

10. REGRESE - POKRAČOVÁNÍ

10.1 Binární logistická regrese

Závisle proměnná je binární, nabývá jen dvou hodnot 0 a 1. Následně se používá stejně jako lineární regrese, jen místo příkazu REGRESS použijeme LOGIT. Samotné koeficienty se interpretují příliš složitě, lepší je interpretovat poměry šancí (odds ratio). Ty dostaneme pomocí parametru OR nebo použitím příkazu LOGISTIC namísto příkazu LOGIT. Poměr šancí pak slouží jako multiplikační konstanta (kolikrát vyšší šanci máme, že nastane vysvětlovaný jev, tedy že závisle proměnná nabývá hodnoty 1).

Výsledné poměry šancí pak interpretujeme buď jako zvýšení/snížení šance, nebo jako procentní nárůst. Např. poměr šancí s hodnotou 1,02 znamená, že pokud se nezávisle proměnná zvýší o jednotku (např. jeden rok), zvýší se pravděpodobnost, že jev nastane přesně 1,02×, neboli o 2 %.

```
logit relig_bin age gender
logit relig_bin age gender, or
logistic relig_bin age gender
```

10.2 Ordinální logistická regrese

Závisle proměnná je ordinální, nabývá několika kategorizovaných hodnot, které lze seřadit, a jsou od sebe stejně vzdáleny. Následně se používá stejně jako lineární regrese, jen místo příkazu REGRESS použijeme OLOGIT. Opět interpretujeme poměry šancí, které dostaneme pomocí parametru OR.

```
ologit relig_trich age gender, or
```

10.3 Multinomická logistická regrese

Závisle proměnná je ordinální, nabývá několika kategorizovaných hodnot, které lze seřadit, ale nejsou od sebe stejně vzdáleny. Následně se používá stejně jako lineární regrese, jen místo příkazu REGRESS použijeme MLOGIT. Opět interpretujeme poměry šancí (v tomto případě se ovšem statisticky korektně označují jako relative risk ratio), které dostaneme pomocí parametru RRR.

```
mlogit relig_trich age gender, rrr
```

10.4 Kvalita logistických modelů

Kvalitu modelu nemůžeme posuzovat pomocí ukazatele R^2 . Ukazatel absolutní kvality modelu neexistuje, vždy je potřeba porovnávat několik modelů mezi sebou a zjišťovat, jestli přidáním dalších vysvětlujících proměnných dochází ke zlepšení kvality modelu. K vypsání většiny ukazatelů slouží příkaz FITSTAT (bez parametru).

Důležité parametry kvality modelu jsou AIC a BIC, které se spolu s počtem případů N prezentuje spolu s modelem.