R: Проверка на хипотези (2020-12-07)

При доверителните интервали използваме данните, за да оценим къде ще попадат търсени параметри. При хипотезите правим предположения за търсени параметри и после изчисляваме колко вероятно е това предположение да е вярно.

H_0-тестова хипотеза
H_A-алтрнативна хипотеза

Максималната вероятност, с която допускаме да отхвърлим H_0, когато H_0 е изпълнена се бележи с α и се нарича грешка от $I^{\text{-ви}}$ род:
$\mathbb{P}(\text{отхвърляме } H_0 \,|\, H_0) \leq \alpha$.
$1 - \alpha$ е ниво на доверие. $\alpha = [0{,}1]$ и обикновено $\alpha = \{0{,}01;\ 0{,}05;\ 0{,}1\}$.

W_α е критична област за H_0 - областта в която отхвърляме нулевата хипотеза, ако попаднем в нея.

Грешка от $II^{\text{-ри}}$ род е когато сме отхвърлили алтернативната хипотеза H_A при положение, че H_A е изпълнена: $\mathbb{P}(\text{отхвърляме } H_A \,|\, H_A) = \beta$.

α и β са обратно пропорционални. $\alpha \uparrow \to \beta \downarrow$ и обратно $\alpha \uparrow \leftarrow \beta \downarrow$.

Тестване на хипотези за параметрите на една популация

Тестване на хипотеза за вероятност за успех:

Пример 1. Хвърляме монета 20 пъти. При 13 от случайте се пада „тура“. Направете тест за хипотеза при ниво на значимост 5%, за да проверите дали при тази монета е по-вероятно да се падат „тури“.
Реш:
\# H0 = монетата е честна p=0.5
\# HA = монетата показва в полза на „тури“
Тук имаме **едностранен** тест:

```
succ=13; N=20;
res=binom.test(x=succ, n=N, p=1/2, alternative="greater", conf.level=0.95)
if(res$p.value>0.05) print("no reason to reject H0") else print("reject H0 in favour of HA")
# [1] "H0"
```

Ако имахме HA = монетата показва в полза на „ези“ или „тура“ - т.е. е нечестна $p \neq 0.5$, тогава щяхме да имаме двустранен тест и да проверим по следния начин хипотезата H0 (дали монетата е честна):
Тук имаме **двустранен** тест:

```
succ=13; N=20;
res=binom.test(x=succ, n=N, p=1/2, alternative="two.sided", conf.level=0.95)
if(res$p.value>0.05) print("no reason to reject H0") else print("reject H0 in favour of HA")
# [1] "H0"
```

Може да използваме и prop.test(13, 20, p=1/2)

Пример 2. Да предположим, че имаме зар и подозираме, че зара е пристрастен в полза на точките 6 от него. Хвърляме зара 25 пъти и преброяваме, че числото 6 се пада 7 пъти. Направете тест за хипотеза при ниво на значимост от 5%, за да проверите дали зара е предубеден и точкитв 6 на зара са по-благоприятни за получаване при хвърлянето му.

Реш:

```
# H0 = зара е честен относно шестиците p=1/6
# HA = зара е по-благоприятно да даде 6-ца
```

Тук имаме **едностранен** тест:

```
succ=7; N=25
res=binom.test(x=7, n=N, p=1/6, alternative="greater", conf.level=0.95)
if(res$p.value>0.05) print("no reason to reject H0") else print("reject H0 in favour of HA")
# [1] "H0"
```

Ако имахме HA = зара е нечестен относно шестиците - т.е. е дава повече или по-малко 6-ци от нормалното за един 6-стенен зар, тогава щяхме да имаме двустранен тест и да проверим по следния начин хипотезата H0:

Тук имаме **двустранен** тест:

```
succ=7; N=25;
res=binom.test(x=succ, n=N, p=1/2, alternative="two.sided", conf.level=0.95)
if(res$p.value>0.05) print("no reason to reject H0") else print("reject H0 in favour of HA")
# [1] "H0"
```

Може да използваме и prop.test(7, 25, p=1/6)

Пример 3. Питате 100 души в анкета а 42 от тях отговарят с „да“ на вашия въпрос. Това подкрепя ли хипотезата, че половината от анкетираните са отговорили с „да“?

Реш:

```
succ=42; N=100
res=binom.test(x=succ, n=N, p=1/2, alternative="two.sided", conf.level=0.95)
if(res$p.value>0.05) print("no reason to reject H0") else print("reject H0 in favour of HA")
# [1] "HO"
```

Аналогично може да използваме и prop.test() с абсолютно същите параметри

Нека увеличим/скалираме извадката и повторим теста:

```
succ=420; N=1000
res=binom.test(x=succ, n=N, p=1/2, alternative="two.sided", conf.level=0.95)
if(res$p.value>0.05) print("H0") else print("reject H0 in favour of HA")
# [1] "reject HO in favour of HA"
succ=420; N=1000
res=prop.test(x=succ, n=N, p=1/2, alternative="two.sided", conf.level=0.95)
if(res$p.value>0.05) print("H0") else print("reject H0 in favour of HA")
# [1] "reject HO in favour of HA"
```

Това е така, тъй като при по-малки извадки - точността на тестовете е по-малка и се допускат по-големи разлики между теста и емпиричната стойност. Но когато увеличаваме размера на пробата се увеличава и точността. Т.е. по-малките разлики между теста и емпиричната стойност стават статистически значими при по големи размери. Това може да се разглежда като следствие от ЗГЧ.

Тестване на хипотези за равенство между популационното **средно** и константа

$I^{-\text{ви}}$ **случай**: Ако $X \in \mathcal{N}(\mu, \sigma^2)$, където $\mathbb{E}X = \mu$ и σ^2 **Е известно**.

Пример 5. Нека приемем, че кола получава $X \in \mathcal{N}(\mu, 4)$ mpg. Производител твърди, че $\mu_0 = 25$ mpg. Потребителска група моли 10 собственика на този модел да изчислят своите mpg и средната стойност е 22 mpg. Поддържа ли се твърдението на производителя? Проверете хипотезата при ниво на значимост 0.05.
Реш:
От извадката олучаваме по-малко средно от 25, а именно 22.
Ще направим едностранен тест:

```
# LEFT SIDED
xbar=22; sigma=2; n=10
mu0=25   # xbar<m0
zemp=(xbar-mu0)/(sigma/sqrt(n))
if(pnorm(zemp, mean=0, sd=1)>0.05) print("no reason to reject H0")  else print("reject H0 in favour of HA")
# [1] "reject H0 in favour of HA"
```

Пример 6. Нека приемем, че стандартното тегло на гъба е най-много 40 грама и е нормално разпределено: $X \in \mathcal{N}(\mu, 9)$. Клиент твърди, че гъбите на производител не са стандартни. Производителя избира 15 гъби на случаен принцип и измерва теглата им. Средното тегло на гъбите от извадката е 41 грама. Може ли твърдението на клиента да бъде отхвърлено?
Реш:

```
# RIGHT SIDED
xbar=41; sigma=3; n=15
mu0=40   # xbar>mu0
zemp=(xbar-mu0)/(sigma/sqrt(n))
if(pnorm(zemp, mean=0, sd=1, lower.tail=F)>0.05) print("no reason to reject H0") else print("reject H0 in favour of HA")
# [1] "no reason to reject H0"
```

Пример 7. Нека приемем, че стандартното тегло на гъба е приблизително 40 грама и е нормално разпределено: $X \in \mathcal{N}(\mu, 9)$. Клиент твърди, че гъбите на производител не са стандартни. Производителя избира 15 гъби на случаен принцип и измерва теглата им. Средното тегло на гъбите от извадката е 41 грама. Може ли твърдението на клиента да бъде отхвърлено?
Реш:

```
# TWO SIDED
xbar=41; sigma=3; n=15
mu0=40   # xbar≠mu0
zemp=abs(xbar-mu0)/(sigma/sqrt(n))
if(2*pnorm(zemp, mean=0, sd=1, lower.tail=F)>0.05) print("no reason to reject H0") else print("reject H0 in favour of HA")
```

[1] "no reason to reject H0"

II-ри случай: Ако $X \in \mathcal{N}(\mu, \sigma^2)$, където $\mathbb{E}X = \mu$ и σ^2 **НЕ е известно**.

Пример 8. Нека приемем кола получава параметър, който е $X \in \mathcal{N}(\mu, \sigma^2)$. Производител твърди, че $\mu_0 = 25$ mpg. Потребителска група моли 10 собственика на този модел да изчислят своите mpg и средната стойност е 22 mpg. Стандартното отклонение на извадката е 3. Поддържа ли се твърдението на производителя? Проверете хипотезата при ниво на значимост 0.05.

Реш:

```
#Едностранен тест
#H0 е нулевата хипотеза: твърдението на производителя
#HA е алтернативната хипотеза
xbar=22; s=3; n=10
mu0=25
temp=(xbar-mu0)/(s/sqrt(n))

alpha=0.05
tcritical=qt(alpha, df=n-1)

if(pt(temp, df=n-1) > alpha) print("no reason to reject H0") else print("reject H0 in favour of HA")
# [1] "reject H0 in favour of HA"
```

Пример 9. Нека приемем, че стандартното тегло на гъба е най-много 40 грама и е нормално разпределено: $X \in \mathcal{N}(\mu, \sigma^2)$. Клиент твърди, че гъбите на производител не са стандартни. Производителя избира 15 гъби на случаен принцип и измерва теглата им. Средното тегло на гъбите от извадката е 41 грама. Стандартното отклонение на извадката е 3. Може ли твърдението на клиента да бъде отхвърлено?

Реш:

```
xbar=41; s=3; n=15
mu0=40 # xbar>mu0 -> RIGHT SIDED, ако е LEFT SIDED ще е с lower.tail=T
temp=(xbar-mu0)/(s/sqrt(n))
alpha=0.05
tcritical=qt(1-alpha, n-1)
tcritical
# [1] 1.76131
if(pt(temp, df=n-1, lower.tail=F) > 0.05) print("no reason to reject H0") else print("reject H0 in favour of HA")
# [1] "no reason to reject H0"
```

Пример 10. Нека приемем, че стандартното тегло на гъба е най-много 40 грама и е нормално разпределено: $X \in \mathcal{N}(\mu, \sigma^2)$. Клиент твърди, че гъбите на производител не са стандартни. Производителя избира 15 гъби на случаен принцип и измерва теглата им. Средното тегло на гъбите от извадката е 41 грама. Стандартното отклонение на извадката е 3. Може ли твърдението на клиента да бъде отхвърлено?

Реш:

```
# TWO SIDED
xbar=41; s=3; n=15
mu0=40
temp=abs(xbar-mu0)/(s/sqrt(n))
alpha=0.05
tcritical=qt(1-alpha/2, n-1)
tcritical
# [1] 2.144787
if(2*pt(temp, df=n-1, lower.tail=F) > 0.05) print("no reason to reject H0") else print("reject H0 in favour of HA")
# [1] "no reason to reject H0"
```

Пример 11. Вече сме използвали данните "puerto". Нека допуснеме, че седмичните приходи на пиертоамериканците в Маями са нормално разпределени и да видим дали може да кажем, че средното е равно на 277.

Реш:

```
# H0: mu=277
# HA: mu≠277
qqnorm(puerto)
qqline(puerto)

sh=shapiro.test(puerto)

library(tseries)
jb=jarque.bera.test(puerto)

library(nortest)
ad=ad.test(puerto)

if(sh$p.value>0.05 && ad$p.value>0.05 && jb$p.value>0.05) print("normal") else print("not normal")
# [1] "not normal"
```

Това, че данните ни не са нормално разпределени можеше да се види и чрез `qqplot.das(puerto, "norm")` и да забележим, че опашките излизат извън доверителния интервал на нормалното разпределение.

Следователно не може да използваме z-test, но:

```
length(puerto)
#[1] 50
res=t.test(puerto, mu = 277, conf.level = 0.95, alternative = "two.sided")
if(res$p.value>0.05) print("no reason to reject H0") else print("reject H0 in favour of HA")
# [1] "no reason to reject H0"
```

$III^{\text{-ти}}$ **случай**: Когато размера на извадката е голям и дисперсията на изследваната слузайна величина X е добре дефинирана (крайна), то предходните методи могат да бъдат използвани, без никаква информация за това дали X е нормално разпределено или не.

Пример 12. Същия като този от пример 11.

Тестване на хипотези за равенство между популационната **медиана** и константа

В случаите, когато нямаме информация дали наблюдаваната случайна величина X има крайна или безкрайна дисперсия, тестът на Wilcox може да бъде полезен.

Пример 12. Имаме изследване за използването на мобилен телефон от потребители за дължин на разговорите.

X=c(12.8, 3.5, 2.9, 9.4, 8.7, 0.7, 0.2, 2.8, 1.9, 2.8, 3.1, 15.8)

Какъв би бил подходящия тест за център?
Реш:
Проверяваме дали X е нормално разпределено.

```
qqnorm(X)
qqline(X)

sh=shapiro.test(X)
jb=jarque.bera.test(X)
ad=ad.test(X)

if(sh$p.value>0.05 && ad$p.value>0.05 && jb$p.value>0.05) print("normal") else print("not normal")
# [1] "not normal"
```

Това, че данните ни не са нормално разпределени можеше да се види и чрез qqplot.das(X, "norm") и да забележим, че опашките излизат извън доверителния интервал на нормалното разпределение.

```
wilcox.test(x, mu = 5, alternative = "greater")
```

Пример 13. Нека се върнем отново на пример 11.

```
wilcox.test(puerto, mu = 273, alternative = "two.sided")
# p-value = 0.9692
```

Ранкови тестове за медиана

```
X=c(12.8, 3.5, 2.9, 9.4, 8.7, 0.7, 0.2, 2.8, 1.9, 2.8, 3.1, 15.8)
simple.median.test(X, median = 5)
# [1] 0.3876953
```