

## WILCOXONŮV ZNAMÉNKOVÝ TEST

*Jedná se o neparametrickou alternativu jednovýběrového t-testu. Použijeme jej v případě, kdy jsme nesplnili předpoklad o normálním rozdělení proměnné, jejíž průměr jsme chtěli testovat vůči konstantě pomocí jednovýběrového t-testu.*

Wilcoxonův znaménkový test lze použít pro jeden výběr, nebo například pro párovaná data (rozdíl v hmotnosti před a po dietě). Jedná se o neparametrický test a ve své nulové hypotéze testuje, zda daný výběr pochází z rozdělení symetrického kolem zadaného teoretického mediánu.

Hlavní výhodou tohoto testu je, že nemá téměř žádné předpoklady. Jedinými předpoklady jsou, že testovaná proměnná musí být číselná (spojitá nebo diskrétní) a pozorování jsou navzájem nezávislá.

### NULOVÁ A ALTERNATIVNÍ HYPOTÉZA

Jak u testů parametrických, tak i u testů neparametrických je možnost zvolit si hned z několika možných formulací nulové a alternativní hypotézy.

#### LEVOSTRANNÁ HYPOTÉZA

---

$$H_0: \tilde{x} \geq \tilde{x}_0$$

Populační medián  $\tilde{x}$  (daný výběrem) je vyšší nebo rovný hypotetickému mediánu  $\tilde{x}_0$  (zadaná konstanta).

$$H_1: \tilde{x} < \tilde{x}_0$$

Populační medián  $\tilde{x}$  (daný výběrem) je menší než hypotetický medián  $\tilde{x}_0$  (zadaná konstanta).

#### PRAVOSTRANNÁ HYPOTÉZA

---

$$H_0: \tilde{x} \leq \tilde{x}_0$$

Populační medián  $\tilde{x}$  (daný výběrem) je nižší nebo rovný hypotetickému mediánu  $\tilde{x}_0$  (zadaná konstanta).

$$H_1: \tilde{x} > \tilde{x}_0$$

Populační medián  $\tilde{x}$  (daný výběrem) je větší než hypotetický medián  $\tilde{x}_0$  (zadaná konstanta).

## OBOUSTRANNÁ HYPOTÉZA

---

$$H_0: \tilde{x} = \tilde{x}_0$$

Populační medián  $\tilde{x}$  (daný výběrem) je rovný hypotetickému mediánu  $\tilde{x}_0$  (zadaná konstanta).

$$H_1: \tilde{x} \neq \tilde{x}_0$$

Populační medián  $\tilde{x}$  (daný výběrem) není rovný hypotetickému mediánu  $\tilde{x}_0$  (zadaná konstanta).

### PŘÍKLAD

#### DATOVÝ SOUBOR: NEONATAL\_WEIGHT.SAV

---

Datový soubor obsahuje údaje o porodní váze cca 1800 dětí. Kromě porodní váhy jsou k dispozici i různé údaje o matce dítěte (věk, počet těhotenství, délka těhotenství, pohlaví dítěte atd.).

#### ZADÁNÍ

---

V tomto příkladu budeme pracovat s proměnnou *Day of delivery*. Ta obsahuje informaci o tom, jak dlouho těhotenství trvalo. Za normu délky trvání těhotenství lze obecně považovat 37 týdnů, tedy 259 dní. Pomocí vybraného statistického testu budeme chtít ověřit, zda v daném výběrovém souboru je tato norma splněna či nikoli. V případě, že by odchylka byla významná, mohlo by to znamenat, že výběrový soubor není z hlediska délky těhotenství reprezentativní. Pro otestování dané hypotézy využijeme Wilcoxonův znaménkový test.

### NULOVÁ A ALTERNATIVNÍ HYPOTÉZA

$$H_0: \tilde{x} = 259$$

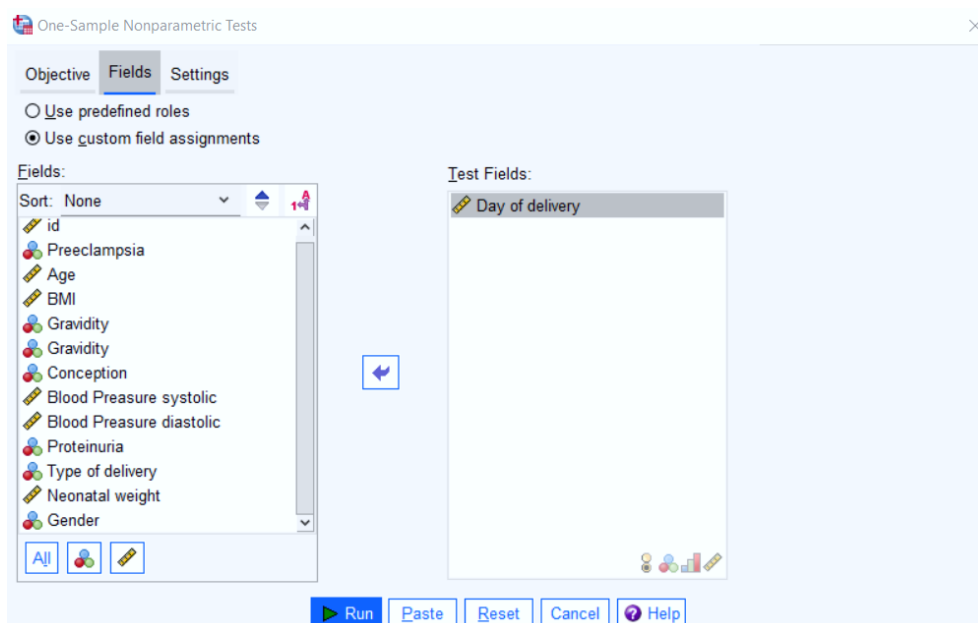
Medián proměnné *Day of delivery* je 257 dní.

$$H_1: \tilde{x} \neq 259$$

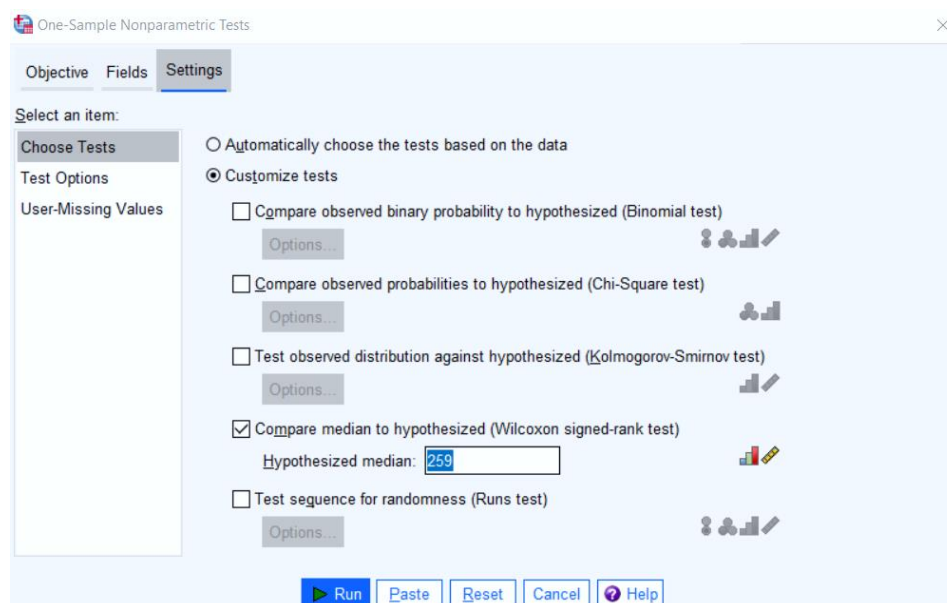
Medián proměnné *Day of delivery* není 257 dní.

### VÝPOČET V SPSS

Klikneme na *Analyze- Nonparametric Tests – One-Sample*. Otevře se nové dialogové okno, které obsahuje tři záložky. Proklikneme se do záložky *Fields*, kde zadáme, kterou proměnnou chceme testovat. Do pole *Test Fields* tedy vložíme proměnnou *Day of delivery*.



Dále klikneme na záložku *Settings*, která slouží k výběru konkrétního testu. Abychom mohli zvolit některý z možných testů v tomto okně, musíme v prvním kroku změnit volbu *Automatically choose the tests based on the data* na *Customize tests*. Následně klikneme na možnost *Wilcoxon signed-rank test* a do okna *Hypothesized median* zadáme hodnotu 259. Následně klikneme na *Run*.



Výsledek testu najdeme v tabulce *Hypothesis Test Summary*. Kromě sloupce *Sig.* zde najdeme také zopakování nulové hypotézy testu a jeho název. Ve sloupci *Decision* dále máme již interpretovanou hodnotu signifikance na 5 % hladině významnosti.

### Hypothesis Test Summary

	Null Hypothesis	Test	Sig. <sup>a,b</sup>	Decision
1	The median of Day of delivery equals 259.	One-Sample Wilcoxon Signed Rank Test	,000	Reject the null hypothesis.

a. The significance level is ,050.

b. Asymptotic significance is displayed.

Na 5 % hladině významnosti se zamítá nulová hypotéza o shodě výběrového a hypotetického mediánu. To znamená, že délka těhotenství žen v daném výběrovém souboru nesplňuje zadanou normu. Z následujícího histogramu dokážeme dále odečítat, že výběrový medián délky těhotenství je delší než hypotetický. Konkrétně je jeho hodnota 274 dní a jedná se o statisticky významnou odchylku.

