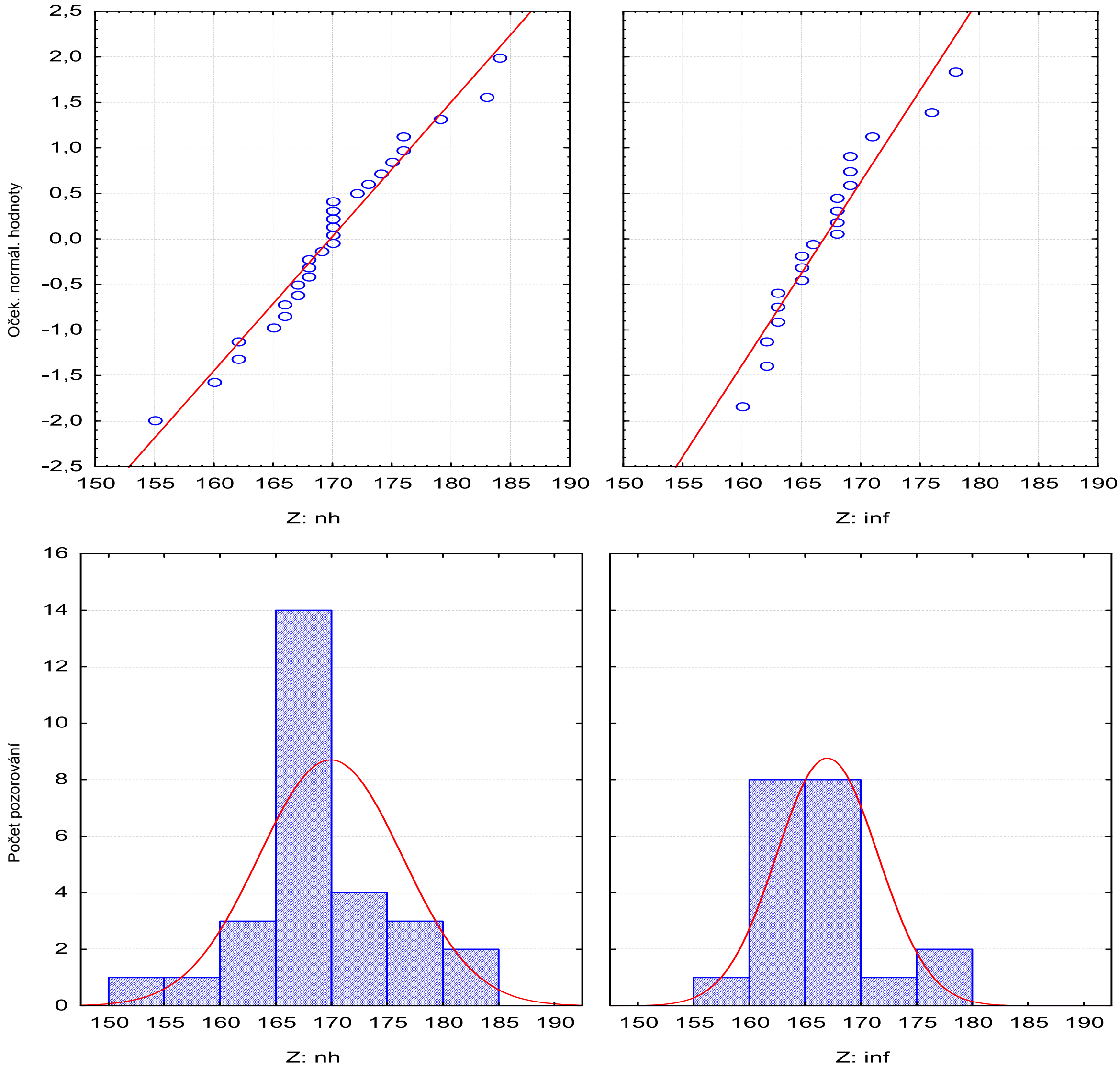
Dvojvýberové testy

Úkol 1.: Do programu STATISTICA načtěte soubor 07_data.sta, který obsahuje údaje o 48 náhodně vybraných studentkách VŠE v Praze:
1. sloupec – výška, 2. sloupec – známka z matematiky v 1. semestru, 3. sloupec – obor studia (1 – národní hospodářství, 2 – informatika).

Úkol 2.: Orientačně ověřte normalitu výšky ve skupině studentek oboru národní hospodářství a oboru informatika vykreslením N-P plotu a histogramu.

Návod:
Grafy – 2D Grafy – Normální pravděpodobnostní grafy – Proměnné X – na záložce Kategorizovaný zaškrtneme Kategorie X Zapnuto – Změnit proměnnou – Z - OK – OK.
Podobně pro histogram.



Komentář: Grafy svědčí o mírném narušení normality, jedná se o mírné kladné zešikmení.

Nyní provedeme testy normality.

Statistiky – Základní statistiky/tabulky – Tabulky četností – OK – Select cases – Zapnout filtr – některé vybrané pomocí Z=1 – OK – Proměnná X – OK - Normalita - zaškrtneme Liliefors test, Shapiro-Wilk's test - Testy normality. Dostaneme tyto výsledky:

Pro studentky oboru nh

Proměnná	Testy normality (studentky.sta) Zhrnout podmínku: Z=1				
	N	max D	Lilliefors p	W	p
X: vyska	28	0,167473	p < ,05	0,970969	0,606793

Pro studentky oboru inf

Proměnná	Testy normality (studentky.sta) Zhrnout podmínku: Z=2				
	N	max D	Lilliefors p	W	p
X: vyska	20	0,172301	p < ,15	0,922747	0,111924

Komentář: Vypočtenou p-hodnotu porovnáváme se zvolenou hladinou významnosti testu (většinou volíme $\alpha = 0{,}05$). Je-li vypočtená p-hodnota $\leq \alpha$, pak hypotézu o normalitě zamítáme na hladině významnosti α. V našem případě dojde k zamítnutí hypotézy o normalitě výšky na hladině významnosti 0,05 pouze u Lilieforsova testu pro studentky oboru nh.

Úkol 3.: Na hladině významnosti 0,05 testujte hypotézu, že rozptyly výšek studentek oboru nh a inf jsou shodné.

Návod:

Jedná se o F-test, kdy testujeme hypotézu $H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1$ proti oboustranné alternativě

$$H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1$$

- F-test je implementován ve STATISTICE.

Statistiky – Základní statistiky/tabulky – t-test, nezávislé, podle skupn - OK, Proměnné – Závislé proměnné X, Grupovací proměnná Z – OK – Výpočet

Proměnná	t-testy; grupováno: Z: obor studia (Tema7) Skup. 1: nh: narodni hospodarstvi Skup. 2: inf: informatika										
	Průměr nh	Průměr inf	t	sv	p	Poč.plat nh	Poč.plat. inf	Sm.odch. nh	Sm.odch. inf	F-poměr rozptyly	p rozptyly
X	169,8214	166,9000	1,744008	46	0,087837	28	20	6,417878	4,552616	1,987288	0,124925

Komentář: Ve výstupní tabulce nás zajímá hodnota testové statistiky F-testu (v našem případě 1,987288) a odpovídající p-hodnota: 0,124925. Protože p-hodnota je větší než hladina významnosti $\alpha = 0{,}05$, nelze na hladině významnosti 0,05 zamítnout nulovou hypotézu.
S rizikem omylu nanejvýš 5% se tedy neprokázalo, že by rozptyly výšek studentek oborů nh a inf byly odlišné.

Úkol 4.: Na hladině významnosti 0,05 testujte hypotézu, že střední hodnoty výšek studentek oboru nh a inf jsou shodné. Výpočet doplňte krabicovými diagramy.

Návod:

Jedná se o dvouvýběrový t-test, kdy testujeme hypotézu $H_0 : \mu_1 - \mu_2 = 0$ proti oboustranné alternativě $H_1 : \mu_1 - \mu_2 \neq 0$

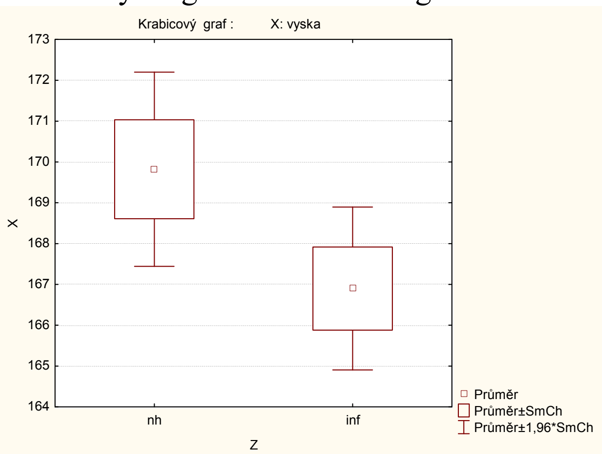
- dvouvýběrový t-test je implementován ve STATISTICE.

Statistiky – Základní statistiky/tabulky – t-test, nezávislé, podle skupin - OK, Proměnné – Závislé proměnné X, Grupovací proměnná Z – OK – Výpočet

Proměnná	t-testy; grupováno: Z: obor studia (Tema7) Skup. 1: nh: narodni hospodarstvi Skup. 2: inf: informatika										
	Průměr nh	Průměr inf	t	sv	p	Poč.plat nh	Poč.plat. inf	Sm.odch. nh	Sm.odch. inf	F-poměr rozptyly	p rozptyly
X	169,8214	166,9000	1,744008	46	0,087837	28	20	6,417878	4,552616	1,987288	0,124925

Komentář: Ve výstupní tabulce najdeme hodnotu testového kritéria (t_0 = 1,744006) a odpovídající p-hodnotu. Protože p-hodnota = 0,087837 je větší než hladina významnosti 0,05, nulovou hypotézu nezamítáme na hladině významnosti 0,05. S rizikem omylu nanejvýš 5% se tedy neprokázal rozdíl mezi středními hodnotami výšek studentek oborů nh a inf.

Konstrukce krabicových diagramů: V tabulce t-test, nezávislé, podle skupin zvolíme Krabicový diagram. Dostaneme graf:



Komentář: Ze vzhledu krabicových diagramů je vidět, že rozložení výšek v obou skupinách je vcelku symetrické kolem průměru, odlehlé ani extrémní hodnoty se nevyskytují, variabilita vyjádřená směrodatnou odchylkou se liší jen nepatrně a průměrná výška ve skupině studentek oboru inf je o něco menší než ve skupině studentek oboru nh.

Poznámka: Protože F-test neprokázal odlišnost rozptylů, mohli jsme ve STATISTICE použít variantu dvouvýběrového t-testu se shodnými rozptyly. Pokud by však F-test zamítl na dané hladině významnosti hypotézu o shodě rozptylů, museli bychom zvolit variantu dvouvýběrového t-testu se separovanými odhady rozptylů.

Úkol 5.: **Párový znaménkový test a párový Wilcoxonův test**

Při zjišťování kvality jedné složky půdy se používají dvě metody označené A a B. Výsledky:

Vzorek	1	2	3	4	5	6	7	8	9	10	11	12
A	0,275	0,312	0,284	0,3	0,365	0,298	0,312	0,315	0,242	0,321	0,335	0,307
B	0,28	0,312	0,288	0,298	0,361	0,307	0,319	0,315	0,242	0,323	0,341	0,315

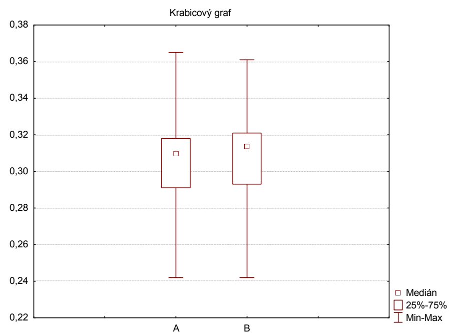
Na hladině významnosti 0,05 testujte pomocí párového znaménkového testu a poté pomocí párového Wilcoxonova testu hypotézu, že metody A a B dávají stejné výsledky.

Návod:
Načteme datový soubor 07_data2.sta. Proměnná A obsahuje výsledky metody A, proměnná B výsledky metody B.
Nejprve budeme testovat nulovou hypotézu pomocí párového znaménkového testu.
Statistiky –Neparametrická statistika – Porovnání dvou závislých vzorků (proměnné) – OK – 1. seznam proměnných A, 2. seznam proměnných B – OK – Znaménkový test.

Dvojice proměnných	Znaménkový test (kvalita_pudy.sta) Označené testy jsou významné na hladině p <,05000			
	Počet různých	procent v < V	Z	Úroveň p
A & B	9	77,77778	1,333333	0,182422

Komentář: Vidíme, že nenulových hodnot n = 9. Z nich záporných je $77,\overline{7}$ %, tj.7. Kladných je tedy 9 – 7 = 2, což je hodnota testové statistiky S_Z^+. Asymptotická testová statistika U_0 (zde označená jako Z) se realizuje hodnotou $1,\overline{3}$. Odpovídající asymptotická p-hodnota je 0,18422, tedy na asymptotické hladině významnosti 0,05 nezamítáme hypotézu, že obě metody dávají stejné výsledky.
Upozornění: V tomto případě není splněna podmínka pro využití asymptotické normality statistiky S_Z^+, tj. n > 20. Je tedy vhodnější najít v tabulkách kritické hodnoty pro znaménkový test. Pro n = 9 a $\alpha = 0{,}05$ jsou kritické hodnoty $k_1 = 1$, $k_2 = 8$. Protože kritický obor $W = \langle 0,1 \rangle \cup \langle 8,9 \rangle$ neobsahuje hodnotu 2, nezamítáme H_0 na hladině významnosti 0,05.
Dostáváme stejný výsledek při použití asymptotického testu

Nyní graficky znázorníme výsledky obou metod: Návrat do Porovnání 2 proměnných - Krabicový graf všech proměnných – OK – A, B – OK.



Komentář: Z krabicových diagramů je vidět, že obě metody se poněkud liší v úrovni, ale neliší se ve variabilitě.

Dále provedeme Wilcoxonův párový test: Návrat do Porovnání 2 proměnných – Wilcoxonův párový test.

Dvojice proměnných	Wilcoxonův párový test (kvalita_pudy) Označené testy jsou významné na hladině p <,05000			
	Počet platných	T	Z	p-hodn.
A & B	9	5,000000	2,073221	0,038153

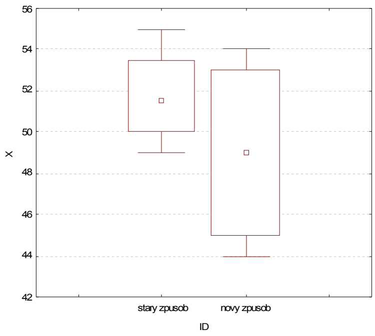
Komentář: Výstupní tabulka poskytne hodnotu testové statistiky S_W^+ (zde označena T), hodnotu asymptotické testové statistiky U_0 a p-hodnotu pro U_0. V tomto případě je p-hodnota 0,038153, tedy nulová hypotéza se zamítá na asymptotické hladině významnosti 0,05. Ze srovnání asymptotických p-hodnot pro znaménkový test a pro Wilcoxonův test plyne, že Wilcoxonův test je silnější.
Upozornění: V tomto případě není splněna podmínka pro využití asymptotické normality statistiky S_W^+, tj. $n \geq 30$. Je tedy vhodnější najít v tabulkách kritickou hodnotu pro Wilcoxonův párový test. Pro $n = 9$ a $\alpha = 0,05$ je kritická hodnota rovna 5. Protože kritický obor $W = \langle 0,5 \rangle$ obsahuje hodnotu 5, zamítáme H_0 na hladině významnosti 0,05. To souhlasí s výsledkem asymptotického testu.

Úkol 6.: **Dvouvýběrový Wilcoxonův test**
Bylo vybráno 10 polí stejné kvality. Na čtyřech z nich se zkoušel nový způsob hnojení, zbylých šest bylo ošetřeno starým způsobem. Pole byla oseta pšenicí a sledoval se její hektarový výnos. Je třeba testovat na hladině významnosti 0,05, zda nový způsob hnojení má týž vliv na průměrné hektarové výnosy pšenice jako starý způsob hnojení.
hektarové výnosy při novém způsobu: 51 52 49 55
hektarové výnosy při starém způsobu: 45 54 48 44 53 50

Návod:
Načteme datový soubor 07_data3.sta. Proměnná X udává výnosy pšenice při obou způsobech hnojení a proměnná ID nabývá hodnoty 1 pro starý způsob hnojení, hodnoty 2 pro nový způsob hnojení.
Dvouvýběrový Wilcoxonův test: Statistiky – Neparametrická statistika – Porovnání dvou nezávislých vzorků (skupiny) – OK - Seznam závislých proměnných X, Grupovací proměnná ID - OK - Mann – Whitneyův U test.

	Mann-Whitneyův U test (hnojeni_poli) Dle proměn. ID Označené testy jsou významné na hladině p <,05000									
Proměnná	Sčt poř. st. zp.	Sčt poř. n. zp.	U	Z	p-hodn.	Z uprav.	p-hodn.	N platn. st. zp.	N platn. n. zp.	2*1str. přesné p
X	27,00000	28,00000	7,000000	0,959403	0,337356	0,959403	0,337356	4	6	0,352381

Komentář: Ve výstupní tabulce jsou součty pořadí T_1, T_2, hodnota testové statistiky $\min(U_1, U_2)$ označená U, hodnota asymptotické testové statistiky U_0 (označená Z), asymptotická p-hodnota pro U_0 a přesná p-hodnota (ozn. 2*1 str. přesné p – ta se používá pro rozsahy výběrů pod 30). V našem případě přesná p-hodnota = 0,352381, tedy H_0 nezamítáme na hladině významnosti 0,05.
Výpočet je vhodné doplnit krabicovým diagramem.



Komentář: Je zřejmé, že výnosy při novém způsobu hnojení jsou vesměs nižší než při starém způsobu a také vykazují mnohem větší variabilitu.

Příklady k samostatnému řešení

Příklad 1.: U osmi osob byl změřen systolický krevní tlak před pokusem a po něm.

č. osoby	1	2	3	4	5	6	7	8
tlak před	130	185	162	136	147	181	128	139
tlak po	139	190	175	135	155	175	158	149

Na hladině významnosti 0,05 testujte hypotézu, že pokus neovlivní systolický krevní tlak.

Řešení:

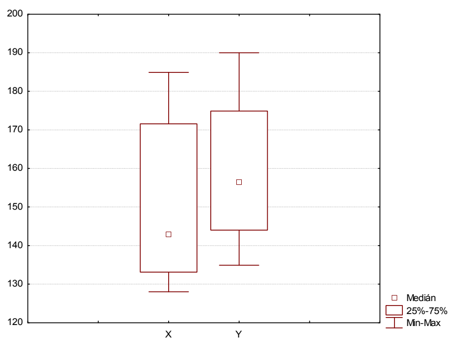
Stejně jako v úkolu 5 provedeme párový znaménkový a párový Wilcoxonův test. Načteme soubor 07_data4.sta. Proměnná X obsahuje hodnoty tlaku před pokusem, proměnná Y po pokusu.

Výstupní tabulka:

	Znaménkový test (tlak.sta) Označené testy jsou významné na hladině p <,05000			
Dvojice proměnných	Počet různých	procent v < V	Z	Úroveň p
X & Y	8	75,00000	1,060660	0,288844

Komentář: Jelikož p-hodnota = 0,288844 > 0,05, nelze nulovou hypotézu zamítnout na hladině významnosti 0,05. Vzhledem k malému rozsahu výběru je však vhodnější najít v tabulkách kritické hodnoty pro znaménkový test (viz skripta Základní statistické metody, tabulka na straně 156). Pro n = 8 a $\alpha = 0{,}05$ jsou kritické hodnoty $k_1 = 0$, $k_2 = 8$. Hodnotu testové statistiky S_Z^+ získáme jako 75% z 8, což je 6. Protože kritický obor neobsahuje hodnotu 6, nezamítáme H_0 na hladině významnosti 0,05. Dostáváme stejný výsledek při použití asymptotického testu

Grafické znázornění výsledků pomocí krabicového diagramu:



Komentář: Úroveň tlaku před pokusem byla poněkud nižší než po pokusu, variabilita je jen nepatrně odlišná.

Výstupní tabulka Wilcoxonova testu:

Dvojice proměnných	Wilcoxonův párový test (tlak.sta) Označené testy jsou významné na hladině p <,05000 Počet platných	T	Z	Úroveň p
X & Y	8	4,000000	1,960392	0,049951

Vidíme, že asymptotická p-hodnota = 0,049951, nulová hypotéza se tedy zamítá na asymptotické hladině významnosti 0,05. Rozsah souboru je pouze 8, není splněna podmínka dobré aproximace standardizovaným normálním rozložením ($n > 30$). Proto zjistíme ve skriptech Základní statistické metody v tabulce na str. 157 kritickou hodnotu pro $n = 8$ a $\alpha = 0{,}05$. Kritická hodnota je rovna 3, hodnota testové statistiky (ve výstupní tabulce označena T) je $4 > 3$, tedy nulovou hypotézu nezamítáme na hladině významnosti 0,05, což je v souladu s výsledkem znaménkového testu.

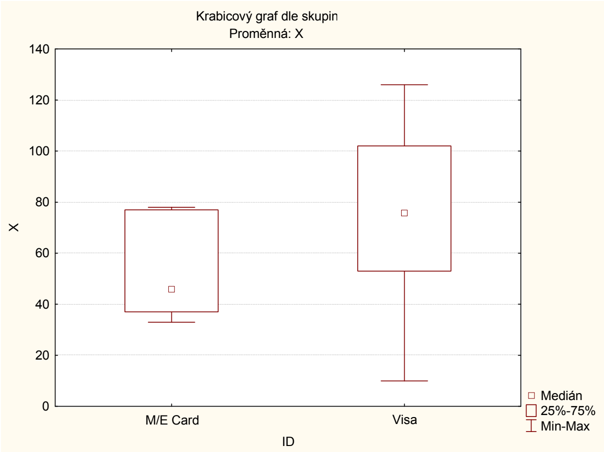
Příklad 2.: Majitel obchodu chtěl zjistit, zda velikost nákupů (v dolarech) placených kreditními kartami Master/EuroCard a Visa jsou přibližně stejné. Náhodně vybral 7 nákupů placených Master/EuroCard: 42 77 46 73 78 33 37 a 9 placených Visou: 39 10 119 68 76 126 53 79 102. Lze na hladině významnosti 0,05 tvrdit, že nákupů placených těmito dvěma typy karet se shodují?

Řešení:
Stejně jako úkolu 6 použijeme dvouvýběrový Wilcoxonův test.
Načteme datový soubor 07_data5.sta. Proměnná X obsahuje hodnoty nákupů, proměnná ID má hodnotu 1 pro kartu Master/EuroCard a hodnotu 2 pro kartu Visa.
Výstupní tabulka pro dvouvýběrový Wilcoxonův test:

Proměnná	Mann-Whitneyův U test (kreditni_karty.sta) Dle proměn. ID Označené testy jsou významné na hladině p <,05000 Sčt poř. M/E Card	Sčt poř. Visa	U	Z	Úroveň p	Z upravené	Úroveň p	N platn. M/E Card	N platn. Visa	2*1str. přesné p
X	48,00000	88,00000	20,00000	-1,21729	0,223495	-1,21729	0,223495	7	9	0,252273

Komentář: Zajímá nás především přesná p-hodnota (ozn. 2*1 sided exact p – ta se používá pro rozsahy výběrů pod 30). V našem případě přesná p-hodnota = 0,252273, tedy H_0 nezamítáme na hladině významnosti 0,05.

Výpočet je vhodné doplnit krabicovým diagramem typu Medián/Kvartily/Rozpětí.



Komentář: Vidíme, že platby za nákupy kartou Master/EuroCard mají nižší úroveň, ale přibližně stejnou variabilitu jako platby kartou Visa.