

Testidest II

Testid mitterandomiseeritud uuringutes

Kolmanda tunnuse mõju

(Pole probleemiks randomiseeritud uuringu korral)

$X \setminus Y$		suri COVIDi tõttu	ei surnud	
	vaktsineeritud	205	199 800	$205 / 200\,005 \approx 0,001$
	pole vakts.	520	1 000 980	$520 / 1\,001\,500 \approx 0,0005$
				suhteline risk: $0,001/0,0005 \approx 2$

vanad	$X \setminus Y$	suri COVIDi tõttu	ei surnud	
	vaktsineeritud	200	99 800	$200 / 100\,000 = 0,002$
	pole vakts.	20	980	$20 / 1\,000 = 0,02$
				suhteline risk: 0,1

noored	$X \setminus Y$	suri COVIDi tõttu	ei surnud	
	vaktsineeritud	5	100 000	$5 / 100\,005 \approx 0,00005$
	pole vakts.	500	1 000 000	$500 / 1\,000\,500 \approx 0,0005$
				suhteline risk: 0,1

Kolmanda tunnuse mõju

(Pole probleemiks randomiseeritud uuringu korral)

$X \setminus Y$		suri COVIDi tõttu	ei surnud	
	vaktsineeritud	205	199 800	$205 / 200\,005 \approx 0,001$
	pole vakts.	520	1 000 980	$520 / 1\,001\,500 \approx 0,0005$
				suhteline risk: $0,001/0,0005 \approx 2$

vanad	$X \setminus Y$	suri COVIDi tõttu	ei surnud	
	vaktsineeritud	200	99 800	$200 / 100\,000 = 0,002$
	pole vakts.	20	980	$20 / 1\,000 = 0,02$
				suhteline risk: 0,1

noored	$X \setminus Y$	suri COVIDi tõttu	ei surnud	
	vaktsineeritud	5	100 000	$5 / 100\,005 \approx 0,00005$
	pole vakts.	500	1 000 000	$500 / 1\,000\,500 \approx 0,0005$
				suhteline risk: 0,1

vanus

vaktsineeritus

COVID-suremus

?

Kas seos (vaktsineerimise ja COVID-suremuse vahel) ikka eksisteerib?

Kui tugev see põhjuslik mõju on?

Cochran-Mantel-Haenzeli test

Y	X			
	eksp+	eksp-		
sündmus toimus	n_{11}	n_{12}	n_{1+}	Fikseeritud marginaalsagedused
ei toimunud	n_{21}	n_{22}	n_{2+}	
	n_{+1}	n_{+2}	n_{++}	

Cochran-Mantel-Haenzeli test

Y	X			
	eksp+	eksp-		
sündmus toimus	n_{11}	n_{12}	n_{1+}	Fikseeritud marginaalsagedused
ei toimunud	n_{21}	n_{22}	n_{2+}	
	n_{+1}	n_{+2}	n_{++}	

$$H_0: \Pr[N_{11} = n_{11}] = \frac{\binom{n_{+1}}{n_{11}} \binom{n_{+2}}{n_{12}}}{\binom{n_{++}}{n_{1+}}}$$

Kuna eeldame, et kehtib H_0 ,
ekspositsioon (nt vaksineerimine) ei
mõju, siis ei sõltu sellest järjekorrast ka
see, kes sureb või ei sure...

Märgime üles ka variandid, millises
järjekorras võivad inimesed surra –
milliseid „suremismustreid“ võib esineda?

Mitu sellist mustrit üldjuhul esineb?

Y	X			
	eksp+	eksp-		
sündmus toimus	n_{11}	n_{12}	2	E+ E- E+ E- S+ S+ S- S- S+ S- S+ S- S+ S- S- S+ S- S+ S+ S- S- S+ S- S+ S- S- S+ S+ S- S- S+ S+
ei toimunud	n_{21}	n_{22}	2	
	2	2	4	

Cochran-Mantel-Haenzeli test

Y	X			
	eksp+	eksp-		
sündmus toimus	n_{11}	n_{12}	n_{1+}	Fikseeritud marginaalsagedused
ei toimunud	n_{21}	n_{22}	n_{2+}	
	n_{+1}	n_{+2}	n_{++}	

$$H_0: \Pr[N_{11} = n_{11}] = \frac{\binom{n_{+1}}{n_{11}} \binom{n_{+2}}{n_{12}}}{\binom{n_{++}}{n_{1+}}}$$

Võimalusi valida n_{1+} inimest, kellel
sündmus toimub

H_0 kehtides (ekspositsioon ja tulemus
pole seotud) on kõik need variandid
võrdvõimalikud; $P(\text{järjestus})=1/6$

Y	X			
	eksp+	eksp-		
sündmus toimus	n_{11}	n_{12}	2	E+ E- E+ E- S+ S+ S- S- S+ S- S+ S- S+ S- S- S+ S- S+ S+ S- S- S+ S- S+ S- S- S+ S+ S- S- S+ S+
ei toimunud	n_{21}	n_{22}	2	
	2	2	4	

$$\binom{n_{++}}{n_{1+}} = C_{n_{++}}^{n_{1+}} = C_4^2 = 6$$

Mitu sellist mustrit üldjuhul esineb?

Cochran-Mantel-Haenzeli test

Y	X			
	eksp+	eksp-		
sündmus toimus	n_{11}	n_{12}	n_{1+}	Fikseeritud marginaalsagedused
ei toimunud	n_{21}	n_{22}	n_{2+}	
	n_{+1}	n_{+2}	n_{++}	

$$H_0: \Pr[N_{11} = n_{11}] = \frac{\binom{n_{+1}}{n_{11}} \binom{n_{+2}}{n_{12}}}{\binom{n_{++}}{n_{1+}}}$$

Võimalusi valida n_{1+} inimest, kellel
sündmus toimub

$$\Pr[N_{11} = 1] = \frac{1}{6} \cdot \text{sobivate võimaluste arv}$$

Y	X			
	eksp+	eksp-		
sündmus toimus	n_{11}	n_{12}	2	E+ E- E+ E- S+ S+ S- S- S+ S- S+ S- S+ S- S- S+ S- S+ S+ S- S- S+ S- S+ S- S- S+ S+ S- S- S+ S+
ei toimunud	n_{21}	n_{22}	2	
	2	2	4	

H_0 kehtides (ekspositsioon ja tulemus
pole seotud) on kõik need variandid
võrdvõimalikud; $P(\text{järjestus})=1/6$

$$\binom{n_{++}}{n_{1+}} = C_{n_{++}}^{n_{1+}} = C_4^2 = 6$$

Mitu sellist mustrit üldjuhul esineb?

Cochran-Mantel-Haenzeli test

Y	X			
	eksp+	eksp-		
sündmus toimus	n_{11}	n_{12}	n_{1+}	Fikseeritud marginaalsagedused
ei toimunud	n_{21}	n_{22}	n_{2+}	
	n_{+1}	n_{+2}	n_{++}	

Millised eksponeeritud
saavad endale sündmuse?
Võimalusi valida välja n_{11}
„surijat“ n_{+1} eksponeeritu
seast

$$\Pr[N_{11} = n_{11}] = \frac{\binom{n_{+1}}{n_{11}} \binom{n_{+2}}{n_{12}}}{\binom{n_{++}}{n_{1+}}}$$

Võimalusi valida n_{1+} inimest, kellel
sündmus toimub

$$\Pr[N_{11} = 1] = \frac{1}{6} \cdot C_2^1 \cdot C_2^1$$

Y	X			
	eksp+	eksp-		
sündmus toimus	n_{11}	n_{12}	2	E+ E- E+ E- S+ S+ S- S- S+ S- S+ S- S+ S- S- S+ S- S+ S+ S- S- S+ S- S+ S- S- S+ S+ S- S- S+ S+
ei toimunud	n_{21}	n_{22}	2	
	2	2	4	

Millised mitteeksponeeritud saavad endale
sündmuse? Võimalusi valida n_{12} surma n_{+2}
mitteeksponeeritu seast

H_0 kehtides (ekspositsioon ja tulemus
pole seotud) on kõik need variandid
võrdvõimalikud; $P(\text{järjestus})=1/6$

$$\binom{n_{++}}{n_{1+}} = C_{n_{++}}^{n_{1+}} = C_4^2 = 6$$

Mitu sellist mustrit üldjuhul esineb?

Cochran-Mantel-Haenzeli test

$$\Pr[N_{11} = n_{11}] = \frac{\binom{n+1}{n_{11}} \binom{n+2}{n_{12}}}{\binom{n++}{n_{1+}}}$$

$$E(N_{11}) = \sum_x \frac{x \binom{n+1}{x} \binom{n+2}{n_{12}}}{\binom{n++}{n_{1+}}} \quad x \binom{m}{x} = \frac{m!}{(x-1)!(m-x)!} = \frac{m(m-1)!}{(x-1)!((m-1)-(x-1))!} = m \binom{m-1}{x-1}$$

$$E(N_{11}) = \sum_x \frac{n_{+1} \binom{n+1-1}{x-1} \binom{n+2}{n_{21}}}{\binom{n++}{n_{1+}}}$$

$$E(N_{11}) = \sum_x \frac{n_{+1} \binom{n+1-1}{x-1} \binom{n+2}{n_{21}}}{\binom{n++}{n_{1+}} \cdot n_{1+}/n_{1+}}$$

$$E(N_{11}) = \sum_x \frac{n_{+1} \binom{n+1-1}{x-1} \binom{n+2}{n_{21}}}{\binom{n_{++}-1}{n_{1+}-1} \cdot n_{++}/n_{1+}}$$

$$E(N_{11}) = \frac{n_{+1}n_{1+}}{n_{++}} \sum_x \frac{\binom{n+1-1}{x-1} \binom{n+2}{n_{21}}}{\binom{n_{++}-1}{n_{1+}-1}}$$

vt näiteks:
<https://planetmath.org/proofofvarianceofthehypergeometricdistribution>

$$D[N_{11}] = \frac{n_{1+}n_{+1}n_{2+}n_{+2}}{n_{++}^2(n_{++}-1)},$$

$$E(N_{11}) = \frac{n_{+1}n_{1+}}{n_{++}}$$

Cochran-Mantel-Haenzeli test

$$N_{111} \xrightarrow{n \rightarrow \infty; H_0} N(E(N_{111}); D(N_{111})) \quad N_{112} \xrightarrow{n \rightarrow \infty; H_0} N(E(N_{112}); D(N_{112}))$$

$$D[N_{111}] \stackrel{H_0}{=} \frac{n_{1+1}n_{+11}n_{2+1}n_{+21}}{n_{++1}^2(n_{++1}-1)}, \quad D[N_{112}] \stackrel{H_0}{=} \frac{n_{1+2}n_{+12}n_{2+2}n_{+22}}{n_{++2}^2(n_{++2}-1)} \quad \dots$$

$$E(N_{111}) \stackrel{H_0}{=} \frac{n_{1+1}n_{+11}}{n_{++1}} \quad E(N_{112}) \stackrel{H_0}{=} \frac{n_{1+2}n_{+12}}{n_{++2}}$$

Cochran-Mantel-Haenzeli test

Kui kõigis gruppides
(noored; vanad; ...)
kehtib H_0

$$\sum_i N_{11i} \xrightarrow{n \rightarrow \infty; H_0} N \left(\sum_i E(N_{11i}); \sum_i D(N_{11i}) \right)$$

$$\frac{\sum_i N_{11i} - \sum_i E(N_{11i})}{\sqrt{\sum_i D(N_{11i})}} \xrightarrow{n \rightarrow \infty; H_0} N(0; 1)$$

$$N_{111} \xrightarrow{n \rightarrow \infty; H_0} N(E(N_{111}); D(N_{111})) \quad N_{112} \xrightarrow{n \rightarrow \infty; H_0} N(E(N_{112}); D(N_{112}))$$

$$D[N_{111}] \stackrel{H_0}{=} \frac{n_{1+1}n_{+11}n_{2+1}n_{+21}}{n_{++1}^2(n_{++1}-1)}, \quad D[N_{112}] \stackrel{H_0}{=} \frac{n_{1+2}n_{+12}n_{2+2}n_{+22}}{n_{++2}^2(n_{++2}-1)} \quad \dots$$

$$E(N_{111}) \stackrel{H_0}{=} \frac{n_{1+1}n_{+11}}{n_{++1}} \quad E(N_{112}) \stackrel{H_0}{=} \frac{n_{1+2}n_{+12}}{n_{++2}}$$

Cochran-Mantel-Haenzeli test

Kui kõigis gruppides
(noored; vanad; ...)
kehtib H_0

$$\sum_i N_{11i} \xrightarrow{n \rightarrow \infty; H_0} N \left(\sum_i E(N_{11i}); \sum_i D(N_{11i}) \right)$$

$$\frac{\sum_i N_{11i} - \sum_i E(N_{11i})}{\sqrt{\sum_i D(N_{11i})}} \xrightarrow{n \rightarrow \infty; H_0} N(0; 1)$$

$$\frac{(\sum_i N_{11i} - \sum_i E(N_{11i}))^2}{\sum_i D(N_{11i})} \xrightarrow{n \rightarrow \infty; H_0} \chi_{df=1}^2$$

$$D[N_{111}] \stackrel{H_0}{=} \frac{n_{1+1}n_{+11}n_{2+1}n_{+21}}{n_{++1}^2(n_{++1}-1)}, \quad D[N_{112}] \stackrel{H_0}{=} \frac{n_{1+2}n_{+12}n_{2+2}n_{+22}}{n_{++2}^2(n_{++2}-1)} \quad \dots$$

$$E(N_{111}) \stackrel{H_0}{=} \frac{n_{1+1}n_{+11}}{n_{++1}} \quad E(N_{112}) \stackrel{H_0}{=} \frac{n_{1+2}n_{+12}}{n_{++2}}$$

vanad	$X \setminus Y$	suri COVIDi tõttu	ei surnud	$E(N_{111}) = \frac{100\,000 \cdot 220}{101\,000} = 217,82 \dots$ $D[N_{111}] = \frac{100\,000 \cdot 220 \cdot 1\,000 \cdot 100\,780}{101\,000^2 \cdot (101\,000 - 1)}$
	vaktsineeritud	200	99 800	
	pole vakts.	20	980	

noored	$X \setminus Y$	suri COVIDi tõttu	ei surnud	$E(N_{112}) = \frac{100\,005 \cdot 505}{110\,0505} = 45,89 \dots$ $D[N_{112}] = \frac{100\,005 \cdot 505 \cdot 1\,000\,500 \cdot 1\,100\,000}{1\,100\,505^2 \cdot (1\,100\,505 - 1)}$
	vaktsineeritud	5	100 000	
	pole vakts.	500	1 000 000	

$$D[N_{111}] \stackrel{H_0}{=} \frac{n_{1+1}n_{+11}n_{2+1}n_{+21}}{n_{++1}^2(n_{++1} - 1)}, \quad D[N_{112}] \stackrel{H_0}{=} \frac{n_{1+2}n_{+12}n_{2+2}n_{+22}}{n_{++2}^2(n_{++2} - 1)}$$

$$E(N_{111}) \stackrel{H_0}{=} \frac{n_{1+1}n_{+11}}{n_{++1}}, \quad E(N_{112}) \stackrel{H_0}{=} \frac{n_{1+2}n_{+12}}{n_{++2}}$$

vanad	$X \setminus Y$	suri COVIDi tõttu	ei surnud	$E(N_{111}) = \frac{100\,000 \cdot 220}{101\,000} = 217,82 \dots$ $D[N_{111}] = \frac{100\,000 \cdot 220 \cdot 1\,000 \cdot 100\,780}{101\,000^2 \cdot (101\,000 - 1)}$
	vaktsineeritud	200	99 800	
	pole vakts.	20	980	

noored	$X \setminus Y$	suri COVIDi tõttu	ei surnud	$E(N_{112}) = \frac{100\,005 \cdot 505}{110\,0505} = 45,89 \dots$ $D[N_{112}] = \frac{100\,005 \cdot 505 \cdot 1\,000\,500 \cdot 1\,100\,000}{1\,100\,505^2 \cdot (1\,100\,505 - 1)}$
	vaktsineeritud	5	100 000	
	pole vakts.	500	1 000 000	

$$\xi_{CMH} = \frac{((200 + 5) - (217,82 + 45,89))^2}{2,15 + 41,7} \approx 78,60596 \text{ >3,84 } \Rightarrow H1$$

> mantelhaen.test(Covid, correct=FALSE)
Mantel-Haenszel chi-squared test without continuity correction
Mantel-Haenszel X-squared = 78.606, df = 1, p-value < 2.2e-16
sample estimates:
common odds ratio
0.09945329

Koondhinnang OR-le

$$\hat{\theta}_{MH} = \frac{\sum_{k=1}^K n_{11k} n_{22k} / n_{++k}}{\sum_{k=1}^K n_{12k} n_{21k} / n_{++k}}.$$

Logistiline regressioon

Olgu edu tõenäosus fikseeritud x_i ja k korral

$$\pi_{ik} = P(Y = 1 | X = x_i, Z = k).$$

Tähistame edu log-shansse tavapäraselt

$$\text{logit}(\pi_{ik}) = \ln \frac{P(Y = 1 | X = x_i, Z = k)}{1 - P(Y = 1 | X = x_i, Z = k)}.$$

Eeldame edu log-shanssidele mudelit

$$\text{logit}(\pi_{ik}) = \alpha + \beta x_i + \beta_k^Z, \quad i = 1, 2, \quad k = 1, \dots, K,$$

mis väljendab eeldust, et shansside suhe on igas osatabelis k sama, nimelt

$$\theta = e^\beta.$$

Logistiline regressioon

```
> m1=glm(cbind(surnud, elus)~factor(vanus)+factor(vakts),
family=binomial())
> summary(m1)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-7.60076	0.04456	-170.58	<2e-16 ***
factor(vanus) vana	3.70536	0.20581	18.00	<2e-16 ***
factor(vakts) +	-2.31685	0.21001	-11.03	<2e-16 ***

```
> exp(-2.31685)
[1] 0.09858364
```

Hinnang ühisele OR-le