

X7: Četnost pozorování jevu, kvetení rostlin.
 Faktor5: typ ošetření, sekaná louka – 1, kontrolní plocha – 2.
 Faktor6: kvetení, druh vykvetl – 1, druh nevykvetl – 2.

X7	Faktor5	Faktor6
13	1	1
2	1	2
6	2	1
9	2	2

Příklad 19: Frekvenční analýza – čtyřpolní tabulky.

Použité proměnné: X7, Faktor5, Faktor6

Máme pokus, kdy sledujeme kvetení určitého druhu rostliny na dvou typech stanovišť – sekané louce a louce nesekané (míněno jako kontrola). Zajímá nás, jestli sekání má pozitivní vliv na reprodukci daného druhu. Máme vždy 15 ploch od každého ošetření a sledujeme, jestli druh během sezóny vykvetne (alespoň jedna kvetoucí rostlina na dané ploše) nebo ne.

Takto strukturovaný pokus, kdy máme dané četnosti nějakého jevu v závislosti na působení dvou či více faktorů, se bude řešit metodami frekvenční analýzy. A protože jsou v tomto případě analyzovány pouze dva faktory (ošetření ploch, kvetení), které navíc nabývají jen dvou hladin, jedná se o nejjednodušší typ frekvenčních analýz, který odpovídá tzv. čtyřpolní tabulce. Tradičním testem spojeným s řešením frekvenčních analýz je *Chi-kvadrát* test nebo *Fischerův exaktní test* (tento je k dispozici pouze v případě čtyřpolních tabulek). Předpokladem a zároveň nulovou hypotézou testu je nezávislost působení obou faktorů (jevů). V našem konkrétním případě nulová hypotéza konstatuje, že kvetení rostlin se v důsledku vysekávání nezmění.

Zadání dat do tabulky je analogické analýze variance, do jednoho sloupceku zadáme zjištěné četnosti, v dalších sloupečcích kódujeme hladiny příslušných faktorů.

Frekvenční analýzy jsou v NCSS zahrnuty pod **Analysis/DescriptiveStatistics/Cross Tabulation**. V zadávacím okně (zelená kostka) je třeba definovat proměnnou s analyzovanými četnostmi (*Variables/Frequency Variable*) a *Column-By Variables* respektive *Row-By Variables* (většinou jsou jako *Discrete Variable*). V našem případě jsou četnosti v proměnné X7 a kódujícími proměnnými jsou Faktor5 a Faktor6.

Předdefinovaná nabídka výsledků zahrnuje pouze vlastní přepis čtyřpolní tabulky (*Count Section*) a výsledek Fischerova exaktního testu. Podrobnější výsledky je třeba specifikovat (příslušné změny se provedou v oddíle *Which Reports/Plots* před vlastním výpočtem).

Cross Tabulation Report

Čtyřpolní tabulka četností i s příslušnými marginálními součty je uvedena v oddíle *Count Section*.

Counts Section

	C2		
C3	1	2	Total
1	13	6	19
2	2	9	11
Total	15	15	30

The number of rows with at least one missing value is 0

Pokud chceme mít ve výsledcích uvedeny také očekávané četnosti (za předpokladu nezávislosti působení obou jevů), je třeba *Show Expected Values* změnit z *Omit* na *Report Only*.

Expected Counts Assuming Independence Section

	C2		
C3	1	2	Total
1	9,5	9,5	19,0
2	5,5	5,5	11,0
Total	15,0	15,0	30,0

The number of rows with at least one missing value is 0

Výpočet vlastního Chi-kvadrát testu musíme opět vyžádat (*Chi-Square Stats* změnit z *None* na *Chi-square Only*, *All Statistics* pak spočítá i celou řadu dalších statistik). Ve výsledcích je uvedena celková hodnota Chi-kvadrát statistiky, příslušný počet stupňů volnosti a dosažená hladina pravděpodobnosti s rozhodnutím o platnosti nulové hypotézy. V našem případě je dosažená hladina významnosti $p = 0.008$ a lze tedy zamítnout nulovou hypotézu o nezávislosti působení obou jevů. Konstatujeme tedy, že sekání louky pozitivně ovlivňuje kvetení sledovaného druhu rostliny. Je třeba ještě uvést, že použití Chi-kvadrát testu pro tento typ příkladů je sice tradiční metodou, ale není zcela oprávněné. Chi-kvadrát je totiž rozdělení spojité, zatímco frekvenční tabulky zahrnují diskrétní data. Tato skutečnost se negativně nejvíce projeví, jsou-li četnosti nízké (proto také upozornění na četnosti <5 v posledním řádku). Analogicky je možné použít tzv. *G-test* (log-likelihood ratio), který má také přibližně Chi-kvadrát rozdělení.

Chi-Square Statistics Section

Chi-Square	7,033493	
Degrees of Freedom	1,000000	
Probability Level	0,008000	Reject Ho

WARNING: At less one cell had a value less than 5.

Speciálně pro případ čtyřpolní tabulky je možné použít Fischerův exaktní test, který spočítá konkrétní pravděpodobnost, že za daných hodnot marginálních součtů dostaneme právě takové četnosti v jednotlivých polích tabulky. Tento test je zvláště vhodný, jsou-li pozorované četnosti nízké a tedy Chi-kvadrát test není příliš přesný. Fisherův exaktní test je v NCSS u výpočtu čtyřpolní tabulky nabízen jako defaultová možnost.

Vlastní test je založen na porovnání proporcí $P1$ a $P2$, které odpovídají danému pozorování a všem ostatním "horším" pozorováním. Většinou nás zajímá oboustranný test, tedy "horší" pozorování nad i pod naším pozorováním (v našem příkladě oboustranný test odpovídá jak pozitivnímu tak i negativnímu působení kosení na kvetení). Výsledek ukazuje, že oboustranný test je průkazný a lze tedy zamítnout nulovou hypotézu, že kosení louky nemá na kvetení rostlin vliv.

Fisher's Exact Test Section

	P1	P2
Proportions	0,866667	0,400000
Difference ($D0 = P1 - P2$)	0,466667	
Correlation Coefficient	0,484200	

Hypothesis	Prob Level	Test Type	Calculation Method
Ho: $P1 = P2$			$D = P1 - P2$ for a table
Ha: $P1 < P2$	0,999150	One-Tailed	Sum of prob's of tables where $D \leq D0$
Ha: $P1 > P2$	0,010470	One-Tailed	Sum of prob's of tables where $D \geq D0$
Ha: $P1 <> P2$	0,020940	Two-Tailed	Sum of prob's of tables where $ D \geq D0 $