



Katedra
aplikované matematiky



Dokážete si vybrat správnou nemocnici?
aneb

~~Ministr zdravotnictví ML varuje~~
Martina Litschmannová

ŠKOMAM 2013

Pán Úzkostný onemocněl chorobou Moribundus. V blízkosti jeho bydliště jsou dvě nemocnice – Dolní Lomná a Nová Dláha. Pán Úzkostný je pečlivý a sežene si statistiky obou nemocnic v léčbě této choroby.

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 (80%)	200 (20%)	1000
Nová Dláha	900 (90%)	100 (10%)	1000
Celkem	1700 (85%)	300 (15%)	2000

Nemocnice v Nové Dláze je úspěšnější!!!



Základní pojmy používané v oblasti analýzy závislosti kategoriálních veličin

Kontingenční tabulka

$X \setminus Y$	$y_{[1]}$	$y_{[2]}$	\dots	$y_{[s]}$	Celkem
$x_{[1]}$	n_{11}	n_{12}	\dots	n_{1s}	$n_{1.}$
$x_{[2]}$	n_{21}	n_{22}	\dots	n_{2s}	$n_{2.}$
\vdots	\vdots	\vdots	\dots	\vdots	\vdots
$x_{[r]}$	n_{r1}	n_{r2}	\dots	n_{rs}	$n_{r.}$
Celkem	$n_{.1}$	$n_{.2}$	\dots	$n_{.s}$	n

Schéma rozšířené kontingenční tabulky

Dvourozměrná tabulka četností, z jejichž hodnot můžeme usoudit na závislost či nezávislost mezi dvěma kategoriálními proměnnými.

Asociační tabulka – kontingenční tabulka typu 2x2

Doplňkové číselné charakteristiky pro kontingenční tabulky

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu		
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen
Dolní Lomná	800	200
Nová Dláha	900	100

Doplňkové číselné charakteristiky pro kontingenční tabulky

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu		
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen
Dolní Lomná	800	200
Nová Dláha	900	100

- četnosti

Doplňkové číselné charakteristiky pro kontingenční tabulky

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800	200	1000
Nová Dláha	900	100	1000
Celkem	1700	300	2000

- četnosti
- marginální četnosti

Doplňkové číselné charakteristiky pro kontingenční tabulky

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 800/2000	200 200/2000	1000
Nová Dláha	900 900/2000	100 100/2000	1000
Celkem	1700	300	2000

- četnosti
- marginální četnosti
- rel. četnosti

Doplňkové číselné charakteristiky pro kontingenční tabulky

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 40 %	200 10 %	1000
Nová Dláha	900 45 %	100 5 %	1000
Celkem	1700	300	2000

- četnosti
- marginální četnosti
- rel. četnosti

Doplňkové číselné charakteristiky pro kontingenční tabulky

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 40 % 800/1000	200 10 % 200/1000	1000
Nová Dláha	900 45 % 900/1000	100 5 % 100/1000	1000
Celkem	1700	300	2000

- četnosti
- marginální četnosti
- rel. četnosti
- řádkové rel. četnosti

Doplňkové číselné charakteristiky pro kontingenční tabulky

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 40 % 80 %	200 10 % 20 %	1000
Nová Dláha	900 45 % 90 %	100 5 % 10 %	1000
Celkem	1700	300	2000

- četnosti
- marginální četnosti
- rel. četnosti
- řádkové rel. četnosti

Doplňkové číselné charakteristiky pro kontingenční tabulky

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 40 % 80 % 800/1700	200 10 % 20 % 200/300	1000
Nová Dláha	900 45 % 90 % 900/1700	100 5 % 10 % 100/300	1000
Celkem	1700	300	2000

- četnosti
- marginální četnosti
- rel. četnosti
- řádkové rel. četnosti
- sloupcové rel. četnosti

Doplňkové číselné charakteristiky pro kontingenční tabulky

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 40 % 80 % 47,1 %	200 10 % 20 % 66,7 %	1000
Nová Dláha	900 45 % 90 % 52,9 %	100 5 % 10 % 33,3 %	1000
Celkem	1700	300	2000

- četnosti
- marginální četnosti
- rel. četnosti
- řádkové rel. četnosti
- sloupcové rel. četnosti

Míry asociace

(pouze pro asociační tabulky)

Poměr šancí (angl. „odds ratio“), nazýván také **křížový poměr** (angl. „cross product ratio“)

Pozorovaný poměr počtu úspěchů k počtu neúspěchů (tzv. pozorovaná **šance**) za okolností I. je $\frac{a}{b}$, za okolností II. $\frac{c}{d}$. Odhad poměru šancí je pak

$$\widehat{OR} = \frac{ad}{bc}.$$

X (okolnosti)\Y(výskyt události)	$y_{[1]}$ (úspěch)	$y_{[2]}$ (neúspěch)	Celkem
$x_{[1]}$ (I.)	a	b	$a + b$
$x_{[2]}$ (II.)	c	d	$c + d$
<i>Celkem</i>	$a + c$	$b + d$	n

Schéma rozšířené asociační tabulky

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 (80%)	200 (20%)	1000
Nová Dláha	900 (90%)	100 (10%)	1000
Celkem	1700 (85%)	300 (15%)	2000

- Odhad šance vyléčení Moribundu v Dolní Lomné je

$$\frac{a}{b} = \frac{800}{200} = 4.$$

- Odhad šance vyléčení Moribundu v Nové Dláze je

$$\frac{c}{d} = \frac{900}{100} = 9.$$

$$\widehat{OR} = \frac{ad}{bc} = \frac{800 \cdot 100}{200 \cdot 900} = 0, \bar{4}$$

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 (80%)	200 (20%)	1000
Nová Dláha	900 (90%)	100 (10%)	1000
Celkem	1700 (85%)	300 (15%)	2000

- Odhad šance vyléčení Moribundu v Dolní Lomné je

$$\frac{a}{b} = \frac{800}{200} = 4.$$

- Odhad šance vyléčení Moribundu v Nové Dláze je

$$\frac{c}{d} = \frac{900}{100} = 9.$$

$\widehat{OR} = \frac{ad}{bc} = \frac{800 \cdot 100}{200 \cdot 900} = 0, \bar{4} \Rightarrow$ šance na vyléčení Moribundu je v Dolní Lomné $0, \bar{4}$ krát vyšší než v Nové Dláze.

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 (80%)	200 (20%)	1000
Nová Dláha	900 (90%)	100 (10%)	1000
Celkem	1700 (85%)	300 (15%)	2000

- Odhad šance vyléčení Moribundu v Dolní Lomné je

$$\frac{a}{b} = \frac{800}{200} = 4.$$

- Odhad šance vyléčení Moribundu v Nové Dláze je

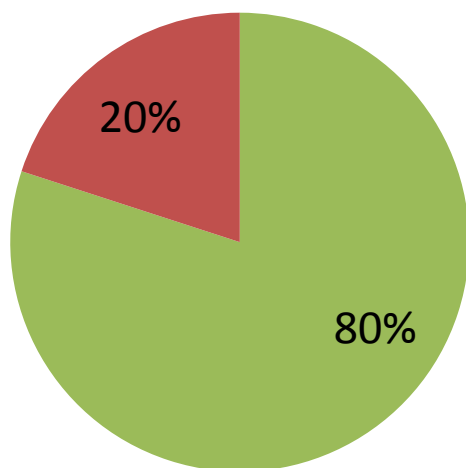
$$\frac{c}{d} = \frac{900}{100} = 9.$$

$\widehat{OR} = \frac{ad}{bc} = \frac{800 \cdot 100}{200 \cdot 900} = 0,44 \Rightarrow$ šance na vyléčení Moribundu je v Dolní Lomné 2,5 krát nižší než v Nové Dláze.

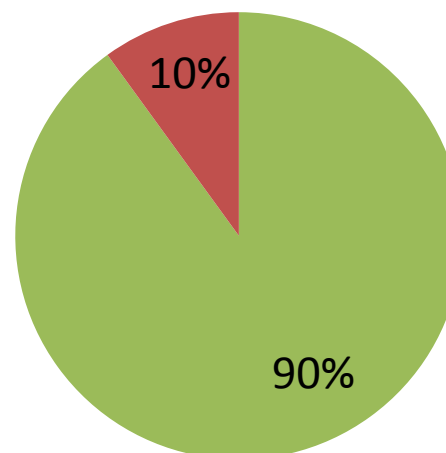
Jak graficky znázornit zjištěné výsledky?

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 (80%)	200 (20%)	1000
Nová Dláha	900 (90%)	100 (10%)	1000
Celkem	1700 (85%)	300 (15%)	2000

Dolní Lomná



Nová Dláha

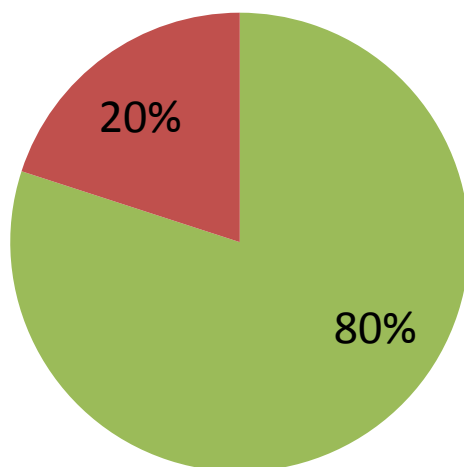


■ Vyléčen ■ Nevyléčen

Jak graficky znázornit zjištěné výsledky?

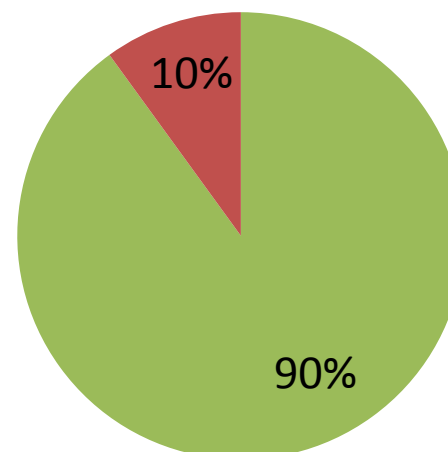
Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 (80%)	200 (20%)	1000
Nová Dláha	900 (90%)	100 (10%)	1000
Celkem	1700 (85%)	300 (15%)	2000

Dolní Lomná



Celkem: 1000 pacientů

Nová Dláha

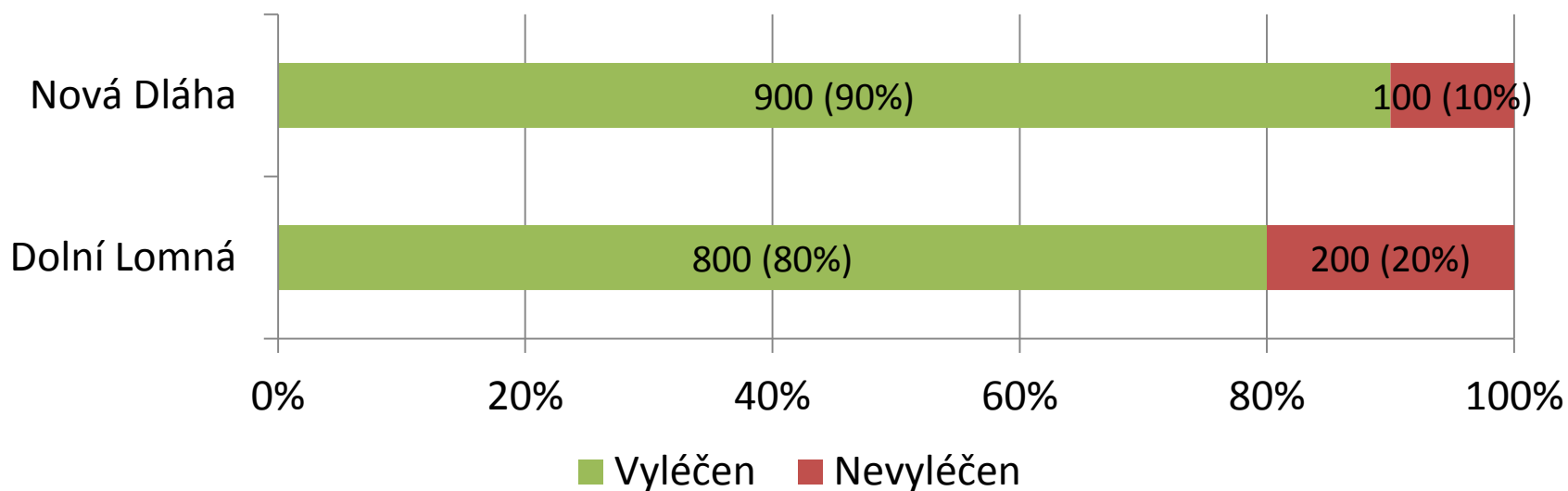


Celkem: 1000 pacientů

■ Vyléčen ■ Nevyléčen

Jak graficky znázornit zjištěné výsledky?

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 (80%)	200 (20%)	1000
Nová Dláha	900 (90%)	100 (10%)	1000
Celkem	1700 (85%)	300 (15%)	2000



100% skládaný pruhový graf

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 (80%)	200 (20%)	1000
Nová Dláha	900 (90%)	100 (10%)	1000
Celkem	1700 (85%)	300 (15%)	2000

Nemocnice v **Nové Dláze** je úspěšnější!!!



Uplynul nějaký čas a pan Úzkostný zjistil
podrobnější informace...



Dolní Lomná

Stav při přijetí/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Lehký	590 (98,3%)	10 (1,7%)	600
Těžký	210 (52,5%)	190 (47,5%)	400
Celkem	800 (80%)	200 (20%)	1000

Nová Dláha

Stav při přijetí/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Lehký	870 (96,7%)	30 (3,3%)	900
Těžký	30 (30%)	70 (70%)	100
Celkem	900 (90%)	100 (10%)	1000

Ve kterém městě je u lehkých pacientů vyšší pravděpodobnost vyléčení?

Dolní Lomná

Stav při přijetí/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Lehký	590 (98,3%)	10 (1,7%)	600
Těžký	210 (52,5%)	190 (47,5%)	400
Celkem	800 (80%)	200 (20%)	1000

Nová Dláha

Stav při přijetí/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Lehký	870 (96,7%)	30 (3,3%)	900
Těžký	30 (30%)	70 (70%)	100
Celkem	900 (90%)	100 (10%)	1000

Ve kterém městě je u těžkých pacientů vyšší pravděpodobnost vyléčení?

Dolní Lomná			
Stav při přijetí/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Lehký	590 (98,3%)	10 (1,7%)	600
Těžký	210 (52,5%)	190 (47,5%)	400
Celkem	800 (80%)	200 (20%)	1000

Nová Dláha			
Stav při přijetí/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Lehký	870 (96,7%)	30 (3,3%)	900
Těžký	30 (30%)	70 (70%)	100
Celkem	900 (90%)	100 (10%)	1000

V praxi prezentované údaje – agregovaná data

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 (80%)	200 (20%)	1000
Nová Dláha	900 (90%)	100 (10%)	1000
Celkem	1700 (85%)	300 (15%)	2000

Ve kterém městě je vyšší pravděpodobnost vyléčení?

Dolní Lomná			
Stav při přijetí/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Lehký	590 (98,3%)	10 (1,7%)	600
Těžký	210 (52,5%)	190 (47,5%)	400
Celkem	800 (80%)	200 (20%)	1000

Nová Dláha			
Stav při přijetí/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Lehký	870 (96,7%)	30 (3,3%)	900
Těžký	30 (30%)	70 (70%)	100
Celkem	900 (90%)	100 (10%)	1000

V praxi prezentované údaje – agregovaná data

Srovnání kvality nemocnic z hlediska úspěšnosti léčby Moribundu			
Nemocnice/Úspěšnost léčby	Vyléčen	Nevyléčen	Celkem
Dolní Lomná	800 (80%)	200 (20%)	1000
Nová Dláha	900 (90%)	100 (10%)	1000
Celkem	1700 (85%)	300 (15%)	2000

Simpsonův paradox

Simpsonův paradox

- Jedná se o situaci, kdy se závislost mezi dvěma znaky kvalitativně změní, jestliže uvážíme vliv znaku třetího (skrytého). **Důvodem je silná závislost mezi jedním z dvou analyzovaných znaků a znakem skrytým.**
- Příkladem může být:
 - Vztah mezi úspěšností léčby a místem léčby (Dolní Lomná vs. Nová Dláha), vezmeme-li v úvahu stav pacienta při přijetí do nemocnice. Důvodem je silná závislost mezi úspěšností léčby a stavem pacienta při přijetí.
 - Závislost procenta tělního tuku na výšce, jenž je rostoucí (čím vyšší člověk, tím více tuku), avšak odstraníme-li vliv hmotnosti, závislost se změní na klesající (při jinak shodné hmotnosti, čím vyšší člověk, tím méně tuku). Důvodem je silná korelace mezi výškou a hmotností.

Dochází k tomuto paradoxu pouze v „učebnicových“ příkladech?



Moc péče škodí. Slavná porodnice neuspěla

22.10.2011 13:58

Zdroj: Mladá fronta Dnes

Autor: Lenka Petrášová

Redakce deníku Mladá fronta Dnes provedla rozsáhlé srovnání kvality zdravotní péče v nemocnicích. Použila údaje dodané Všeobecnou zdravotní pojišťovnou. V hodnocení porodnic došlo k překvapivým zjištěním. Porodnice, které jsou u matek oblíbené, neskončily na nejvyšších místech v žebříčku.

„Pro rodičky to bude překvapení. Nejlepší porodnicí v zemi je nenápadná nemocnice v Hradci Králové či ještě nenápadnější Český Krumlov. Naopak vyhlášené porodnice v Praze či Brně zaostávají.“

[4]

[Hodnocení kvality nemocnic](#)

Další příklady výskytu Simpsonova paradoxu

Studie o souvislosti velikosti ledvinových kamenů, zvolené léčebné metody a vyléčení [2]

	Metoda 1	Metoda 2
Malé kameny	93 % (81/87)	87 % (234/270)
Velké kameny	73 % (192/263)	69 % (55/80)
Celkem	78 % (273/350)	83 % (289/350)

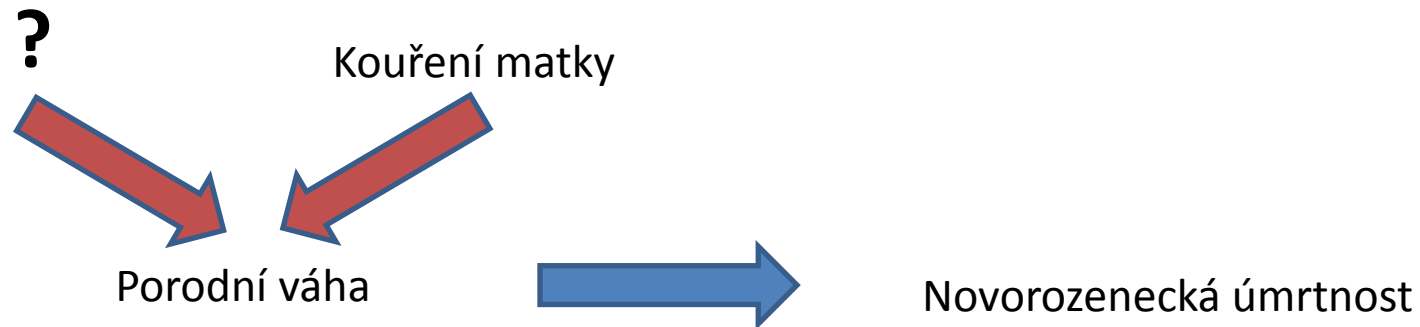
Účinnost dvou léčebných metod v závislosti na velikosti ledvinových kamenů
(zdroj: [1])

Příčina paradoxu: Účinnější metoda se častěji využívá u těžších případů.

Řešení: „Párování pacientů“.

Vliv kouření matky na novorozeneckou úmrtnost [4]

Mezi dětmi narozenými s nízkou porodní vahou je kojenecká úmrtnost nižší, pokud je matka kuřačka.



Příčina paradoxu: Kouření přispívá k nízké porodní váze.

Řešení: Komplexní analýza.

1973

Universita v Berkeley obviněna z diskriminace žen
při přijímacím řízení

Úspěšnost mužů: **43%**

X

Úspěšnost žen: **30%**

1973

Universita v Berkeley obviněna z diskriminace žen při přijímacím řízení

obor	muži			ženy		
	zájemců	přijatých	úspěšnost	zájemkyň	přijatých	úspěšnost
A	825	512	62%	108	89	82%
B	560	313	56%	25	17	68%
C	325	120	37%	593	202	34%
D	417	138	33%	375	131	35%
E	191	53	28%	393	94	24%
F	373	22	6%	341	24	7%
A-F	2691	1158	43%	1835	557	30%

[4]

Příčina paradoxu: Nízký počet zájemkyň v oborech (A a B), v nichž byla vysoká úspěšnost v přijímacím řízení, oproti tomu vysoký počet zájemkyň v oboru F...

Řešení: Neposuzovat dle agregovaných dat.

Srovnání kvality studentů na základě srovnávacího testu (SAT Reasoning Test) v USA

- Testu se účastní převážně zájemci o studium na vysoké škole

Známka	1992	2002
A+	619 (5 %)	607 (7 %)
A	575 (12 %)	565 (17 %)
A-	546 (14 %)	538 (17 %)
B	486 (52 %)	479 (47 %)
C	434 (17 %)	424 (11 %)
Celkem	500 (100 %)	504 (100 %)

Průměrný bodový zisk v testu SAT v závislosti na známce (Zdroj: [1])

Příčina paradoxu: „Inflace známek“ - na „dobrou“ známku dosáhne více studentů, než v minulosti.

Řešení: Neposuzovat dle agregovaných dat? Komplexní analýza.

Literatura

1. Balhar, J. (2011): Simpsonův paradox, bakalářská práce, Vedoucí bakalářské práce: RNDr. Arnošt Komárek, Ph.D., MFF, Karlova univerzita, Praha.
2. Charig, C.R., Webb, D.R., Payne, S.R., Wickham, O.E. (1986), Comparison o treatment of renal calculi by operative surgery, percutaneous nephrolithotomy, and extracor poreal shock wawe lithotripsy, British Medical Journal, 292, 879-88.
3. Wilcox, Allen (2006). „[The Perils of Birth Weight – A Lesson from Directed Acyclic Graphs](#)“. American Journal of Epidemiology. 164(11):1121-1123.
4. Blog: Kvantová koroptev, příspěvek: [Zločin a statistika](#)

Další zajímavé odkazy

- [1] Agresti, A. (2002). *Categorical Data Analysis, Second Edition*. Hoboken: John Wiley and Sons. ISBN 0-471-36093-7.
- [2] Blyth, C. R. (1972). On Simpson's paradox and the sure-thing principle. *Journal of the American Statistical Association*, 67, 364-366.
- [3] Davis, L. J. (1989). Intersection union tests for strictly collapsibility in three-dimensional contingency tables. *Annals of Statistics*, 17, 1693-1708.
- [4] Dong, J. (1998). Simpson's paradox. Pp. 4108-4110 in *Encyclopedia of Biostatistics*, vol. 5. Chichester: John Wiley and Sons.
- [5] Pavlides, M. G., Perlman, M. D. (2009). How likely is Simpson's paradox? *The American Statistician*, 63, 226-233.
- [6] Samuels, M. L. (1993). Simpson's paradox and related phenomena. *Journal of the American Statistical Association*, 88, 81-88.
- [7] Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Series B*, 13, 238-241.
- [8] Wagner, C. H. (1982). Simpson's paradox in real life. *The American Statistician*, 36, 46-48.
- [9] Wardrop, R. L. (1995). Simpson's paradox and the hot hand in basketball. *The American Statistician*, 49, 24-28.

Děkuji za pozornost!