

Podobná situace nastává u soudu, kde roli nulové hypotézy hraje presumpce nevinu obžalovaného. Soudce na základě předložených důkazů zamítne jeho nevinu a odsoudí ho k trestu nebo naopak nezamítne jeho nevinu a neodsoudí ho, čímž však nijak netvrdí, že obžalovaný je skutečně nevinen. Buď je nevinen, nebo k prokázání jeho viny nemá soudce dostatek důkazů.

Stejně ve statistice, jestliže nulovou hypotézu nezamítneme, neznamena to ještě, že  $H_0$  skutečně platí. Buď je pravdivá, nebo pro její zamítnutí nemáme dostatek potřebných měření, dostatek informací.

Při rozhodování se můžeme s jistou pravděpodobností dopustit chyby.

Uvažujme čtyři možnosti:

1. hypotéza  $H_0$  platí, hypotézu  $H_0$  zamítneme (chyba 1. druhu),
2. hypotéza  $H_0$  platí, hypotézu  $H_0$  nezamítneme,
3. hypotéza  $H_0$  neplatí, hypotézu  $H_0$  zamítneme,
4. hypotéza  $H_0$  neplatí, hypotézu  $H_0$  nezamítneme (chyba 2. druhu).

Možnosti 2 a 3 jsou v pořádku. K omylu dochází v případě 1, kdy zamítneme

hypotézu, která platí (při soudním líčení tato situace odpovídá odsouzení nevinného). Tomuto závažnému omylu říkáme chyba 1. druhu a chceme ho mít pod dostatečnou kontrolou. Požadujeme, aby pravděpodobnost chyby 1. druhu nepřekročila námi předem zvolenou mez  $\alpha$ , tzv. hladinu testu, volíme zpravidla  $\alpha = 0,05$  nebo  $0,01$ . Příklad 4 nazýváme chybou 2. druhu (odpovídá neodsouzení viníka) a snažíme se ji minimalizovat. Obě chyby jsou vzájemně nepřímo úměrné. Jestliže  $H_0$  platí (tedy), pravděpodobnost zamítnutí  $H_0$  má být menší než  $\alpha$

Poté, co zformulujeme nulovou hypotézu a nasbíráme data, spočteme pravděpodobnost, s jakou bychom mohli obdržet pozorovaná data nebo data stejné, či ještě více odporující nulové hypotéze, za předpokladu, že je nulová hypotéza pravdivá. Tato pravděpodobnost se nazývá dosažená hladina významnosti a značí se  $p$  ( $p$ -value,  $p$ -level). Čím menší je  $p$ , tím neudržitelnější čili méně důvěryhodná je nulová hypotéza.