

# Domácí práce z předmětu 01M6f - Statistika

Pavel Dědourek

9. května 2006

# 1 Zadání

Pomocí dat [1] chci otestovat hypotézu, že počet odbavených cestujících letecké dopravy v České republice má rovnoměrné rozdělení pro roky 2004 a 2005.

# 2 Hypotéza

$H_0$ : rozložení počtu odbavených cestujících v letecké dopravě dle ročního období má rovnoměrné rozdělení v roce 2004 (2005).

# 3 Vstupní data

měsíc	rok 2004	rok 2005
leden	398802	514573
únor	384934	469561
březen	453636	555490
duben	484168	581965
květen	533964	640323
červen	626859	700792
červenec	669813	753570
srpen	676706	749901
září	636022	709369
říjen	592168	633718
listopad	552985	545977
prosinec	538658	496595

Tabulka 1: Počet odbavených cestujících v letecké dopravě

# 4 Řešení

Pro testování hypotézy budu využívat  $\chi^2$ -test dobré shody. Vztahy převzaty z [2].

$$T := \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i}$$

$n$  - rozsah výběru, počet odbavených cestujících za rok

$n_i$  - četnosti jevu, počet odbavených cestujících v daném měsíci, ( $i = 1, \dots, k$ )

$k$  - počet hodnot, tj. počet měsíců v roce

$p_i$  - pravděpodobnosti jevu, že nastal právě v daném měsíci je  $p_i = \frac{\text{počet dnů v měsíci}}{365}$

$np_i$  - teoretické četnost

Pro  $n \rightarrow \infty$  se blíží k  $\chi^2(k-1)$  Dosažené významnost:  $1 - F_{\chi^2(k-1)}(t)$ .  
 Hypotézu  $H_0$  zamítáme pro  $t > Q_{\chi^2(k-1)}(1 - \alpha)$  tj.  $1 - F_{\chi^2(k-1)}(t) < \alpha$ .

## 5 Výpočet

### 5.1 Testování hypotézy pro rok 2004

měsíc	$n_i$	$p_i$	$mp_i$	$\frac{(n_i - np_i)^2}{np_i}$
leden	398802	0,085	54672,58	43801,77
únor	3384934	0,079	518887,25	34580,68
březen	453636	0,085	554672,58	18404,35
duben	484168	0,082	536779,92	5156,7
květen	533964	0,085	554672,58	773,15
červen	626859	0,082	536779,92	15116,51
červenec	669813	0,085	554672,58	23901,16
srpen	676706	0,085	554672,58	26848,55
září	6636022	0,082	536779,92	18348,28
říjen	592168	0,085	554672,58	2534,66
listopad	552985	0,082	536779,92	489,22
prosinec	538658	0,085	554672,58	462,38

Tabulka 2: Testování hypotézi pro rok 2004,  $T := 190417,42$

Hypotézu  $H_0$  zamítneme na hladině významnosti 95% za předpokladu, že platí  $T > \chi_{0,95}^2(k-1)$ , kde  $k = 12$ .

$190417,42 > 19,675 \Rightarrow$  Zamítáme hypotézu  $H_0$ .

### 5.2 Testování hypotézy pro rok 2005

Hypotézu  $H_0$  zamítneme na hladině významnosti 95% za předpokladu, že platí  $T > \chi_{0,95}^2(k-1)$ , kde  $k = 12$ .

měsíc	$n_i$	$p_i$	$mp_i$	$\frac{(n_i - np_i)^2}{np_i}$
leden	514573	0,085	624402,34	19318,45
únor	469561	0,077	563976,31	15806,07
březen	555490	0,085	624402,34	7605,53
duben	581965	0,082	604260,33	822,63
květen	640323	0,085	624402,34	405,94
červen	700792	0,082	604260,33	15421,11
červenec	753570	0,085	624402,34	26720,41
srpen	749901	0,085	624402,34	25223,98
září	709369	0,082	604260,33	18283,23
říjen	633718	0,085	624402,34	138,98
listopad	545977	0,082	604260,33	5621,66
prosinec	496595	0,085	624402,34	26160,56

Tabulka 3: Testování hypotézi pro rok 2005,  $T := 161528,55$

$161528,55 > 19,675 \Rightarrow$  Zamítáme hypotézu  $H_0$ .

## 6 Závěr

V dané úloze jsem vyvrátil v obou případech hypotézu, že rozložení počtu odbavených cestujících v letecké dopravě má rovnoměrné rozdělení. Je to nejspíš způsobeno tím, že v naší republice, není cestování letadlem zas tak všední záležitost. To je dáno její velikostí. Z dat je vidět, že cestování letadlem, je více méně sezóní záležitost. Během léta se počet cestujících zvýší a na zimu sníží.

## Literatura

- [1] Oficiální webové stránky Českého statistického úřadu: *Cestovní ruch - časové řady*  
[http://www.czso.cz/csu/redakce.nsf/i/cru\\_cr](http://www.czso.cz/csu/redakce.nsf/i/cru_cr)
- [2] Mirko Navara: *Matematika pro výpočetní techniku*  
[http://cmp.felk.cvut.cz/~navara/MVT/MVT\\_pr.pdf](http://cmp.felk.cvut.cz/~navara/MVT/MVT_pr.pdf)