Pravděpodobnost vs. statistika

Teorie pravděpodobnosti

- pracuje s jednou nebo více teoretickými náhodnými veličinami, jejichž rozdělení je známo
- odvozovali jsme charakteristiky těchto rozdělení atd.

Statistika

- pracuje s pozorováními (daty) $\leftrightsquigarrow$ náhodný výběr z nějakého neznámého rozdělení
- na základě dat se snažíme něco říci o rozdělení, z něhož pocházejí (např. o střední hodnotě apod.)
- někdy pozorujeme více náhodných veličin (více náhodných výběrů) a chceme něco usoudit o jejich vzájemném vztahu

Příklad datového souboru

Studie zkoumající účinky nového léku pro snižování krevního tlaku:

id	lék	tlak pred	tlak po	pohl.	váha	...	kuřák
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
103	T	145	120	M	82	...	ano
104	C	155	130	M	97	...	ano
105	T	140	135	Z	74	...	ne
106	C	160	150	M	123	...	ano
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Dva typy problémů:

- odhady neznámých kvantit $\leftrightsquigarrow$ odhady parametrů
- rozhodování o platnosti nějakého výroku $\leftrightsquigarrow$ testování hypotéz

Teorie odhadu

- máme data $x_1, \dots, x_n$ (např. hodnoty výšky studentů)
- považujeme je za realizaci náhodného výběru $X_1, \dots, X_n$ z nějakého neznámého rozdělení
- chceme něco usuzovat o charakteristikách tohoto rozdělení (střední hodnota, rozptyl, hustota ...) $\rightsquigarrow$ budeme konstruovat jejich odhady (tzv. bodové odhady)
- odhadů je mnoho, chceme vybrat ty „dobré“

Jak by měl vypadat „dobrý odhad“?

- Neměl by mít žádnou systematickou výchylku (v průměru by měl odhadovat to, co chceme odhadovat).
- S přibývajícím počtem pozorování by měl být „přesnější a přesnější“.

Formální definice

Definice

Odhadem neznámé charakteristiky θ rozumíme jakoukoli funkci $\widehat{\theta}_n$ pozorování $X_1, \dots, X_n$.

1. Odhad $\widehat{\theta}_n$ nazýváme *nestranný (nevychýlený)*, pokud $E\widehat{\theta}_n = \theta$.
2. Odhad $\widehat{\theta}_n$ nazýváme *konzistentní*, pokud $\lim_{n\to\infty} \widehat{\theta}_n = \theta$.

Závěr: Rozumné odhady by měly být konzistentní a pokud možno nestranné (ale malá výchylka nevadí).

Už víme: průměr je dobrý odhad střední hodnoty (je nestranný i konzistentní).

Popisné statistiky jako bodové odhady

Teorie

- náhodná veličina X
- hustota f
- střední hodnota EX
- rozptyl var X
- medián, kvantily
- pravděpodobnost jevu

Odhady

- data $\leftrightsquigarrow$ realizace náh.výběru
- histogram
- výběrový průměr $\overline{X}_n$
- výběrový rozptyl S_X^2
- výběrový medián, výběrové kvantily
- relativní četnost případů

Odhad distribuční funkce

Problém: $X_1, \dots, X_n$ náhodný výběr, chceme odhadnout distribuční funkci $F(x) = P(X \le x)$

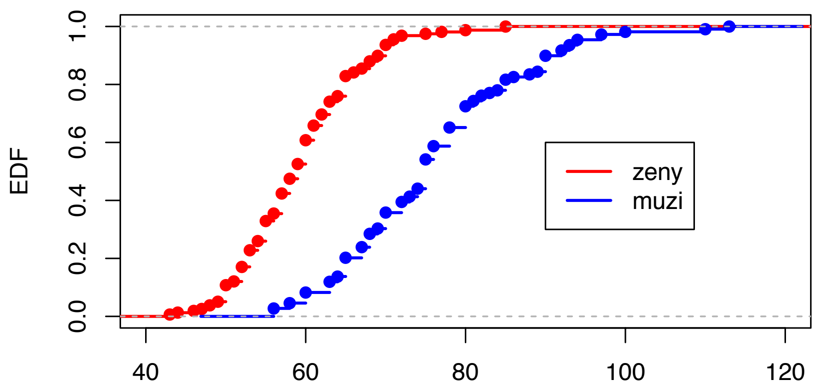
Empirická distribuční funkce definovaná jako

$$\widehat{F}_n(x) = \frac{\#[i : X_i \le x]}{n}$$

- lze ukázat, že má „dobré“ vlastnosti
- hodnota funkce $\widehat{F}_n$ v bodě x je odhadem pravděpodobnosti $P[X_i \le x]$ pomocí relativní četnosti jevu $[X_i \le x]$

Odhad distribuční funkce

Empirická distribuční funkce váhy studentů 1. ročníku PřF (muži a ženy zvlášť).



Odhad kovariance a korelace

Problém: náhodný výběr $\binom{X_1}{Y_1}, \dots, \binom{X_n}{Y_n}$ z dvourozměrného rozdělení, chceme odhadnout kovarianci a korelaci znaků X a Y

Výběrová kovariance:

$$S_{XY} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})$$

- $\overline{X}$ je výběrový průměr $X_1, \dots, X_n$
- $\overline{Y}$ je výběrový průměr $Y_1, \dots, Y_n$
- S_{XY} má stejnou struktura jako teoretická kovariance, jen střední hodnoty nahrazeny průměry
- S_{XY} je „dobrý“ odhad cov(X, Y)

Odhad korelace

Korelace: $\rho_{XY} = \dfrac{\text{cov}\,(X,Y)}{\sqrt{\text{var}\,X\ \text{var}\,Y}}$

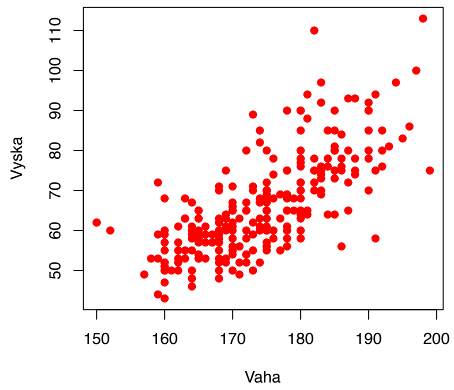
Výběrový korelační koeficient

$$r_{XY} = \frac{S_{XY}}{S_X\,S_Y} = \frac{\sum_{i=1}^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^n (X_i - \overline{X})^2 \sum_{i=1}^n (Y_i - \overline{Y})^2}}.$$

- S_X^2 je výběrový rozptyl $X_1, \ldots, X_n$
- S_Y^2 je výběrový rozptyl $Y_1, \ldots, Y_n$
- r_{XY} je „dobrý“ (konzistentní ale ne nestranný) odhad ρ_{XY}

Odhad kovariance a korelace: příklad

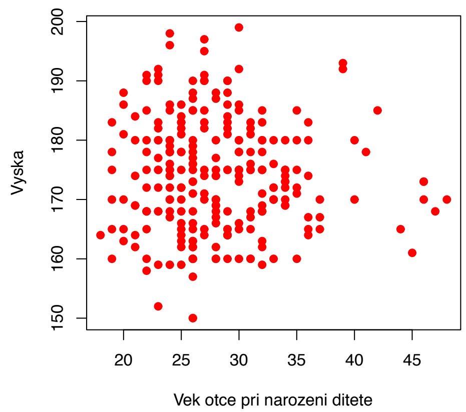
Graf váhy proti výšce ($r_{XY} = 0.72$):



hodnota r_{XY} koresponduje s obrázkem $\leftrightsquigarrow$ zdá se, že větší výška se pojí s vyšší hmotností

Odhad kovariance a korelace — příklad

Graf výšky proti věku otce při narození dítěte ($r_{XY} = -0.04$):



nic nenaznačuje, že by výška nějak souvisela s věkem otce při narození dítěte

Intervalové odhady

- doteď: tzv. bodové odhady (odhadem charakteristiky je číslo)
- nevyjadřuje nic o přesnosti odhadu

Intervalový odhad parametru θ

- interval s náhodnými mezemi, který překryje θ s předepsanou pravděpodobností $1 - \alpha$
- např. 95% interval spolehlivosti, interval na hladině 99% apod.
- konkrétněji si uvedeme později

Testování hypotéz – motivační příklad

Příklad: Provádíme průzkum, jaký skutečný objem piva točí v nejmenované hospodě. Zakoupeno bylo 10 piv a jejich objem byl (v litrech):

$$0.510,\ 0.462,\ 0.491,\ 0.466,\ 0.461,\ 0.503,\ 0.495,\ 0.488,\ 0.512,\ 0.505.$$

Z pohledu zákazníka bychom chtěli otestovat, zda hostinský netočí pod míru.

Průměr vychází $\overline{X} = 0.4893 < 0.5$. Otázkou je, zda je to vlivem náhody, nebo už se jedná o systematickou výchylku?

Testování hypotéz – motivační příklad

Je rozdíl $\overline{X} = 0.4893 < 0.5$ dostatečný na to, abychom mohli tvrdit, že hostinský točí pod míru?

- vezmeme-li přímo výběr se střední hodnotou 0.5 $\rightsquigarrow$ průměr nebude nikdy přesně 0.5, ale neměl by se lišit velmi
- je potřeba zohlednit:
  - ➔ počet pozorování (více dat $\leftrightsquigarrow$ větší přesnost odhadů)
  - ➔ variabilitu (vysoký rozptyl = větší nejistota)
- vezmeme v úvahu rozdělení $\overline{X} - 0.5$ a z něho zjistíme, jaké hodnoty jsou již „extrémní a málo pravděpodobné“ $\leftrightsquigarrow$ statistické testy

Testování hypotéz

Testování hypotéz

= vyhodnocování pravdivostní hodnoty výroků na základě náhodného výběru
(tj. ověřování platnosti nějakého výroku)

– provádíme pomocí statistických testů

Hypotéza

= výrok, o jehož pravdivosti chceme rozhodnout

- nulová hypotéza H_0
  - tvrzení o populaci, o jehož platnosti rozhodujeme
  - (není rozdíl, nezávisí, neliší se, ...)
- alternativní hypotéza H_1:
  - alternativa (doplňující možnost) k H_0
  - často tvrzení, které chceme prokázat

Statistický test

Statistický test = rozhodovací pravidlo, na jehož základě zamítáme nebo nezamítáme H_0

- *testová statistika* $T_n = T_n(X_1, \dots, X_n)$ = náhodná veličina, která je funkcí pozorování $X_1, \dots, X_n$
- *kritický obor* C = možné výsledky pokusu, kdy H_0 zamítáme

Rozhodovací pravidlo:

- pokud $T_n \in C$ ➔ zamítáme hypotézu H_0 ve prospěch alternativy H_1
  (*naše data svědčí proti H_0 a prokazujeme tak platnost H_1*)
- pokud $T_n \notin C$ ➔ hypotézu H_0 nezamítáme
  (*na základě našich dat nelze zamítnout H_0, naše data nejsou v rozporu s H_0*)

Pozor: nesymetrie mezi H_0 a H_1

Chyba I. a II. druhu

- rozhodujeme na základě náhodného výběru ➔ nemůžeme testovanou otázku zodpovědět s absolutní jistotou
- můžeme se dopustit chyby ➔ tyto chyby se budeme snažit omezit (resp. kontrolovat jejich pravděpodobnosti)

	H_0 zamítáme	H_0 nezamítáme
H_0 platí	chyba 1. druhu	OK
H_0 neplatí	OK	chyba 2. druhu

Označíme:

$$\alpha = \mathrm{P}(\text{chyba 1. druhu}) = \mathrm{P}(\text{zamítáme } H_0 |\ H_0 \text{ platí})$$

$$\beta = \mathrm{P}(\text{chyba 2. druhu}) = \mathrm{P}(\text{nezamítáme } H_0 |\ H_0 \text{ neplatí})$$

Přirozený požadavek: $\alpha, \beta \to \min \leftrightsquigarrow$ bohužel nelze současně

Chyba I. a II. druhu

- zvolíme hladinu testu α (zpravidla $\alpha = 0.05$)
  - maximální dovolená pst chyby 1. druhu
  - maximální pst falešného prokázání vědecké hypotézy
  - chyba 1.druhu je závažnější (falešně něco prokazujeme)
  - volíme před pokusem, nezávisle na jeho výsledku
- pro dané α chceme minimální $\beta \leftrightsquigarrow$ maximální $1 - \beta$
- síla testu $1 - \beta$
  - pst zamítnutí neplatné H_0
  - pst, s jakou prokážeme platnou vědeckou hypotézu H_1
  - nemáme pod kontrolou (závisí na tom, co opravdu platí)
  - můžeme ovlivnit volbou statistického testu, počtem pozorování, ...

➔ α máme plně pod kontrolou, o β toho moc nevíme

Dosažená hladina testu

Dosažená hladina testu p-hodnota (angl. p-value)

- pravděpodobnost, že dostaneme výsledek, který stejně nebo ještě méně podporuje H_0, jestliže H_0 platí
- nejmenší hladina α, na které lze ještě H_0 zamítnout
- „stupeň důvěry" v platnost H_0
- výsledek provedení statistického testu pomocí softwaru

Pravidlo:

- je-li $p \leq \alpha \rightsquigarrow$ zamítáme H_0
- je-li $p > \alpha \rightsquigarrow$ nezamítáme H_0

(Zapamatovat!)

Testování hypotéz – shrnutí

Při testování můžeme

- dojít ke správnému rozhodnutí
- udělat chybu 1. druhu (zamítnout platnou H_0) anebo 2. druhu (nezamítnout neplatnou H_0)

Postup:

- stanovíme dostatečně nízkou pravděpodobnost chyby 1. druhu α (hladinu testu)
  (chyba 1. druhu je závažnější $\rightsquigarrow$ její pst máme pod kotrolou)
- pravděpodobnost chyby 2. druhu β závisí na okolnostech (volba testu, počet pozorování, variabilita, ...) a její hodnotu v praxi neznáme

Proto: H_0 a H_1 nejsou posuzovány symetricky

Nesymetrie H_0 a H_1

H_0 a H_1 nejsou posuzovány symetricky:

- H_0 považujeme a priori za platnou a zamítáme ji jen tehdy, pokud k tomu máme dostatečně silné důvody
- pokud jsme zamítli H_0 $\rightsquigarrow$ můžeme tvrdit, že data svědčí o tom, že H_0 neplatí (a prokazujeme platnost H_1)
- pokud jsme H_0 nezamítli $\rightsquigarrow$ pak
  - buď H_0 opravdu platí
  - anebo H_0 neplatí, ale data neposkytují dostatečné „důkazy“ k jejímu zamítnutí (malá síla testu)

➔ nutné volit opatrné formulace závěrů (*hypotézu H_0 nelze na základě našich dat zamítnout apod.)*

Závěr

Hypotézu H_0 nemůžeme prokázat, ale pouze vyvrátit

Filozofie testování hypotéz

soudní proces ctící princip presumpce neviny.

H_0: „Obžalovaný je nevinen“.
H_1: „Obžalovaný je vinen“.

- zamítneme-li H_0 $\rightsquigarrow$ odsoudíme obžalovaného k trestu
- nezamítneme-li H_0 $\rightsquigarrow$ obžalovaný je propuštěn
- data $\leftrightsquigarrow$ důkazy svědčící pro nebo proti vině
- testová statistika $\leftrightsquigarrow$ soudce
- kritický obor $\leftrightsquigarrow$ důkazní břemeno = množství důkazů nutné k odsouzení
- hladina testu α $\leftrightsquigarrow$ pst odsouzení nevinného
- síla testu $1 - \beta$ $\leftrightsquigarrow$ pst odsouzení skutečného pachatele

Filozofie testování hypotéz

Hladinu α zvolíme malou (např. 5 %)

- chráníme nevinné před odsouzením
- čím menší α $\rightsquigarrow$ vyšší důkazní břemeno $\rightsquigarrow$ řada viníků propuštěna pro nedostatek důkazů

Výsledek procesu:

- odsouzení obžalovaného $\rightsquigarrow$ lze tvrdit, že je vinen
- propuštění obžalovaného $\rightsquigarrow$ buď opravdu nevinen, anebo vinen, ale soud neměl dostatečné důkazy k jeho odsouzení

K odsouzení viníka může dopomoci (bez porušení presumpce neviny):

- schopný soudce (vhodná testová statistika),
- dodatečné množství důkazů (více dat)
- málo chyb a omylů v důkazním materiálu (malý rozptyl)

Statistické testy

- statistických testů je obrovské množství
- uvedeme si jen několik vybraných základních
- podrobné odvození testové statistiky a kritického oboru jen u některých
- v praxi důležité:
  - výběr vhodného testu pro daný problém
  - provedení testu
  - ověření předpokladů testu (např. normalita dat, . . .)
  - správná interpretace výsledku

Test o střední hodnotě v normálním rozdělení

Situace: $X_1,\dots,X_n$ náhodný výběr z normálního rozdělení N(μ,σ^2), kde $\sigma>0$ známe; chceme testovat

$$H_0:\mu=\mu_0,$$

kde μ_0 je nějaká známá (předepsaná) hodnota, proti alternativě

$$H_1:\mu>\mu_0.$$

Chceme: statistický test $\leftrightsquigarrow$ testová statistka a kritický obor

Jiné možné alternativy:

- $H_1:\mu\neq\mu_0$ tzv. oboustranná alternativa
- $H_1:\mu<\mu_0$ tzv. jednostranná alternativa

(volba alternativy dle zadání úlohy – vyplývá z její formulace)

Odvození testu

Víme:

- střední hodnotu μ odhadujeme pomocí $\overline{X}_n$
- rozdělení $\overline{X}_n$ je N$(\mu,\frac{\sigma^2}{n})$, a proto

$$\sqrt{n}\frac{\overline{X}_n-\mu}{\sigma}\sim \mathrm{N}(0,1)$$

Proti H_0 a ve prospěch H_1 svědčí hodnoty $\overline{X}_n>>\mu_0$

- zvolíme hladinu testu α
- budeme hledat „mezní hodnotu“ K tak, že pro $\overline{X}_n>K$ už zamítneme H_0
- musí platit

$$\mathrm{P}(\text{ chyba 1.druhu })=\mathrm{P}(\overline{X}_n>K|H_0\ \text{ platí})=\alpha$$

Odvození testu – pokrač.

Postupně upravíme:

$$\alpha=\mathrm{P}(\overline{X}_n>K|H_0\text{ platí})=\mathrm{P}(\overline{X}_n>K|\mu=\mu_0)$$
$$=\mathrm{P}(\underbrace{\sqrt{n}\frac{\overline{X}-\mu_0}{\sigma}}_{T_n\sim\mathrm{N}(0,1)}>\underbrace{\sqrt{n}\frac{K-\mu_0}{\sigma}}_{=C}|\mu=\mu_0)=1-\Phi(C)$$

Odtud

$$C=u_{1-\alpha},\quad K=\frac{\sigma}{\sqrt{n}}u_{1-\alpha}+\mu_0$$

(Připomenutí: $u_{1-\alpha}$ je $1-\alpha$ kvantil normálního rozdělení, $u_{1-\alpha}=z_\alpha$, kde z_α je kritická hodnota)

Test o střední hodnotě normálního rozdělení při známém σ^2

Odvodili jsme test:

- testová statistika
$$T_n=\sqrt{n}\frac{\overline{X}-\mu_0}{\sigma}$$
- kritický obor $C=(u_{1-\alpha},\infty)$
- je-li $T_n>u_{1-\alpha}\rightsquigarrow$ zamítáme H_0
- je-li $T_n\leq u_{1-\alpha}\rightsquigarrow$ nezamítáme H_0

Podobným postupem bychom odvodili test H_0 proti

- $H_1:\mu<\mu_0\rightsquigarrow$ kritický obor $T_n<-u_{1-\alpha}$
- $H_1:\mu\neq\mu_0\rightsquigarrow$ kritický obor $|T_n|>u_{1-\frac{\alpha}{2}}$

Zvolíme-li hladinu $\alpha=0.05$ dostaneme $u_{1-\alpha}=u_{0.95}=1.64$ a $u_{1-\alpha/2}=u_{0.975}=1.96$.

Příklad – pivo

Máme zakoupeno 10 piv, chceme otestovat, zda hostinský netočí pod míru. Předpokládejme, že z předchozí zkušenosti víme, že směrodatná odchylka naměřených objemů piva je 0.02 l.

Model: Předpokládáme, že data jsou realizací náhodného výběru z N(μ, σ^2), kde $\sigma = 0.02$ (známe)

Hypotézy: $H_0 : \mu = 0.5$ proti $H_1 : \mu < 0.5$

Hladina testu: zvolíme test na hladině 5 %, tj. $\alpha = 0.05$

Test: musíme zjistit, zda platí $T_n = \sqrt{n}\frac{\overline{X}-0.05}{\sigma} < -u_{0.95} = -1.645$

Provedení testu

- máme $\overline{X} = 0.4893$, odtud testová statistika

$$T_n = \sqrt{10}\,\frac{0.4893 - 0.5}{0.02} = -1.691$$

- test: platí $T_n < -u_{0.95} = -1.645 \rightsquigarrow$ zamítáme nulovou hypotézu a prokazujeme alternativu
- podařilo se nám tedy prokázat, že hostinský točí pod míru
- riskovali jsme 5%, že budeme nesprávně tvrdit, že hostinský točí pod míru, pokud to tak není

Příklad – pokrač.

Dosažená hladina testu (p-hodnota)

- pravděpodobnost za platnosti H_0, že dostaneme výsledek ještě „extrémněji“ svědčící proti H_0 a ve prospěch H_1
- P($\sqrt{n}\frac{\overline{X}_n-0.5}{\sigma} < -1.691) = \Phi(1.691) = 0.045$
- p-hodnota je rovna 0.045 (je tedy nižší než $\alpha = 0.05$)

- podařilo se nám tedy prokázat, že hostinský točí pod míru
- ale p-hodnota je jen těsně menší než $\alpha = 0.05 \rightsquigarrow$ ne úplně přesvědčivý výsledek

Jiné alternativy

- kdybychom se zajímali o test $H_0 : \mu = 0.5$ proti $H_1 : \mu > 0.5$, pak bychom se zkoumali, zda platí

$$T_n > 1.64$$

(neplatí)

- kdybychom se zajímali o test $H_0 : \mu = 0.5$ proti $H_1 : \mu \neq 0.5$, pak bychom se zkoumali, zda platí

$$|T_n| > u_{0.975} = 1.96$$

(neplatí)

Jednovýběrový t-test – úvod

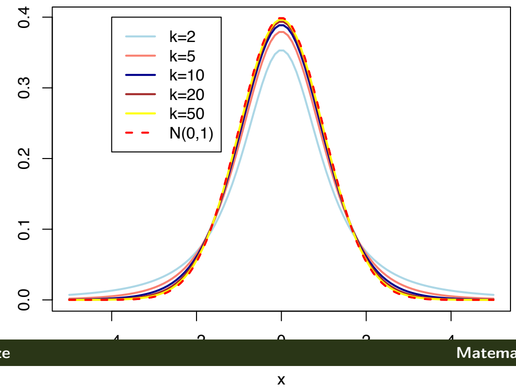
- dosud jsme předpokládali, že σ^2 je známé
- v praxi většinou σ^2 neznáme $\rightsquigarrow$ odhadujeme pomocí S_n^2
- v testové statistice místo σ použijeme S_n:

$$T_n = \sqrt{n}\,\frac{\overline{X}_n - \mu_0}{S_n}$$

- rozdělení T_n za H_0 není normální $N(0,1)$, ale nazývá se t-rozdělení s $n-1$ stupni volnosti, značíme jej t_{n-1} rozdělení

t-rozdělení

- hustota rozdělení t_k má pro velké k tvar velmi podobný normovanému normálnímu rozdělení
- pro malé k má rozptyl větší než 1 a kritické hodnoty jsou vzdálenější od 0 (cena, kterou platíme za neznalost σ)



Kvantily t-rozdělení

- α-kvantil budeme značit $t_k(\alpha)$

	α				
k	0.9	0.95	0.975	0.99	0.995
5	1.476	2.015	2.571	3.365	4.032
10	1.372	1.812	2.228	2.764	3.169
20	1.325	1.725	2.086	2.528	2.845
50	1.299	1.676	2.009	2.403	2.678
100	1.290	1.660	1.984	2.364	2.626
$N(0,1)$	1.282	1.645	1.960	2.326	2.576

Jednovýběrový t-test

Situace: $X_1, \dots, X_n$ náhodný výběr z rozdělení $N(\mu, \sigma^2)$, chceme otestovat hypotézu $H_0 : \mu = \mu_0$ proti (oboustranné) alternativě $H_1 : \mu \neq \mu_0$.

Testová statistika:

$$T_n = \sqrt{n}\frac{\overline{X}_n - \mu_0}{S_n}$$

Kritický obor

$$|T_n| > t_{n-1}(1 - \frac{\alpha}{2}) \implies \text{zamítáme } H_0,$$

kde $t_{n-1}(1 - \frac{\alpha}{2})$ je $(1 - \frac{\alpha}{2})$-kvantil rozdělení t_{n-1}

Tento test nazýváme *jednovýběrový t-test* na hladině α.

Další možné alternativy

Lze uvažovat také jednostranné alternativy:

- $H_1 : \mu > \mu_0 \rightsquigarrow$ kritický obor $T_n > t_{n-1}(1-\alpha)$
- $H_1 : \mu < \mu_0 \rightsquigarrow$ kritický obor $T_n < -t_{n-1}(1-\alpha)$

Poznámky

- t-rozdělení se někdy také nazývá *Studentovo* t-rozdělení
- William Sealy Gosset (1876–1937) $\leftrightsquigarrow$ chemik pracující v pivovaru Guinness
- 1908 odvození jednovýběrového t-testu (pseudonym Student)

Příklad — pivo

Máme zakoupeno 10 piv, chceme otestovat, zda hostinský netočí pod míru:

Model: Předpokládejme, že datům odpovídají nezávislé náhodné veličiny s normálním rozdělením N(μ, σ^2), σ^2 neznáme

Hypotézy:

$$H_0 : \mu = 0.5 \text{ proti } H_1 : \mu < 0.5$$

Hladina testu: opět budeme uvažovat test na hladině 5%

Test: zajímá nás, zda platí

$$T_n = \sqrt{n}\frac{\overline{X} - 0.5}{S} < -t_9(0.95) = -1.833$$

(pak budeme zamítat H_0)

Provedení testu

- spočteme
$$\overline{X} = 0.4893, \quad S = 0.0197,$$
- odtud
$$T_n = \sqrt{n}\frac{\overline{X} - 0.5}{S} = \sqrt{10}\frac{0.4893 - 0.5}{0.0197} = -1.7148$$
- máme $T_n > -t_9(0.95) = -1.833 \rightsquigarrow H_0$ nelze na hladině významnosti 5 % zamítnout
- nelze prokázat, že by hostinský točil pivo pod míru (buď skutečně pod míru netočí nebo tak málo, že tuto odchylku nemůžeme na základě našich dat prokázat)

Příklad – výpočet v programu R

```
>t.test(pivo,mu=0.5,alternative=‘‘less‘‘)
                    One Sample t-test
data: pivo

t = -1.7148, df = 9, p-value = 0.06026
alternative hypothesis: true mean is less than 0.5
95 percent confidence interval:
-Inf 0.5007382
sample estimates:
mean of x
0.4893
```

p-hodnota $> 0.05 \rightsquigarrow$ nezamítáme H_0 na hladině 5 %

Ověření předpokladu o normálním rozdělení

Jak ověříme, že data pocházejí skutečně z normálního rozdělení

- graficky
  - porovnáme histogram s hustotou normálního rozdělení
  - Q-Q graf (diagram normality) $\rightsquigarrow$ porovnává výběrové kvantily s teoretickými kvantily normálního rozdělení
- statistickým testem
  - např. Shapirův-Wilkův test
  - H_0 : výběr pochází z normálního rozdělení
    H_1 : výběr nepochází z normálního rozdělení
  - `shapiro.test`: p-hodnota $\rightsquigarrow$ zamítáme nebo nezamítáme normalitu $\rightsquigarrow$ lze nebo nelze použít jednovýběrový t-test

Úvod do statistiky Bodové odhady Testování hypotéz Test o μ v $N(\mu, \sigma^2)$ při známém σ^2 Jednovýběrový t-test

Porušení normality

Situace: $X_1, \ldots, X_n$ výběr z (jiného než normálního) rozdělení se střední hodnotou μ a konečným rozptylem, chceme testovat $H_0 : \mu = \mu_0$ proti alternativě $H_1 : \mu \neq \mu_0$

Jednovýběrový t-test

- předpoklad normality není splněn $\rightsquigarrow$ skutečná hladina testu není rovna předepsanému α (a můžeme se tedy dopouštět daleko větší chyby)
- použijeme-li $\rightsquigarrow$ hladinu testu se blíží k α pro $n \to \infty$ (zdůvodnění: centrální limitní věta)
- při velkém počtu pozorování n je hladina testu $\approx \alpha$

Závěr: Jednovýběrový t-test je možné použít k testování střední hodnoty nenormálních dat, pokud je pozorování dostatečně mnoho (řekněme alespoň 50).

Úvod do statistiky Bodové odhady Testování hypotéz Test o μ v $N(\mu, \sigma^2)$ při známém σ^2 Jednovýběrový t-test

Porušení normality

Co dělat v případě, že normální rozdělení nelze pro naše data uvažovat?

- použijeme t-test
  - výsledek interpretujeme „asymptoticky“ (ve smyslu, že test je proveden na hladině, která je asymptoticky rovna α)
  - nutné mít dostatečný počet pozorování
- použijeme jiný (např. neparametrický) test