

Základní informace o programovém systému STATISTICA

Systém má modulární stavbu. V multilicenci pro Masarykovu univerzitu jsou k dispozici moduly: Basic Statistics/Tables, Multiple Regression, ANOVA, Nonparametrics, Distribution Fitting, Advanced Linear / Nonlinear Models, Multivariate Exploratory Techniques, Industrial Statistics & Six Sigma. Instalační CD lze získat v celouniverzitní počítačové studovně MU na Komenského náměstí 2. Lze také přímo stáhnout ISO obraz instalačního CD. Bližší informace viz

https://wiki.ics.muni.cz/doku.php?id=softwarove_licence:licence#statistica

Velké množství informací o systému STATISTICA lze najít na webové stránce společnosti StatSoft, která je jejím distributorem v České republice (www.statsoft.cz). Z této stránky vede rovněž odkaz na elektronickou učebnici statistiky.

STATISTICA má několik typů oken:

- **spreadsheet** (datové okno, má příponu sta, jeho obsah však lze exportovat i v jiných formátech). Do datového okna lze načítat datové soubory nejrůznějších typů (např. z tabulkových procesorů, databázové soubory, ASCII soubory).
- **workbook** (má příponu stw). Do workbooku ukládají výstupy, tj. tabulky a grafy. Skládá se ze dvou oken, v levém okně je znázorněna stromová struktura výstupů, v pravém jsou samotné výstupy. V levém okně se lze pohybovat myší nebo kurzorem, mazat, přesouvat, editovat apod. Výstupy mohou sloužit jako vstupy pro další analýzy a grafy.
- **report** (má příponu str, lze ho uložit i ve formátu rtf, txt či htm). Pokud požadujeme, aby se výstupy ukládaly nejen do workbooku, ale i do reportu, postupujeme takto: Tools – Options – Output Manager – zaškrtneme Also send to Report Window – OK. Report se podobně jako workbook skládá ze dvou oken. Do reportu můžeme vkládat vlastní text, vysvětlující komentáře, poznámky apod. Tabulky a grafy lze v reportu i workbooku dále upravovat.
- **okno grafů** (přípona stg, lze ho uložit i jako bmp, jpg, png a wmf). Získá se tak, že ve workbooku klikneme pravým tlačítkem na graf a vybereme Clone Graph.
- **programovací okno** (přípona svb). Slouží pro zápis programů v jazyku STATISTICA Visual Basic.

Mezi jednotlivými typy oken se přepínáme pomocí položky Window v hlavním menu.

Bodové zpracování četností

1. Zapište do datového okna programu STATISTICA datový soubor, který bude obsahovat známky z matematiky, angličtiny a údaje o pohlaví dvaceti studentů.
Návod: File – New – Number of variables 3, Number of cases 20, OK.
2. ZnakY nazvěte X, Y, Z, vytvořte jim návěští (X - známka z matematiky, Y - známka z angličtiny, Z - pohlaví studenta) a popište, co znamenají jednotlivé varianty (u znaků X a Y: 1 - výborně, 2 - velmi dobře, 3 - dobře, 4 - neprospěl, u znaku Z: 0 - žena, 1 - muž). Soubor uložte pod názvem znamky.sta
Návod: Kurzor nastavíme na Var1 – 2x klikneme myší – Name X – Long Name známka z matematiky, Text label – 1 výborně, 2 velmi dobře, 3 dobře, 4 neprospěl, OK. U proměnné Y lze text label okopírovat z proměnné X – v Text Labels Editor zvolíme Copy from variable X.
Přepínání mezi číselnými hodnotami a jejich textovým popisem se děje pomocí tlačítka s obrázkem štítku.
3. U znaků X a Y vypočtete absolutní četnosti, relativní četnosti a relativní kumulativní četnosti.
Návod: Statistics - Basic Statistics/Tables – Frequency tables – OK – Variables X, Y, OK – Summary. Všechny tři tabulky se uloží do workbooku a listovat v nich můžeme pomocí stromové struktury v levém okně.
4. Vytvořte sloupkový diagram absolutních četností znaků X a Y.
Návod: Graphs – Histograms – Variables X, Y – OK- vypneme Normal fit – Advanced – zaškrtneme Breaks between Columns, OK.
Vytvořte výsečový diagram absolutních četností znaků X a Y.
Návod: Graphs – 2D Graphs – Pie Charts – Variables X, Y – OK – Advanced – Pie legend Text and Percent (nebo Text and Value) – OK.
Vytvořte polygon absolutních četností znaků X a Y.
Návod: ve workbooku vstoupíme do tabulky rozložení četností proměnné X. Pomocí Edit – Delete - Cases vymažeme řádek označený Missing. Nastavíme se kurzorem na Count – Graphs – Graphs of Block Data - Line Plot:Entire Columns. Vykreslí se polygon četností.
5. Vytvořte graf empirické distribuční funkce znaku X.
Návod: Při tvorbě histogramu zadáme v Advanced volbu Showing Type Cumulative, Y axis % - 2 x klikneme myší na pozadí grafu – otevře se okno All Options – vybereme Plot:Bars – Type Rectangles. V tomto grafu jsou však svislé čáry až k vodorovné ose. Lze použít i jiný typ grafu: vytvoříme nový datový soubor, který bude mít dvě proměnné a případů o dva víc než je počet variant znaku X. Do 1. proměnné zapíšeme do 1. řádku hodnotu o 1 menší než je 1. varianta znaku X, pak varianty znaku X a nakonec hodnotu o 1 větší než je poslední varianta znaku X. Do 2. proměnné zapíšeme 0, pak relativní kumulativní četnosti znaku X (v procentech) a nakonec 100. Graphs – Scatterplots – Variables V1, V2 – OK – vypneme Linear fit – OK – 2x klikneme na pozadí grafu – Plot:General – vypneme Markers, zaškrtneme Line – Line Type:Step – OK.
Vytvořte graf četnostní funkce znaku X.
Návod: Při tvorbě histogramu zadáme v Advanced Y axis % - 2 x klikneme myší na pozadí grafu – vybereme Plot General – zaškrtneme Markers – vybereme Plot:Bars – Type Lines.

5. Z datového souboru vyberte pouze ženy (pouze muže) a úkol 3 proveďte pro ženy (pro muže).

Návod: Statistics - Basic Statistics/Tables – Frequency tables – OK – Variables X, Y, OK – Select Cases – zaškrtneme Enable Selection Conditions – zaškrtneme Specific, selected by: By Expression $Z = 0$, OK.

6. Nadále pracujte s celým datovým souborem. Vytvořte kontingenční tabulku absolutních četností znaků X a Y a graf simultánní četností funkce.

Návod: Statistics - Basic Statistics/Tables – Tables and banners – OK – Select cases – All – OK – Specify tables - List 1 X, List 2 Y, OK, Summary.

Vytvoření grafu simultánní četnostní funkce: Návrat do Crosstabulation Tables Result – 3D histograms – vybereme Axis Scaling – Mode Manual – Minimum 0 (a totéž provedeme pro Axis Y) – dále vybereme Graph Layout – Type – Spikes – OK. Graf lze natáčet pomocí Point of View.

Vytvořte kontingenční tabulku sloupcově a řádkově podmíněných relativních četností znaků X a Y.

Návod: Návrat do Crosstabulation Tables Result – Options - zaškrtneme ve sloupci Compute tables volbu Percentages of column counts (resp. Percentages of row counts).

Intervalové zpracování četností

1. Zapište do datového okna programu STATISTICA datový soubor, který bude obsahovat údaje o mezi plasticity oceli a mezi pevnosti (viz př. 2.13.). Proměnným X a Y vytvořte návěští „mez plasticity“ a „mez pevnosti“. Soubor pak uložte pod názvem ocel.sta.
Návod: viz 1. cvičení, bod 1.
2. Pro X a Y použijeme intervalové zpracování četností. Pro aplikaci Sturgersova pravidla potřebujeme znát počet variant proměnné X a Y.
Návod: Zjištění absolutních četností - viz 1. cvičení, bod 3. Zjištění počtu variant: ve workbooku se nastavíme kurzorem na sloupec Count – 2x klikneme myši – vybereme Values/Stats – ve výstupní tabulce se objeví mj. N. Počet variant je N-1.
(X má 50 variant, Y má 52 variant, v obou případech volíme 7 třídicích intervalů.)
Dále musíme zjistit minimum a maximum, abychom vhodně stanovili třídicí intervaly.
Návod: Statistics – Basic Statistics/Tables – Descriptive statistics - Variables X,Y – zaškrtneme Minimum&maximum – Summary. (Pro X je minimum 33 a maximum 160, tedy vhodná volba třídicích intervalů je (30,50>, (50,70>, ..., (150,170> - viz př. 2.13., pro Y je minimum 52 a maximum 189, tedy třídicí intervaly zvolíme (50,70>, (70,90>, ... (170,190> - viz př. 2.19.)
3. Vytvořte histogram pro X a pro Y.
Návod: Graphs – Histograms – Variables X – vypneme Normal fit – Advanced – zaškrtneme Boundaries – Specify Boundaries – 50 70 90 110 130 150 170 OK – Y Axis %.
2 x klikneme na pozadí grafu a ve volbě All Options můžeme měnit různé vlastnosti grafu.
Upozornění: STATISTICA v histogramu znázorňuje relativní četnost výškou obdélníku, nikoliv jeho plochou, což není v souladu s definicí 2.14.
4. Proveďte zakódování hodnot proměnných X a Y do příslušných třídicích intervalů.
Návod: Insert – Add Variables – 2 – After Y – OK – přejmenujeme je na RX a RY.
Nastavíme se kurzorem na RX – Data – Recode - vyplníme podmínky pro všech 7 kategorií. (Pozor – podmínky se musí psát ve tvaru $X > 30$ and $X \leq 50$ atd.). Pak klepneme na OK. Analogicky pro Y.
5. Vytvořte graf intervalové empirické distribuční funkce pro X.
Návod: Vytvoříme Frequency table pro RX. Před 1. případ vložíme řádek, kde do Category napíšeme 0 a do Cumulative Count také 0. Nastavíme se kurzorem na Cumulative Percent – Graphs – Graphs of Block Data – Custom Graph from Block by Column – Line Plots (Variables) – OK. 2x klikneme na pozadí grafu – Plot: General – vypneme Markers – Axis: Scaling – Mode Manual – Minimum 1, Maximum 9 – Axis: Custom Units – Position 1, Text 30 atd až Position 9, Text 190 – OK.
6. Sestavte kontingenční tabulky absolutních četností (relativních četností, sloupcově a řádkově podmíněných relativních četností) dvourozměrných třídicích intervalů pro (X,Y).
Návod: Viz úkol č. 6 ve cvičení 1, kde budeme pracovat s proměnnými RX a RY.
7. Vytvořte stereogram pro (RX,RY).
Návod: V tabulce Crosstabulation Tables Result zaškrtneme 3D histograms. Ve volbě Axis Scaling (pro RX i pro RY) změním Mode na Manual – Minimum 0. Pomocí Axis: Custom Units (viz úkol č. 5) můžeme změnit popis hodnot na x-ové a y-ové ose.

Výpočet číselných charakteristik jednorozměrného a dvourozměrného souboru, regresní přímka

1. Načtete soubor znamky.sta. Pro známky z matematiky a angličtiny vypočtete medián, dolní a horní kvartil a kvartilovou odchylku. Výsledky porovnejte s příkladem 3.5.
Návod: Statistics – Basic Statistics/Tables – Descriptive Statistics – NK - Variables X, Y, OK – zaškrtneme Median, Lower & upper quartiles, Quartile range – Summary.
2. Načtete soubor ocel.sta. Ůro mez plasticity a mez pevnosti vypčtete aritmetické průměry, směrodatné odchylky a rozptyly. Výsledky porovnejte s příkladem 15.
Návod: Návod: Statistics – Basic Statistics/Tables – Descriptive Statistics – OK - Variables X, Y, OK – zaškrtneme Mean, Standard Deviation, Variance – Summary.
Vysvětlení: Rozptyl a směrodatná odchylka vyjdou ve STATISTICE jinak než v př. 3.15., protože STATISTICA ve vzorci pro výpočet rozptylu nepoužívá $1/n$, ale $1/(n-1)$ - bude objasněno později v matematické statistice.
3. Nakreslete dvourozměrný tečkový diagram pro (X,Y).
Návod: Graphs – Scatterplots – Variables X,Y – OK - vypne Linear fit – OK.
4. Vypočtete kovarianční a koeficient korelace meze plasticity a meze pevnosti. Výsledky porovnejte s příkladem 3.15.
Návod: Statistics - Multiple Regression - Variables Independent X, Dependent Y – OK – OK – Residuals/assumption-prediction – Descriptive statistics – Covariances. Pro získání korelačního koeficientu zvolíme Correlation místo Covariances.
Vysvětlení: Kovariance vyjde ve STATISTICE jinak než v př. 3.15., protože ve STATISTICE se ve vzorci pro výpočet kovariance nepoužívá $1/n$, ale $1/(n-1)$ - bude objasněno později.
5. Určete koeficienty regresní přímky meze pevnosti na mez plasticity a stanovte index determinace. Určete regresní odhad meze pevnosti, je-li mez plasticity 110. Nakreslete regresní přímku do dvourozměrného tečkového diagramu.
Návod: V tabulce Multiple Regression zvolíme Variables Independent X, Dependent Y – OK – Summary:Regression results. Ve výstupní tabulce najdeme koeficient b_0 ve sloupci B na řádce označeném Intercept, koeficient b_1 ve sloupci B na řádce označeném X, index determinace pod označením R2.
Pro výpočet predikované hodnoty zvolíme Residuals/assumption/prediction Predict dependent variable X:110 - OK. Ve výstupní tabulce je hledaná hodnota označena jako Predictd.
Nakreslení regresní přímky: Návrat do Multiple Regression – Residuals / assumption / prediction – Perform residuals analysis – Scatterplots – Bivariate correlation – X, Y – OK.
Jiný způsob: Do dvourozměrného tečkového diagramu nakreslíme regresní přímku tak, že v tabulce 2D Scatterplots zvolíme Fit Linear, OK.

Výpočty pravděpodobností pomocí distribuční funkce binomického rozložení

Označme X náhodnou veličinu. Její distribuční funkci zavedeme vztahem $\Phi(x) = P(X \leq x)$. Pokud náhodná veličina X nabývá pouze konečně nebo spočetně mnoha hodnot, lze pomocí $\Phi(x)$ vyjádřit následující pravděpodobnosti:

- a) $P(X=x) = P(X \leq x) - P(X \leq x-1) = \Phi(x) - \Phi(x-1)$;
- b) $P(X \geq x) = 1 - P(X < x) = 1 - P(X \leq x-1) = 1 - \Phi(x-1)$;
- c) $P(x_1 \leq X \leq x_2) = P(x_1-1 < X \leq x_2) = \Phi(x_2) - \Phi(x_1)$.

STATISTICA poskytuje hodnoty distribučních funkcí mnoha rozložení. Omezíme se na **binomické rozložení** (funkce $\text{IBinom}(x,p,n)$, kde x ... počet úspěchů, p ... pravděpodobnost úspěchu v jednom pokusu, n ... celkový počet pokusů).

Vzorový příklad na binomické rozložení: Pojišťovna zjistila, že 12% pojistných událostí je způsobeno vloupáním. Jaká je pravděpodobnost, že mezi 30 náhodně vybranými pojistnými událostmi bude způsobeno vloupáním a) nejvýše 6, b) aspoň 6, c) právě 6, d) od dvou do pěti?

Řešení: X ... počet pojistných událostí způsobených vloupáním, $n = 30$, $p = 0,12$.

ad a) $P(X \leq 6) = \Phi(6) = 0,9393$,

ad b) $P(X \geq 6) = 1 - P(X \leq 5) = 1 - \Phi(5) = 0,1431$,

ad c) $P(X=6) = \Phi(6) - \Phi(5) = 0,0825$,

ad d) $P(2 \leq X \leq 5) = \Phi(5) - \Phi(1) = 0,7469$.

Postup ve STATISTICE: Otevřeme nový datový soubor se čtyřmi proměnnými a o jednom případě.

Řešení: Do Long Name 1. proměnné napíšeme $=\text{IBinom}(6;0,12;30)$.

Do Long Name 2. proměnné napíšeme $=1-\text{IBinom}(5;0,12;30)$.

Do Long Name 3. proměnné napíšeme $=\text{IBinom}(6;0,12;30)-\text{IBinom}(5;0,12;30)$.

Do Long Name 3. proměnné napíšeme $=\text{IBinom}(5;0,12;30)-\text{IBinom}(1;0,12;30)$.

(Do Lange Name proměnné vstoupíme tak, že v datovém okně 2x klikneme myší na název proměnné.)

Kreslení grafů distribuční funkce a pravděpodobnostní funkce binomického rozložení

Vzorový příklad: Nakreslete graf distribuční funkce a pravděpodobnostní funkce náhodné veličiny $X \sim \text{Bi}(12, 0.3)$

Postup ve STATISTICE: Vytvoříme nový datový soubor o 3 proměnných a 13 případech.

První proměnnou nazveme X a uložíme do ní hodnoty 0, 1, ..., 12 (do Long Name napíšeme $=v0-1$). Druhou proměnnou nazveme DF a uložíme do ní hodnoty distribuční funkce (do

Long Name napíšeme příkaz $=\text{IBinom}(x;0,3;12)$). Třetí proměnnou nazveme PF a uložíme do ní hodnoty pravděpodobnostní funkce (do Long Name napíšeme příkaz $=\text{Binom}(x;0,3;12)$).

Graf distribuční funkce: Graphs – Scatterplots – Variables X , DF – OK – vypneme Linear fit – OK – 2x klikneme na pozadí grafu – Plot:General – zaškrtneme Line – Line Type: Step – OK.

Graf pravděpodobnostní funkce: Graphs – Scatterplots – Variables X , PF – OK – vypneme Linear fit – OK.

Podle tohoto návodu nakreslete grafy distribučních a pravděpodobnostních funkcí binomického rozložení pro různá n a p , např. $n=5$, $p=0,5$ (resp. 0,75) apod. Sledujte vliv parametrů na vzhled grafů.

Grafy hustot a distribučních funkcí, výpočet kvantilů

STATISTICA umí kreslit grafy hustot a distribučních funkcí mnoha spojitých rozložení a počítat kvantily těchto rozložení. Slouží k tomu Probability Calculator v menu Statistics. Zaměříme se na rozložení uvedená definici 8.6.

1. Rovnoměrné spojité rozložení $Rs(0, 1)$
Statistics – Probability Calculator – Distributions – Beta – shape 1 - napíšeme 1, shape 2 – napíšeme 1. STATISTICA vykreslí graf hustoty a distribuční funkce. Hodnotu α -kvantilu zjistíme tak, že do okénka označeného p napíšeme dané α a po kliknutí na Compute se v okénku Beta objeví hodnota tohoto kvantilu.
2. Exponenciální rozložení $Ex(\lambda)$
Ve volbě Distributions vybereme Exponential a do okénka lambda napíšeme patřičnou hodnotu. Hodnotu α -kvantilu zjistíme tak, že do okénka označeného p napíšeme dané α a po kliknutí na Compute se v okénku exp objeví hodnota tohoto kvantilu.
3. Normální rozložení $N(\mu, \sigma^2)$
Ve volbě Distributions vybereme Z (Normal), do okénka mean napíšeme hodnotu μ a do okénka st. dev. napíšeme hodnotu σ . Hodnotu α -kvantilu zjistíme tak, že do okénka označeného p napíšeme dané α a po kliknutí na Compute se v okénku X objeví hodnota tohoto kvantilu.
4. Pearsonovo rozložení chí-kvadrát s n stupni volnosti $\chi^2(n)$
Ve volbě Distributions vybereme Chi 2 a do okénka df napíšeme patřičný počet stupňů volnosti. Hodnotu α -kvantilu zjistíme tak, že do okénka označeného p napíšeme dané α a po kliknutí na Compute se v okénku Chi 2 objeví hodnota tohoto kvantilu.
5. Studentovo rozložení s n stupni volnosti $t(n)$
Ve volbě Distributions vybereme t (Student) a do okénka df napíšeme patřičný počet stupňů volnosti. Hodnotu α -kvantilu zjistíme tak, že do okénka označeného p napíšeme dané α a po kliknutí na Compute se v okénku t objeví hodnota tohoto kvantilu.
6. Fisherovo-Snedecorovo rozložení s n_1 a n_2 stupni volnosti $F(n_1, n_2)$
Ve volbě Distributions vybereme F (Fisher) a do okének df_1 a df_2 napíšeme počet stupňů volnosti čitatele a jmenovatele. Hodnotu α -kvantilu zjistíme tak, že do okénka označeného p napíšeme dané α a po kliknutí na Compute se v okénku F objeví hodnota tohoto kvantilu.

Intervaly spolehlivosti pro parametry normálního rozložení

1. Interval spolehlivosti pro střední hodnotu, když neznáme rozptyl: pro tuto situaci umí STATISTICA vypočítat meze intervalu spolehlivosti sama.
Příklad: Při kontrole pěti balíčků cukru o deklarované hmotnosti 1000 g byly zjištěny tyto odchylky: -3, 2, -2, 0, 1. Odchylky považujeme za realizace náhodného výběru rozsahu 5 z rozložení $N(\mu, \sigma^2)$. Sestrojte 90% interval spolehlivosti pro μ .
Návod: Vytvoříme nový datový soubor o jedné proměnné a pěti případech. Zapišeme do něj uvedené odchylky. Statistics - Basic Statistics/Tables – Descriptive statistics – OK – Advanced – Variables v1, OK, zaškrtněte Conf. limits for mean – Interval 90%, Summary.
2. Ve všech ostatních případech postupujeme podle vzorců uvedených ve větách 12.9. a 12.13. Uvedme postup pro situaci, kdy hledáme interval spolehlivosti pro rozdíl středních hodnot dvou nezávislých normálně rozložených náhodných výběrů, když neznáme rozptyly, ale víme, že jsou shodné.
Příklad: Na jisté velké americké univerzitě bylo v r. 1969 náhodně vybráno 5 profesorek a nezávisle na tom 5 profesorů a byl zjištěn jejich roční příjem v tisících dolarů. Ženy: 9 12 8 10 16, muži: 16 19 12 11 22. Předpokládáme, že uvedené hodnoty jsou realizace dvou nezávislých náhodných výběrů, první z rozložení $N(\mu_1, \sigma^2)$, druhý z rozložení $N(\mu_2, \sigma^2)$. Sestrojte 95% interval spolehlivosti. Sestrojte 95% interval spolehlivosti pro rozdíl středních hodnot.
Návod: Vytvoříme nový datový soubor o čtyřech proměnných (Plat, Sex, HorniMez, DolniMez) a 10 případech. Do proměnné Plat napíšeme příjmy žen, pak příjmy mužů. Do proměnné Sex napíšeme 5x jedničku a 5x dvojku (1 = žena, 2 = muž). Pomocí Descriptive statistics zjistíme průměry a rozptyly platů žen a mužů. (Výběr žen či mužů: viz cvičení 1, úkol 5.). Výsledky: $m_1 = 11$, $s_1^2 = 10$, $n_1 = 5$, $m_2 = 16$, $s_2^2 = 21,5$, $n_2 = 5$. Do Long Name proměnné DolniMez napíšeme vzorec pro dolní mez (viz věta 12.13. (b)):
$$=11-16-\text{sqrt}((4*10+4*21,5)/8)*\text{sqrt}(1/5+1/5)*\text{VStudent}(0,975;8)$$

Do proměnné DolniMez se 10x uloží hodnota -10,79. Do Long Name proměnné HorniMez napíšeme vzorec pro horní mez (viz věta 12.13. (b)):
$$=11-16+\text{sqrt}((4*10+4*21,5)/8)*\text{sqrt}(1/5+1/5)*\text{VStudent}(0,975;8)$$

Do proměnné HorniMez se 10x uloží hodnota 0,79.
Znamená to, že s pravděpodobností aspoň 0,95 leží rozdíl středních hodnot platů žen a mužů v intervalu (-10,79, 0,79). Tento výsledek však nemá praktický význam, protože rozsahy obou výběrů byly příliš malé.

Příklad: Vyřešte pomocí STATISTIKY příklad 12.16.

Návod: Vytvoříme nový datový soubor o třech proměnných (Leva, Prava, Rozdil) a šesti případech. Do prvních dvou proměnných zapišeme zjištěné hodnoty. Do LongName proměnné Rozdil napíšeme =Leva – Prava a nyní postupujeme stejně jako v úkolu 1.

Testování hypotéz o parametrech normálního rozložení

Jednovýběrový t-test

Příklad: Při kontrole balicího automatu, který má plnit cukrem balíčky o hmotnosti 1000 g, byly při přesném převážení pěti balíčků zjištěny tyto odchylky (v gramech) od požadované hodnoty: 3, -2, 2, 0, 1. Na hladině významnosti 0,05 testujte hypotézu, že automat nemá systematickou odchylku od požadované hodnoty.

Návod pro provedení t-testu: Vytvořte soubor o jedné proměnné X a pěti případech. Do X запиšte naměřené hodnoty. V menu Basic Statistics/Tables vyberte volbu t-test, single sample, OK, Variables X, zaškrtněte Test all means against 0, Summary. Ve výstupní tabulce najdete hodnotu testového kritéria a p-hodnotu. Pokud p-hodnota nabude hodnoty $\leq \alpha$, pak se nulovou hypotézu zamítá na hladině významnosti α .

Dvouvýběrový t-test

Příklad: Na jisté velké americké univerzitě bylo v r. 1969 náhodně vybráno 5 profesorů a nezávisle na tom 5 profesorek a byl zjištěn jejich roční příjem v tisících dolarů.

Ženy: 9 12 8 10 16 Muži: 16 19 12 11 22.

Na hladině významnosti 0,05 testujte hypotézu, že střední hodnota příjmu žen je stejná jako střední hodnota příjmu mužů.

Návod: Vytvořte soubor o dvou proměnných (Plat a Sex) a 10 případech. Do proměnné Plat napište příjmy žen a mužů a do proměnné Sex dejte 5x jedničku a 5x dvojku. V menu Basic Statistics/Tables vyberte volbu t-test, independent, by groups, OK, Variables – Grouping Sex, Dependent Plat, OK, Summary T-tests. Ve výstupní tabulce se nejprve podívejte na p-hodnotu pro test homogenity rozptylů. Je-li větší než zvolená hladina významnosti, zjistěte hodnotu testového kritéria a p-hodnotu pro test shody středních hodnot. V opačném případě zaškrtněte v Options volbu t-test with separate variance estimates.

Párový t-test

Příklad: Na hladině významnosti 0,05 rozhodněte, zda se u osobního vozu určité značky při správném seřízení geometrie vozu sjíždějí obě přední pneumatiky stejně rychle. Bylo vybráno šest nových vozů a po určité době bylo zjištěno, o kolik mm se sjely jejich levé a pravé přední pneumatiky.

Poř. č. vozu	1	2	3	4	5	6
Levá pneumatika	1,8	1	2,2	0,9	1,5	1,6
Pravá pneumatika	1,5	1,1	2	1,1	1,4	1,4

Návod: Vytvořte soubor o dvou proměnných (Leva a Prava) a šesti případech. V menu Basic Statistics/Tables vyberte volbu t-test, dependent samples, OK, Variables Leva, Prava – Summary.