

# Pravděpodobnostní metody ve strojírenství

## D.3 - Testy o kvalitativních výsledcích



# D.3 - Testy o kvalitativních výsledcích

- Klíčové pojmy:**
- poměry (proporce)
  - Kontingenční tabulka, čtyřpolní tabulka
- Klíčové vztahy:**
- Test chí-kvadrát
  - Fisherův test

# Test poměrů

## Příklady:

- Statistický odhad pravděpodobnosti
- Odhad volebních preferencí
- Poměr neshodných výrobků
- Poměr žen v určité funkci
- ... atd.

odhadujeme pomocí relativní četnosti  $\hat{p} = \frac{n_A}{N}$

**Jednovýběrový test:**  $H_0 : p = p_0$   
 $H_A : p \neq p_0$

$$T = \frac{\hat{p} - Np_0}{\sqrt{Np_0(1 - p_0)}} \sim N(0, 1)$$

> `prop.test(36, 100, 0.3)`

1-sample proportions test with continuity correction

```
data: 36 out of 100, null probability 0.3
X-squared = 1.4405, df = 1, p-value = 0.2301
alternative hypothesis: true p is not equal to 0.3
95 percent confidence interval:
 0.2681721 0.4627255
sample estimates:      p      0.36
```



# Test poměrů

**Dvouvýběrový test:**

$$H_0 : p_1 = p_2 = p$$

$$H_A : p_1 \neq p_2$$

$$T = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{p_0(1 - p_0)(1/N_1 + 1/N_2)}} \sim N(0, 1)$$

## Příklad:

Bylo dotazováno 800 absolventů ČVUT a UK. Z výzkumu plyne, že 322 dotazovaných jsou absolventi ČVUT kteří pracují dále ve svém oboru, 234 dotázaných absolvovalo UK a pracují také ve svém oboru. Ve 158 případech odpovědí bylo zjištěno, že absolventi ČVUT ve svém oboru nepracují a zbývajících 86 jsou absolventi UK, kteří také nepracují ve svém oboru.

**Liší se statisticky významně poměr absolventů pracujících v oboru pro tyto dvě univerzity?**

$$\hat{p}_1 = 322/480 = 0,67, \quad \hat{p}_2 = 234/320 = 0,73$$

	ČVUT	UK	celkem
v oboru	322	234	556
mimo obor	158	86	244
celkem	480	320	800



# Test poměrů

**Dvouvýběrový test:**  $H_0 : p_1 = p_2 = p$   
 $H_A : p_1 \neq p_2$

$$T = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{p_0(1 - p_0)(1/N_1 + 1/N_2)}} \sim N(0, 1)$$

## Příklad:

```
> prop.test(c(322,234),c(480,320))
```

2-sample test for equality of proportions with continuity correction

```
data: c(322, 234) out of c(480, 320)
X-squared = 3.0273, df = 1, p-value = 0.08187
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.127257708  0.006424375
sample estimates:
 prop 1      prop 2
0.6708333 0.7312500
```

	ČVUT	UK	celkem
v oboru	322	234	556
mimo obor	158	86	244
celkem	480	320	800



# Test nezávislosti v kontingenční tabulce

## Příklad:

Bylo dotazováno 800 absolventů ČVUT a UK. Z výzkumu plyne, že 322 dotazovaných jsou absolventi ČVUT kteří pracují dále ve svém oboru, 234 dotázaných absolvovalo UK a pracují také ve svém oboru. Ve 158 případech odpovědí bylo zjištěno, že absolventi ČVUT ve svém oboru nepracují a zbývajících 86 jsou absolventi UK, kteří také nepracují ve svém oboru.

Lze považovat uplatnění se v oboru studia stochasticky závislé na absolvované univerzitě?

	ČVUT	UK	celkem
v oboru	0,4025	0,2925	0,695
mimo obor	0,1975	0,1075	0,305
celkem	0,6000	0,4000	1,000

$$0,695 \times 0,6 = 0,417 \neq 0,4025$$



# Test nezávislosti v kontingenční tabulce

## Příklad:

Bylo dotazováno 800 absolventů ČVUT a UK. Z výzkumu plyne, že 322 dotazovaných jsou absolventi ČVUT kteří pracují dále ve svém oboru, 234 dotázaných absolvovalo UK a pracují také ve svém oboru. Ve 158 případech odpovědí bylo zjištěno, že absolventi ČVUT ve svém oboru nepracují a zbývajících 86 jsou absolventi UK, kteří také nepracují ve svém oboru.

Lze považovat uplatnění se v oboru studia stochasticky závislé na absolvované univerzitě?

## Test chí-kvadrát:

$$H_0 : n_{ij} = \frac{n_{i.} \cdot n_{.j}}{n} \quad \text{pro všechny dvojice } i, j$$

$$H_A : n_{ij} \neq \frac{n_{i.} \cdot n_{.j}}{n} \quad \text{pro nějakou dvojici } i, j$$

$$\text{kde } n_{i.} = \sum_{j=1}^n n_{ij}, \quad n_{.j} = \sum_{i=1}^n n_{ij}.$$

	ČVUT	UK	celkem
v oboru	0,4025	0,2925	0,695
mimo obor	0,1975	0,1075	0,305
celkem	0,6000	0,4000	1,000

$$0,695 \times 0,6 = 0,417 \neq 0,4025$$



# Test nezávislosti v kontingenční tabulce

## Příklad:

Bylo dotazováno 800 absolventů ČVUT a UK. Z výzkumu plyne, že 322 dotazovaných jsou absolventi ČVUT kteří pracují dále ve svém oboru, 234 dotázaných absolvovalo UK a pracují také ve svém oboru. Ve 158 případech odpovědí bylo zjištěno, že absolventi ČVUT ve svém oboru nepracují a zbývajících 86 jsou absolventi UK, kteří také nepracují ve svém oboru.

Lze považovat uplatnění se v oboru studia stochasticky závislé na absolvované univerzitě?

## Test chí-kvadrát:

$$H_0 : n_{ij} = \frac{n_{i.} \cdot n_{.j}}{n} \quad \text{pro všechny dvojice } i, j$$

$$H_A : n_{ij} \neq \frac{n_{i.} \cdot n_{.j}}{n} \quad \text{pro nějakou dvojici } i, j$$

$$\text{kde } n_{i.} = \sum_{j=1}^n n_{ij}, \quad n_{.j} = \sum_{i=1}^n n_{ij}.$$

	ČVUT	UK	celkem
v oboru	$n_{11}$	$n_{12}$	$n_{1.}$
mimo obor	$n_{21}$	$n_{22}$	$n_{2.}$
celkem	$n_{.1}$	$n_{.2}$	$n$



# Test nezávislosti v kontingenční tabulce

## Test chí-kvadrát:

Označme  $n_{ij}^* = \frac{n_{i.} \cdot n_{.j}}{n}$  tzv. teoretické (očekávané) četnosti. Potom testová statistika bude mít tvar:

$$X^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - n_{ij}^*)^2}{n_{ij}^*} \sim \chi^2(r-1)(s-1)$$

Hypotézu  $H_0$  o nezávislosti znaků zamítáme na hladině významnosti  $\alpha$ , pokud hodnota testové statistiky  $X^2$  překročí  $(1-\alpha)$ -kvantil rozdělení chí-kvadrát o  $(r-1)(s-1)$  stupních volnosti.

## Podmínky pro použití testu nezávislosti v kontingenční tabulce:

- žádná teoretická četnost nesmí být menší než 1
- nejvíce 20 % teoretických četností může být menších než 5
- pro tabulku 2x2 (čtyřpolní) musí navíc platit, že:
  - $n > 40$ , jinak
  - pro  $20 < n < 40$ , je nutná úprava test. kritéria pomocí tzv. Yatesovy korekce
  - pro  $n < 20$  je třeba použít tzv. přesný Fisherův test.



# Test nezávislosti v kontingenční tabulce

## Příklad:

Bylo dotazováno 800 absolventů ČVUT a UK. Z výzkumu plyne, že 322 dotazovaných jsou absolventi ČVUT kteří pracují dále ve svém oboru, 234 dotázaných absolvovalo UK a pracují také ve svém oboru. Ve 158 případech odpovědí bylo zjištěno, že absolventi ČVUT ve svém oboru nepracují a zbývajících 86 jsou absolventi UK, kteří také nepracují ve svém oboru.

Lze považovat uplatnění se v oboru studia stochasticky závislé na absolvované univerzitě?

```
> pruzkum <- matrix(c(322, 234, 158, 86),2)
> colnames(pruzkum)= c("CVUT","UK")
> rownames(pruzkum)= c("v oboru", "mimo obor")
> pruzkum
```

```
          CVUT  UK
v oboru   322 158
mimo obor 234  86
```

	ČVUT	UK	celkem
v oboru	322	234	556
mimo obor	158	86	244
celkem	480	320	800

```
> chisq.test(pruzkum)
```

Pearson's Chi-squared test with Yates' continuity correction

data: pruzkum, X-squared = 3.0273, df = 1, p-value = 0.08187



# Test nezávislosti v kontingenční tabulce

**Fisherův test** je založen na tzv. poměru šancí (Odds ratio) výskytu určité události v závislosti na události druhé (v našem příkladu „CVUT“ a „v oboru“). Kvantifikuje tedy sílu vztahu mezi těmito dvěma veličinami.

$$OR = \frac{P(\text{jev nastane})}{P(\text{jev nenastane})} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

OR = 1 : žádná závislost mezi zkoumanými událostmi – pravděpodobnost nastání jedné události je stejná nezávisle na přítomnosti či nepřítomnosti druhé

OR > 1 : přítomnost jedné události navyšuje šanci na nastání události druhé

OR < 1 : přítomnost jedné události snižuje šanci na nastání události druhé

> `fisher.test(pruzkum)`

Fisher's Exact Test for Count Data

	ČVUT ano	ČVUT ne	celkem
v oboru ano	322	234	556
v oboru ne	158	86	244
celkem	480	320	800

data: pruzkum, p-value = 0.07187

alternative hypothesis: true odds ratio is not equal to 1

95 percent confidence interval: 0.5407144 1.0343715

sample estimates: odds ratio 0.7492708



# Pravděpodobnostní metody ve strojírenství

## 13. Síla testu a rozsah výběru



# 13. Síla testu a rozsah výběru

- Klíčové pojmy:**
- Síla testu
  - Rozsah výběru
- Klíčové vztahy:**
- Test chí-kvadrát
  - Fisherův test

# Síla testu

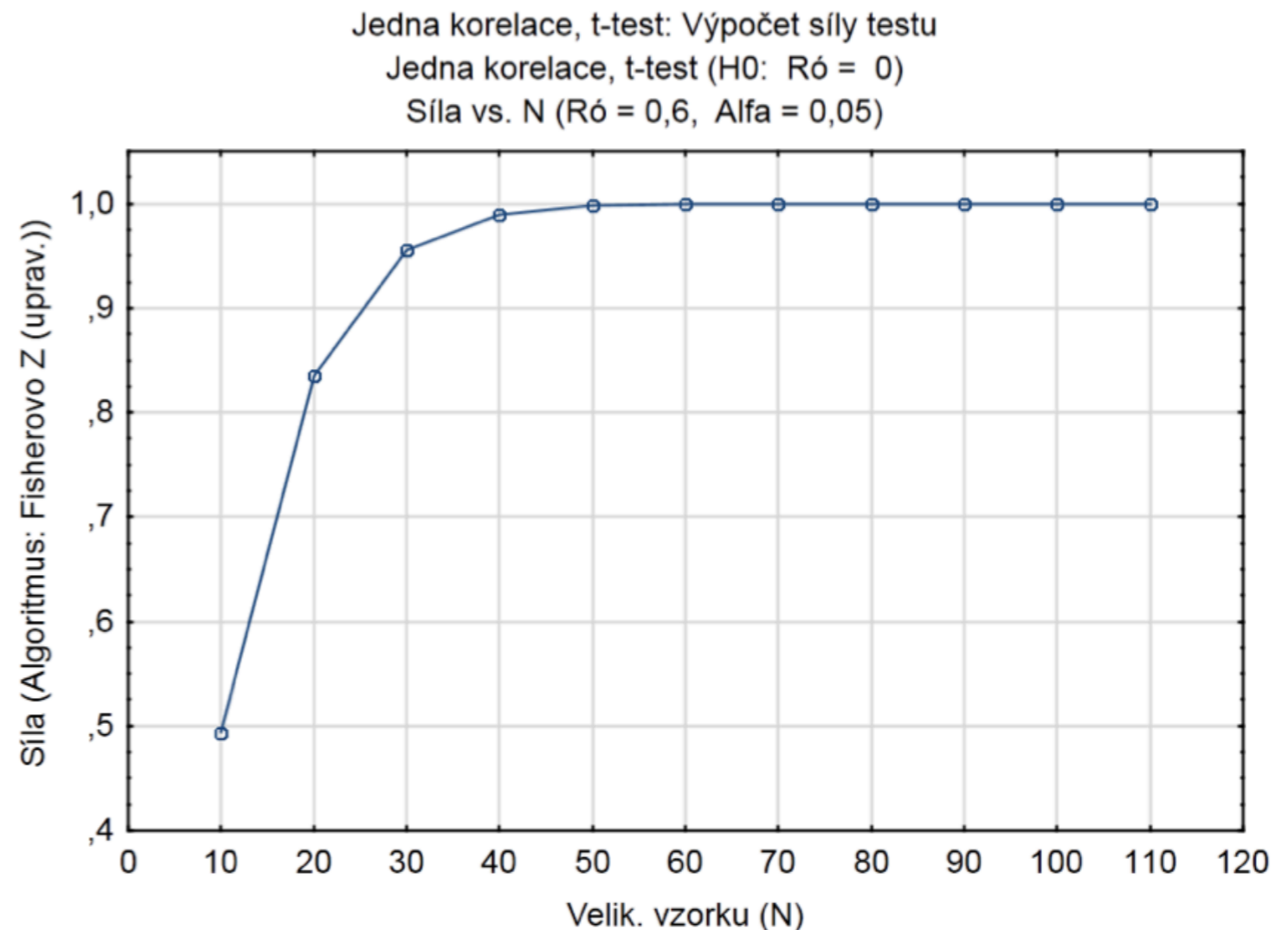
**Síla testu** je pravděpodobnost toho, že test zamítne hypotézu když neplatí, tedy je rovna  $(1 - \beta)$ , kde  $\beta$  je pravděpodobnost chyby II. druhu (= nezamítáme nulovou hypotézu když nepatí)

**Síla testu** je vlastnost testu, tedy závisí především na:

- testové statistice
- alternativní hypotéze
- na velikosti výběru
- no a samozřejmě na datech

**V případě t-testu** závisí na

- rozptylu dat  $\sigma^2$
- předpokládané diferenci  $\delta$
- na velikosti výběru  $n$



# Síla testu

**Jednovýběrový t-test (nebo párový t-test):** Za předpokladu alternativní hypotézy  $H_A: \mu = m_0 + \delta$  má testová statistika tzv. necentrální t-rozdělení s parametrem

necentrality  $\nu = \frac{\delta}{\sigma} \sqrt{n}$

```
> davky <- data.matrix(read.table("davkovac.txt"))
```

```
> n = 50; delta = 0.2
```

```
> noncentral = delta*sqrt(n)/sd(davky)
```

```
> curve(pt(x,n-1,ncp=noncentral), from=0, to=10)
```

```
> abline(v=qt(0.975, n-1))
```

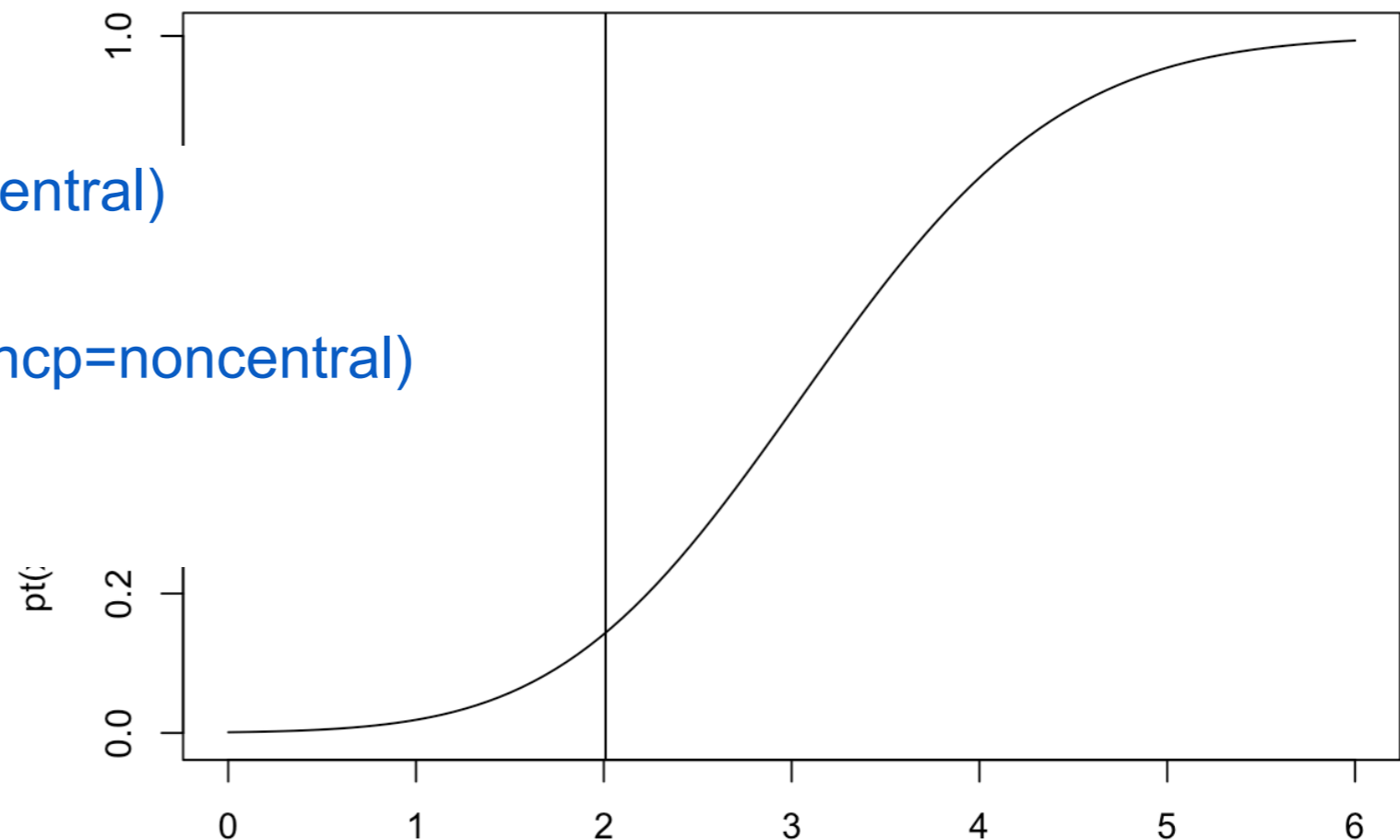
```
> pt(qt(0.975,n-1), n-1,ncp=noncentral)
```

```
[1] 0.1438221
```

```
> power=1-pt(qt(0.975,n-1), n-1,ncp=noncentral)
```

```
> power
```

```
[1] 0.8561779
```



# Síla testu

## Jednovýběrový t-test (nebo párový t-test):

```
> power.t.test(delta=0.2, sd=sd(davky), sig.level=0.05, power=0.9,  
  type="one.sample")
```

One-sample t test power calculation

```
      n = 57.18447  
delta = 0.2  
  sd = 0.458528  
sig.level = 0.05  
  power = 0.9  
alternative = two.sided
```

Při ručním počítání obvykle předpokládáme znalost směrodatné odchylky  $\sigma$  a namísto t-rozdělení použijeme standardní normální rozdělení (a jeho kvantily  $u_p$ ). Požadovaná síla testu je zde označena  $\beta$ . Potom lze odvodit přibližný vztah:

$$n = \left[ \left( \frac{u_{\alpha/2} + u_{\beta}}{\delta/\sigma} \right)^2 \right] + 1$$



# Síla testu

**Dvouvýběrový t-test:** Za předpokladu alternativní hypotézy  $H_A: \mu_X - \mu_Y = \delta$  má testová statistika necentrální t-rozdělení s parametrem necentrality

$$\nu = \frac{\delta}{\sigma \sqrt{1/n_X + 1/n_Y}}$$

```
> x <- data.matrix(read.table("dodavatelX.txt"))
> y <- data.matrix(read.table(„dodavatelY.txt“))
> z=c(x,y)

> power.t.test(delta=2, sd=sd(z), sig.level=0.05, power=0.9)
```

Two-sample t test power calculation

```
      n = 43.59226
delta = 2
      sd = 2.847857
sig.level = 0.05
      power = 0.9
alternative = two.sided
```

NOTE: n is number in \*each\* group



# Síla testu

**Dvouvýběrový t-test:** Za předpokladu alternativní hypotézy  $H_A: \mu_X - \mu_Y = \delta$  má testová statistika necentrální t-rozdělení s parametrem necentrality

$$\nu = \frac{\delta}{\sigma \sqrt{1/n_X + 1/n_Y}}$$

```
> x <- data.matrix(read.table("dodavatelX.txt"))
> y <- data.matrix(read.table(,"dodavatelY.txt"))
> z=c(x,y)

> power.t.test(n=100, delta=2, sd=sd(z), sig.level=0.05)
```

Two-sample t test power calculation

```
      n = 100
  delta = 2
     sd = 2.847857
sig.level = 0.05
  power = 0.9985668
alternative = two.sided
```

NOTE: n is number in \*each\* group



# Síla testu

**Dvouvýběrový t-test:** Za předpokladu alternativní hypotézy  $H_A: \mu_X - \mu_Y = \delta$  má testová statistika necentrální t-rozdělení s parametrem necentrality

$$\nu = \frac{\delta}{\sigma \sqrt{1/n_X + 1/n_Y}}$$

```
> x <- data.matrix(read.table("dodavatelX.txt"))
> y <- data.matrix(read.table(„dodavatelY.txt“))
> z=c(x,y)

> power.t.test(delta=2, sd=sd(z), sig.level=0.05, power=0.9,
alt="one.sided")
```

Two-sample t test power calculation

```
      n = 35.42511
delta = 2
      sd = 2.847857
sig.level = 0.05
      power = 0.9
alternative = one.sided
```

NOTE: n is number in \*each\* group



# Síla testu

**Dvouvýběrový t-test:** Za předpokladu alternativní hypotézy  $H_A: \mu_X - \mu_Y = \delta$  má testová statistika necentrální t-rozdělení s parametrem necentrality

$$\nu = \frac{\delta}{\sigma \sqrt{1/n_X + 1/n_Y}}$$

```
> x <- data.matrix(read.table("dodavatelX.txt"))
> y <- data.matrix(read.table(„dodavatelY.txt“))
> z=c(x,y)

> power.t.test(delta=2, sd=5.0, sig.level=0.05, power=0.9,
alt="one.sided")
```

Two-sample t test power calculation

```
      n = 107.7313
delta = 2
  sd = 5
sig.level = 0.05
  power = 0.9
alternative = one.sided
```

NOTE: n is number in \*each\* group



# Síla testu

**Dvouvýběrový t-test:** Za předpokladu alternativní hypotézy  $H_A: \mu_X - \mu_Y = \delta$  má testová statistika necentrální t-rozdělení s parametrem necentrality

$$\nu = \frac{\delta}{\sigma \sqrt{1/n_X + 1/n_Y}}$$

Při ručním počítání obvykle předpokládáme znalost směrodatné odchylky  $\sigma$  a namísto t-rozdělení použijeme standardní normální rozdělení (a jeho kvantily  $u_p$ ). Požadovaná síla testu je zde označena  $\beta$ . Potom lze odvodit přibližný vztah:

$$n = 2 \cdot \left[ \left( \frac{u_{\alpha/2} + u_{\beta}}{\delta/\sigma} \right)^2 \right] + 1$$

(Toto  $n$  je pro každou skupinu.)



# Síla testu

## Porovnání dvou poměrů:

```
> power.prop.test(n=400, p1=0.67, p2=0.73)
```

Two-sample comparison of proportions power calculation

```
      n = 400
     p1 = 0.67
     p2 = 0.73
sig.level = 0.05
  power = 0.4567771
alternative = two.sided
```

NOTE: n is number in \*each\* group



# Síla testu

## Porovnání dvou poměrů:

```
> power.prop.test(power=0.9, p1=0.67, p2=0.73)
```

Two-sample comparison of proportions power calculation

```
      n = 1223.788
    p1 = 0.67
    p2 = 0.73
sig.level = 0.05
  power = 0.9
alternative = two.sided
```

NOTE: n is number in \*each\* group



# Síla testu

## Porovnání dvou poměrů:

```
> power.prop.test(power=0.85, p1=0.67, p2=0.73)
```

Two-sample comparison of proportions power calculation

```
      n = 1045.926
    p1 = 0.67
    p2 = 0.73
sig.level = 0.05
  power = 0.85
alternative = two.sided
```

NOTE: n is number in \*each\* group

