

Odhad parametrů normálního rozdělení a testy hypotéz o těchto parametrech*

Věty o výběru z normálního rozdělení

Nechť $X_1, X_2, \dots, X_n$ je náhodný výběr z rozdělení $N(\mu, \sigma^2)$. Označme

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

výběrový průměr a

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

výběrový rozptyl. Veličina $\bar{X}$ je nestranným a konzistentním odhadem parametru μ a veličina S^2 je nestranným odhadem parametru σ^2 (viz Principy, tvrzení 8). Chceme-li vyjádřit chybu, jaké se při odhadech parametrů μ a σ^2 veličinami $\bar{X}$ a S^2 dopouštíme, je třeba znát rozdělení těchto veličin. O tom pojednává následující věta.

Věta 1. Necht' $X_1, X_2, \dots, X_n$ je náhodný výběr z rozdělení $N(\mu, \sigma^2)$. Pak

(a) $\bar{X} \sim N(\mu, \sigma^2/n)$.

(b) Jestliže $\sigma^2 > 0$ a $n \geq 2$, pak náhodná veličina $(n-1)S^2/\sigma^2$ má rozdělení χ_{n-1}^2 .

(c) Jestliže $n \geq 2$, pak veličiny $\bar{X}$ a S^2 jsou nezávislé.

Poznámky. Tvrzení (b) je důsledkem věty 2 z Principů, nikoliv však důsledkem zcela bezprostředním (přemýšlejte o tom). Z tohoto tvrzení vyplývá, že v případě výběru z normálního rozdělení je výběrový rozptyl konzistentním odhadem parametru σ^2 (obecně to neplatí). Vskutku, rozptyl rozdělení χ_{n-1}^2 je roven $2(n-1)$, a proto

$$2(n-1) = D\left[\frac{(n-1)S^2}{\sigma^2}\right] = \frac{(n-1)^2}{\sigma^4} \cdot D(S^2),$$

odkud plyne, že

$$D(S^2) = 2\sigma^4/(n-1) \xrightarrow{n \rightarrow \infty} 0.$$

Veličina S^2 je tedy konzistentním odhadem parametru σ^2 dle věty 5 z Principů.

Nezávislost výběrového průměru a rozptylu normálního rozdělení zformulovaná v tvrzení (c) je zcela netriviální vlastností tohoto rozdělení, kterou navíc žádné jiné rozdělení kromě normálního nemá.

Z tvrzení (a) předchozí věty vyplývá, že

$$(*) \quad \frac{\bar{X} - \mu}{\sigma} \cdot \sqrt{n} \sim N(0, 1).$$

Nahradíme-li v (*) neznámou hodnotu směrodatné odchylky σ jejím výběrovým ekvivalentem S , změní se rozdělení $N(0, 1)$ v Studentovo t rozdělení. Přesněji řečeno platí následující věta.

* Na úvodní pojednání o principech matematické statistiky se budeme v tomto textu odkazovat jako na *Principy*.

Věta 2. Necht' $X_1, X_2, \dots, X_n$ je náhodný výběr z rozdělení $N(\mu, \sigma^2)$. Pak

$$(**) \quad \frac{\bar{X} - \mu}{S} \cdot \sqrt{n} \sim t_{n-1}.$$

Důkaz. Je

$$\frac{\bar{X} - \mu}{S} \cdot \sqrt{n} = \frac{\frac{\bar{X} - \mu}{\sigma} \cdot \sqrt{n}}{\frac{S}{\sigma}} = \frac{\frac{\bar{X} - \mu}{\sigma} \cdot \sqrt{n}}{\sqrt{\frac{(n-1)S^2/\sigma^2}{n-1}}}.$$

Dle věty 1 jsou veličiny $(\bar{X} - \mu)\sqrt{n}/\sigma$ a $(n-1)S^2/\sigma^2$ vzájemně nezávislé, první z nich má rozdělení $N(0, 1)$ a druhá rozdělení χ^2_{n-1} . Tvrzení věty 2 tedy vyplývá z věty 3 z Principů.

Odhad střední hodnoty, test hypotézy o střední hodnotě

Označení. Necht' $u(\alpha)$, resp. $t_n(\alpha)$ jsou čísla vyhovující následujícím podmínkám: má-li veličina X rozdělení $N(0, 1)$, pak

$$P[X > u(\alpha)] = \alpha;$$

má-li veličina X rozdělení t_n , pak

$$P[|X| > t_n(\alpha)] = \alpha.$$

Tato čísla nazýváme *kritickými hodnotami rozdělení* $N(0, 1)$, resp. t_n . Povšimněte si, že označení kritických hodnot rozdělení $N(0, 1)$ a t_n není konzistentní, důvody této nekonzistence jsou však zcela tradiční a nemají žádný hlubší smysl.

Necht' $X_1, X_2, \dots, X_n$ je náhodný výběr z rozdělení $N(\mu, \sigma^2)$. Ze vztahu (*) vyplývá, že s pravděpodobností $1 - \alpha$ je

$$\left| \frac{\bar{X} - \mu}{\sigma} \cdot \sqrt{n} \right| < u(\alpha/2).$$

Podobně z (**) plyne, že s pravděpodobností $1 - \alpha$ je

$$\left| \frac{\bar{X} - \mu}{S} \cdot \sqrt{n} \right| < t_{n-1}(\alpha).$$

Odtud jednoduchou aritmetickou úpravou dospějeme k následujícímu výsledku.

Věta 3. Necht' $X_1, X_2, \dots, X_n$ je náhodný výběr z rozdělení $N(\mu, \sigma^2)$. Pak

(a) *Interval*

$$\left(\bar{X} - u(\alpha/2) \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + u(\alpha/2) \cdot \frac{\sigma}{\sqrt{n}} \right)$$

pokrývá hodnotu parametru μ s pravděpodobností $1 - \alpha$.

(b) *Interval*

$$\left(\bar{X} - t_{n-1}(\alpha) \cdot \frac{S}{\sqrt{n}}, \bar{X} + t_{n-1}(\alpha) \cdot \frac{S}{\sqrt{n}} \right)$$

pokrývá hodnotu parametru μ s pravděpodobností $1 - \alpha$.

Intervaly z předchozí věty představují *intervalové odhady (intervaly spolehlivosti) pro střední hodnotu normálního rozdělení*. První interval je přitom (*oboustranným*) $(1-\alpha) \cdot 100\%$ *intervalem spolehlivosti pro parametr μ při známé hodnotě parametru σ^2* , druhý z nich je *intervalovým odhadem parametru μ při neznámé hodnotě parametru σ^2* . (V aplikacích se zpravidla použije druhý z obou intervalů.) Číslo α se nazývá *koefficient spolehlivosti odhadu*.

Podobně je možné zkonstruovat *jednostranné intervaly spolehlivosti*. Tak například platí:
Interval

$$\left(-\infty, \bar{X} + t_{n-1}(2\alpha) \cdot \frac{S}{\sqrt{n}} \right),$$

a rovněž tak interval

$$\left(\bar{X} - t_{n-1}(2\alpha) \cdot \frac{S}{\sqrt{n}}, +\infty \right)$$

pokrývá hodnotu parametru μ s pravděpodobností $1-\alpha$.

Konstrukce intervalů spolehlivosti pro parametr μ úzce souvisí s testem hypotézy $H_0 : \mu = \mu_0$, kde μ_0 je dané reálné číslo. Je přitom nutné předem stanovit proti jaké alternativě budeme hypotézu H_0 testovat. Přirozená alternativa je $H_1 : \mu \neq \mu_0$; budeme jí říkat *oboustranná alternativa*. V některých případech lze ale na základě nějaké předběžné informace o parametru μ určitě jeho alternativní hodnoty předem vyloučit. V aplikacích se setkáváme s alternativami $H_1 : \mu > \mu_0$ a $H_1 : \mu < \mu_0$, které se nazývají *jednostranné*. Hypotézu H_0 zamítneme, jestliže nalezený interval spolehlivosti pro parametr μ hypotetickou hodnotu μ_0 neobsahuje; v případě oboustranné alternativy použijeme interval oboustranný, v případě jednostranné alternativy odpovídající interval jednostranný. Snadno nahlédneme, že hladina významnosti takového testu (tj. pravděpodobnost, že při aplikaci testu bude zamítnuta správná hypotéza) je rovna α , kde α je koefficient spolehlivosti odhadu parametru μ . Testujeme-li tedy hypotézu H_0 proti oboustranné alternativě, zamítneme ji tehdy, když interval

$$\left(\bar{X} - t_{n-1}(\alpha) \cdot \frac{S}{\sqrt{n}}, \bar{X} + t_{n-1}(\alpha) \cdot \frac{S}{\sqrt{n}} \right)$$

neobsahuje číslo μ_0 . K zamítání hypotézy H_0 dochází právě tehdy, když

$$(***) \quad |\bar{X} - \mu_0| \geq t_{n-1}(\alpha) \cdot \frac{S}{\sqrt{n}}.$$

V případě, že je splněna nerovnost (***), říkáme že rozdíl mezi výběrovým průměrem $\bar{X}$ a hypotetickou hodnotou μ_0 parametru μ je *statisticky významný (signifikantní)*. Číslo $t_{n-1}(\alpha) \cdot \frac{S}{\sqrt{n}}$ přitom představuje *nejmenší statisticky významný rozdíl* (angl. „Least Significant Difference“, zkráceně *LSD*). Kritérium (***) lze přepsat ve tvaru

$$\left| \frac{\bar{X} - \mu_0}{S} \cdot \sqrt{n} \right| \geq t_{n-1}(\alpha).$$

To znamená, že hypotéza H_0 se zamítá, jestliže $|T| \geq t_{n-1}(\alpha)$, kde

$$T = \frac{\bar{X} - \mu_0}{S} \cdot \sqrt{n}.$$

Náhodná veličina T představuje tzv. *testovací statistiku* pro hypotézu H_0 .

Kritérium pro zamítání hypotézy H_0 se vyjadřuje též následujícím způsobem. Množina

$$(-\infty, -t_{n-1}(\alpha)] \cup [t_{n-1}(\alpha), \infty)$$

se nazývá *kritickou oblastí (oblastí zamítání)* pro hypotézu H_0 ; hypotéza se zamítá, jestliže hodnota testovací statistiky T padne do kritické oblasti.

Analogicky lze zkonstruovat testy hypotézy $H_0: \mu = \mu_0$ proti jednostranným alternativám $H_1: \mu > \mu_0$ a $H_1: \mu < \mu_0$. Popis testu při všech třech možných alternativách (za použití testovací statistiky T) je shrnut v následující větě:

Věta 4. *Testujme hypotézu $H_0: \mu = \mu_0$ o parametru μ rozdělení $N(\mu, \sigma^2)$ prostřednictvím konfrontace této hypotézy s empiricky zjištěnou hodnotou výběrového průměru $\bar{X}$ a výběrového rozptylu S^2 . Položme*

$$T = \frac{\bar{X} - \mu_0}{S} \cdot \sqrt{n}$$

a stanovme následující kritéria pro zamítání hypotézy H_0 :

- (1) *Hypotéza $H_0: \mu = \mu_0$ se proti alternativě $H_1: \mu \neq \mu_0$ zamítá, jestliže $|T| \geq t_{n-1}(\alpha)$.*
- (2) *Hypotéza $H_0: \mu = \mu_0$ se proti alternativě $H_1: \mu > \mu_0$ zamítá, jestliže $T \geq t_{n-1}(2\alpha)$.*
- (3) *Hypotéza $H_0: \mu = \mu_0$ se proti alternativě $H_1: \mu < \mu_0$ zamítá, jestliže $T \leq -t_{n-1}(2\alpha)$.*

Hladina významnosti tohoto testu je rovna α . Jde přitom o test, který má ze všech možných testů hypotézy $H_0: \mu = \mu_0$ největší možnou sílu. Jinak řečeno, pravděpodobnost zamítnutí nesprávné hypotézy je u tohoto testu při libovolné předem zadané hladině významnosti α největší možná.

Definice 1. Test popsaný v předchozí větě se nazývá *jednovýběrový t test*.

Poznámka. K testovací statistice jednovýběrového t testu lze dospět přímo, aniž bychom předtím odvozovali vzorec pro interval spolehlivosti pro parametr μ . Jestliže je totiž hypotéza $H_0: \mu = \mu_0$ správná, pak v důsledku věty 2 má veličina

$$T = \frac{\bar{X} - \mu_0}{S} \cdot \sqrt{n}$$

rozdělení t_{n-1} . Použijeme-li tedy pro zamítání hypotézy H_0 kritérií popsaných ve větě 4, bude hypotéza zamítnuta právě tehdy, když absolutní hodnota testovací statistiky T je tak velká, že takto velká hodnota je u této statistiky za předpokladu platnosti hypotézy H_0 jen velmi málo pravděpodobná. Skutečnost, že hodnota statistiky T je velká přitom znamená, že mezi výběrovým průměrem $\bar{X}$ a hypotetickou hodnotou μ_0 parametru μ je veliký rozdíl; právě proto má popsaný test velkou sílu a s velkou pravděpodobností odhalí skutečnost, že hypotéza H_0 správná není.

ÚLOHY

1. Předpokládejme, že výčetní tloušťka stromu se řídí normálním zákonem rozdělení pravděpodobností. Byla změřena tloušťka šestnácti náhodně vybraných stromů (v centimetrech); výsledky jsou zaznamenány v následující tabulce:

Tloušťka	18	20	24	28	32	38
Četnost	1	2	4	5	3	1

- a) Nalezněte interval, který pokrývá hodnotu střední výčetní tloušťky s pravděpodobností 0,95.
 b) Předpokládejme (např. na základě předchozí zkušenosti či s přihlédnutím k výše naměřeným hodnotám), že směrodatná odchylka tloušťky není větší než osm centimetrů. Jaký nejmenší počet stromů musíme vybrat, jestliže požadujeme, aby chyba, které se dopustíme při odhadu střední tloušťky aritmetickým průměrem tlouštěk vybraných stromů, byla s 95% pravděpodobností menší než jeden centimetr?

Řešení. Označme $X_1, X_2, \dots, X_n$ tloušťky vybraných stromů a předpokládejme, že jde o náhodný výběr z rozdělení $N(\mu, \sigma^2)$.

- a) Je třeba odhadnout hodnotu parametru μ . Průměrná tloušťka změřených stromů je $\bar{X} = 26,75$, výběrový rozptyl je $S^2 = 27,67$. Veličina $\bar{X}$ je maximálně věrohodným odhadem parametru μ , přičemž 95% interval spolehlivosti pro parametr μ je dán předpisem

$$\bar{X} \pm t_{n-1}(0,05) \cdot \frac{S}{\sqrt{n}},$$

kde $n = 16$ je velikost výběru a $t_{n-1}(0,05) = t_{15}(0,05) = 2,131$ je kritická hodnota rozdělení t_{n-1} . Po dosažení dospějeme k intervalu

$$26,75 \pm 2,80.$$

Tento interval pokrývá hodnotu střední tloušťky μ s pravděpodobností 0,95.

- b) Je-li číslo σ^2 známé, pak $(1 - \alpha) \cdot 100\%$ interval spolehlivosti pro parametr μ má meze

$$\bar{X} \pm u(\alpha/2) \cdot \frac{\sigma}{\sqrt{n}},$$

kde $u(\alpha/2)$ je kritická hodnota rozdělení $N(0, 1)$. Pro určení minimálního počtu vybraných stromů potřebného k dosažení požadované přesnosti odhadu střední tloušťky počítejme s největší možnou předpokládanou variabilitou tlouštěk, tj. s hodnotou $\sigma = 8$. Požadujeme, aby

$$u(0,025) \cdot \frac{\sigma}{\sqrt{n}} < 1.$$

Odtud plyne, že musí být $n > [u(0,025) \cdot \sigma]^2 = (1,96 \cdot 8)^2 = 246$. Je tedy nutno vybrat a změřit řádově alespoň dvě stě padesát stromů.

2. Necht' chyba, které se dopouštíme při měření výšky stromu určitým typem výškoměru, má rozdělení $N(0, \sigma^2)$, přičemž σ^2 je neznámý parametr. Dejme tomu, že určitý strom má výšku h metrů. Za účelem přesného určení této výšky bylo provedeno devět nezávislých opakovaných měření s následujícím výsledkem (v metrech):

20,29 20,11 19,84 19,81 19,83 20,16 19,79 19,96 19,72

- a) Určete 99 % interval spolehlivosti pro výšku h .
 b) Jak se změní řešení úlohy a), jestliže číslo σ , vyjadřující velikost náhodné chyby měření, znáte? Konkrétně předpokládejte, že $\sigma = 0,15$.

Řešení. Necht' n je počet měření a X_i , kde $i = 1, 2, \dots, n$, je výsledek i -tého měření. Označíme-li e_i chybu i -tého měření, pak $X_i = h + e_i$. Předpokládáme přitom, že tyto chyby mají normální rozdělení, nulovou střední hodnotu a stejný rozptyl σ^2 . Předpoklad nulové střední hodnoty veličin e_i znamená, že během měření nedochází k systematickým chybám, stejný rozptyl těchto veličin pak vyjadřuje stálost podmínek, za nichž jsou měření prováděna. Normalita rozdělení chyb přitom odpovídá implicitnímu předpokladu o jejich zcela náhodné povaze. Je tedy

$$e_i \sim N(0, \sigma^2),$$

odkud plyne, že

$$X_i \sim N(h, \sigma^2)$$

pro každé $i = 1, 2, \dots, n$. Přitom σ^2 je neznámá konstanta nezávislá na i a charakterizující zvolený proces měření. Předpokládáme-li dále, že chyby $e_1, e_2, \dots, e_n$ jsou vzájemně nezávislé, pak vektor $X_1, X_2, \dots, X_n$ je náhodným výběrem z rozdělení $N(h, \sigma^2)$. Chceme přitom odhadnout střední hodnotu h tohoto rozdělení. Výběrový průměr je $\bar{X} \doteq 19,95$, výběrový rozptyl $S^2 \doteq 0,04$ a směrodatná odchylka $S \doteq 0,2$.

a) Hledaný 99% interval spolehlivosti pro parametr h (skutečnou výšku stromu) je dán předpisem

$$\bar{X} \pm t_{n-1}(0,01) \cdot \frac{S}{\sqrt{n}},$$

kde $n = 9$ je počet provedených měření. Z tabulek zjistíme, že $t_{n-1}(0,01) = t_8(0,01) = 3,355$. Po dosazení obdržíme výsledek $19,95 \pm 0,22$.

b) Předpokládáme-li, že $\sigma = 0,15$, pak 99% interval spolehlivosti pro parametr h je

$$\bar{X} \pm u(0,005) \cdot \frac{\sigma}{\sqrt{n}}.$$

Z tabulek zjistíme, že $u(0,005) = 2,576$. Po dosazení dostaneme výsledek $19,95 \pm 0,13$.

3. Při výrobě mincí je stanovena hmotnost mince pět gramů. Je podezření, že na materiálu se systematicky šetří. Testujte hypotézu, že tomu tak není. Použijte výsledků namátkové kontroly, při níž bylo náhodně vybráno deset mincí, a poté zjištěna jejich hmotnost; zaznamenané hmotnosti vybraných mincí (v gramech) jsou přitom následující:

4,91 5,02 4,88 4,79 4,89 4,72 5,01 4,97 4,86 4,93

Řešení. Označme $X_1, X_2, \dots, X_n$ hmotnosti vybraných mincí a předpokládejme, že tyto hmotnosti představují náhodný výběr z rozdělení $N(\mu, \sigma^2)$. Testujme hypotézu $H_0: \mu = 5$, tj. hypotézu, že střední hmotnost vyráběných mincí je rovna stanovené normě, a to nejprve proti oboustranné alternativě $H_1: \mu \neq 5$. Provedeme jednovýběrový t test, a to na základě testovací statistiky

$$T = \frac{\bar{X} - 5}{S} \cdot \sqrt{n},$$

o níž víme, že má za předpokladu správnosti hypotézy H_0 rozdělení t_{n-1} . V našem případě je

$$n = 10, \bar{X} = 4,898, S \doteq 0,0937, |T| \doteq 3,44 > 3,25 = t_9(0,01).$$

Hypotéza $H_0: \mu = 5$ se proto zamítá proti alternativě $H_1: \mu \neq 5$ na hladině významnosti $\alpha = 0,01$. Vzhledem k tomu, $3,44 \doteq t_9(0,0074)$, je hladina významnosti testu ve skutečnosti o něco málo menší než 0,01.

V naší úloze však a priori vylučujeme možnost, že se mince vyrábějí systematicky těžší než je stanovená norma, správná alternativa k testované hypotéze je tudíž jednostranná alternativa $H_1: \mu < 5$. Hladina významnosti se přechodem k této jednostranné alternativě snižuje na polovinu, je tedy menší než 0,004.

Jednovýběrový t test lze na předem dané hladině významnosti α realizovat též tak, že určíme $(1 - \alpha) \cdot 100\%$ interval spolehlivosti pro parametr μ a povšimneme si, zda tento interval obsahuje či neobsahuje testovanou hodnotu (zde číslo pět). Například oboustranný 99% interval spolehlivosti pro parametr μ je $(4,802; 4,994)$. Tento interval neobsahuje číslo 5 (nýbrž leží celý nalevo od tohoto čísla), což odpovídá skutečnosti, že na hladině významnosti $\alpha = 0,01$ se hypotéza $H_0: \mu = 5$ proti alternativě $H_1: \mu \neq 5$ zamítá. Jednostranné alternativě pak odpovídá jednostranný interval spolehlivosti $(-\infty; 4,982)$.

4. *Párový t test.* U devíti vzorků půdy byla stanovena koncentrace jisté škodlivé látky. Použilo se přitom dvou metod. Jedna z nich je klasická, zatímco druhá je nová. Na základě zjištěných výsledků (viz následující tabulka) posuďte, zda první metoda dává systematicky větší nebo menší výsledky než druhá.

Klasická metoda	2,68	2,60	2,43	2,85	2,94	2,70	2,68	2,98	2,85
Nová metoda	2,36	2,41	2,39	2,90	2,82	2,73	2,58	2,89	2,78
Rozdíl	0,32	0,19	0,04	-0,05	0,12	-0,03	0,10	0,09	0,07

Řešení. Předpokládejme, že rozdíl mezi hodnotami výsledků jednotlivých metod se řídí normálním zákonem rozdělení pravděpodobností a testujme hypotézu, že střední hodnota tohoto rozdílu je rovna nule. Tím je úloha převedena na jednovýběrový t test. Hodnota testovací statistiky je $T = 2,53 \geq 2,306 = t_8(0,05)$. Systematický rozdíl ve výsledcích obou metod je tím prokázán.

Odhad rozptylu, test hypotézy o rozptylu

Označení. Symbolem $\chi_n^2(\alpha)$ budeme označovat *kritickou hodnotu rozdělení* χ_n^2 , tj. takovou hodnotu, která je veličinou s rozdělením χ_n^2 překročena s pravděpodobností α .

Nechť $X_1, X_2, \dots, X_n$ je náhodný výběr z rozdělení $N(\mu, \sigma^2)$. Z věty 1(b) vyplývá, že s pravděpodobností $1 - \alpha$ je

$$\chi_{n-1}^2(1 - \alpha/2) < \frac{(n-1)S^2}{\sigma^2} < \chi_{n-1}^2(\alpha/2).$$

Odtud pak lehce dospějeme k následujícímu výsledku.

Věta 5. *Nechť $X_1, X_2, \dots, X_n$ je náhodný výběr z rozdělení $N(\mu, \sigma^2)$. Pak interval*

$$\left(\frac{(n-1)S^2}{\chi_{n-1}^2(\alpha/2)}, \frac{(n-1)S^2}{\chi_{n-1}^2(1 - \alpha/2)} \right)$$

pokrývá hodnotu parametru σ^2 s pravděpodobností $1 - \alpha$.

Interval z předchozí věty představuje intervalový odhad pro rozptyl normálního rozdělení; přesněji jde o oboustranný $(1 - \alpha) \cdot 100\%$ interval spolehlivosti. Analogicky lze zkonstruovat jednostranné intervaly spolehlivosti. Konkrétně:

Interval

$$\left(-\infty, \frac{(n-1)S^2}{\chi_{n-1}^2(1 - \alpha)} \right),$$

a rovněž tak interval

$$\left(\frac{(n-1)S^2}{\chi_{n-1}^2(\alpha)}, +\infty \right)$$

pokrývá hodnotu parametru σ^2 s pravděpodobností $1 - \alpha$.

Konstrukce intervalu spolehlivosti pro parametr σ^2 souvisí s testem hypotézy $H_0 : \sigma^2 = \sigma_0^2$, kde

σ_0^2 je dané reálné číslo. Tuto hypotézu zamítneme, jestliže nalezený interval spolehlivosti pro parametr σ^2 hypotetickou hodnotu σ_0^2 neobsahuje. Testujeme-li tedy např. hypotézu H_0 proti oboustranné alternativě $H_1 : \sigma^2 \neq \sigma_0^2$, zamítneme ji tehdy, když interval

$$\left(\frac{(n-1)S^2}{\chi_{n-1}^2(\alpha/2)}, \frac{(n-1)S^2}{\chi_{n-1}^2(1-\alpha/2)} \right)$$

neobsahuje číslo σ_0^2 . Snadno nahlédneme, že hladina významnosti takového testu je rovna α . Je přitom zjevné, že k zamítání hypotézy H_0 dochází právě tehdy, když

$$\frac{(n-1)S^2}{\sigma_0^2} < \chi_{n-1}^2(1-\alpha/2) \quad \text{nebo} \quad \frac{(n-1)S^2}{\sigma_0^2} > \chi_{n-1}^2(\alpha/2).$$

Náhodná veličina

$$T = \frac{(n-1)S^2}{\sigma_0^2}$$

představuje tedy testovací statistiku pro hypotézu H_0 . Hypotéza se zamítá, jestliže $T < \chi_{n-1}^2(1-\alpha/2)$ nebo $T > \chi_{n-1}^2(\alpha/2)$. Analogicky lze zkonstruovat testy hypotézy $H_0 : \sigma^2 = \sigma_0^2$ proti jednostranným alternativám $H_1 : \sigma^2 > \sigma_0^2$ a $H_1 : \sigma^2 < \sigma_0^2$. Popis testu při všech třech možných alternativách (za použití testovací statistiky T) je shrnut v následující větě:

Věta 6. *Testujme hypotézu $H_0 : \sigma^2 = \sigma_0^2$ o parametru σ^2 normálního rozdělení prostřednictvím konfrontace této hypotézy s empiricky zjištěnou hodnotou výběrového rozptylu S^2 . Položme*

$$T = \frac{(n-1)S^2}{\sigma_0^2}$$

a stanovme následující kritéria pro zamítání hypotézy H_0 :

(1) *Hypotéza $H_0 : \sigma^2 = \sigma_0^2$ se proti alternativě $H_1 : \sigma^2 \neq \sigma_0^2$ zamítá, jestliže*

$$T < \chi_{n-1}^2(1-\alpha/2) \quad \text{nebo} \quad T > \chi_{n-1}^2(\alpha/2).$$

(2) *Hypotéza $H_0 : \sigma^2 = \sigma_0^2$ se proti alternativě $H_1 : \sigma^2 > \sigma_0^2$ zamítá, jestliže $T > \chi_{n-1}^2(\alpha)$.*

(3) *Hypotéza $H_0 : \sigma^2 = \sigma_0^2$ se proti alternativě $H_1 : \sigma^2 < \sigma_0^2$ zamítá, jestliže $T < \chi_{n-1}^2(1-\alpha)$.*

Hladina významnosti tohoto testu je rovna α . Jde přitom o nejsilnější možný test.

Poznámka. Kritéria pro zamítání hypotézy $H_0 : \sigma^2 = \sigma_0^2$ zformulovaná v předchozí větě se lehce zapamatují. Hypotéza se totiž zamítá tehdy, jestliže poměr empiricky zjištěné hodnoty výběrového rozptylu S^2 a hypotetické hodnoty σ_0^2 rozptylu σ^2 je příliš malý (resp. příliš velký).

ÚLOHY

5. a) Necht' chyba, které se dopouštíme při zjišťování výšky stromu určitým typem výškoměru, má rozdělení $N(0, \sigma^2)$. Chceme odhadnout hodnotu parametru σ^2 . Za tím účelem provedeme $n = 100$ opakovaných měření jedné konkrétní výšky; předpokládáme přitom, že výsledky jednotlivých

vých měření jsou vzájemně nezávislé. Z naměřených hodnot vypočítáme výběrovou směrodatnou odchylku; její hodnota necht' je $S = 0,137$. Určete 95% oboustranný interval spolehlivosti pro parametr σ^2 . Zcela speciálně testujte hypotézu, že $\sigma = 0,15$.

b) Řešte předchozí úlohu pro $n = 1000$ měření s výběrovou směrodatnou odchylkou $S = 0,144$. (Provedeme-li více měření, obdržíme přesnější odhad.)

Řešení. a) Podobně jako v úloze 2 předpokládáme, že chyby $e_1, e_2, \dots, e_n$ jednotlivých měření jsou vzájemně nezávislé veličiny s rozdělením $N(0, \sigma^2)$. Jinak řečeno vektor $e_1, e_2, \dots, e_n$ je náhodným výběrem z rozdělení $N(0, \sigma^2)$ a naším úkolem je odhadnout rozptyl σ^2 tohoto rozdělení. Bodovým odhadem parametru σ^2 je veličina $S^2 = 0,137^2$; jde přitom o nestraný a maximálně věrohodný odhad. (Je třeba si uvědomit, že výběrový rozptyl naměřených hodnot je stejný jako výběrový rozptyl chyb, kterých se při měření dopustíme.) Intervalový odhad parametru σ^2 zkonstruujeme následujícím způsobem. Platí, že

$$\frac{99S^2}{\sigma^2} \sim \chi_{99}^2,$$

a tudíž s pravděpodobností 0,95 je

$$\chi_{99}^2(0,975) < \frac{99S^2}{\sigma^2} < \chi_{99}^2(0,025).$$

Příslušné kritické hodnoty rozdělení χ_{99}^2 jsou

$$\chi_{99}^2(0,025) = 128,42 \quad \text{a} \quad \chi_{99}^2(0,975) = 73,36.$$

To znamená, že s pravděpodobností 0,95 je

$$73,36 < \frac{99S^2}{\sigma^2} < 128,42.$$

Odtud vyplývá, že interval

$$\left(\frac{99S^2}{128,42}, \frac{99S^2}{73,36} \right) = (0,01447; 0,02533)$$

pokrývá neznámou hodnotu parametru σ^2 s pravděpodobností 0,95. Tudíž 95% oboustranný interval spolehlivosti pro směrodatnou odchylku σ je $(0,12; 0,16)$. To ale znamená, že hypotéza $H_0: \sigma^2 = 0,15^2$ se proti alternativě $H_1: \sigma^2 \neq 0,15^2$ na hladině významnosti $\alpha = 0,05$ nezamítá.

b) Podobně jako v úloze a) zjistíme, že 95% oboustranný interval spolehlivosti pro rozptyl σ^2 je

$$\left(\frac{999S^2}{\chi_{999}^2(0,025)}, \frac{999S^2}{\chi_{999}^2(0,975)} \right).$$

Kritické hodnoty $\chi_{999}^2(0,025)$ a $\chi_{999}^2(0,975)$ rozdělení χ_{999}^2 lze s poměrně velikou přesností určit pomocí aproximace tohoto rozdělení normálním rozdělením $N(999, 2 \cdot 999)$. Obdržíme tak kritické hodnoty $999 \pm 1,96 \cdot \sqrt{2 \cdot 999}$, tj. $999 \pm 87,61$. Po dosazení za tyto kritické hodnoty a za empiricky zjištěnou hodnotu rozptylu $S^2 = 0,137^2$ získáme pro 95% interval spolehlivosti pro parametr σ^2 vyjádření $(0,019; 0,023)$. Odpovídající intervalový odhad parametru σ je $(0,138; 0,151)$.

6. Pevnost vlákna bavlněné příze lze pokládat za náhodnou veličinu s rozdělením $N(\mu, \sigma^2)$. Je-li $\sigma^2 > 0,36 \text{ kg}^2$, vznikají potíže při tkaní. Při zkoušce jedenácti náhodně vybraných vláken byly zjištěny tyto hodnoty jejich pevnosti:

5,3 3,0 4,8 3,6 4,1 2,5 4,7 2,4 3,2 3,8 4,4

Je třeba posoudit, zda je příze vyhovující.

Řešení. Testujme hypotézu $H_0 : \sigma^2 = 0,36$ proti alternativě $H_1 : \sigma^2 > 0,36$. Test lze realizovat buď tak, že vypočteme jednostranný interval spolehlivosti pro parametr σ^2 , nebo pomocí testovací statistiky. Bodovým odhadem parametru σ^2 je výběrový rozptyl $S^2 = 0,92$. Přitom 95 % jednostranný interval spolehlivosti pro parametr σ^2 je roven

$$\left(\frac{10 \cdot S^2}{\chi_{10}^2(0,05)}, +\infty \right),$$

zatímco testovací statistika pro hypotézu $H_0 : \sigma^2 = 0,36$ je

$$T = \frac{10 \cdot S^2}{0,36}.$$

Po dosazení za číslo S^2 obdržíme pro interval spolehlivosti pro parametr σ^2 vyjádření $(0,5; +\infty)$; tento interval pokrývá neznámou hodnotu parametru σ^2 s pravděpodobností 0,95. Hypotéza $H_0 : \sigma^2 = 0,36$ se tudíž proti alternativě $H_1 : \sigma^2 > 0,36$ zamítá na hladině významnosti $\alpha = 0,05$, neboť interval $(0,5; +\infty)$ neobsahuje číslo 0,36. Dále

$$T \doteq 25,6 > 23,209 = \chi_{10}^2(0,01),$$

což znamená, že hypotéza H_0 se proti alternativě H_1 zamítá i na hladině významnosti $\alpha = 0,01$. Označíme-li tedy přízi za nevyhovující, pak pravděpodobnost, že je tento závěr špatný, nepřekračuje hodnotu 0,01.

7. V úloze 1b) předpokládáme, že směrodatná odchylka výčetní tloušťky stromu nepřekračuje hodnotu osm centimetrů. Ukažte, že tento předpoklad je v souladu se získanými experimentálními daty.

Řešení. Interval $(-\infty; 7,56)$ pokrývá hodnotu směrodatné odchylky σ s pravděpodobností 0,95.

Test hypotézy o rovnosti středních hodnot dvou normálních rozdělení

Nechť $X_1, X_2, \dots, X_m$ je náhodný výběr z rozdělení $N(\mu_1, \sigma_1^2)$ a $Y_1, Y_2, \dots, Y_n$ je náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$. To znamená, že $X_1, X_2, \dots, X_m$ jsou vzájemně nezávislé náhodné veličiny s rozdělením $N(\mu_1, \sigma_1^2)$ a $Y_1, Y_2, \dots, Y_n$ jsou vzájemně nezávislé náhodné veličiny s rozdělením $N(\mu_2, \sigma_2^2)$. Předpokládejme dále, že hodnoty veličin $Y_1, Y_2, \dots, Y_n$ nezávisí na hodnotách veličin $X_1, X_2, \dots, X_m$ (a naopak), tj. že jde o vzájemně *nezávislé náhodné výběry*. Chceme testovat hypotézu $H_0 : \mu_1 = \mu_2$. Označme

$$\bar{X} = \frac{1}{m} \sum_{i=1}^m X_i, \quad \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i, \quad S_1^2 = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X})^2, \quad S_2^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2.$$

Věta 7. Náhodná veličina

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

má rozdělení $N(0, 1)$.

Důkaz. Víme, že

$$\bar{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{m}\right) \quad \text{a} \quad \bar{Y} \sim N\left(\mu_2, \frac{\sigma_2^2}{n}\right).$$

Vzhledem k tomu, že výběry $X_1, X_2, \dots, X_m$ a $Y_1, Y_2, \dots, Y_n$ jsou vzájemně nezávislé, jsou nezávislé též veličiny $\bar{X}$ a $\bar{Y}$, a proto

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}\right).$$

Odtud pak plyne, že

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0, 1),$$

což bylo dokázat. $\square$

Důsledek. Necht' $H_0 : \mu_1 = \mu_2$ je správná hypotéza. Pak náhodná veličina

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

má rozdělení $N(0, 1)$.

Náhodnou veličinu T z předchozího důsledku lze považovat za testovací statistiku pro hypotézu H_0 . Pokud budeme zamítat hypotézu H_0 tehdy, když hodnoty statistiky T padnou do kritické oblasti rozdělení $N(0, 1)$, tj. tehdy, když $|T| \geq u(\alpha/2)$, obdržíme test s hladinou významnosti α . Při aplikaci tohoto testu se ovšem předpokládá, že hodnoty rozptylů σ_1^2 a σ_2^2 jsou známy; v běžných aplikacích tomu tak ale zpravidla nebývá. Podobně jako v jednovýběrovém t testu se přitom nabízí myšlenka nahradit ve statistice T hodnoty neznámých rozptylů σ_1^2 a σ_2^2 jejich odhady S_1^2 a S_2^2 a zjistit jak se přitom rozdělení statistiky T změní. Bohužel však v případě, že o parametrech σ_1^2 a σ_2^2 nemáme vůbec žádnou informaci, neumíme toto rozdělení exaktně vyčíslit.

Kdybychom však předpokládali, že hodnoty rozptylů σ_1^2 a σ_2^2 jsou stejné, pak lze postupovat následujícím způsobem. Označme $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Veličina T z věty 7 pak nabývá tvaru

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sigma \cdot \sqrt{\frac{1}{m} + \frac{1}{n}}}.$$

Lze přitom ukázat, že veličina

$$S^2 = \frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2}$$

je nestranným (a maximálně věrohodným) odhadem parametru σ^2 a že veličiny $\bar{X} - \bar{Y}$ a S^2 jsou vzájemně nezávislé. Odtud lze pomocí věty 1 z tohoto pojednání a věty 3 z Principů vyvodit následu-

jící výsledek:

Věta 8. *Nechť $\sigma_1^2 = \sigma_2^2$. Pak náhodná veličina*

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S \cdot \sqrt{\frac{1}{m} + \frac{1}{n}}}$$

má rozdělení t_{m+n-2} .

Důsledek. *Nechť $H_0 : \mu_1 = \mu_2$ je správná hypotéza (a necht' navíc $\sigma_1^2 = \sigma_2^2$). Pak náhodná veličina*

$$T = \frac{\bar{X} - \bar{Y}}{S \cdot \sqrt{\frac{1}{m} + \frac{1}{n}}}$$

má rozdělení t_{m+n-2} .

Veličinu T z předchozího důsledku lze tedy za předpokladu, že $\sigma_1^2 = \sigma_2^2$, považovat za testovací statistiku pro hypotézu $H_0 : \mu_1 = \mu_2$. Hypotéza $H_0 : \mu_1 = \mu_2$ se zamítá proti oboustranné alternativě $H_1 : \mu_1 \neq \mu_2$, pokud $|T| \geq t_{m+n-2}(\alpha)$. Proti alternativě $H_1 : \mu_1 > \mu_2$ se hypotéza H_0 zamítá, jestliže $T \geq t_{m+n-2}(2\alpha)$ a proti alternativě $H_1 : \mu_1 < \mu_2$ se zamítá tehdy, když $T \leq -t_{m+n-2}(2\alpha)$. Souhrnně řečeno, proti jednostranné alternativě se hypotéza H_0 zamítá, jestliže $|T| \geq t_{m+n-2}(2\alpha)$. Hladina významnosti tohoto testu je rovna α . Lze přitom ukázat, že za předpokladu $\sigma_1^2 = \sigma_2^2$ je popsáný test nejsilnějším možným testem hypotézy $H_0 : \mu_1 = \mu_2$.

Poznámka. Jestliže $m = n$, pak

$$S^2 = \frac{S_1^2 + S_2^2}{2}.$$

Definice 2. Výše popsáný test se nazývá *dvouvýběrový t test*.

Prohlédneme-li si pozorně kritérium pro zamítání hypotézy $H_0 : \mu_1 = \mu_2$ v právě popsaném dvouvýběrovém t testu, vidíme, že proti oboustranné alternativě $H_1 : \mu_1 \neq \mu_2$ se tato hypotéza zamítá právě tehdy, když

$$|\bar{X} - \bar{Y}| \geq t_{m+n-2}(\alpha) \cdot S \cdot \sqrt{\frac{1}{m} + \frac{1}{n}}.$$

V takovém případě říkáme, že rozdíl $|\bar{X} - \bar{Y}|$ je *statisticky významný*. Číslo

$$LSD = t_{m+n-2}(\alpha) \cdot S \cdot \sqrt{\frac{1}{m} + \frac{1}{n}}$$

přitom představuje *nejmenší (statisticky) významný rozdíl*.

Dvouvýběrový t test hypotézy $H_0 : \mu_1 = \mu_2$ proti alternativě $H_1 : \mu_1 \neq \mu_2$ lze při dané hladině významnosti α realizovat též tak, že zkonstruujeme oboustranný $(1 - \alpha) \cdot 100\%$ interval spolehlivosti pro rozdíl $\mu_1 - \mu_2$. Z věty 8 vyplývá, že takový interval je dán předpisem

$$\bar{X} - \bar{Y} \pm t_{m+n-2}(\alpha) \cdot S \cdot \sqrt{\frac{1}{m} + \frac{1}{n}}.$$

Hypotéza H_0 se přitom zamítá, jestliže tento interval neobsahuje nulu.

Pokud máme pádný důvod předpokládat, že $\sigma_1^2 \neq \sigma_2^2$, je možno použít k testu hypotézy $H_0 : \mu_1 = \mu_2$ nějakou obdobu dvouvýběrového t testu, v níž se nerovnost rozptylů explicitně předpokládá. Jako testovací statistika pak obvykle slouží veličina

$$T^* = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}}.$$

Rozdělení pravděpodobností veličiny T^* (za platnosti hypotézy H_0), a tedy ani kritické hodnoty tohoto rozdělení však neumíme, jak již bylo poznamenáno, přesně vyčíslit. Jestliže však jsou čísla m a n dost velká, lze tyto kritické hodnoty určit alespoň přibližně. Často bývá používán tzv. *Cochranův-Coxův test*, který pracuje s modifikovanou kritickou hodnotou

$$t^*(\alpha) = \frac{v_1 t_{m-1}(\alpha) + v_2 t_{n-1}(\alpha)}{v_1 + v_2},$$

kde

$$v_1 = \frac{s_1^2}{m} \quad \text{a} \quad v_2 = \frac{s_2^2}{n}.$$

Hypotéza H_0 se zamítá, jestliže $|T^*| \geq t^*(\alpha)$. *Satterthwaiteův test* a *Welchův test* pak pracují s modifikovaným počtem stupňů volnosti, a to

$$f = \frac{(v_1 + v_2)^2}{\frac{v_1^2}{m-1} + \frac{v_2^2}{n-1}}$$

u Satterthwaiteova a

$$h = \frac{(v_1 + v_2)^2}{\frac{v_1^2}{m+1} + \frac{v_2^2}{n+1}} - 2$$

u Welchova testu. Hypotéza H_0 se zamítá, jestliže $|T^*| \geq t_f(\alpha)$, resp. $|T^*| \geq t_h(\alpha)$. Hladiny významnosti všech tří testů jsou při oboustranné alternativě asymptoticky (tj. pro $m, n \rightarrow \infty$) rovny α .

Test hypotézy o rovnosti rozptylů dvou normálních rozdělení

Předpokládejme opět, že $X_1, X_2, \dots, X_m$ je náhodný výběr z rozdělení $N(\mu_1, \sigma_1^2)$, $Y_1, Y_2, \dots, Y_n$ je náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$ a že tyto výběry jsou vzájemně nezávislé. Chceme testovat hypotézu $H_0 : \sigma_1^2 = \sigma_2^2$. Tak jako u dvouvýběrového t testu označme

$$S_1^2 = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X})^2 \quad \text{a} \quad S_2^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2.$$

Řešení úlohy spočívá ve vyčíslení rozdělení veličiny S_1^2/S_2^2 za předpokladu, že H_0 je správná hypotéza; konkrétně se ukáže, že jde o F rozdělení (stačí spojit větu 1(b) z tohoto pojednání s větou 4 z Principů). Důvodem k zamítnutí hypotézy H_0 jsou zřejmě příliš malé či příliš velké hodnoty poměru S_1^2/S_2^2 (menší či větší než hodnoty kritické). Popíšeme postup tohoto testování podrobněji. Nejprve přesně zformulujeme základní věty, na nichž je toto testování založeno.

Označení. Symbolem $F_{m,n}(\alpha)$ budeme označovat *kritickou hodnotu rozdělení* $F_{m,n}$, tj. takovou hodnotu, která je veličinou s rozdělením $F_{m,n}$ překročena s pravděpodobností α . Lze ukázat, že $F_{m,n}(1-\alpha) = F_{n,m}^{-1}(\alpha)$.

Věta 9. *Náhodná veličina*

$$T = \frac{\frac{S_1^2}{S_2^2}}{\frac{\sigma_1^2}{\sigma_2^2}}$$

má rozdělení $F_{m-1,n-1}$.

Důsledek. *Necht' $H_0 : \sigma_1^2 = \sigma_2^2$ je správná hypotéza. Pak náhodná veličina $T = S_1^2/S_2^2$ má rozdělení $F_{m-1,n-1}$.*

Hodnota veličiny T z předchozí věty se zřejmě s pravděpodobností $1-\alpha$ nachází v intervalu $(F_{m,n}(1-\alpha/2), F_{m,n}(\alpha/2))$. Odtud bezprostředně vyplývá následující věta.

Věta 10. *Interval*

$$\left(\frac{S_1^2/S_2^2}{F_{m-1,n-1}(\alpha/2)}, \frac{S_1^2/S_2^2}{F_{m-1,n-1}(1-\alpha/2)} \right) = \left(\frac{S_1^2/S_2^2}{F_{m-1,n-1}(\alpha/2)}, \frac{S_1^2/S_2^2}{F_{n-1,m-1}^{-1}(\alpha/2)} \right)$$

pokrývá hodnotu podílu σ_1^2/σ_2^2 s pravděpodobností $1-\alpha$. Tento interval tedy představuje $(1-\alpha) \cdot 100\%$ interval spolehlivosti pro parametr σ_1^2/σ_2^2 .

Nyní můžeme přistoupit k vlastnímu popisu testu hypotézy $H_0 : \sigma_1^2 = \sigma_2^2$. Hypotéza $H_0 : \sigma_1^2 = \sigma_2^2$ se zamítá proti alternativě $H_1 : \sigma_1^2 \neq \sigma_2^2$ tehdy, jestliže interval

$$\left(\frac{S_1^2/S_2^2}{F_{m-1,n-1}(\alpha/2)}, \frac{S_1^2/S_2^2}{F_{m-1,n-1}(1-\alpha/2)} \right)$$

neobsahuje číslo 1 (tj. tu hodnotu, které podíl σ_1^2/σ_2^2 nabývá tehdy, když H_0 je správná hypotéza.) To nastane právě tehdy, když

$$(\text{****}) \quad T < F_{m-1,n-1}(1-\alpha/2) \quad \text{nebo} \quad T > F_{m-1,n-1}(\alpha/2),$$

kde $T = S_1^2/S_2^2$. Veličinu $T = S_1^2/S_2^2$ lze tedy považovat za testovací statistiku pro hypotézu H_0 . Hladina významnosti tohoto testu je rovna α .

Ukazuje se přitom, že při běžně používaných hladinách významnosti a při nepřiliš velikém rozdílu mezi čísly m a n se nemůže stát, aby jedna z nerovností (****) byla splněna a druhá nikoliv. Proto v praxi postupujeme zpravidla tak, že za S_1^2 zvolíme větší a za S_2^2 menší z obou výběrových rozptylů a hypotézu H_0 zamítneme, jestliže poměr S_1^2/S_2^2 je příliš veliký, tj. jestliže $T > F_{m-1,n-1}(\alpha/2)$.

Definice 3. Výše popsany test je znám jako *F test*.

ÚLOHY

8. Bylo vybráno třináct polí stejné kvality. Na osmi z nich se zkoušel nový způsob hnojení, zbývajících pět bylo ošetřeno běžným způsobem. Výnosy pšenice v tunách na hektar jsou uvedeny v následující tabulce.

Nový způsob hnojení	5,7	5,5	4,3	5,9	5,2	5,6	5,8	5,1
Běžný způsob hnojení	5,0	4,5	4,2	5,4	4,4			

Je třeba zjistit, zda způsob hnojení má vliv na výnos pšenice.

Řešení. Označme $X_1, X_2, \dots, X_m$ výnosy pšenice v tunách na hektar při novém a $Y_1, Y_2, \dots, Y_n$ při běžném způsobu hnojení. Je tedy $m = 8$ a $n = 5$. Předpokládejme dále, že $X_1, X_2, \dots, X_m$ je náhodný výběr z rozdělení $N(\mu_1, \sigma_1^2)$, $Y_1, Y_2, \dots, Y_n$ je náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$ a testujme hypotézu $H_0 : \mu_1 = \mu_2$.

Nejprve předpokládejme, že $\sigma_1^2 = \sigma_2^2$ a provedme dvouvýběrový t test. Je $\bar{X} = 5,3875$, $\bar{Y} = 4,7$, $S_1^2 = 0,2698$, $S_2^2 = 0,24$. Dále pak

$$S^2 = \frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2} = \frac{7 \cdot S_1^2 + 4 \cdot S_2^2}{11} = 0,259.$$

Testovací statistika pro dvouvýběrový t test je

$$T = \frac{\bar{X} - \bar{Y}}{S \cdot \sqrt{\frac{1}{m} + \frac{1}{n}}}.$$

Jelikož $|T| \doteq 2,37 > 2,201 = t_{11}(0,05)$, hypotéza $H_0 : \mu_1 = \mu_2$ se proti alternativě $H_1 : \mu_1 \neq \mu_2$ zamítá na hladině významnosti $\alpha = 0,05$. V daném kontextu by ovšem možná bylo přirozenější zvolit alternativu $H_1 : \mu_1 > \mu_2$ a porovnávat hodnotu statistiky T s kritickou hodnotou $t_{11}(0,10)$.

Interval spolehlivosti pro rozdíl $\mu_1 - \mu_2$ je dán předpisem

$$\bar{X} - \bar{Y} \pm t_{m+n-2}(\alpha) \cdot S \cdot \sqrt{\frac{1}{m} + \frac{1}{n}}.$$

V našem případě shledáváme, že 95% interval spolehlivosti pro rozdíl středních výnosů při novém a běžném způsobu hnojení je $0,6875 \pm 0,6386$, tj. $(0,05; 1,33)$. Hypotéza $H_0 : \mu_1 = \mu_2$ byla proti alternativě $H_1 : \mu_1 \neq \mu_2$ zamítnuta na hladině významnosti $\alpha = 0,05$ z toho důvodu, že tento interval neobsahuje nulu.

Při posouzení, zda interval spolehlivosti pro rozdíl $\mu_1 - \mu_2$ obsahuje či neobsahuje nulu, lze postupovat též následujícím způsobem. Nejprve vypočítáme číslo

$$LSD = t_{m+n-2}(\alpha) \cdot S \cdot \sqrt{\frac{1}{m} + \frac{1}{n}},$$

tj. nejmenší (statisticky) významný rozdíl mezi průměrnými výnosy při novém a běžném způsobu hnojení a ověříme, zda rozdíl $|\bar{X} - \bar{Y}|$ je statisticky významný, tj. zda

$$|\bar{X} - \bar{Y}| \geq LSD.$$

V našem případě je $|\bar{X} - \bar{Y}| = 0,6875 \geq 0,6386 = LSD$, tudíž rozdíl $|\bar{X} - \bar{Y}|$ je při zadané hladině významnosti $\alpha = 0,05$ statisticky významný.

Otestujeme ještě hypotézu $H_0 : \sigma_1^2 = \sigma_2^2$, což je předpoklad výše provedeného dvouvýběrového t testu. Hodnota testovací statistiky je

$$T = S_1^2 / S_2^2 = 1,124.$$

(Připomeňme, že je třeba dělit větší rozptyl menším!) Vzhledem k tomu, že $T < F_{7,4}(0,025) = 9,074$, hypotéza $H_0 : \sigma_1^2 = \sigma_2^2$ se proti alternativě $H_1 : \sigma_1^2 \neq \sigma_2^2$ na hladině významnosti $\alpha = 0,05$ nezamítá. Tento tzv. F test lze realizovat též tak, že vyčíslíme 95 % interval spolehlivosti pro podíl σ_1^2 / σ_2^2 , tj. interval

$$\left(\frac{S_1^2 / S_2^2}{F_{7,4}(0,025)}, \frac{S_1^2 / S_2^2}{F_{7,4}(0,975)} \right).$$

Dosadíme-li

$$F_{7,4}(0,025) = 9,074 ; F_{7,4}(0,975) = F_{4,7}^{-1}(0,025) = 0,181,$$

obdržíme interval $(0,124; 6,209)$. Tento interval pokrývá hodnotu σ_1^2 / σ_2^2 podílu rozptylů σ_1^2 a σ_2^2 s pravděpodobností 0,95. Hypotéza $H_0 : \sigma_1^2 = \sigma_2^2$ se proti alternativě $H_1 : \sigma_1^2 \neq \sigma_2^2$ na hladině významnosti $\alpha = 0,05$ nezamítla proto, že tento interval obsahuje jedničku (a připouští tedy možnost, že poměr σ_1^2 / σ_2^2 je roven jedné).

Testujeme nyní hypotézu $H_0 : \mu_1 = \mu_2$ za předpokladu, že $\sigma_1^2 \neq \sigma_2^2$. Jako testovací statistika se pak používá veličina

$$T^* = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}}.$$

V našem případě je

$$v_1 = \frac{S_1^2}{m} = 0,033725, \quad v_2 = \frac{S_2^2}{n} = 0,048, \quad T^* \doteq 2,405.$$

Cochranův-Coxův test: Je

$$t^*(0,05) = \frac{v_1 t_{m-1}(0,05) + v_2 t_{n-1}(0,05)}{v_1 + v_2} \doteq 2,606.$$

Jelikož $T^* < t^*(0,05)$, hypotéza $H_0 : \mu_1 = \mu_2$ se proti alternativě $H_1 : \mu_1 \neq \mu_2$ na hladině významnosti $\alpha = 0,05$ nezamítá.

Satterthwaiteův test: Je

$$f = \frac{(v_1 + v_2)^2}{\frac{v_1^2}{m-1} + \frac{v_2^2}{n-1}} = 9,044$$

a $t_f(0,05) \doteq 2,260$. Jelikož $T^* > t_f(0,05)$, hypotéza $H_0 : \mu_1 = \mu_2$ se proti alternativě $H_1 : \mu_1 \neq \mu_2$ na hladině významnosti $\alpha = 0,05$ zamítá.

Welchův test: Je

$$h = \frac{(v_1 + v_2)^2}{\frac{v_1^2}{m+1} + \frac{v_2^2}{n+1}} - 2 = 11,086$$

a $t_h(0,05) \doteq 2,199$. Jelikož $T^* > t_h(0,05)$, hypotéza $H_0 : \mu_1 = \mu_2$ se proti alternativě $H_1 : \mu_1 \neq \mu_2$ na hladině významnosti $\alpha = 0,05$ zamítá.

9. Chceme prokázat, že studenti lesnické fakulty jsou v „průměru“ hmotnější než studentky. Za tím účelem vybereme náhodně deset studentů a deset studentek a zjistíme jejich hmotnost (v kilogramech). Výsledky tohoto průzkumu jsou uvedeny v následující tabulce.

Studenti	73	51	68	69	85	97	75	77	73	72
Studentky	49	41	70	76	54	52	62	46	47	63

Je rozdíl v průměrných hmotnostech vybraných studentů a studentek statisticky významný?

Řešení. Označme $X_1, X_2, \dots, X_{10}$ hmotnosti vybraných studentů a $Y_1, Y_2, \dots, Y_{10}$ hmotnosti vybraných studentek. Předpokládejme, že $X_1, X_2, \dots, X_{10}$ je náhodný výběr z rozdělení $N(\mu_1, \sigma_1^2)$ a $Y_1, Y_2, \dots, Y_{10}$ je náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$.

Je $\bar{X} = 74$, $\bar{Y} = 56$, $S_1 \doteq 11,81$, $S_2 \doteq 11,33$. Odtud plyne, že 95% interval spolehlivosti pro střední hmotnost studentů je (65,55; 82,45) a 95% interval spolehlivosti pro střední hmotnost studentek je (47,89; 64,11). Tyto intervaly mají prázdný průnik, pokud by tedy skutečně pokrývaly hodnoty středních hmotností μ_1 a μ_2 , znamenalo by to, že $\mu_1 \neq \mu_2$, tj. že střední hmotnosti studentů a studentek nejsou stejné. Bohužel však nelze přesně určit pravděpodobnost, že takovýmto postupem dospějeme k chybnému závěru (tj. zamítneme hypotézu $H_0: \mu_1 = \mu_2$). Víme pouze, že tato pravděpodobnost je větší než 0,05 a menší než 0,10. Chceme-li takovou pravděpodobnost určit přesně, je třeba provést dvouvýběrový t test.

Předpokládejme tedy navíc, že $\sigma_1^2 = \sigma_2^2$ a testujme hypotézu $H_0: \mu_1 = \mu_2$ proti alternativě $H_1: \mu_1 > \mu_2$ (možnost, že studentky jsou v průměru hmotnější než studenti a priori vylučujeme.). Je $S = \sqrt{(S_1^2 + S_2^2)/2} \doteq 11,57$ a $|T| \doteq 3,48 > 2,552 = t_{18}(0,02)$. Hypotéza $H_0: \mu_1 = \mu_2$ se tedy proti alternativě $H_1: \mu_1 > \mu_2$ zamítá na hladině významnosti $\alpha = 0,01$.

Předpokládejme nyní, že $\sigma_1^2 \neq \sigma_2^2$. Je $T^* \doteq 3,48$ a $t^*(0,02) = t_9(0,02) = 2,821$. Cochranův-Coxův test tedy rovněž vede k zamítnutí hypotézy $H_0: \mu_1 = \mu_2$ proti alternativě $H_1: \mu_1 > \mu_2$ na hladině významnosti $\alpha = 0,01$. Dále je $f \doteq 18$ a $t_{18}(0,02) = 2,552$, a konečně $h \doteq 20$ a $t_{20}(0,02) = 2,528$. Tudíž jak Satterthwaiteův, tak Welchův test vedou k zamítnutí hypotézy $H_0: \mu_1 = \mu_2$ proti alternativě $H_1: \mu_1 > \mu_2$ na hladině významnosti $\alpha = 0,01$.

10. Na stroji vyrábějícím dřevěné polotovary byla provedena jistá úprava za účelem zvýšení přesnosti obrábění. Chceme posoudit, zda tato úprava byla účinná. Bylo proto náhodně vybráno osm polotovarů vyrobených na stroji bez úpravy a osm na upraveném stroji. Délky výrobků v milimetrech jsou obsaženy v následující tabulce:

Stroj bez úpravy	114	114	118	122	124	130	116	116
Upravený stroj	116	118	118	118	120	120	120	124

- a) Testujte hypotézu, že střední délka vyráběných polotovarů zůstala po úpravě stroje stejná.
b) Testujte hypotézu, že přesnost obrábění se nezměnila, a to proti alternativě, že tato přesnost se po úpravě stroje zvýšila.

Řešení. Označme $X_1, X_2, \dots, X_8$ délky polotovarů vyrobených na stroji bez úpravy a $Y_1, Y_2, \dots, Y_8$ délky polotovarů vyrobených na upraveném stroji. Předpokládejme přitom, že na obou strojích se odchylka délky polotovaru od jeho střední délky se řídí normálním zákonem rozdělení pravděpodobností.

a) Jelikož $\bar{X} = \bar{Y} = 119,25$, nelze zamítnout hypotézu, že střední délka polotovarů je pro oba stroje stejná.

b) Je $S_1^2 = 31,93$, $S_2^2 = 5,64$ a $T = S_1^2/S_2^2 = 5,66 > 3,787 = F_{7,7}(0,05)$. Hypotéza, že přesnost ob-

rábění se nezměnila, se tedy zamítá na hladině významnosti $\alpha = 0,05$. Jelikož však $T < 6,993 = F_{7,7}(0,01)$, hypotéza se nezamítá na hladině významnosti $\alpha = 0,01$.