**Kontingenční tabulka s malými četnostmi – řešení.**

**Datový soubor MOTÝLI**

Provedli jsme tento experiment: 20 samečků a 20 samiček jistého druhu motýla bylo jednotlivě vpouštěno do testovacího boxu s bílým a červeným terčem. Zaznamenali jsme, kterou barvu si jedinec vybral k prvnímu usednutí. Zajímá nás, jestli je výběr barvy nějak ovlivněn pohlavím motýla. Experiment dopadl takto: samečci: 8 bílý terč 12 červený terč

Statistiky 🡪 Základní statistiky 🡪 Kontingenční tabulky.

Specifikace tabulky: *pohlaví* a *terč*; proměnnou *počet* zadávám přes tlačítko váhy (vpravo dole).

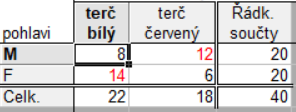
samičky: 14 bílý terč 6 červený terč.

V řádkovém zápisu by data vypadala takto:

**Číslo jedince pohlaví terč**

01 M červený

02 M červený

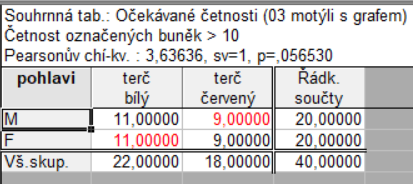
 03 F bílý

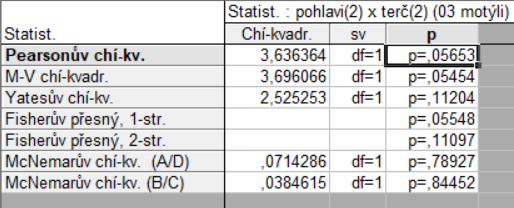
04 F červený

05 M bílý atd.

Jedná se tedy o kategoriální data, hodnotíme kontingenční tabulkou.

**Nulová hypotéza:** pravděpodobnost výběru barvy terče nezávisí na pohlaví, je to **test nezávislosti**: . Nezávislost prakticky znamená, že pravděpodobnost kombinace např. „bílé barvy vybrané samečkem“ spočtu jako pravděpodobnost výběru bílé barvy (bez ohledu na pohlaví) krát pravděpodobnost, že jedinec je sameček (v tomto příkladu 0.5).

**Podmínka:** všechny očekávané četnosti jsou alespoň 5, *oij* ≥ 5. Je splněna. Tabulka očekávaných četností:

**Výsledek chí-kvadrát testu:** těsně nezamítám hypotézu o nezávislosti výběru barvy a příslušnosti k pohlaví. Chí-kvadrát statistika = 3,64, jeden stupeň volnosti, p-hodnota = 0,0565.

Vzhledem k malým četnostem bychom mohli použít i statistiku s Yatesovou opravou na spojitost, která vyšla 2,53 s p-hodnotou = 0,112. Podobně vyšel i Fisherův přesný oboustranný test: p = 0,111.

Zde dokonce můžeme uvažovat takto: chí-kvadrát test bez Yatesovy opravy na spojitost má poloviční p-hodnotu (0,0565) proti chí-kvadrát testu s  opravou (p = 0,112). Protože Fisherův test spočítal pravděpodobnost p = 0,111 přímou metodou, bude věrohodnější chí-kvadrát test s Yatesovou opravou, jehož p-hodnota je srovnatelná s Fisherovým přímým testem.

**Odhad pravděpodobnosti, že bude vybrán červený terč**, za předpokladu, že není rozdíl mezi preferencí samiček a samečků (tedy bez ohledu na pohlaví): červený terč vybralo (12+6) jedinců ze 40 testovaných, proto P = 0,45. Podobně bílý terč vybralo 22 ze 40 jedinců, tedy P = 0,55.

Mohu vyjít i z tabulky očekávaných četností, kde je zapracován právě předpoklad nezávislosti, tedy že není rozdíl mezi preferencemi samiček a samečků (jinými slovy: výběr barvy je nezávislý na pohlaví).