Úvod do testování hypotéz

Motivace: Častým úkolem statistika je na základě dat ověřit předpoklady o parametrech nebo typu rozložení, z něhož pochází náhodný výběr. Takovému předpokladu se říká nulová hypotéza. Nulová hypotéza vyjadřuje nějaký teoretický předpoklad, často skeptického rázu a uživatel ji musí stanovit předem, bez přihlédnutí k datovému souboru. Proti nulové hypotéze stavíme alternativní hypotézu, která říká, co platí, když neplatí nulová hypotéza. Alternativní hypotéza je formulována tak, aby mohla platit jenom jedna z těchto dvou hypotéz. Pravdivost alternativní hypotézy by znamenala objevení nějakých nových skutečností, nebo zásadnější změnu v dosavadních představách.

Např. výzkumník by chtěl na základě dat prověřit tezi (nový objev), že pasivní kouření škodí zdraví. Jako nulovou hypotézu tedy položí tvrzení, že pasivní kouření neškodí zdraví a proti nulové hypotéze postaví alternativní, že pasivní kouření škodí zdraví.

Testováním hypotéz se myslí rozhodovací postup, který je založen na daném náhodném výběru a s jehož pomocí rozhodneme o zamítnutí či nezamítnutí nulové hypotézy.

Nulová a alternativní hypotéza

Nechť $X_1, ..., X_n$ je náhodný výběr z rozložení $L(\vartheta)$, kde parametr $\vartheta \in \Xi$ neznáme. Nechť $h(\vartheta)$ je parametrická funkce a c daná reálná konstanta.

a) Oboustranná alternativa: Tvrzení H_0: $h(\vartheta) = c$ se nazývá jednoduchá nulová hypotéza. Proti nulové hypotéze postavíme složenou oboustrannou alternativní hypotézu H_1: $h(\vartheta) \neq c$.

b) Levostranná alternativa: Tvrzení H_0: $h(\vartheta) \geq c$ se nazývá složená pravostranná nulová hypotéza. Proti jednoduché nebo složené pravostranné nulové hypotéze postavíme složenou levostrannou alternativní hypotézu H_1: $h(\vartheta) < c$.

c) Pravostranná alternativa: Tvrzení H_0: $h(\vartheta) \leq c$ se nazývá složená levostranná nulová hypotéza. Proti jednoduché nebo složené levostranné nulové hypotéze postavíme složenou pravostrannou alternativní hypotézu H_1: $h(\vartheta) > c$.

Testováním H_0 proti H_1 rozumíme rozhodovací postup založený na náhodném výběru $X_1, ..., X_n$, s jehož pomocí zamítneme či nezamítneme platnost nulové hypotézy.

Chyba 1. a 2. druhu

Při testování H_0 proti H_1 se můžeme dopustit jedné ze dvou chyb: chyba 1. druhu spočívá v tom, že H_0 zamítneme, ač ve skutečnosti platí a chyba 2. druhu spočívá v tom, že H_0 nezamítneme, ač ve skutečnosti neplatí. Situaci přehledně znázorňuje tabulka:

skutečnost	rozhodnutí	
	H_0 nezamítáme	H_0 zamítáme
H_0 platí	správné rozhodnutí	chyba 1. druhu
H_0 neplatí	chyba 2. druhu	správné rozhodnutí

Pravděpodobnost chyby 1. druhu se značí α a nazývá se hladina významnosti testu (většinou bývá $\alpha = 0{,}05$, méně často 0,1 či 0,01). Pravděpodobnost chyby 2. druhu se značí β. Číslo $1–\beta$ se nazývá síla testu a vyjadřuje pravděpodobnost, že bude H_0 zamítnuta za předpokladu, že neplatí. Obvykle se snažíme, aby síla testu byla aspoň 0,8. Obě hodnoty, α i $1–\beta$, závisí na velikosti efektu, který se snažíme detekovat. Čím drobnější efekt, tím musí být větší rozsah náhodného výběru.

skutečnost	rozhodnutí	
	zdravý	nemocný
jsem zdravý	zdravý a neléčený	zdravý a léčený
jsem nemocný	nemocný a neléčený	nemocný a léčený

Testování pomocí kritického oboru

Najdeme statistiku $T_0 = T_0(X_1, ..., X_n)$, kterou nazveme testovým kritériem. Množina všech hodnot, jichž může testové kritérium nabýt, se rozpadá na obor nezamítnutí nulové hypotézy (značí se V) a obor zamítnutí nulové hypotézy (značí se W a nazývá se též kritický obor). Tyto dva obory jsou odděleny kritickými hodnotami (pro danou hladinu významnosti α je lze najít ve statistických tabulkách).

Jestliže číselná realizace t_0 testového kritéria T_0 padne do kritického oboru W, pak nulovou hypotézu zamítáme na hladině významnosti α a znamená to skutečné vyvrácení testované hypotézy. Jestliže t_0 padne do oboru nezamítnutí V, pak jde o pouhé mlčení, které platnost nulové hypotézy jenom připouští.

Pravděpodobnosti chyb 1. a 2. druhu nyní zapíšeme takto:

$P(T_0 \in W/H_0 \text{ platí}) = \alpha$, $P(T_0 \in V/H_1 \text{ platí}) = \beta$.

Stanovení kritického oboru pro danou hladinu významnosti α:

Označme t_{min} (resp. t_{max}) nejmenší (resp. největší) hodnotu testového kritéria.

Kritický obor v případě oboustranné alternativy má tvar

$W = \left(t_{min}, K_{\alpha/2}(T)\right\rangle \cup \left\langle K_{1-\alpha/2}(T), t_{max}\right)$, kde $K_{\alpha/2}(T)$ a $K_{1-\alpha/2}(T)$ jsou kvantily rozložení, jímž se řídí testové kritérium T_0, je-li nulová hypotéza pravdivá.

Kritický obor v případě levostranné alternativy má tvar:

$W = \left(t_{min}, K_{\alpha}(T)\right\rangle$.

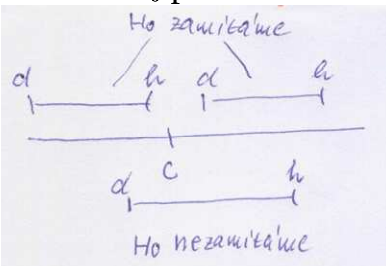
Kritický obor v případě pravostranné alternativy má tvar:

$W = \left\langle K_{1-\alpha}(T), t_{max}\right)$.

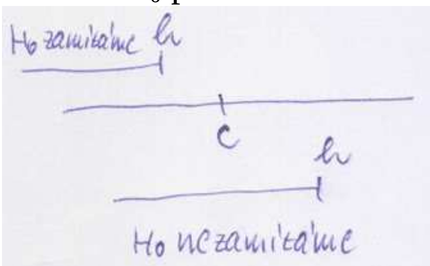
Testování pomocí intervalu spolehlivosti

Sestrojíme 100(1-α)% empirický interval spolehlivosti pro parametrickou funkci h(ϑ). Pokryje-li tento interval hodnotu c, pak H_0 nezamítáme na hladině významnosti α, v opačném případě H_0 zamítáme na hladině významnosti α.

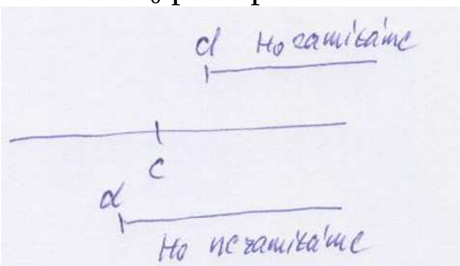
Pro test H_0 proti oboustranné alternativě sestrojíme oboustranný interval spolehlivosti.



Pro test H_0 proti levostranné alternativě sestrojíme pravostranný interval spolehlivosti.



Pro test H_0 proti pravostranné alternativě sestrojíme levostranný interval spolehlivosti.



Testování pomocí p-hodnoty

p-hodnota udává nejnižší možnou hladinu významnosti pro zamítnutí nulové hypotézy. Je to riziko, že bude zamítnuta H_0 za předpokladu, že platí (riziko planého poplachu). Jestliže p-hodnota $\leq \alpha$, pak H_0 zamítáme na hladině významnosti α, je-li p-hodnota $> \alpha$, pak H_0 nezamítáme na hladině významnosti α.

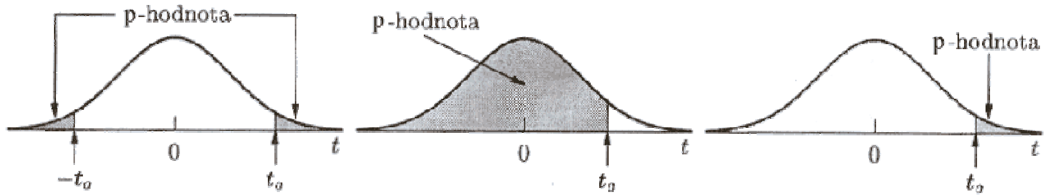
Způsob výpočtu p-hodnoty:

Pro oboustrannou alternativu $p = 2 \min\{P(T_0 \leq t_0), P(T_0 \geq t_0)\}$.

Pro levostrannou alternativu $p = P(T_0 \leq t_0)$.

Pro pravostrannou alternativu $p = P(T_0 \geq t_0)$.

Ilustrace významu p-hodnoty pro test nulové hypotézy proti oboustranné, levostranné a pravostranné alternativě:



(Zvonovitá křivka reprezentuje hustotu rozložení, kterým se řídí testové kritérium, je-li nulová hypotéza pravdivá.)

p-hodnota vyjadřuje pravděpodobnost, s jakou číselné realizace $x_1, ..., x_n$ náhodného výběru $X_1, ..., X_n$ podporují H_0, je-li pravdivá. Statistické programové systémy poskytují ve svých výstupech p-hodnotu. Její výpočet vyžaduje znalost distribuční funkce rozložení, kterým se řídí testové kritérium T_0, je-li H_0 pravdivá.

Doporučený postup při testování hypotéz

1. Stanovíme nulovou hypotézu a alternativní hypotézu. Přitom je vhodné zvolit jako alternativní hypotézu ten předpoklad, jehož přijetí znamená závažné opatření a mělo by k němu dojít jen s malým rizikem omylu.

2. Zvolíme hladinu významnosti α. Zpravidla volíme $\alpha = 0{,}05$, méně často 0,1 nebo 0,01.

3. Najdeme vhodné testové kritérium a na základě zjištěných dat vypočítáme jeho realizaci.

4.
a) Testujeme-li pomocí kritického oboru, pak ho stanovíme. Jestliže realizace testového kritéria padla do kritického oboru, nulovou hypotézu zamítáme na hladině významnosti α a přijímáme alternativní hypotézu. V opačném případě nulovou hypotézu nezamítáme na hladině významnosti α.
b) Testujeme-li pomocí intervalu spolehlivosti, vypočteme empirický $100(1-\alpha)\%$ interval spolehlivosti pro parametrickou funkci $h(\vartheta)$. Pokud číslo c padne do tohoto intervalu, nulovou hypotézu nezamítáme na hladině významnosti α. V opačném případě nulovou hypotézu zamítáme na hladině významnosti α a přijímáme alternativní hypotézu.
c) Testujeme-li pomocí p-hodnoty, vypočteme ji a porovnáme ji s hladinou významnosti α. Jestliže $p \le \alpha$, pak nulovou hypotézu zamítáme na hladině významnosti α a přijímáme alternativní hypotézu. Je-li $p > \alpha$, pak nulovou hypotézu nezamítáme na hladině významnosti α.

5. Na základě rozhodnutí, které jsme učinili o nulové hypotéze, provedeme nějaké konkrétní opatření, např. seřídíme obráběcí stroj.

(Při testování hypotéz musíme mít k dispozici odpovídající nástroje, nejlépe vhodný statistický software. Nemáme-li ho k dispozici, musíme znát příslušné vzorce. Dále potřebujeme statistické tabulky a kalkulačku.)

Příklad: 10 x nezávisle na sobě byla změřena jistá konstanta μ. Výsledky měření byly: 2 1,8 2,1 2,4 1,9 2,1 2 1,8 2,3 2,2. Tyto výsledky považujeme za číselné realizace náhodného výběru $X_1, \ldots, X_{10}$ z rozložení $N(\mu, 0{,}04)$. Nějaká teorie tvrdí, že $\mu = 1{,}95$.

1. Oboustranná alternativa

Proti nulové hypotéze H_0: $\mu = 1{,}95$ postavíme oboustrannou alternativu H_1: $\mu \neq 1{,}95$. Na hladině významnosti 0,05 testujte H_0 proti H_1 všemi třemi popsanými způsoby.

Řešení:

$m = \frac{1}{10}(2 + \ldots + 2{,}2) = 2{,}06$, $\sigma^2 = 0{,}04$, $n = 10$, $\alpha = 0{,}05$, $c = 1{,}95$

a) Test provedeme pomocí kritického oboru.

Pro úlohy o střední hodnotě normálního rozložení při známém rozptylu používáme pivotovou statistiku $U = \frac{M-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$.

Testové kritérium tedy bude

$T_0 = \frac{M-c}{\frac{\sigma}{\sqrt{n}}}$ a bude mít rozložení N(0, 1), pokud je nulová hypotéza pravdivá. Vypočítáme realizaci testového kritéria:

$t_0 = \frac{2{,}06-1{,}95}{\frac{0{,}2}{\sqrt{10}}} = 1{,}74$. Stanovíme kritický obor:

$W = \left(t_{min}, K_{\alpha/2}(T)\right\rangle \cup \left\langle K_{1-\alpha/2}(T), t_{max}\right) = \left(-\infty, u_{\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}, \infty\right) = \left(-\infty, -u_{1-\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}, \infty\right) = \left(-\infty, -u_{0.975}\right\rangle \cup \left\langle u_{0,975}, \infty\right) = \left(-\infty, -1{,}96\right\rangle \cup \left\langle 1{,}96, \infty\right)$.

Protože $1{,}74 \notin W$, H_0 nezamítáme na hladině významnosti 0,05.

b) Test provedeme pomocí intervalu spolehlivosti.

Meze 100(1-α)% empirického intervalu spolehlivosti pro střední hodnotu μ při známém rozptylu σ^2 jsou:

$$(d, h) = \left(m - \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}, m + \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}\right).$$

V našem případě dostáváme:

$$d = 2{,}06 - \frac{0{,}2}{\sqrt{10}} u_{0{,}975} = 2{,}06 - \frac{0{,}2}{\sqrt{10}} .1{,}96 = 1{,}936,$$

$$h = 2{,}06 + \frac{0{,}2}{\sqrt{10}} u_{0{,}975} = 2{,}06 + \frac{0{,}2}{\sqrt{10}} .1{,}96 = 2{,}184.$$

Protože 1,95 $\in$ (1,936; 2,184), H_0 nezamítáme na hladině významnosti 0,05.

c) Test provedeme pomocí p-hodnoty.

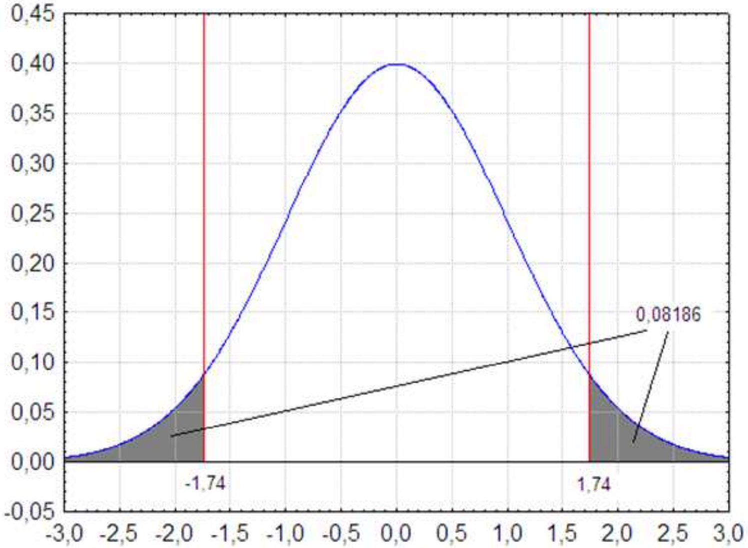
Protože proti nulové hypotéze stavíme oboustrannou alternativu, použijeme vzorec

$p = 2 \min\{P(T_0 \leq t_0), P(T_0 \geq t_0)\} = 2 \min \{P(T_0 \leq 1{,}74), P(T_0 \geq 1{,}74)\} =$

$= 2 \min \{ \Phi(1{,}74), 1 - \Phi(1{,}74) \} = 2 \min \{ 0{,}95907, 1 - 0{,}95907 \} = 0{,}08186.$

Jelikož 0,08186 > 0,05, nulovou hypotézu nezamítáme na hladině významnosti 0,05.

Ilustrace významu p-hodnoty pro oboustranný test



2. Levostranná alternativa

Proti nulové hypotéze H_0: $\mu = 1{,}95$ postavíme levostrannou alternativu
H_1: $\mu < 1{,}95$. Na hladině významnosti 0,05 testujte H_0 proti H_1 všemi třemi popsanými způsoby.

Řešení:

a) Test provedeme pomocí kritického oboru.

Na rozdíl od oboustranné alternativy bude mít kritický obor tvar

$W = \langle -\infty, u_\alpha) = \langle -\infty, u_{0,05}) = \langle -\infty, -1{,}645)$.

Protože $1{,}74 \notin W$, H_0 nezamítáme na hladině významnosti 0,05.

b) Test provedeme pomocí intervalu spolehlivosti.

Meze $100(1-\alpha)\%$ empirického pravostranného intervalu spolehlivosti pro střední hodnotu μ při známém rozptylu σ^2 jsou:

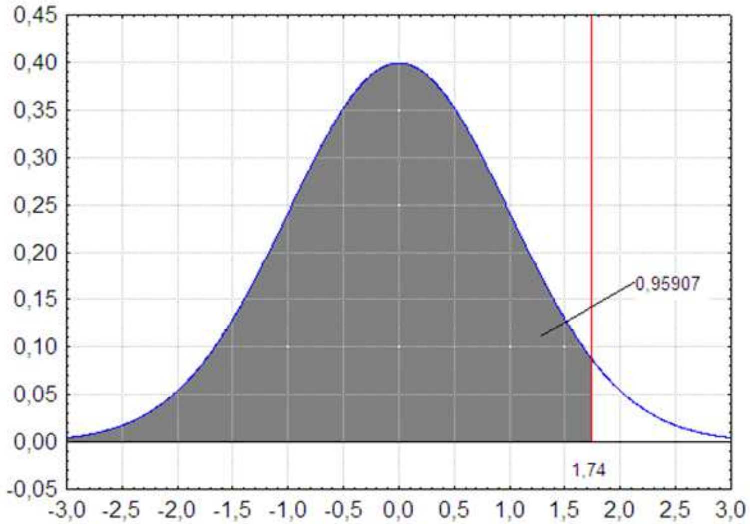
$(-\infty, h) = (-\infty, m + \frac{\sigma}{\sqrt{n}} u_{1-\alpha})$.

V našem případě dostáváme: $h = 2{,}06 + \frac{0{,}2}{\sqrt{10}} u_{0,95} = 2{,}06 + \frac{0{,}2}{\sqrt{10}} .1{,}645 = 2{,}164$.

Protože $1{,}95 \in (-\infty; 2{,}164)$, H_0 nezamítáme na hladině významnosti 0,05.

c) Test provedeme pomocí p-hodnoty.
Protože proti nulové hypotéze stavíme levostrannou alternativu, použijeme vzorec
$p = P(T_0 \leq t_0) = \Phi(1{,}74) = 0{,}95907$.
Jelikož 0,95907 > 0,05, nulovou hypotézu nezamítáme na hladině významnosti 0,05.

Ilustrace významu p-hodnoty pro levostranný test



3. Pravostranná alternativa

Proti nulové hypotéze H_0: $\mu = 1{,}95$ postavíme pravostrannou alternativu
H_1: $\mu > 1{,}95$. Na hladině významnosti 0,05 testujte H_0 proti H_1 všemi třemi popsanými způsoby.

Řešení:

a) Test provedeme pomocí kritického oboru.

Na rozdíl od oboustranné alternativy bude mít kritický obor tvar

W = $\langle u_{1-\alpha}, \infty) = \langle u_{0,95}, \infty) = \langle 1{,}645, \infty)$.

Protože $1{,}74 \in$ W, H_0 zamítáme na hladině významnosti 0,05 ve prospěch pravostranné alternativy.

b) Test provedeme pomocí intervalu spolehlivosti.

Meze $100(1-\alpha)\%$ empirického levostranného intervalu spolehlivosti pro střední hodnotu μ při známém rozptylu σ^2 jsou:

$(d, \infty) = (m - \frac{\sigma}{\sqrt{n}} u_{1-\alpha}, \infty)$.

V našem případě dostáváme: $d = 2{,}06 - \frac{0{,}2}{\sqrt{10}} u_{0,95} = 2{,}06 - \frac{0{,}2}{\sqrt{10}} . 1{,}645 = 1{,}956$.

Protože $1{,}95 \notin (1{,}956, \infty)$, H_0 zamítáme na hladině významnosti 0,05 ve prospěch pravostranné alternativy.

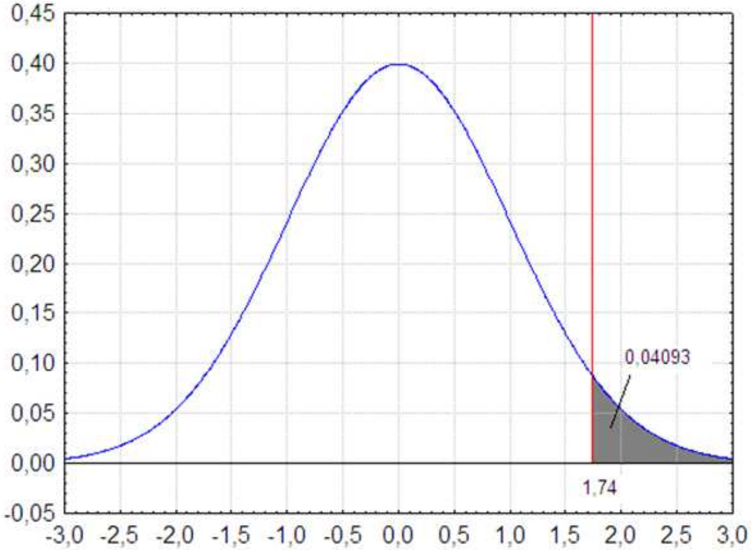
c) Test provedeme pomocí p-hodnoty.

Protože proti nulové hypotéze stavíme pravostrannou alternativu, použijeme vzorec

$p = P(T_0 \geq t_0) = 1 - \Phi(1{,}74) = 1 - 0{,}95907 = 0{,}04093$.

Jelikož $0{,}04093 \leq 0{,}05$, nulovou hypotézu zamítáme na hladině významnosti 0,05 ve prospěch pravostranné alternativy.

Ilustrace významu p-hodnoty pro pravostranný test



Testy normality dat

K ověřování normality dat slouží celá řada testů, které jsou podrobně popsány ve statistické literatuře. Zde se omezíme na tři testy, které jsou implementovány v systému STATISTICA, a to Kolmogorovův – Smirnovův test a jeho Lilieforsovu variantu, Shapirův – Wilkův test a Andersonův – Darlingův test.

K závěrům těchto testů však přistupujeme s určitou opatrností. Máme-li k dispozici rozsáhlejší datový soubor (orientačně $n > 30$) a test zamítne na obvyklé hladině významnosti 0,01 nebo 0,05 hypotézu o normalitě, i když vzhled diagnostických grafů svědčí jenom o lehkém porušení normality, nedopustíme se závažné chyby, pokud použijeme statistickou metodu založenou na normalitě dat.

Kolmogorovův – Smirnovův test a jeho Lilieforsova varianta

Testujeme hypotézu, která tvrdí, že náhodný výběr $X_1, ..., X_n$ pochází z normálního rozložení s parametry μ a σ^2.
Distribuční funkci tohoto rozložení označme $\Phi_T(x)$.
Nechť $F_n(x)$ je výběrová distribuční funkce.
Testovou statistikou je statistika $D_n = \sup\limits_{-\infty<x<\infty} \left|F_n(x) - \Phi_T(x)\right|$.
Nulovou hypotézu zamítáme na hladině významnosti α, když $D_n \geq D_n(\alpha)$, kde $D_n(\alpha)$ je tabelovaná kritická hodnota.
Pro $n \geq 30$ lze $D_n(\alpha)$ aproximovat výrazem $\sqrt{\frac{1}{2n} \ln \frac{2}{\alpha}}$.

Upozornění: Nulová hypotéza musí specifikovat distribuční funkci zcela přesně, včetně všech jejích případných parametrů. Např. K-S test lze použít pro testování hypotézy, že náhodný výběr $X_1, ..., X_n$ pochází z rozložení Rs(0,1), což se využívá při testování generátorů náhodných čísel. Pokud však parametry distribuční funkce odhadujeme z výběru, změní se rozložení testové statistiky D_n a jde o Lilieforsův test. Příslušné modifikované kvantily byly určeny pomocí simulačních studií.

Příklad: Jsou dány hodnoty 10, 12, 8, 9, 16. Pomocí K- S testu zjistěte na hladině významnosti 0,05, zda tato data pocházejí z normálního rozložení.

Řešení: Odhadem střední hodnoty je výběrový průměr m = 11, odhadem rozptylu je výběrový rozptyl $s^2 = 10$. Uspořádaný náhodný výběr je (8, 9, 10, 12, 16). Vypočteme hodnoty výběrové distribuční funkce:

$$x < 8 : F_5(x) = 0$$

$$8 \leq x < 9 : F_5(x) = \frac{1}{5} = 0{,}2$$

$$9 \leq x < 10 : F_5(x) = \frac{2}{5} = 0{,}4$$

$$10 \leq x < 12 : F_5(x) = \frac{3}{5} = 0{,}6$$

$$12 \leq x < 16 : F_5(x) = \frac{4}{5} = 0{,}8$$

$$x \geq 16 : F_5(x) = 1$$

Hodnoty teoretické distribuční funkce $\Phi_T(x)$ v bodech 8, 9, 10, 12, 16:

$$\Phi_T(8) = \Phi\left(\frac{8-11}{\sqrt{10}}\right) = \Phi(-0{,}95) = 1 - \Phi(0{,}95) = 1 - 0{,}82894 = 0{,}17106$$

$$\Phi_T(9) = \Phi\left(\frac{9-11}{\sqrt{10}}\right) = \Phi(-0{,}63) = 1 - \Phi(0{,}63) = 1 - 0{,}73565 = 0{,}26435$$

$$\Phi_T(10) = \Phi\left(\frac{10-11}{\sqrt{10}}\right) = \Phi(-0{,}32) = 1 - \Phi(0{,}32) = 1 - 0{,}62552 = 0{,}37448$$

$$\Phi_T(12) = \Phi\left(\frac{12-11}{\sqrt{10}}\right) = \Phi(0{,}32) = 0{,}62552$$

$$\Phi_T(16) = \Phi\left(\frac{16-11}{\sqrt{10}}\right) = \Phi(1{,}58) = 0{,}94295$$

(Φ je distribuční funkce rozložení N(0,1).)

Rozdíly mezi výběrovou distribuční funkcí $F_5(x)$ a teoretickou distribuční funkcí $\Phi_T(x)$:

$d_1 = 0{,}2 - 0{,}17106 = 0{,}02894$;
$d_2 = 0{,}4 - 0{,}26435 = 0{,}13565$;
$d_3 = 0{,}6 - 0{,}37448 = 0{,}22552$;
$d_4 = 0{,}8 - 0{,}62552 = 0{,}17448$;
$d_5 = 1 - 0{,}94295 = 0{,}05705$.

Testová statistika: $D_5 = 0{,}22552$, modifikovaná kritická hodnota pro $n = 5$, $\alpha = 0{,}05$ je 0,343.
Protože $0{,}22552 < 0{,}343$, hypotézu o normalitě nezamítáme na hladině významnosti 0,05.

Shapirův – Wilkův test

Testujeme hypotézu, že náhodný výběr $X_1, ..., X_n$ pochází z normálního rozložení $N(\mu, \sigma^2)$.

Testová statistika má tvar:

$$W = \frac{\sum_{i=1}^{m} a_i^{(n)} \left[X_{(n-i+1)} - X_{(i)}\right]^2}{\sum_{i=1}^{m} \left(X_i - M\right)^2},$$

kde m = n/2 pro n sudé a m = (n-1)/2 pro n liché. Koeficienty $a_i^{(n)}$ jsou tabelovány.

Na testovou statistiku W lze pohlížet jako na korelační koeficient mezi uspořádanými pozorováními a jim odpovídajícími kvantily standardizovaného normálního rozložení.

V případě, že data vykazují perfektní shodu s normálním rozložením, bude mít W hodnotu 1.

Hypotézu o normalitě tedy zamítneme na hladině významnosti α, když se na této hladině neprokáže korelace mezi daty a jim odpovídajícími kvantily rozložení N(0,1).

Lze také říci, že S – W test je založen na zjištění, zda body v Q-Q grafu jsou významně odlišné od regresní přímky proložené těmito body.

Andersonův – Darlingův test

Testujeme hypotézu, že náhodný výběr $X_1, \dots, X_n$ pochází z normálního rozložení $N(\mu, \sigma^2)$.

Testová statistika má tvar:

$$AD = -\frac{1}{n}\left[\sum_{i=1}^{n}(2i-1)\left\{\ln\Phi\left(\frac{x_{(i)}-m}{s}\right)+\ln\left(1-\Phi\left(\frac{x_{n+1-(i)}-m}{s}\right)\right)\right\}\right]-n\,,$$

kde $x_{(i)}$ jsou vzestupně uspořádané realizace náhodného výběru, Φ je distribuční funkce rozložení N(0,1).

Hypotéza H_0 se zamítá na hladině významnosti α, je-li vypočítaná hodnota testové statistiky AD větší než kritická hodnota $D_{1-\alpha}$. Pro velký rozsah výběru se přibližná 95% kritická hodnota počítá podle vzorce

$$D_{0,95} = 1{,}0348\left(1-\frac{1{,}013}{n}-\frac{0{,}93}{n^2}\right)$$

Příklad:

Jsou dány hodnoty 10, 12, 8, 9, 16. Pomocí Lilieforsova testu, S – W testu a A – D testu testujte na hladině významnosti 0,05 hypotézu, že tato data pocházejí z normálního rozložení.

Řešení:

Vytvoříme nový datový soubor o jedné proměnné nazvané X a pěti případech. Do proměnné X zapíšeme uvedené hodnoty.

Provedení Lilieforsova a S-W testu:

V menu vybereme Statistiky – Základní statistiky/tabulky – Tabulky četností – OK, Proměnné X – OK. Na záložce zvolíme Normalita a zaškrtneme Lilieforsův test a Shapiro – Wilkův W test – Testy normality.

	Testy normality (Tabulka1)				
Proměnná	N	max D	Lilliefors p	W	p
X	5	0,224085	p > .20	0,912401	0,482151

Vidíme, že testová statistika K-S testu je d = 0,22409, odpovídající Lilieforsova p-hodnota je větší než 0,2, tedy hypotézu o normalitě nezamítáme na hladině významnosti 0,05.

Testová statistika S-W testu je W = 0,9124, odpovídající p-hodnota je 0,48215, tedy hypotézu o normalitě nezamítáme na hladině významnosti 0,05.

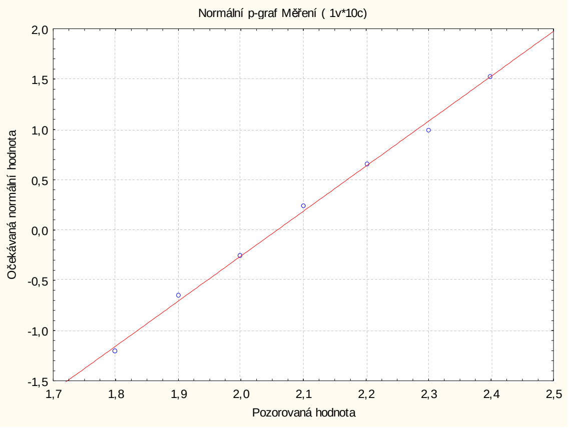
Provedení A - D testu:

Statistiky – Rozdělení & simulace – proložení dat rozděleními – OK – Proměnné Spojité: X – na záložce Spojité proměnné ponecháme zaškrtnuté pouze Normální, na záložce Možnosti vybereme Anderson – Darling – OK – Souhrnné statistiky rozdělení.

	Souhrn rozdělení for Proměnná: x (Tabulka4)							
	K-S d	K-S p-hodn.	AD stat.	AD p-hodn.	Chí-kvadrát	Chí-kvadr. p-hodn.	Chí-kvadr. SV	Posun (práh/poloha)
Normální (poloha,měřítko)	0,224085	0,915101	0,295219	0,940172				

Testová statistika A – D testu je 0,2952, odpovídající p-hodnota je 0,9402, tedy hypotézu o normalitě nezamítáme na hladině významnosti 0,05.

Výpočet doplníme NP plotem vytořeným pomocí systému STATISTICA:
Grafy – 2D Grafy – Normální pravděpodobnostní grafy – Proměnné X, zrušíme volbu Neurčovat průměrnou pozici svázaných pozorování – OK.



Poznámka o dalších testech normality

Existují testy normality založené na výběrové šikmosti a špičatosti. Pro náhodnou veličinu s normálním rozložením platí, že její šikmost i špičatost jsou nulové. Pro výběr z normálního rozložení by tedy výběrová šikmost a špičatost měly být blízké 0.

Nechť $X_1, \ldots, X_n$ je náhodný výběr.

Výběrová šikmost: $A_3 = \dfrac{\frac{1}{n}\sum_{i=1}^{n}(X_i - M)^3}{\left[\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - M)^2}\right]^3}$

Výběrová špičatost: $A_4 = \dfrac{\frac{1}{n}\sum_{i=1}^{n}(X_i - M)^4}{\left[\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - M)^2}\right]^4} - 3$

Lze dokázat, že pro výběr z normálního rozložení platí:

$E(A_3) = 0$, $D(A_3) = \frac{6(n-2)}{(n+1)(n+3)}$, $E(A_4) = -\frac{6}{n+1}$, $D(A_4) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}$.

Test založený na šikmosti zamítne hypotézu o normalitě na asymptotické hladině významnosti α, když $U_3 = \frac{|A_3|}{\sqrt{D(A_3)}} \geq u_{1-\alpha/2}$.

D'Agostinův test: zavedeme pomocné veličiny

$$b = \frac{3(n^2+27n-70)(n+1)(n+3)}{(n-2)(n+5)(n+7)(n+9)},$$

$$W^2 = \sqrt{2(b-1)} - 1,$$

$$d = \frac{1}{\sqrt{\ln W}}, \quad a = \sqrt{\frac{2}{W^2-1}}.$$

Testová statistika má tvar $Z_3 = d \cdot \ln\left[\frac{U_3}{a} + \sqrt{(\frac{U_3}{a})^2+1}\right]$ a platí, že má přibližně rozložení N(0,1). Pro n>8 zamítáme hypotézu o normalitě pokud $|Z_3| \geq u_{1-\alpha/2}$.

Test založený na špičatosti zamítne hypotézu o normalitě na asymptotické hladině významnosti α, když $U_4 = \frac{|A_4-E(A_4)|}{\sqrt{D(A_4)}} \geq u_{1-\alpha/2}$.