

Marián Rimarčík

# Štatistika pre prax

**Ing. Marián Rimarčík, PhD.: Štatistika pre prax**

Jazyková korektúra: Mgr. Erika Rimarčíková

Obal: Ing. Rastislav Bajus

Vydané nákladom vlastným

Rok vydania: 2007

ISBN 978-80-969813-1-1

© 2007 Marián Rimarčík, <http://rimarcik.com>

Žiadna časť tejto knihy nesmie byť publikovaná a šírená žiadnym spôsobom a v žiadnej podobe bez písomného súhlasu autora.

**OBSAH**

<b>Úvod</b>	<b>9</b>
<b>1 Štatistika ako veda</b>	<b>11</b>
1.1 Postavenie štatistiky medzi vedami	11
1.2 Oblasti využitia štatistiky	12
1.3 Vývoj štatistiky	13
1.4 Základné štatistické pojmy	15
<b>2 Výskumné štúdie</b>	<b>21</b>
2.1 Špecifikácia problému a stanovenie cieľov výskumu	21
2.1.1 Druhy výskumných štúdií	23
2.2 Plán výskumu	28
2.2.1 Populácia a vzorka	28
2.2.2 Metódy výberu vzorky	29
2.2.3 Dotazník	35
2.3 Zber údajov	38
2.4 Analýza údajov	42
2.5 Prezentácia výsledkov	43
<b>3 Deskriptívna štatistika</b>	<b>44</b>
3.1 Jednorozmerná deskriptívna štatistika	46
3.1.1 Frekvenčná tabuľka	46
3.1.2 Koláčový graf	47
3.1.3 Stĺpcový graf	48
3.1.4 Kumulatívny stĺpcový graf	48
3.1.5 Histogram	50
3.1.6 Opisné charakteristiky	51
3.1.7 Škatuľový graf	62

3.2 Dvojrozmerná deskriptívna štatistika	64
3.2.1 Bivariačný histogram	64
3.2.2 X–Y graf	64
3.2.3 Párová regresia a korelácia	65
3.2.4 Opisná tabuľka	73
3.2.5 Eta koeficient	74
3.2.6 Priemerový graf	75
3.2.7 Viacnásobný škatuľový graf	76
3.2.8 Kontingenčná tabuľka	78
3.2.9 Rozdiel podielov, pomer podielov, pomer šancí	79
3.2.10 Senzitivita, Specificita, ROC krivka	81
3.2.11 Kontingenčný koeficient	84
3.2.12 Viacnásobný stĺpcový graf	86
3.2.13 Viacnásobný kumulatívny stĺpcový graf	86
3.2.14 3-D stĺpcový graf	87
3.3 Viacrozmerná deskriptívna štatistika	88
3.3.1 Viaccestné tabuľky a grafy	90
3.3.2 Parciálne korelácie	92
3.3.3 Viacnásobná lineárna regresia	92
3.3.4 Analýza spoľahlivosti, Cronbachova alfa	96
3.3.5 Faktorová analýza	101
3.3.6 Analýza hlavných komponentov	106
3.3.7 Zhluková analýza	107
3.3.8 Multidimenzionálne škálovanie	111
<b>4 Základy pravdepodobnosti</b>	<b>114</b>
4.1 Náhoda, náhodný jav a náhodný pokus	114
4.2 Pravdepodobnosť	115
4.3 Náhodná premenná a rozdelenie pravdepodobnosti	118
4.4 Charakteristiky náhodnej premennej	123

4.5	Teoretické rozdelenia pravdepodobnosti.....	124
4.5.1	Bernoulliho rozdelenie .....	124
4.5.2	Binomické rozdelenie.....	125
4.5.3	Záporné binomické rozdelenie .....	126
4.5.4	Geometrické rozdelenie.....	126
4.5.5	Hypergeometrické rozdelenie.....	127
4.5.6	Poissonovo rozdelenie .....	128
4.5.7	Rovnomerné rozdelenie.....	129
4.5.8	Weibullovo rozdelenie.....	130
4.5.9	Exponenciálne rozdelenie.....	130
4.5.10	Normálne rozdelenie .....	132
4.5.11	Chi-kvadrát rozdelenie .....	137
4.5.12	Studentovo <i>t</i> -rozdelenie.....	138
4.5.13	Fisherovo <i>F</i> -rozdelenie.....	139
<b>5</b>	<b>Induktívna štatistika .....</b>	<b>140</b>
5.1	Odhady parametrov .....	141
5.1.1	Bodové odhady parametrov.....	141
5.1.2	Intervalové odhady parametrov.....	145
5.2	Testovanie hypotéz.....	147
5.2.1	Postup testovania štatistických hypotéz .....	147
5.2.2	Štatistická hypotéza a výskumná hypotéza .....	152
5.2.3	Druhy chýb pri testovaní štatistických hypotéz.....	153
5.2.4	Nevýhody testovania štatistických hypotéz.....	156
5.3	Metódy induktívnej štatistiky .....	157
5.3.1	Typy vzoriek.....	157
5.3.2	Parametrické a neparametrické metódy.....	159
5.4	Jednorozmerná induktívna štatistika .....	160
5.4.1	Testy normality.....	160
5.4.2	Jednovzorkový <i>t</i> -test.....	161
5.4.3	Chi-kvadrát test dobrej zhody .....	162
5.4.4	Binomický test.....	163

5.5	Dvojrozmerná induktívna štatistika.....	163
5.5.1	Dvojvzorkový <i>t</i> -test.....	163
5.5.2	Jednocestná analýza rozptylu .....	165
5.5.3	Jednocestná analýza rozptylu pre opakované merania .....	168
5.5.4	Korelačné koeficienty.....	170
5.5.5	Chi-kvadrát test nezávislosti.....	171
5.5.6	Cochranov <i>Q</i> test .....	172
5.5.7	Stuart-Maxwellov test, McNemarov test symetrie .....	173
5.5.8	Kapa.....	175
5.5.9	Fisherov presný test.....	177
5.5.10	Liddelov presný test .....	179
5.6	Viacrozmerná induktívna štatistika .....	181
5.6.1	Viacnásobná regresia.....	181
5.6.2	Všeobecné lineárne modely.....	184
5.6.3	Kanonická korelácia .....	185
5.6.4	MANOVA, MANCOVA .....	187
5.6.5	Diskriminačná analýza .....	190
5.6.6	Logistická regresia.....	191
5.6.7	Log-lineárne modely .....	194
5.6.8	Modelovanie štruktúrálnych rovníc.....	196

<b>Literatúra.....</b>	<b>199</b>
------------------------	------------

---

## ÚVOD

Minimálne vedomosti zo štatistiky sú pre správne pochopenie množstva informácií s ktorými každý denne prichádza do kontaktu nevyhnutné. Preto sa štatistika vyučuje na takmer každej vysokej škole.

Výučba štatistiky je často zameraná na odvodzovanie a memorovanie vzorcov, prácu so štatistickými tabuľkami a mechanickým dosadzovaním do vzorcov. Dnes k tomu (našťastie) pribúdajú základy práce so štatistickým programom (alebo aspoň tabuľkovým procesorom Excel). Na to najpodstatnejšie už väčšinou neostáva čas. Najdôležitejšie je dokonalé osvojenie si štatistických pojmov a princípov, získanie uceleného prehľadu o štatistických metódach a širokých možnostiach ich praktického využitia, schopnosti výberu správnej štatistickej metódy pre konkrétnu situáciu a správnej interpretácie získaných výsledkov. Následky sa prejavujú neskôr pri diplomových, rigorózných a dizertačných prácach a samozrejme v praxi. V redakčných radách mnohých časopisov chýba kvalifikovaný štatistik, ktorý by upozornil na nesprávnu metodiku a zabránil vydaniu takéhoto článku. Neskôr čitatelia metodiku použitú v článku z ich oblasti jednoducho prevezmú bez toho, aby si jej správnosť overili v kvalitnej štatistickej knihe.

Cieľom knihy je v ucelenej podobe poskytnúť základné vedomosti zo štatistiky. Dôraz je kladený na podrobné vysvetlenie kľúčových pojmov, východísk a predpokladov, na ktorých je štatistika postavená. Nájdete tu prehľad širokej škály štatistických metód s príkladmi ich aplikácie v medicíne, psychológii, sociológii a sociálnej práci. Kniha poskytuje „prvú pomoc“ pri výbere vhodnej štatistickej metódy pre konkrétnu situáciu a správnej interpretácii výsledkov, ktoré táto metóda poskytuje. Okrem štatistických vzorcov v knihe nájdete podrobný postup na vytvorenie tabuliek, grafov a výpočtov v programe Microsoft

Excel, ktorý možno použiť aj v jeho voľne šíriteľnom príbuznom OpenOffice.org Calc. Podrobný návod len k najrozšírenejším štatistickým programom (SPSS, SAS, Statistica, Statgraphics) by zabral tisíce strán, preto ho v knihe nenájdete. Každý program sa ovláda trochu odlišne, logika je však rovnaká. Pokiaľ si nájdete, ktorú metódu máte použiť, nájsť v dokumentácii konkrétneho programu jednoduchý postup ako ju spustiť, už nie je veľký problém. Výsledky analýz a ich interpretácia sú samozrejme vo všetkých programoch rovnaké.

Kniha je rozdelená do piatich kapitol. V prvej kapitole sú odpovede na otázky, čím sa štatistika zaoberá, aký bol jej vývoj a hlavne si osvojíte kľúčové štatistické pojmy. Druhá časť knihy je venovaná výskumným štúdiám, ich druhom a jednotlivým fázam výskumu. V tretej kapitole sú prezentované metódy deskriptívnej štatistiky. Štvrtá kapitola je venovaná základom pravdepodobnosti, ktoré sú východiskom k pochopeniu indukčnej štatistiky. Záverečná kapitola sa zaoberá metódami indukčnej štatistiky: odhadmi parametrov a testami štatistických hypotéz. Princípy, východiská a jednoduché štatistické metódy sú vysvetlené podrobne, zložitejšie sú z priestorových dôvodov iba stručne načrtnuté, dôraz je kladený na praktické možnosti ich využitia.

Recenzentovi doc. Ing. Vladimírovi Gazdovi, PhD. ďakujem za starostlivé preštudovanie a podnetné pripomienky. Dúfam, že moje mnohoročné skúsenosti so štatistickými analýzami vrátane medzinárodných projektov a pedagogickej praxe, ktoré sa zmestili na 200 strán, pomôžu všetkým ľuďom so záujmom o pochopenie základov štatistiky.

*Autor*