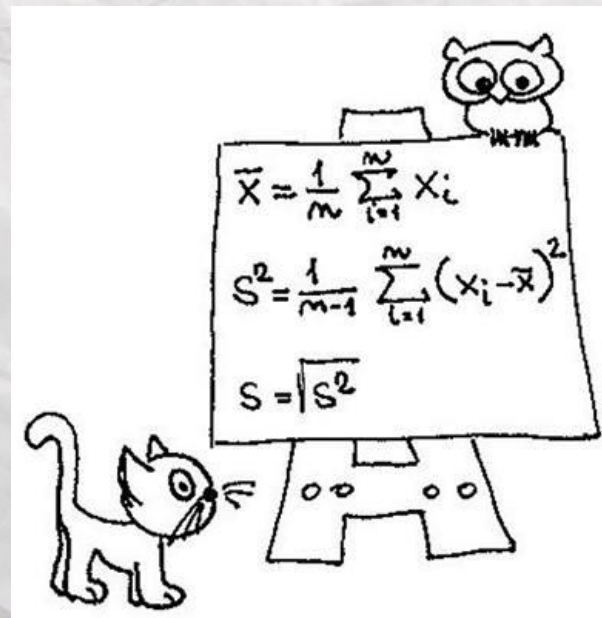


# Statistické metody v medicíně II. - p-hodnota

Martin Hynek

Gennet, Centre for Fetal Medicine, Prague

EuroMISE Centre, First Faculty of Medicine of Charles University in Prague



# p-hodnota

p-value (p-level)

dosažená hladina významnosti

# P

statistická významnost při testování hypotéz

# p-hodnota

široce rozšířený, ale velmi záludný  
koncept, velmi často špatně chápán  
a interpretován, často kritizován



# P

Nutné pro : publikace, granty  
porozumění odborné literatuře



# p-hodnota

## ULTRASOUND in Obstetrics & Gynecology



12 z 14 článků

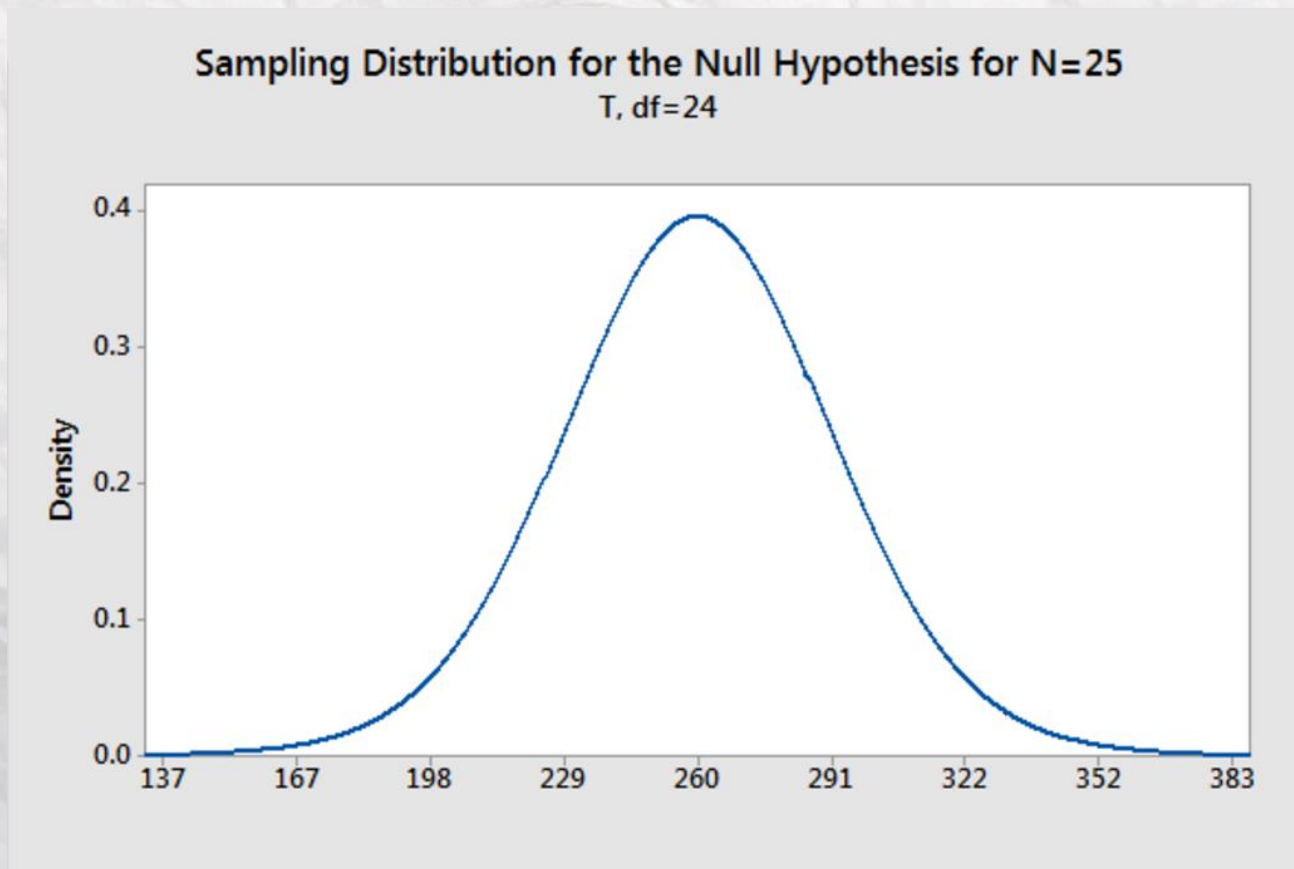
5 z 10 článků

# Testování hypotéz

velmi rozšířený koncept v medicíně  
porovnávání (typu léčby, metod, skupin) – **efekt**

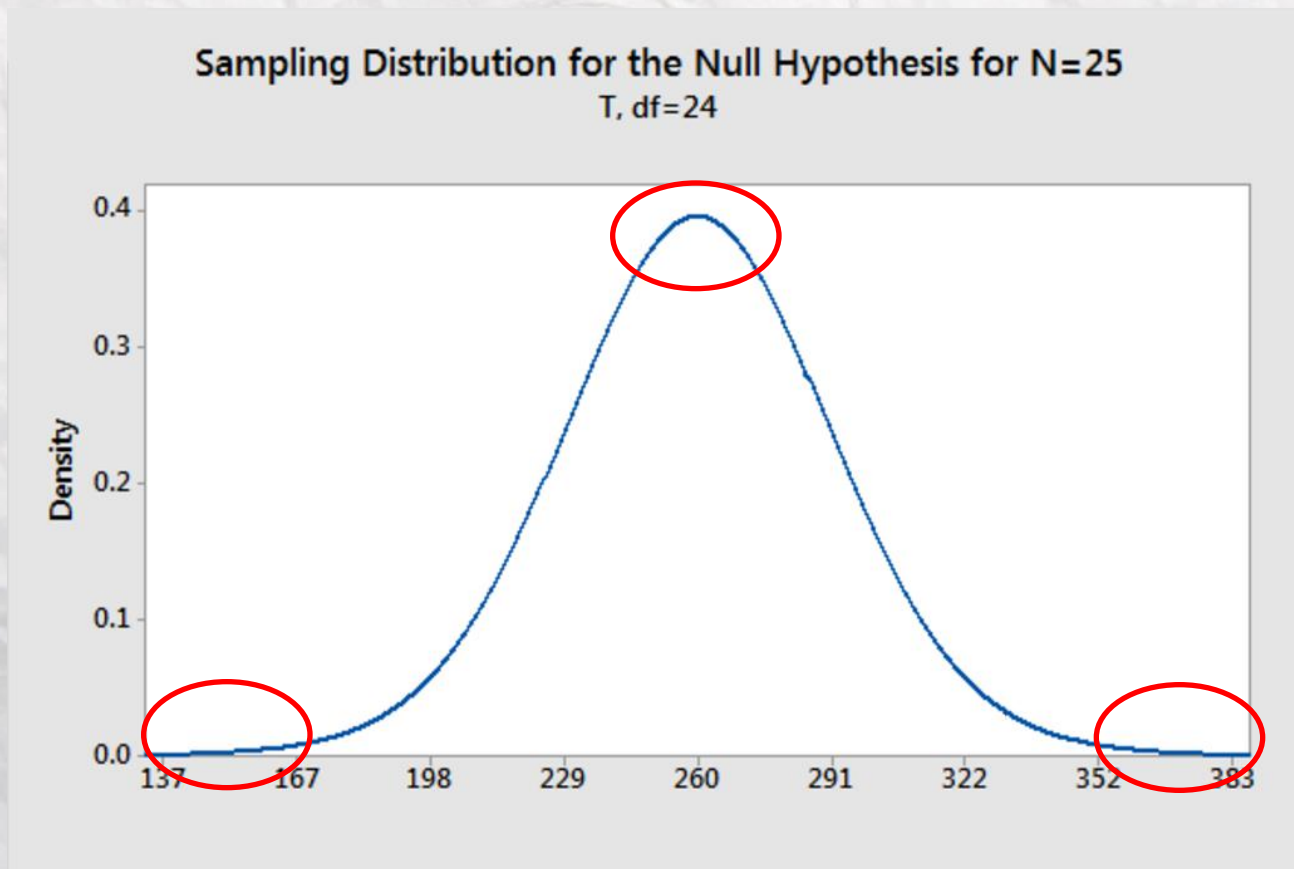
# Testování hypotéz

velmi rozšířený koncept v medicíně  
 porovnávání (typu léčby, metod, skupin) – **efekt**



# Testování hypotéz

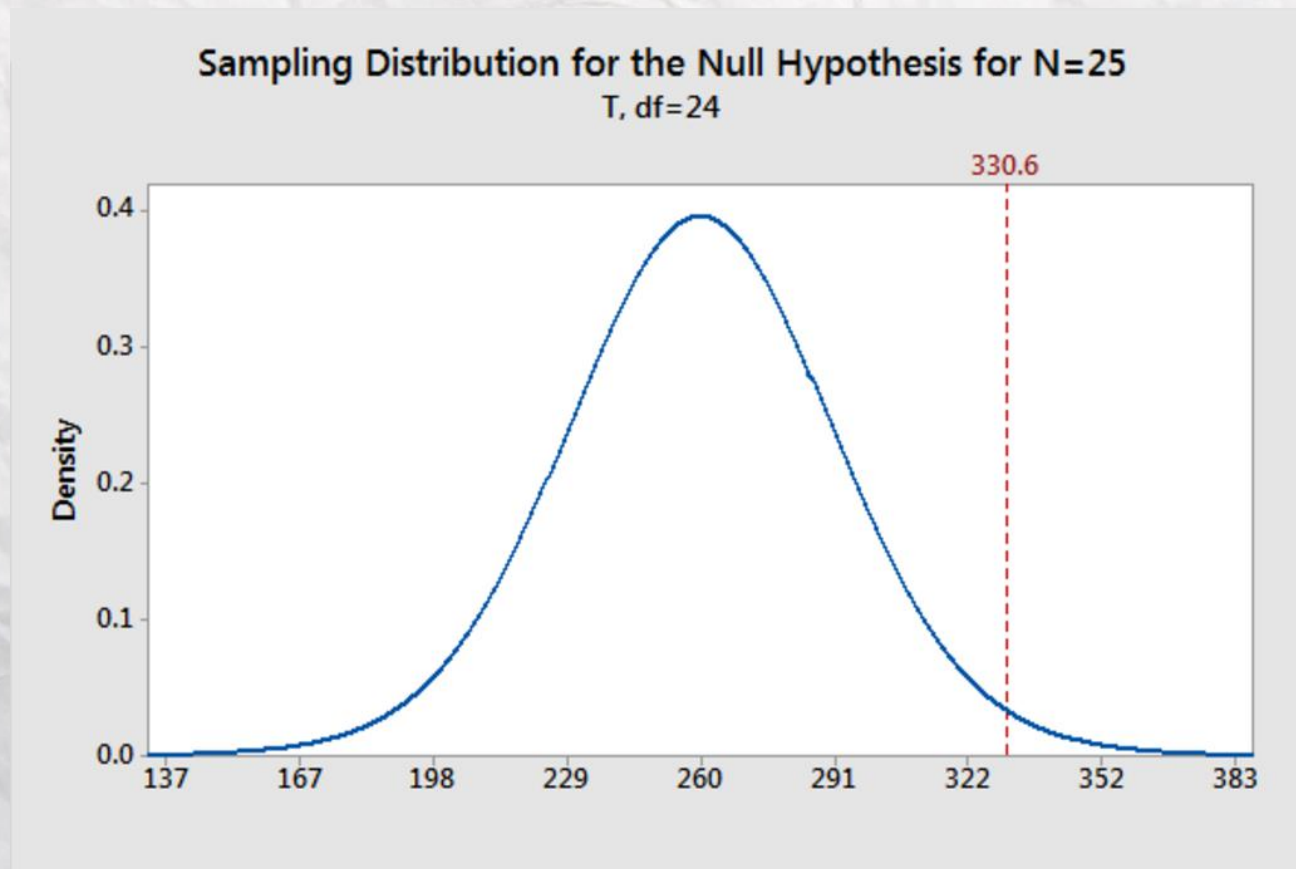
velmi rozšířený koncept v medicíně  
 porovnávání (typu léčby, metod, skupin) – **efekt**





# Testování hypotéz

velmi rozšířený koncept v medicíně  
 porovnávání (typu léčby, metod, skupin) – **efekt**





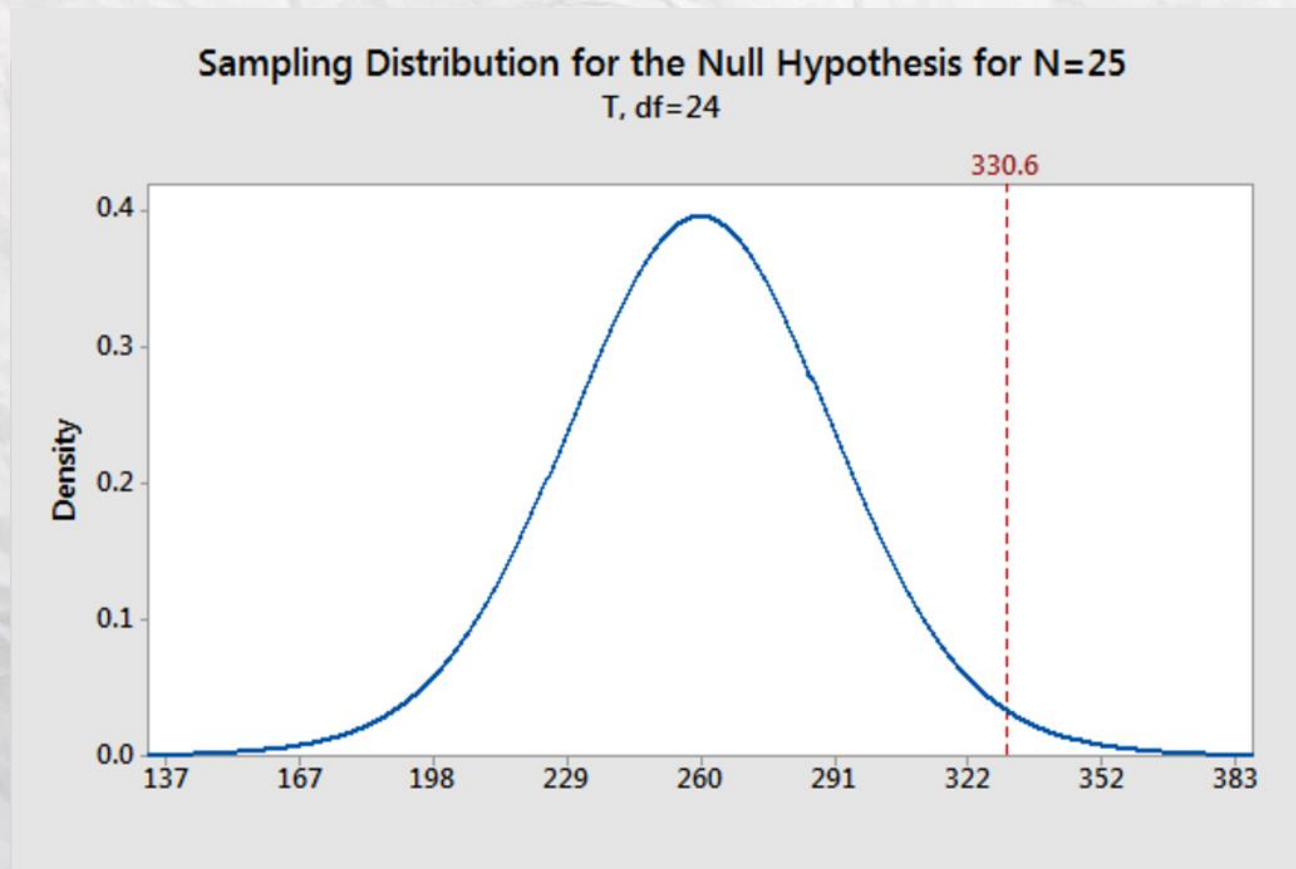
# Testování hypotéz

nulový efekt

=> nulová hypotéza  $H_0$

rozdíl existuje

=> alternativní hypotéza  $H_A$



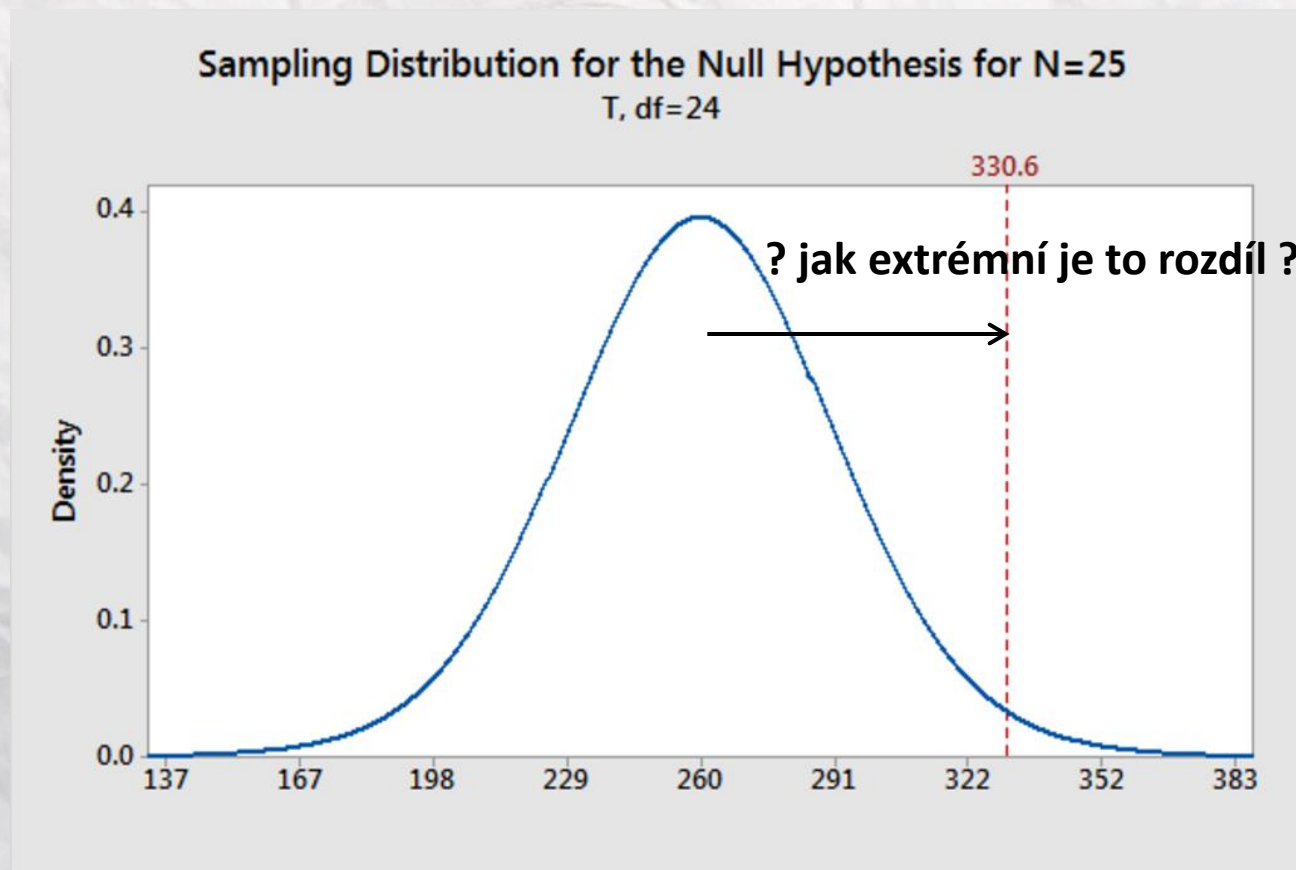
# Testování hypotéz

nulový efekt

=> nulová hypotéza  $H_0$

rozdíl existuje

=> alternativní hypotéza  $H_A$



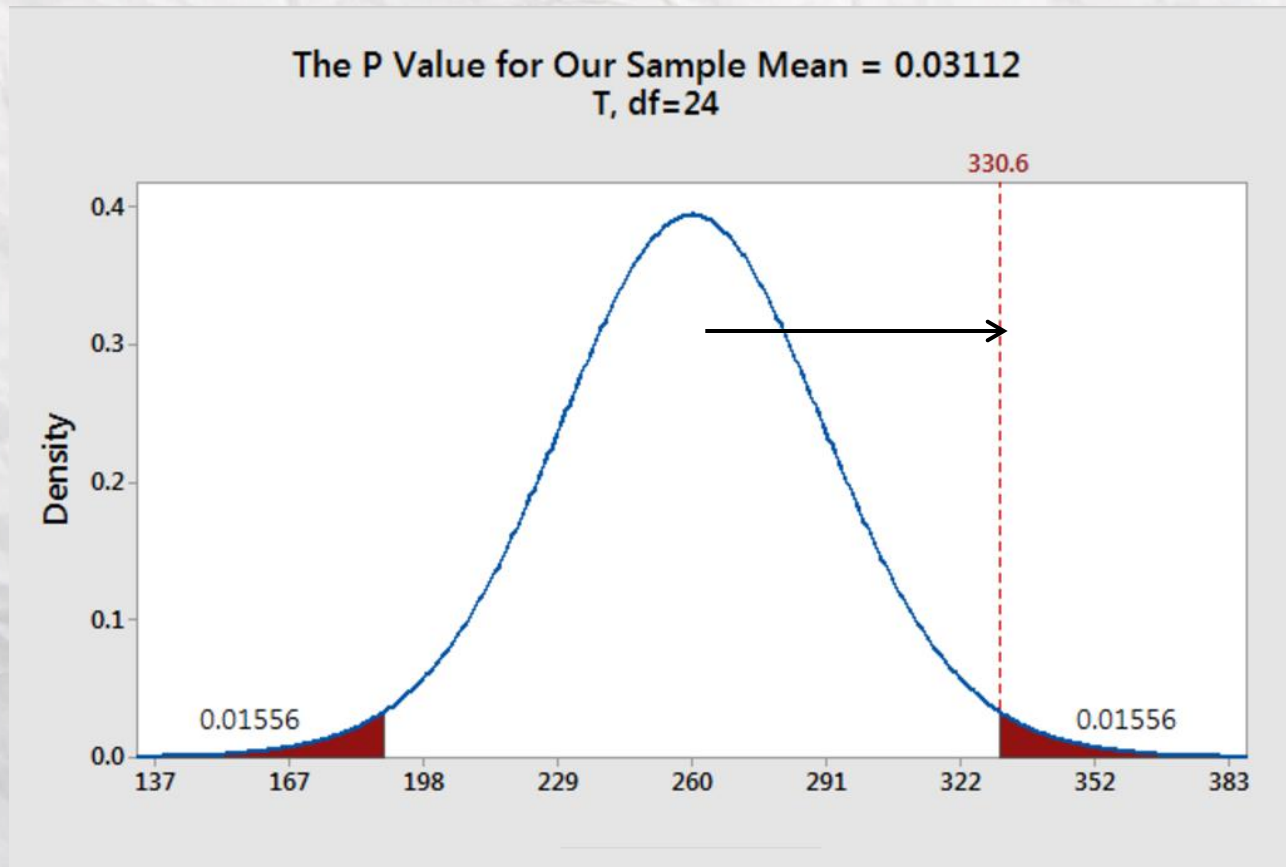
# Testování hypotéz

nulový efekt

=> nulová hypotéza  $H_0$

rozdíl existuje

=> alternativní hypotéza  $H_A$



**p = 0,03112**

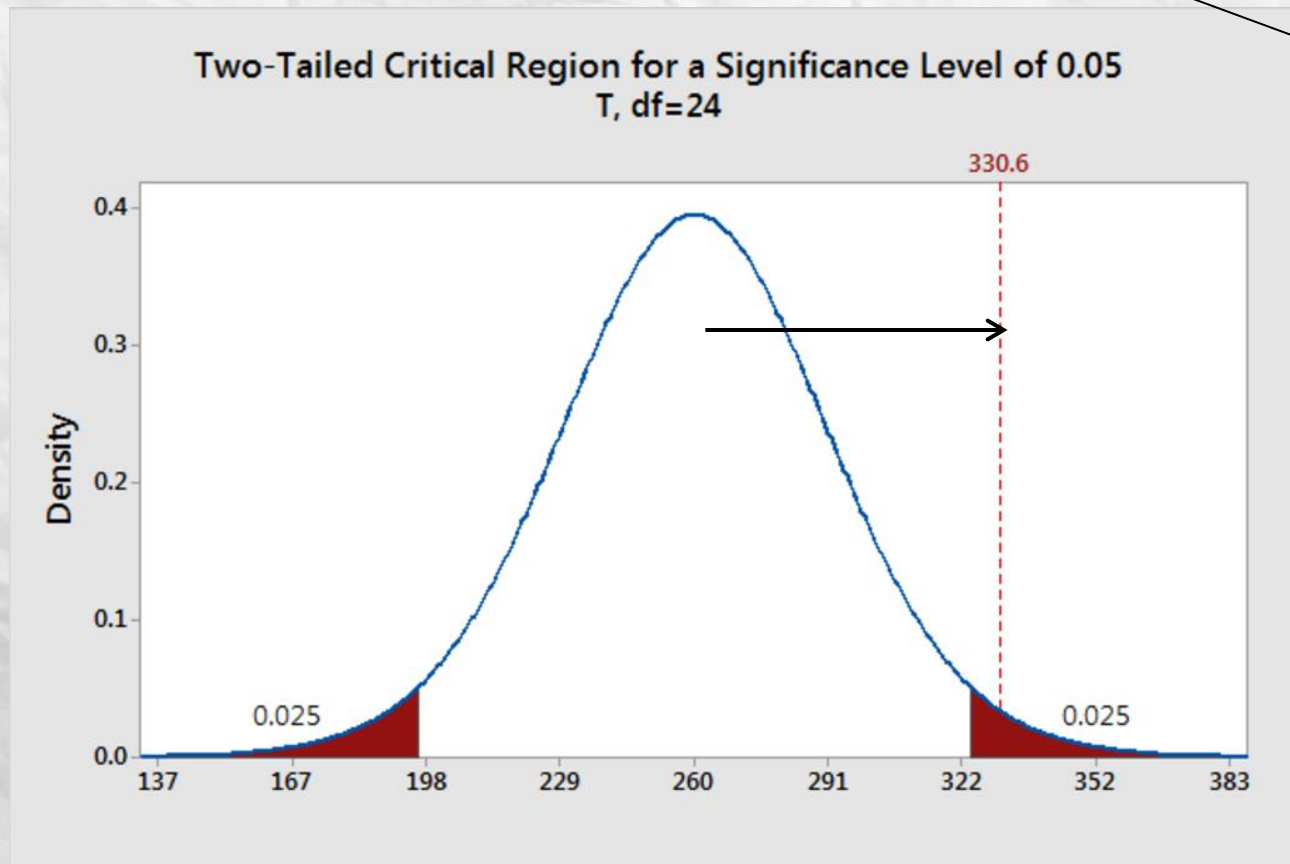
# Testování hypotéz

hladina významnosti  $\alpha = 0,05$  (5%)

$p < \alpha$

$p > \alpha$

~~$H_0$~~   
 $H_A$



$H_0$   
 ~~$H_A$~~

$p = 0,03112$





# Testování hypotéz

**p-hodnota** = pravděpodobnost, s jakou bychom mohli obdržet pozorovaná data nebo data „extrémnější“ (více odporující nulové hypotéze), *za předpokladu, že nulová hypotéza je pravdivá*

**CAVE**

**!! p není pravděpodobnost platnosti nulové hypotézy**



**p = 0,03** neznamená že pravděpodobnost  $H_0$  je 3%  
a pravděpodobnost  $H_A$  je 97%

# Nepochopení a úskalí a p-hodnoty

navzdory všudypřítomnosti p-hodnoty je tento koncept velmi hojně kritizovaný

únor 2015 – Journal of Basic and Applied Social Psychology – editoři vykázali testování hypotéz a p-hodnotu z časopisu

plyne z nepochopení a špatného použití



# Nepochopení a úskalí a p-hodnoty

p-hodnota není pravděpodobnost, že nulová hypotéza je pravdivá ani že alternativní hypotéza je nepravdivá

? jaká tato pravděpodobnost vlastně je



# Nepochopení a úskalí a p-hodnoty

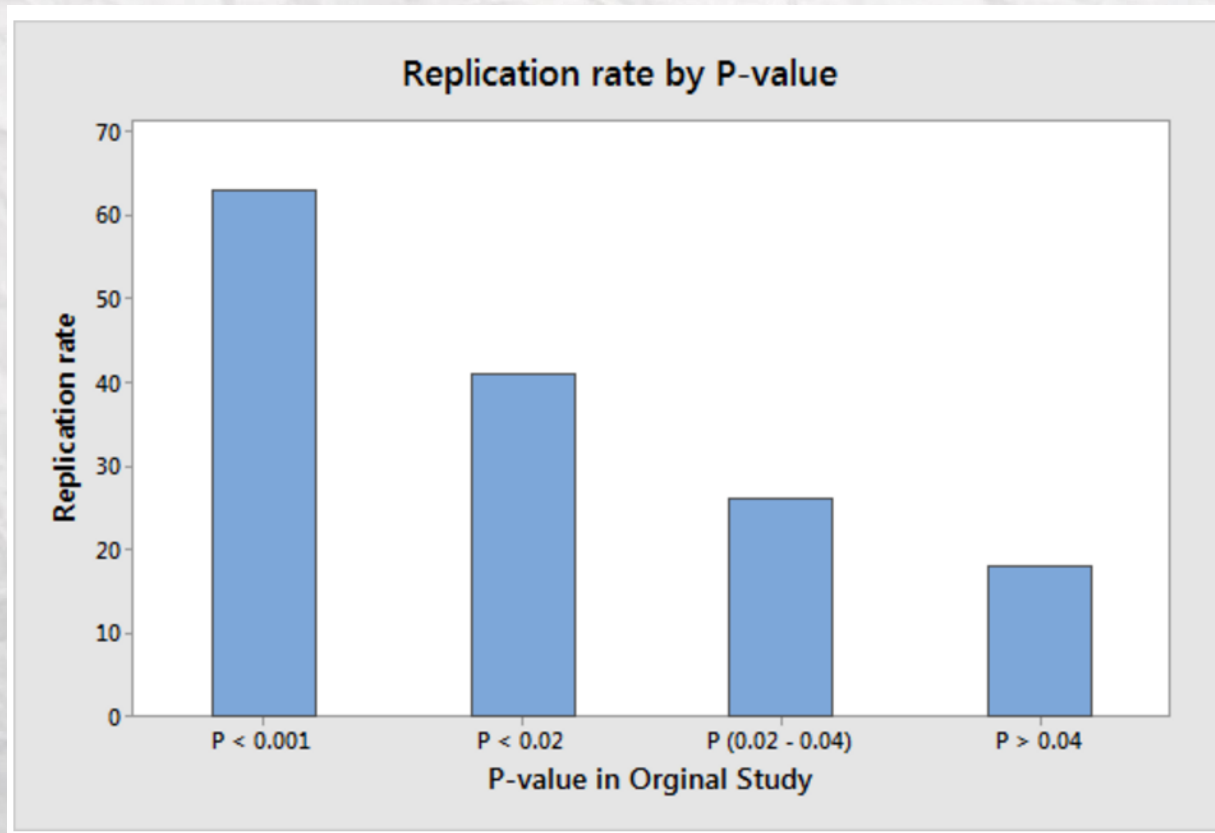
p-hodnota není pravděpodobnost že nulová hypotéza je pravdivá ani že alternativní hypotéza je nepravdivá

? jaká tato pravděpodobnost vlastně je

<b>P value</b>	<b>Probability of incorrectly rejecting a true null hypothesis</b>
0.05	At least 23% (and typically close to 50%)
0.01	At least 7% (and typically close to 15%)

# Nepochopení a úskalí a p-hodnoty

p-hodnota není pravděpodobnost, že při opakování experimentu dospějeme ke stejnému závěru



# Nepochopení a úskalí a p-hodnoty

p-hodnota se týká pouze nulové hypotézy

neříká absolutně nic o platnosti alternativní hypotézy

počítám p-hodnotu, která se týká  $H_0$ ,

ale to, co mě zajímá je  $H_A$  !!!

nesmyslnost alternativní hypotézy nebo existence více alternativních hypotéz



# Nepochopení a úskalí a p-hodnoty

**hladina významnosti  $\alpha$  není určena p-hodnotou**

historická konvence  $\alpha = 0,05$

**dichotomizace** – signifikantní vs.  
nesignifikantní

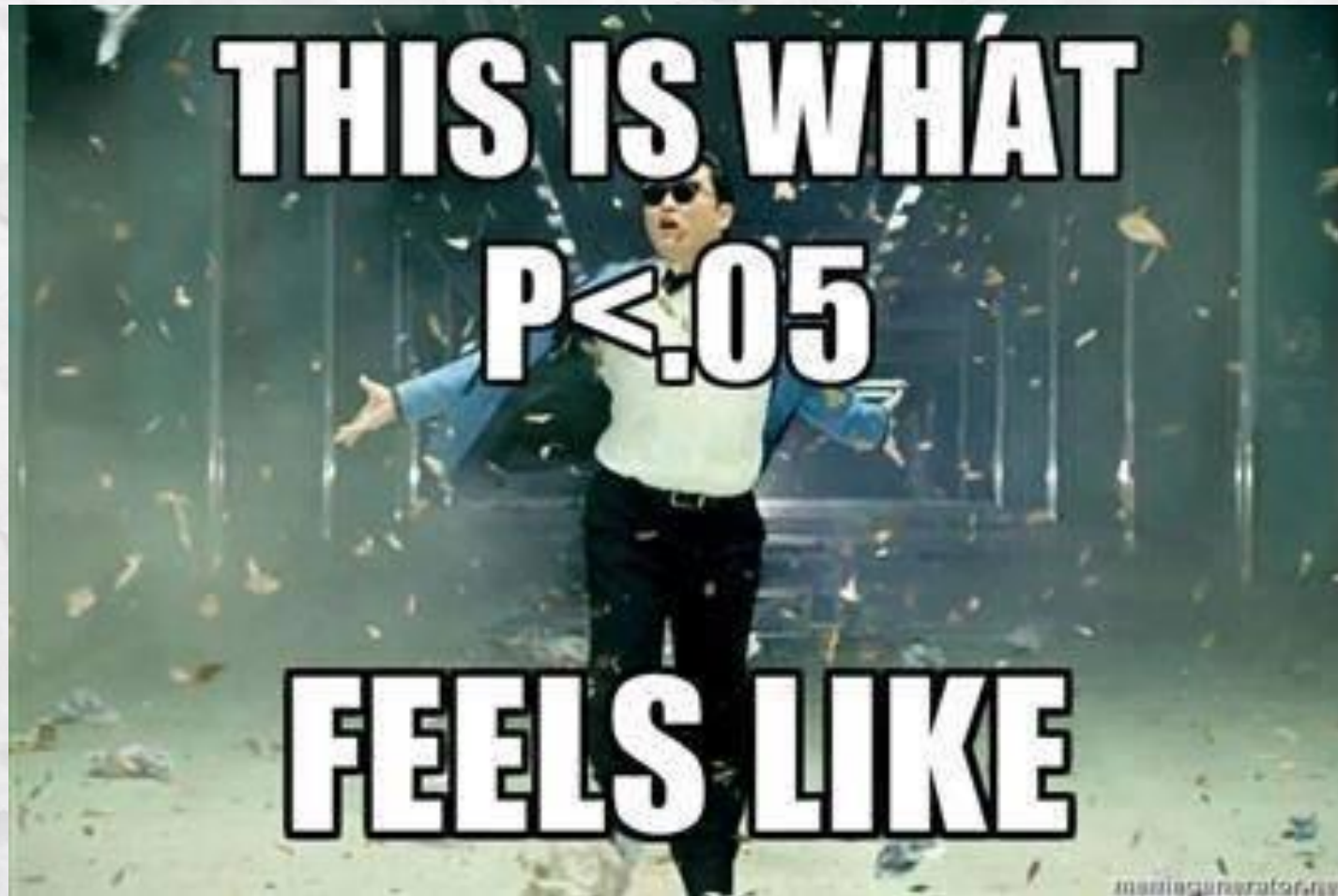


vždy uvádět přesnou p-hodnotu

- dlouhodobě argumentováno, že takto arbitrárně stanovená meze zpomaluje vědecký pokrok a má škodlivý vliv na odbornou literaturu

(Anderson 2000, Berkson 2006, Cohen 1994, Hubbard and Lindsey 2008, Johnstone 1986, Smedslund 2006)





jednoznačný publikační bias ve prospěch studií se  
signifikantním závěrem !!

# Nepochopení a úskalí a p-hodnoty

$p > 0,05$  „failed to reach statistical significance“



# Nepochopení a úskalí a p-hodnoty

Freiman et al. analyzovali 71 studií, kde  $p > 0,1$  a spočítali konfidenční intervaly

→ v téměř polovině studií byly výsledky kompatibilní s 50% terapeutickým efektem



# Nepochopení a úskalí a p-hodnoty

p-hodnota nic neříká o velikosti pozorovaného efektu

- 1. léčba  
p=0,02
- 2. léčba  
p=0,002
- 3. léčba  
p=0,02
- studie s AIDS vakcínou v Thajsku 2009  
p=0,039 !!  
ale 95% konfidenční interval 1-52%



# Testování hypotéz

**p-hodnota** = pravděpodobnost, s jakou bychom mohli obdržet pozorovaná data nebo data „extrémnější“ (více odporující nulové hypotéze), *za předpokladu, že nulová hypotéza je pravdivá*

**CAVE**

**!! p není pravděpodobnost platnosti nulové hypotézy**

**p = 0,03** neznamená že pravděpodobnost  $H_0$  je 3%  
a pravděpodobnost  $H_A$  je 97%

# FOUR keypoints for using p-value

- testování hypotéz a p-hodnota jsou užitečné nástroje a poskytují důležité informace, ale nikdy ne samostatně (nástroje deskriptivní statistiky)
- nutnost správné interpretace
- signifikantní výsledky musí být vždy potvrzeny z opakovaných studií, abychom jim mohli věřit
- abychom zhodnotili kvalitu studie, je nutné se dívat na celý všechny její dílčí kroky a ne jen na p-hodnotu

Děkuji za pozornost

