УНИВЕРЗИТЕТ У БЕОГРАДУ
МАТЕМАТИЧКИ ФАКУЛТЕТ

Анита Златић

# Неки тестови засновани на биномној расподели

— мастер–рад —

Београд, 2017.

# Садржај

**1 Увод** **1**
1.1 Биномна расподела . . . 1
1.2 Тестирање хипотеза . . . 7

**2 Статистички тестови** **10**
2.1 Биномни тест . . . 12
2.1.1 Примери . . . 18
2.2 Квантил тест . . . 25
2.2.1 Примери . . . 33

**3 Моћ теста** **39**
3.1 Биномни тест-симулација у R-у . . . 42
3.2 Биномни тест-теоријски начин . . . 44
3.3 Квантил тест-симулација у R-у . . . 46

**4 Закључак** **52**

# Поглавље 1

# Увод

## 1.1 Биномна расподела

За опис овог поглавља користили смо књигу "Статистичке медоте у метеорологији и инжењерству", чији су аутори В. Јевремовић, Ј. Малишић.

Посматрајмо извођење неког експеримента, у коме се једно извођење назива опит. Претпоставимо да су у сваком појединачном опиту могућа само два, узајамно искључива исхода, које називамо "успех" и "неуспех". Ова два термина (успех и неуспех) се користе само ради разграничавања два могућа исхода и не означавају "квалитет" неког догађаја. Велики број експеримената може се статистички класификовати на овај начин. На пример, приликом бацања новчића могућа су само два догађаја: пад писма или грба, произведени артикал може бити исправан или неисправан, испит студент може положити или пасти, и још доста других примера. У сваком од ових примера, у зависности шта је предмет нашег истраживања, један догађај означавамо као "успех", а други као "неуспех". Такав опит, који може имати само два исхода назива се Бернулијев опит, по швајцарском математичару Јакобу Бернулију (1654 - 1705).

*Јакоб Бернули 1671.године магистрирао је филозофију на Универзитету у Базелу, а 1676. дипломирао теологију, по жељи својих родитеља. У исто време студирао је математику и астрономију против жеље родитеља. Јакоб је започео математичарску традицију у својој породици.*

Кажемо да случајна промењива има Бернулијеву расподелу са параметром $p$, $0 < p < 1$, ако је:

$$X : \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}.$$

Тачније ако је $P(X=0) = 1-p$ и $P(X=1) = p$. Тако случајна промењива добија вредност 1 при позитивној реализацији опита, а 0 при негативној реализацији. Обично се вероватноћа $1-p$ обележава са $q$.

Међутим, нас неће интересовати само један Бернулијев опит, већ низ независних, поновљених Бернулијевих опита. Такав низ опита назива се Бернулијев процес, уколико су испуњени следећи услови:

(1) Сваки опит има само два могућа исхода: успех(U) и неуспех(N).

(2) Вероватноћа успеха $p = P(U)$ је константна од опита до опита. Вероватноћа неуспеха, $P(N) = 1 - p$, означава се са $q$, тако да је $p + q = 1$.

(3) Опити су независни, односно, реализација опита било ког исхода у једном опиту нема утицаја на вероватноћу исхода у било ком другом опиту.

Уколико извршимо $n$ поновљених Бернулијевих опита, тада број успеха може износити $0, 1, 2 \ldots, n$. Циљ нам је да одредимо формулу за израчунавање вероватноће остваривања сваког могућег броја успеха унутар $n$ независних, узастопно поновљених опита, који формирају Бернулијев процес. Расподела до које долазимо на овакав начин назива се биномна расподела.

*Дефинишимо сада биномну расподелу на следећи начин.*

Нека се, под истим условима и независно један од другог, изводи $n$ експеримената и нека је вероватноћа реализације догађаја A у сваком од тих експеримената константна и једнака $p$. Ако је случајна променљива X једнака броју реализација догађаја A при описаним условима, тада кажемо да X има биномну расподелу са параметрима $n$ и $p$: $X \sim \mathrm{B}(n, p)$.

Дакле, случајна променљива са биномном расподелом:

$$X : \begin{pmatrix} 0 & \cdots & n \\ p_{n:0} & \cdots & p_{n:n} \end{pmatrix}, \; p_{n,k} = \binom{n}{k} p^k (1-p)^{n-k}, \; k = 0, 1, \ldots, n.$$

Приметимо да је биномна расподела дискретна расподела, јер случајна променљива може узимати само целобројне вредности из интервала $0, 1, \ldots, n$. Она задовољава оба услова које мора да испуни било која расподела вероватноћа – свака вероватноћа која фигурише у расподели вероватноћа је ненегативна и сума свих вероватноћа једнака је 1. Видимо да у формули вредности $n$ и $p$ нису унапред одређене. Јасно је да ћемо мењањем појединачних вредности $n$ и $p$ добијати различите биномне расподеле, од којих свака задовољава исту математичку формулу коју смо изнад навели, али придодаје различите вероватноће појединим вредностима случајне променљиве. Са повећањем вредности броја $n$, рачунање биномних вероватноћа може захтевати доста времена, стога су статистичари конструисали специјалне таблице са већ израчунатим вредностима. Ако је број опита велики, и налази се изван опсега датим у таблицама израчунавање вероватноће за одговарајуће вредности случајне променљиве применом изнад наведене формуле постаје компликовано. Уколико је вероватноћа успеха $p$ веома мала (најчешће се узима да је $p \leq 0.05$) и када је $n \geq 20$, уместо биномног модела можемо користити Пуасонов модел. Приметимо да када је $p$ блиско нули заправо испитујемо неке ретке догађаје. Наведимо пример биномне расподеле:

**Пример**. Три нумерисане коцкице за игру се бацају и у сваком бацању се бележи који је највећи од добијених бројева. Нека је А догађај: највећи од добијених бројева је једнак 2. Одредити закон расподеле случајне променљиве $S_5$ која је једнака броју реализација догађаја А у 5 бацања те три коцкице.

**Решење**: Најпре треба да одредимо вероватноћу догађаја А, нека је једнака $p$, па ће случајна променљива $S_5$ имати биномну расподелу са параметрима 5 и $p$. Догађај А да је највећи од добијених бројева једнак је 2 може се остварити ако се добијају бројеви:

1,1,2; 1,2,1; 2,1,1; 1,2,2; 2,1,2; 2,2,1 или 2,2,2;
па је вероватноћа $P(A) = p = \frac{7}{6^3} = 0.0324$.
Закон расподеле случајне променљиве $S_5$ је:
$P[S_5 = k] = \binom{5}{k} p^k (1-p)^{5-k}$ где је $k = 0, 1, \ldots, 5$.

Када све вероватноће израчунамо на три децимале добијамо:

$$S_5 : \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 0.848 & 0.142 & 0.009 & 0.000 & 0.000 & 0.000 \end{pmatrix}.$$

*Поступак*: користећи формулу $P[S_5 = k] = \binom{5}{k} p^k (1-p)^{5-k}$ добијамо:
k=0

$$P[S_5 = k] = \binom{5}{k} p^k (1-p)^{5-k} = \binom{5}{0} \cdot 0.0324^0 \cdot (1-0.0324)^{5-0} = \frac{5!}{0!(5-0)!} \cdot 0.9676^5 = 0.9676^5 = 0.848163$$

заокруживањем броја на три децимале добијамо 0.848.
к=1

$$P[S_5 = k] = \binom{5}{k} p^k (1-p)^{5-k} = \binom{5}{1} \cdot 0.0324^1 \cdot (1-0.0324)^{5-1} = \frac{5!}{1!(5-1)!} \cdot 0.0324 \cdot 0.9676^4 = 5 \cdot 0.0324^1 \cdot 0.8765636 = 0.1420033$$

заокруживањем броја на три децимале добијамо 0.142.
к=2

$$P[S_5 = k] = \binom{5}{k} p^k (1-p)^{5-k} = \binom{5}{2} \cdot 0.0324^2 \cdot (1-0.0324)^{5-2} = \frac{5!}{2!(5-2)!} \cdot 0.0324^2 \cdot 0.9676^3 = 10 \cdot 0.0324^2 \cdot 0.9059153 = 0.009409936$$

заокруживањем броја на три децимале добијамо 0.009.
На исти начин се ради за вредности $k = 3$, $k = 4$, $k = 5$.

Резултати су следећи:
к=3

$$P[S_5 = k] = 0.0003184393$$

заокруживањем броја на три децимале добијамо 0.000.
к=4

$$P[S_5 = k] = 0.000005331$$

заокруживањем броја на три децимале добијамо 0.000.
к=5

$$P[S_5 = k] = 0.00000003570467$$

заокруживањем броја на три децимале добијамо 0.000.

Вероватноћа догађаја ($S_5 = 3$), ($S_5 = 4$) и ($S_5 = 5$) нису једнаке 0, али су мање од 0.0005, па стога имамо наведени резултат. Збир свих вероватноћа у закону расподеле мора бити једнак 1, што значи да разлика од 0.001 која овде постоји, резултат заокруживања на три децимале.

Ако случајна променљива $X$ има $B(n_1, p)$ расподелу, а случајна променљива $Y$ има $B(n_2, p)$, ако су $X$ и $Y$ независне, тада случајна променљива $Z = X + Y$ има $B(n_1 + n_2, p)$ расподелу. Ово можемо објаснити на следећи начин: Ако је вероватноћа извлачења беле куглице из кутије једнака $p$ и $n_1$ пута бирамо са враћањем по једну куглицу, и ако је $X$ укупан број изабраних белих куглица, тада $X$ има $B(n_1, p)$ расподелу. Ако наставимо експеримент и изведемо још $n_2$ бирања (такође са враћањем), укупан број белих куглица који тада добијемо је случајна променљива $Y$ са $B(n_2, p)$ расподелом. Обједињени, ови експерименти дају укупан број белих куглица при $n_1 + n_2$ бирања, што одговара случајној променљивој $Z$ са $B(n_1 + n_2, p)$ расподелом.

Покажимо овај пример у R-у:
Користимо функцију **dbinom**($x, size, prob$), затим стављамо наше вредности у ову формулу и добијамо следећи резултат:

> **dbinom(x=0:5, size=5, prob=0.0324)**
> [1] 8.481630e-01 1.420033e-01 9.509936e-03 3.184394e-04 5.331457e-06 3.570467e-08

Ако добијени резултат редом заокружимо на три децимале добијамо:

> **round(dbinom(x=0:5, 5, prob=0.0324), digits=3)**
> [1] 0.848 0.142 0.010 0.000 0.000 0.000

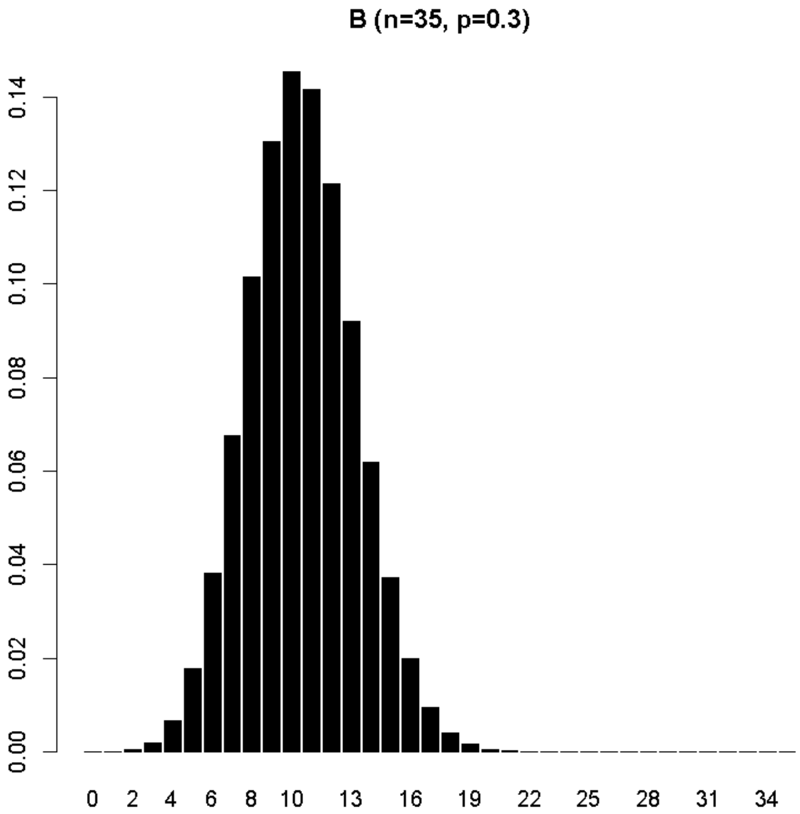
Можемо приметити да смо добили исте резултате.

*Графички приказ биномне расподеле.*

Расподеле дискретне случајне променљиве могу се приказати графички. Покажимо сада неке од графика биномне расподеле:

Узећемо да је обим узорка $n = 35$ и да је вероватноћа $p = 0.3$.
```
x <- 0:35
> x
[1] 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
27 28 29 30 31 32 33 34 35
> prob <- dbinom(x, 35, 0.3)
> barplot(prob, names.arg = x, col = "black", main="B(n=35, p=0.3)")
```

Види функцију dbinom(x, 35, 0.3) приказану на графику испод.

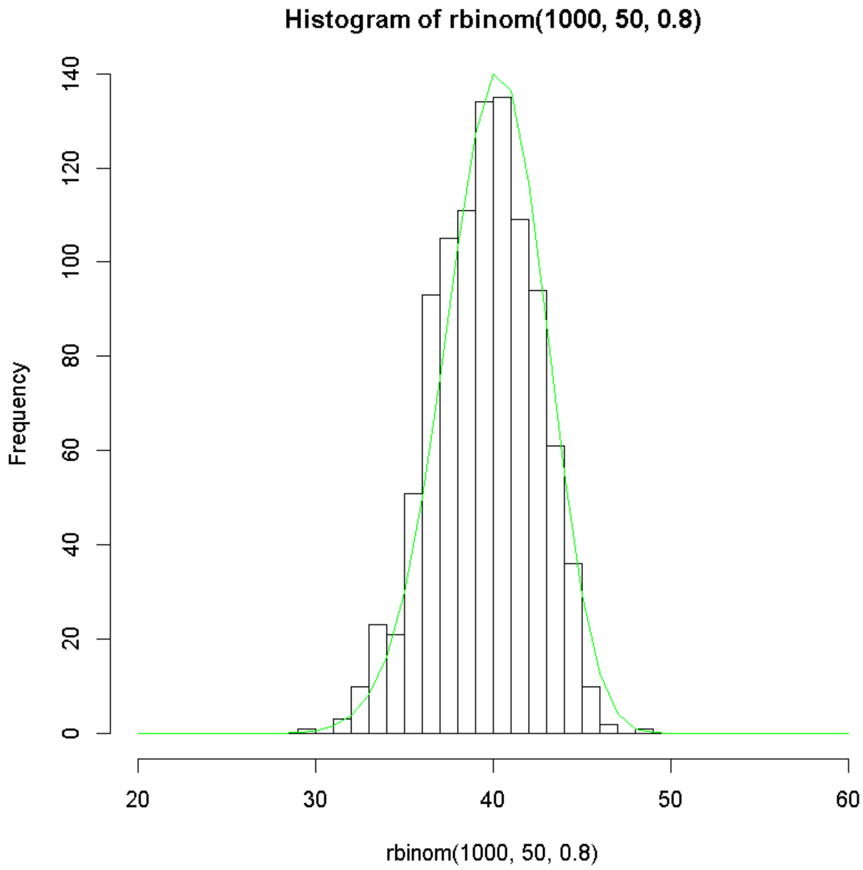



Слика 1.1: Биномна расподела

Најчешће се дискретне случајне променљиве које узимају велики број различитих вредности, графички приказују у облику хистограма.

$> hist(rbinom(1000, 50, 0.8), breaks = seq(20, 60, 1))$
$> lines(seq(20, 60, 1), 1e{+}3{*}dbinom(seq(20, 60, 1), 50, 0.8), col = "green")$

Види на слици испод rbinom(1000,50,0.8) функцију приказану помоћу хистограма.




Слика 1.2: Биномна расподела

Математичко очекивање, дисперзија и стандардно одступање дискретне случајне променљиве дате биномним законом расподеле износи:

- $E(X) = n \cdot p$
- $D(X) = n \cdot p \cdot q$
- $\sigma = \sqrt{D(X)} = \sqrt{n \cdot p \cdot q}$

*Апроксимација биномне расподеле Пуасоновом.*

Наиме, споменули смо да је Пуасонова расподела настала као гранични случај биномне расподеле, тачније када у биномној расподели са параметрима $n$ и $p$ proizvod $np \to \lambda$ ако је $n$ велико а $p$ мало, тада биномна расподела тежи Пуасоновој расподели са параметром $\lambda$. У пракси замена се врши ако је $n \cdot p \leq 10$.

- $P(X = k) \approx e^{-\lambda} \cdot \frac{\lambda^k}{k!}$
- $P(X \leq m) \approx \sum_{k=0}^{m} e^{-\lambda} \cdot \frac{\lambda^k}{k!}$
- $P(X > m) \approx 1 - \sum_{k=0}^{m} e^{-\lambda} \cdot \frac{\lambda^k}{k!}$

О тестирању хипотеза у следећем поглављу, додатне информације смо преузели из материјала за наставу, Економског факултета Универзитета у Крагујевцу.

## 1.2 Тестирање хипотеза

Тестирање хипотеза има велику примену како у практичним тако и теоријским истраживањима. Данас је, на пример у SAD незамисливо добити одобрење за дистрибуцију новог лека од FDA (Food and Drug Administration) пре детаљног испитивања ефикасности и сигурности тог лека, у коме кључну улогу има тестирање хипотеза. Такође публиковање радова у референтим стручним часописима из различитих области као што су: медицина, фармација, психологија, биологија, педагогија, и друге области, често подразумева да су тврдње и хипотезе поткрепљене резултатима неког статистичког теста. Циљ тестирања је да се испита прихватљивост неког тврђења или претпоставке која се тиче карактеристика једног или више основних скупова. На пример, тестирањем можемо проверити следеће ставке:

- Да ли је просечна дужина студирања на приватним и државним факултетима једнака?
- Постоји ли веза између незапослених и броја криминалних дела?
- Да ли узимање антиоксиданта успорава старење, и слично?

Можемо да приметимо да на свако постављено питање, постоје два могућа одговора: "да" или "не". Наиме, можемо рећи да се тестирање хипотезе врши ако су испуњена следећа три услова:

(1) Располажемо неким претходним сазнањем (претпоставком) о карактеристици скупа;

(2) На истраживачко питање (проблем), постоје само два могућа одговора, "да" или "не";

(3) Проблем се решава коришћењем случајног узорка.

Статистичка хипотеза је прецизно формулисана претпоставка о некој важној особини једног или више скупова. Статистички метод који на систематски начин емпиријске податке узорка користи ради утврђивања прихватљивости хипотезе назива се тестирањем хипотезе, а сама процедура статистичким тестом. Наведимо кораке при тестирању хипотезе:

(1) Формулишу се **нулта и алтернативна хипотеза** и бира **ниво значајности** $\alpha$.

(2) **Узима се случајан узорак**, и на основу анализирања постављеног проблема **бира по истраживачевом мишљењу оптималан тест.**

(3) **Проверава се да ли су испуњени предуслови** на којима се заснива изабрани тест. Ако нису прекида се процес тестирања и бира неки други тест, који почива на блажим, испуњеним условима.

(4) Израчунава се **тест статистика**. Одређује се **p-вредност** (алтернативно одређују се **критичне вредности** из таблица одговарајућих расподела).

(5) Поставља се **правило одлучивања**. На основу p-вредност (алтернативно критичних вредности) доноси се одлука о одбацивању или неодбацивању нулте хипотезе.

(6) Формулише се **закључак**, односно даје одговор на постављени проблем.

*Постављање нулте и алтернативне хипотезе.*

Нулта и алтернативна хипотеза представљају два прецизна, међу собом искључива исказа или тврђења. Нулта хипотеза може бити проста и сложена. Нулта хипотеза је проста ако се њом тврди да је параметар једнак тачно једној, унапред познатој нумеричкој вредности, тзв. хипотетичној вредности. Нулту хипотезу можемо написати речима, али је, по правилу, пишемо симболима, на пример, $H_0 : \mu = 300$. Ако нулта хипотеза обухвата већи број могућих вредности, на пример, $H_0 \leq 300$, она је сложена. Свакој нултој хипотези супротстављамо (придружујемо) алтернативну хипотезу и означавамо је са $H_1$ или $H_a$.

**Алтернативна хипотеза** ($H_1$ или $H_a$) је исказ (тврђење) о карактеристици једног или више скупова које је супротно тврђењу садржаном у нултој хипотези. Алтернативна хипотеза садржи вредности које параметар може имати, а које нису обухваћене нултом хипотезом. Због тога је дата у облику сложене хипотезе. Алтернативну хипотезу, којом могућа одступања стварне од хипотетичке вредности параметра пратимо у оба смера називамо двосмерном или двостраном хипотезом. Тест који се примењује у оваквој ситуацији назива се двосмерним или двостраним тестом.

**Ниво значајности теста** (назива се још ризиком прве врсте) је вероватноћа да ћемо одбацити истиниту нулту хипотезу. Обележава се са $\alpha$. Међутим, како узорак никад није савршено репрезентативан, могућа су следећа два исхода: да информација из узорка противречи истинитој нултој хипотези, или да је сагласна са неистинитом нултом хипотезом. Очигледно је да у оба случаја доносимо погрешну одлуку. То значи да постоји могућност да направимо две различите грешке у одлучивању:

(1) Прва грешка је да одбацимо истиниту нулту хипотезу. Таква грешка се назива грешка прве врсте и означавамо је са $\alpha$.

(2) Другу грешку у тестирању чинимо ако не одбацимо нетачну нулту хипотезу. То је грешка друге врсте и означавамо је са $\beta$.

Поступак тестирања спроводимо тако што унапред фиксирамо ризик грешке прве врсте, тј. ниво значајности $\alpha$. При томе бирамо релативно мали ниво значајности, али не толико мали да он онемогућава одбацивање сваке нулте хипотезе. Већина истраживача користи углавном само два нивоа значајности: $0,05$ и $0,01$.
Наиме, Фишер је сматрао да ако при тестирању већ морамо да се суочимо са могућношћу јављања грешке, онда је потребно да се негде повуче демаркациона линија, а он је лично преферирао да то буде 5%.

*Избор теста.*

Оптималан тест је онај тест који за фиксни ниво значајности $\alpha$ има највећу моћ. Моћ теста је вероватноћа да се одбаци погрешна нулта хипотеза. Моћ теста зависи од идентичних фактора као и ризик грешке $\beta$ па је јасно да њен ниво не можемо прецизно израчунати за конкретну хипотезу. При избору теста треба да изаберемо онај који највише одговара емпиријским подацима (тј. тест за који су задовољени полазни услови), али и тест који ће за изабрани ниво грешке прве врсте имати највећу моћ, и то за унапред изабрану величину узорка, $n$.

*Испитивање да ли су услови на којима се тест заснива испуњени.*

Будући да сваки тест почива на специфичним условима који морају бити испуњени да би резултат био валидан, треба изабрати онај тест који је оптималан у датој ситуацији, а то значи тест за који су предуслови примене задовољени. Наиме, током већег дела XX века коришћени су искључиво параметарски тестови, при чему је услов нормалности скупа ретко провераван због компликоване и непоуздане процедуре. Ситуација се потпуно изменила, са општом доступношћу рачунара и статистичког софтвера, предуслови примене тестова се могу лако проверити, а сам тест применити на податке из узорка. Због тога нагласак у модерном тестирању хипотеза више није пука примена неког теста, већ пре свега:

- Испитивање да ли су предуслови на којима се тест заснива испуњени.
- Конкретна интерпретација резултата.

*Тест статистика и p-вредност*

**Статистика теста** је критеријум на основу којег вршимо тестирање. Израчунавање вредности статистике теста из доступних података и поређење са регионом прихватања и одбацивања који су већ дефинисани. **p-вредност** је вероватноћа да статистика теста узме вредност једнаку или још екстремнију од вредности која се управо реализовала у узорку, под условом да је нулта хипотеза тачна. Што је мања **p-вредност**, јачи су докази против нулте хипотезе.

*Правило одлучивања*

**Правило одлучивања на основу p-вредности гласи**: Ако је p-вредност мања од нивоа значајности $\alpha$, нулта хипотеза се одбацује. У супротном, кажемо да немамо довољно аргумената да одбацимо $H_0$. Можемо рећи, у случајевима када користимо статистички софтвер, односно, када располажемо са p-вредношћу, правило одлучивања постаје веома лако. Ако је вредност мања од нивоа значајности $\alpha$, кажемо да је добијени резултат статистички значајан, и нулту хипотезу одбацујемо са ризиком $\alpha$.

За опис наредног поглавља користили смо књигу "Practical Nonparametric Statistics", аутора W.J. Conover. Поред ове књиге користили смо још и материјале за наставу Математичког факултета Универзитета у Београду, које је објавила Марија Радичевић.

# Поглавље 2

# Статистички тестови

Статистичке тестове можемо класификовати по три основа:

(1) Према природи проблема које решавамо: тестови о параметрима скупа (аритметичкој средини, пропорцији, медијани, варијанси, ...), о облику расподела скупа, о независности два обележја, о случајности узорка и других.

(2) Према броју узорка на којима заснивамо тестирање делимо у три групе:

   – Тестови засновани на једном узорку
   – Тестови засновани на два узорка
   – Тестови засновани на три и више узорака

(3) Према врсти и јачини предуслова на којима се заснивају:

   – Параметарски тестови
   – Непараметарски тестови

Задржимо се на трећој подели.

Непараметарски тестови, представљају логичку реакцију статистичара на чињеницу да се претпоставка о „нормалности" свих појава показала нетачном. Ови тестови се заснивају на много блажим условима о облику основног скупа, а могу се примењивати и на квалитативна обележја. Непараметарски тестови могу бити примењени у свим условима у којима нису испуњене претпоставке за примену параметарских тестова. Такође могу бити примењени и у условима када јесу испуњене претпоставке за примену непараметарских тестова, али тада предност треба дати параметарским тестовима јер су они снажнији. За сваки параметарски тест постоји најмање један еквивалентан непараметарски тест.

Опште претпоставке које захтевају:

- Случајност узорка
- Независност опсервација (изузев код поновљених мерења)

*Неки непараметарски тестови:*

(1) Непараметарски тестови за један узорак.

   - Хи-квадрат тест
   - **Биномни тест**
   - Тест корака
   - Колмогоров - Смирнов тест
   - Вилкоксонов тест ранга са знаком

(2) Непараметарски тестови за два независна узорка.

   - Тест медијане
   - Ман - Витни U тест
   - Колмогоров - Смирнов тест
   - Валд - Волфовитцов тест
   - Мозесов тест

(3) Непараметарски тестови за два зависна узорка.

   - Мек Немаров тест
   - Тест знакова
   - Маргинални тест хомогености

(4) Непараметарски тестови за K независних узорака.

   - Тест медијане
   - Крускал - Валисов тест

(5) Непараметарски тестови за K зависних узорака.

   - Фридманов тест
   - Кендалов тест $W$

Непараметарски тестови имају мању моћ од параметарских. У овом раду описаћемо **биномни тест** и **квантил тест**. Такође, видећемо њихову примену у различитим областима.

## 2.1 Биномни тест

Код биномног теста узорак се састоји од $n$ исхода независних експеримената. Сваки исход припада или класи 1 или класи 2 али никако не може припадати обема истовремено. Број исхода у класи 1 означимо са $O_1$, а број исхода у класи 2 означимо са $O_2 = n - O_1$. Претпостављамо такође да је вероватноћа сваког исхода да буде у класи 1 једнака $p$.

*Претпоставке.* Биномни тест се примењује ако су задовољени следећи услови:

(1) Свих $n$ експеримената међусобно су независни.

(2) Сваки покушај са вероватноћом $p$ резултира у класи 1, где је $p$ исто за све покушаје.

*Хипотезе.* Могуће је тестирати три различите хипотезе. Нека $p^*$ буде нека специфична константа, таква да је $0 \leq p^* \leq 1$. Хипотеза може да има један од следећа три облика:

А. (Двострани тест)

$H_0 : p = p^*$
$H_1 : p \neq p^*$

Б. (Једнострани тест)

$H_0 : p \leq p^*$
$H_1 : p > p^*$

В. (Једнострани тест)

$H_0 : p \geq p^*$
$H_1 : p < p^*$

*Статистички тест*: Тест статистика ће представљати број елемената класе 1, тачније тест статистика је:

$$T = O_1$$

*Расподела тест статистике.* Присетимо се Бернулијеве расподеле. Вероватноћа реализације неког догађаја А у сваком понављању експеримената је иста и износи $p$. Експерименти су међусобно независни. Случајна промењлива $I$ има Бернулијеву расподелу ако је њен закон вероватноће:

$$I : \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$$

Дефинишемо случајну промењливу $Y$ која представља број реализација догађаја А у $n$ понављања експеримената.

Дакле, случајну промењљиву $Y$ можемо дефинисати као:

$$Y = \sum_{j=0}^{n} I_j$$

где свако $I_j$ има Бернулијеву расподелу са параметром $p$. Нека је за наш пример вероватноћа догађаја А заправо вероватноћа да експеримент има резултат класе 1. Тада тај догађај поново има Бернулијеву расподелу са параметром $p$, тачније $p$ је вероватноћа да експеримент има резултат у класи 1. У том случају промењљива $Y$ представља бројач експеримената који имају резултат у класи 1, тачније представља број исхода класе 1 од $n$ независних експеримената.
Сума независних случајних величина са Бернулијевом расподелом, има биномну расподелу са параметрима $n$ и $p$. Односно имамо:

$$Y \sim \mathrm{B}\,(n,p)$$
$$P\{Y=k\} = \binom{n}{k} p^k (1-p)^{n-k}$$

*Правило одлучивања.* Зависно од тога која се хипотеза тестира, А, Б или В, следе различита правила одлучивања.

А. (двострани тест) Критична област величине $\alpha$ одговара двостраној биномној расподели са параметрима $p^*$ и $n$, где је $\alpha_2$ величина горњег исхода, а $\alpha_1$ величина доњег исхода, па је $\alpha_1 + \alpha_2$ једнако $\alpha$. За одређене вредности $p^*$ и $n$, потребно је наћи број $t_1$:

$$P(Y \leq t_1) = \alpha_1,$$

и наћи број $t_2$ тако да

$$P(Y > t_2) = \alpha_2,$$

или еквивалентно,

$$P(Y \leq t_2) = 1 - \alpha_2$$

где је $Y$ случајна промењљива са параметрима $p^*$ и $n$.

Вредности $\alpha_1$ и $\alpha_2$ треба да буду приближно једнаке. Одбацити $H_0$ ако Т превазилази $t_2$, или ако је Т мање или једнако од $t_1$. У супротном, треба прихватити $H_0$.

Б. (једнострани тест) Велике вредности Т указују да је $H_0$ нетачна, а величина критичне области $\alpha$, састоји се од свих вредности Т које су веће од $t$, где је $t$ број добијен из таблице, користећи $p^*$ и $n$, тако да:

$$P(Y > t) = \alpha,$$

или еквивалентно,

$$P(Y \leq t) = 1 - \alpha$$

где $Y$ има биномну расподелу са параметрима $p^*$ и $n$. $H_0$ се одбацује ако је Т веће од $t$. $H_0$ се прихвата ако је Т мање или једнако од $t$.

В. (једнострани тест) Мале вредности Т указују да је $H_0$ нетачна, а величина критичне области $\alpha$ састоји се од свих вредности Т мање или једнако од $t$, где је $t$ број добијен из таблице, користећи $p^*$ и $n$, тако да:

$$P(Y \leq t) = \alpha$$

где $Y$ има биномну расподелу са параметрима $p^*$ и $n$. $H_0$ се одбацује ако је Т мање или једнако од $t$. У супротном се $H_0$ прихвата.

**Пример**. Под једноставним Менделовим наслеђем може се очекивати веза између биљака два појединa генотипа за производњу потомства једне четвртине који су "патуљасте" и три четвртине који су "високе". У експерименту је потребно утврдити да ли је претпоставка простог Менделовог наслеђа логична у одређеној ситуацији, а унакрсни резултати у потомству имају 243 патуљастих биљака и 682 високих биљака. Ако "класа 1" означава "високе", $p^* = \frac{3}{4}$ и Т представља број високих биљака. Нулта хипотеза о једноставном Менделевом наслеђу еквивалентна је моделу хипотезе:

$$H_0 : p = \tfrac{3}{4},$$

алтернатива је:

$$H_1 : p \neq \tfrac{3}{4}.$$

Од $n = 952(243 + 682)$, ако је n$\geq$ 30, као што јесте, можемо користити апроксимацију нормалном расподелом, коју нам обезбеђује централна гранична теорема, тако да имамо следеће :

$$t_1 = n \cdot p^* + \omega_{0.025} \cdot \sqrt{n \cdot p^* \cdot (1 - p^*)} = 925 \cdot \tfrac{3}{4} + (-1.960) \cdot \sqrt{925 \cdot \tfrac{3}{4} \cdot \tfrac{1}{4}} = 667.94,$$

где је $\omega_{0.025}$ квантил нормалне расподеле за ниво 0.025.

И све вредности Т веће од $t_2$ где је:

$$t_2 = n \cdot p^* + \omega_{0.975} \cdot \sqrt{n \cdot p^* \cdot (1 - p^*)} = 925 \cdot \tfrac{3}{4} + (1.960) \cdot \sqrt{925 \cdot \tfrac{3}{4} \cdot \tfrac{1}{4}} = 719.56,$$

где је $\omega_{0.975}$ квантил нормалне расподеле за ниво 0.975.

Подсетимо се да централна гранична теорема за последицу има једнакост:

$$\lim_{n \to \infty} P_{H_0}\{\tfrac{S_n - n \cdot p^*}{\sqrt{n \cdot p^* \cdot (1-p^*)}} \leq x\} = \tfrac{1}{2 \cdot \pi} \cdot \int_{-\infty}^{x} e^{\frac{-t^2}{2} dt}.$$

Вредност Т добијена у овом експерименту је 682. Стога је нулта хипотеза прихваћена. Може се наћи критичан ниво $\hat{\alpha}$, прихватањем области теста за исту област са обе стране за $n \cdot p^* = 693.75$. Како је вредност 682 мања од $n \cdot p^*$, половина од $\hat{\alpha}$ се израчунава из $P(T \leq 682)$, претпостављајући да је нулта хипотеза тачна, добијамо:

$P(T \leq 682) = P(\frac{T-n\cdot p^*}{\sqrt{n\cdot p^*\cdot(1-p^*)}} \leq \frac{682-n\cdot p^*}{\sqrt{n\cdot p^*\cdot(1-p^*)}}) \approx P(Z \leq \frac{682-693.75}{13.17})$

где Z има стандардну нормалну расподелу.

$\frac{\hat{\alpha}}{2} = P(T \leq 682) \approx P(Z \leq -0.8922) \approx 0.186$

Због тога је $\hat{\alpha} = 2 \cdot (0.186) = 0.372$. Ниво значајности од најмање 0.372 мора бити потребан да се одбије $H_0$. Тако да су подаци у доброј сагласности са нултом хипотезом.

Претходни пример приказује двострани облик биномног теста. Сада ћемо претходни пример показати у R-у:

**binom.test (c(682, 243), p = 3/4)**

Exact binomial test

data: c(682, 243)
number of successes = 682, number of trials = 925, p-value = 0.3825
alternative hypothesis: true probability of success is not equal to 0.75
95 percent confidence interval:
0.7076683 0.7654066
sample estimates:
probability of success
0.7372973

На основу p-вредности теста која износи p-вредност=0.3825 и већа је од 0.05 (праг значајности), можемо закључити да су подаци у складу са нултом хипотезом.

Да статистички тест у биномном тесту има биномну расподелу, лако се може видети. Ако је T једнак броју понављања који доводе до резултата класе 1, где су испитивања међусобно независна и где свако понављање има вероватноћу $p$ у тим резултатима (као што је наведено у претпоставкама), онда T има биномну расподелу са параметрима $p$ и $n$. Величина критичног региона је максимална када је $p$ једнако $p^*$, под нултом хипотезом, па је уведена таблица са $n$ и $p^*$, да утврди тачну вредност $\alpha$. Као што је раније поменуто, тестирање хипотеза је само један огранак статистичког закључивања.

Сада ћемо разматрати друге гране, **о интервалу поверења**. Ако покушавамо да направимо неке закључке у вези са непознатим параметром који је повезан са неким становништвом, логично је да се испита случајни узорак из те популације, и на основу тог узорка да се има увид у вези са параметром популације. Таква изјава може бити "параметар популације лежи између a и b", где су a и b два реална броја добијена из узорка. Бројеви a и b су израчунати из узорка и реализација су две статистике. Две статистике које нам показују доње и горње граничне тачке за интервал ће бити означене са L и U, односно, "доњи" и "горњи". Интервал од L до U назива се **интервал поверења**. Вероватноћа да непознати параметар лежи у његовом интервалу поверења зове се **коефицијент поверења**. Метода за проналажење интервала поверења за $p$, непознате вероватноће за неки посебан исход, је уско повезан са биномним тестом.

*Интервал поверења.*

*Подаци.* Узорак који се састоји од запажања $n$ независних студија је испитан, а број $Y$ напомиње колико се пута јавља наведени догађај. Претпоставке:

(1) $n$ исходи су међусобно независни.

(2) Вероватноћа $p$ од наведеног исхода остаје константна од једног испитивања до следећег.

Наведимо сада једну од метода интервала поверења: За $n$ веће од 20 може се користити нормална апроксимација. Употреба:

$$L = \frac{Y}{n} - x_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{Y \cdot (n-Y)}{n^3}}$$

и,

$$U = \frac{Y}{n} + x_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{Y \cdot (n-Y)}{n^3}}$$

где је $x_{1-\alpha/2}$ квантил за нормалну расподелу случајне променљиве, добијен из таблице. Објаснимо како смо дошли до овог интервала. Треба пронаћи L и U тако да важи:

$$1 - \alpha = P\{L < p < U\}$$

Ако је $Y$ случајна променљива са биномном расподелом са параметрима $p$ и $n$, тада је на основу централне граничне теореме:

$$Z = \frac{Y - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}} : N(0,1)$$

Сада, ако са $x_{1-\frac{\alpha}{2}}$ означимо квантил нормалне расподеле за одговарајући ниво, те ако приметимо да је $x_{\frac{\alpha}{2}} = -x_{1-\frac{\alpha}{2}}$ имамо,

$$\begin{aligned} 1 - \alpha &= P(-x_{1-\frac{\alpha}{2}} < \frac{Y - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}} < x_{1-\frac{\alpha}{2}} \\ &= P(-x_{1-\frac{\alpha}{2}} \cdot \sqrt{n \cdot p \cdot (1-p)} < Y - n \cdot p < x_{1-\frac{\alpha}{2}} \cdot \sqrt{n \cdot p \cdot (1-p)}). \end{aligned}$$

Помножимо претходну неједнакост са -1. Добијамо:

$$\begin{aligned} 1 - \alpha &= P(-x_{1-\frac{\alpha}{2}} \cdot \sqrt{n \cdot p \cdot (1-p)} < Y - n \cdot p < x_{1-\frac{\alpha}{2}} \cdot \sqrt{n \cdot p \cdot (1-p)}) \\ &= P(Y - x_{1-\frac{\alpha}{2}} \cdot \sqrt{n \cdot p \cdot (1-p)} < n \cdot p < Y + x_{1-\frac{\alpha}{2}} \cdot \sqrt{n \cdot p \cdot (1-p)}). \end{aligned}$$

Поделимо сада све са $n$, имамо:

$$1 - \alpha = P(\frac{Y}{n} - x_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p \cdot (1-p)}{n}} < p < \frac{Y}{n} + x_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p \cdot (1-p)}{n}})$$

Уведемо сада апроксимацију $\frac{Y}{n} \approx p$ коју оправдава закон великих бројева, претходна једначина постаје:

$$\beta \approx P(\frac{Y}{n} - x_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\frac{Y}{n}\cdot(1-\frac{Y}{n})}{n}} < p < \frac{Y}{n} + x_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\frac{Y}{n}\cdot(1-\frac{Y}{n})}{n}}).$$

Ако претходни израз средимо добијамо: $\beta = 1 - \alpha \approx P(L < p < U)$ где је:

$$L = \frac{Y}{n} - x_{1-\frac{\alpha}{2}} \cdot \sqrt{Y \cdot \frac{(n-Y)}{n^3}}$$

и,

$$U = \frac{Y}{n} + x_{1-\frac{\alpha}{2}} \cdot \sqrt{Y \cdot \frac{(n-Y)}{n^3}}.$$

Са једначинама за L и U, за дати ниво значајности $\alpha$ можемо добити границе интервала поверења за вероватноћу $p$. Ради илустрације ове методе користимо следећи пример.

**Пример**. У одређеним државама насумично је одабрано 20 средњих школа да се провери да ли испуњавају стандардне норме, које је предложио национални комитет за образовање. Утврђено је да је 7 школа квалификовано и према томе су означене са "одличан". Који је интервал поверења од 95% за $p$, тј. проценат свих средњих школа у држави која би се квалификовала за обележје "одличан"? Прво, претпоставимо да је број средњих школа у држави довољно велики да се средње школе разврставају са "одличан" или "није одличан" независно једна од друге. (Заправо, за одређен број школа, чињеница да је једна школа класификована на један начин где тежи да повећа шансу да се наредне школске године класификује другачије, јер би онда био нешто већи проценат школа у другој категорији међу онима које још нису изабране). Претпостављајући да је избор био случајан, $p$ је исти за све школе и представља вероватноћу да се изабрана школа обележи са "одличан". Видимо да је $n$ једнако 20 и $Y$ једнако 7, односно $\frac{Y}{n}$ је 0.35 Дакле, наведена метода користи нормалну апроксимацију засновану на централном граничном правилу, која даје:

$$L = \frac{Y}{n} - x_{.975} \cdot \sqrt{\frac{Y\cdot(n-Y)}{n^3}} = 0.35 - (1.960) \cdot \sqrt{7 \cdot \frac{13}{20^3}} = 0.35 - 0.21 = 0.14,$$

и сада имамо да је,

$$U = 0.35 + 0.21 = 0.56.$$

Интервал поверења добијен овом методом је:

$$P(0.14 < p < 0.56) = 0.95.$$

За ову методу је процењено да је боља уколико је $n$ веће ($\geq 30$) и за вредност $p$ близу 0.5, него за вредност $p$ ближе 0 или 1.

Покажимо претходни пример користећи R:

**binom.test (7, 20, p=0.5, conf.level=0.95)**

Exact binomial test

data: 7 and 20
number of successes = 7, number of trials = 20, p-value = 0.2632
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
0.1539092 0.5921885
sample estimates:
probability of success
0.35

Друге методе за добијање биномне границе дали су Андерсон и Бурстеин (1967. и 1968.).

Покажимо још један пример интервала поверења користећи R:

**binom.test (7, 20, p=0.5, conf.level=0.95, alternative="less")**

Exact binomial test

data: 7 and 20
number of successes = 7, number of trials = 20, p-value = 0.1316
alternative hypothesis: true probability of success is less than 0.5
95 percent confidence interval:
0.0000000 0.5580345 sample estimates:
probability of success
0.35

### 2.1.1 Примери

**Пример (1)**.

Зна се да 20% одређене врсте инсеката испољавају специфичну карактеристику А. Осамнаест инсеката те врсте су добијене из необичног окружења, и нико од њих нема карактеристику А. Да ли је разумно претпоставити да инсекти из тог окружења имају исту вероватноћу од 0.20 као инсекти из генералног окружења?

**Решење (1)**.

Нека је нулта хипотеза $H_0(p = 0.2)$, а алтернативна $H_1(p \neq 0.2)$, и $\alpha = 0.05$. Радићемо двострани биномни тест. Постоји $n = 18$ посматрања, и карактеристику А гледамо као на "класу 1", а карактеристику Б као на "класу 2". Посматрана вредност тест статистика Т је $t = 0$, и Т има биномну расподелу $T \sim \mathrm{B}(18, 0.2)$. Добијамо следеће:

$$P(T \leq 0) = P(\tfrac{T-n\cdot p^*}{\sqrt{n\cdot p^*}} \leq \tfrac{0-18\cdot 0.2}{\sqrt{18\cdot 0.2\cdot 0.8}}) \approx P(Z \leq \tfrac{0-3.6}{1.697})$$

где Z има стандардну нормалну расподелу, као што је дато у таблици за нормалну расподелу.

$\frac{\alpha}{2} = P(T \leq 0) \approx P(Z \leq -2.121) \approx 0.017$

Због тога је $\hat{\alpha} = 2 \cdot (0.017) = 0.034$. Тако да подаци нису у сагласности са нултом хипотезом.

Користимо биномни тест у R-у, добијамо следеће резултате:

**binom.test(0,18,0.2)**

Exact binomial test

data: 0 and 18
number of successes = 0, number of trials = 18, p-value = 0.03429
alternative hypothesis: true probability of success is not equal to 0.2
95 percent confidence interval:
0.000000 0.185302
sample estimates:
probability of success
0

На основу p-вредности теста, p=0.0343, која је мања од 0.05, одбацујемо нулту хипотезу да инсекти из специјалне околине имају исту вероватноћу да имају карактеристику А као и инсекти из генералних услова.

**Пример (2)**.

Од 16 возила која су прегледана током кампање о безбедности, утврђено је да је 6 возила несигурно. Тестирати хипотезу да је мање од 10% аутомобила у популацији несигурно. (Претпоставка је да су вероватно биле лажне пријаве).

**Решење (2)**.

Користићемо биномни тест за тестирање да је мање од 10% од аутомобила у узорку неисправно. Поставићемо хипотезу тако да је нулта $H_0(p < 0.1)$, и алтернативну хипотезу $H_1(p \geq 0.1)$.

Статистика T има биномну расподелу са параметром $p$, и са величином узорка $n = 16$. Из таблице за биномну расподелу, $n = 16$, видимо да је $H_0(p < 0.1)$ тачна, односно:

| | | p | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | c | 0.05 | 0.10 | 0.20 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 | 0.95 |
| n = 15 | 0 | 0.463 | 0.206 | 0.035 | 0.005 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 1 | 0.829 | 0.549 | 0.167 | 0.035 | 0.005 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 2 | 0.964 | 0.816 | 0.398 | 0.127 | 0.027 | 0.004 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 3 | 0.995 | 0.944 | 0.648 | 0.297 | 0.091 | 0.018 | 0.002 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 4 | 0.999 | 0.987 | 0.836 | 0.515 | 0.217 | 0.059 | 0.009 | 0.001 | 0.000 | 0.000 | 0.000 |
| | 5 | 1.000 | 0.998 | 0.939 | 0.722 | 0.403 | 0.151 | 0.034 | 0.004 | 0.000 | 0.000 | 0.000 |
| | 6 | 1.000 | 1.000 | 0.982 | 0.869 | 0.610 | 0.304 | 0.095 | 0.015 | 0.001 | 0.000 | 0.000 |
| | 7 | 1.000 | 1.000 | 0.996 | 0.950 | 0.787 | 0.500 | 0.213 | 0.050 | 0.004 | 0.000 | 0.000 |
| | 8 | 1.000 | 1.000 | 0.999 | 0.985 | 0.905 | 0.696 | 0.390 | 0.131 | 0.018 | 0.000 | 0.000 |
| | 9 | 1.000 | 1.000 | 1.000 | 0.996 | 0.966 | 0.849 | 0.597 | 0.278 | 0.061 | 0.002 | 0.000 |
| | 10 | 1.000 | 1.000 | 1.000 | 0.999 | 0.991 | 0.941 | 0.783 | 0.485 | 0.164 | 0.013 | 0.001 |
| | 11 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 | 0.982 | 0.909 | 0.703 | 0.352 | 0.056 | 0.005 |
| | 12 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.996 | 0.973 | 0.873 | 0.602 | 0.184 | 0.036 |
| | 13 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.995 | 0.965 | 0.833 | 0.451 | 0.171 |
| | 14 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.995 | 0.965 | 0.794 | 0.537 |
| | 15 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| n = 16 | 0 | 0.440 | 0.185 | 0.028 | 0.003 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 1 | 0.811 | 0.515 | 0.141 | 0.026 | 0.003 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 2 | 0.957 | 0.789 | 0.352 | 0.099 | 0.018 | 0.002 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 3 | 0.993 | 0.932 | 0.598 | 0.246 | 0.065 | 0.011 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 4 | 0.999 | 0.983 | 0.798 | 0.450 | 0.167 | 0.038 | 0.005 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 5 | 1.000 | 0.997 | 0.918 | 0.660 | 0.329 | 0.105 | 0.019 | 0.002 | 0.000 | 0.000 | 0.000 |
| | 6 | 1.000 | 0.999 | 0.973 | 0.825 | 0.527 | 0.227 | 0.058 | 0.007 | 0.000 | 0.000 | 0.000 |
| | 7 | 1.000 | 1.000 | 0.993 | 0.926 | 0.716 | 0.402 | 0.142 | 0.026 | 0.001 | 0.000 | 0.000 |
| | 8 | 1.000 | 1.000 | 0.999 | 0.974 | 0.858 | 0.598 | 0.284 | 0.074 | 0.007 | 0.000 | 0.000 |
| | 9 | 1.000 | 1.000 | 1.000 | 0.993 | 0.942 | 0.773 | 0.473 | 0.175 | 0.027 | 0.001 | 0.000 |
| | 10 | 1.000 | 1.000 | 1.000 | 0.998 | 0.981 | 0.895 | 0.671 | 0.340 | 0.082 | 0.003 | 0.000 |
| | 11 | 1.000 | 1.000 | 1.000 | 1.000 | 0.995 | 0.962 | 0.833 | 0.550 | 0.202 | 0.017 | 0.001 |
| | 12 | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 | 0.989 | 0.935 | 0.754 | 0.402 | 0.068 | 0.007 |
| | 13 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 | 0.982 | 0.901 | 0.648 | 0.211 | 0.043 |
| | 14 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.997 | 0.974 | 0.859 | 0.485 | 0.189 |
| | 15 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.997 | 0.972 | 0.815 | 0.560 |
| | 16 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

Слика 2.1: Биномна расподела.

Дакле, прихватамо нулту хипотезу да је мање од 10% аутомобила неисправно. Пример је урађен у R-у, добијени су резултати:

**binom.test (6,16,1/10,alternative="less",conf.level=0.90)**

Exact binomial test
data: 6 and 16
number of successes = 6, number of trials = 16, p-value = 0.9995
alternative hypothesis: true probability of success is less than 0.1
90 percent confidence interval:
0.0000000 0.5654405
sample estimates:
probability of success
0.375

На основу p-вредности теста закључујемо да прихватамо нулту хипотезу, зато што је p-вредност већа од 0.05 (праг значајности). Дакле, мање од 10% аутомобила је неисправно.

**Пример (3)**.

У игри коцкица, пар коцкица је бачен 180 пута. Резултат "седам" догодио се 38 пута.

(а) Ако је вероватноћа "седам", шта би било да су коцкице биле другачије?

(б) Наћи интервал поверења од 95% за P (седам).

(в) Наћи интервал поверења од 95% за P (седам), користећи велики узорак.

**Решење (3)**.

Пар нумерисаних коцкица приликом игре бачен је 180 пута. Број 7 се поновио 38 пута.

**(а)** Користићемо биномни тест како бисмо тестирали хипотезу $H_0(p = \frac{1}{6})$, против алтернативне $H_1(p \neq \frac{1}{6})$. Вероватноћу да се понови број 7 приликом бацања коцкица израчунавамо на следећи начин. Вероватноћа добијања "седмице", уколико претпоставимо да су сви бројеви на коцки једнако вероватни је:

$$P(1,6) + P(2,5) + P(3,4) + P(4,3) + P(5,2) + P(6,1) = \frac{6}{36} = \frac{1}{6}$$

Па је вероватноћа да исход буде "седмица" једнака $p = 6 \cdot \frac{1}{6}\frac{1}{6}$:

Од $n = 180$ критична област приближне величине $\alpha = 0.05$ могу се добити помоћу великог узорка приближно датим у таблици. Стога критичан регион одговара свим вредностима T мањим или једнаким од $t_1$, где је:

$$t_1 = n \cdot p^* + \omega_{0.025}\sqrt{n \cdot p^* \cdot (1 - p^*)} = 180 \cdot \frac{1}{6} + (-1.960) \cdot \sqrt{180 \cdot \frac{1}{6}\frac{5}{6}} = 20.26$$

и све вредности T веће од $t_2$, где је

$$t_2 = n \cdot p^* + \omega_{0.975}\sqrt{n \cdot p^* \cdot (1 - p^*)} = 180 \cdot \frac{1}{6} + (1.960) \cdot \sqrt{180 \cdot \frac{1}{6}\frac{5}{6}} = 39.86$$

Вредност T добијена у овом експерименту је 38. Стога је нулта хипотеза прихваћена. Покажимо како изгледа пример у R-у:

**binom.test (38, 180, 1/6)**

Exact binomial test

data: 38 and 180
number of successes = 38, number of trials = 180, p-value = 0.11
alternative hypothesis: true probability of success is not equal to 0.1666667
95 percent confidence interval:
0.1539369 0.2780534
sample estimates:
probability of success
0.2111111

На основу добијене p-вредности теста која је 0.11 и већа је од 0.05 (праг значајности), можемо прихватити нулту хипотезу да су коцкице биле исправне током игре и да овај узорак одговара нултој хипотези.

**(б)** Овај део задатка решавамо користећи R, резултат је следећи:

**binom.test (38, 180, 1/6, conf.level=0.95)**

Exact binomial test

data: 38 and 180
number of successes = 38, number of trials = 180, p-value = 0.11
alternative hypothesis: true probability of success is not equal to 0.1666667
percent confidence interval:
0.1539369 0.2780534
sample estimates:
probability of success
0.2111111

Можемо да закључимо да је интервал поверења: $[0.1539369, 0.2780534]$, или $P[0.15 < p < 0.28] = 0.095$.

**(в)** За $n$ веће од 20 може се користити нормална апроксимација. Употреба методе које смо описали у овом раду, доводи до следећих резултата:
Подсетимо се прво једначина за L и U.

$$L = \frac{Y}{n} - x_{1-\frac{\alpha}{2}} \cdot \sqrt{Y \cdot \frac{(n-Y)}{n^3}}$$
$$U = \frac{Y}{n} + x_{1-\frac{\alpha}{2}} \cdot \sqrt{Y \cdot \frac{(n-Y)}{n^3}}$$

где је $x_{1-\frac{\alpha}{2}}$ квантил за нормалну расподелу случајне променљиве, добијен из таблице.

$$L = \frac{Y}{n} - x_{1-\frac{\alpha}{2}} \cdot \sqrt{Y \cdot \frac{(n-Y)}{n^3}} = \frac{38}{180} - (1.960)\sqrt{38 \cdot \frac{180-38}{180^3}} = 0.1514 \approx 0.15$$
$$U = \frac{Y}{n} + x_{1-\frac{\alpha}{2}} \cdot \sqrt{Y \cdot \frac{(n-Y)}{n^3}} = \frac{38}{180} + (1.960)\sqrt{38 \cdot \frac{180-38}{180^3}} = 0.27$$

Интервал поверења урађен према методи коју смо описали у овом поглављу је:

$$P[0.15 < p < 0.27] = 0.95$$

Исти резултат добијамо када користимо биномни тест у R-у:

**binom.test(38,180,1/6,conf.level=0.95)**

Exact binomial test

data: 38 and 180
number of successes = 38, number of trials = 180, p-value = 0.11
alternative hypothesis: true probability of success is not equal to 0.1666667
percent confidence interval:
0.1539369 0.2780534
sample estimates:
probability of success
0.2111111

**Пример (4).**

У примеру (2), шта је 90% интервал поверења за тачну пропорцију несигурних аутомобила у популацији?

**Решење (4).**

За решавање овог примера можемо користити поступак из примера (2). Пример је урађен у R-у, добијени су резултати:

**binom.test (6,16,1/10,alternative="less",conf.level=0.90)**

Exact binomial test
data: 6 and 16
number of successes = 6, number of trials = 16, p-value = 0.9995
alternative hypothesis: true probability of success is less than 0.1
90 percent confidence interval:
0.0000000 0.5654405
sample estimates:
probability of success
0.375

На основу p-вредности теста закључујемо да прихватимо нулту хипотезу, зато што је p-вредност већа од 0.05 (праг значајности). Дакле, мање од 10% аутомобила је неисправно.

Доносимо закључак да је 90% интервал поверења $[0.0000000, 0.5654405]$.

**Пример (5).**

Двадесет независних запажања случајне промењиве X са непознатом функцијом расподеле F(x) довело је до следећих бројева:

| | | | | |
|---|---|---|---|---|
| 141 | 134 | 98 | 119 | 131 |
| 103 | 154 | 122 | 93 | 137 |
| 86 | 119 | 161 | 144 | 158 |
| 165 | 81 | 117 | 128 | 103 |

Наћи интервал поверења 95% за F(100).

**Решење (5).**

Да бисмо решили овај задатак поставимо да је:

$$F(X) = P(X \leq x),\ F(100) = P(X \leq 100)$$

Како имамо двадесет независних понављања, то значи да нам је $n = 20$. Решење можемо тражити применом апроксимације нормалном расподелом. Такође, можемо користити као у претходном примеру:

$$L = \frac{Y}{n} - x_{1-\frac{\alpha}{2}} \cdot \sqrt{Y \cdot \frac{(n-Y)}{n^3}} = \frac{4}{20} - (1.960)\sqrt{4 \cdot \frac{20-4}{20^3}} = 0.025$$
$$U = \frac{Y}{n} + x_{1-\frac{\alpha}{2}} \cdot \sqrt{Y \cdot \frac{(n-Y)}{n^3}} = \frac{4}{20} + (1.960)\sqrt{4 \cdot \frac{20-4}{20^3}} = 0.375$$

Дакле, 95% интервал је $(0.025, 0.375)$

**Пример (6)**.

Група грађана је пријавила градском већу да су најмање 60% градског становништва корисници одређених обвезница. Градско веће је затим питало случајни узорак од 100 становника, да ли су корисници обвезница. Четрдесет осам је рекло да. Да ли је извештај групе становника разуман?

**Решење (6)**.

Поставићемо нулту хипотезу $H_0(p \geq 0.6)$, против алтернативне хипотезе која је $H_1(p < 0.6)$. Имамо да је $n = 48$.

$$1 - P(Z \leq \frac{48-60}{4.899}) = 1 - P(Z \leq -2.45) = 0.993$$

где $Z$ има стандардну нормалну расподелу. Због тога је $\hat{\alpha} = 0.994$. Тако да су подаци у доброј сагласности са нултом хипотезом.
Решење ћемо представити и у R-у:

**binom.test(48, 100, 0.6, alternative="greater")**

Exact binomial test
data: 48 and 100
number of successes = 48, number of trials = 100, p-value = 0.9942
alternative hypothesis: true probability of success is greater than 0.6
95 percent confidence interval:
0.3941407 1.0000000
sample estimates:
probability of success
0.48

Можемо закључити на основу p-вредности која је 0.0994 и која је већа од 0.05 (праг значајности), да прихватамо нулту хипотезу. На оба начина добијамо исти резултат. Овај узорак одговара нултој хипотези, и да је извештај групе становника добар.

У овом раду описаћемо још један тест који је заснован на биномној расподели, квантил тест.

## 2.2 Квантил тест

Биномни тест се може користити за тестирање хипотезе у вези квантила на случајне променљиве, када га зовемо квантил тест. На пример, можемо испитати случајни узорак вредности неке случајне променљиве $X$, да видимо да ли је средина од $X$ већа од нуле, или једнака 17 (рецимо). Ако се случајна променљива гледа као непрекидна случајна променљива, тестира се хипотеза,

$H_0 : p^*$ квантил од X је $x^*$

исто је као,

$H_0 : P(X \leq x^*) = p^*$

из дефиниције квантила.
Ако $P(X \leq x^*)$ представља непознату вероватноћу за $p$, $H_0$ постаје:

$H_0 : p = p^*$

што је исто нултој хипотези тестираној биномним тестом. Статистички тест једнак је броју вредности узорка које су мање или једнаке од $x^*$, па се може користити двострани биномни тест. Није једноставно ако се претпоставља да је случајна променљива дискретна. Онда је нулта хипотеза:

$H_0 : p^*$ квантил од X је $x^*$

исто је као,

$H_0 : P(X \leq x^*) \geq p^*$ и $P(X < x^*) \leq p^*$

**Сада се може користити биномни тест, али прилагођавање теста овој хипотези је мало незгодно, па ћемо представити поступак као посебан тест.**

*Хипотезе.* Нека $x^*$ и $p^*$ представљају неке специфичне бројеве, $0<p^*<1$. Хипотеза може да узме један од следећа три облика.

**А.** (двострани тест)

$H_0 : p^*$ популациони квантил је $x^*$.

(Ово је еквивалентно са $H_0 : P(X \leq x^*) \geq p^*$, и $P(X < x^*) \leq p^*$ )

$H_1 : x^*$ није $p^*$ популациони квантил.

(Ово је еквивалентно са $H_1 : P(X \leq x^*) < p^*$ или $H_1 : P(X < x^*) > p^*$.

**Б.** (једнострани тест)

$H_0 : p^*$ популациони квантил је бар као $x^*$

(Ово је еквивалентно са $H_0 : P(X < x^*) \leq p^*$).

$H_1 : p^*$ популациони квантил је мањи од $x^*$.

Ово је исто као $H_1 : P(X < x^*) > p^*$.

**В.** (једнострани тест)

$H_0 : p^*$ популациони квантил није већи од $x^*$.

(или је $H_0 : P(X \leq x^*) \geq p^*$)

$H_1 : p^*$ популациони квантил већи од $x^*$.

(или је $H_1 : P(X \leq x^*) < p^*$)

*Тест статистика.* Користићемо две тест статистике. Нека је $T_2$ једнак броју запажања мање него $x^*$, и нека је $T_1$ једнак броју запажања мање или исто $x^*$. Тада је $T_1 = T_2$ ако ниједан од бројева из података није једнако $x^*$. Иначе, $T_1$ је веће од $T_2$.

*Правило одлучивања.* Као у биномном тесту, статистички тест има дискретну расподелу. Дата су различита правила одлучивања, која одговарају хипотези А, Б, или В.

**А.** (двострани тест). Критична област одговара вредности $T_2$ које су превелике (што значи да је можда $P(X < x^*)$ већа од $p^*$), и вредности $T_1$ које су премале (што значи да је можда $P(X < x^*)$ мање од $p^*$). Критична област је нађена уносом у таблицу (таблица за биномну расподелу), са величином узорка $n$ и хипотезом вероватноће $p^*$, као у двостраном биномном тесту. Затим се нађе број $t_1$ тако да,

(1) $P(Y \leq t_1) = \alpha_1$

где $Y$ има биномну расподелу са параметрима $n$ и $p^*$, и где је $\alpha_1$ око половине жељеног нивоа значајности. Затим се нађе број тако да,

(2) $P(Y > t_2) = \alpha_2$

или,

(3) $P(Y \leq t_2) = 1 - \alpha_2$

где је $\alpha_2$ изабрано тако да је $\alpha_1 + \alpha_2$ отприлике једнако за жељени ниво значајности. Одбацити $H_0$ ако је $T_1$ мањи или исти од $t_1$ , или ако је $T_2$ већи од $t_2$ . У супротном, прихватамо $H_0$. Ниво значајности износи $\alpha_1 + \alpha_2$.

**Б.** (једнострани тест) Будући да велике вредности за $T_2$ указују да $H_0$ није тачна, уносом величине узорка $n$ у таблицу (таблица за биномну расподелу), претпоставимо да је $p^*$ као $p$. Затим се нађе број $t_2$ тако да

(4) $P(Y > t_2) = \alpha$

што је исто као

(5) $P(Y \leq t_2) = 1 - \alpha$

за неки прихватљиви ниво значајности $\alpha$. Одбацити $H_0$ ако $T_2$ надмашује $t_2$. Прихватити $H_0$ ако је $T_2$ мање или једнако од $t_2$.

**B.** (једнострани тест) Мале вредности $T_1$ указују да $H_0$ није тачна, па би уносом величине узорка $n$ у таблицу (таблица за биномну расоделу) и наведене вероватноће $p^*$ пронашли $t_1$ тако да,

(6) $P(Y \leq t_1) = \alpha$

за прихватљиви ниво $\alpha$, где $Y$ има биномну расподелу са параметрима $n$ и $p^*$. Одбацити $H_0$ ако је $T_1$ мање или једнако од $t_1$. Прихватити $H_0$ ако $T_1$ прелази $t_1$.

**Пример**. Уписујући колеџ, бруцоши су из средње школе достигли висок ниво образовања за дуги низ година, а горња граница је успостављена резултатом 193. Средња школа шаље петнаест својих дипломираних студената на студије, где су полагањем испита добијени следећи резултати

| | | | | |
|---|---|---|---|---|
| 189 | 233 | 195 | 160 | 212 |
| 176 | 231 | 185 | 199 | 213. |
| 202 | 193 | 174 | 166 | 248 |

Претпоставимо да ових петнаест ученика представљају случајан узорак свих ученика те школе који иду на колеџ. Један од начина поређења студената из средње школе са осталим студентима је тестирање хипотезе, да наведени резултати долазе из популације чија је горња граница 193.
Тако да је:

$H_0$: горња граница је 193,

је испитан против алтернативе,

$H_1$: горња граница није 193,

где се мисли на горњу границу теста за све студенте те средње школе, прошле, садашње или будуће. Примењује се двострани квантил тест. Критична област приближне величине 0.05 се добија уносом у таблицу (таблица за биномну расподелу) за $n = 15$ и $p = 0.75$. За тест статистику $Y$ важи следеће,

(7) $P(Y \leq 7) = 0.0173$

и,

(8) $P(Y \leq 14) = 0.9866$

Критична област величине

(9) $\alpha = 0.173 + 0.0134 = 0.0307$

одговара вредности $T_1$ мање или једнако од $t_1 = 7$, а вредности $T_2$ веће од $t_2 = 14$. У овом примеру $T_1 = 7$, број запажања је мањи од 193 или једнако, и $T_2 = 6$, пошто је једно запажање једнако 193. Стога је $T_1$ премало, а $H_0$ одбачена. Горња граница за студенте из те средње школе није 193.

Пошто је посматрана вредност тест статистике $T_1$ једва у области одбацивања, ниво значајности је мали, што би могао бити резултат одбацивања $H_0$. Зато критичан ниво у овом примеру износи 0.0307, исто као ниво значајности. Једнострани квантилни тест, са приближно великим узорцима, је илустрован у следећем примеру. Покажимо овај пример у R-у:

$> quantile.test < -function(x, xstar = 0, quantile = 0.5, alternative = "two.sided")\{$
$n < -length(x)$
$p < -quantile$
$T1 < -sum(x \leq xstar)$
$T2 < -sum(x < xstar)$
$if(alternative == "quantile.less")\{$
$p.value < -1 - pbinom(T2 - 1, n, p)\}$
$if(alternative == "quantile.greater")\{$
$p.value < -pbinom(T1, n, p)\}$
$if(alternative == "two.sided")\{$
$p.value < -2 * min(1 - pbinom(T2 - 1, n, p), pbinom(T1, n, p))\}$
$list(xstar = xstar, alternative = alternative, T1 = T1, T2 = T2, p.value = p.value)\}$
$> testscores$
$< -c(189, 233, 195, 160, 212, 176, 231, 185, 199, 213, 202, 193, 174, 166, 248)$
$> quantile.test(testscores, xstar = 193, quantile = 0.75, alternative = "two.sided")$
$\$xstar$
$[1]193$
$\$alternative$
$[1]"two.sided"$
$\$T1$
$[1]7$
$\$T2$
$[1]6$
$\$p.value$
$[1]0.03459968$

Приметимо да одбацујемо нулту хипотезу и да узорак није у сагласности са нултом хипотезом. Дакле, закључујемо да горња граница за студенте из средње школе није 193.

**Пример**. Временски интервал између ерупција гејзира је забележен 112 пута, да се види да ли је средњи интервал мањи или једнак од 60 минута (нулта хипотеза) или да ли је средњи интервал већи од 60 минута (алтернативна хипотеза).

Ако је средњи интервал 60, 60 је $x_{0.05}$ или средња вредност. Ако је средњи интервал мањи од 60, 60 је квантил за неке $p \geq 0.50$. Тако је $H_0 = P(X \leq 60) \geq 0.50$, и $H_1 = P(X \leq 60) < 0.50$, где је $X$ временски интервал између ерупција. Претпоставимо да су различити интервали независни и идентично расподељени, може се користити једнострани квантилни тест, са избором правила В. Статистички тест $T_1$ једнак броју интервала који су мање или једнаки од 60 минута, а критични регион величине 0.05 одговара вредности $T_1$ мање од:

$$t_1 = n \cdot p^* + \omega_{0.05} \cdot \sqrt{n \cdot p^* \cdot (1 - p^*)} = 112 \cdot 0.5 - (1.645)\sqrt{122 \cdot 0.5 \cdot 0.5}$$
$$(10) = 47.3.$$

Од 112 пута, њих 8 су 60 минута или мање, тако да је $T_1 = 8$, а $H_0$ је чврсто одбијена у корист алтернативе "средњи временски интервал између ерупција је већи од 60 минута". Критичан ниво је:

$$(11)\ \hat{\alpha} = P(T_8 \leq 8) = P(\tfrac{T-n \cdot p}{\sqrt{n \cdot p \cdot q}}) \cong P[Z \leq \tfrac{8-112 \cdot 0.50}{\sqrt{112 \cdot 0.50 \cdot 0.50}}]$$

где је $Z$ стандардна нормална случајна променљива. Затим, из таблице (таблица за нормалну расподелу),

$$(12)\ \hat{\alpha} = P(Z \leq \tfrac{-48}{5.3}) = P(Z \leq -9.05) \ll 0.0001$$

које читамо "много мање него 0.0001."

*Интервал поверења*

Подаци. Подаци се састоје од запажања за $X_1, \ldots, X_n$ које су независне и идентично распоређене случајне променљиве.

Нека $X^{(1)} \leq X^{(2)} \leq \cdots \leq X^{(r)} \leq \cdots \leq X^{(s)} \leq \cdots \leq X^{(n)}$ представља варијациони низ, где је $1 \leq r < s \leq n$ . Желимо да пронађемо интервал поверења (непознати) за $p^*$ квантил, где је $p^*$ неки одређени број између нула и један.
Претпоставке:

(1) Узорак $X_1, \ldots, X_n$ је случајан узорак,

(2) $X_i s$ је најмањег реда

**МЕТОДА А** (мали узорци). За $n \leq 20$ може се користити таблица (таблица за биномну расподелу) за проналажење $r$ и $s$. Уноси се се са величином узорка $n$ и вероватноћом $p = p^*$. Читати колону за $p = p^*$, све док унос не буде приближно једнак $\frac{\alpha}{2}$, где је $1 - \alpha$ приближно жељени коефицијент поверења. Нека тај унос буде $\alpha_1$, а одговарајућа вредност за $y$ (остатак од $\alpha_1$) је $r-1$. Додамо 1 да добијемо $r$. Настављамо до колоне $p = p^*$, све до постизања уноса који је приближно једнак $1 - (\frac{\alpha}{2})$, коју ћемо назвати $1 - \alpha_2$. Вредност $y$ одговара уносу $1 - \alpha_2$ који се зове $s - 1$, а 1 се додаје како би добили $s$. Тако смо одредили $\alpha_1, \alpha_2$, $r$ и $s$. Тачан коефицијент поверења је $1 - \alpha_1 - \alpha_2$. Онда:

$$(13)\ P(X^{(r)} \leq x_{p^*} \leq X^{(s)}) \geq 1 - \alpha_1 - \alpha_2$$

обезбеђује интервал поверења. Ако претпоставимо да је непозната функција расподеле непрекидна, тада је:

(14) $P(X^{(r)} \leq x_{p^*} \leq X^{(s)}) = 1 - \alpha_1 - \alpha_2$

**МЕТОДА B**. За $n$ већи од 20 може се користити апроксимација заснована на централној граничној теореми. Израчунавањем:

(15) $r^* = n \cdot p^* + \omega_{\frac{\alpha}{2}} \sqrt{n \cdot p^* \cdot (1 - p^*)}$

и,

(16) $s^* = n \cdot p^* + \omega_{1-\frac{\alpha}{2}} \sqrt{n \cdot p^* \cdot (1 - p^*)}$

где се квантиле $\omega_p$ добијају из таблице (таблица за нормалну расподелу), и где је $1 - \alpha$ жељени коефицијент поверења. У принципу, $r^*$ и $s^*$ неће бити цели бројеви. Нека $r$ и $s$ буду цели бројеви добијени заокруживањем $r^*$ и $s^*$ навише у наредне више целе бројеве. Онда је приближан интервал поверења дат једначином (13), или једначином (14), ако је непозната функција расподеле непрекидна. Једнострани интервал поверења се може формирати проналажењем само $r$ или $s$. Једнострани интервали поверења имају облик:

(17) $P(X^{(r)} \leq x_{p^*}) = 1 - \alpha_1$

и,

(18) $P(x_{p^*} \leq X^{(s)}) = 1 - \alpha_2$

ако је функција расподеле непрекидна, или

(19) $P(X^{(r)} \leq x_{p^*}) \geq 1 - \alpha_1$

и,

(20) $P(x_{p^*} \leq X^{(s)}) \geq 1 - \alpha_2$.

**Пример**. Шеснаест цеви је изабрано насумице из једне велике гомиле цеви да се тестирају. Број часова до квара је забележен за сваки од њих. Желимо да пронађемо интервал поверења за горњу границу, са коефицијентом поверења близу 90%. У таблицу (таблица за биномну расподелу) се уноси $n = 16$ и $p = 0.75$. Читањем низ колону за вероватноћу 0.0271 је изабрана близу 0.05. Вредност $y$ повезан са $\alpha_1 = 0.0271$ је $y = 8$, па је $r$ једнако 9. Вероватноћа најближа 0.95 је $0.9365 = 1 - \alpha_2$, који има одговарајући $y$ за 14. Према томе $s = 15$. Интервал поверења је:

(21) $P(X^{(9)} \leq x_{0.75} \leq X^{(15)}) = 0.9094$

(Разумно је претпоставити да је време квара непрекидна случајна промењива, тако да можемо користити једначину(14)).
Следе резултати тестирања у растућем редоследу.

$$\begin{array}{llll}
X^{(1)} = 46.9 & X^{(5)} = 56.8 & X^{(9)} = 63.3 & X^{(13)} = 67.1 \\
X^{(2)} = 47.2 & X^{(6)} = 59.2 & X^{(10)} = 63.4 & X^{(14)} = 67.7 \\
X^{(3)} = 49.1 & X^{(7)} = 59.9 & X^{(11)} = 63.7 & X^{(15)} = 73.3 \\
X^{(4)} = 56.5 & X^{(8)} = 63.2 & X^{(12)} = 64.1 & X^{(16)} = 78.5
\end{array}$$

Зато што су $X^{(9)} = 63.3$ и $X^{(15)} = 73.3$ , можемо рећи "интервал од 63.3 h до $73,3$ h, је интервал поверења од 90.94% за горњи квартил". Коришћењем једначина (15) и (16) добијамо:

$$(22)\ r^* = 16 \cdot 0.75 + (-1.645) \cdot \sqrt{16 \cdot 0.75 \cdot 0.25} = 12 - 2.86 = 9.14$$

и,

$$(23)\ s^* = 12 + 2.86 = 14.86$$

Због тога је $r = 10$ и $s = 15$, па интервал поверења од 90% постаје $(63.4, 73.3)$. Разматрамо прво једноставнији случај, где је непрекидна функција расподеле. Ако $x_{p^*}$ је $p^*$ квантил, имамо:

$$(24)\ P(X \geq x_{p^*}) = P(X > x_{p^*}) = 1 - p^*$$

где је расподела функције X иста као код случајног узорка. Редослед статистике низа 1, $X^{(1)}$, преузима већу вредност од неки одређених константи, само ако је најмања вредност у узорку, већа него константа. Због тога је $X^{(1)}$ већа него што је константа, само ако су све $n$ вредности у узорку веће него константа. Избором $x_{p^*}$ као константе, можемо закључити:

$$\begin{aligned}
&(25)\ P(x_{p^*} < X^{(1)}) = P \text{ (све вредности премашују } x_{p^*})\\
&= P(x_{p^*} < X_1, x_{p^*} < X_2, \ldots, x_{p^*} < X_p)\\
&= P(x_{p^*} < X_1) \cdot P(x_{p^*} < X_2) \ldots P(x_{p^*} < X_p)\\
&= (1 - p^*)^n
\end{aligned}$$

јер су $X_i s$ независне, и све имају исте $p^*$ квантиле $x_{p^*}$. Ако је $x_{p^*}$ мање од $X^{(2)}$, где су $n-1$ запажања већа него $x_{p^*}$, у том случају $X^{(1)} \leq x_{p^*} < X^{(2)}$, или су $n$ запажања већа од $x_{p^*}$, у том случају $x_{p^*} < X^{(1)} < X^{(2)}$. Дакле,

$$\begin{aligned}
&(26)\ P(x_{p^*} < X^{(2)}) = P(x_{p^*} < X^{(1)}) + P(X^{(1)} \leq x_{p^*} < X^{(2)})\\
&= P \text{ (барем } n-1 \text{ од } X_i \text{ превазилази } x_{p^*})\\
&= P \text{ (1 или мање од } Xi \text{ су } \leq x_{p^*})
\end{aligned}$$

Сада је вероватноћа у једначини (26) дата као биномна функција расподеле, јер свака $X_i$ има вероватноћу $p^*$ која је мања или једнака од $x_{p^*}$, а $X_i$ су независне. Стога, претходна једначина доводи до:

$$(27)\ P(x_{p^*} < X^{(2)}) = \textstyle\sum_{i=0}^{1} \binom{n}{i} (p^*)^i (1-p^*)^{n-i}$$

Претходни доказ може бити продужен:

$$\begin{aligned}
&(28)\ P(x_{p^*} < X^{(r)}) = P \text{ (барем } n-r+1 \text{ од } X_i \text{ превазилази } x_{p^*})\\
&= P\ (r-1 \text{ или мање од } X_i \text{ су } \leq x_{p^*})\\
&= \textstyle\sum_{i=0}^{r-1} \binom{n}{i} (p^*)^i (1-p^*)^{n-i}
\end{aligned}$$

Коефицијент поверења даје,

$$(29)\ 1 - \alpha \cong P(X^{(r)} \leq x_{p^*} \leq X^{(s)}) = P(x_{p^*} \leq X^{(s)}) - P(x_{p^*} < X^{(r)})$$

Зато $r$ и $s$ могу бити изабрани, уз помоћ једначине 28 и таблице (таблица за биномну расподелу), тако да:

(30) $1 - \alpha_2 = P(x_{p^*} \leq X^{(s)}) = 1 - \frac{\alpha}{2}$

и,

(31) $\alpha_1 = P(X_{p^*} < X^{(r)}) = \frac{\alpha}{2}$

Онда ће коефицијент поверења бити $1 - \alpha_2 - \alpha_1 \cong 1 - \alpha$. Претпостављa се да је функција расподеле континуирана, онда имамо

(32) $P(x_{p^*} \leq X^{(s)}) = P(x_{p^*} < X^{(s)})$

Прво ћемо размотрити како једначина $P(X > X_{p^*}) \leq 1 - p^*$ утиче на једначину (28), и методу за проналажење $r$ у једначини (31). Вероватноћа $P(x_{p^*} < X^{(r)})$ може бити мања него што је била у непрекидном случају, која је затим дата једначином (28). Дакле, генерално, уместо једначине (28) важи следећа:

(33) $P(x_{p^*} < X^{(r)}) \leq \sum_{i=0}^{r-1} \binom{n}{i} (p^*)^i (1 - p^*)^{n-i}$

Ако се таблица користи да пронађе $r$, на начин који је описан, онда:

(34) $P(x_{p^*} < X^{(r)}) \leq \alpha_1$.

Сада ћемо размотрити како једначина $P(X \geq X_{p^*}) \geq 1 - p^*$ утиче на вероватноћу $1 - \alpha_2$, која произилази из нашег начина избора $s$. Као резултат, једначина (28) је модификована у општем случају:

(35) $P(x_{p^*} \leq X^{(s)}) \geq \sum_{i=0}^{s-1} \binom{n}{i} (p^*)^i (1 - p^*)^{n-i}$

Ако се таблица користи да нађе $s$ на описан начин, имамо,

(36) $P(x_{p^*} \leq X^{(s)}) \geq 1 - \alpha_2$

Једначине (34) и (36), које су тачне за све расподеле, могу се користити као следеће:

(37) $P(X^{(r)} \leq x_{p^*} \leq X^{(s)}) = P(x_{p^*} \leq X^{(s)}) - P(x_{p^*} < X^{(r)})$
$\geq P(x_{p^*} \leq X^{(s)}) - \alpha_1 \geq 1 - \alpha_1 - \alpha_2$

Добијање великог узорака за $r$ и $s$ заснива се на коришћењу стандардне нормалне расподеле за приближну биномну расподелу. Различите чињенице могу да се пласирају на различите начине претварањем $r^*$ и $s^*$ у целе бројеве $r$ и $s$, али метода која је овде дата је једноставно заокруживање горе на следећи виши цео број, како бисмо обезбедио довољно приближавање.

### 2.2.1 Примери

**Пример (1)**.

Случајни узорак обима 20 тежина (у неким јединицама мерења) дечака који похађају 10.разред.

| | | | | |
|---|---|---|---|---|
| 142 | 134 | 98 | 119 | 131 |
| 103 | 154 | 122 | 93 | 137 |
| 86 | 119 | 161 | 144 | 158 |
| 165 | 81 | 117 | 128 | 103 |

Тестирати хипотезу да је средња тежина 103.

**Решење (1)**.

Нека је нулта хипотеза $H_0(X_{0.5} = 103)$, и алтернативна $H_1(X_{0.5} \neq 103)$. Ради лакшег проналажења статистика $T_1$ и $T_2$, написаћемо бројеве у растућем поретку:

81, 86, 93, 98, 103, 103, 117, 119, 119, 122, 128, 131, 134, 137, 142, 144, 154, 158, 161 ,165.
$T_1$= број исхода $\leq 103$
$T_2$= број исхода $< 103$

Па је: $T_1 = 6$, $T_2 = 4$. Пошто имамо двострани тест, p-вредност теста рачунамо на следећи начин:

$$p = 2 \cdot \min\{P(Y \leq T_1), P(Y \geq T_2)\}$$
$$= 2 \cdot \min\{P(Y \leq 6), P(Y \geq 4)\} = 2 \cdot \min\{P(Y \leq 6), 1 - P(Y < 4)\}$$
$$= 2 \cdot \min\{0.0588, 0.9982\} = 0.1176$$

Вредности 0.0588 и 0.9982 смо добили из таблице за биномну расподелу за $n = 20$ јер:

$$Y \sim \mathrm{B}\,(n = 20, p = 0.5)$$

Kako је $p = 0.1176 > 0.05$ прихватамо нулту хипотезу. Дакле, средња тежина дечака јесте 103.
Сада ћемо тражити решење применом апроксимације нормалном расподелом.

$$Z_1 = \frac{T_1 - n \cdot p + 0.5}{\sqrt{n \cdot p \cdot (1-p)}} = \frac{6 - 20 \cdot 0.5 + 0.5}{\sqrt{20 \cdot 0.5 \cdot (1 - 0.5)}} = -1.565$$
$$Z_2 = \frac{T_2 - n \cdot p + 0.5}{\sqrt{n \cdot p \cdot (1-p)}} = \frac{4 - 20 \cdot 0.5 + 0.5}{\sqrt{20 \cdot 0.5 \cdot (1 - 0.5)}} = -2.9089$$
$$p = 2 \cdot \min\{P(Z \leq Z_1), P(Z \geq Z_2)\}$$
$$= 2 \cdot \min\{P(Z \leq -1.565), P(Z \geq -2.9089)\}$$
$$= 2 \cdot \min\{0.0588, 0.9982\} = 0.1176$$

где $Z$ има стандардну нормалну расподелу. Тако да прихватамо нулту хипотезу јер је 0.1176 веће од 0.05 (праг значајности).

Решили смо задатак на два начина и добили да је у оба случаја p-вредност 0.1176. Сада ћемо пример решити у R-у:

Како не постоји стандардна функција у R-у за квантил тест, морамо направити.

$> quantile.test <- function(x, xstar = 103, quantile = 0.5, alternative = "two.sided")\{$
$n <- length(x)$
$p <- quantile$
$T1 <- sum(x \leq xstar)$
$T2 <- sum(x < xstar)$
$if(alternative == "quantile.less")\{$
$p.value <- 1 - pbinom(T2 - 1, n, p)$
$\}$
$if(alternative == "quantile.greater")\{$
$p.value <- pbinom(T1, n, p)$
$\}$
$if(alternative == "two.sided")\{$
$p.value <- 2 \cdot \min(1 - pbinom(T2 - 1, n, p), pbinom(T1, n, p))$
$\}$
$list(xstar = xstar, alternative = alternative, T1 = T1,$
$T2 = T2, p.value = p.value)$
$\}$
$> quantile.test(x, xstar = 103, quantile = 0.5, alternative = "two.sided")$
$\$xstar$
$[1]103$
$\$alternative$
$[1]"two.sided"$
$\$T1$
$[1]6$
$\$T2$
$[1]4$
$\$p.value$
$[1]0.1153183$

Из добијених резултата, можемо видети да је $T_1 = 6$, $T_2 = 4$, а p-вредност теста је 0.115 и већа је од 0.05 (праг значајности). Закључујемо, можемо да прихватимо нулту хипотезу.

**Пример (2)**.

У примеру (1) тестирати хипотезу да је горњи квартил бар 150.

**Решење (2)**.

Нека је сад нулта хипотеза $H_0(X_{0.75} \geq 150)$, а алтернативна $H_1(X_{0.75} < 150)$. Тражимо тест статистику за $T_2$.

$T_2 =$број исхода $<$150, а то је $T_2 = 16$.
$p = P(Y \geq T_2) = 1 - P(Y \leq 15) = 0.4148$

$Y$ има биномну расподелу са параметрима $n = 20$ і $p = 0.75$. Тако да прихватамо нулту хипотезу јер је 0.4148 веће од 0.05 (праг значајности). Тражимо решење у R-у:

$> \quad quantile.test(x, xstar \quad = \quad 150, quantile \quad = \quad 0.75, alternative \quad =$
$"quantile.less")$
$\$xstar$
$[1]150$
$\$alternative$
$[1]"quantile.less"$
$\$T1$
$[1]16$
$\$T2$
$[1]16$
$\$p.value$
$[1]0.4148415$

Како смо и на овај начин добили да је p-вредност 0.4148, доносимо закључак да прихватамо нулту хипотезу. Дакле, да је горњи квартил бар 150.

**Пример (3)**.

У примеру (1) тестирати хипотезу да трећи децил није већи од 100.

**Решење (3)**.

Нека је сад нулта хипотеза $H_0(X_{0.30} \leq 100)$, а алтернативна $H_1(X_{0.30} > 100)$. Тражимо тест статистику $T_1$.

$T_1$ који су $\leq 100$. $T_1 = 4$.
$p = P(Y \leq T_1) = P(Y \leq 4) = 0.2375$

$Y$ има биномну расподелу са параметрима $n = 20$ і $p = 0.30$. Тако да прихватамо нулту хипотезу јер је p-вредност 0.2375 веће од 0.05 (праг значајности). Тражимо решење у R-у:

$> \quad quantile.test(x, xstar \quad = \quad 100, quantile \quad = \quad 0.3, alternative \quad =$
$"quantile.greater")$
$\$xstar$
$[1]100$
$\$alternative$
$[1]"quantile.greater"$

$\$T1$
$[1]4$
$\$T2$
$[1]4$
$\$p.value$
$[1]0.2375078$

Како смо добили да је p-вредност 0.2375 већа од 0.05 (праг значајности) и прихватамо нулту хипотезу, закључујемо да трећи децил није већи од 100.

**Пример (4).**

У примеру (1) пронађи приближно 90% интервал поверења за средњу вредност.

**Решење (4).**

Направићемо прво функцију у R-у која рачуна квантил теста, пошто то није стандардна функција.

$> quantile.interval < -function(x, quantile = 0.5, conf.level = 0.95)\{$
$n < -length(x)$
$p < -quantile$
$alpha < -1 - conf.level$
$rmin1 < -qbinom(alpha/2, n, p) - 1$
$r < -rmin1 + 1$
$alpha1 < -pbinom(r - 1, n, p)$
$smin1 < -qbinom(1 - alpha/2, n, p)$
$s < -smin1 + 1$
$alpha2 < -1 - pbinom(s - 1, n, p)$
$clo < -sort(x)[r]$
$chi < -sort(x)[s]$
$conf.level < -1 - alpha1 - alpha2$
$list(quantile = quantile, conf.level = conf.level, r = r, s = s, interval = c(clo, chi))\}$
$> quantile.interval(x, quantile = 0.5, conf.level = 0.90)$
$\$quantile$
$[1]0.5$
$\$conf.level$
$[1]0.9586105$
$\$r$
$[1]6$
$\$s$
$[1]15$
$\$interval$
$[1]103\ 142$

Можемо приметити да је 90% интервал поверења за средњу вредност $[103, 142]$.

**Пример (5).**

Пожељно је дизајнирати аутомобил који би омогућио довољан простор за главу, да се удобно прилагоди свима, али највише 5% високим људима који возе. Раније студије указују да је 95.перцентил 70.3 инча. Да би видели да ли раније студије још увек важе, изабран је случајни узорак величине 100. Утврђено је да 12 највиших особа у узорку имају следеће висине.

| | | | |
|---|---|---|---|
| 72.6 | 70.0 | 71.3 | 70.5 |
| 70.8 | 76.0 | 70.1 | 72.5 |
| 71.1 | 70.6 | 71.9 | 72.8 |

Да ли је разумно да се користи 70.3 као 95. перцентила?

**Решење (5).**

Тестирајмо нулту хипотезу $H_0$ (да је 95ти перцентил 70.3), против алтернативне хипотезе да 95ти перцентил није 70.3 - двострани тест. Тест ћемо радити у R-у.

$x < -c(72.6, 70.0, 71.3, 70.5, 70.8, 76.0, 70.1, 72.5, 71.1, 70.6, 71.9, 72.8)$
$n < -100$
$percentil < -0.95$
$t1 < -(100 - 12) + sum(x \leq 70.3)$
$t2 < -(100 - 12) + sum(x < 70.3)$
$c(n, t1, t2)$
$[1]1009090$
$> quantile.test.counts < -function(t1, t2, n, quantile = 0.5, alternative = "two.sided")\{$
$p < -quantile$
$if(alternative == "less")\{$
$p.value < -1 - pbinom(t2 - 1, n, p)\}$
$elseif(alternative == "greater")\{$
$p.value < -pbinom(t1, n, p)\}$
$elseif(alternative == "two.sided")\{$
$p.value < -2 * min(1 - pbinom(t2 - 1, n, p), pbinom(t1, n, p))\}$
$list(alternative = alternative, t1 = t1, t2 = t2, p.value = p.value)\}$
$> quantile.test.counts(t1, t2, n, pstar)$
$\$alternative$
$[1]"two.sided"$
$\$t1$
$[1]90$
$\$t2$
$[1]90$
$\$p.value$
$[1]0.05637659$

Како је p-вредност 0.0564 већа од 0.05 (праг значајности), можемо прихватити нулту хипотезу. Дакле 70.3 се може користити као 95ти перцентил.

**Пример (6)**.

У задатку (5), шта је 95% интервал поверења 95. перцентила возача од којих је изабран узорак?

**Решење (6)**.

x <- c(72.6, 70.0, 71.3, 70.5, 70.8, 76.0, 70.1, 72.5, 71.1, 70.6, 71.9, 72.8)

$> quantile.interval(x, quantile = 0.95, conf.level = 0.95)$
$\$quantile$
[1]0.95
$\$conf.level$
[1]0.9804317
$\$r$
[1]10
$\$s$
[1]13
$\$interval$
[1]72.6 $NA$

Кроз ове примере покушали смо да боље сагледамо у којим све областима и на који начин можемо користити наведене тестове.

У наредном поглављу за додатне информације, користили смо материјал за предавања, Медицинског факултета Универзитета у Нишу.

# Поглавље 3

# Моћ теста

У овом раду описујемо биномни и квантил тест. Када користимо ове тестове при тестирању, добро је знати каква је моћ поменутих тестова.

Желимо да видимо колико су добри тестови описани у претходном поглављу. Дакле, желимо да испитамо моћ теста. При тестирању нулте хипотезе можемо направити две грешке:

1) Грешка прве врсте ($\alpha$), која настаје када одбацимо истиниту нулту хипотезу.

2) Грешка друге врсте ($\beta$), која настаје када прихватимо нетачну нулту хипотезу.

Вероватноћа остварења грешке прве врсте су прагови значајности. Значи праг или прагови значајности је вероватноћа ризика да нулта хипотеза буде одбачена иако је истинита. Што је p-вредност мања, то је мања вероватноћа да је нулта хипотеза тачна. Вероватноћа грешке друге врсте повезана је са моћи теста. Моћ теста $\gamma$ је вероватноћа да се одбаци нетачна $H_0$, и једнака је разлици између 1 и вероватноће грешке друге врсте. Однос вероватноћа је $\gamma = 1 - \beta$. Моћ теста је већа, када је вероватноћа грешке друге врсте мања. Што је моћ теста већа тест је бољи.

**Пример.** Сматрамо да ако барем 60% студената положи испит, онда је испит успешно изведен. У супротном требало би радити на бољој припреми студената за испит, или испит боље прилагодити. Узет је узорак од 500 студената да би се утврдило да ли је испит добро спроведен. Добијено је да је 270 студената положило испит.

(1) Да ли је потребно на основу добијених резултата порадити на припреми испита?

(2) Написати функцију у R која генерише истраживање на $n$ студената и приказати резултате тестирања а да студенти полажу испит са вероватноћом 0.6. Одредити 95% интервал поверења ако се зна да је 294 студената положило испит.

(3) Показати моћ теста.

**Решење.**

Поставићемо нулту хипотезу да је испит успешно изведен ако је бар 60% студената положило испит, против алтернативне хипотезе да добијени резултати показују да се мора радити на бољој припреми испита.
Нулта хипотеза:

$$H_0 : p \geq 0.6.$$

Алтернатива:

$$H_1 : p < 0.6.$$

Односно, нека је тест статистика T укупан број положених испита. T има биномну расподелу са параметром $p$ и $n$=500.

За $n$= 500, критична област приближне величине $\alpha = 0.05$ могу се добити помоћу великог узорка приближно датим у таблици. Стога критичан регион одговара свим вредностима T мањим или једнаким од $t_1$, где је:

$$t_1 = n \cdot p^* + \omega_{.05} \cdot \sqrt{n \cdot p^* \cdot (1 - p^*)} = 500 \cdot 0.6 + (-1.64)\sqrt{500 \cdot 0.6 \cdot 0.4} = 281.92$$

Вредност T добијена у овом експерименту је 270. Како ова вредност упада у критичну област, тачније 270 < 281.92, кажемо да одбацујемо нулту хипотезу.

Како је вредност 270 је мања од $n \cdot p^* = 500 \cdot 0.6 = 300$, $\hat{\alpha}$ се израчунава из $P(T \leq 270)$, претпостављајући да је нулта хипотеза тачна,

$$P(T \leq 270) = P\left(\frac{T - n \cdot p^*}{\sqrt{n \cdot p^* \cdot (1 - p^*)}} \leq \frac{270 - n \cdot p^*}{\sqrt{n \cdot p^* \cdot (1 - p^*)}}\right)$$
$$\approx P(Z \leq \tfrac{270 - 300}{10.95}) = P(Z \leq -2.7397)$$

где $Z$ има стандардну нормалну расподелу, као што се користи у таблицама или користећи стандардну функцију у R-у **pnorm**.

$$\hat{\alpha} = P(T \leq 270) \approx P(Z \leq -2.7397) \approx 0.0031$$

Ниво значајности је мањи од 0.05, тако да одбацујемо нулту хипотезу. Резултат добијен у R-у је:

$> pnorm(-2.7397)$
[1]0.003074764

Представимо решење примера коришћењем биномног теста у R-у:

**binom.test(270, 500, p = 0.6, alternative = "less")**

Exact binomial test

data: 270 and 500
number of successes = 270, number of trials = 500, p-value = 0.003714
alternative hypothesis: true probability of success is less than 0.6

95 percent confidence interval:
0.0000000 0.5774139
sample estimates:
probability of success
0.54

На основу добијене p-вредности теста, p-value = 0.003714, која је мања од 0.05 (праг значајности), дакле можемо рећи да одбацујемо нулту хипотезу.

Дакле, можемо закључити да се мора радити на бољој припреми испита.

(2) Написаћемо функцију у R-у која генерише случајан низ:

$Slucajan\_Uzorak\_binom$ <-rbinom(500, size=1, prob=0.6)
$Broj\_Uspeha = sum(Slucajan\_Uzorak\_binom)$
**binom.test (*$Broj\_Uspeha$*, 500, p=0.6, alternative="less")**

Exact binomial test

data: $Broj\_Uspeha$ and 500
number of successes = 294, number of trials = 500, p-value = 0.307
alternative hypothesis: true probability of success is less than 0.6
95 percent confidence interval
0.0000000 0.6247475
sample estimates:
probability of success
0.588

На основу добијене p-вредности теста, p-вредност = 0.307, која је већа од 0.05 (праг значајности), дакле можемо закључити да прихватамо нулту хипотезу.

Овај узорак је у складу са нултом хипотезом. То значи да су студенти положили са вероватноћом 0.6.

**Сада ћемо одредити 95% интервал поверења, ако знамо да је 294 студената положило испит.**

Како је $n$ веће од 20 тражимо решење применом апроксимације нормалном расподелом. Употреба:

$$L = \frac{Y}{n} - x_{1-\frac{\alpha}{2}} \cdot \sqrt{Y \cdot \frac{(n-Y)}{n^3}}$$
$$U = \frac{Y}{n} + x_{1-\frac{\alpha}{2}} \cdot \sqrt{Y \cdot \frac{(n-Y)}{n^3}}$$

где је $x_{1-\frac{\alpha}{2}}$ квантил за нормалну расподелу случајне промењиве, добијен из таблице.

$$L = \frac{Y}{n} - x_{.05} \cdot \sqrt{Y \cdot \frac{(n-Y)}{n^3}} = \frac{294}{500} + 1.64 \cdot \sqrt{294 \cdot \frac{(500-294)}{500^3}} = 0.587 + 0.036 = 0.0623$$

Интервал поверења је:

$[0, 0.0623]$

Интервал поверења можемо добити у R-у:

**binom.test(294 , 500, p=0.6, alternative="less")**
Exact binomial test

data: 294 and 500
number of successes = 294, number of trials = 500, p-value = 0.307
alternative hypothesis: true probability of success is less than 0.6
95 percent confidence interval:
0.0000000 0.6247475
sample estimates:
probability of success
0.588

95% **интервал поверења** $[0.0000000, 0.6247475]$**.**

Ставку (3) у овом примеру можемо показати на два начина:

- Симулацијом у R-у
- Теоријским приступом

## 3.1 Биномни тест-симулација у R-у

(3) Дакле, када генеришемо један узорак ми очекујемо да ће биномни тест, да га прихватили или да га одбаци.

Када генеришемо 10000 узорака, ми тада очекујемо да ће отприлике 500 узорака да одбаци, јер мора имати грешку прве врсте, за $\alpha$=0.05, а осталих 9500 да не одбаци. Онда кажемо да је то добар тест.

Снага истраживања расте са величином узорка. Генеришимо сада 10000 узорака са обимом узорка $n = 500$ , вероватноћом $p = 0.6$.

$pval < -vector(length = 10000)$
$for(n\ in\ 1 : 10000)\{$
$Slucajan\_Uzorak\_binom < -rbinom(500, size = 1, prob = 0.6)$
$Broj\_Uspeha < -sum(Slucajan\_Uzorak\_binom)$
$pval[n] < -binom.test(Broj\_Uspeha, 500, p = 0.6, alternative = "less")\$p.value$
$\}$
$sum(pval >= 0.05)$
$[1]9545$

Можемо закључити да од 10000 узорака, у 9545 узорака испит је успешно изведен, као и то да је у тим узорцима бар 60% студената положило испит.

Као што смо и очекивали, наш тест одбацује 455 узорака, јер је направљена грешка прве врсте, $\alpha = 0.05$, а осталих 9545 узорака прихвата. Сада можемо рећи да је биномни тест добар.

Сада за вредност $p$ узмемо вредности из алтернативног скупа. Колико узорака ће од 10000 узорака тест одбацити? За неке узорке ће погрешити и прихватити, али то је грешка друге врсте ($\beta$), али оно што одбаци или колико узорака одбаци, јесте моћ теста.

```
for(n in 1 : 10000){
Slucajan_Uzorak_binom < −rbinom(500, size = 1, prob = 0.55)
Broj_Uspeha < −sum(Slucajan_Uzorak_binom)
pval[n] < −binom.test(Broj_Uspeha, 500, p = 0.6, alternative =
"less")$p.value
}
sum(pval <= 0.05)
[1]7289
```

Овде можемо приметити да тест 7289 узорака одбацује, што је приближно да је моћ теста 73%. За остале узорке је направљена грешка друге врсте $\beta$.

У наставку текста видећемо да смо и на основу теоријског приступа добили сличан резултат.
Можемо закључити на основу одбачених узорака, да је моћ теста 73%.

Сада за вредност $p$ узмемо нову вредности из алтернативног скупа. Колико узорака ће од 10000 узорака тест одбацити?

```
for(n in 1 : 10000){
Slucajan_Uzorak_binom < −rbinom(500, size = 1, prob = 0.5)
Broj_Uspeha < −sum(Slucajan_Uzorak_binom)
pval[n] < −binom.test(Broj_Uspeha, 500, p = 0.6, alternative =
"less")$p.value
}
sum(pval <= 0.05)
[1]9977
```

Овде можемо приметити да тест 9977 узорака одбацује, што значи да је моћ теста 99.73%. За остале узорке је направљена грешка друге врсте $\beta$.

У наставку текста видећемо да смо и на основу теоријског приступа добили сличан резултат.

Можемо закључити да је овде моћ теста велика, $\gamma = 99.77\%$. Како је моћ теста велика, тако је грешка друге врсте мања.

Доносимо закључак да што се вредност више одаљава од вредности $p = 0.6$ то биномни тест јаче одбацује нетачну хипотезу. Показали смо моћ биномног теста симулацијом у R-у. Сада ћемо видети како то изгледа кроз теоријски приступ.

## 3.2 Биномни тест-теоријски начин

Користимо сада теоријски приступ да прикажемо моћ биномног теста.

Из примера имамо да је:

$H_0 : p \geq 0.6.$

Алтернатива:

$H_1 : p < 0.6.$

Тражимо сада вероватноћу $P(T \leq t_1)$ при чему T статистика има биномну расподелу са параметрима $n = 500$ и $p = 0.6$. Треба да израчунамо $t_1$. То постижемо на следећи начин:

$$t_1 = n \cdot p^* + \omega_{.05} \cdot \sqrt{n \cdot p^* \cdot (1 - p^*)} = 500 \cdot 0.6 + (-1.64)\sqrt{500 \cdot 0.6 \cdot 0.4} = 281.92$$

Можемо заокружити вредност $t_1$ на 282.

Ми желимо да покажемо колика је грешка прве врсте, тачније да се одбаци нулта хипотеза која је тачна.

Како смо израчунали вредност за критичну област, сада рачунамо вероватноћу. Грешка прве врсте једнака је $P(T \leq 282) = \alpha$. Она треба да постоји, и она је једнака прагу значајности.

$$P(T \leq 282) = P\left(\frac{T - n \cdot p^*}{\sqrt{n \cdot p^* \cdot (1 - p^*)}} \leq \frac{282 - n \cdot p^*}{\sqrt{n \cdot p^* \cdot (1 - p^*)}}\right)$$
$$\approx P\left(Z \leq \frac{282 - 300}{10.95}\right) = P(Z \leq -1.6438)$$

где $Z$ има стандардну нормалну расподелу, као што се користи у таблицама за нормалну расподелу или користећи стандардну функцију у R-у **pnorm**.

$$\hat{\alpha} = P(T \leq 282) \approx P(Z \leq -1.6438) \approx 0.0503.$$

Што значи да је вероватноћа грешке прве врсте једнака 0.05. Закључујемо, када бисмо узели 10000 узорака, биномни тест би 9500 узорака прихватио, а одбацио 500 узорака.

Сада желимо да израчунамо моћ биномног теста. Тачније вероватноћу $P_{H_1}(T \in 282) = M$

Сада рачунамо колика је моћ теста за ново $p$ из алтернативног скупа. На пример узећемо за $p = 0.55$.

$$P(T \leq 282) = P(\frac{T-n\cdot p^*}{\sqrt{n\cdot p^*\cdot(1-p^*)}} \leq \frac{282-n\cdot p^*}{\sqrt{n\cdot p^*\cdot(1-p^*)}})$$
$$\approx P(Z \leq \frac{282-275}{11.1243}) = P(Z \leq 0.6292)$$

где $Z$ има стандардну нормалну расподелу, као што се користи у таблицама за нормалну расподелу или користећи стандардну функцију у R-у **pnorm**.

$$\hat{\alpha} = P(T \leq 282) \approx P(Z \leq 0.6292) \approx 0.7353.$$

Што значи да је моћ теста 73.53%

Сада рачунамо моћ за вредност $p = 0.5$.

$$P(T \leq 282) = P(\frac{T-n\cdot p^*}{\sqrt{n\cdot p^*\cdot(1-p^*)}} \leq \frac{282-n\cdot p^*}{\sqrt{n\cdot p^*\cdot(1-p^*)}})$$
$$\approx P(Z \leq \frac{282-250}{11.18}) = P(Z \leq 2.8621)$$

где $Z$ има стандардну нормалну расподелу, као што се користи у таблицама за нормалну расподелу или користећи стандардну функцију у R-у **pnorm**.

$$\hat{\alpha} = P(T \leq 282) \approx P(Z \leq 2.8621) \approx 0.9978.$$

Што значи да је моћ теста 99.78%

Закључујемо, да што се више одаљавамо, моћ теста јаче одбицује нетачну нулту хипотезу.

Можемо да приметимо да смо симулацијом и теоријским приступом добили исте резултате за моћ теста, што је и требало показати.

Ово је био још један пример како и на који начин користимо биномни тест као и моћ теста.

## 3.3 Квантил тест-симулација у R-у

Покажимо сада моћ квантил теста кроз следећи пример.

**Пример.**

Генеришемо један случајан узорак који има нормалну расподелу $Y : \mathrm{N}(120, 10)$. **Тестирати хипотезу да је медијана једнака 118**.

Добили смо следећи узорак за $n = 20$:

$x < -rnorm(20, 120, 10)$
$> x$
[1] 123.4984 116.8789 107.5167 116.6251 135.6093 128.4512
115.7726 124.7335 125.4494 138.6808 104.9128 121.8711 126.9321
118.1623 118.0682 125.2085 133.4363 118.6484 116.0040
118.3202.

Ради лакше употребе, сортираћемо низ у растућем поретку:

$y < -zapsmall(x, digits = 0)$
$> y$ [1] 123 117 108 117 136 128 116 125 125 139 105 122 127 118 118 125
133 119 116 118
$ysort < -sort(y)$
$> ysort$
[1] 105 108 116 116 117 117 118 118 118 119 122 123 125 125 125 127 128
133 136 139.

Нека је нулта хипотеза $H_0(X_{0.5} = 118)$, и алтернативна $H_1(X_{0.5} \neq 118)$. Ради проналажења статистика $T_1$ и $T_2$, написаћемо бројеве у растућем поретку:

105 108 116 116 117 117 118 118 118 119 122 123 125 125 125 127 128 133
136 139
$T_1 =$ број исхода $\leq 118$
$T_2 =$ број исхода $< 118$
Па је: $T_1 = 9$, $T_2 = 6$.

Пошто имамо двострани тест, p-вредност теста рачунамо на следећи начин:

$$p = 2 \cdot \min\{P(Y \leq T_1), P(Y \geq T_2)$$
$$= 2 \cdot \min\{P(Y \leq 9), P(Y \geq 6)\}$$
$$= 2 \cdot \min P(Y \leq 9), 1 - P(Y < 6)$$
$$= 2 \cdot \min\{0.4119, 1 - 0.0207\}$$
$$= 2 \cdot \min\{0.4119, 0.9793\} = 0.8238$$

Вредности 0.4119 и 0.0207 смо добили из таблице за биномну расподелу за $n = 20$ јер $Y \sim \mathrm{B}\,(n = 20, p = 0.5)$.

Како је p=0.8238 > 0.05, можемо донети закључак да прихватамо нулту хипотезу. Подаци су у складу са нултом хипотезом. Као закључак доносимо да медијана јесте једнака 118.

Сада ћемо тражити решење применом апроксимације нормалном расподелом.

$$Z_1 = \frac{T_1 - n \cdot p + 0.5}{\sqrt{n \cdot p \cdot (1-p)}} = \frac{9 - 20 \cdot 0.5 + 0.5}{\sqrt{20 \cdot 0.5 \cdot (1-0.5)}} = -0.2232$$
$$Z_2 = \frac{T_2 - n \cdot p + 0.5}{\sqrt{n \cdot p \cdot (1-p)}} = \frac{6 - 20 \cdot 0.5 + 0.5}{\sqrt{20 \cdot 0.5 \cdot (1-0.5)}} = -1.565$$
$$p = 2 \cdot \min\{P(Z \leq Z_1), P(Z \geq Z_2)\}$$
$$= 2 \cdot \min\{P(Z \leq -0.2232), P(Z \geq -1.565)\}$$
$$= 2 \cdot \min\{0.4129, 0.94\} = 0.824$$

где $Z$ има стандардну нормалну расподелу. Тако да прихватамо нулту хипотезу јер је 0.824 веће од 0.05. Дакле, овај узорак одговара нултој хипотези.

Решење примера у R-у:

$> quantile.test < -function(x, xstar = 118, quantile = .5, alternative = "two.sided")\{$
$n < -length(x)$
$p < -quantile$
$T1 < -sum(x \leq xstar)$
$T2 < -sum(x < xstar)$
$if(alternative == "quantile.less")\{$
$p.value < -1 - pbinom(T2 - 1, n, p)\}$
$if(alternative == "quantile.greater")\{$
$p.value < -pbinom(T1, n, p)\}$
$if(alternative == "two.sided")\{$
$p.value < -2 \cdot \min(1 - pbinom(T2 - 1, n, p), pbinom(T1, n, p))\}$
$list(xstar = xstar, alternative = alternative, T1 = T1, T2 = T2, p.value = p.value)\}$

$> quantile.test(x, xstar = 118, quantile = 0.5, alternative = "two.sided")$
$\$xstar$
$[1]118$
$\$alternative$
$[1]"two.sided"$
$\$T1$
$[1]9$
$\$T2$
$[1]6$
$\$p.value$
$[1]0.8238029$

Приметимо да је p-вредност 0.8238029 већа од 0.05, зато прихватамо нулту хипотезу.

Како смо на основу p-вредноси теста закључили да прихватамо нулту хипотезу, тако можемо рећи да узорак одговара нултој хипотези, да је медијана једнака 118.

Сада желимо да видимо како се понаша тест ако генеришемо на основу једног узорка који одговара нултој хипотези, 10000 узорака са прагом значајности 0.05 (грешка прве врсте ($\alpha$)) и медијаном 118.

Прикажимо користећи R:

$> quantile.test <- function(x, xstar = 118, quantile = .5, alternative =$ $"two.sided"$$)\{$
$n <- length(x)$
$p <- quantile$
$T1 <- sum(x \leq xstar)$
$T2 <- sum(x < xstar)$
$if(alternative == "quantile.less")\{$
$p.value <- 1 - pbinom(T2 - 1, n, p)\}$
$if(alternative == "quantile.greater")\{$
$p.value <- pbinom(T1, n, p)\}$
$if(alternative == "two.sided")\{$
$p.value <- 2 \cdot \min(1 - pbinom(T2 - 1, n, p), pbinom(T1, n, p))\}$
$list(xstar = xstar, alternative = alternative, T1 = T1, T2 = T2, p.value = p.value)\}$

Сада ћемо генерисати 10000 узорака на основу узорка који одговара нултој хипотези, тачније када је медијана 118.

$pval <- vector(length = 10000)$
$for(n\ in\ 1 : 10000)\{$
$x <- rnorm(20, 120, 10)$
$pval[n] <- quantile.test(x, xstar = 118, quantile = 0.5, alternative = "two.sided")\$p.value$
$\}$
$sum(pval >= 0.05)$
$[1]9065$

Можемо да кажемо да наш тест прихвата 9065 узорака од 10000, а одбацује 935 узорака. Када анализирамо резултат, видимо да 90.65% узорака одговара нултој хипотези. Тачније да 90.65% узорака има медијану 118.

Поновимо сада исти поступак само за вредност медијане 128. Како се понаша тест?

$pval <- vector(length = 10000)$
$for(n\ in\ 1 : 10000)\{$
$x <- rnorm(20, 120, 10)$
$pval[n] <- quantile.test(x, xstar = 128, quantile = 0.5, alternative = "two.sided")\$p.value$
$\}$
$sum(pval >= 0.05)$
$[1]2397$

Као резултат добијамо да је тест за 2397 узорака направио грешку друге врсте, јер је за ове узорке прихватио нетачну нулту хипотезу.

Моћ теста се огледа у томе, колико узорака је тест одбацио, тај број је 7603. Тачније моћ теста је 76.03%

Да бисмо могли да донесемо бољи закључак о моћи квантил теста, можемо применити исти поступак тако што померимо вредност медијане на 130.

$for(n\ in\ 1:10000)\{$
$x <- rnorm(20, 120, 10)$
$pval[n] <- quantile.test(x, xstar = 130, quantile = 0.5, alternative = "two.sided")\$p.value$
$\}$
$sum(pval >= 0.05)$
$[1]851$

Моћ теста је 91.49%, јер 9149 узорака тест је одбацио. За остале узорке, 8.51%, направио је грешку друге врсте.

Приметимо, како је проценат прихватања нетачне нулте хипотезе мањи, то је моћ теста боља.

Можемо закључити, што се више одаљавамо од вредност медијане 118, то је одбацивање нетачне нулте хипотезе јаче.

**Још нека израчунавања везана за квантил тест.**

Користећи претходни пример, тестирати следеће:

1) Тестирати нулту хипотезу да је горњи квартил барем 127.

Нека је нулта хипотеза $H_0(X_{0.75} \geq 127)$, а алтернативна $H_1(X_{0.75} < 127)$. Тражимо тест статистику $T_2$.

$T_2$ број исхода <127, а то је $T_2 = 15$.
$p = P(Y \geq T_2) = 1 - P(Y \leq 14)$
$= 1 - 0.383 = 0.617$

$Y$ има биномну расподелу са параметрима $n = 20$ и $p = 0.75$. Тако да прихватамо нулту хипотезу јер је 0.617 веће од 0.05.

Тражимо решење у R-у:

$> quantile.test(x, xstar = 127, quantile = 0.75, alternative = "quantile.less")$
$\$xstar$
$[1]127$
$\$alternative$
$[1]"quantile.less"$
$\$T1$
$[1]16$
$\$T2$
$[1]15$
$\$p.value$
$[1]0.6171727$

Како је p-вредност теста већа од 0.05, закључујемо да прихватамо нулту хипотезу. Узорак одговара нултој хипотези.
Дакле, вредност горњег квартила је бар 127.

2) Тестирати нулту хипотезу да је доњи квартил не већи од 117.

Нека је сада нулта хипотеза $H_0(X_{0.25} \leq 117)$, а алтернативна $H_1(X_{0.25} > 117)$. Тражимо тест статистику $T_1$.

$T_1$ који су $\leq 117$. $T_1 = 6$.
$p = P(Y \leq T_1) = P(Y \leq 6) = 0.786$

$Y$ има биномну расподелу са параметрима $n = 20$ и $p = 0.25$. Закључак је да прихватамо нулту хипотезу јер је 0.786 веће од 0.05.
Дакле, вредност доњег квартила није већа од 117.

Тражимо решење у R-у:

$> quantile.test(x, xstar = 117, quantile = 0.25, alternative = "quantile.greater")$
$\$xstar$
$[1]117$
$\$alternative$
$[1]"quantile.greater"$
$\$T1$
$[1]6$
$\$T2$
$[1]4$
$\$p.value$
$[1]0.7857819$

Применом квантил теста у R-у, добили смо исто решење, да је узорак у складу са нултом хипотезом и да доњи квартил није већи од 117.

3) Рачунамо 95% интервал поверења за медијану.

$> quantile.interval < -function(x, quantile = 0.5, conf.level = 0.95)\{$
$n < -length(x)$
$p < -quantile$
$alpha < -1 - conf.level$
$rmin1 < -qbinom(alpha/2, n, p) - 1$
$r < -rmin1 + 1$
$alpha1 < -pbinom(r - 1, n, p)$
$smin1 < -qbinom(1 - alpha/2, n, p)$
$s < -smin1 + 1$
$alpha2 < -1 - pbinom(s - 1, n, p)$
$clo < -sort(x)[r]$
$chi < -sort(x)[s]$
$conf.level < -1 - alpha1 - alpha2$

$list(quantile = quantile, conf.level = conf.level, r = r, s = s, interval = c(clo, chi))\}$

$> quantile.interval(x, quantile = 0.5, conf.level = 0.90)$
$\$quantile$
[1]0.5
$\$conf.level$
[1]0.9586105
$\$r$
[1]6
$\$s$
[1]15
$\$interval$
[1]117 125

Можемо закључити да 95% интервал поверења за медијану је $[117, 125]$.

# Поглавље 4

# Закључак

У овом раду, бавили смо се тестовима који су засновани на биномној расподели. Одабрали смо два теста, биномни и квантил тест.

Можемо закључити да биномни тест користимо када испитујемо популацију у којој постоје само две могућности. Једна се може назвати успехом, а друга неуспехом. На основу узорка потребно је испитати вероватноћу појављивања догађаја који је означен као успех. Тест статистика T има биномну расподелу.

Приметимо да када тестирамо хипотезе у вези квантила, тада биномни тест називамо квантил тест.

Кроз сам опис и примере које смо навели у овом раду, можемо закључити да се наведени тестови могу користити у различитим областима. Своју примену могу наћи у области медицине, економије, фармације и многим другим областима.

Поред самог описа тестова, у овом раду могли смо да видимо колико су тестови добри при тестирању различитих хипотеза. Колико су они добри, оценили смо користећи моћ теста. Један од значаја рада је да смо моћ теста представили на два начина, симулацијом у R-у и теоријски.

# Литература


[1] В. Јевремовић, Ј. Малишић, Статистичке медоте у метеорологији и инжењерству, Савезни хидрометеоролошки завод , Београд, 2002.

[2] Економски факултет Универзитета у Крагујевцу, материјал за наставу, Основи статистике, 2016.

[3] М. Радичевић, Математички факултет Универзитета у Београду материјал за наставу, статистички софтвер 4, 2015.

[4] W.J. Conover, Practical Nonparametric Statistics, Wiley, 1980.

[5] З. Милошевић, Медицински факултет Универзитета у Нишу, материјал за предавања, Медицинска статистика и информатика, 2017.

# Биографија аутора

Анита Златић је рођена 01.08.1990. године у Чачку. Основну школу "Милан Благојевић" завршила је у Лучанима. 2005. године уписује средњу школу, гимназију у Београду. По завршетку средње школе, уписује Математички факултет, Универзитета у Београду. Дипломирала је на смеру "Статистика, актуарска и финансијска математика" 2015. године. Своју професионалну афирмацију започела је 2015. године у фирми "Msg global solutions South East Europe d.o.o" где и данас ради.