

DVOUVÝBĚROVÝ T TEST

Dvouvýběrový t-test se používá ke srovnání průměrů dvou nezávislých skupin.

Dvouvýběrový t-test se používá například v následujících případech:

Porovnání průměrného výkonu dvou skupin studentů v testech z různých předmětů.

Porovnání průměrného počtu prodaných výrobků dvou skupin obchodníků.

Porovnání průměrného času, který dvě skupiny lidí tráví cvičením denně.

Dvouvýběrový t-test se používá v případech, kdy máme k dispozici dvě nezávislé skupiny. Používá se například jako alternativa k párovému t-testu, který se používá k porovnání průměrů pro dvě skupiny, které jsou párovány (například pacienti před a po ošetření).

PŘEDPOKLADY

1. Nezávislost pozorování a skupin
2. Normální rozdělení – data musí být normálně rozdělena v obou skupinách.
3. Shoda rozptylů ve skupinách

HYPOTÉZY

Obě pozorované veličiny pochází z rozdělení se stejnou střední hodnotou.

$$H_0: \mu_1 = \mu_2$$

Střední hodnoty rozdělení pozorovaných veličin se liší.

$$H_1: \mu_1 \neq \mu_2$$

Nulová a alternativní hypotéza jsou stejné jako u párového t-testu. Hlavní rozdíl těchto dvou testů spočívá v předpokladu o nezávislosti skupin.

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 prvním kroku analýzy nás bude zajímat, zda je průměrná porodní váha stejná pro chlapce a dívky. K tomu využijeme dvouvýběrový t-test, protože testujeme shodu středních hodnot ve dvou skupinách.

Nulová a alternativní hypotéza tohoto testu bude vypadat pro tento příklad následovně:

$$H_0: \mu_{chlapci} = \mu_{dívky}$$

Průměrná porodní váha chlapců a dívek je stejná.

$$H_1: \mu_{chlapci} \neq \mu_{dívky}$$

Průměrná porodní váha chlapců a dívek je odlišná.

VÝPOČET V SPSS

V prvním kroku je nutné ověřit předpoklady testu.

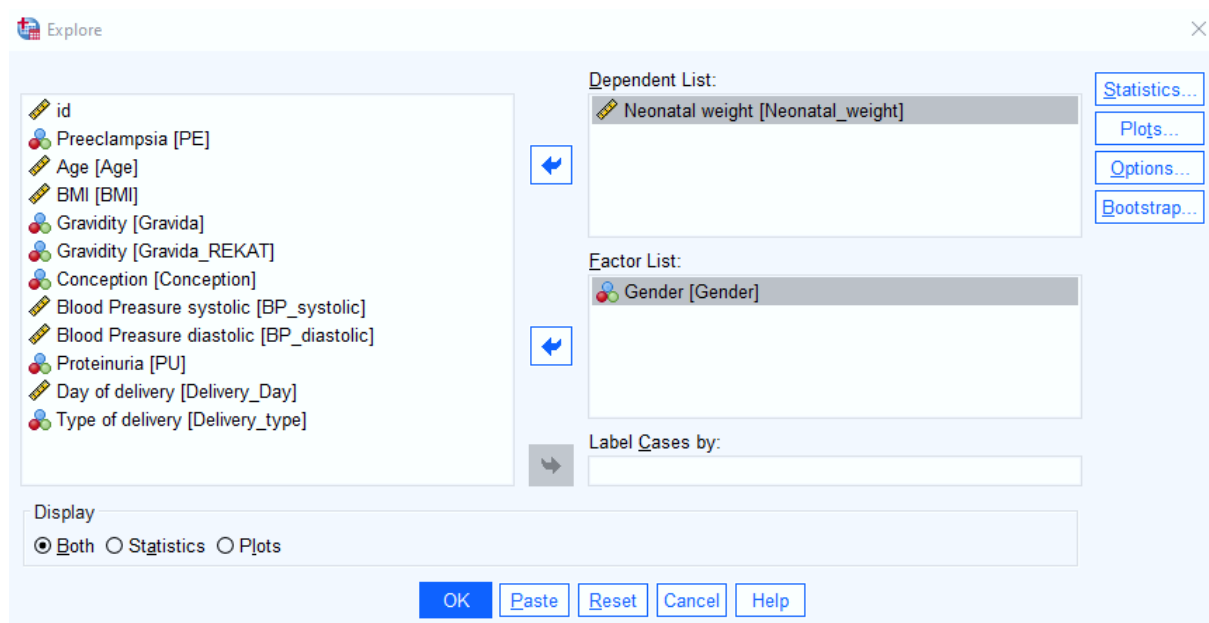
PŘEDPOKLAD NEZÁVISLOSTI POZOROVÁNÍ A SKUPIN

Jednotlivá pozorování musí být vzájemně nezávislá a zároveň skupiny musí být nezávislé. Tím se myslí, že porodní váha dívek nijak nezávisí na porodní váze chlapců a zároveň v rámci jednotlivých skupin neexistuje mezi porodní váhou jednotlivých dětí žádná závislost. Oba tyto předpoklady můžeme považovat za splněné.

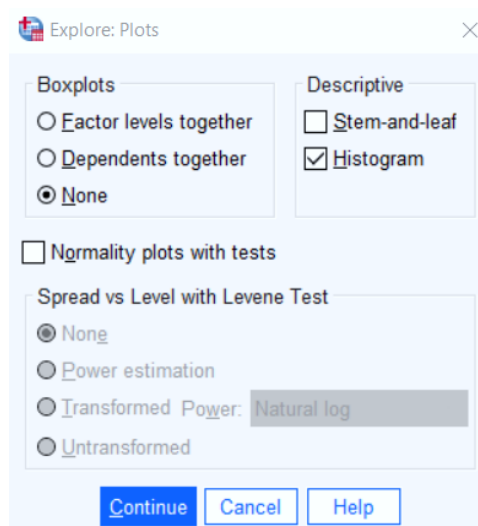
NORMALITA

Normální rozdělení musíme ověřovat samostatně pro obě skupiny, tedy nejprve pro porodní hmotnost chlapců a poté pro hmotnost dívek. Vzhledem k většímu počtu pozorování bude nejvhodnější normalitu posoudit na základě histogramu.

Klikneme na *Analyze*, dále na *Descriptive Statistics* a *Explore*. Do pole *Dependent List* zadáme proměnnou *Neonatal weight*. Do *Factor List* zadáme proměnnou *Gender*, která definuje dvě testované skupiny.



Následně v záložce *Plots* zvolíme histogram, zrušíme volbu *Stem-and-leaf* a v nabídce *Boxplots* klikneme na možnost *None*.



Potvrdíme tlačítkem *Continue* a v hlavním okně klikneme na OK. Otevře se nové výstupové okno, ve kterém se kromě histogramů zobrazí ještě tabulka s přehledem o počtu pozorování a popisnou statistikou.

Vzhledem k tomu, že v nulové hypotéze testujeme shodu průměrů může nám jistou představu o výsledku testu poskytnout i popisná statistika. V tabulce *Descriptives* najdeme průměrnou porodní váhu pro obě skupiny. Chlapci vážili v průměru 3258 gramů a dívky 3107 gramů. Rozdíl mezi nimi je tedy 150 gramů. Zda je tento rozdíl statisticky významný nám ale rozhodne až samotný test. Kromě průměru máme v tabulce také jeho interval spolehlivosti. Vzhledem k tomu, že analyzovaný datový soubor je výběrový, je i hodnota průměru jen odhad skutečné populační hodnoty. V takových situacích je vhodné tento tzv. bodový odhad průměru doplnit i o odhad intervalový, tedy interval spolehlivosti.

Interval spolehlivosti nám poskytuje informaci o tom, v jakém intervalu se se zvolenou hladinou spolehlivosti bude nacházet skutečná populační hodnota daného parametru. Nejčastěji se počítá právě pro průměr a volí se hladina spolehlivosti 95 %. Je ale možné si zvolit i jakoukoli jinou hladinu spolehlivosti. Platí ale, že při zvyšování hladiny spolehlivosti se bude interval spolehlivosti postupně rozšiřovat, což není vždy žádoucí.

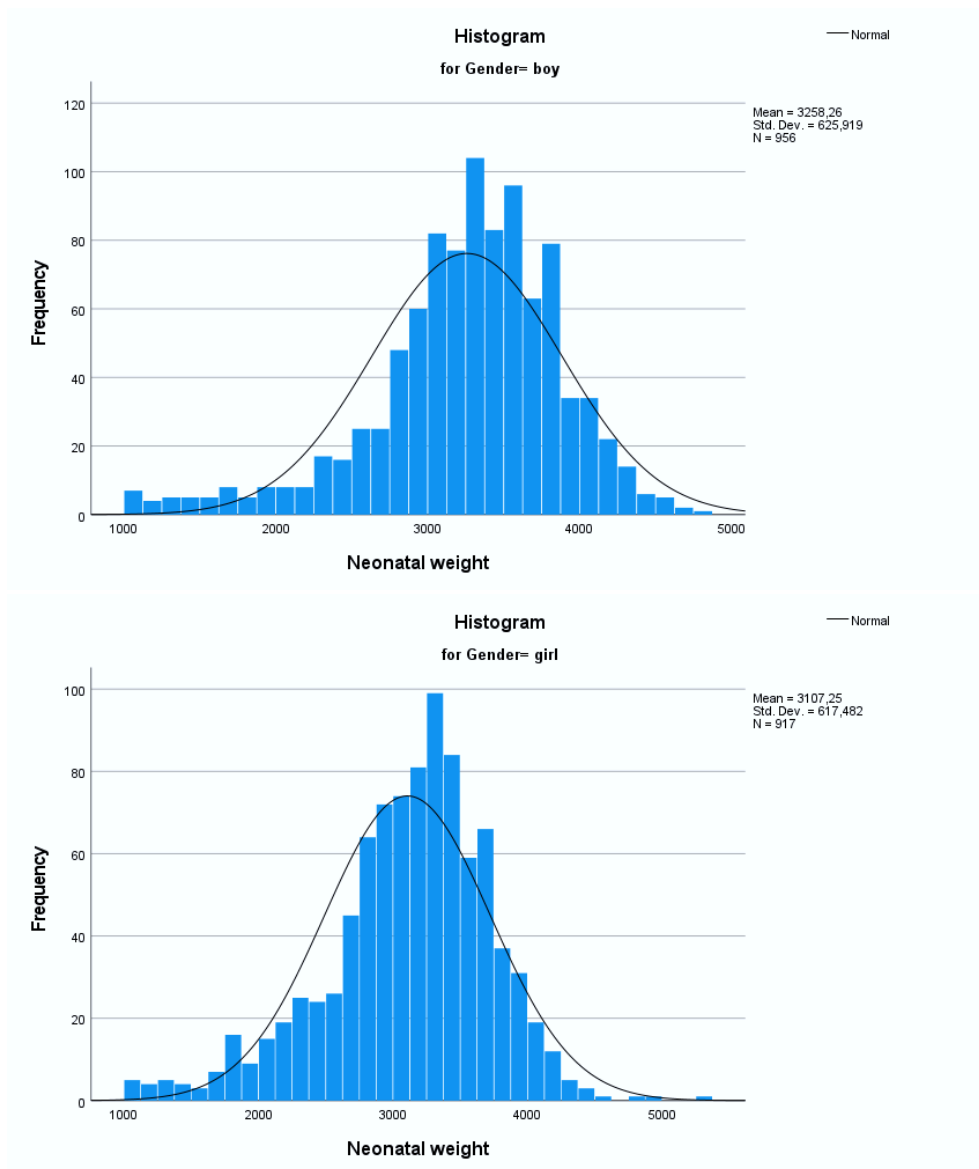
Dolní a horní hranici 95 % intervalu spolehlivosti pro průměrnou váhu chlapců a dívek najdeme v řádcích pod průměrem. U chlapců by se dal interpretovat následovně: průměrná porodní váha chlapců se bude s 95 % spolehlivostí pohybovat v intervalu 3218 gramů až 3297 gramů. U dívek se průměrná porodní váha bude s 95 % spolehlivostí pohybovat v intervalu 3067 gramů až 3147 gramů. Můžeme si všimnout, že intervaly spolehlivosti chlapců a dívek se ani v malé části nepřekrývají, mají tedy nulový průnik, což může být první indikátor toho, že by se nám mohlo pomoci dvouvýběrového t-testu podařit zamítnout nulovou hypotézu.

Abychom mohli přistoupit k výpočtu testu musíme ale ověřit zbývající předpoklady, v prvním kroku pomocí histogramu ověříme normální rozdělení v obou skupinách.

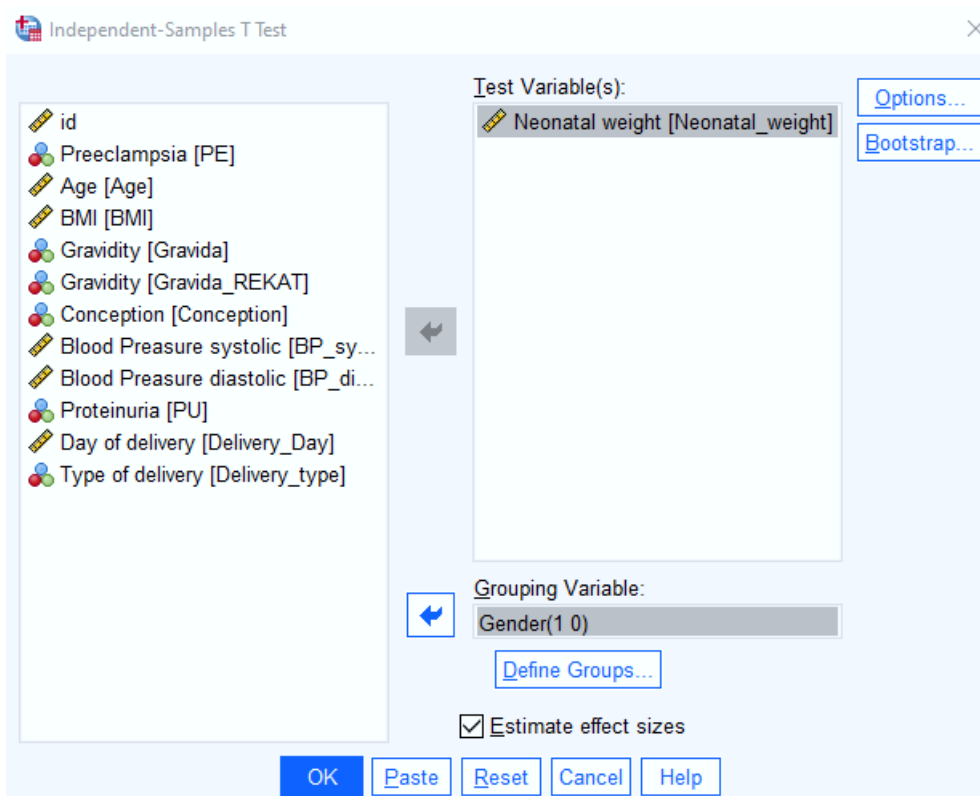
Descriptives				
Gender		Statistic		Std. Error
Neonatal weight	boy	Mean		3258,26
		95% Confidence Interval for Mean	Lower Bound	3218,54
			Upper Bound	3297,99
		5% Trimmed Mean		3294,43
		Median		3330,00
		Variance		391774,051
		Std. Deviation		625,919
		Minimum		1010
		Maximum		4870
		Range		3860
		Interquartile Range		700
		Skewness		-,932
		Kurtosis		,158
	girl	Mean		3107,25
		95% Confidence Interval for Mean	Lower Bound	3067,23
			Upper Bound	3147,27
		5% Trimmed Mean		3131,35
		Median		3170,00
		Variance		381283,508
		Std. Deviation		617,482
		Minimum		1025
		Maximum		5350
		Range		4325
		Interquartile Range		720
		Skewness		-,626
		Kurtosis		,842

Aby bylo grafické posouzení normality jednodušší, je vhodné si do jednotlivých histogramů vykreslit křivku normálního rozdělení. V prvním kroku 2x klikneme na konkrétní histogram, čímž otevřeme editační okno. Následně v editačním okně klikneme na záložku *Elements*, ve které najdeme volbu *Show Distribution Curve*. Po této úpravě by se v histogramu měla objevit černá křivka reprezentující normální rozdělení.

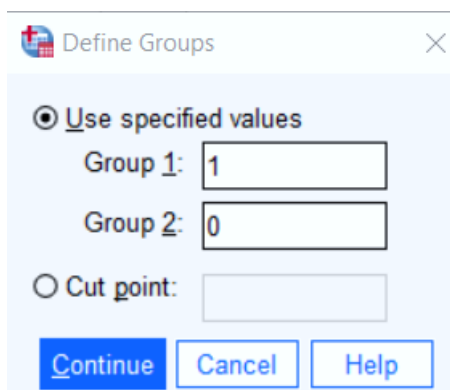
V obou případech vidíme, že odchylky od normality nejsou moc výrazné a je tedy možné použít parametrický dvouvýběrový t-test. V případě, že by rozdělení bylo výrazně odchýleno od normality, museli bychom použít neparametrickou alternativu tohoto testu (Mann-Whitney).



Dialog pro dvouvýběrový t-test najdeme v záložce Analyze-Compare Means – Independent samples T-test.



Do okna *Test Variable(s)* zadáme číselnou proměnnou, jejíž průměrnou hodnotu chceme testovat, tedy proměnnou *Neonatal weight*. Do okna *Grouping Variable* zadáme proměnnou *Gender*. Abychom mohli spustit výpočet testu musíme ještě v nastavení *Define Groups* zadat jakými hodnotami jsou v datové matici kódovány jednotlivé skupiny, tedy chlapci a dívky.



Následně potvrdíme *Continue* a v hlavním okně *OK*. Ve výstupovém okně se objeví 2 nové tabulky s výsledkem dvouvýběrového t-testu.

Tabulka *Independent Samples Test* má dva řádky. Je tomu tak proto, že se v ní nacházejí dvě varianty dvouvýběrového t-testu. V prvním řádku je jeho výsledek pro variantu, ve které jsou rozptyly obou skupin stejné, což je jeden z předpokladů testu. V případě, že tento předpoklad porušíme, tak je nutné použít robustní alternativu testu, která ale testuje stejnou nulovou hypotézu. To, zda jsme daný předpoklad splnili nebo ne se dozvíme v druhém sloupci *Sig.*, který obsahuje signifikanci pro Leveneho test.

Independent Samples Test											
		Levene's Test for Equality of Variances				t-test for Equality of Means					
		F	Sig.	t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						One-Sided p	Two-Sided p			Lower	Upper
Neonatal weight	Equal variances assumed	,136	,712	-5,254	1871	<,001	<,001	-151,010	28,741	-207,378	-94,641
	Equal variances not assumed			-5,256	1869,523	<,001	<,001	-151,010	28,733	-207,362	-94,657

Nulová hypotéza Leveného testu předpokládá shodu rozptylů ve skupinách, tedy abychom ji nezamítali měla by být signifikance větší než zvolená hladina významnosti. V tomto příkladu bude zvolena hladina významnosti 5 % (0,05). Signifikace testu je 0,712. **To znamená, že na 5 % hladině významnosti nezamítáme nulovou hypotézu o shodě rozptylů, splnili jsme tedy předpoklad dvouvýběrového t-testu** a můžeme pro jeho interpretaci dále pokračovat v prvním řádku tabulky Independent Samples Test. Signifikace pro dvouvýběrový t-testu je ve sloupci *Significance - Two-Sided p*. Je velmi blízká nule. **Na 5 % hladině významnosti tedy zamítáme nulovou hypotézu, která tvrdí, že průměrná porodní váha chlapců a dívek je stejná. Potvrdili jsme tedy, že rozdíl v průměrné porodní váze je statisticky významný (není způsoben jen náhodnými odchylkami) a porodní váha dívek je v průměru nižší.**