

Biostatistika
2. loeng

Riskide hindamine ja võrdlemine

Märt Möls
martm@ut.ee

Näide: kas aspiriin on hea?

Harvardi meditsiinikooli inimesed jagasid 22071 inimest juhuslikult kahte gruppi. Pooled neist inimestest võtsid ülepäeva sisse tableti aspiriini, ülejäänud võtsid ülepäeva sisse tableti platseebot. Inimesed ise ei teadnud, kas nad võtavad platseebot või aspiriini. Katsealuseid jälgiti viis aastat ja loeti siis kokku, kui paljud katsealustest said infarkti ühes ja kui paljud teises rühmas.

	Infarkt	ei saanud Infarkti	Kokku
Aspiriin	104	10 933	11 037
Platseebo	189	10 845	11 034
Kokku	293	21 778	22 071

Fisher'i täpne test

```
> Truth  
[1] "Milk" "Tea" "Tea" "Milk" "Milk" "Milk" "Tea" "Tea"
```

```
> table(Truth, Guess)  
      Guess  
Truth Milk Tea  
Milk   3   1  
Tea    1   3
```

Teststatistiku jaotus H_0 kehtides

0	0.014
1	0.229
2	0.514
3	0.229
4	0.014

P-väärtus: $0,014+0,229+0,229+0,014 = 0,486$

```
> Guess  
[1] "Milk" "Tea" "Tea" "Tea" "Milk" "Milk" "Tea" "Milk"
```

Fisher'i täpne test

```
> Truth  
[1] "Milk" "Tea" "Tea" "Milk" "Milk" "Milk" "Tea" "Tea"
```

```
> fisher.test(table(Guess, Truth))  
      Fisher's Exact Test for Count Data  
p-value = 0.4857  
alternative hypothesis: true odds ratio is not equal to 1  
95 percent confidence interval:  
 0.2117329 621.9337505  
sample estimates:  
odds ratio  
 6.408309
```

P-väärtus: $0,014+0,229+0,229+0,014 = 0,486$

```
> Guess  
[1] "Milk" "Tea" "Tea" "Tea" "Milk" "Milk" "Tea" "Milk"
```

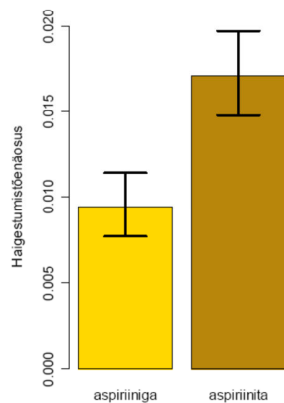
Kui palju aitab aspiriin? Kas erinevus haigestumisriskide vahel on meditsiiniliselt oluline?

p_T – risk haigestuda eksponeeritute seas (aspiriini tarvitades);

p_C – risk haigestuda kontrollrühmas (ilma aspiriinita);

Näites:

$$\hat{p}_T = 0,00942; \hat{p}_C = 0,01713$$



Ekspositsiooni mõju tugevust iseloomustavad statistikud:

- Riskide vahe (Risk difference; Absolute Risk Reduction; Absolute Risk Increase);

$$ARR = p_C - p_T$$

Näites: $\widehat{ARR} = 0,00771$

Interpretatsioon: eksponeerides 100 000 inimest näeme $ARR * 100\ 000 = 771$ haigusjuhtu vähem.

- Suhteline risk (Relative Risk):

$$RR = p_T / p_C = 0,55$$

- Suhteline riski muut (Relative Risk Reduction/Increase):

$$RRR = (p_C - p_T) / p_C = 0,45 \text{ (45\%)}$$

- Number needed to treat (NNT)

$$NNT = 1/ARR \\ = 1/0,00771 = 130$$

Mugav kasutada näiteks rahalistes arvutustes: ravi maksumus x krooni, ühe haigusjuhu ärahoidmise maksumus $x * NNT$ krooni.

- Šansside suhe (Odds ratio)

$$OR = [p_T / (1 - p_T)] / [p_C / (1 - p_C)] \\ = 0,546$$

Kõige levinum ekspositsiooni mõju tugevust kirjeldav statistik

Usaldusintervallid

$$D(\widehat{ARR}) = D(\hat{p}_C - \hat{p}_T) = p_C(1 - p_C) / n_C + p_T(1 - p_T) / n_T$$

$$\widehat{D}(\widehat{ARR}) = \hat{p}_C(1 - \hat{p}_C) / n_C + \hat{p}_T(1 - \hat{p}_T) / n_T$$

ligikaudne $(1 - \alpha)$ -usaldusintervall ARR -le:

$$\widehat{ARR} \pm z_{\alpha/2} \sqrt{\widehat{D}(\widehat{ARR})}$$

Usaldusintervalli NNT -le saame leida kasutades ARR -le leitud usaldusintervalli:

$$P \left(\widehat{ARR} + z_{\alpha/2} \sqrt{\widehat{D}(\widehat{ARR})} \leq ARR \leq \widehat{ARR} - z_{\alpha/2} \sqrt{\widehat{D}(\widehat{ARR})} \right) = 1 - \alpha \\ P \left(\frac{1}{\widehat{ARR} - z_{\alpha/2} \sqrt{\widehat{D}(\widehat{ARR})}} \leq NNT \leq \frac{1}{\widehat{ARR} + z_{\alpha/2} \sqrt{\widehat{D}(\widehat{ARR})}} \right) = 1 - \alpha$$

Usaldusintervall suhtelisele riskile (RR)

Usaldusintervalli konstrueerimise idee: leiame esmalt usaldusintervalli $\log(RR)$ -le ja konstrueerime seada kasutades usaldusintervalli RR -le eksponentfunksiooni abil.

$$\log(RR) = \log(p_T) - \log(p_C).$$

Teame (tsentraalne piirteoreem) et:

$$\frac{\widehat{p}_T - p_T}{\sqrt{D(\widehat{p}_T)}} \xrightarrow{n_T \rightarrow \infty} N(0; 1)$$

Milline on aga $\log(\widehat{p}_T)$ asümptootiline jaotus? Selgub, et selleks jaotuseks on samuti normaaljaotus...

Siinkohal on sobilik tutvuda nn. delta-meetodiga:

(usaldusintervalli leidmine RR-le jätkub hiljem...)

Delta meetod

(ehk kasutame kavalalt ära Taylori rittaarenduse esimesi liikmeid)

Olgu $X \sim N(\mu, \sigma^2)$, ja $Y = g(X)$. Kui $g'(\mu) \neq 0$, siis ka Y on ligikaudu normaaljaotusega juhuslik suurus:

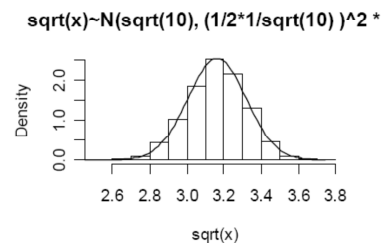
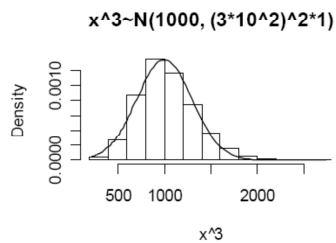
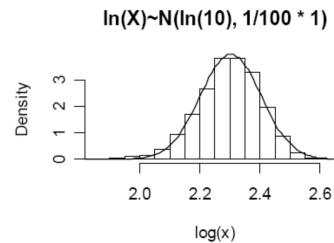
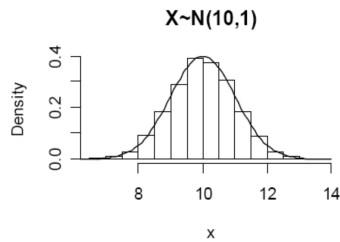
$$Y \sim N(g(\mu), [g'(\mu)]^2 \sigma^2).$$

Delta-meetodi mitmemõõtmeline variant:

Kui $X \sim N(\mu, \Sigma)$, $Y = g(X)$

Siis, ligikaudu,

$$Y \sim N(g(\mu), [g'(\mu)]^T \Sigma [g'(\mu)])$$



(jätkub usaldusintervalli leidmine RR-le...)

Seega Delta-meetodit kasutades on (asümptootiliselt, ligikaudu):

$$\log(\widehat{p}_T) \sim N\left(\log(p_T); \frac{1}{p_T^2} \cdot \frac{p_T(1-p_T)}{n_T}\right)$$

ja

$$\log(\widehat{p}_T) - \log(\widehat{p}_C) \sim N\left(\log(p_T) - \log(p_C); \frac{1-p_T}{p_T n_T} + \frac{1-p_C}{p_C n_C}\right)$$

ehk

$$\log(\widehat{RR}) \sim N\left(\log(RR); \frac{1-p_T}{p_T n_T} + \frac{1-p_C}{p_C n_C}\right)$$

Sagedustabelit kiigates:

	ravi	kontroll
suri	a	c
ei surnud	b	d

$$\log(\widehat{RR}) \sim N\left(\log(RR); \frac{1-p_T}{p_T n_T} + \frac{1-p_C}{p_C n_C}\right)$$

$$\widehat{p}_T = a/(a+b) \quad 1 - \widehat{p}_T = b/(a+b)$$
$$\widehat{p}_C = c/(c+d) \quad 1 - \widehat{p}_C = d/(c+d)$$

$$\frac{1 - \widehat{p}_T}{\widehat{p}_T n_T} = \frac{b}{a(a+b)} \quad \frac{1 - \widehat{p}_C}{\widehat{p}_C n_C} = \frac{d}{c(c+d)}$$

Ehk $\log(RR)$ hinnangu dispersiooni saame hinnata valemiga:

$$\widehat{D}(\log(\widehat{RR})) = \frac{b}{a(a+b)} + \frac{d}{c(c+d)}$$

95%-usaldusintervall $\log(RR)$ -le on seega:

$$\log(\widehat{RR}) \pm 1.96 \sqrt{\frac{b}{a(a+b)} + \frac{d}{c(c+d)}}$$

ja usaldusintervalli RR -le saame eksponentfunktsiooni kasutades:

$$\widehat{RR} \cdot \exp\left(\pm 1.96 \sqrt{\frac{b}{a(a+b)} + \frac{d}{c(c+d)}}\right)$$

Esimesed kliinilised randomiseeritud uuringud:

1835 – Nürnbergi soolatest (pimetest; testiti tol ajal äärmiselt populaarset homöopaatia ravimeetodit ja leiti see olevat mitteefektiivne. 100 katsealust, 50 said destilleeritud lumesulamisvett; 50 nn homöopaatilist soolalahust kontsentratsioonis 1:10⁶⁰)

<http://www.jameslindlibrary.org/articles/inventing-the-randomized-double-blind-trial-the-numberg-salt-test-of-1835/>

1946 – tuberkuloosi ravi streptomütsiini abil (antibiootikum)