Testování statistických hypotéz

Úvodní poznámky

- **Statistickou hypotézou** rozumíme hypotézu o populaci (základním souboru) – např.:
  - Střední hodnota základního souboru je rovna 100.
  - Střední hodnota prvního základního souboru se rovná střední hodnotě druhého základního souboru.
- Statistické hypotézy dělíme na **parametrické** a **neparametrické** hypotézy.

Úvodní poznámky

- Parametrická hypotéza je hypotéza o parametrech rozdělení základního souboru, zde patří:
  - Hypotézy o parametru jednoho základního souboru – o střední hodnotě, mediánu, rozptylu atd.
  - Hypotézy o parametrech dvou základních souborů (srovnávací testy) – rovnost středních hodnot, rovnost rozptylů atd.
  - Hypotézy o parametrech tří a více základních souborů.

Úvodní poznámky

- Neparametrická hypotéza je hypotéza o jiných vlastnostech základního souboru (tvar rozdělení, závislost proměnných atd.)
- Statistické testy dělíme na **parametrické** a **neparametrické** testy.
- Parametrickým testem rozumíme takový test, pro jehož odvození je nutno specifikovat typ rozdělení, případně jeho parametry. Nejčastěji se setkáváme s předpokladem **normality** dat.

Úvodní poznámky

- Neparametrickým testem rozumíme takový test, pro jehož odvození není nutno specifikovat typ rozdělení.
- Při testování hypotéz proti sobě stojí 2 hypotézy – **nulová** a **alternativní hypotéza**.
- Nulová hypotéza H_0 vyjadřuje tvrzení o základním souboru, které je bráno jako předpoklad při testování (rovnovážný stav).

Úvodní poznámky

- Alternativní hypotéza H_1 (resp. H_A) stojí proti nulové hypotéze a představuje porušení rovnovážného stavu. Rozlišujeme 3 typy alternativních hypotéz:
  - **Levostranná** alternativní hypotéza.
  - **Pravostranná** alternativní hypotéza.
  - **Oboustranná** alternativní hypotéza.

Úvodní poznámky

- Např.
  - H_0 – střední hodnota základního souboru $\mu = 100$.
  - Levostranná H_A – $\mu < 100$.
  - Pravostranná H_A – $\mu > 100$.
  - Oboustranná H_A – $\mu \neq 100$.
- Příslušná alternativní hypotéza se volí na základě pozorování chování výběrového souboru.

Úvodní poznámky

- Testování hypotéz je založeno na následujícím principu: Pokud výběrový soubor neukáže na statisticky významný rozpor s nulovou hypotézou, pak nesmíme nulovou hypotézu zamítnout.
- Jelikož na základě chování výběrového souboru (tedy pouze vzorku populace) usuzujeme o chování celé populace (základního souboru), můžeme se při rozhodování dopustit chyby.

Úvodní poznámky

		Výsledek testu	
		Platí H_0	Platí H_1
Skutečnost	Platí H_0	Správné rozhodnutí, pravděpodobnost **1 – α** (**spolehlivost testu**)	Chyba I. druhu, pravděpodobnost **α** (**hladina významnosti**)
	Platí H_1	Chyba II. druhu, pravděpodobnost **β**	Správné rozhodnutí, pravděpodobnost **1 – β** (**síla testu**)

Úvodní poznámky

- Snahou je samozřejmě minimalizovat obě chyby, což však není možné, neboť snížením β vzroste α a naopak. Při statistickém testování hypotéz se volí hodnota α (nejčastěji 0,05 nebo 0,01), neboť chyba I. druhu je významnější než chyba II. druhu.
- Chybu II. druhu lze snížit volbou vhodného testu nebo zvětšením rozsahu výběrového souboru.

Klasický test

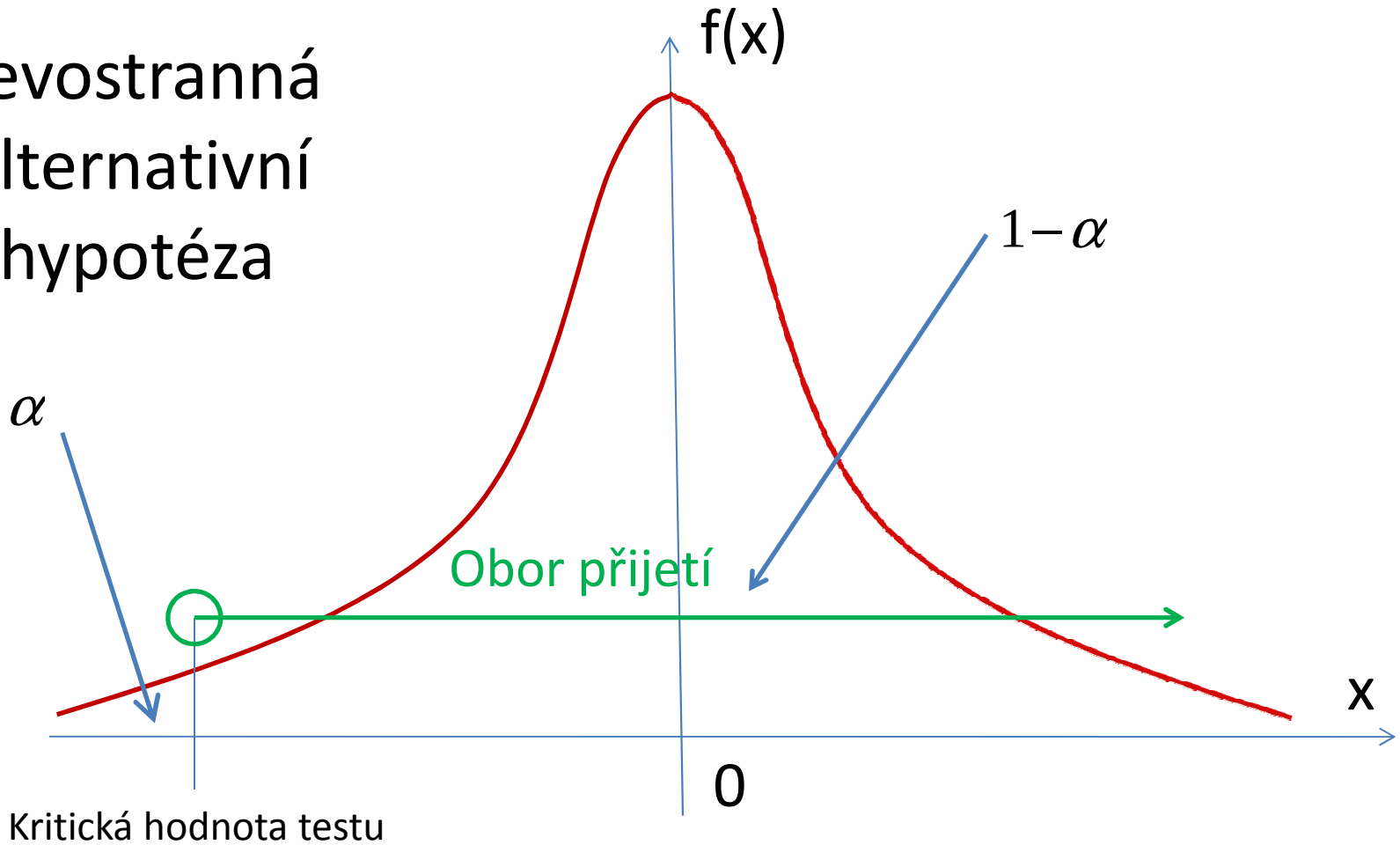
- Při klasickém testu hypotéz postupujeme v několika krocích:
  1) **Formulace nulové a alternativní hypotézy**.
  2) **Volba testové statistiky a jejího rozdělení** při platnosti nulové hypotézy (tzv. nulové rozdělení). Testová statistika a její nulové rozdělení je dána pro konkrétní test.

Klasický test

3) **Sestrojení kritického oboru a oboru přijetí** – obor všech možných hodnot testové statistiky rozdělíme na dva disjunktní obory – obor přijetí (takové hodnoty testové statistiky svědčící pro nezamítnutí nulové hypotézy) a kritický obor (takové hodnoty testové statistiky, které svědčí pro zamítnutí nulové hypotézy). Hranice mezi obory se nazývá **kritická hodnota testu** (x_{krit}). Kritický obor je tak velký, aby pravděpodobnost, že testová statistika leží v kritickém oboru při předpokladu platnosti nulové hypotézy, byla rovna α.

Klasický test

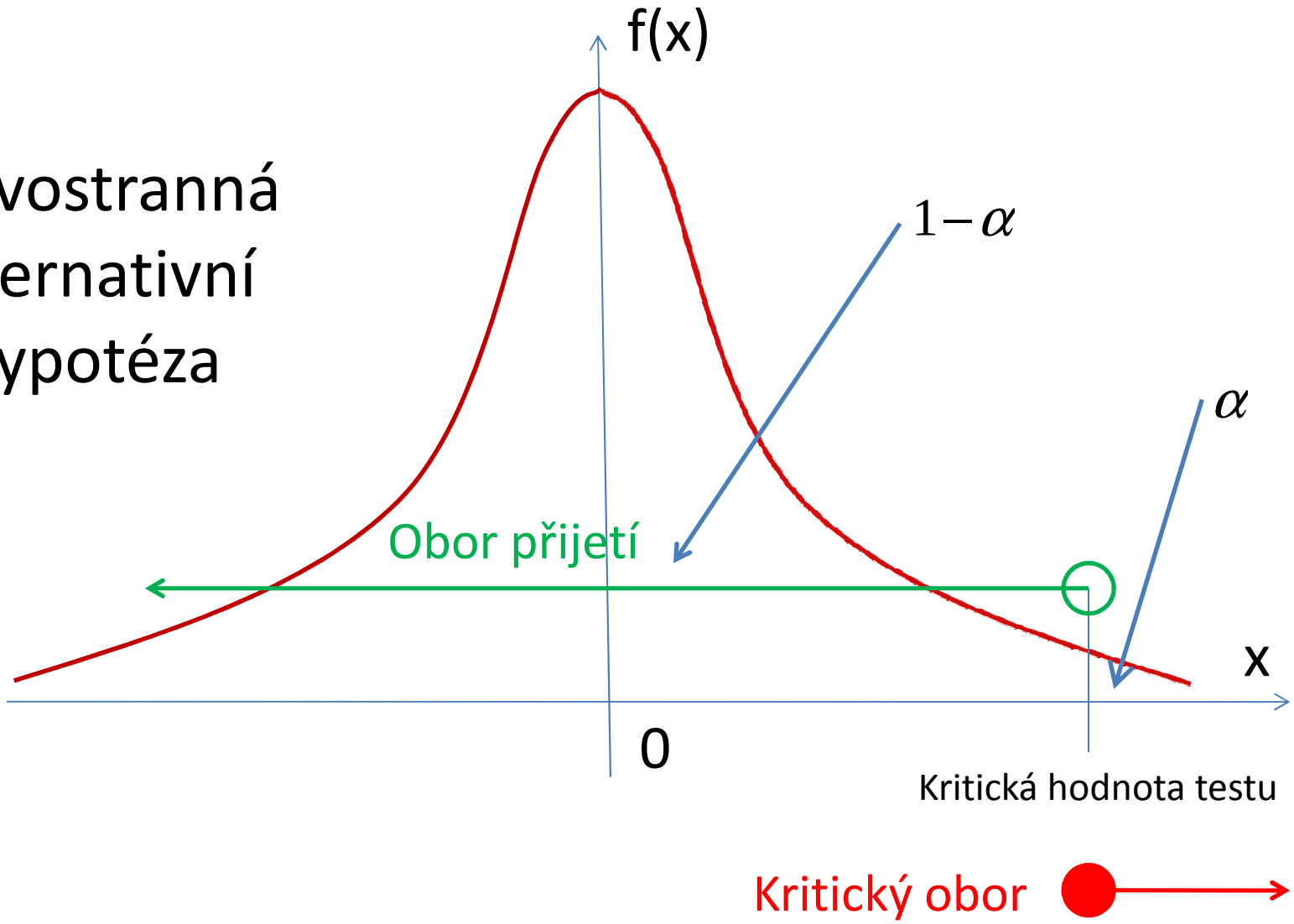
Levostranná alternativní hypotéza



Kritický obor

Klasický test

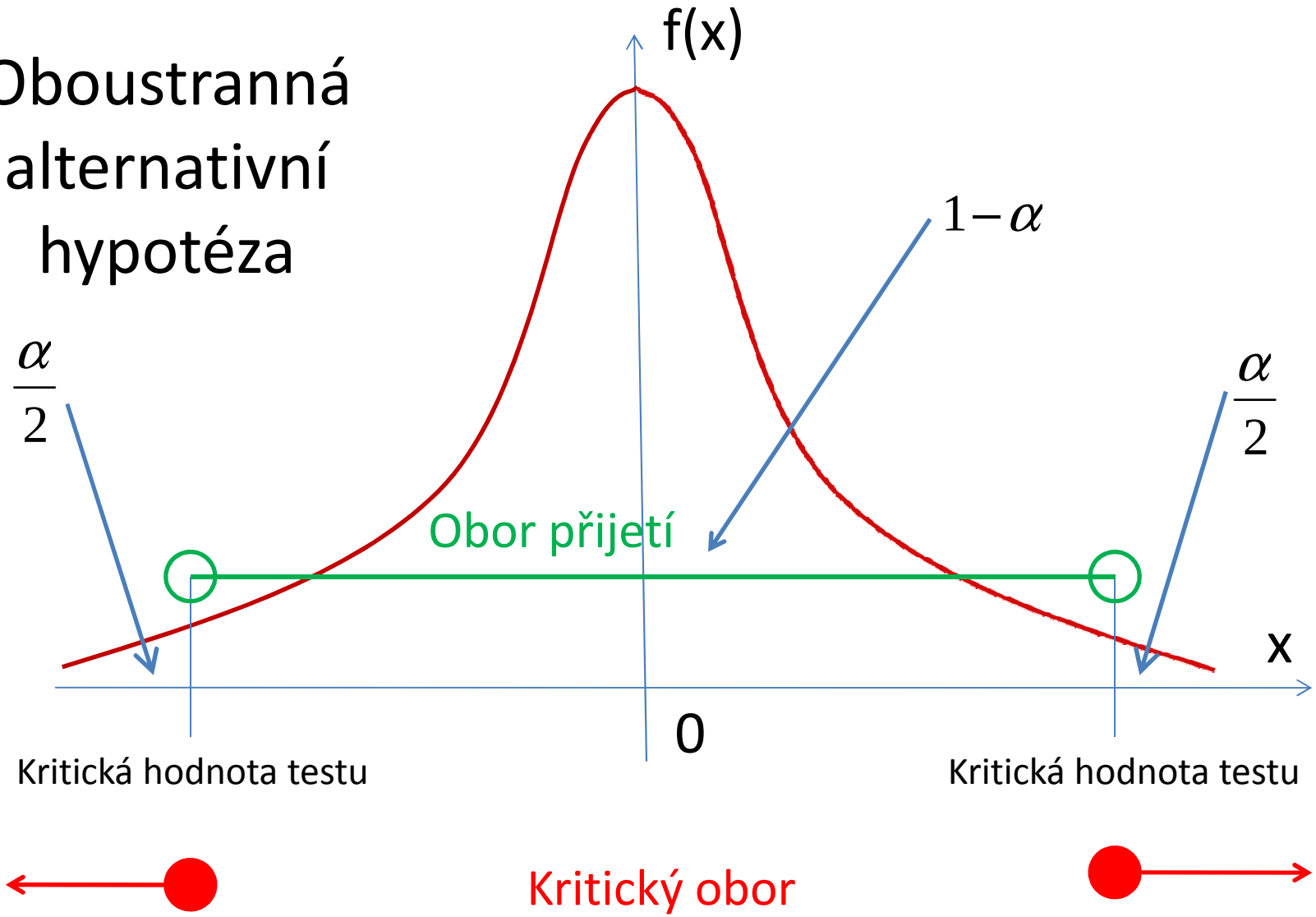
Pravostranná alternativní hypotéza



Kritický obor

Klasický test

Oboustranná alternativní hypotéza



Klasický test

4) **Výpočet pozorované hodnoty testové statistiky x_{obs}.**
5) **Formulace závěru testu**:
   - Leží-li x_{obs} v oboru přijetí, potom nezamítáme nulovou hypotézu.
   - Leží-li x_{obs} v kritickém oboru, potom zamítáme nulovou hypotézu ve prospěch alternativní hypotézy.

Testování normality dat

- Jelikož nejpoužívanější parametrické testy předpokládají normalitu dat, je nejprve vhodné se zabývat otázkou, jak otestovat, že data získaná náhodným výběrem pocházejí z populace řídící se normálním rozdělením s parametry μ a σ^2. K tomuto slouží tzv. **testy dobré shody**, zaměřme se na **Pearsonův χ^2 test dobré shody**.

Pearsonův χ^2 test dobré shody

- Tento test slouží k testování nulové hypotézy v obecném tvaru:
  - Náhodný výběr pochází z konkrétního rozdělení pravděpodobnosti s konkrétními parametry.
- Alternativní hypotéza neguje nulovou hypotézu:
  - Náhodný výběr nepochází z konkrétního rozdělení pravděpodobnosti s konkrétními parametry.
- Neznáme-li parametry příslušného rozdělení, je nutno je na základě náhodného výběru odhadnout, např. pomocí metody maximální věrohodnosti.

Pearsonův χ^2 test dobré shody

- Tento test umožňuje otestovat náhodný výběr i na jiná rozdělení než je normální rozdělení!
- Pro testování normality lze dále použít i jiné testy, např. Kolmogorovův – Smirnovův test dobré shody, Shapiro – Wilkův test, testy založené na šikmosti a špičatosti atd. Tyto testy jsou součástí specializovaných statistických software.

Pearsonův χ^2 test dobré shody

- Pro testovou statistiku G platí:

$$G = \sum_{i=1}^{k} \frac{\left(n_i - n \cdot \pi_{0,i}\right)^2}{n \cdot \pi_{0,i}} \rightarrow \chi^2_{k-h-1},$$

kde: k… počet tříd, n… rozsah souboru,

n_i… počet pozorování v i-té třídě (pozorované četnosti),

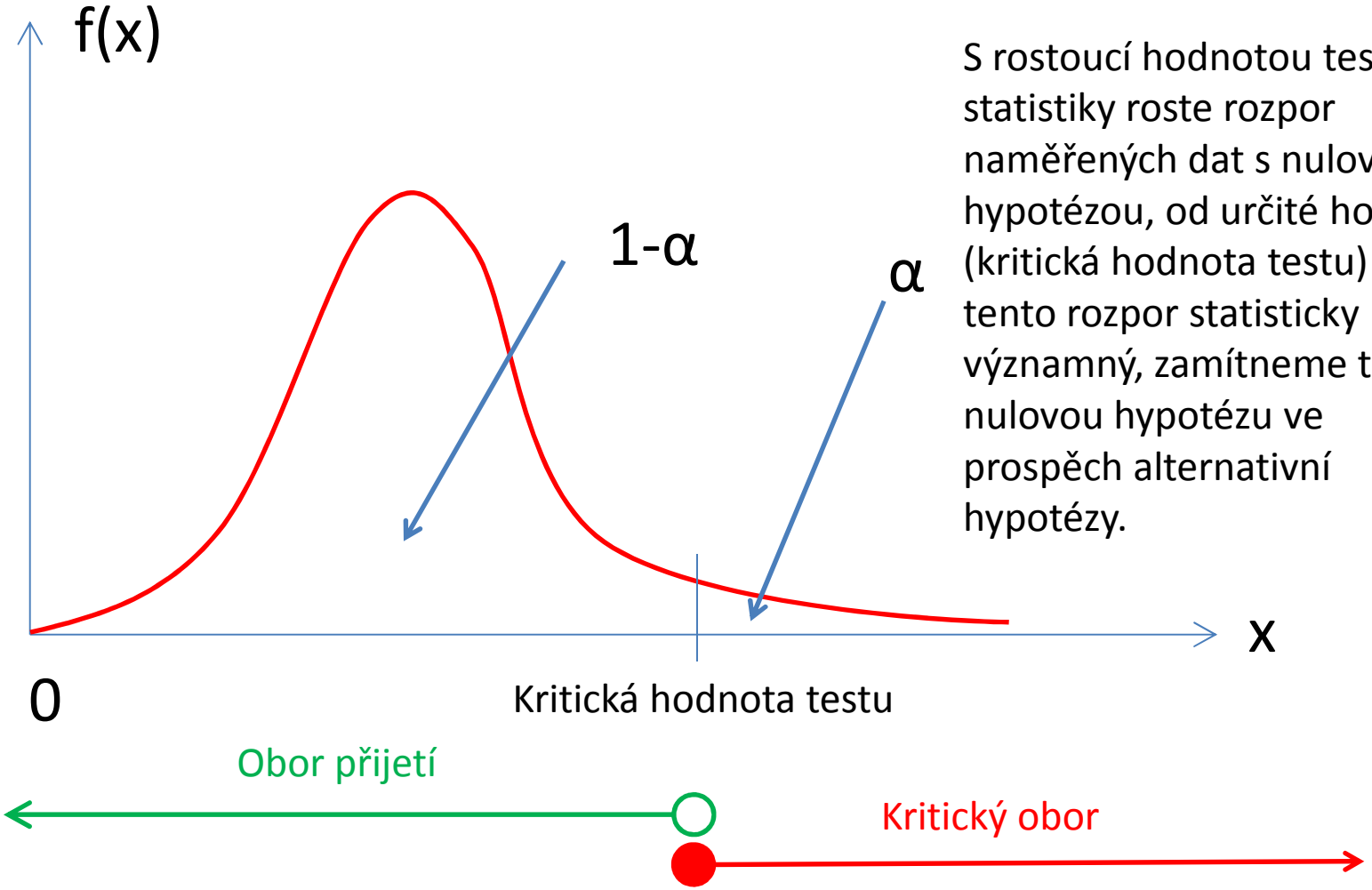
$n \cdot \pi_{0,i}$… teoretické (očekávané) četnosti,

h… počet odhadovaných parametrů rozdělení.

Pearsonův χ^2 test dobré shody

- Aby bylo nulové rozdělení dobře aproximováno rozdělením χ^2, je třeba, aby byly teoretické četnosti ve všech třídách větší než 5. Není-li tento předpoklad pro všechny třídy splněn, je nutno příslušné třídy vhodně sloučit (toto má za následek pokles stupňů volnosti rozdělení χ^2).

Pearsonův χ^2 test dobré shody



S rostoucí hodnotou testové statistiky roste rozpor naměřených dat s nulovou hypotézou, od určité hodnoty (kritická hodnota testu) je tento rozpor statisticky významný, zamítneme tedy nulovou hypotézu ve prospěch alternativní hypotézy.

Pearsonův χ^2 test dobré shody

- Kritickou hodnotu testu získáme jako 100·(1 – α)%-ní kvantil rozdělení χ^2 s příslušným počtem stupňů volnosti.
- Hodnotu můžeme odečíst z tabulek nebo ji můžeme vypočítat pomocí Excelu:

$$x_{krit} = \chi^2_{(1-\alpha);k-h-1} = CHIINV(\alpha; k-h-1).$$

Test hypotézy o střední hodnotě

- Test lze použít, pokud se základní soubor řídí normálním rozdělením (je tedy splněna normalita), jedná se tedy o parametrický test.
- Test slouží k ověření hypotézy, že střední hodnota populace se rovná konkrétní hodnotě, tedy:

$$\mu = \mu_0 .$$

Test hypotézy o střední hodnotě

- U alternativní hypotézy máme na výběr tři možnosti (pro konkrétní alternativní hypotézu se rozhodujeme na základě výběrového průměru):

$$1)\ \mu < \mu_0\,,$$

$$2)\ \mu > \mu_0\,,$$

$$3)\ \mu \neq \mu_0\,.$$

Test hypotézy o střední hodnotě

- Volba testové statistiky závisí na tom, zda známe či neznáme směrodatnou odchylku základního souboru σ.

  1) Známe-li σ, potom:

$$Z = \frac{\bar{x} - \mu_0}{\sigma} \cdot \sqrt{n} \rightarrow N(0,1),$$

testová statistika se tedy řídí normovaným normálním rozdělením.

Test hypotézy o střední hodnotě

2) Neznáme-li σ, potom:

$$T_{n-1} = \frac{\bar{x} - \mu_0}{s} \cdot \sqrt{n} \rightarrow t_{n-1},$$

kde *s* je výběrová směrodatná odchylka. Testová statistika se v tomto případě řídí Studentovým rozdělením s *n* – 1 stupni volnosti. V případě výběru velkého rozsahu (*n* ≥ 30) lze Studentovo rozdělení aproximovat normovaným rozdělením pravděpodobnosti.

Test hypotézy o střední hodnotě

- Kritickou hodnotu testu získáme:
  1) Jako příslušný kvantil normovaného rozdělení pravděpodobnosti buď z tabulek a nebo s využitím funkce Excel NORMSINV:
     - Levostranná alternativa

       $$x_{krit} = z_{\alpha} = NORMSINV(\alpha) = -z_{1-\alpha},$$

     - Pravostranná alternativa

       $$x_{krit} = z_{1-\alpha} = NORMSINV(1-\alpha),$$

     - Oboustranná alternativa

       $$x_{krit,L} = z_{\frac{\alpha}{2}} = NORMSINV\left(\frac{\alpha}{2}\right) = -z_{1-\frac{\alpha}{2}}, x_{krit,P} = z_{1-\frac{\alpha}{2}} = NORMSINV\left(1-\frac{\alpha}{2}\right).$$

Test hypotézy o střední hodnotě

2) Jako příslušný kvantil Studentova rozdělení s *n* – 1 stupni volnosti buď z tabulek a nebo s využitím funkce Excel TINV:

- Levostranná alternativa

$$x_{krit} = t_{\alpha;n-1} = -TINV(2\cdot\alpha; n-1) = -t_{1-\alpha;n-1},$$

- Pravostranná alternativa

$$x_{krit} = t_{1-\alpha;n-1} = TINV(2\cdot\alpha; n-1),$$

- Oboustranná alternativa

$$x_{krit,L} = t_{\frac{\alpha}{2};n-1} = -TINV(\alpha; n-1) = -t_{1-\frac{\alpha}{2};n-1}, x_{krit,P} = t_{1-\frac{\alpha}{2};n-1} = TINV(\alpha; n-1).$$

Test hypotézy o střední hodnotě

- **Př. 1:** Simulací byl získán statistický soubor o rozsahu 40 pozorování. Výpočtem bylo zjištěno, že výběrový průměr je roven 36,7 a výběrová směrodatná odchylka 10. Na hladině významnosti 0,05 a za předpokladu, že data pochází z normálního rozdělení, otestujte hypotézu, že střední hodnota je rovna 30.

Test hypotézy o střední hodnotě

- V tomto případě sice neznáme směrodatnou odchylku celé populace, ale máme výběr velkého rozsahu, proto můžeme Studentovo rozdělení aproximovat normálním rozdělením (výběrová směrodatná odchylka bude odhadem směrodatné odchylky populace).

Test hypotézy o střední hodnotě

- Stanovme si obě hypotézy:
  - H_0 – Střední hodnota populace je rovna 30.
  - H_1 – Střední hodnota populace je vyšší než 30 (vypovídá o tom provedený náhodný výběr).
- Dosazením do předpisu pro testovou statistiku získáme pozorovanou hodnotu testu:

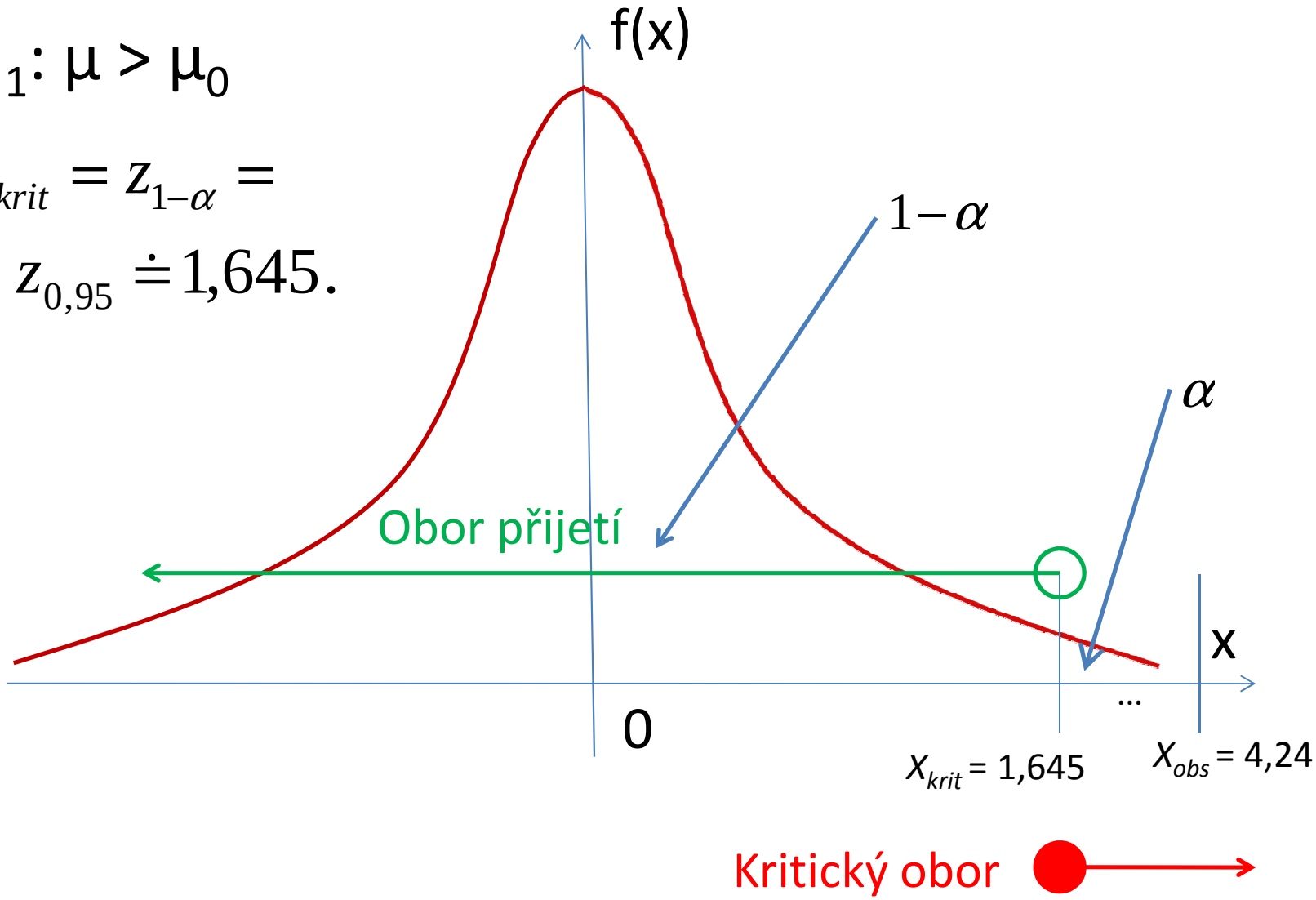
$$Z = \frac{\bar{x} - \mu_0}{s} \cdot \sqrt{n} = \frac{36{,}7 - 30}{10} \cdot \sqrt{40} \doteq 4{,}24.$$

Test hypotézy o střední hodnotě

$H_1: \mu > \mu_0$

$$x_{krit} = z_{1-\alpha} =$$

$$= z_{0,95} \doteq 1{,}645.$$



Test hypotézy o střední hodnotě

- Z obrázku vidíme, že pozorovaná hodnota testové statistiky leží v kritickém oboru, výsledek testu je tedy takový, že na hladině významnosti 0,05 zamítáme nulovou hypotézu ve prospěch alternativní, střední hodnota je tedy vyšší než 30.

Test hypotézy o shodě dvou středních hodnot

- Tento test slouží ke srovnání středních hodnot dvou populací, test opět předpokládá normalitu obou populací, jedná se tedy o parametrický test.
- Nulová hypotéza je ve tvaru:

$$\mu_1 = \mu_2 \left(\mu_1 - \mu_2 = 0\right).$$

Test hypotézy o shodě dvou středních hodnot

- U alternativní hypotézy máme na výběr ze 3 variant (pro vhodnou variantu se rozhodujeme na základě obou náhodných výběrů):

$$1)\ \mu_1 < \mu_2 \left(\mu_1 - \mu_2 < 0\right),$$
$$2)\ \mu_1 > \mu_2 \left(\mu_1 - \mu_2 > 0\right),$$
$$3)\ \mu_1 \neq \mu_2 \left(\mu_1 - \mu_2 \neq 0\right).$$

Test hypotézy o shodě dvou středních hodnot

- Volba testové statistiky opět závisí na tom, zda známe či neznáme směrodatnou odchylku obou základních souborů σ_1 a σ_2.
  1) Známe-li směrodatné odchylky, potom:

$$Z_2 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \rightarrow N(0,1),$$

testová statistika se tedy řídí normovaným normálním rozdělením.

Test hypotézy o shodě dvou středních hodnot

2) Neznáme-li směrodatné odchylky, ale víme, že platí $\sigma_1^2 = \sigma_2^2$ (příslušný test bude uveden dále), potom:

$$T_2 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_r \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightarrow t_{n_1+n_2-2},$$

$$\text{kde } s_r = \sqrt{\frac{(n_1 - 1)\cdot s_1^2 + (n_2 - 1)\cdot s_2^2}{n_1 + n_2 - 2}},$$

testová statistika se tedy řídí Studentovým rozdělením pravděpodobnosti s počtem volnosti rovným součtu rozsahů obou náhodných výběrů sníženým o 2.

Test hypotézy o shodě dvou středních hodnot

3) Neznáme-li směrodatné odchylky a víme, že nelze předpokládat rovnost rozptylů $\sigma_1^2 \neq \sigma_2^2$, potom lze použít testovou statistiku

$$T_2 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \to t_v ,$$

$$\text{kde } v \cong \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1} \cdot \left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1} \cdot \left(\frac{s_2^2}{n_2}\right)^2} .$$

Test hypotézy o shodě dvou středních hodnot

- Počet stupňů volnosti je nutno zaokrouhlit na celé číslo. Testová statistika se tedy řídí Studentovým rozdělením pravděpodobnosti s počtem volnosti rovným zaokrouhlenému číslu ν.
- Kritické hodnoty testu stanovíme analogickým způsobem jako u jednovýběrového testu střední hodnoty.

Test o shodě dvou rozptylů (F-test)

- Mějme dva náhodné výběry pocházející ze dvou populací s normálním rozdělením s neznámými parametry.
- Naším úkolem je otestovat rovnost rozptylů těchto dvou populací, nulovou hypotézu tedy formulujeme ve tvaru:

$$\sigma_1^2 = \sigma_2^2 \left(\frac{\sigma_1^2}{\sigma_2^2} = 1 \right).$$

Test o shodě dvou rozptylů (F-test)

- Jelikož testové Fischer-Snedecorovo rozdělení není symetrické, přichází v úvahu pouze jednostranné alternativní hypotézy:

$$1)\ \sigma_1^2 < \sigma_2^2 \left(\frac{\sigma_1^2}{\sigma_2^2} < 1 \right),$$

$$2)\ \sigma_1^2 > \sigma_2^2 \left(\frac{\sigma_1^2}{\sigma_2^2} > 1 \right).$$

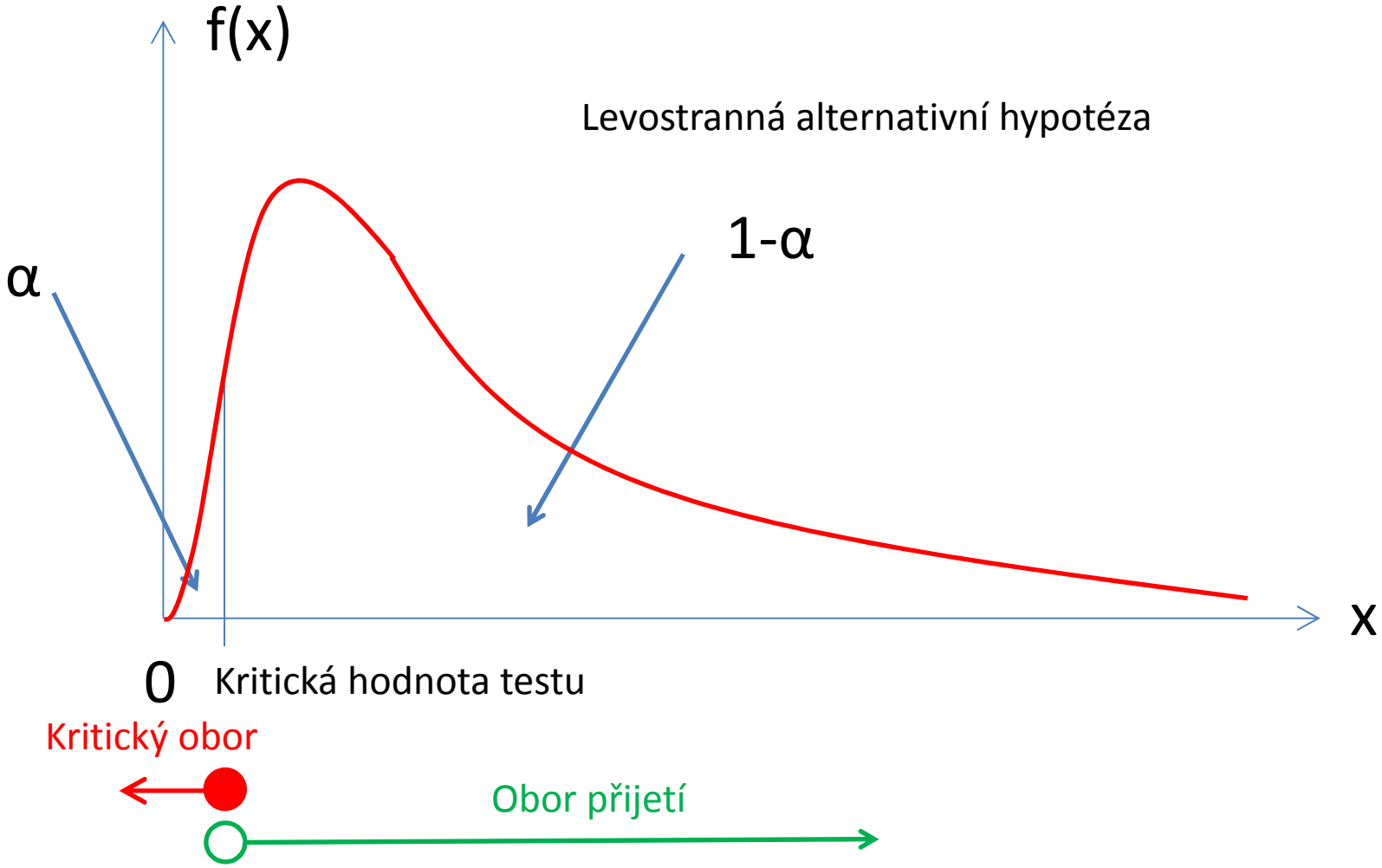
Test o shodě dvou rozptylů (F-test)

- Testová statistika je ve tvaru:

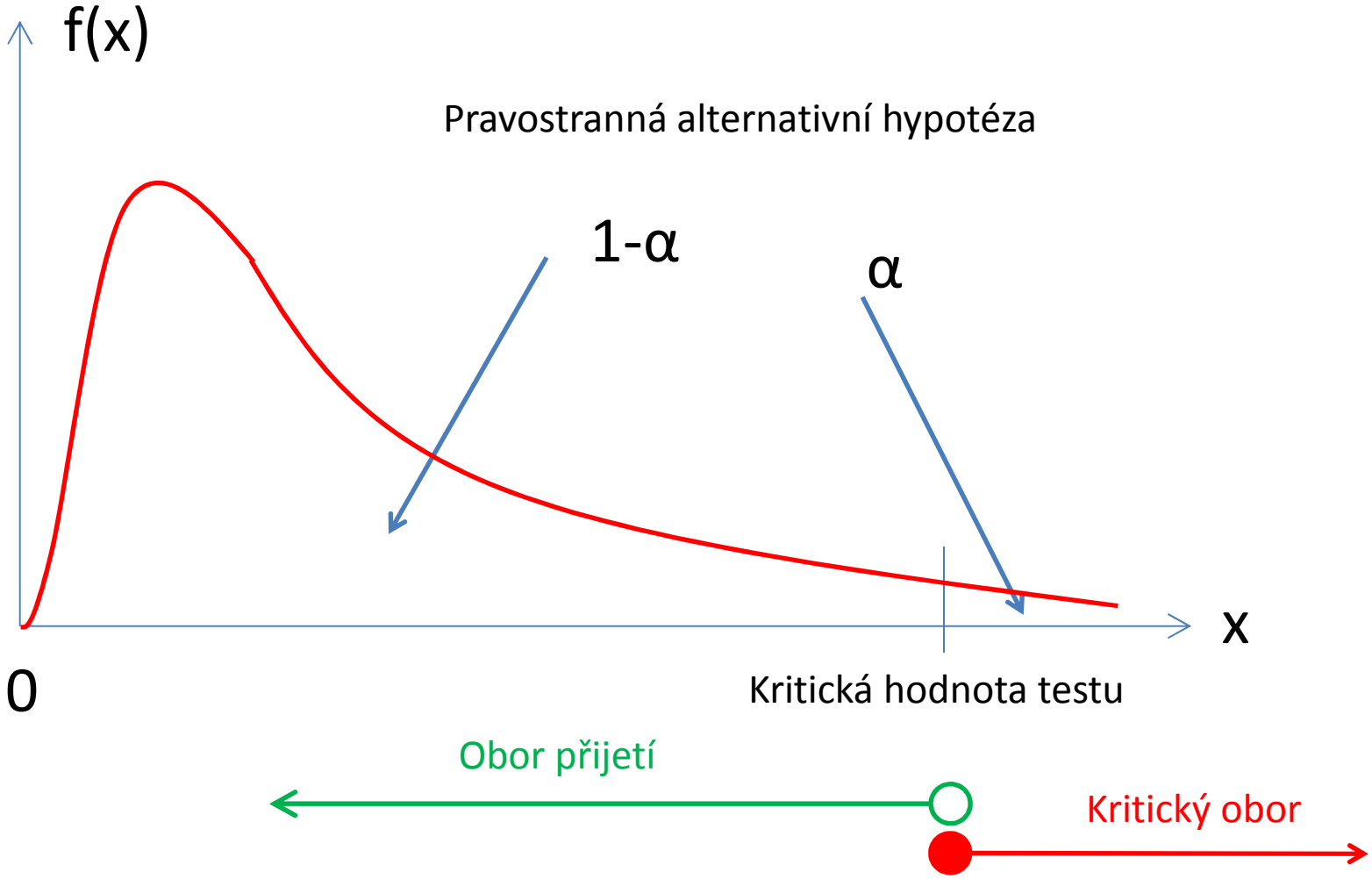
$$F_2 = \frac{s_1^2}{s_2^2} \rightarrow F_{n_1-1,n_2-1},$$

testová statistika se tedy řídí Fisher-Snedecorovým rozdělením s n_1-1 stupni volnosti v čitateli a n_2-1 stupni volnosti ve jmenovateli.

Test o shodě dvou rozptylů (F-test)



Test o shodě dvou rozptylů (F-test)



Test o shodě dvou rozptylů (F-test)

- Kritickou hodnotu testu získáme jako příslušný kvantil Fisher-Snedecorova rozdělení s příslušnými stupni volnosti v čitateli a jmenovateli buď pomocí tabulek nebo pomocí funkce Excel FINV:
  - Levostranná alternativa

$$x_{krit} = F_{\alpha;n_1-1;n_2-1} = FINV(1-\alpha; n_1-1; n_2-1),$$

  - Pravostranná alternativa

$$x_{krit} = F_{1-\alpha;n_1-1;n_2-1} = FINV(\alpha; n_1-1; n_2-1).$$

Test hypotézy o shodě dvou středních hodnot

- **Př. 2:** Simulací jsme získali dva statistické soubory.

	Soubor 1	Soubor 2
Rozsah	35	45
Výběrový průměr	38,7	42,4
Výběrová směrodatná odchylka	7,9	10,1

Za předpokladu, že oba výběry pocházejí z normálního rozdělení, otestujte na hladině významnosti 0,05 hypotézu o rovnosti středních hodnot.

Test hypotézy o shodě dvou středních hodnot

- V tomto případě je zřejmé, že neznáme směrodatné odchylky populací (máme pouze jejich odhady). Proto musíme nejdříve přistoupit k otestování, zda lze očekávat shodné rozptyly obou populací či nikoliv. K tomu použijeme F-test, budeme uvažovat hladinu významnosti rovnu 0,05.

Test hypotézy o shodě dvou středních hodnot

- Stanovme si obě hypotézy k F-testu:
  - H_0 – Rozptyl populace 1 je roven rozptylu populace 2, tedy $\sigma_1^2 = \sigma_2^2$.
  - H_1 – Rozptyl populace 1 je menší než rozptyl populace 2 (vypovídají o tom oba náhodné výběry), tedy $\sigma_1^2 < \sigma_2^2$.
- Dosazením do vztahu pro testovou statistiku dostaneme:

$$F_2 = \frac{s_1^2}{s_2^2} = \frac{7{,}9^2}{10{,}1^2} \doteq 0{,}61.$$

Test hypotézy o shodě dvou středních hodnot

- Nyní musíme stanovit kritickou hodnotu testu, pro kterou v případě zvolené levostranné alternativy platí:

$$x_{krit} = F_{\alpha;n_1-1;n_2-1} = FINV(1-\alpha;n_1-1;n_2-1) =$$
$$= FINV(0{,}95;34;44) \doteq 0{,}58.$$

- Jelikož z výsledků testu vidíme, že pozorovaná hodnota testové statistiky leží v oboru přijetí, nezamítáme nulovou hypotézu o rovnosti rozptylů obou populací.

Test hypotézy o shodě dvou středních hodnot

- Nyní můžeme přistoupit k testování rovnosti středních hodnot dvou populací, použijeme variantu 2.
- Stanovme si obě hypotézy:
  - H_0 – Střední hodnota populace 1 je rovna střední hodnotě populace 2, tedy $\mu_1 = \mu_2$.
  - H_1 – Střední hodnota populace 1 je menší než střední hodnota populace 2 (vypovídají o tom oba náhodné výběry), tedy $\mu_1 < \mu_2$.

Test hypotézy o shodě dvou středních hodnot

- Dosazením do předpisu pro testovou statistiku získáme pozorovanou hodnotu testu:

$$T_2 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1 - 1)\cdot s_1^2 + (n_2 - 1)\cdot s_2^2}{n_1 + n_2 - 2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} =$$

$$= \frac{(38{,}7 - 42{,}4) - 0}{\sqrt{\frac{(35-1)\cdot 7{,}9^2 + (45-1)\cdot 10{,}1^2}{35+45-2}} \cdot \sqrt{\frac{1}{35} + \frac{1}{45}}} \doteq -1{,}78.$$

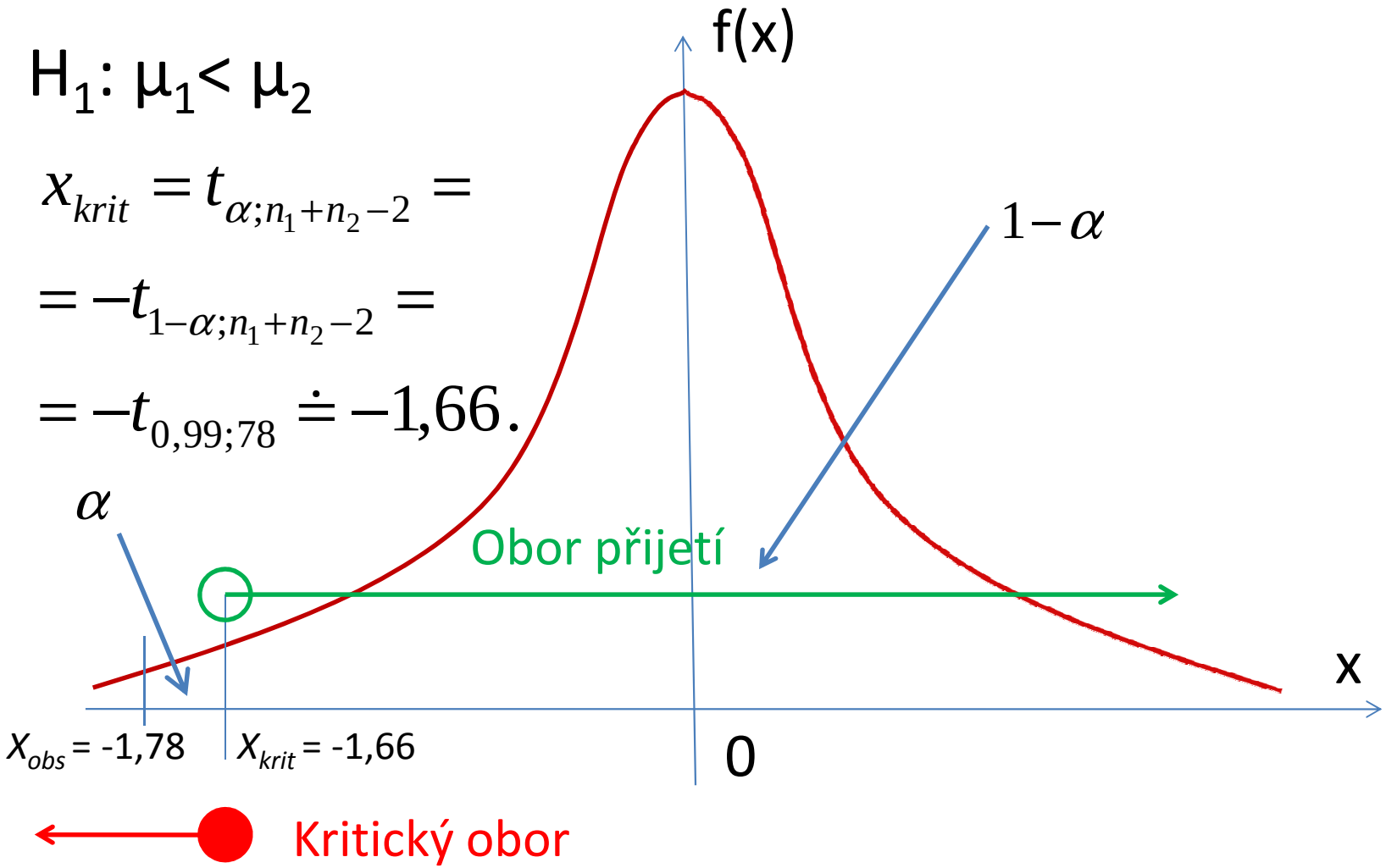
Test hypotézy o shodě dvou středních hodnot

$H_1: \mu_1 < \mu_2$

$$x_{krit} = t_{\alpha;n_1+n_2-2} =$$

$$= -t_{1-\alpha;n_1+n_2-2} =$$

$$= -t_{0,99;78} \doteq -1{,}66.$$



Test hypotézy o shodě dvou středních hodnot

- Z obrázku vidíme, že pozorovaná hodnota testové statistiky leží v kritickém oboru, výsledek testu je tedy takový, že na hladině významnosti 0,05 zamítáme nulovou hypotézu, střední hodnoty obou populací se nerovnají, střední hodnota první populace je nižší než střední hodnota druhé populace.