

Soov märgata muutust Testid sagedustabelite jaoks.

Märt Möls
martm@ut.ee

Jaotus ja selle muutus

Loopealsel kasvavad kukeharjad

suur kukehari (<i>Sedum maximum</i>)	valge kukehari (<i>Sedum album</i>)	harilik kukehari (<i>Sedum acre</i>)	
10%	35%	55%	
20%	60%	20%	lisasoojus

Eksisteerib seos temperatuuri ja loopealsete
liigilise koostise vahel...

Seos – jaotuse muutumine – võib olla näiv

Loopealsel kasvavad kukeharjad

suur kukehari (<i>Sedum maximum</i>)	valge kukehari (<i>Sedum album</i>)	harilik kukehari (<i>Sedum acre</i>)	
4 (10%)	14 (35%)	22 (55%)	
12 (20%)	36 (60%)	12 (20%)	lisasoojus

Jaotuse hindamisel tehtav viga
võib tekitada näiva erinevuse...

Kuidas saaksime testida, kas seos võib olla
näiv (valimi juhuslikkusest / hindamisvigadest
tulenev muutus) või eksisteerib päriselt?

Üks võimalus: χ^2 -test

Seos – jaotuse muutumine – võib olla näiv

Loopealsel kasvavad kukeharjad

suur kukehari (<i>Sedum maximum</i>)	valge kukehari (<i>Sedum album</i>)	harilik kukehari (<i>Sedum acre</i>)	
4 (10%)	14 (35%)	22 (55%)	40
12 (20%)	36 (60%)	12 (20%)	60
16 (16%)	50 (50%)	34 (34%)	

$$\chi^2 = \sum_{i,j} \frac{(N_{ij} - E_{ij})^2}{E_{ij}} = \frac{(4 - 6,4)^2}{6,4} + \frac{(14 - 20)^2}{20} + \frac{(22 - 13,6)^2}{13,6} + \frac{(12 - 9,6)^2}{9,6} + \dots$$

= 18,3...

Seos – jaotuse muutumine – võib olla näiv

Loopealsel kasvavad kukeharjad

suur kukehari (<i>Sedum maximum</i>)	valge kukehari (<i>Sedum album</i>)	harilik kukehari (<i>Sedum acre</i>)
4 (10%)	14 (35%)	22 (55%)
12 (20%)	36 (60%)	12 (20%)

df= (ridade arv -1) (veergude arv - 1)

$$\chi^2 = \sum_{i,j} \frac{(N_{ij} - E_{ij})^2}{E_{ij}} = \frac{(4 - 6,4)^2}{6,4} + \frac{(14 - 20)^2}{20} + \frac{(22 - 13,6)^2}{13,6} + \frac{(12 - 9,4)^2}{9,4} = 18,3 \dots$$

df	0,95-kvantiil
1	3,84
2	5,99
3	7,81
4	9,49
5	11,07

Seos – jaotuse muutumine – võib olla näiv

Loopealsel kasvavad kukeharjad

suur kukehari (<i>Sedum maximum</i>)	valge kukehari (<i>Sedum album</i>)	harilik kukehari (<i>Sedum acre</i>)
4 (10%)	14 (35%)	22 (55%)
12 (20%)	36 (60%)	12 (20%)

18,3 > 5,99 seos eksisteerib, kukeharjade liigiline jaotus tõepoolest muutub

$$\chi^2 = \sum_{i,j} \frac{(N_{ij} - E_{ij})^2}{E_{ij}} = \frac{(4 - 6,4)^2}{6,4} + \frac{(14 - 20)^2}{20} + \frac{(22 - 13,6)^2}{13,6} + \frac{(12 - 9,4)^2}{9,4} = 18,3 \dots$$

df	0,95-kvantiil
1	3,84
2	5,99
3	7,81
4	9,49
5	11,07

Seose olemasolu testimine

Ligikaudsed testid:

- χ^2 -test
- Tõepärasuhte test (likelihood ratio test; G-test)

Täpsed testid

- Fisheri täpne test
- Barnardi test

Tõepärasuhte test (*Likelihood Ratio test, G-test*)

Suurima tõepära meetod (*Maximum Likelihood, ML*)

Suurima tõepära meetodi korral valitakse tundmatute parameetrite väärtused selliselt, et meie poolt nähtud valimi (andmete) saamine oleks kõige tõenäolisem.

Andmed (sõltumatud mündivisked):

KULL, KULL, KIRI

Mis on kulli saamise tõenäosus, P(KULL)=?

Arvame tõenäosuse P(KULL) olevat sellise, mille puhul oleks vaadeldud andmete nägemise tõenäosus kõige suurem...

Tõepärasuhte test (Likelihood Ratio test, G-test)

Suurima tõepära meetod (Maximum Likelihood, ML)

Andmed (sõltumatud mündivisked):

KULL, KULL, KIRI

Mis on kulli saamise tõenäosus, $P(KULL)=?$

Kui $P(KULL) = p$

siis $L = p \cdot p \cdot (1-p) = p^2 - p^3$

$L' = 2p - 3p^2$

maksimum: $2p - 3p^2 = 0$

$2p = 3p^2$

$2 = 3p \rightarrow p = 2/3$

Tõepärasuhte test (Likelihood Ratio test, G-test)

Andmed (sõltumatud mündivisked):

KULL, KULL, KIRI

$H_0: P(KULL)=0,5?$

$$G = 2 \ln \left(\frac{L}{L_0} \right) = 2 \ln \left(\frac{\frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3}}{0,5 \cdot 0,5 \cdot 0,5} \right) = 2 \ln (1,185 \dots) \approx 0,34$$

Alati suurem ühest

Suure valimi korral:

$$G \stackrel{H_0}{\sim} \chi^2_{\text{hinnatavate parameetrite arvu erinevus}}$$

$1-0 = 1$

Tõepärasuhte test (Likelihood Ratio test, G-test)

Andmed (sõltumatud mündivisked):

KULL, KULL, KIRI

$H_0: P(KULL)=0,5?$

$$G = 2 \ln \left(\frac{L}{L_0} \right) = 2 \ln \left(\frac{\frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3}}{0,5 \cdot 0,5 \cdot 0,5} \right) = 2 \ln (1,185 \dots) \approx 0,34$$

juhuslikul suurusel k võimalikkude väärtust ja $k-1$ etteantud kontrollitavat tõenäosust (k . tõenäosus alati teiste kaudu leitav)

$$G = 2 \ln \left(\frac{L}{L_0} \right) = 2 \ln \left(\frac{(n_1/n)^{n_1} \cdot (n_2/n)^{n_2} \cdot \dots \cdot (n_k/n)^{n_k}}{p_1^{n_1} \cdot p_2^{n_2} \cdot \dots \cdot p_k^{n_k}} \right)$$

$$= 2 \ln \left(\left(\frac{n_1}{np_1} \right)^{n_1} \cdot \dots \cdot \left(\frac{n_k}{np_k} \right)^{n_k} \right) = 2 \sum_{i=1}^k n_i \ln \left(\frac{n_i}{N p_i} \right) \sim \chi^2_{k-1}$$

Tõepärasuhte test vs Hii-ruut test

hübriid + hübriid

AB AB

AA 25% AB 50% BB 25%

$N_1=50$ $N_2=100$ $N_3=50$

$n_1=50$ $n_2=115$ $n_3=35$ $n=200$

$$\chi^2 = \frac{(50-50)^2}{50} + \frac{(100-115)^2}{100} + \frac{(50-35)^2}{50} = 6,75$$

$$G = 2 \left(50 \ln \left(\frac{50}{50} \right) + 115 \ln \left(\frac{115}{100} \right) + 35 \ln \left(\frac{35}{50} \right) \right) = 7,178$$

kriitilised väärtused

df 0,95-kvantiil

1	3,84
2	5,99
3	7,81
4	9,49
5	11,07

Tõepärasuhte test – sõltumatuse testimine

sugu	hinnang tervisele			kokku
	v.hea	hea	keskmine/halb	
naine	83 (13%)	404 (62%)	161 (25%)	648 (100%)
mees	35 (18%)	105 (55%)	50 (26%)	190 (100%)
				838

$$G = 2 \ln \left(\frac{L}{L_0} \right) = 2 \ln \left(\frac{(n_{11}/n)^{n_{11}} \cdot (n_{12}/n)^{n_{12}} \cdot \dots \cdot (n_{23}/n)^{n_{23}}}{(\hat{p}_{1.} \cdot \hat{p}_{.1})^{n_{11}} \cdot \dots \cdot (\hat{p}_{2.} \cdot \hat{p}_{.3})^{n_{23}}} \right) \approx 4,45$$

$$\chi^2 = \frac{(83 - \hat{p}_{.1} \cdot 648)^2}{\hat{p}_{.1} \cdot 648} + \frac{(404 - \hat{p}_{.2} \cdot 648)^2}{\hat{p}_{.2} \cdot 648} + \dots \approx 4,6$$

5-3 →

df	kriitilised väärtused 0,95-kvantiil
1	3,84
2	5,99
3	7,81
4	9,49
5	11,07

Ligikaudsed testid: tõepärasuhte ja hii-ruut test

Mõlemad testid eeldavad, et vaatluste arv on suur. Väikeste vaatluste arvu puhul pole nende teststatistikute jaotus nullhüpoteesi kehtides hii-ruut jaotusega.

Kui suur valim on piisavalt suur? Sõltub sellest, kui suur viga on Teie jaoks veel aktsepteeritav. Sage soovitus: ootuspärased sagedused ei tohiks olla väiksemad 5-st.

tõepärasuhte teststatistik (G-test)

$$G = 2 \sum_{i,j} N_{ij} \ln \left(\frac{N_{ij}}{E_{ij}} \right) \stackrel{H_0}{\sim} \chi^2_{(\text{ridade arv}-1) \cdot (\text{veergude arv}-1)}$$

Hii-ruut test

$$\chi^2 = \sum_{i,j} \frac{(N_{ij} - E_{ij})^2}{E_{ij}} \stackrel{H_0}{\sim} \chi^2_{(\text{ridade arv}-1) \cdot (\text{veergude arv}-1)}$$

Fisheri täpne test

Eksperiment - kas alguses valati tassi tee või piim?



```
> table(Tegelik, Oletus)
```

Oletus	arv	tõenäosus
Tegelik Piim Tee	0	
Piim	1	
Tee	2	
	3	
Teststatistikuks õigesti äraarvatud piim-enne tasside arv	4	0,014

Võimalikke teetasside järjestusi 8! tükki
Neist sobivaid: 4! · 4! (niipalju eri järjestuste korral läheks kõik otsused täppi)
Tõenäosus: 4! · 4! / 8! = 0,01428571

Fisheri täpne test

Eksperiment - kas alguses valati tassi tee või piim?



```
> table(Tegelik, Oletus)
```

Oletus	arv	tõenäosus
Tegelik Piim Tee	0	
Piim	1	
Tee	2	
	3	0,229
Teststatistikuks õigesti äraarvatud piim-enne tasside arv	4	0,014

Võimalikke teetasside järjestusi 8! tükki
Neist sobivaid: C₄¹ · C₄¹ · 4! · 4!
Tõenäosus: 4 · 4 · 4! / 8! = 0,2285714

Fisheri täpne test

Tegelik:



Blanche Muriel:

piim tee tee piim piim piim tee tee

```
> table(Tegelik, Oletus)
```

Oletus	Piim	Tee
Tegelik Piim	4	0
Tegelik Tee	0	4

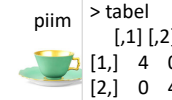
Teststatistiku jaotus H_0 kehtides

arv	tõenäosus
0	0,014
1	0,229
2	0,514
3	0,229
4	0,014

P-väärtus: $0,014 + 0,014 = 0,028$

Fisheri täpne test

Tegelik:



Blanche Muriel:

```
> tabel=rbind(c(4,0),c(0,4))
> tabel
      [,1] [,2]
[1,]  4  0
[2,]  0  4
> fisher.test(tabel)
```

Fisher's Exact Test for Count Data

```
data: tabel
p-value = 0.02857
alternative hypothesis: true odds ratio is not equal to 1
```

Oletus	Piim	Tee
Tegelik Piim	4	0
Tegelik Tee	0	4

Teststatistiku jaotus H_0 kehtides

arv	tõenäosus
0	0,014
1	0,229
2	0,514
3	0,229
4	0,014

P-väärtus: $0,014 + 0,014 = 0,028$

Fisheri täpne test – suuremad tabelid

Fisheri täpset testi saab kasutada ka suuremate sagedustabelite korral.

		tegelik		
		piim	tee	kohvi
arvamus	piim	3	0	0
	tee	0	3	0
	kohvi	0	0	3

H_0 kehtides on sama ekstreemseteks – sama väikese esinemistõenäosusega tabeliteks (vähetõenäolisemaid pole):

0	3	0	3	0	0
3	0	0	0	0	3
0	0	3	0	3	0
0	0	3	3	0	0
3	0	0	0	0	3
0	3	0	0	3	0
0	0	3	0	0	3
0	3	0	0	0	3
3	0	0	0	0	3

Kõik teised sagedustabelid esinevad vastuste juhuslikul pakkumisel suurema tõenäosusega

Fisheri täpse testi p-väärtus on seega $0,000595... \cdot 6 = 0,003571...$

Täpselt sellise tabeli nägemise tõenäosus on H_0 kehtides (inimene pakub juhuslikult) $3!3!3!/9! = 0,000595...$

Fisheri täpne test – suuremad tabelid

```
> sagedustabel=table(X,Y)
> sagedustabel
      Y
X     a b c
A     3 0 0
B     0 3 0
C     0 0 3
> fisher.test(sagedustabel)
```

Fisher's Exact Test for Count Data

```
data: sagedustabel
p-value = 0.003571
alternative hypothesis: two.sided
```

		tegelik		
		a	b	c
arvamus	A	3	0	0
	B	0	3	0
	C	0	0	3

Ekstreemseteks tabeliteks (vähetõenäolisemaid pole):

3	0	0
0	0	3
0	3	0
3	0	0
0	0	3
0	3	0
3	0	0
0	0	3
0	3	0

Fisheri täpse testi p-väärtus on seega $0,000595... \cdot 6 = 0,003571...$

Fisheri täpne test - tähelepanekud

Fisheri täpse testi puhul eeldati, et sagedustabeli marginaalid on fikseeritud (uurijad otsustavad, mitmesse tassi lisada esmalt piim ja mitme tassi korral esmalt valada teed... ja ka daam teadis, mitmesse tassi valati esmalt piima).

Võimalike korduseksperimentide tulemused
(sama valimi mahu – 8 tassi – korral:

piim	tee		
3	1	4	
1	3	4	
4	4		

piim	tee		
4	0	4	
0	4	4	
4	4		

piim	tee		
2	2	4	
2	2	4	
4	4		

Uuringut korrates jäävad sagedustabeli marginaalid samaks.

	suri	elab	
uus ravi	3	97	100
vana ravi	9	91	100
	12	188	

	suri	elab	
uus ravi	2	98	100
vana ravi	13	87	100
	15	185	

Kõigi uuringute korral ei jää uuringut korrates (mõlemad) marginaaljaotused samaks – kas see on probleem?

Fisheri täpne test - tähelepanekud

Fisheri täpse testi puhul eeldati, et sagedustabeli marginaalid on fikseeritud (uurijad otsustavad, mitmesse tassi lisada esmalt piim ja mitme tassi korral esmalt valada teed... ja ka daam teadis, mitmesse tassi valati esmalt piima).

Kõigi uuringute korral ei jää uuringut korrates (mõlemad) marginaaljaotused samaks – kas see on probleem?

$$P(\text{I liiki viga}) = \sum_{n_{+1}+n_{+2}=n} P(\text{I liiki viga} | N_{+1} = n_{+1}, N_{+2} = n_{+2}) P(N_{+1} = n_{+1}, N_{+2} = n_{+2})$$

$$< \sum_{n_{+1}+n_{+2}=n} 0,05 \cdot P(N_{+1} = n_{+1}, N_{+2} = n_{+2})$$

$$= 0,05 \sum_{n_{+1}+n_{+2}=n} P(N_{+1} = n_{+1}, N_{+2} = n_{+2}) = 0,05 \cdot 1 = 0,05$$

Järelikult võib Fisheri täpset testi kasutada ka siis, kui mõlemad marginaalid pole fikseeritud – esimest liiki viga ei tehta ikkagi lubatust sagedamini.

See ei tähenda, et Fisheri täpne test oleks sellisel juhul parim valik!

Barnardi täpne test

Barnardi täpse testi puhul eeldatakse, et 2x2 sagedustabelil on fikseeritud vaid üks marginaal. Näiteks fikseerime nende patsientide arvu, kellele antakse uut ravi ja samuti fikseerime uuringut palneeritades vana ravi või platseebot saavate patsientide arvu (kontrollgrupi suuruse). Kui paljudel neist patsientidest läheb hästi ja kui paljudel halvasti – see on juhuslik:

	suri	elab		
uus ravi	3	97		fikseeritud $n_{1+}=100$ $n_{2+}=100$
vana ravi	9	91		
	$N_{+1}=12$	$N_{+2}=188$		

Juhuslikud (uuringut korrates võiksime näha teistsuguseid numbreid)

Barnardi täpne test

Barnardi täpse testi puhul eeldatakse, et 2x2 sagedustabelil on fikseeritud vaid üks marginaal. Näiteks fikseerime nende patsientide arvu, kellele antakse uut ravi ja samuti fikseerime uuringut palneeritades vana ravi või platseebot saavate patsientide arvu (kontrollgrupi suuruse). Kui paljudel neist patsientidest läheb hästi ja kui paljudel halvasti – see on juhuslik:

	suri	elab		
uus ravi	3	97		fikseeritud $n_{1+}=100$ $n_{2+}=100$
vana ravi	9	91		
	$N_{+1}=12$	$N_{+2}=188$		

$N_{11} \sim B(n_{1+}; p_1)$
 $N_{21} \sim B(n_{2+}; p_2)$

Barnardi test:

$$H_0 : p_1 = p_2$$

Ei oska koostada teststatistikut, mille jaotus oleks nullhüpoteesi kehtides alati sama (sõltumata tõenäosuste p_1 ja p_2 väärtusest). Käsitleme kui liithüpoteesi.

Liithüpotees:

$H_0: p_1=0,1$ ja $p_2=0,1$ või

$p_1=0,2$ ja $p_2=0,2$ või

$p_1=0,23$ ja $p_2=0,23$ või ...

Kui üks neist väidetest kehtib, siis kehtib H_0 .

Barnardi täpne test

Liithüpoteesi

$$H_0 : p_1 = p_2$$

testimisest (idee)

H_0 : $p_1=0,1$ ja $p_2=0,1$ või
 $p_1=0,2$ ja $p_2=0,2$ või
 $p_1=0,23$ ja $p_2=0,23$ või ...

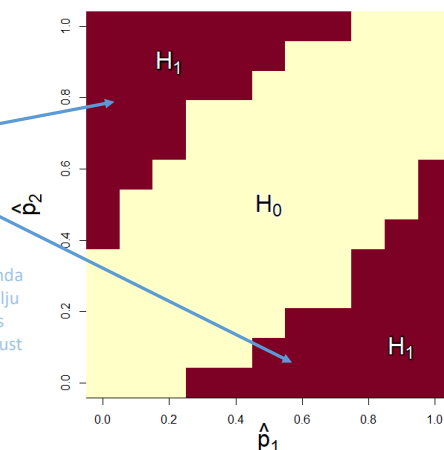
Kui tõenäosuste p_1 ja p_2 hinnangud sattuvad kuhugi siia piirkonda (kriitilisse regioon) siis kummutame nullhüpoteesi H_0 .

Kriitiline piirkond valitakse nii, et:

a) Mistahes $p_1=p_2$ väärtuste korral kriitilisse piirkonda sattumise tõenäosus poleks suurem kui α

b) Kõigist tingimust a) rahuldavatest kriitilistest piirkondadest valitakse välja suurim

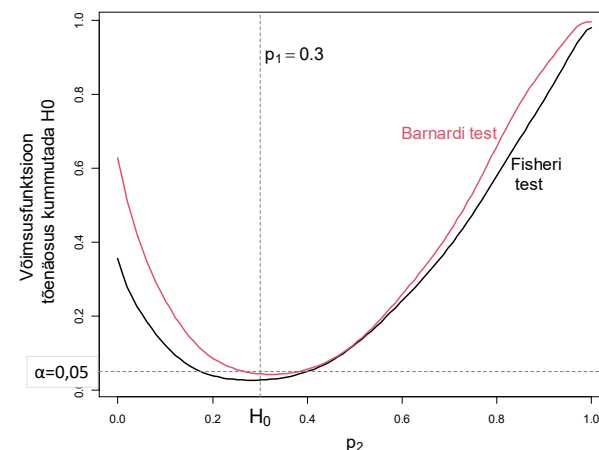
Ei määra kriitilist piirkonda üheselt – eksisteerib palju võimalikke piirkondi mis rahuldavad seda tingimust



Barnardi täpne test

Juhul, kui fikseeritud on vaid üks marginaaljaotus, on Barnardi täpne test suurema võimsusega – suudab tõenäolisemalt märgata, et tegelikult kehtib alternatiivne hüpotees (samas tegemata liiga sageli I-liiki viga siis, kui tegelikult kehtib nullhüpotees H_0).

Siintoodud joonis kirjeldab Barnardi täpse testi ja Fisheri täpse testi võimsuseid siis, kui ravime 10 patsienti ühe raviga (ravi A) ja 12 patsienti teise raviga (ravi B). Ravi A saanutel on hea tulemuse tõenäosus 0,3.



Barnardi täpne test

Otsime geenivaramu abil välja 10 teatud geenimutatsiooni kandjat ja 12 „kontrolli“ – inimest, kellel pole mutatsiooni. Seejärel korraldame väljavalitutele spordivõistluse ja vaatame, kas inimene sattus 6 parema sekka või mitte.

Antud näite puhul on **mõlemad marginaaljaotused fikseeritud**.

	parimate		
	seas	kaotajad	
mutatsioon	...	...	$n_{1+}=10$
pole mutant	...	...	$n_{2+}=12$
	$N_{+1}=6$	$N_{+2}=16$	

Oletame, et kehtib nullhüpotees – geenimutatsioon ei mõjuta kuidagi inimeste füüsilist võimekust. Teeme Barnardi täpse testi kasutades olulisuse nivood 0,05.

Barnardi täpne test otsustaks H_0 alternatiivse hüpoteesi kasuks ~5,6% valimistest

Barnardi täpset testi kirjeldatud olukorras kasutada EI TOHI!

Tehnilised probleemid Fisheri / Barnardi täpse testiga

Fisheri testi jaoks arvutused võivad olla aeglased (eriti suuremate kui 2x2 sagedustabelite korral); Barnardi täpse testi jaoks arvutused (näiteks kriitilise piirkonna leidmine) sageli üliaeglased – eriti suuremate valimite korral.

```
fisher.test(sagedustabel,
  simulate.p.value=TRUE,
  B=1000000)
```

Ligikaudne p-väärtuse arvutus Fisheri täpse testi abil (Monte-Carlo meetodit kasutades, simuleerides 1000000 valimit eeldades H_0 kehtimist)

Vahel eelistame ligikaudseid arvutusi täpsetele arvutustele mille tulemusi peaksime ootama päevi või nädalaid.

```
library(Barnard)
barnard.test(10, 3, 40, 47)
```

Arvutus võtab aega paar minutit

```
library(DescTools)
BarnardTest(rbind(c(10, 40), c(3, 47)), method="z-pooled")
```

Arvutus kestab ehk sekundi

Kriitilise piirkonna paikapanekut lihtsustatakse – tulemuseks pole „parim“ kriitiline piirkond, aga siiski „hea“ ja valiidne kriitiline piirkond (I-liiki viga ei tehta liiga sageli).