chętnie.

H0: $p_1 = p_2$ – odsetek gier zagranych przez Jankosa jest taki sam jak Perkza

H1: $p_1 > p_2$ - odsetek gier zagranych przez Jankosa jest większy niż Perkza

```
2-sample test for equality of proportions with continuity correction

data:  c(Jankos$suma, perkz$suma) out of c(Jankos$nc, perkz$nc)
X-squared = 0.15882, df = 1, p-value = 0.6549
alternative hypothesis: greater
95 percent confidence interval:
 -0.02459942  1.00000000
sample estimates:
   prop 1    prop 2 
0.05573770 0.06120219
```

Wniosek: W wykonanym teście wartości oczekiwanej $p\text{-value} > 0,05$ zatem nie ma podstaw do odrzucenia hipotezy głównej. Na poziomie ufności 95% możemy stwierdzić, że odsetek gier zagranych przez Jankosa jest taki sam jak odsetek gier zagranych przez Perkza.

Test istotności wariancji

H0: wariancja zdobytego złota w 15 minucie gry przez ADC jest równa wariancji zdobytego złota w 15 minucie gry przez Supporta

H1: wariancja zdobytego złota w 15 minucie gry przez ADC jest większa niż wariancja zdobytego złota w 15 minucie gry przez Supporta

```
F test to compare two variances
data: probaf$min_15.x and probaf$min_15.y
F = 1.876, num df = 914, denom df = 914, p-value < 2.2e-16
alternative hypothesis: true ratio of variances is greater than 1
95 percent confidence interval:
 1.682519      Inf
sample estimates:
ratio of variances
      1.876039
```

Rysunek 6. Test istotności dwóch wariancji

Wniosek: W wykonanym teście wartości oczekiwanej $p\text{-value} < 0,05$ zatem odrzucamy hipotezę główną na rzecz hipotezy alternatywnej. Na poziomie ufności 95% możemy stwierdzić, że wariancja zdobytego złota w 15 minucie gry przez ADC jest większa niż wariancja zdobytego złota w 15 minucie gry przez Supporta.

Analiza wariancji

$H_0: m_1=m_2=m_3$ – wartości średnie w badanych regionach są równe

$H_1: m_i=m_j$ – dla pewnych i,j – wartości średnie czasu rozgrywki w danych regionach nie są równe

Sprawdzam założenie o normalności rozkładu reszt z ANOVY

H_0 : Reszty rozkładu ANOVA mają rozkład normalny

H_1 : Reszty rozkładu ANOVA nie mają rozkładu normalnego

Używam testu Shapiro-Wilka

```
Shapiro-Wilk normality test
data:  aov_residuals
W = 0.98227, p-value = 4.148e-09
> |
```

test Shapiro-Wilka dla reszt ANOVY

Wniosek: W wykonanym teście $p\text{-value} < 0,05$ zatem odrzucamy hipotezę główną na rzecz hipotezy alternatywnej. Na poziomie ufności 95% możemy powiedzieć, że reszty ANOVA nie pochodzą z rozkładu normalnego.

Założenia ANOVY nie zostały spełnione dlatego stosuję jej nieparametryczny odpowiednik a mianowicie test Kruskala-Wallisa.

$H_0: m_1=m_2=m_3$ – wartości średnie w badanych regionach są równe

$H_1: m_i=m_j$ – dla pewnych i,j – wartości średnie czasu rozgrywki w danych regionach nie są równe

```

Kruskal-Wallis rank sum test

data:  gamelength by League
Kruskal-Wallis chi-squared = 0.33083, df = 2, p-value = 0.8475
```

test Kruskala-Wallisa

Wniosek: W przeprowadzonym teście $p\text{-value} > 0,05$ dlatego nie ma podstaw do odrzucenia hipotezy głównej. Na poziomie ufności 95% możemy stwierdzić, że wartości średnie czasu gry w poszczególnych regionach są sobie równe.

Wnioski

1. Czy można przyjąć że średnia czasu gry jest równa 36,1?

Analizując wybrane przez mnie dane pod kątem średniej długości możemy stwierdzić, że jest ona równa 36,1 minuty. Większość gier kończy się w okolicy tej średniej pomijając odstające dane, gdzie czasem zdarzało się, że rozgrywka trwała znacznie dłużej albo bardzo krótko.

2. Czy można przyjąć że średnia czasu gry jest wyższa w Europie niż w Korei?

Z moich badań wynika, że średnia długość czasu gry w Europie i w Korei jest taka sama co oznacza, że nie ma znaczenia w jakim regionie aktualnie miała miejsce rozgrywka i region nie wpływa na długość czasu rozgrywki. Jest to istotne z punktu widzenia zawodników, ponieważ dzięki temu mają oni świadomość, że to w jakim miejscu się znajdują nie wpływa na przebieg ich rozgrywki. Czy będą oni grać w Europie czy w Korei średni czas rozgrywki będzie taki sam.

3. Czy można przyjąć że odsetek gier postacią Caitlyn jest taki sam jak odsetek gier postacią Xayah?

Gracze równie chętnie wybierają postać Caitlyn i Xayah. Możemy więc stwierdzić, że wartość tych postaci jest taka sama. Gracze chcą mieć możliwość zagrania jedną z tych postaci. Odsetek częstotliwości wybierania tych postaci jest dość wysoki jest to około 17% zważając na to, że liczba postaci, które mają do wyboru jest równa około 140.

4. Czy można stwierdzić, że odsetek zagranych liczby gier przez Jankosa jest taki większy jak Perkza?

Według moich badań odsetek jest taki sam co znaczy, że na przestrzeni wszystkich rozegranych gier zegrali oni taką samą ilość. Możemy stwierdzić więc, że grają oni w tej samej drużynie. Po dokonaniu sprawdzenia czy rzeczywiście grali oni w 2017 roku w tej samej drużynie jest prawdą.

5. Czy istnieją statystyczne różnice średnich czasów rozgrywki w 3 badanych regionach?

Na podstawie moich badań wynika, że średnia długość czasu gry w poszczególnych regionach jest taka sama, dlatego można stwierdzić, że przebieg rozgrywki nie zależy od regionu w jakim jest rozgrywana i zawodowi gracze nie powinni brać tego czynnika za istotny aspekt przebiegu gry.

6. Czy można przyjąć że wariancja złota w 15 minucie jest niższa dla ADC niż Supporta?

Po dokonaniu testu okazało się, że wariancja złota zdobytego przez ADC w 15 minucie jest większa niż złoto zdobyte przez Supporta w 15 minucie. Zróznicowanie pod względem zdobytego złota przez ADC jest większe niż Supporta. Mówi to nam o tym, że rozbieżność złota ADC jest duża, więc może on mieć bardzo małą ilość złota albo bardzo dużą, gdzie w przypadku Supporta jest większe prawdopodobieństwo utrzymania poziomu złota na jednakowym poziomie.

Podsumowanie:

W grze League of Legends jest wiele zależności. Możemy mówić o tym, że pewnie czynniki są dla graczy bardziej lub mniej istotne. Czynniki zewnętrzny taki jak region nie ma znaczenia dla graczy w aspekcie czasu trwania rozgrywki. Natomiast czynnik związany z wyborem postaci już jest ważny. Nie bez powodu w moim teście wykazałam, że granie tymi dwoma postaciami jest taka sama. Wskazuje to na wartość tych postaci w tamtym okresie, które miały wpływ na przebieg rozgrywki. Gra League

of Legends to bardzo zaawansowana statycznie gra i można by dokonywać analizy na różnych płaszczyznach i patrzeć na tą grę pod różnym kątem. W dzisiejszym świecie jest coraz większy nacisk na analizę prowadzonych rozgrywek aby pomóc ulepszyć swoje umiejętności i zrozumieć co tak naprawdę ma znaczenie w tej grze.

Bibliografia

- „Statystyka od podstaw” Janina Józwiak, Jarosław Podgórski
- „przewodnik po pakiecie R” Biecek P.
- „Statystyka w zarządzaniu – pełny wykład” Amir D. Aczel

Spis tabel i rysunków

- DANE
- Rysunek – wykres pudełkowy zmiennej czas gry
- Rysunek – wykres pudełkowy zmiennej min_15.x
- Rysunek – wykres pudełkowy zmiennej min_15.y
- czas gry, złoto w 15 minucie dla ADC, złoto w 15 minucie dla Supporta
- Wskaźniki zmiennej czas gry
- histogram – czas gry
- histogram – złoto w 15 minucie dla ADC
- histogram – złoto w 15 minucie dla Supporta
- Czas gry w Europie i Korei
- Rysunek 1. Wyniki testu wartości oczekiwanej w R
- Rysunek 2. Przedział ufności wartości oczekiwanej
- Rysunek 3. Test dwóch wartości oczekiwanych
- Rysunek 4. Test istotności dwóch frakcji
- Rysunek 5. Test istotności dla dwóch frakcji pytania badawczego 5
- Rysunek 6. Test istotności dwóch wariancji
- test Shapiro-Wilka dla reszt ANOVY
- test Kruskala-Wallisa