1 Úvod do statistického myšlení

Cílem tématu je pochopení principu výběrových šetření, úvodní seznámení s některými statistickými metodami a počáteční nácvik práce v Excelu a ve Statistice.

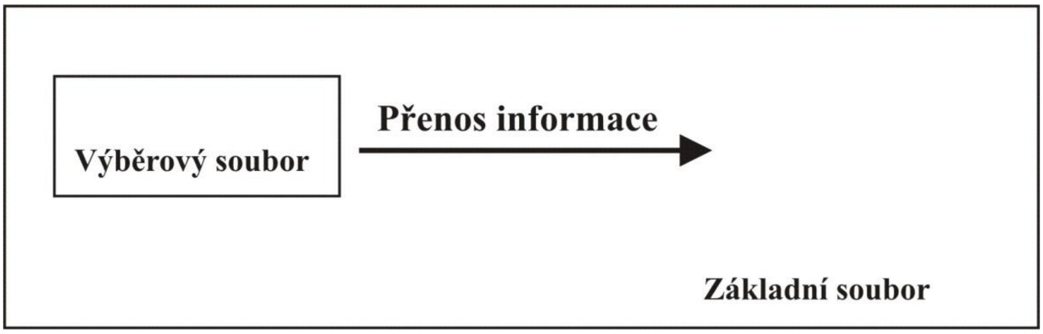
1.1 Základní soubor a náhodný výběr, měření jako realizace náhodného výběru

Základní idea statistiky, týkající se výběrových šetření, je tato:

Získaná či naměřená data, tvořící tzv. *výběrový soubor*, považujeme za hodnoty náhodných veličin. O těchto veličinách předpokládáme, že v *základním souboru (populaci)* mají určité *teoretické rozdělení pravděpodobnosti*.

Statistickými metodami pak získáváme poznatky o těchto teoretických rozděleních a **přenášíme informaci z výběrového souboru na základní soubor**.

Výběrový soubor musí být pro základní soubor *reprezentativní* a musí mít vzhledem k cílům a použitým metodám dostatečně velký rozsah.



Podrobně se touto základní myšlenkou budeme zabývat v odstavci 1.1.2, její pochopení nám usnadní práce s excelovským souborem Náhodné veličiny, bodové odhady, intervaly spolehlivosti.
Nejdříve se ale v následujícím odstavci 1.1.1 seznámíme s velmi často užívaným modelem – tzv. *normálním rozdělením pravděpodobnosti*.

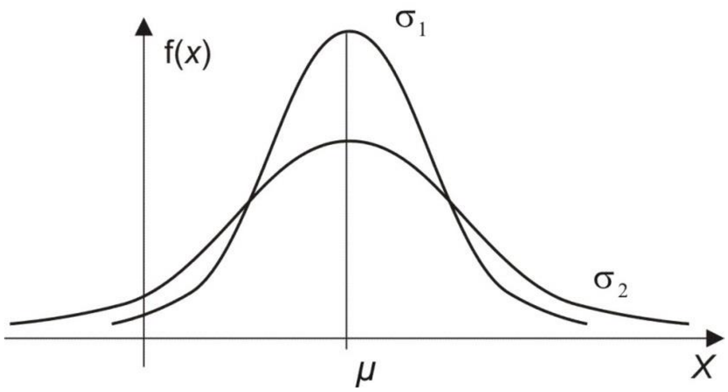
1.1.1 Normální rozdělení pravděpodobnosti a jeho parametry

Často se v aplikacích setkáváme s veličinami, jejichž hodnoty se většinově kumulují okolo tzv. *střední hodnoty*, přičemž větší odchylky od střední hodnoty jsou málo pravděpodobné.

Příkladem může být třeba výška lidí: tam jsou poměrně časté hodnoty kolem 178 cm (střední hodnota) a velmi malých lidí či velmi velkých lidí je relativně málo.

Dalším příkladem mohou být třeba hodnoty získané měřením jisté fyzikální či chemické konstanty. Vlivem náhodných okolností se dopouštíme chyb, jejichž střední hodnota je rovna 0, malé chyby jsou častější a větší chyby se vyskytují zřídka.

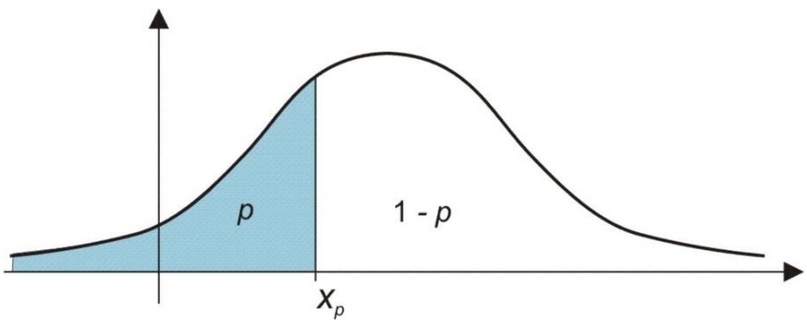
Normální rozdělení pravděpodobnosti je často užívaným modelem pro řadu náhodných veličin. Pravděpodobnosti výskytu různě velkých hodnot v tomto modelu určuje tzv. *hustota normálního rozdělení* (viz následující obrázek), která má kromě střední hodnoty (označujeme ji obvykle μ) ještě jeden parametr, a to tzv. *směrodatnou odchylku* (označujeme ji obvykle σ). Ta je malá, když jsou hodnoty „nakumulovány velmi blízko střední hodnotě" (tomu odpovídá užší a vyšší křivka s parametrem σ_1) a je větší, když se hodnoty „odlišují od střední hodnoty více" (tomu odpovídá širší a nižší křivka s parametrem σ_2).



Velikost celé plochy ohraničené zdola osou *x* a shora hustotou je rovna číslu 1. Velikost plochy pod hustotou nad nějakým intervalem je rovna pravděpodobnosti toho, že hodnota náhodné veličiny je v tomto intervalu.

Nejprve se seznámíme s tzv. *kvantily x_p normálního rozdělení* (viz následující obrázek).
Pro danou hodnotu pravděpodobnosti *p*, kde $0 < p < 1$, má kvantil x_p tuto vlastnost:

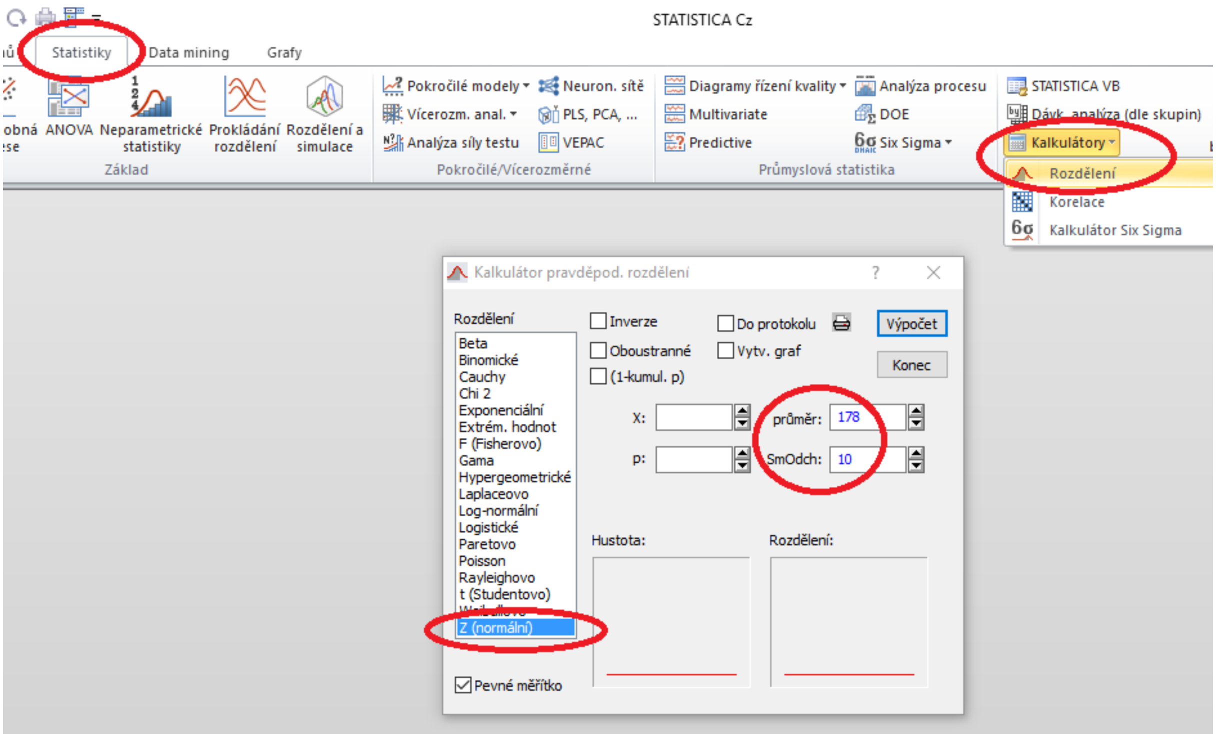
- hodnota náhodné veličiny bude menší než x_p s pravděpodobností *p*,
- a hodnota náhodné veličiny bude větší než x_p s pravděpodobností $1 - p$.



Podrobněji si ukážeme význam hustoty rozdělení a kvantilů pomocí programu Statistica.

Nejprve si aktivujeme pravděpodobnostní kalkulátor normálního rozdělení.
Cesta: Statistiky → Kalkulátory → Rozdělení → Z (normální)

Nastavíme si střední hodnotu μ = 178 a směrodatnou odchylku σ = 10 (bude to pro nás model výšky lidí v základním souboru).



Tento kalkulátor můžeme využívat dvěma způsoby:

- buď zadáme hodnotu *p* a klikneme na výpočet – v poli *X* se objeví příslušný kvantil,
- anebo zadáme hodnotu *X* a klikneme na výpočet – pak se v poli *p* objeví pravděpodobnost toho, že výška libovolného člověka v základním souboru je menší než zadané *X* (což po vynásobení číslem 100 dává procento lidí v základním souboru, jejichž výška je menší než zadané *X*).

Úloha 1

- Zjistěte hodnotu kvantilu $x_{0,70}$ (tato výška má tu vlastnost, že 70% lidí v základním souboru je menších a 30% lidí je větších). Zjistěte ještě kvantily $x_{0,50}$ a $x_{0,15}$.
  Výsledky: 183,244; 178,000; 167,636
- Zjistěte, kolik procent lidí má menší výšku než 170 cm, resp. Větší výšku než 185 cm.
  Výsledky: 0,211855 ~ 21,2 %; 1 – 0,758036 = 0,241964 ~ 24,2 %
- Zjistěte, kolik procent lidí má výšku v intervalu od 160 do 190 cm, resp. v intervalu od 148 do 208 cm (viz následující *Pravidlo* 3σ).
  Výsledky: 0,884930 – 0,035930 = 0,849 ~ 84,9 %; 0,9973 ~ 99,7 %

Pravidlo 3σ

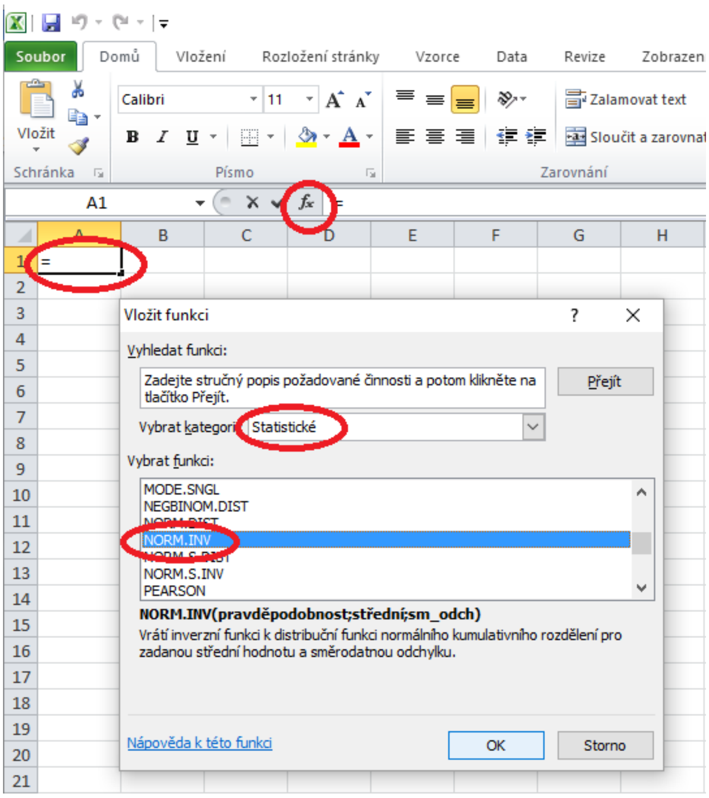
Prakticky všechny hodnoty náhodné veličiny s normálním rozdělením se střední hodnotou μ a směrodatnou odchylkou σ leží v intervalu od hodnoty μ – 3σ do hodnoty μ + 3σ. Mimo tento interval leží jen 0,3 % hodnot náhodné veličiny.

Pokud chceme pracovat s kvantily normálního rozdělení v Excelu, můžeme používat statistické funkce NORM.INV a NORM.DIST, které aktivujeme ikonkou f_x na začátku příkazového řádku a výběrem potřebné funkce.

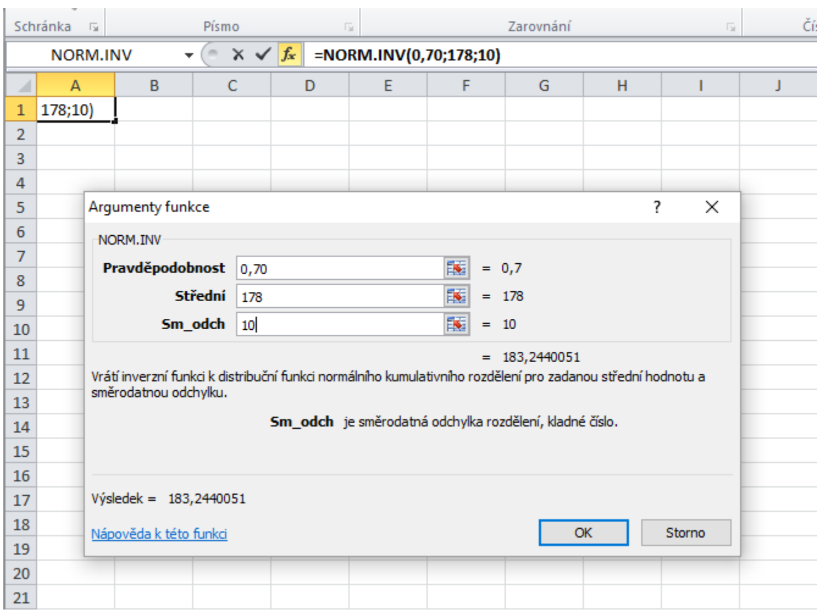
Pomocí první z těchto funkcí získáme pro zadanou pravděpodobnost odpovídající kvantil, druhá z těchto funkcí nám k zadané hodnotě x určí pravděpodobnost toho, že normálně rozdělená veličina nabývá hodnoty menší než zadaná hodnota.

Chceme-li tedy například zjistit hodnotu $x_{0,70}$ z předchozí úlohy, můžeme postupovat takto:

po otevření Excelu označíme buňku (například A1) a ikonou f_x aktivujeme výběr funkce (je zařazena v kategorii Statistických funkcí).



Po potvrzení OK pak zadáme požadované parametry, tedy pravděpodobnost 0,70 a střední hodnotu a směrodatnou odchylku normálního rozdělení, tedy 178 a 10.



Po potvrzení se v buňce A1 objeví příslušný kvantil 183,224, tedy stejná hodnota, jakou jsme získali v úloze 1.

Úloha 2
Zjistěte všechny hodnoty kvantilů a pravděpodobností z úlohy 1 pomocí Excelu.

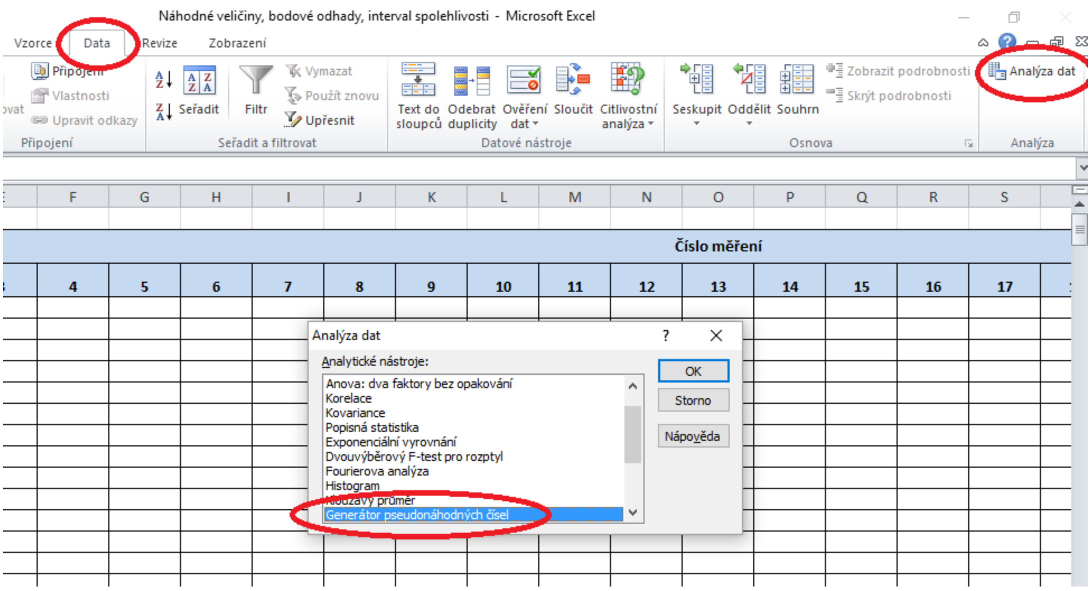
Známe-li tedy u normálního rozdělení střední hodnotu μ a směrodatnou odchylku σ, můžeme pro libovolný interval vypočítat pravděpodobnost toho, že hodnota náhodné veličiny je v tomto intervalu.

1.1.2 Náhodné veličiny

Připravíme si teď fiktivní situaci, kdy ze základního souboru výšek lidí (budeme předpokládat, že výška má normální rozdělení se střední hodnotou 178 a směrodatnou odchylkou 10) neuděláme jen jeden výběr, ale uděláme jich mnoho, konkrétně 1000. Každý výběr bude mít rozsah 25 lidí. Budeme si třeba představovat, že výšku lidí zkoumá 1000 výzkumníků a každý z nich změří 25 lidí.

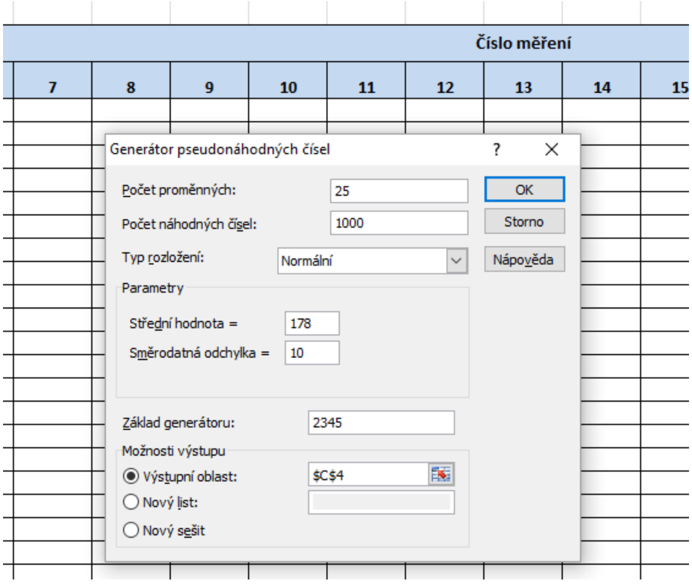
Zdůrazněme, že naše situace je opravdu jen fiktivní. Při skutečných výběrových šetřeních samozřejmě ani střední hodnotu ani směrodatnou odchylku neznáme, a děláme ze základního souboru jen jeden výběr. Náš další myšlenkový postup nám ale umožní zjistit, jak jsou některé vybrané statistické metody efektivní.

Budeme pracovat v Excelu se souborem Náhodné veličiny, bodové odhady, intervaly spolehlivosti a nejprve si zaplníme připravenou tabulku fiktivními daty.
Použijeme Generátor pseudonáhodných čísel (cesta: Data → Analýza dat → Generátor pseudonáhodných čísel).

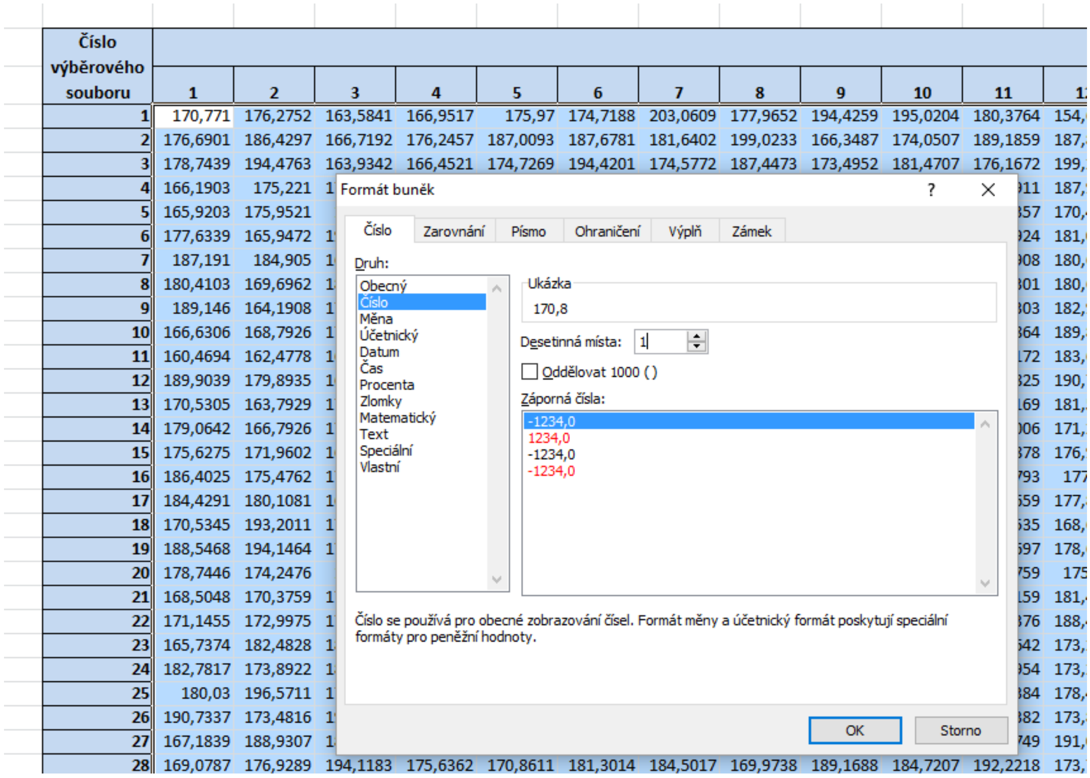


Po potvrzení OK generátor nastavíme takto:

Počet proměnných	25	(je to počet zaplňovaných sloupců)
Počet náhodných čísel	1000	(je to počet zaplňovaných řádků)
Typ rozložení	normální	
Střední hodnota	178	
Směrodatná odchylka	10	
Základ generátoru	2345	(abychom získali všichni stejné výsledky)
Výstupní oblast	ťukneme do tabulky na levou horní buňku C4	



Po potvrzení OK se nám tabulka zaplní fiktivními daty. Protože čísla mají příliš mnoho desetinných míst, stiskněme pravé tlačítko a v nabídce Formát buněk zvolíme Číslo a v Desetinných místech si nastavíme 1.



Po potvrzení kliknutím na OK se objeví zaokrouhlená fiktivní data. Některý z řádků si označte barevně – bude představovat vámi uskutečněné měření, ostatní řádky budou patřit zbylým 999 výzkumníkům.

Číslo výběrového souboru	1	2	3	4	5	6	7	8	9	10	11	12	13 Číslo měření
1	170,8	176,3	163,6	167,0	176,0	174,7	203,1	178,0	194,4	195,0	180,4	154,6	174,8
2	176,7	186,4	166,7	176,2	187,0	187,7	181,6	199,0	166,3	174,1	189,2	187,8	177,1
3	178,7	194,5	163,9	166,5	174,7	194,4	174,6	187,4	173,5	181,5	176,2	199,3	184,7
4	166,2	175,2	176,9	176,3	159,2	161,9	190,0	164,7	181,7	165,5	191,1	187,9	194,3
5	165,9	176,0	181,5	183,4	177,4	172,5	185,6	167,5	170,5	191,2	165,8	170,5	185,1
6	177,6	165,9	192,0	187,9	174,1	183,6	175,1	171,4	193,3	182,1	184,3	181,0	181,1
7	187,2	184,9	160,2	191,4	183,5	187,4	166,9	167,5	164,2	177,4	185,9	180,6	186,1
8	180,4	169,7	185,0	192,9	177,8	171,7	194,1	173,9	161,5	166,0	171,3	180,7	176,7
9	189,1	164,2	175,7	176,2	165,4	169,3	168,4	172,8	187,9	179,0	181,9	183,0	188,5
10	166,6	168,8	172,5	175,5	177,8	168,8	165,7	170,6	193,6	181,6	175,9	189,8	174,2

Překontrolujte si, že tabulka fiktivních dat má opravdu požadované rozměry, tedy že máme opravdu k dispozici 1000 výběrů po 25 naměřených výškách.

Soustřeďme se nyní jen na sloupec C, který obsahuje 1. měření, tady výšku prvního měřeného člověka ve výběru každého z výzkumníků. Porovnejte své první měření s měřeními ostatních. Je patrné, že váš výsledek závisí na náhodě, jistě mohl být o něco menší či větší. Přehled o možných hodnotách vašeho prvního měření nám poskytne tzv. *histogram*, který nám graficky vyznačí počet hodnot ze sloupce C, které leží ve vhodně zvolených intervalech (třídách).

Nejprve si určíme vhodné intervaly výšek:

- prvním intervalem budou výšky do 150 cm (včetně),
- druhým intervalem budou výšky větší než 150 cm a menší nebo rovny 155 cm,

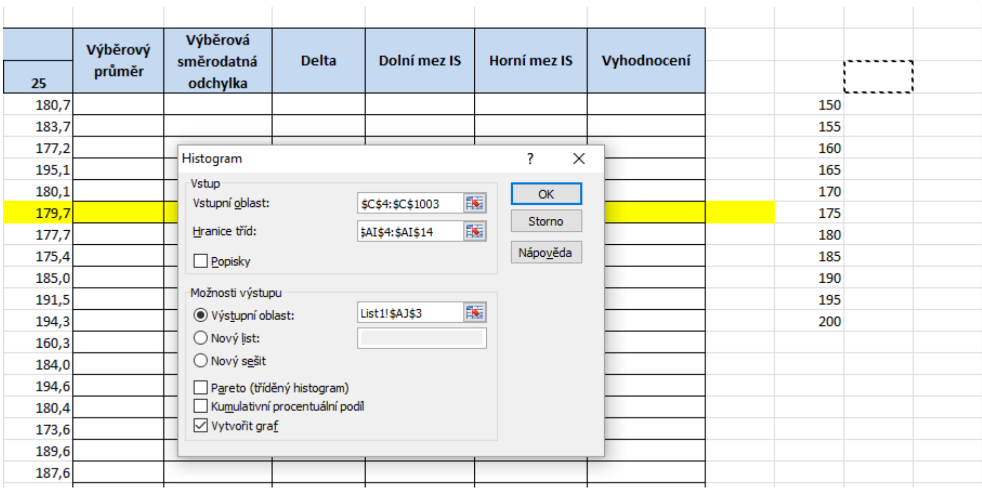
atd.

- jedenáctým intervalem budou výšky větší než 195 cm a menší nebo rovny 200 cm,
- a konečně dvanáctým intervalem budou výšky větší než 200 cm.

Těchto 12 intervalů můžeme určit jedenácti hranicemi mezi nimi, tedy čísly:

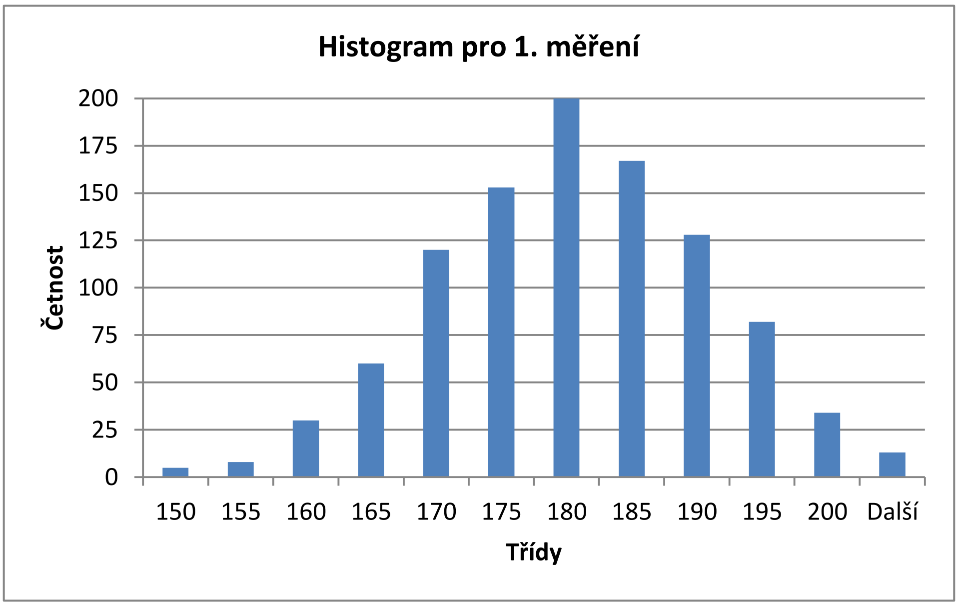
150, 155, 160, 165, 170, 175, 180, 185, 190, 195 a 200.

Tyto třídní hranice si předem připravíme do sloupečku, například do buněk AI4 až AI14. Dále aktivujeme nástroj Histogram z doplňku Analýza dat, zadáme Vstupní oblast, Hranice tříd, Výstupní oblast (například ťukneme na buňku AJ3) a zaškrtneme volbu Vytvořit graf:



Po potvrzení získáme tuto tabulku a graf, který můžeme ještě upravit do přehlednějšího tvaru:

	Třídy	*Četnost*
150	150	5
155	155	8
160	160	30
165	165	60
170	170	120
175	175	153
180	180	200
185	185	167
190	190	128
195	195	82
200	200	34
	Další	13



Úloha 3

Sestrojte si ještě histogram pro 25. měření ve sloupci AA a porovnejte jeho charakter se zkonstruovaným histogramem pro 1. měření.

Uvědomte si, že první, druhé, až dvacáté páté měření jsou vlastně **náhodné veličiny**, a že konkrétní číselné hodnoty ve vašem vyznačeném řádku, tvořící výběrový soubor, jsou pouze jejich realizacemi.

1.1.3 Důležité statistiky pro odhady parametrů

Veličinám, které se počítají z jednotlivých měření pomocí vhodných vzorců, říkáme *statistiky*. Statistiky *výběrový průměr* a *výběrová směrodatná odchylka* se počítají podle těchto vzorců:

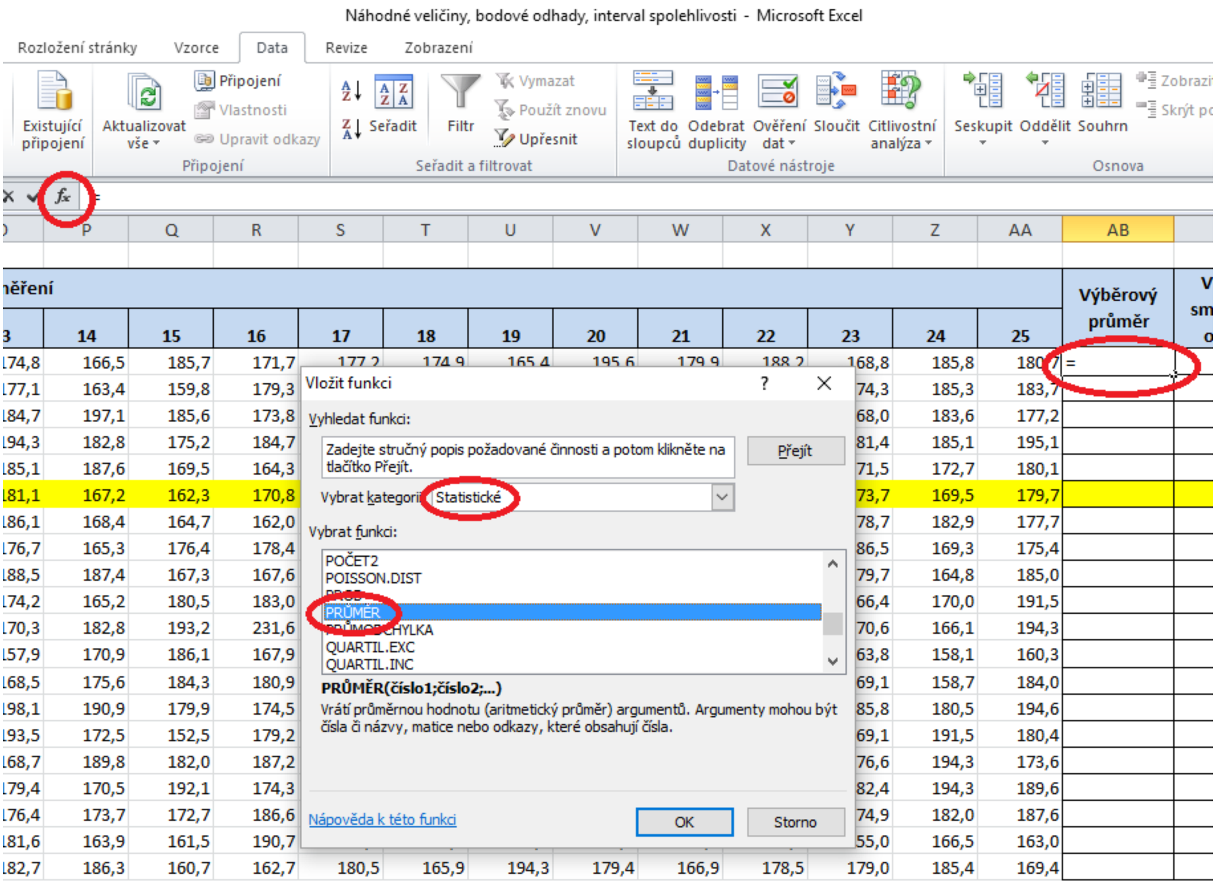
$$\overline{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i \qquad S = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \overline{X})^2}$$

V dalším textu ukážeme, že hodnota výběrového průměru je vhodným odhadem střední hodnoty μ normálního rozdělení, a že hodnota výběrové směrodatné odchylky je vhodným odhadem směrodatné odchylky σ normálního rozdělení.

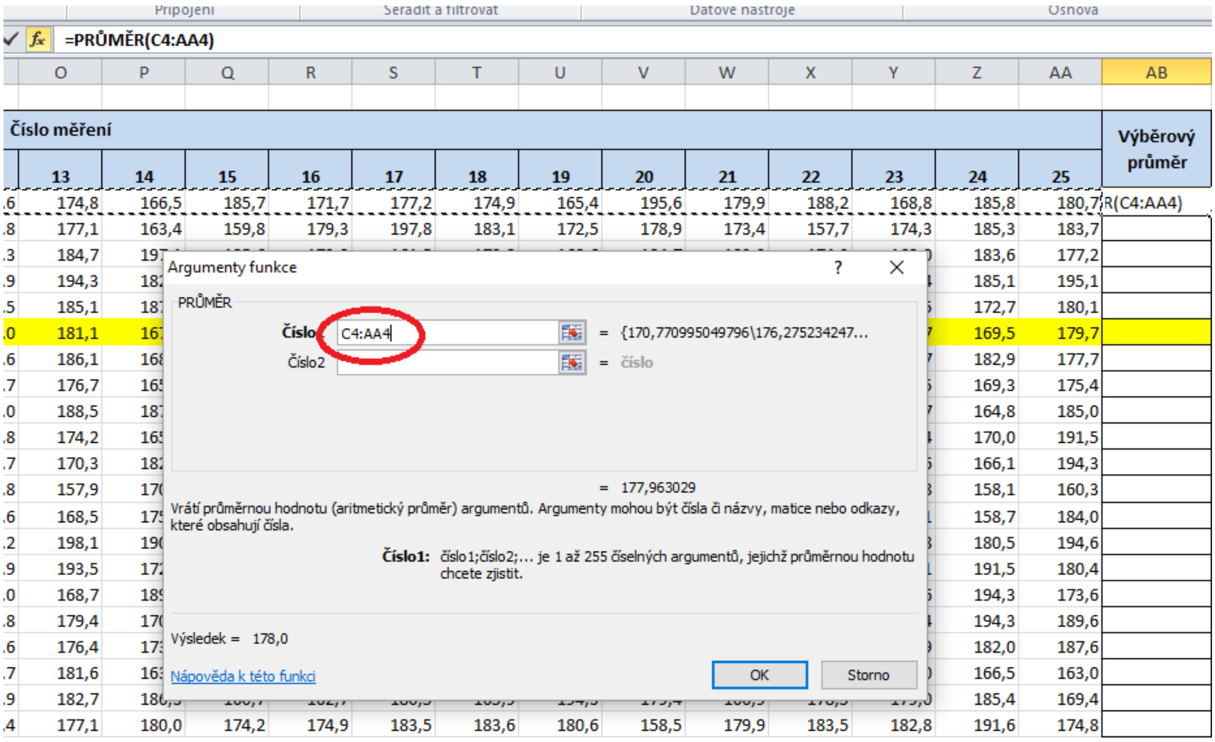
Pro výpočet těchto dvou statistik budeme používat excelovské statistické funkce PRŮMĚR a SMODCH.VÝBĚR.S, které se opět aktivují ikonou f_x na začátku příkazového řádku a výběrem potřebné funkce.

Nejprve vypočítáme výběrový průměr a výběrovou směrodatnou odchylku pro první výběrový soubor v prvním řádku tabulky (buňky C4 až AA4). Výsledky pak uložíme do připravených buněk AB4 a AC4.

Označíme tedy buňku AB4 a aktivujeme statistickou funkci PRŮMĚR:



Po potvrzení OK se objeví následující nabídka, kde musíme pouze zadat, z kterých hodnot se má průměr počítat (zadáme to tak, že jen myší označíme příslušný blok buněk):



Po potvrzení OK získáme v buňce AB4 hodnotu výběrového průměru 178,0 (je zaokrouhlena na jedno desetinné místo).

Zcela analogicky můžeme získat v buňce AC4 hodnotu výběrové směrodatné odchylky pro první řádek tabulky – vyjde 11,4.

Označíme-li si dvojblok buněk AB4 a AC4, pak roztáhnutím dvojbloku dolů vypočteme hodnoty obou těchto statistik i pro všechny další řádky, tedy pro všech 1000 výběrů jednotlivých výzkumníků. Několik prvních hodnot vypadá takto:

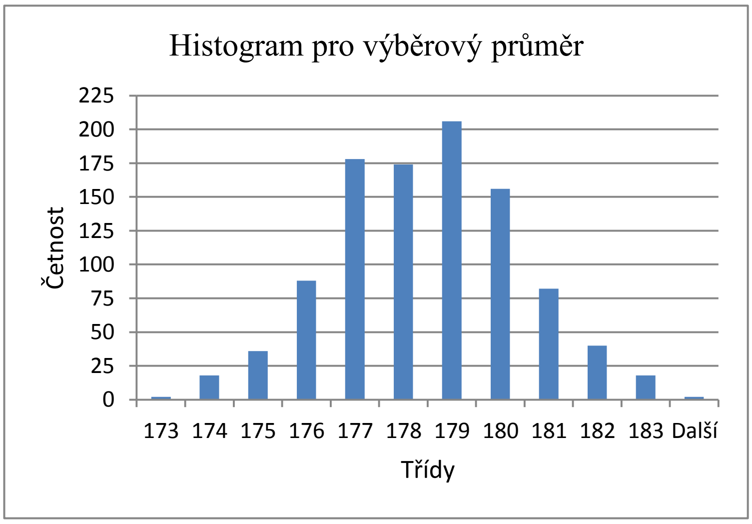
Y	Z	AA	AB	AC	AD
23	24	25	Výběrový průměr	Výběrová směrodatná odchylka	Delta
168,8	185,8	180,7	178,0	11,4	
174,3	185,3	183,7	178,6	10,7	
168,0	183,6	177,2	179,3	10,3	
181,4	185,1	195,1	178,5	10,2	
171,5	172,7	180,1	176,1	9,7	
173,7	169,5	179,7	177,7	8,3	
178,7	182,9	177,7	177,3	9,5	
186,5	169,3	175,4	176,2	9,2	
179,7	164,8	185,0	177,5	9,0	
166,4	170,0	191,5	175,4	8,9	

Protože hodnoty jednotlivých měření závisely na náhodě (jednotlivá měření jsou náhodné veličiny), dá se očekávat, že i statistiky z nich počítané, tedy výběrový průměr a výběrová směrodatná odchylka budou náhodnými veličinami. Je otázka, jaké budou mít tyto statistiky rozdělení pravděpodobnosti – k tomu bude potřeba zkonstruovat jejich histogramy (podobně, jako jsme to dělali pro jednotlivá měření).

Prozradíme předem, že z teorie matematické statistiky plyne, že výběrový průměr má zase normální rozdělení, jeho střední hodnota je stejná jako střední hodnota původní veličiny (v našem případě tedy 178), ale směrodatná odchylka je $\sqrt{n}$ –krát menší než směrodatná odchylka původní veličiny (přitom n je rozsah výběrového souboru). V našem případě byla směrodatná odchylka původní veličiny 10, rozsah výběru je 25, tedy směrodatná odchylka výběrového průměru bude $\frac{10}{\sqrt{25}} = \frac{10}{5} = 2$. Podle pravidla 3σ budou tedy skoro všechny hodnoty výběrového průměru ležet v intervalu od hodnoty 178 – 3.2 = 172 do hodnoty 178 + 3.2 = 184.

Pro konstrukci histogramu pro výběrový průměr můžeme tedy za hranice tříd zvolit hodnoty 173, 174, 175, 176, 177, 178, 179, 180, 181,182 a 183. Těchto jedenáct hranic nám definuje dvanáct třídních intervalů – histogram bude mít 12 sloupečků. Získáme tyto výsledky:

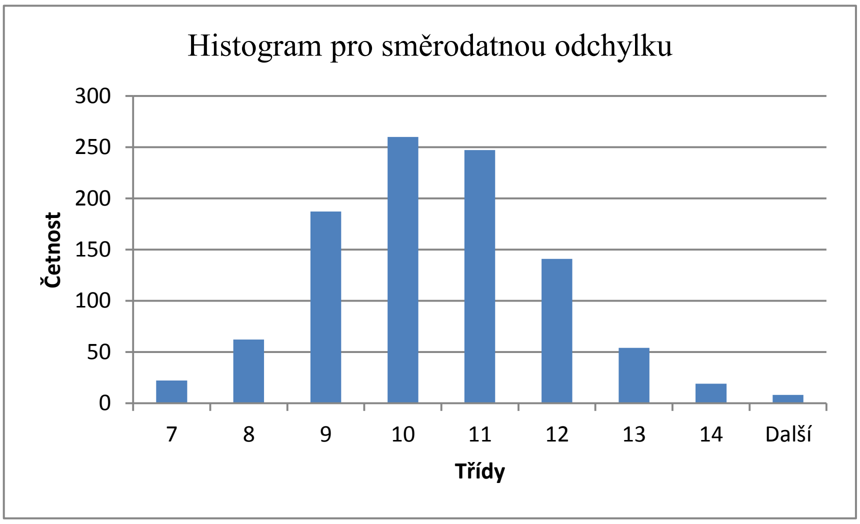
Třídy	*Četnost*
173	2
174	18
175	36
176	88
177	178
178	174
179	206
180	156
181	82
182	40
183	18
Další	2



Úloha 4

Sestrojte si ještě histogram pro výběrovou směrodatnou odchylku (hranice tříd 7, 8, …, 14).

Výsledek:



Předcházející histogramy nám ukazují, že hodnoty výběrových průměrů v jednotlivých řádcích (výběrových souborech) se pohybují „poblíž“ hodnoty parametru $\mu = 178$, a že hodnoty výběrových směrodatných odchylek se pohybují „poblíž“ hodnoty parametru $\sigma = 10$ normálního rozdělení (což je model pro základní soubor).

Náhodná veličina **výběrový průměr** je
vhodným bodovým odhadem **střední hodnoty** μ normálního rozdělení.
Náhodná veličina **výběrová směrodatná odchylka** je
vhodným bodovým odhadem **směrodatné odchylky** σ normálního rozdělení.

To prakticky znamená, že provádí-li výzkumník (pochopitelně) jen jeden náhodný výběr, může očekávat, že skutečná (jemu neznámá) hodnota střední hodnoty μ v celé populaci bude někde „poblíž“ jím zjištěné hodnoty výběrového průměru. Podobně může očekávat, že skutečná hodnota směrodatné odchylky σ v celé populaci bude někde „poblíž“ jím zjištěné hodnoty výběrové směrodatné odchylky.

Úloha 5

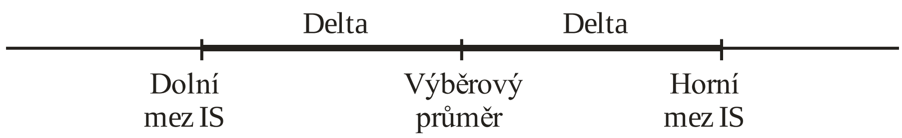
- Podívejte se do vašeho předem vyznačeného řádku v tabulce fiktivních dat a zjistěte, jak velkých chyb jste se dopustili při odhadování střední hodnoty a směrodatné odchylky v celé populaci.
- Nalezněte výzkumníky, kteří měli větší štěstí a dopustili se menších chyb, a nalezněte také výzkumníky, kteří měli větší smůlu a mají vůči vám odhady nepřesnější.

1.1.4 Interval spolehlivosti pro střední hodnotu normálního rozdělení

Víme-li již, že střední hodnota μ (u celé populace) je poblíž námi zjištěné hodnoty výběrového průměru, je přirozené se ptát, „jak daleko“ od výběrového průměru může μ být.

Odpověď na tuto otázku přináší konstrukce tzv. *intervalu spolehlivosti*, který má tu vlastnost, že **pokrývá neznámou hodnotu μ s pravděpodobností $1 - \alpha$**, kde α je malé číslo. Standardně se volí $\alpha = 0{,}10$, resp. $\alpha = 0{,}05$, resp. $\alpha = 0{,}01$, pak hovoříme o 90% intervalu spolehlivosti, resp. o 95% intervalu, resp. o 99% intervalu spolehlivosti.

Středem tohoto intervalu je výběrový průměr a číslo Delta říká, o kolik se střední hodnota μ může (s danou spolehlivostí) od výběrového průměru lišit na jednu či druhou stranu. Dolní mez intervalu spolehlivosti se tedy získá odečtením Delta a horní mez přičtením Delta.



Klíčový je tedy výpočet čísla Delta, které je polovinou délky intervalu spolehlivosti. Vypočítáme ho pomocí výběrové směrodatné odchylky S a rozsahu výběru n takto:

$$\delta = t_{1-\frac{\alpha}{2}}(n-1).\frac{s}{\sqrt{n}}\,.$$

Číslo $t_{1-\frac{\alpha}{2}}(n-1)$, které se ve výrazu vyskytuje, je kvantil *tzv. Studentova rozdělení (rozdělení t)*, který nalezneme ve statistických tabulkách či pomocí Kalkulátoru ve Statistice.

Vypočítejme teď například 95% interval spolehlivosti pro střední hodnotu μ pro našeho prvního výzkumníka (zaplníme tedy buňky AD4 až AG4). Chceme-li tedy 95% interval, máme volbu α = 0,05. Odtud plyne, že $\frac{\alpha}{2} = 0{,}025$, a tedy index kvantilu je $1 - \frac{\alpha}{2} = 0{,}975$. Druhé číslo n – 1 určující kvantil se nazývá *stupeň volnosti* a v našem případě je 25 – 1 = 24. Ve statistických tabulkách pak nalezneme v Tabulce VI příslušnou hodnotu kvantilu, tedy 2,064.

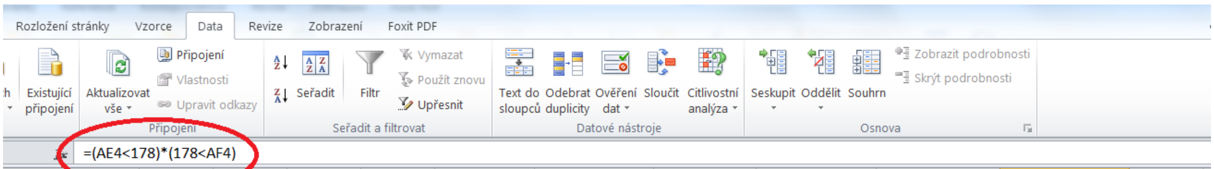
Tabulka VI. **Kvantily rozdělení *t***

	P				
ν	0,90	0,95	0,975	0,99	0,995
1	3,078	6,314	12,706	31,821	63,657
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,553	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,355
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898
18	1,330	1,734	2,101	2,552	2,878
19	1,328	1,729	2,093	2,539	2,861
20	1,325	1,725	2,086	2,528	2,845
21	1,323	1,721	2,080	2,518	2,831
22	1,321	1,717	2,074	2,508	2,819
23	1,319	1,714	2,069	2,500	2,807
24	1,318	1,711	2,064	2,492	2,797
25	1,316	1,708	2,060	2,485	2,787
26	1,315	1,706	2,056	2,479	2,779
27	1,314	1,703	2,052	2,473	2,771

Pokud použijeme nám již známý kalkulátor v programu Statistica, můžeme získat ještě přesnější hodnotu, a to 2,063899. Ověřte si to. My budeme dále počítat s tabelovanou hodnotou. Nyní již můžeme vypočítat příslušné číslo Delta a umístit ho do buňky AD4, pak odečtením od výběrového průměru a přičtením získat meze intervalu v buňkách AE4 a AF4. Získáme tyto hodnoty:

	Y	Z	AA	AB	AC	AD	AE	AF	AG
	23	24	25	Výběrový průměr	Výběrová směrodatná odchylka	Delta	Dolní mez IS	Horní mez IS	Vyhodnocení
2	168,8	185,8	180,7	178,0	11,4	4,7	173,3	182,7	
7	174,3	185,3	183,7	178,6	10,7				
9	168,0	183,6	177,2	179,3	10,3				
4	181,4	185,1	195,1	178,5	10,2				
6	171,5	172,7	180,1	176,1	9,7				
1	173,7	169,5	179,7	177,7	8,3				
8	178,7	182,9	177,7	177,3	9,5				
2	186,5	169,3	175,4	176,2	9,2				

Vypočítaný interval spolehlivosti od 173,3 do 182,7 pokrývá střední hodnotu μ = 178, první výzkumník byl tedy úspěšný. Abychom se rychle zorientovali v tom, jak jsou na tom i ostatní výzkumníci, zaplníme ještě buňku AG4 logickou hodnotou tak, že do příkazového řádku dáme instrukci, která vyhodnotí, že platí současně dvě podmínky – dolní mez je menší než 178 a horní mez je větší než 178. Oba výrazy v závorce se logicky vyhodnotí (nepravda … 0, resp. pravda … 1) a vynásobením zaručíme, že výsledná hodnota 1 se v buňce AG4 objeví právě tehdy, když obě meze svírají hodnotu 178 mezi sebou.



V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG
20	21	22	23	24	25	Výběrový průměr	Výběrová směrodatná odchylka	Delta	Dolní mez IS	Horní mez IS	Vyhodnocení
195,6	179,9	188,2	168,8	185,8	180,7	178,0	11,4	4,7	173,3	182,7	1
178,9	173,4	157,7	174,3	185,3	183,7	178,6	10,7				
184,7	183,9	174,9	168,0	183,6	177,2	179,3	10,3				

Teď již můžeme čtyřblok AD4 až AG4 označit a roztažením dolů až po poslední řádek zjistit, jak jsou na tom ostatní výzkumníci. Několik prvních řádků má tyto výsledky:

Z	AA	AB	AC	AD	AE	AF	AG
24	25	Výběrový průměr	Výběrová směrodatná odchylka	Delta	Dolní mez IS	Horní mez IS	Vyhodnocení
185,8	180,7	178,0	11,4	4,7	173,3	182,7	1
185,3	183,7	178,6	10,7	4,4	174,2	183,0	1
183,6	177,2	179,3	10,3	4,2	175,0	183,5	1
185,1	195,1	178,5	10,2	4,2	174,3	182,7	1
172,7	180,1	176,1	9,7	4,0	172,1	180,1	1
169,5	179,7	177,7	8,3	3,4	174,3	181,1	1
182,9	177,7	177,3	9,5	3,9	173,4	181,2	1
169,3	175,4	176,2	9,2	3,8	172,4	180,0	1
164,8	185,0	177,5	9,0	3,7	173,8	181,3	1
170,0	191,5	175,4	8,9	3,7	171,7	179,0	1
166,1	194,3	179,7	15,2	6,3	173,4	186,0	1
158,1	160,3	175,5	10,5	4,3	171,1	179,8	1
158,7	184,0	179,2	10,3	4,2	174,9	183,4	1
180,5	194,6	179,3	10,2	4,2	175,1	183,5	1

Úloha 6

Sečtěte všechna čísla ve sloupci Vyhodnocení. Co nám říká toto číslo? Výsledek: 959

Dolní a horní mez intervalu spolehlivosti pro střední hodnotu μ se vypočítají podle vzorců

$$DM = \bar{X} - \delta\,, \qquad HM = \bar{X} + \delta\,, \qquad \text{kde} \quad \delta = t_{1-\frac{\alpha}{2}}(n-1).\frac{S}{\sqrt{n}}\,.$$

Pak pravděpodobnost toho, že platí $\left(\bar{X} - \delta < \mu < \bar{X} + \delta\right)$ je rovna číslu 1 – α .

1.1.5 Problematika rozsahu náhodného výběru

Z poslední tabulky můžeme zjistit, že polovina délky 95% intervalu spolehlivosti pro μ byla u všech výzkumníků přibližně δ = 4 cm, tedy že hodnota výběrového průměru se od hledané hodnoty μ s pravděpodobností 95% neliší více než o 4 cm. To není velká přesnost, je to ale způsobeno tím, že rozsah náhodného výběru byl pouze $n = 25$.

Vzhledem k tomu, že ve vzorci pro číslo δ se ve jmenovateli zlomku vyskytuje $\sqrt{n}$, dá se očekávat, že s rostoucím rozsahem výběru n se bude délka intervalu spolehlivosti zkracovat. Přirozeně se vynořuje otázka, jak velký by měl být rozsah výběru, abychom získali odhad pro μ s předem danou velikostí chyby, například pro náš případ velikostí lidí třeba 1 cm.

Pak by muselo platit, že $\delta = t_{1-\frac{\alpha}{2}}(n-1).\frac{S}{\sqrt{n}} < 1$. Nalézt takové n, které splňuje tuto nerovnici, by bylo obtížné, ale může nám pomoci tento fakt: pro větší hodnoty n jsou kvantily rozdělení t přibližně stejně velké, jako kvantily u_P normálního rozdělení se střední hodnotou 0 a směrodatnou odchylkou 1 (tzv. *normované normální rozdělení*).

Úloha 7

Potvrďte si pravdivost poslední věty tak, že pomocí Kalkulátorů ve Statistice naleznete hodnoty kvantilů $t_{0,95}(300)$, $t_{0,975}(300)$, $t_{0,995}(300)$ a $u_{0,95}$, $u_{0,975}$, $u_{0,995}$.

Výsledky:	$t_{0,95}(300) = 1{,}649949$	a	$u_{0,95}, = 1{,}644854$,
	$t_{0,975}(300) = 1{,}967903$	a	$u_{0,975} = 1{,}959964$,
	$t_{0,995}(300) = 2{,}592316$	a	$u_{0,995} = 2{,}575829$.

Vzhledem k tomuto faktu, a protože předpokládáme, že rozsah výběru vyjde velký, můžeme naši nerovnici nahradit jednodušší nerovnicí: $u_{1-\frac{\alpha}{2}}.\frac{S}{\sqrt{n}} < 1$.

Pro 95% interval spolehlivosti tedy máme $1{,}96.\frac{S}{\sqrt{n}} < 1$,

a závěrem $n > (1{,}96.S)^2$.

Tedy k tomu, aby výzkumník se spolehlivostí 95% odhadl svým výběrovým průměrem parametr μ s chybou menší než 1 cm, potřeboval by uskutečnit náhodný výběr s rozsahem větším, než je číslo $(1{,}96.S)^2$. Například pro prvního výzkumníka, kde je výběrová směrodatná odchylka $S = 11{,}4$, by potřebný rozsah výběru byl

$$(1{,}96\,.\,11{,}4\,)^2 = (22{,}344\,)^2 = 499{,}254336.$$

Místo 25 lidí by jich tedy musel změřit asi 500. Odhad rozsahu souboru u dalších výzkumníků by pochopitelně byl jiný, protože z jejich měření vyšly jiné hodnoty výběrových směrodatných odchylek.

Při reálných výběrových šetřeních, kde si předem stanovujeme potřebnou přesnost měření, tedy pro plánování rozsahu výběru potřebujeme předem znát odhad velikosti výběrové směrodatné odchylky – ten obvykle získáváme menším pilotním šetřením.

1.2 Organizace dat, datový soubor, typy náhodných veličin

Pro počítačové zpracování je velmi důležité všechna naměřená data vhodně uspořádat.

Data vždy uspořádáme tak, že každé **statistické jednotce** z výběrového souboru bude v tabulce příslušet jeden **řádek** a každé zkoumané či **měřené veličině** pak jeden **sloupec**.

Důležité je také vhodné označení jednotlivých veličin.

Všechna naměřená data budou tedy obsažena jen v jedné tabulce.

Typy sledovaných náhodných veličin

Náhodné veličiny dělíme podle možností práce s jejich hodnotami na *nominální*, *ordinální* a *metrické*, resp. podle druhu teoretického rozdělení pravděpodobnosti jejich modelu v základním souboru na *diskrétní* a *spojité*.

- Hodnoty **nominálních** veličin nelze ani uspořádávat ani s nimi aritmeticky operovat (tedy sčítat je, násobit, dělit, atd.)
  Příklady: pohlaví, barva květu, druh rostliny, plemeno u skotu, atd.

- Hodnoty **ordinálních** veličin lze uspořádávat a porovnávat, ale nelze s nimi aritmeticky operovat.
  Příklady: věková kategorie, vědecká hodnost, známka z předmětu, stupeň poškození hnilobou, atd.

- Hodnoty **metrických (kvantitativních)** veličin lze jak uspořádávat, tak s nimi lze aritmeticky operovat.
  Příklady: hmotnost, délka, čas, počet písmen ve slově, výše platu státních zaměstnanců, atd.

Nominální a ordinální veličiny nazýváme souhrnně *veličiny kvalitativní (*resp. *kategoriální).*

Je třeba poznamenat, že se někdy tyto výše uvedené typy nesnadno odlišují a někdy se tomuto striktnímu rozlišování zpronevěřujeme (například se počítá průměr známek žáka, apod.)

- **Diskrétní** náhodné veličiny nabývají jen konečně mnoha, resp. spočetně mnoha (obvykle celočíselných) hodnot.
- **Spojité** náhodné veličiny mohou teoreticky nabývat nespočetně mnoha reálných hodnot, které zjišťujeme s předem danou přesností.

Úloha 8

Prohlédněte si uspořádání dat v souboru Data_deti_min a rozhodněte, jakých typů jsou jednotlivé veličiny.

Výsledky:	pohlaví	nominální, diskrétní
	třída, známka, výkon	ordinální, diskrétní
	výška, hmotnost, BMI	metrická, spojitá
	atd.	

1.3 Princip testování statistických hypotéz, ukázka testování

Testování statistických hypotéz je jednou z nejčastěji používaných metod. Statistických testů je velmi mnoho, s nejdůležitějšími z nich se seznámíte v dalších tématech. Vždy je ale třeba dbát na to, abychom pro danou situaci použili „správný“ test – použití každého testu je vždy vázáno na splnění určitých předpokladů (typ veličiny, normalita či nenormalita rozdělení v základním souboru, atd.)

U všech testů se však objevují tytéž základní myšlenky a při testování se užívají tytéž základní pojmy – univerzální princip testování, s kterým se zde seznámíme, je společný všem testům.

1.3.1 Problém kuliček v osudí

Máme před sebou tento problém: víme zcela jistě, že v osudí máme deset kuliček, ale nevíme, která z následujících variant je pravdivá:

buď je tam pět kuliček bílých a pět kuliček černých,

anebo tam je osm bílých kuliček a dvě černé kuličky.

Nemůžeme se do osudí podívat a spatřit všechny kuličky najednou, smíme ale opakovaně vytahovat jednu kuličku, zjistit její barvu, a pak ji do osudí zase vrátit. Počet vytažení jedné kuličky není nijak omezen. Intuitivně cítíme, že pro druhou variantu (osm bílých a dvě černé) bychom se rozhodli, když budeme bílou kuličku vytahovat častěji, je ale zřejmé, že i při této variantě se může stát, že „náhodou“ budeme vytahovat pouze černé kuličky. Navrhneme teď vhodný postup pro rozhodování o tom, která varianta počtů kuliček v osudí je ve skutečnosti pravdivá, a tento postup prakticky ověříme.

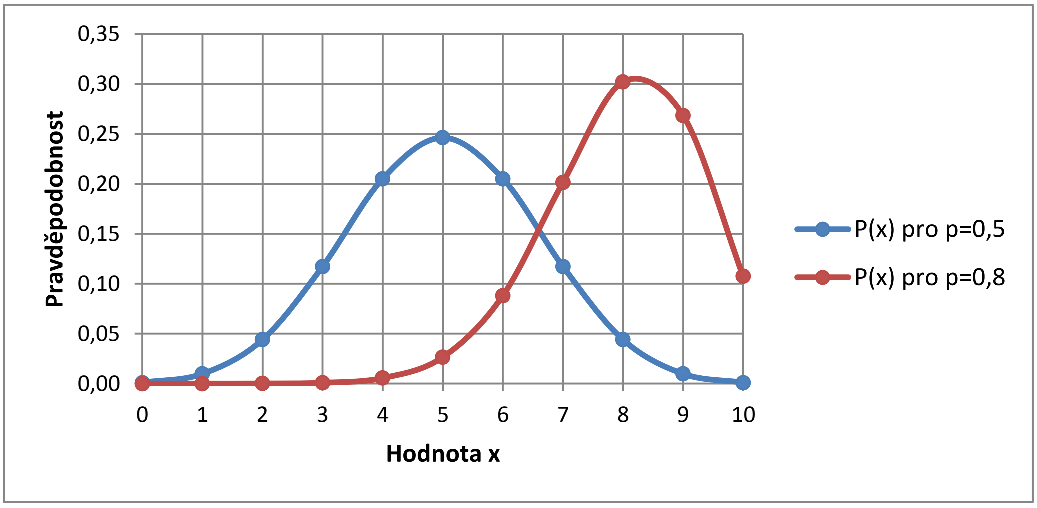
Máme se tedy rozhodnout mezi dvěma hypotézami, nazveme je *nulová* a *alternativní*:

nulová hypotéza H_0 : pravděpodobnost vytažení bílé kuličky z osudí je $p = 0{,}5$ a

alternativní hypotéza H_a : pravděpodobnost vytažení bílé kuličky z osudí je $p = 0{,}8$.

Z osudí budeme opakovaně vytahovat vždy jednu kuličku (a zase ji budeme do osudí vracet). Takových vytažení uskutečníme n. Mezi hypotézami se rozhodneme podle toho, jaké hodnoty nabude náhodná veličina „počet vytažených bílých kuliček při n pokusech“ – tuto veličinu nazveme *testovým kritériem*. Bude-li **testové kritérium** nabývat „relativně malých hodnot“, přikloníme se k nulové hypotéze, bude-li nabývat „relativně velkých hodnot“, zamítneme nulovou hypotézu a přikloníme se k alternativní hypotéze. Co jsou ale „malé“ a „velké“ hodnoty? Tady nám musí pomoci teorie.

Otevřete si cvičný excelovský soubor Princip testování statistických hypotéz. V prvním jeho listu s názvem n = 10 je tabulka, která obsahuje pravděpodobnosti toho, že počet x vytažených bílých kuliček bude roven 0, 1, 2, atd. až 10, a to pro obě možné situace (když platí nulová hypotéza, resp. když platí alternativní hypotéza). Hodnoty z tabulky jsou znázorněny na tomto grafu:



Protože se potřebujeme mezi hypotézami rozhodnout jednoznačně, bude vhodné obor všech hodnot, které může testové kritérium nabývat (tedy množinu 0, 1, 2, až 10), rozdělit na dvě části: na „velké“ hodnoty tvořící tzv. *kritický obor W*, a „malé“ hodnoty tvořící tzv. *doplňkový obor*. Je přirozené obory volit tak, že rozhraní mezi nimi bude určovat průsečík křivek v grafu, tedy: kritickým oborem *W* bude množina čísel 7, 8, 9 a 10,
doplňkovým oborem bude množina čísel 0, 1, 2, 3, 4, 5 a 6.

Připomeňme, že kritický obor je „kritický“ pro nulovou hypotézu H_0, jestliže tedy testové kritérium nabyde hodnoty alespoň 7 (vytáhli jsme při 10 pokusech alespoň 7 bílých kuliček), přestaneme věřit, že platí nulová hypotéza (tedy že v osudí je pět bílých a pět černých kuliček). Naopak, pokud testové kritérium nabude hodnoty z doplňkového oboru, nemáme důvod zamítat nulovou hypotézu.

Co je ale nejdůležitější – při testování se sice můžeme dopustit chyb, ale jsme schopni pravděpodobnost těchto chyb vypočítat a test upravit tak, aby pravděpodobnost chybování byla „dostatečně malá“.

Úloha 9

Vyznačte si v horní tabulce barevně řádky příslušné kritickému oboru a pomocí funkce SUMA si vypočtěte čtyři pravděpodobnosti z dolní tabulky (jedná se o součty čísel v blocích C10 až C13, C3 až C9, atd.)

Výsledky:

Pravděpodobnost	Platí H_0: p=0,5	Platí H_a: p=0,8
x náleží do W	0,17187500	0,87912612
x nenáleží do W	0,82812500	0,12087388

Čísla v červeně označených polích tabulky jsou pravděpodobnosti chyb při testování. Proč?

Číslo vlevo nahoře je pravděpodobnost toho, že nulová hypotéza platí, a přitom testové kritérium náleží do kritického oboru, což znamená, že nulovou hypotézu zamítneme. Jedná se o tzv. *chybu prvního druhu*, která spočívá v nesprávném zamítnutí nulové hypotézy, která ve skutečnosti platí. V našem případě je pravděpodobnost chyby prvního druhu asi 17,2 %.

Číslo vpravo dole je pravděpodobnost toho, že nulová hypotéza neplatí (platí alternativní hypotéza), a přitom testové kritérium nenáleží do kritického oboru, což znamená, že nulovou hypotézu nezamítneme. Jedná se o tzv. *chybu druhého druhu*, která spočívá v nesprávném nezamítnutí nulové hypotézy, která ve skutečnosti neplatí. V našem případě je pravděpodobnost chyby druhého druhu asi 12,1 %.

Pravděpodobnost obou chyb je příliš velká, takovýto test by nebyl vhodný. Kdybychom se snažili test vylepšit tak, že bychom hranici mezi kritickým a doplňkovým oborem nestanovovali podle průsečíku křivek, ale jen ji posunuli doleva či doprava, pak by se sice pravděpodobnost jedné chyby mohla zmenšit, ale pravděpodobnost druhé chyby by se pochopitelně zvětšila. Posunutí hranice kritického oboru tedy situaci nevyřeší. Naštěstí se však ukáže, že když testovací postup změníme tak, že z osudí budeme kuličky opakovaně tahat vícekrát, speciálně 20krát, resp. 30krát, situace se zlepší.

Úloha 10

Využijte podkladů v dalších listech souboru Princip testování statistických hypotéz, vyznačte si kritické obory a vypočítejte všechny pravděpodobnosti v malých tabulkách, speciálně pravděpodobnosti chyb prvního a druhého druhu.

Výsledky: pro $n = 20$: W je množina čísel 14, 15, 16, až 20.

Pravděpodobnost	Platí H_0: p=0,5	Platí H_a: p=0,8
x náleží do W	0,05765915	0,91330749
x nenáleží do W	0,94234085	0,08669251

pro $n = 30$: W je množina čísel 20, 21, 22, až 30.

Pravděpodobnost	Platí H_0: p=0,5	Platí H_a: p=0,8
x náleží do W	0,04936857	0,97438374
x nenáleží do W	0,95063143	0,02561626

Podle závažnosti důsledků testování hypotéz volíme předem velikost pravděpodobnosti chyb při testování. Protože pravděpodobnost chyby druhého druhu v reálných testech, kdy obě hypotézy nejsou tak jednoduché, jako v našem případě, se počítá obtížně (souvisí to s tzv. *silou testu*, o čemž se více dozvíte v dalších tématech), soustřeďujeme se obvykle jen na chybu prvního druhu. Její pravděpodobnost se značí α , nazývá se *hladina významnosti* a standardně se užívají hodnoty $\alpha = 0{,}10$, resp. $\alpha = 0{,}05$, resp. $\alpha = 0{,}01$. Podle výsledků předchozí úlohy 10 bychom mohli nulovou hypotézu testovat na hladině významnosti 10% pro $n = 20$ a na hladině významnosti 5% pro $n = 30$.

Úloha 11

Udělejte reálný experiment s osudím. Situaci zorganizujte tak, aby žádný z účastníků kurzu (ani lektor) nevěděl, kolik bílých kuliček je v osudí. Postupně pak vytahujte dvacetkrát, resp. třicetkrát kuličku z osudí a podle výsledku úlohy 10 testujte nulovou hypotézu. Pohledem do osudí pak porovnejte skutečnou platnost nulové hypotézy s vaším závěrem.

Shrňme si postup při testování a při něm používané pojmy:

- Zkoumanou hypotézu formulujeme „v řeči pravděpodobnosti či statistiky“ jako tzv. nulovou hypotézu H_0.
- Proti této hypotéze stavíme tzv. alternativní hypotézu H_a, která může, ale nemusí být negací nulové hypotézy.
- Nalezneme vhodnou náhodnou veličinu *G*, kterou budeme nazývat *testovým kritériem*.
- Zvolíme malé kladné číslo α (bývá zvykem volit zejména hodnoty $\alpha = 0{,}10$, resp. $\alpha = 0{,}05$, resp. $\alpha = 0{,}01$), které nazýváme hladinou významnosti.
- Určíme tzv. kritický obor *W* tak, aby platilo:

  Jestliže platí hypotéza H_0, pak pravděpodobnost toho, že $G \in W$, je rovna číslu α.
- Z tohoto faktu pak plyne závěrečná interpretace testu:

  zjistíme hodnotu testového kritéria, a

  nastane-li případ, že $G \in W$, pak nulovou hypotézu H_0 zamítneme,

  nastane-li případ, že $G \notin W$, pak nulovou hypotézu H_0 nezamítneme.

Při tomto způsobu testování hypotéz se můžeme dopustit dvojího druhu chyb, jak ukazuje následující tabulka:

Druhy chyb při testování		**Skutečnost** nulová hypotéza H_0 **platí**	**Skutečnost** nulová hypotéza H_0 **neplatí**
Výsledek testu	$G \in W$ nulovou hypotézu H_0 **zamítáme**	**chyba prvního druhu**	SPRÁVNĚ
Výsledek testu	$G \notin W$ nulovou hypotézu H_0 **nezamítáme**	SPRÁVNĚ	**chyba druhého druhu**

Chyba prvního druhu znamená, že zamítneme nulovou hypotézu, i když ve skutečnosti platí. Pravděpodobnost této chyby je rovna hladině významnosti α.

Chyba druhého druhu znamená, že nulovou hypotézu nezamítneme, i když ve skutečnosti neplatí. Pravděpodobnost této chyby bývá označována β a souvisí s pojmem „síla testu“.

1.3.2 Ukázka testování (F-test)

Rozptylem normálního rozdělení se střední hodnotou μ a směrodatnou odchylkou σ nazýváme číslo σ^2 (je to tedy druhá mocnina směrodatné odchylky). Odhaduje se pomocí tzv. *výběrového rozptylu* (což je druhá mocnina výběrové směrodatné odchylky *S*), počítá se tedy pomocí vzorce

$$S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \overline{X})^2 .$$

V Excelu vypočítáme hodnotu výběrového rozptylu pomocí statistické funkce VAR.S .

Pro první ukázku testování použijeme cvičný datový excelovský soubor Koně, který obsahuje data o 53 náhodně vybraných koních, u nichž bylo zjištěno pohlaví (v tabulce jsou označeny klisny růžově a hodnota Pohlaví je rovna 0, hřebci jsou označeni modře a hodnota Pohlaví je rovna 1) a změřeny další tři metrické údaje: kohoutková výška, sedlová výška a délka pánve. Budeme předpokládat, že náhodná veličina kohoutková výška má v základním souboru normální rozdělení. Protože ale nevíme, zda výška v kohoutku bude mít u klisen a u hřebců stejné parametry, bude třeba uvažovat dvě normální rozdělení:

normální rozdělení pro klisny s parametry μ_1 a σ_1^2

a normální rozdělení pro hřebce s parametry μ_2 a σ_2^2.

Bude nás zajímat otázka, jestli mají tyto dvě náhodné veličiny v základním souboru stejný rozptyl (tedy stejnou i směrodatnou odchylku), tedy zda jsou jim příslušné hustoty pravděpodobnosti „stejně široké“. Nulová hypotéza má tedy tvar: $H_0 : \sigma_1^2 = \sigma_2^2$, budeme ji testovat tzv. *F-testem.*

Nejprve se pokusíme neznámá čísla σ_1^2 a σ_2^2 odhadnout – jak bylo výše uvedeno – pomocí hodnot výběrových rozptylů. Použijeme-li dvakrát funkci VAR.S, získáme tyto odhady:

neznámý rozptyl klisen σ_1^2 leží „poblíž“ hodnoty výběrového rozptylu S_1^2 = 8,705645

a neznámý rozptyl klisen σ_2^2 leží „poblíž“ hodnoty výběrového rozptylu S_2^2 = 5,747619.

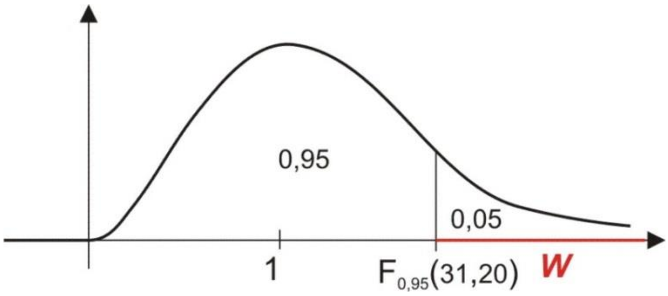
Poznámka: Vzhledem k dalšímu postupu je vhodné, aby první výběrový rozptyl S_1^2 byl větší než druhý výběrový rozptyl S_2^2, kdyby tomu tak nebylo, přečíslovali bychom oba výběrové soubory.

Zdá se tedy, že hodnota σ_1^2 by mohla být větší než hodnota σ_2^2, ale protože odhady jsou jen přibližné, nejsme si tím jisti. Proti nulové hypotéze ale přesto postavíme alternativní hypotézu ve tvaru $H_a : \sigma_1^2 > \sigma_2^2$ (je to tzv. jednostranná alternativa, oboustranná alternativa by měla tvar logické negace nulové hypotézy, tedy $\sigma_1^2 \neq \sigma_2^2$).

Jak nyní zvolit vhodné testové kritérium? Ukazuje se vhodné za testové kritérium zvolit náhodnou veličinu $G = \dfrac{{S_1}^2}{{S_2}^2}$. Je totiž zřejmé, že když platí nulová hypotéza, budou hodnoty veličiny G přibližně rovny jedné (čitatel totiž bude přibližně stejně velký jako jmenovatel), a „velké“ hodnoty G budou svědčit proti nulové hypotéze. Navíc je z teorie statistiky známo, že za předpokladu, že platí nulová hypotéza o rovnosti obou rozptylů, má tato náhodná veličina G tzv. *Fisherovo rozdělení pravděpodobnosti* $F(n_1 - 1; n_2 - 1)$, u kterého můžeme zjišťovat jeho kvantily (čísla n_1 a n_2 jsou rozsahy obou výběrů, v našem případě počty klisen a hřebců). Protože ve výběrech máme n_1 = 32 klisen a n_2 = 21 hřebců, bude mít veličina G rozdělení $F(31;20)$. Čísla 31 a 20 jsou parametry tohoto rozdělení, nazývají se opět stupně volnosti.

A jak zvolit kritický obor W? Musíme dodržet jeho základní vlastnost, tedy že:
Jestliže platí nulová hypotéza H_0, pak pravděpodobnost toho, že $G \in W$, je rovna číslu α.

Pomůže nám následující obrázek hustoty pravděpodobnosti Fisherova rozdělení (protože veličina G je podílem druhých mocnin, její hodnoty jsou tedy kladné, je hustota pravděpodobnosti v levé části vodorovné osy rovna nule):



Zvolíme-li hladinu významnosti α = 0,05 = 5 % , má kritický obor obsahovat „velké" hodnoty a přitom má hodnota G rozdělení $F(31;20)$. Stačí tedy nalézt hodnotu kvantilu $F_{0,95}(31;20)$ a kritický obor W budou tvořit hodnoty větší než tento kvantil. Pak bude zaručeno, že hodnoty G budou náležet kritickému oboru s pravděpodobností α = 0,05.

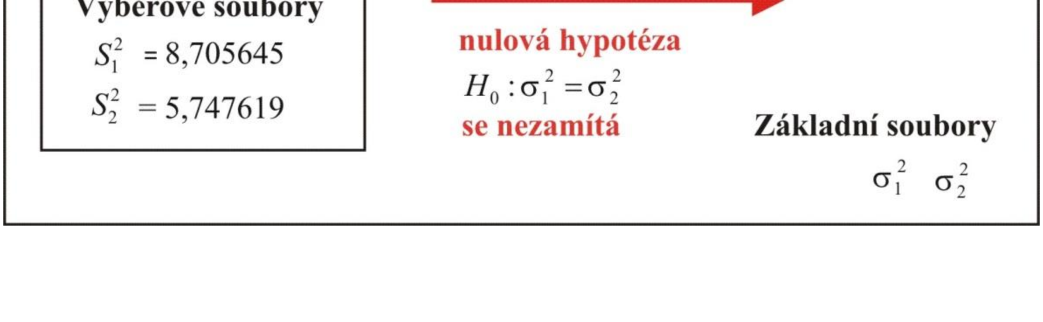
Nalezněme teď pomocí Kalkulátoru ve Statistice potřebný kvantil: je $F_{0,95}(31;20) = 2{,}033357$, a kritický obor je tedy $W = (2{,}033357\ ;\ +\infty)$.

A zbývá již jen vypočítat hodnotu testového kritéria: $G = \frac{S_1^{\,2}}{S_2^{\,2}} = \frac{8{,}705645}{5{,}747619} = 1{,}514652$.

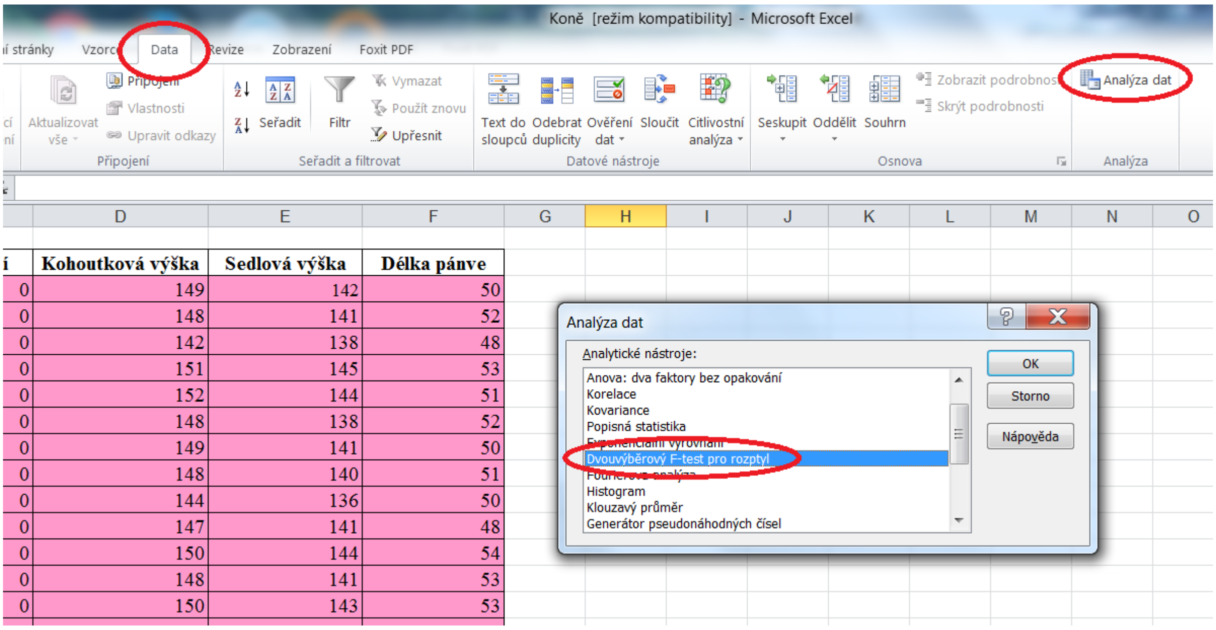
Závěr: protože testové kritérium G nenáleží kritickému oboru W, nezamítáme nulovou hypotézu $H_0 : \sigma_1^2 = \sigma_2^2$ o rovnosti obou rozptylů náhodných veličin v základních souborech (kohoutkových výšek u klisen a hřebců).

Poznámka:

Ze známých hodnot výběrových rozptylů S_1^2 a S_2^2 by se „na první pohled" mohlo zdát, že rozptyl u klisen je větší než u hřebců. Musíme si znovu uvědomit, že výběrové rozptyly charakterizují poměry ve výběrových souborech, a jsou to jen odhady neznámých rozptylů σ_1^2 a σ_2^2, o nichž v nulové hypotéze hovoříme. Chceme-li pak dostát zvolené hladině významnosti 5 %, tedy chceme-li, aby pravděpodobnost chyby při testování nepřekročila 5 % (což je dosti silná podmínka), nulovou hypotézu nezamítáme. Netvrdíme tedy, že nulová hypotéza určitě platí, jen konstatujeme, že na základě naměřených dat (výběrových souborů) nemůžeme nulovou hypotézu o poměrech v základním souboru zamítnout.



Poté, co jsme kvůli dobrému pochopení provedli celý F-test „ručně“, ukažme si, jak jej můžeme rychle zpracovat v Excelu. Budeme používat nám již známý doplněk Analýza dat a označíme, že chceme použít Dvouvýběrový F-test:



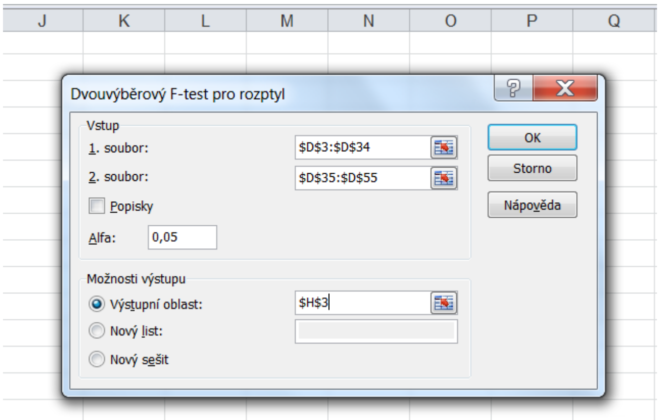
Po potvrzení zadáme do nabídky potřebné údaje:

jako 1. soubor označíme ten, který má větší výběrový rozptyl (to musíme zjistit předem!), v našem případě je to blok kohoutkových výšek klisen,

jako 2. soubor označíme blok kohoutkových výšek hřebců,

ponecháme implicitně nastavenou hladinu významnosti α

a zvolíme výstupní oblast (buňku H3).

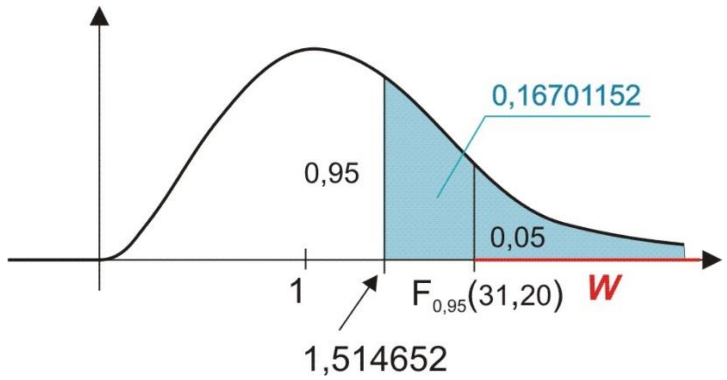


Po potvrzení nám vyjdou v tabulce všechny hodnoty, které jsme získali při ručním testování:

Dvouvýběrový F-test pro rozptyl

	Soubor 1	*Soubor 2*
Stř. hodnota	148,4375	149,047619
Rozptyl	8,70564516	5,74761905
Pozorování	32	21
Rozdíl	31	20
F	1,51465243	
P(F<=f) (1)	0,16701152	
F krit (1)	2,03335727	

Zbývá jen zjistit, jaký je význam číselné hodnoty v předposledním řádku (tzv. *p-hodnota).* Je to *„skutečně dosažená" hladina významnosti*, tedy pravděpodobnost toho, že testové kritérium *G* nabude větší hodnoty, než je vypočítaná hodnota testového kritéria 1,514652. K pochopení nám pomůže obrázek.



Protože testové kritérium nepadlo do kritického oboru, je pochopitelně *p*-hodnota větší než 0,05 (zvolená hladina významnosti určující kritický obor), pokud by testové kritérium padlo do kritického oboru (a my jsme tedy zamítli nulovou hypotézu), byla by *p*-hodnota menší než 0,05. Tak je tomu u všech testů prováděných počítačem, lze tedy (s jistou obezřetností) používat toto jednoduché pravidlo:

Je-li *p*-hodnota menší než hladina významnosti (např. 0,05), pak nulovou hypotézu zamítáme. Je-li *p*-hodnota větší než hladina významnosti, pak nulovou hypotézu nezamítáme.

1.3.3 Další ukázka testování (porovnání úspěšnosti žáků při řešení úloh)

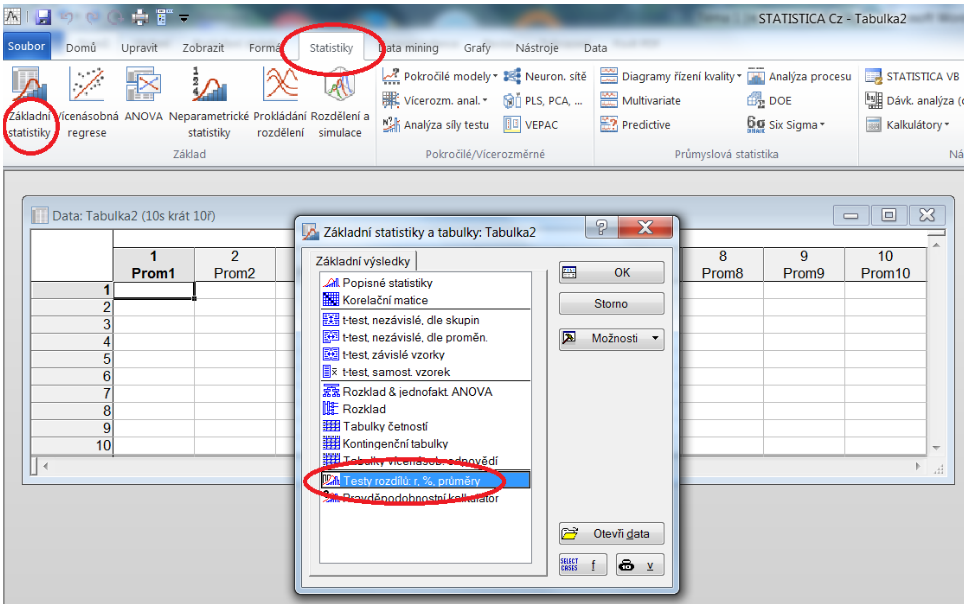
Při pedagogickém výzkumu často řešíme situace, kdy máme rozhodnout o tom, zda určitý nový postup přináší lepší výsledky, než doposud aplikovaný tradiční postup. Obvykle výzkum probíhá tak, že nový postup aplikujeme v experimentální skupině (například třídě) a tradiční postup použijeme v kontrolní skupině (pochopitelně je třeba zaručit, že obě skupiny jsou „stejně kvalitní").

Dejme tomu, že výsledky obou skupin chceme hodnotit jednoduchým způsobem – a to procentem žáků ve skupině, kteří správně vyřešili určitou úlohu. Jakým myšlenkovým postupem se můžeme dostat k závěru, že činnosti, které jsme prováděli v experimentální skupině před testováním, významně přispěly k úspěšnosti žáků při řešení úlohy?

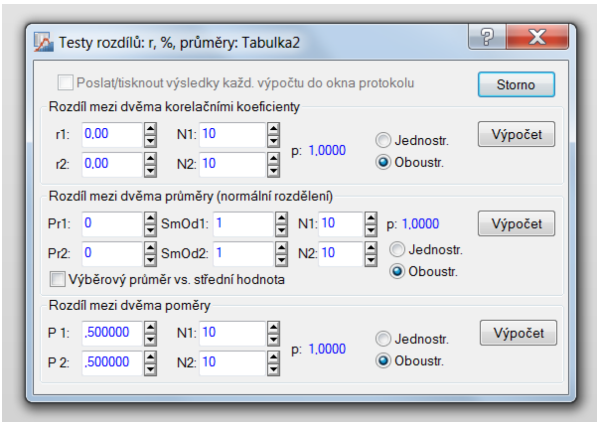
Představíme si, že experimentální i kontrolní skupina jsou náhodné výběry ze dvou hypotetických základních souborů – první z nich tvoří žáci, kteří se učí (či se budou učit) podle nového postupu, a analogicky druhý tvoří žáci, kteří se učí (či se budou učit) klasickým postupem. Budeme zřejmě testovat tuto nulovou hypotézu: pravděpodobnost toho, že žák úspěšně vyřeší danou úlohu je v obou těchto základních souborech stejná (výzkumník musí být vždy pesimista).

Pro tuto hypotézu existuje tzv. *Test rozdílu mezi dvěma poměry* – naučíme se jej používat. Nebudeme však počítat testové kritérium ani stanovovat kritický obor, to vše za nás udělá program Statistica, budeme postupovat jen podle počítačem vypočtené *p*-hodnoty tak, jak je uvedeno v závěru předchozího oddílu 1.3.2.

Náš test nalezneme cestou: Statistiky → Základní statistiky → Testy rozdílů.



Po potvrzení se objeví nabídka, do jejíž části Rozdíl mezi dvěma poměry budeme jen zadávat hodnoty potřebné pro náš test:

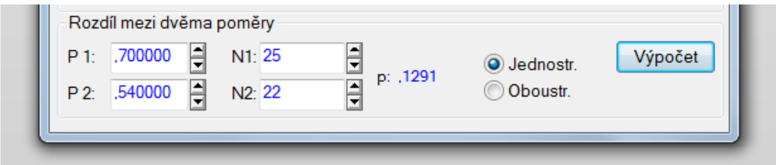


Do P1 zadáme procento úspěšnosti v experimentální skupině,
do P2 zadáme procento úspěšnosti v kontrolní skupině (přitom předpokládáme, že P1 > P2),
do N1 zadáme počet žáků v experimentální skupině,
do N2 zadáme počet žáků v kontrolní skupině.
Dále zvolíme jednostrannou alternativní hypotézu, která říká, že pravděpodobnost úspěšného vyřešení úlohy je větší v základním souboru žáků vyučovaných novým přístupem.

Úloha 12

V našem fiktivním výzkumu má experimentální třída 25 žáků a kontrolní třída 22 žáků. Úspěšnost v experimentální třídě je 70 % a v kontrolní třídě 54 %. Proveďte test nulové hypotézy a rozhodněte, zda nový postup vede k lepším výsledkům či nikoliv.

Výsledky:



Je patrné, že *p*-hodnota 0,1291 = 12,91 % je příliš vysoká, pravděpodobnost neoprávněného zamítnutí nulové hypotézy kolem 13 % je nepřípustná. I když se rozdíl v odhadech pravděpodobností jeví velký – 16 % – nemůžeme vzhledem k malému počtu žáků ve třídách nulovou hypotézu zamítnout.

Úloha 13

- Předpokládejme stejně jako v úloze 12, že v našem výzkumu je úspěšnost v experimentální skupině 70 % a v kontrolní skupině 54 %. Kolik žáků by muselo být v každé skupině (předpokládejme je stejně početné), abychom mohli nulovou hypotézu zamítnout na 5% hladině významnosti?
- Předpokládejme stejně jako v úloze 12, že v našem výzkumu má experimentální třída 25 žáků a kontrolní třída 22 žáků. Předpokládejme znovu, že úspěšnost v kontrolní skupině je 54 %. Jaká by musela být úspěšnost v experimentální skupině, abychom mohli nulovou hypotézu zamítnout na 5% hladině významnosti?

Výsledky: Pro N1 = N2 = 50 je *p*-hodnota rovna 0,0497.
Úspěšnost v experimentální skupině by musela být alespoň 77 %.