# Test pro podíl u dvou výběrů

**Příklad 1:** Máme 60 studentů matematické biologie a mezi nimi 17 s modrýma očima, přičemž 11 z nich je současných a 6 je již vystudovaných. Celkový počet současných studentů matematické biologie je 42 a bývalých 18. Testujeme tedy $H\_{0}: π\_{1}=π\_{2}=π$.

Postup: Použijeme kalkulátor <http://vassarstats.net/propdiff_ind.html>, do něhož zadáme příslušné hodnoty (ka=11, na=42, kb=6, nb=18) a stiskneme tlačítko „Calculate“ (viz. obrázek níže). Získáme testovou statistiku Z=-0,563, která je v absolutní hodnotě menší než 1,96, tedy nezamítáme $H\_{0}$. Nezamítnutí nulové hypotézy lze odvodit i z p-hodnoty, která je rovna 0,573 a není tedy menší než α=0,05.



# Test pro podíl u jednoho výběru

**Příklad 2:** Máme 60 studentů matematické biologie, z nichž 17 má modré oči. Chceme testovat na hladině významnosti α=0,05, zda je podíl modrookých studentů mezi všemi matematickými biology (tedy $π={17}/{60}=0,283$) roven dané hodnotě π0=0,40. Nulová hypotéza je tedy $H\_{0}: π=0,04$.

Postup: Použijeme kalkulátor <http://vassarstats.net/propdiff_ind.html>, do něhož zadáme příslušné hodnoty pro „Sample A“ (ka=17, na=60). Protože π0=0,40 je chápáno jako populační hodnota, zvolíme co největší nb (např. 1 000 000) a kb tedy musí být 400 000. Následně stiskneme tlačítko „Calculate“ (viz. obrázek níže). Získáme testovou statistiku Z=-1,845, která je v absolutní hodnotě menší než 1,96, tedy nezamítáme $H\_{0}$. Nezamítnutí nulové hypotézy lze odvodit i z p-hodnoty, která je rovna 0,065 a není tedy menší než α=0,05.



# Relativní riziko (relative risk) a poměr šancí (odds ratio)

**Příklad 3:** Sledujeme souvislost pohlaví (pohlavi\_rek) a kardiovaskulárního onemocnění (kv\_nemoc).

Postup: Analyze – Descriptive Statistics – Crosstabs... – na záložce Statistics zatrhnout Risk

Interpretace: Relativní riziko vzniku kardiovaskulárního onemocnění je 5,3-krát vyšší u mužů než u žen. Šance na vznik kardiovaskulárního onemocnění je 6,2-krát vyšší u mužů než u žen.

IS pro poměr šancí neobsahuje jedničku → zamítáme nulovou hypotézu o nezávislosti vzniku kardiovaskulárního onemocnění na pohlaví.

**Úkol 1.** Sledujeme souvislost pohlaví (pohlavi\_rek) a vysokého HDL cholesterolu (hdl\_chol\_norma\_II; 1..v normě, 2..vysoký cholesterol). Vypočtěte relativní riziko a poměr šancí výskytu vysokého HDL cholesterolu u mužů a žen. Vzhledem k tomu, že u třech lidí chybí údaj o vysokém HDL cholesterolu (hodnota 99), tyto tři lidi nejdříve odfiltrujte (pomocí Data – Select Cases).

Řešení: Relativní riziko výskytu vysokého HDL cholesterolu je 1,12-krát vyšší u mužů než u žen. Šance na výskyt vysokého HDL cholesterolu je 1,18-krát vyšší u mužů než u žen.

# Korelace

**Příklad 4.** Chceme hodnotit vztah systolického tlaku a věku u mužů.

Vykreslení tečkového grafu (scatter plot): Graphs – Legacy Dialogs – Scatter/Dot – Simple Scatter – Define

Výpočet Pearsonova a Spearmanova korelačního koeficientu: Analyze – Correlate – Bivariate

**Úkol 2.** Ověřte, zda je při hodnocení vztahu systolického tlaku a věku u mužů použití Pearsonova korelačního koeficientu vhodné.

**Úkol 3.** Zhodnoťte vztah systolického tlaku (sys\_tlak) a celkového cholesterolu (cel\_cholesterol) u aktivních kuřáků (koureni=0). Zamyslete se, zda je vhodnější použít Pearsonův nebo Spearmanův korelační koeficient.

Řešení: Pearsonův korelační koeficient = 0,034; Spearmanův korelační koeficient = 0,025.