

## **2. DOTAZNÍKOVÉ PRŮZKUMY A JEJICH ZPRACOVÁNÍ**

Zjišťování a vyhodnocování názorů, postojů a motivací obyvatelstva je podkladem pro rozhodování v sociální, ekonomické a manažerské sféře. Účelným a účinným nástrojem v tomto směru se jeví využití různých dotazníkových a anketních šetření. Tato šetření umožňují analyzovat i skutečnosti a indikátory, které se nedají či dají velmi obtížně kvantifikovat či zjišťovat jiným způsobem. Aby zmíněná šetření přinesla žádoucí informaci, ze které by bylo možné činit konkrétní závěry a rozhodnutí, je třeba věnovat pozornost jak přípravě šetření, tvorbě dotazníku i vlastnímu statistickému zpracování.

### **2.1. Příprava a provádění dotazníkových průzkumů**

Přípravu a hodnocení anketního šetření lze popsat pomocí několika etap:

1. Vytvoření projektu výzkumu
2. Definování jednotek, stanovení nutného rozsahu výběru a způsobu výběru
3. Vlastní rozpracování dotazníku
4. Ověření dotazníku, provedení pilotního průzkumu
5. Vlastní shromažďování materiálu
6. Analýza získaného materiálu a jeho zobecnění.

První etapa začíná stanovením tzv. programové otázky. Programová otázka přesně definuje cíl celého výzkumu. Jednoduše ji lze charakterizovat dotazy: Co chceme zkoumat? A proč? Velmi často se stává, že se pozornost obrací především na rychlé sestavení dotazníku a přesné definování cíle zůstává na okraji zájmu. To potom vede k tomu, že jsou zanedbány formulace problémů, které mají být objasněny a rovněž tak formulace důvodů, proč právě tyto problémy mají být šetřeny. Tím je samozřejmě znehodnocena kvalita celého řešení.

Z hlediska dalšího zpracování a zevšeobecnování výsledků je nutné věnovat náležitou pozornost druhé etapě. Zde je třeba jasně definovat způsoby výběru jednotek a stanovit rozsah výběru tak, aby byl dodržen požadavek reprezentativnosti. Výběr jednotek může být i neúmyslně zkreslen nedostatečným zhodnocením skutečnosti – například nevhodnou dobou, nevhodným místem dotazování, či podvědomým výběrem osob, které podle názoru tazatele budou ochotny odpovídat. Při písemném dotazování to může být například nedostatečně zajištěná návratnost dotazníků. Všechny tyto skutečnosti je třeba při přípravě šetření uvážit a stanovit jasný a přesný postup dotazování.

## 2.2. Stanovení rozsahu výběru

Velmi důležitá je otázka stanovení nutného rozsahu výběru. Při správném postupu obvykle není porušena reprezentativnost výběru a zejména je potom možné využít všech dostupných technik zpracování. V praxi se obvykle setkáváme se třemi odlišnými přístupy:

- a) Slepý odhad – zde se odhad provádí subjektivně na základě dosavadních zkušeností. Čím je zkušenost výzkumníka menší, tím je horší přesnost a spolehlivost takového výzkumu.
- b) Nákladový přístup - zde se vychází z kalkulace nákladů na šetření a zpracování.
- c) Statistický přístup – velikost vzorku je stanovena na základě statistických metod s ohledem na požadovanou míru přesnosti a spolehlivosti. Nutný rozsah lze vyjádřit ze

$$\text{vztahu: } n = \frac{u_{\alpha}^2 \cdot p \cdot q}{\Delta^2}, \text{ kde} \quad (2.1)$$

$\alpha$  = hladina významnosti (obvykle volíme 0,05 či 0,01)

$u_{\alpha}$  = tabelovaná hodnota normovaného normálního rozdělení

$p$  = podíl počtu respondentů znalých problematiky, přiklánějících se k variantě jedné

$q$  = podíl počtu respondentů znalých problematiky, přiklánějících se k variantě druhé

Obvykle tyto počty dopředu neznáme a v takovém případě volíme  $p = q = 0,5$

$\Delta$  = námi stanovená maximální přípustná chyba (podíl)

Pokud tedy například neznáme ani neodhadujeme dopředu podíl respondentů přiklánějících se k některé variantě odpovědi a požadujeme-li spolehlivost odhadu 95 % a připouštíme-li chybu ve velikosti 10 %, určíme velikost vzorku následujícím způsobem:

$$n = \frac{1,98^2 \cdot 0,5 \cdot 0,5}{0,1^2} = 98,01 = 99$$

V některých případech se ne vždy podaří dodržet požadavek reprezentativnosti vzorku, přesto však je žádoucí výzkum provést. V takovém případě je nutno tuto skutečnost uvést a vlastní analýzy a zejména závěry analýz tomuto faktu přizpůsobit.

## 2.3. Programová otázka a otázky zjišťovací

Vlastní příprava dotazování pak představuje převod programové otázky na otázky zjišťovací. Programová otázka je často abstraktní povahy, na níž by bylo obtížné odpovědět přímo a rovněž hodnocení odpovědí na jednu komplikovanou otázku by neumožnilo detailnější rozbor. Z toho důvodů používáme celou řadu otázek zjišťovacích. Tyto otázky jsou

formulovány konkrétněji a odpovědi na ně pak představují výchozí bázi pro vlastní analýzu hlavního problému.

Při přípravě šetření je třeba nejprve stanovit:

- 1) formu dotazování (anketa, dotazník či rozhovor)
- 2) druhy a počet použitých otázek
- 3) formu odpovědí na otázky.

Při přípravě otázek musíme vycházet z toho, komu je šetření určeno. Otázka musí být srozumitelná, aby byla dotazovaným správně pochopena a aby odpověď na ní byla lehce formulovatelná a jednoznačná. Rovněž je třeba uvážit, zda otázky nejsou příliš osobní, na které respondenti nebudou chtít upřímně odpovídat, či zda odpovědi nebudou z prestižních důvodů zkreslené. Otázky pak mohou mít různou formu. Nejčastěji se používají tyto druhy otázek:

- uzavřené - mají předem dané možnosti odpovědi, z hlediska dalších analýz jsou pak odpovědi snadno zpracovatelné
- otevřené - nechávají odpověď volně na dotazovaných. Získáme tak obvykle široké ale velmi těžko hodnotitelné spektrum názorů
- polootevřené - představují spojení uzavřených otázek a možnosti jedné volné odpovědi
- identifikační - tyto otázky jsou důležité z hlediska dalšího zpracování, kdy na jejich základě je možné třídění daného souboru dotazovaných a různé skupiny
- kontaktní a tréninkové- většina takových otázek sleduje pouze jeden cíl - navázat lepší kontakt s dotazovaným, lépe mu připomenout skutečnost, na níž chcí slyšet jeho názor. Odpovědi na tyto otázky se obvykle dále nezpracovávají
- filtrační otázky - umožní vyloučit ze souboru ty jednotky, u nichž další dotazování ztrácí význam a nemělo by smysl jejich odpovědi hodnotit.

Z hlediska využití testovacích postupů jsou nejvhodnější otázky uzavřené s menším počtem variant odpovědí. Otázky s velkým množstvím variant odpovědí se obvykle hodnotí pomocí základních charakteristik - četností, relativních četností či graficky. Otevřené otázky vhodně doplňují a dále rozvíjejí základní poznatky, neměly by však v dotazníku mít převahu. Pro provedení různých třídění a zjištění souvislostí mezi znaky, je vhodné, aby dotazník obsahoval dostatečné množství otázek identifikačních (pohlaví, věk, profese apod.). Rovněž

je důležité zamyslet se nad volbou takových otázek, aby bylo možné sledovat vzájemné vztahy a souvislosti mezi odpověďmi.

Důležité je stanovení počtu otázek. Příliš málo nebo naopak příliš velké množství otázek může ohrozit kvalitu následných analýz. V zásadě jde o to, aby byl počet otázek takový, aby bylo možné dobře analyzovat daný problém a přitom neunavoval respondenta.

Po ukončení všech předchozích etap, nastává etapa ověření dotazníku. Tato část bývá někdy opomíjena, což může později vést k znehodnocení celé práce. V této etapě by měl být znovu celý navrhovaný postup prozkoumán a diskutován – zda byly splněny všechny požadavky kvalitního výzkumu. Ověření by pak mělo proběhnout na pilotním průzkumu – to je na průzkumu malého počtu respondentů. Vlastní dotazování prověří konstrukci jednotlivých otázek: zda jsou dostatečně konkrétní, zda jim respondenti dobře rozumějí a nepotřebují k tomu další vysvětlení, zda jsou ochotni na dané téma odpovídat apod. Po tomto prověření by mělo být provedeno i statistické zpracování, aby bylo možno včas zjistit, zda je možné všechny otázky odborně zpracovat pomocí vybraných postupů a získat tak skutečně odpovídající závěry. Pokud předvýzkum neobjeví nedostatky dotazníku, je možné navázat vlastním šetřením. V případě zjištěných nedostatků na tomto malém vzorku je možné udělat jednoduše a rychle nápravu, aby vlastní šetření proběhlo správně a bylo z něj možno činit odpovídající závěry.

#### **2.4. Statistické metody vhodné pro hodnocení dotazníkových průzkumů**

Pro hodnocení jednotlivých otázek a zjištění základních názorů a postojů využíváme **základní popisné charakteristiky**, zejména vyjádření pomocí absolutních a relativních četností. Lepší představu o struktuře odpovědí má pro čtenáře určitě vyjádření relativní - v procentech. To však může být v případech malých počtů zavádějící. Jestliže uvedeme, že 66,6% respondentů nesouhlasí s daným nějakým tvrzením, máme obvykle představu o velkém množství nesouhlasících. Toto tvrzení však bude poněkud oslabeno, jestliže dále vyplyne, že ze tří dotazovaných, dva nesouhlasili. Z této informace by bylo jasné, že není možné činit z odpovědi nějaké větší závěry. Proto při korektním postupu je nutné uvádět jak vyjádření absolutní, tak relativní. Velmi vhodné a přehledné je potom grafické znázornění a to jak ve formě sloupcových grafů (tyto lze použít jak pro absolutní četnosti, tak pro relativní četnosti), tak ve formě různých diagramů, které jsou vhodné pro znázornění četností relativních.

Pro další zpracování a zjištění vzájemných vztahů a souvislostí mezi jednotlivými odpověďmi využíváme **metody analýz kvalitativních znaků**. Odpovědi jsou zpracovány ve formě kontingenčních tabulek. Jejich analýza pak umožňuje řešit dva okruhy problémů:

- a) test nezávislosti kvalitativních znaků
- b) posouzení síly závislosti.

Popis výše zmíněných metod je obsahem kapitoly 1. Analýza kontingenčních tabulek přinese detailnější informace a velmi vhodně doplní předchozí jednoduché zpracování. Tyto detailnější informace pak představují kvalitní podporu při rozhodování.

Pro hodnocení tzv. závislých pozorování, kdy sledujeme vliv určitého faktoru v čase (efekt před a po provedení určitého opatření), můžeme využít **Mc Nemarův test**.

Chceme-li posuzovat míry rizika či ohrožení, lze opět pro data uspořádaná v asociační tabulce 2 x 2 využít následující charakteristiky: relativní riziko, křížový poměr, atributivní riziko a relativní atributivní riziko.

Podrobnější hodnocení znaků uspořádaných v kontingenci tabulce umožňuje tzv. **znaménkové schéma odchylek**. Umožní určit, ve kterých kombinacích znaků se sdružené četnosti statisticky významně odlišují od hypotetických četností, konstruovaných pro předpoklad nezávislosti, tedy která kombinace znaků největší měrou ovlivňuje výsledek.

## 2.5. Příklad hodnocení dotazníku

Z dotazníku týkajícího se přípravy na přijímací zkoušky a následné úspěšnosti studia na vysoké škole, do kterého bylo zařazeno 380 respondentů, vybereme pro hodnocení odpovědi na dvě otázky s nabídkou odpovědí :

1. Jaká byla vaše příprava na přijímací zkoušku?
  - Návštěva intenzivních přípravných kurzů
  - Návštěva základních opakovacích kurzů
  - Nenavštěvoval/a jsem žádný přípravný kurz
2. Jaký byl výsledek přijímací zkoušky?
  - Výborný
  - Dobrý
  - Nevyhovující

Hodnocení odpovědí:

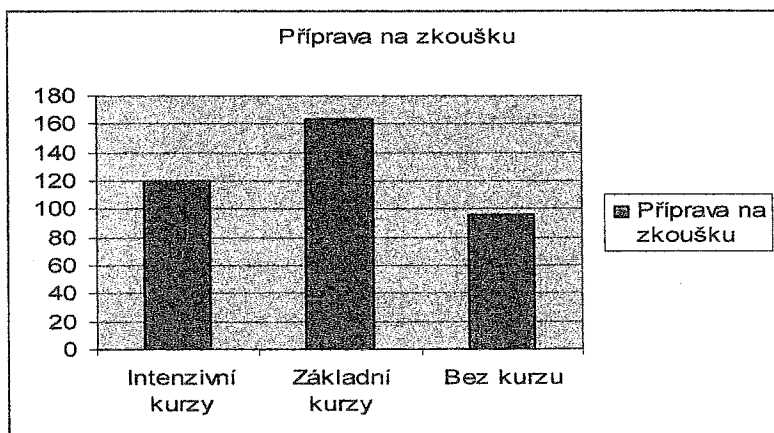
a) V první fázi provedeme zhodnocení odpovědí na jednotlivé otázky. Zjistíme počty na jednotlivé varianty odpovědí, ve vyjádření absolutním i relativním (%), pro lepší názornost využijeme i možnosti grafického znázornění:

Příprava na přijímací zkoušku:

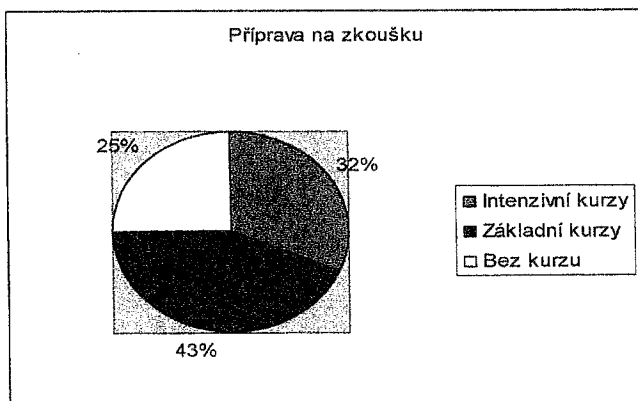
	Počet	Podíl (%)
Intenzivní kurzy	120	31,5
Základní kurzy	164	43,2
Bez kurzů	96	25,3

Grafické výstupy:

a) Sloupcový graf absolutních četností



b) Kruhový diagram relativních četností



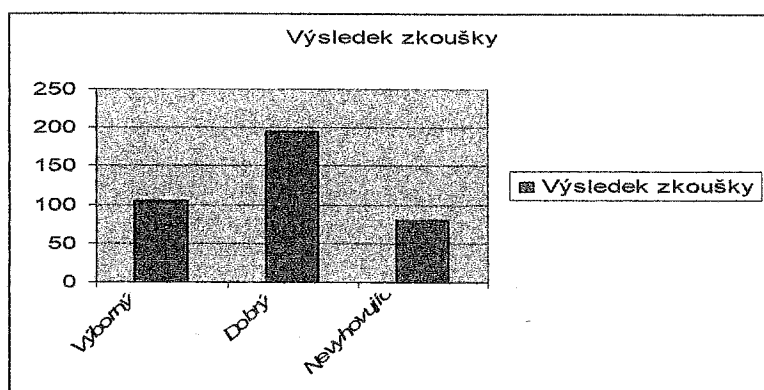
Z hodnocení je patrné, že největší část respondentů využila možnosti návštěvy základních opakovacích kurzů – 164 respondentů, to je 43,2% . 31,5 % respondentů navštěvovalo intenzivní přípravné kurzy a 25,3% respondentů nenavštěvovalo žádný kurz.

## Výsledek přijímací zkoušky

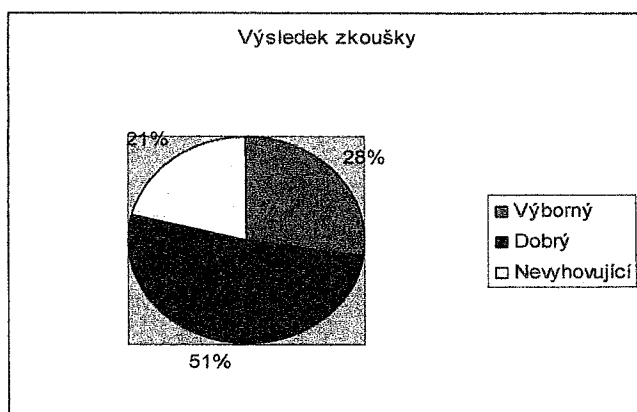
	Počet	Podíl (%)
Výborný	105	27,6
Dobry	195	51,3
Nevyhovující	80	21,1

Grafické výstupy:

a) Sloupcový graf absolutních četností



b) Kruhový diagram relativních četností



Z odpovědi na tuto otázku je patrné, že největší podíl uchazečů - 51,3% (tj. 195 respondentů) měl výsledek hodnocen jako dobrý, 27,6% respondentů mělo výsledek přijímací zkoušky výborný a 21,1 % mělo výsledek nevyhovující pro přijetí.

**b) Posouzení vztahu mezi návštěvou kurzů a výsledkem zkoušky**

Z obou jednotlivé hodnocených odpovědi však nevyplývá vztah mezi návštěvou přípravných kurzů a výsledkem přijímací zkoušky. Nabízí se možnost výkladu, že ti, co nenavštěvovali kurzy nebyli dobře připraveni a tedy jejich výsledek pak byl nevyhovující.

Můžeme však dospět i k jiné hypotéze : ti, co nenavštěvovali kurzy, byli již z předchozího studia dobře připraveni a proto návštěvu kurzů nepovažovali za potřebnou. Naopak, ti co měli velmi malé znalosti, si je v kurzech potřebovali doplnit a výsledek zkoušky by pak nemusel být dobrý. Odpovědi získáme spojením obou otázek do kontingenční tabulky. Zde pak můžeme hodnotit, zda existuje závislost mezi návštěvou kurzů a výsledkem zkoušky a v případě, že závislost existuje, můžeme určit i její sílu.

### Existuje významná závislost mezi návštěvou kurzů a výsledkem zkoušky?

Výsledek	Kurzy			Celkem
	Intenzivní	Základní	Žádný	
Výborný	29	62	14	105
Dobrý	79	85	31	195
Nevyhověl	12	17	51	80
Celkem	120	164	96	380

Provedeme  $\chi^2$  test o nezávislosti (viz kapitola 1).

#### Teoretické četnosti

Výsledek	Kurzy		
	Intenzivní	Základní	Žádný
Výborný	33,16	45,31	26,52
Dobrý	61,58	84,15	49,26
Nevyhověl	25,26	34,53	20,21

$$\chi^2 = \sum \sum \frac{(n_{ij} - n_{oj})^2}{n_{oj}} = 98,22 \quad \chi_{0,05,4}^2 = 9,488$$

Nulovou hypotézu o nezávislosti zamítáme, mezi návštěvou kurzů a výsledkem zkoušky existuje významná závislost. Z rozložení četností v kontingenční tabulce, lze pak usuzovat na to, že návštěva kurzů má pozitivní vliv na výsledek přijímací zkoušky.



### Jak silná je tato závislost?

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}} = 0,453 \qquad C_n = \frac{C}{C_{\max}} = 0,555$$

Pomocí Pearsonova normalizovaného koeficientu kontingence lze závislost charakterizovat jako středně silnou.

V hodnocení můžeme pokračovat pomocí znaménkového schématu, kde můžeme odhalit, která kombinace odpovědí se na závislosti projevuje nejsilněji.

**Znaménkové schéma** (postup popsán v kapitole 1) : Vytvoření tabulek 2 x 2, výpočet hodnoty testového kritéria, přidělení znamének.

Kritické hodnoty:  $\chi^2_{0,05(1)} = 3,84$ ,  $\chi^2_{0,01(1)} = 6,62$ ,  $\chi^2_{0,001(1)} = 10,83$

29	76		62	43		14	91
91	184		102	173		82	193
$\chi^2=12,0$	---		$\chi^2=42,7$	---		$\chi^2=3,5$	
79	116		85	110		31	164
41	144		79	106		65	120
$\chi^2=43,7$	+++		$\chi^2=24,1$	+++		$\chi^2=5,3$	-
12	68		17	63		51	29
108	192		147	153		45	255
$\chi^2=2,7$			$\chi^2=3,0$			$\chi^2=98,2$	+++

Výsledek:

Výsledek	Kurzy		
	Intenzivní	Základní	Žádný
Výborný	---	---	
Dobrý	+++	+++	-
Nevyhověl			+++

Při návštěvě intenzivních i základních kurzů lze nejpravděpodobněji očekávat výsledek dobrý, při přípravě bez kurzů je nejpravděpodobnější výsledek nevyhověl.

### Kolikrát větší riziko neúspěchu má ten, kdo nenavštěvoval žádný kurz?

Pro odpověď na otázku, kolikrát větší riziko neúspěchu má ten, kdo žádné Kurzy nenavštěvoval, můžeme využít charakteristiku relativní riziko :

	Návštěva kurzů	Bez kurzů
Vyhověl	255	45
Nevyhověl	29	51

$$RR = \frac{a(c+d)}{c(a+b)} = \frac{255 \cdot (29 + 51)}{29 \cdot (255 + 45)} = 8,54$$

U toho kdo nenavštěvoval žádný kurz, je riziko neúspěchu 8,5-krát vyšší než toho, kdo nějaký kurz navštěvoval.

Otázku můžeme také formulovat jiným způsobem: **O kolik se zvětší pravděpodobnost neúspěchu u toho, kdo kurzy nenavštěvoval?** V takovém případě využijeme charakteristiku atributivní riziko či relativní atributivní riziko:

$$AR = \frac{a}{a+b} - \frac{c}{c+d} = 0,487$$

$$AF = \frac{\frac{a}{a+b} - \frac{c}{c+d}}{\frac{a}{a+b}} \cdot 100 = 57,3 \%$$

U uchazeče, který nenavštěvoval přípravné kurzy se oproti uchazečům navštěvujícím přípravné kurzy pravděpodobnost neúspěchu zvýší o 0,48, tj. o 57,3%.