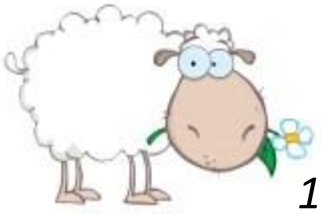
Přednáška 7



Příklad : Nepárový dvouvýběrový t-test

1. skupina, N=30

- Průměrná hmotnost ovcí v čase páření byla srovnávána pro kontrolní skupinu a skupinu krmenou zvýšenou dávkou potravy. Kontrolní skupina obsahuje 30 ovcí, skupina se zvýšeným příjmem potravy pak 24 ovcí.

- Vlastní experiment byl prováděn tak, že na začátku máme 54 ovcí (ideálně stejného plemene, stejně staré atd.), které náhodně rozdělíme do dvou skupin (náhodné rozdělování objektů do pokusných skupin je objektem celého specializovaného odvětví statistiky nazývaného randomizace). Poté co experiment proběhne, musíme nejprve ověřit teoretický předpoklad pro využití nepárového t-testu. Pro obě skupiny jsou vykresleny grafy (můžeme též spočítat základní popisnou statistiku), na kterých můžeme posoudit normalitu a homogenitu rozptylu, kromě okometrického pohledu můžeme pro <u>ověření normality použít testy normality, pro ověření homogenity rozptylu pak F-test.</u>
- Pokud platí všechny předpoklady dvouvýběrového nepárového t-testu, můžeme spočítat testovou statistiku, výsledné *t* je 2,43 s 52 stupni volnosti, podle tabulek je a $t_{0,975\,(52)} = 2{,}01$, tedy $|t| > t_{0,975\,(52)}$ a nulovou hypotézu můžeme zamítnout, skutečná pravděpodobnost je pak 0,018. Rozdíl mezi skupinami je 1,59 kg ve prospěch skupiny se zvýšeným příjmem.

$$t = \frac{Rozdil.prumeru}{SE(rozdil.prumeru)} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \qquad \upsilon = n_1 + n_2 - 2$$

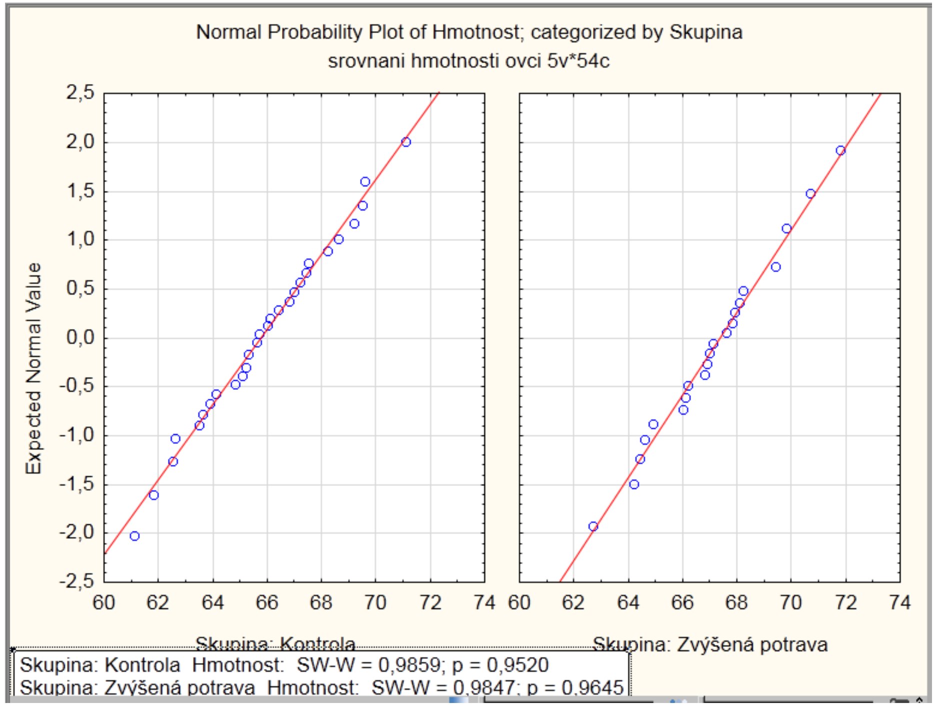
- Pro rozdíl mezi oběma soubory jsou spočítány 95% intervaly spolehlivosti jako 1,59±2.01*(0,655) kg, což odpovídá rozsahu 0,28 až 2,91 kg. To, že interval spolehlivosti nezahrnuje 0 je dalším potvrzením, že mezi skupinami je významný rozdíl – jde o další způsob testování významnosti rozdílů mezi skupinami dat – nulovou hypotézu o tom, že rozdíl průměrů dvou skupin dat je roven nějaké hodnotě zamítáme v případě, kdy 95% interval spolehlivosti rozdílu nezahrnuje tuto hodnotu (v tomto případě 0).

$$(\bar{x}_1 - \bar{x}_2) \pm t_{0,975} SE(\bar{x}_1 - \bar{x}_2) = (\bar{x}_1 - \bar{x}_2) \pm t_{0,975} \sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

2. skupina, N=24

Příklad : Nepárový dvouvýběrový t-test

- Nejprve ověřte normalitu hmotnosti jednak ve skupině kontroly a ve skupině se zvýšenou potravou



- V obou případech se tečky odchylují od přímky jenom málo a p-hodnoty S-W testu převyšují 0,05. Předpoklad o normálním rozložení dat v obou skupinách je oprávněný.

Příklad : Nepárový dvouvýběrový t-test

•***<u>POZOR: Výstupní tabulku vyhodnocujeme zezadu!!!</u>***

Výběrový průměr u 1. skupiny

Výběrový průměr u 2. skupiny

Výběrová směrodatná odchylka u 2. skupiny

Rozsah výběru 1. skupiny

Rozsah výběru 2. skupiny

T-tests; Grouping: Skupina (srovnani hmotnosti ovci)
Group 1: Kontrola
Group 2: Zvýšená potrava

Variable	Mean Kontrola	Mean Zvýšená potrava	t-value	df	p	Valid N Kontrola	Valid N Zvýšená potrava	Std.Dev. Kontrola	Std.Dev. Zvýšená potrava	F-ratio Variances	p Variances
Hmotnost	65,77333	67,36667	-2,43226	52	0,018483	30	24	2,497162	2,25247	1,229066	0,617383

Hodnota testové statistiky
(pro test shody středních hodnot)

Počet stupňů volnosti

Testová statistika pro test shody rozptylů
(F-test)

<u>Tyto sloupce lze interpretovat pouze pokud rozdíl mezi rozptyly byl neprůkazný !!!</u>

Párový dvouvýběrový t-test

- Skupiny dat jsou spojeny přes objekt měření, příkladem může být měření parametrů pacienta před léčbou a po léčbě (nemusí jít přímo o stejný objekt, dalším příkladem mohou být např. krysy ze stejné linie).
- Oba soubory musí mít shodný počet hodnot, protože všechna měření v jednom souboru musí být spárována s měřením v druhém souboru. Při vlastním výpočtu se potom počítá se změnou hodnot (diferencí) subjektů v obou souborech.
- V případě, že se nejedná o měření na témže subjektu je vhodné si před párovým testem ověřit si, zda existuje vazba mezi oběma skupinami – vynesení do grafu, korelace.

Existuje několik možných designů experimentu, stručně lze sumarizovat:

- pokus je párový a jako párový se projeví
- párové provedení pokusu – párově se neprojeví
    - možná párovost není
    - špatně provedený pokus – malé n, velká variabilita, špatný výběr jedinců
- čekali jsme nezávislé a jsou
- čekali jsem nezávislé a nejsou
    - vazba
    - náhoda

Párový dvouvýběrový t-test

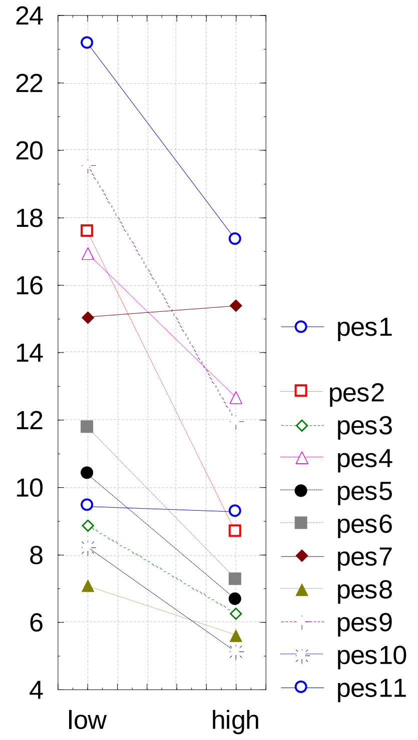
- Tento test nemá žádné předpoklady o rozložení vstupních dat, protože je počítán až na základě jejich diferencí.
- Tyto diference by měly být normálně rozloženy a otázkou v párovém t-testu je, zda se průměrná hodnota diferencí rovná nějakému číslu, typicky jde o srovnání s nulou jako důkaz neexistence změny mezi oběma spárovanými skupinami.
- V podstatě jde o one sample t-test, kde místo rozdílu průměru vzorku a cílové populace je uveden průměr diferencí a srovnávané číslo (0 v případě otázky, zda není rozdíl mezi vzorky).
- Pro srovnání s 0 (testovou statistikou je t rozložení): $$t = \frac{\bar{D}}{s}\sqrt{n} \qquad \upsilon = n-1$$
- Někdy je obtížné rozhodnout, zda jde nebo nejde o párové uspořádání, párový test by měl být použit pouze v případě, že můžeme potvrdit vazbu (korelace, vynesení do grafu), jedním z důvodů proč toto ověřovat je fakt, že v případě párového t-testu není nutné brát ohled na variabilitu původních dvou souborů, tento předpoklad však platí pouze v případě vazby mezi proměnnými. Výpočet obou typů testů se vlastně liší v použité s, jednou jde o s diferencí, v druhém případě o složený odhad rozptylu obou souborů.
- Zda je párové uspořádání efektivnější lze určit na základě:
  - Síly vazby
  - Je-li s_D výrazně menší než s_{x1-x2}
- Závislost je možné rozepsat pomocí vzorce: $$s_D^2 \cong \sigma_{x_1}^2 + \sigma_{x_2}^2 - 2Cov(x_1; x_2)$$
- v případě Cov=0, tedy v případě neexistence vazby pak s_D^2 odpovídá součtu původních rozptylů, tedy přibližně S_{x1-x2}.

Příklad 1: Párový dvouvýběrový test

- Byl prováděn pokus s dietou 11 diabetických psů, každý pes byl vystaven dvěma dietám s odlišným typem sacharidů (snadno vstřebatelné X pozvolna se rozkládající na glukózu), hodnoty krevní glukózy v průběhu jednotlivých diet mají být srovnány pro zjištění vlivu diety na hladinu krevní glukózy. Protože každý pes absolvoval obě diety, jde o párové uspořádání, kdy výsledky hodnoty v obou pokusech jsou spojeny přes pokusné zvíře.

1. Nulová hypotéza zní, že skutečný průměrný rozdíl mezi oběma dietami je 0, alternativní hypotéza zní, že to není 0.
2. Pro každého psa je spočítán rozdíl mezi jeho hladinou glukózy při obou dietách a měly by být ověřeny předpoklady pro one sample t-test – tedy alespoň přibližně normální rozložení.
3. Je spočítána testová charakteristika, výpočet vlastně probíhá jako one-sample t-test, kde je zjišťována významnost průměru diferencí obou souborů jako rozdíl mezi touto hodnotou a nulou (nula je hodnota, kterou by průměrná diference měla nabývat, pokud platí nulová hypotéza). T=4.37 s 10 stupni volnosti, skutečná hodnota p=0,0014 a tedy na hladině p=0,05 můžeme nulovou hypotézu zamítnou

$$t = \frac{rozdíl\_průměru\_vzorku\_a\_populace}{SE(průměru)} = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \frac{\bar{x}-\mu}{s}\sqrt{n}$$



4. Závěrem můžeme říci, že nulová hypotéza neexistence rozdílu mezi oběma dietami byla zamítnuta, což znamená, že high-fibre dieta má významný vliv na snížení hladiny krevní glukózy.

Příklad 2: Párový dvouvýběrový test

- Byl prováděn pokus s dietou u 18 diabetických krys, každá krysa byla vystavena dvěma dietám (jedné nové speciální a jedné kontrolní dietě). Protože každá krysa absolvovala obě diety, jde o párové uspořádání, kdy hodnoty v obou pokusech jsou spojeny přes pokusné zvíře. Zjistěte, zda testovaná dieta způsobí změnu hmotnosti u krys (zda se liší hmotnost krys po nové speciální a po kontrolní dietě).

1. Nulová hypotéza zní, že skutečný průměrný rozdíl v hmotnosti krys po speciální a kontrolní dietě je nulový (speciální dieta nevedla ke změně hmotnosti ve srovnání s kontrolní dietou), alternativní hypotéza zní, že rozdíl hmotností je odlišný od nuly (speciální dieta vedla ke změně hmotnosti ve srovnání s kontrolní dietou).
2. Pro každou krysu je spočítán rozdíl hmotností naměřených po obou dietách a měly by být ověřeny předpoklady pro jednovýběrový t-test – alespoň přibližně normální rozložení diferencí.
3. Je spočítána testová statistika, výpočet vlastně probíhá jako jednovýběrový t-test, kde je zjišťována významnost průměru diferencí obou souborů jako rozdíl mezi touto hodnotou a nulou (0 je hodnota, kterou by průměrná diference měla nabývat, pokud platí nulová hypotéza). T=-1,72 s 17 stupni volnosti, skutečná p-hodnota=0,102 a tedy na hladině významnosti α=0,05 nemůžeme nulovou hypotézu zamítnout.

$$t = \frac{rozdíl\_průměru\_vzorku\_a\_populace}{SE(průměru)} = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \frac{\bar{x}-\mu}{s}\sqrt{n}$$

4. Závěrem můžeme říci, že nulová hypotéza neexistence rozdílu vlivu na snížení váhy mezi oběma dietami nebyla zamítnuta.

Příklad: Párový dvouvýběrový test

Výběrový průměr

Výběrová směrodatná odchylka

Počet pozorování

T-test for Dependent Samples (efektivita diety pro krysy)
Marked differences are significant at p < ,05000

Variable	Mean	Std.Dv.	N	Diff.	Std.Dv. Diff.	t	df	p
testovaná dieta	186,0556	59,52011						
kontrolní dieta	191,7222	69,65022	18	-5,66667	13,91994	-1,72714	17	0,102266

Hodnota testovacího kritéria

Průměrná hodnota diferencí

Výběrová směrodatná odchylka diferencí

Neparametrické dvouvýběrové statistické testy

Nepárový Mannův-Whitneyův test

Párový Wilcoxonův a znaménkový test

Mannův-Whitneyův U test

- Neparametrická alternativa dvouvýběrového t-testu.
- Počítá s pořadím dat v souborech namísto s originálními daty.
- Předpoklad: rozdělení pravděpodobnosti veličiny ve skupinách se může lišit pouze posunutím.

Postup:

1. Stanovíme nulovou a alternativní hypotézu (F(x)=distribuční funkce):

$$H_0: F(x_1)=F(x_2)$$
$$H_1: F(x_1)\neq F(x_2).$$

2. Čísla obou souborů jsou sloučena a je určeno jejich pořadí v tomto sloučeném souboru.
3. Pro oba výběry zvlášť je spočítán součet pořadí (T_1 a T_2).
4. Ze součtů pořadí ve skupinách je určena finální hodnota testové statistiky U.

$$U_1 = n_1 n_2 + \frac{n_1-(n_1+1)}{2} - T_1 \qquad U_2 = n_1 n_2 + \frac{n_2-(n_2+1)}{2} - T_2$$

$$U = \min(U_1, U_2)$$

5. Hodnotu testové statistiky U porovnáme s kritickou hodnotou testu, pokud je tato hodnota menší než kritická hodnota testu, zamítáme nulovou hypotézu shody distribučních funkcí obou skupin.

Mannův-Whitneyův U test – asymptotická varianta

5. Pro velká n_1 a n_2 (>30) lze využít asymptotické normality statistiky U.

$$E(U) = \frac{n_1 n_2}{2} \qquad D(U) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

6. Pro testování lze využít Z-statistiky:

$$Z = \frac{U - E(U)}{\sqrt{D(U)}} \approx N(0,1)$$

7. Pokud $|Z| \geq u_{1-\alpha/2}$ zamítáme nulovou hypotézu o shodnosti distribučních funkcí

Mannův-Whitneyův U test

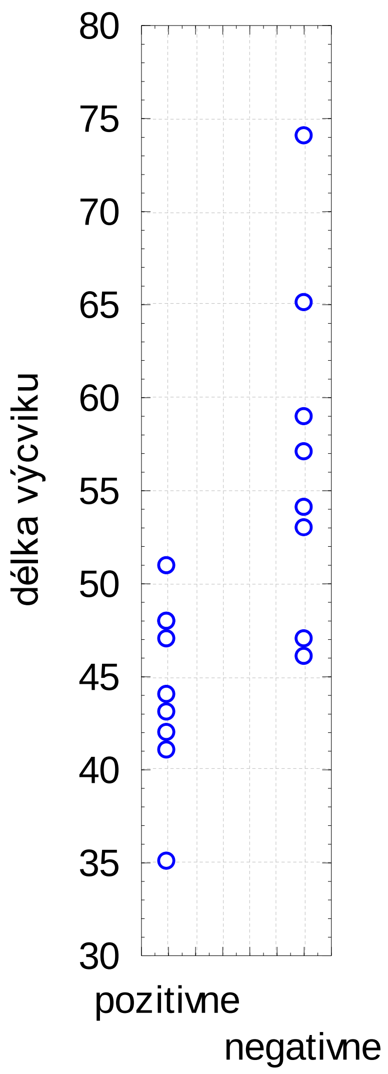
- Stejně jako řada jiných neparametrických testů počítá i tento test s pořadím dat v souborech namísto s originálními daty. Jde o neparametrickou obdobu nepárového t-testu a z těchto neparametrických testů má nejvyšší sílu testu (95% párového t-testu).
- V případě Mann-Whitney testu jsou nejprve čísla obou souborů sloučena a je vytvořeno jejich pořadí v tomto sloučeném souboru, pak jsou hodnoty vráceny do původních souborů a nadále se pracuje již jen s jejich pořadím.
- Pro oba soubory je tedy vytvořen součet pořadí a menší z obou součtů je porovnán s kritickou hodnotou testu, pokud je tato hodnota menší než kritická hodnota testu, zamítáme nulovou hypotézu shody distribučních funkcí obou skupin.

X1	X2	ALL	Rank ALL	X1 rank	X2 rank
27	25	25	5	6	5
35	29	29	7,5	11	7,5
38	31	31	9	13	9
37	23	23	4	12	4
39	18	18	2	14	2
29	17	17	1	7,5	1
41	32	32	10	15	10
	19	19	3		3
		27	6		
		35	11		
		38	13		
		37	12		
		39	14		
		29	7,5		
		41	15		

Příklad: Mannův-Whitneyův U test

- 17 štěňat bylo trénováno v chození na záchod metodou pozitivní motivace (pochvala, když jde na záchod venku) nebo negativní motivace (trest, když jde na záchod doma). Jako parametr bylo měřeno, za kolik dní je štěně vycvičeno.
- Nulová hypotéza je, že není rozdíl v metodách tréninku, tedy, že oběma metodami je štěně vycvičeno za stejnou dobu.
- Po srovnání rozložení + kvůli nízkému počtu hodnot je vhodné použít neparametrický test.
- Je vytvořeno pořadí hodnot v kompletním souboru.
- Hodnota testové statistiky je určena ze součtu pořadí hodnot v jednotlivých skupinách.

- **Jak dopadne testování?**



Příklad: Řešení v softwaru

Součet pořadí T_1

Součet pořadí T_2

Hodnota Z statistiky

Mann-Whitney U Test (Spreadsheet15)
By variable skupina
Marked tests are significant at p <,05000

variable	Rank Sum Group 1	Rank Sum Group 2	U	Z	p-value	Z adjusted	p-value	Valid N Group 1	Valid N Group 2	2*1sided exact p
delka	49,50000	103,5000	13,50000	-2,11695	0,034265	-2,11955	0,034045	8	9	0,027396

Hodnota testové statistiky

Asymptotická p-hodnota

Přesná p-hodnota
(použít, jestliže rozsah výběru je menší než 30)

Párový Wilcoxonův a znaménkový test

- Vycházíme z rozdílů párových hodnot a přecházíme na design jednovýběrových testů
- Testuje, zda je medián diferencí (D) párových hodnot roven hodnotě *c*

  H_0: $D_{0.5}$=c proti H_1: $D_{0.5}$≠ c.

Wilcoxonův párový test

1. Spočítáme rozdíly **diferencí** výběru s testovanou hodnotou mediánu = c.
2. Absolutní hodnoty rozdílů uspořádáme vzestupně a přiřadíme jim pořadí.
3. Spočítáme statistiky S_w^+ a S_w^-, které odpovídají součtu pořadí kladných (S_w^+) a záporných rozdílů (S_w^-). Jako finální hodnotu testové statistiky bereme minimum z S_w^+ a S_w^-. Nulovou hypotézu zamítáme, pokud hodnota testové statistiky menší nebo rovna tabelované kritické hodnotě (při dané hladině významnosti a počtu nenulových rozdílů).

Znaménkový párový test

1. Spočítáme rozdíly **diferencí** výběru s testovanou hodnotou mediánu = c.
2. Spočítáme statistiku S_z^+, která odpovídá počtu kladných rozdílů → test nevyužívá hodnot pořadí původních dat ale pouze informaci, zda se hodnota realizuje nad nebo pod mediánem → dochází ke snížení síly testu
3. Nulovou hypotézu zamítáme, pokud statistika S_z^+ realizuje v kritickém oboru hodnot $W=(0,k_1)U(k_2,n)$, kde *n* odpovídá počtu nenulový rozdílů a hodnoty k_1 a k_2 lze dohledat v matematických tabulkách.

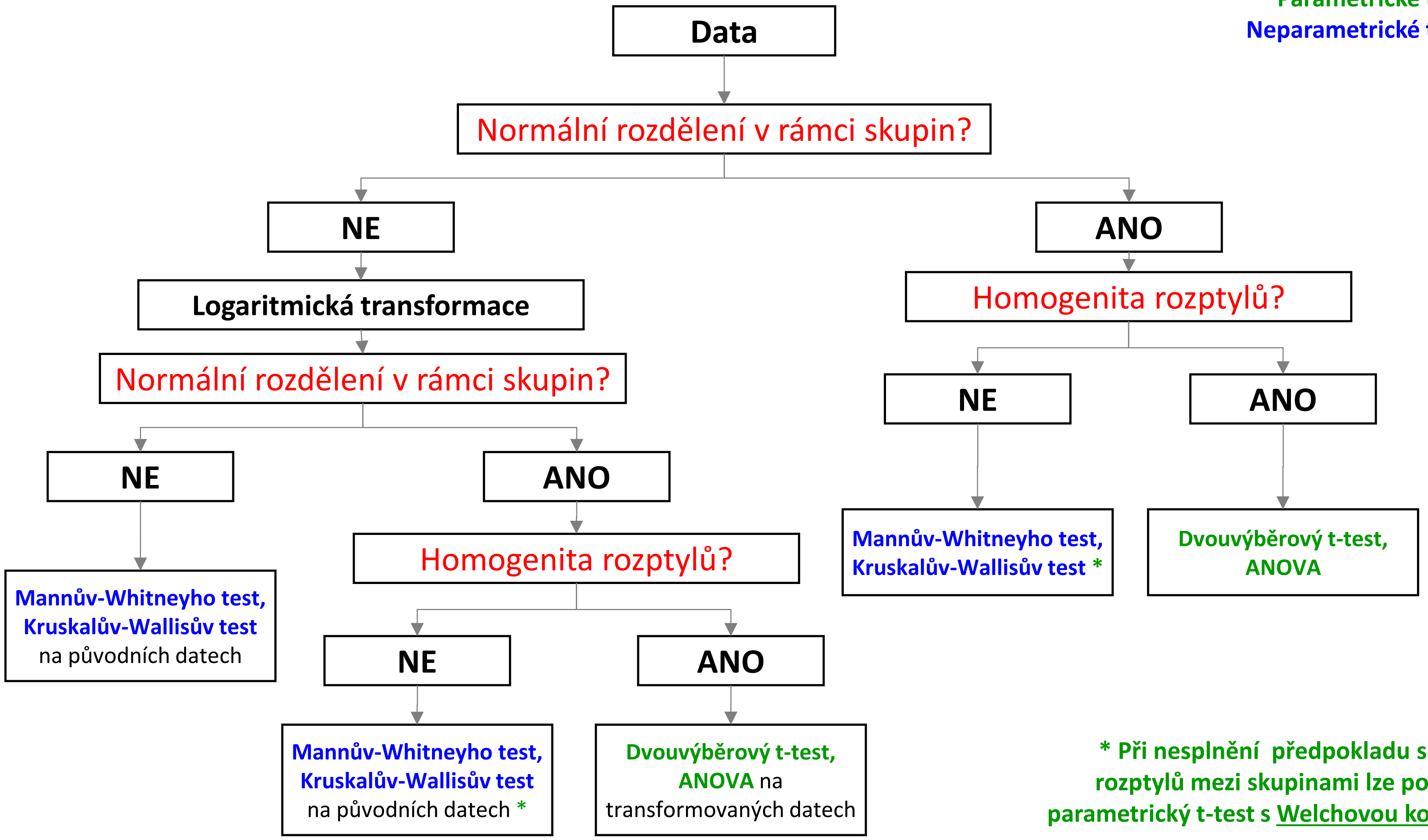
Příklad 2: Párový dvouvýběrový test

Byla testována nová dieta pro laboratorní krysy, při pokusu byl zjišťován její vliv na hmotnost v různých liniích krys, bylo proto zvoleno párové uspořádání kdy krysy v obou dietách jsou spojeny přes svoji linii, tj. na začátku byly dvojice krys stejné linie, jedna z nich byla náhodně přiřazena k dietě, druhá z dvojice pak do druhé diety.

1. nulová hypotéza je, že váha krys není ovlivněna použitou dietou, alternativní, že ovlivnění dietou existuje
2. spočítáme diference – tyto diference jsou nenormální a proto je vhodné využít neparametrický test
3. Spočítáme sumu pořadí kladných a záporných diferencí, zde je menší suma záporných diferencí – 31
4. výsledkem výpočtu je $p>0{,}05$ a tedy nemáme dostatečné důkazy pro zamítnutí nulové hypotézy, nelze říci, že by nová dieta byla efektivnější než stará
5. pro doplnění výsledků je vhodné zjistit také skutečnou velikost rozdílu hmotností ve skupinách, např. ve formě mediánu

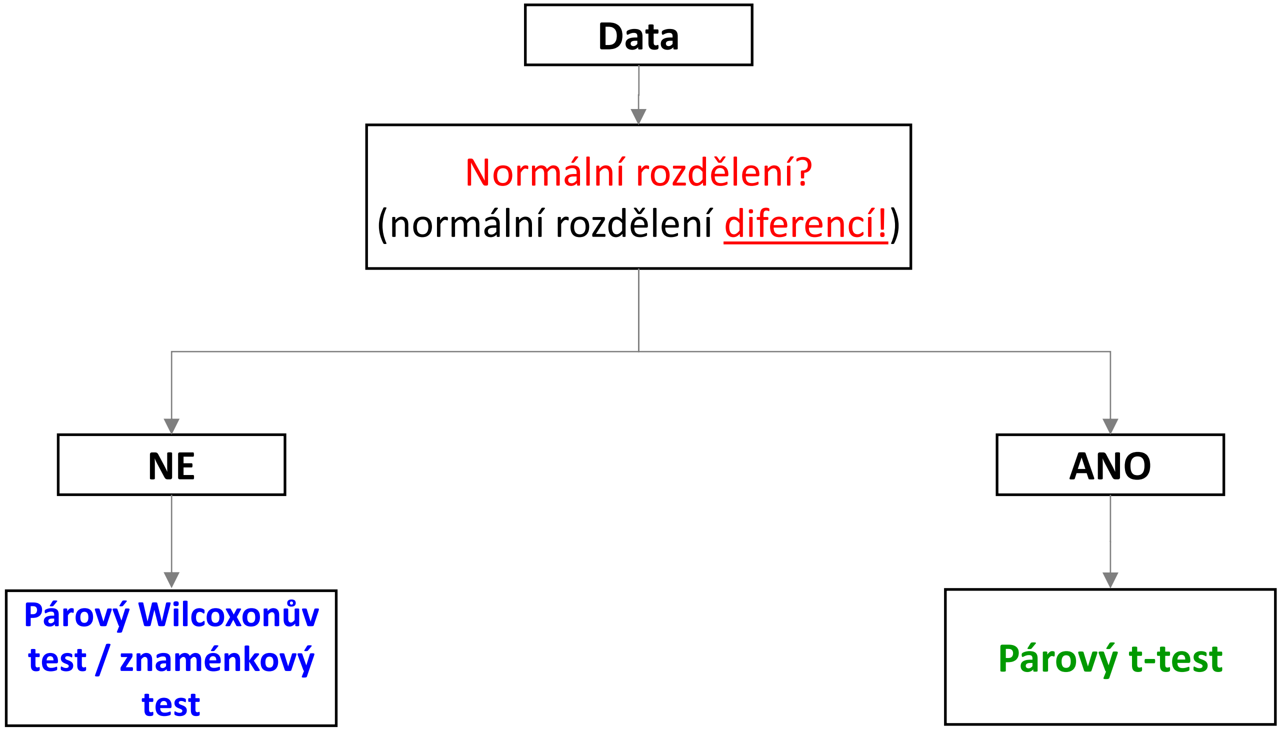
Schéma při testování 2 a více skupin

Parametrické testy
Neparametrické testy



* Při nesplnění předpokladu shody rozptylů mezi skupinami lze použít i parametrický t-test s Welchovou korekcí

Schéma při testování pomocí párových testů



Parametrické testy
Neparametrické testy

Binomické rozdělení

Popis binomického rozložení

Testování hypotéz binomicky rozložených dat

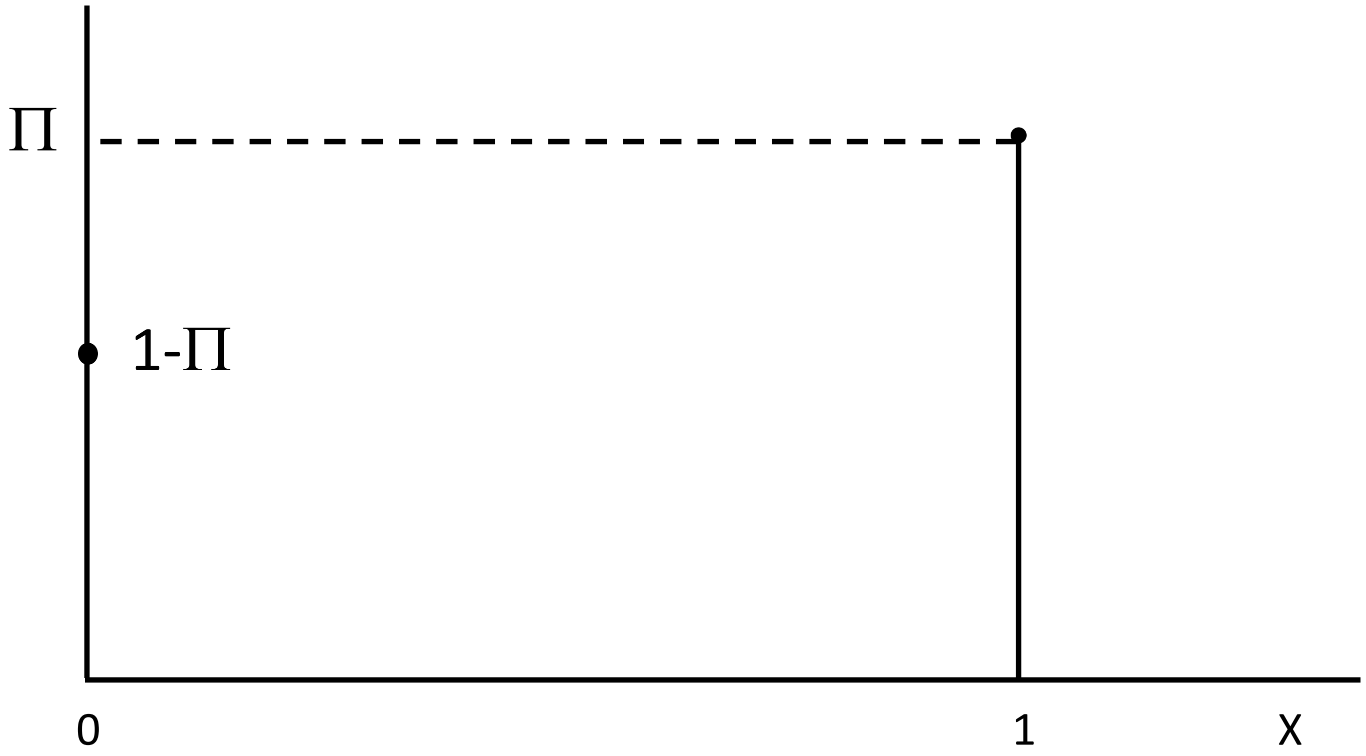
Anotace

- Kromě spojitých dat se setkáváme také s daty kategoriálními, jejichž nejjednodušším případem jsou data binární.
- Binární data jsou popsána binomickým rozdělením, od chování binomického rozdělení je odvozena popisná statistika binárních dat (procento výskytu jevu), její interval spolehlivosti a binomické testy pro srovnání procentuálního výskytů jevů v různých skupinách.

Alternativní rozdělení

- Nastane jedna ze dvou možných varianta

$$\left.\begin{array}{l}\Pi(x) = \Pi \text{ pro } X = 1 \\ \Pi(x) = 1 - \Pi \text{ pro } X = 0 \\ \Pi(x) = 0 \text{ jinak}\end{array}\right\} \quad X = 1 \text{ ......jev}$$



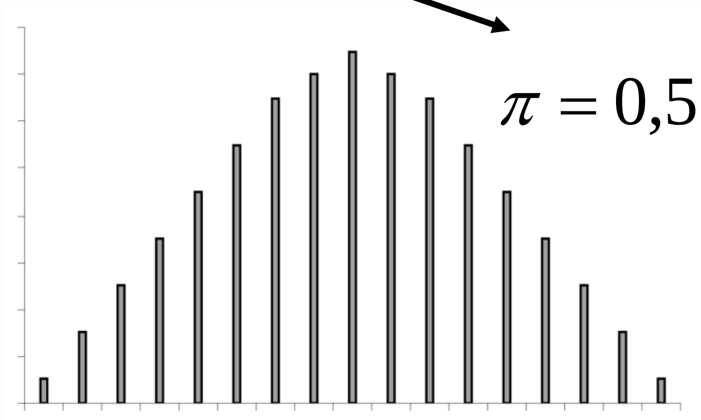
Binomické rozdělení

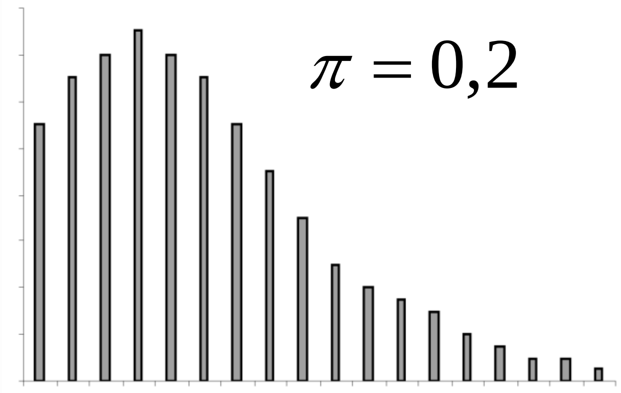
X ..... celkový počet nastání jevu v n nezávislých pokusech

E(x)= n . Π
D(x)= n . Π (1-Π)

Π ~ p

jediný parametr distribuce
určuje tvar distribuce





Binomické rozdělení jako model pro zkoumání výskytu sledovaného jevu

n ..... počet nezávislých opakování (dotazů)

X ..... počet lidí s jistým symptomem

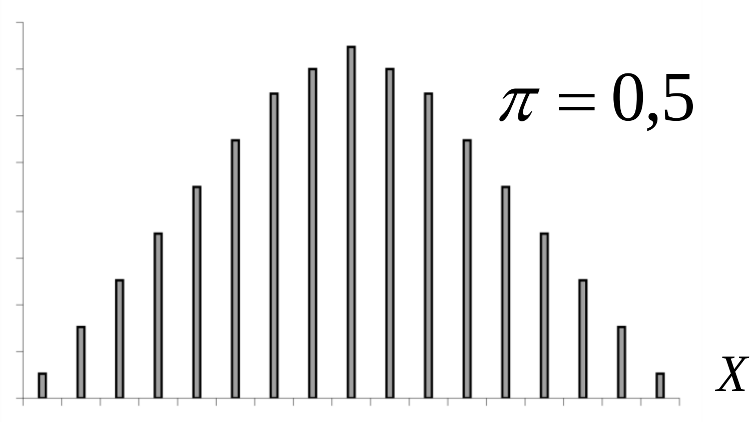
r znamená celkový počet nastání jevu v n nezávislých experimentech

r : 0 …… n

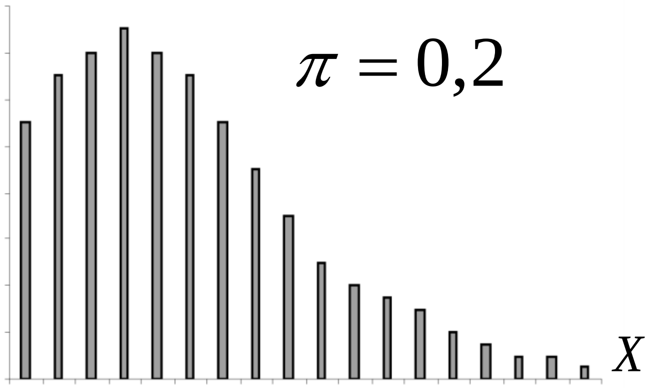
p ~ π .. jediný parametr binomického rozložení

p .... relativní četnost nastání jevu

p ………. určuje tvar distribuce



$$p = \frac{r}{n}$$



Binomická proměnná X

Binomické rozdělení jako model

Jev: narození chlapce $\Pi = 0{,}5$

n : rodina s 5 dětmi

r: 0,1,2,3,4,5 chlapců

$$\mathbf{P}(\mathbf{r}) = \binom{\mathbf{n}}{\mathbf{r}} \cdot \mathbf{p}^{\mathbf{r}} \cdot (\mathbf{1}-\mathbf{p})^{(\mathbf{n}-\mathbf{r})} = \frac{\mathbf{n}!}{\mathbf{r}!(\mathbf{n}-\mathbf{r})!} \cdot \mathbf{p}^{\mathbf{r}} \cdot \mathbf{q}^{(\mathbf{n}-\mathbf{r})}$$

r = 0 : $\frac{5!}{(0!\ 5!)} \cdot (0{,}5)^0 \cdot (0{,}5)^5 = 0{,}031$

r = 1 : $\frac{5!}{(1!\ 4!)} \cdot (0{,}5)^1 \cdot (0{,}5)^4 = 0{,}15625$

r = 2: P(r) = 0,3125

r = 3: P(r) = 0,3125

r = 4: P(r) = 0,15625

r = 5: P(r) = 0,031

X: Binomická proměnná

Střed rozložení:

$$\mathrm{E}(x) = n \cdot p$$

Rozptyl:

$$\mathrm{D}(x) = n \cdot p \cdot (1 - p)$$

Příklad: n = 100 respondentů
r = 20 má symptom

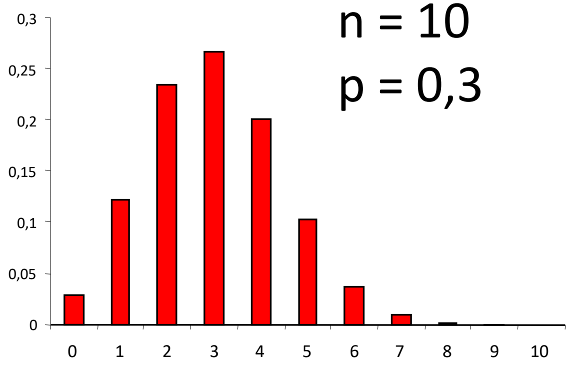
$$\mathrm{E}(x) = n \cdot p = 20$$

je střed rozložení
a nejpravděpodobnější hodnota

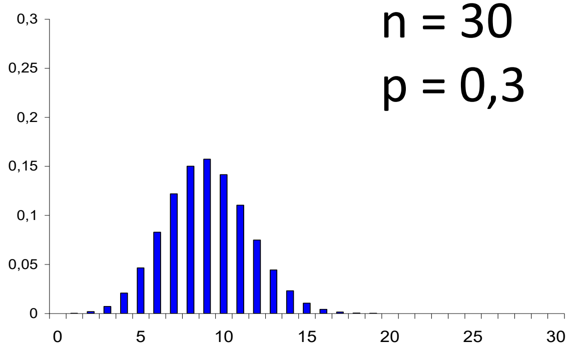
Binomické rozdělení jako model

$$P(x=r)=\frac{n!}{r!(n-r)!}\cdot p^{r}\cdot q^{(n-r)}$$

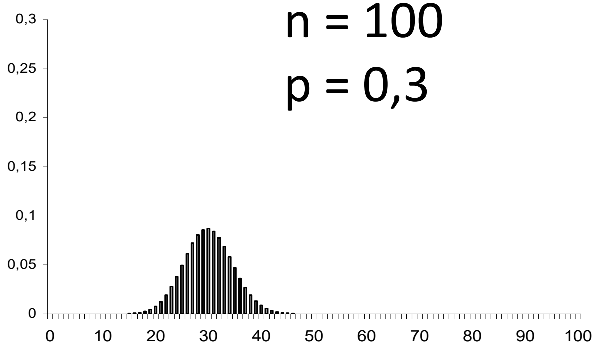
$q = 1 - p$



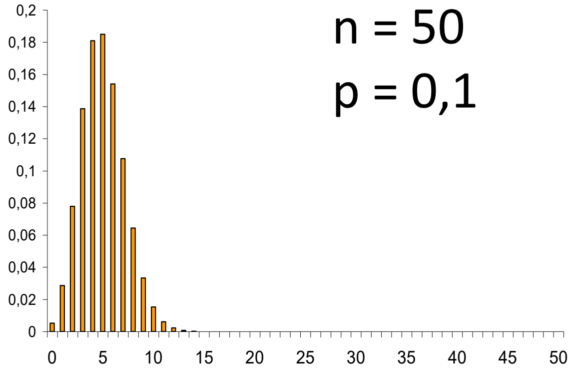
n = 10
p = 0,3



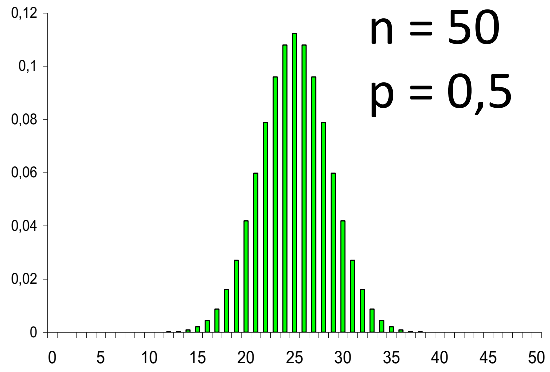
n = 30
p = 0,3



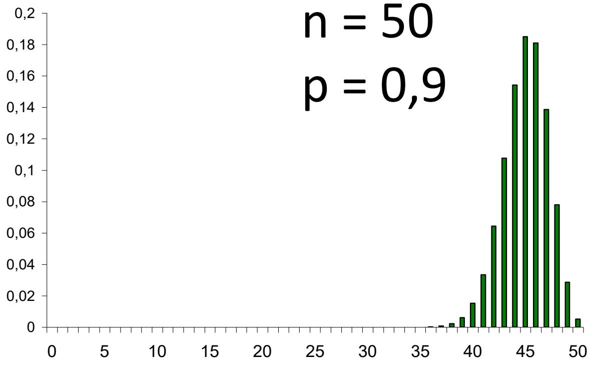
n = 100
p = 0,3



n = 50
p = 0,1



n = 50
p = 0,5

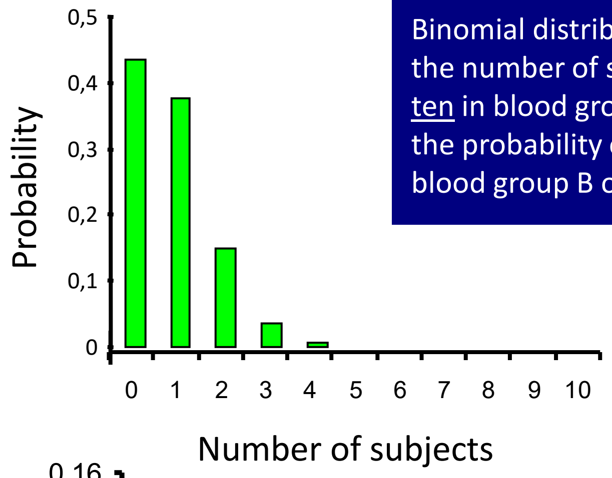


n = 50
p = 0,9

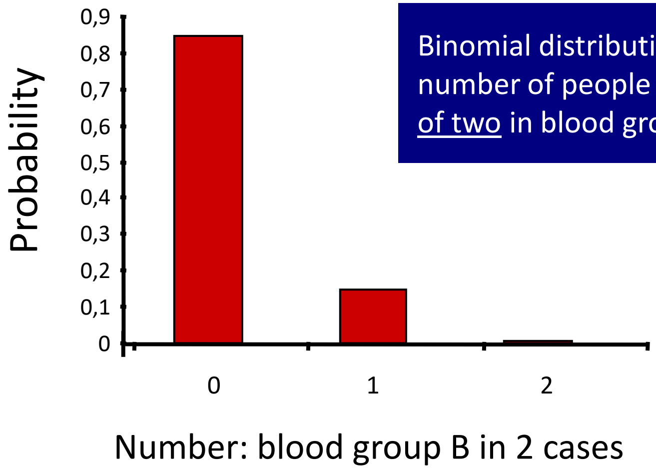
Aplikace binomického rozdělení

- Výskyt krevní skupiny B v určité populaci: p = 0,08

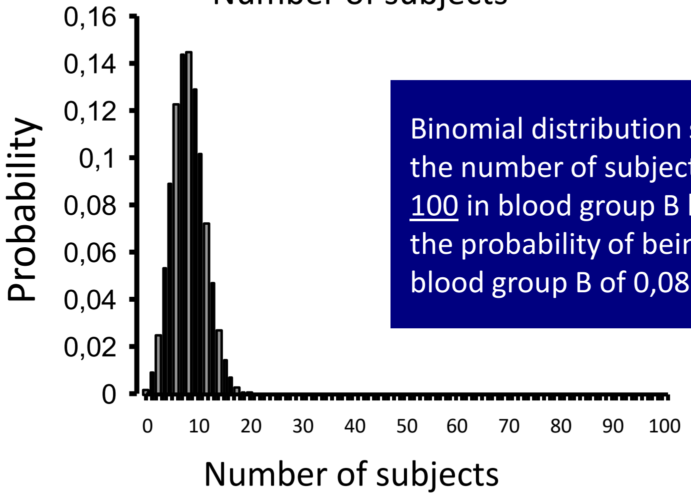
		Number in blood group B	Probability
B	B	2	0,0064
not B	B	1	0,0736
B	not B	1	0,0736
not B	not B	0	0,8464



Binomial distribution showing the number of subjects out of ten in blood group B based on the probability of being in in blood group B of 0,08.



Binomial distribution of number of people out of two in blood group B

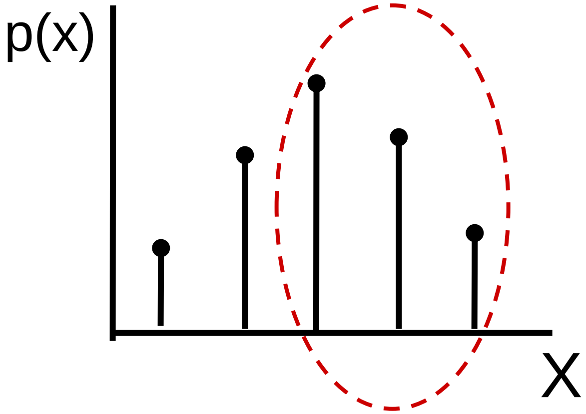


Binomial distribution showing the number of subjects out of 100 in blood group B based on the probability of being in in blood group B of 0,08.

Aplikace binomického rozdělení

- Populace: 60% jedinců má zvýšenou hladinu cholesterolu; výběr: 5 lidí
- Kolik lidí má ve výběru vyšší hladinu cholesterolu ?
  - n. p = 5 . 0,6 = 3 lidé   ~ E(x)
- Jaká je P, že právě 3 lidé budou mít vyšší hladinu cholesterolu ? ~   Tzn. Výběr přesně odpovídá dané populaci ?

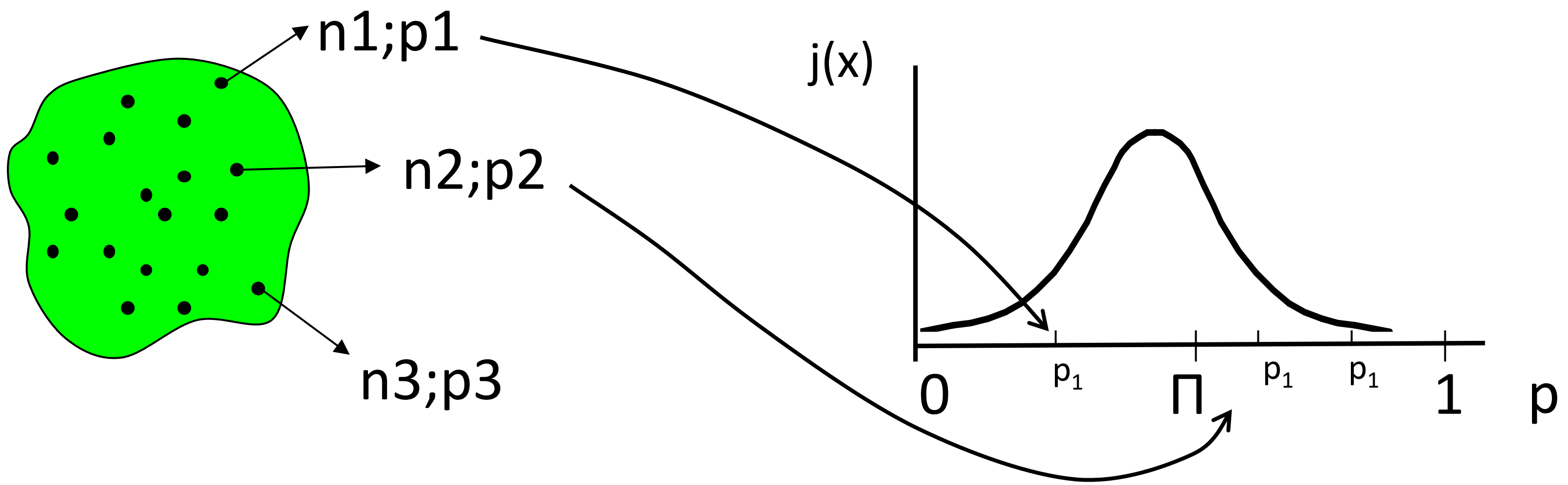
$$P_{(3)} = \frac{5!}{3!(5-3)!} \cdot (0,6)^3 \cdot (0,4)^2 = 0,346 = 35\%$$



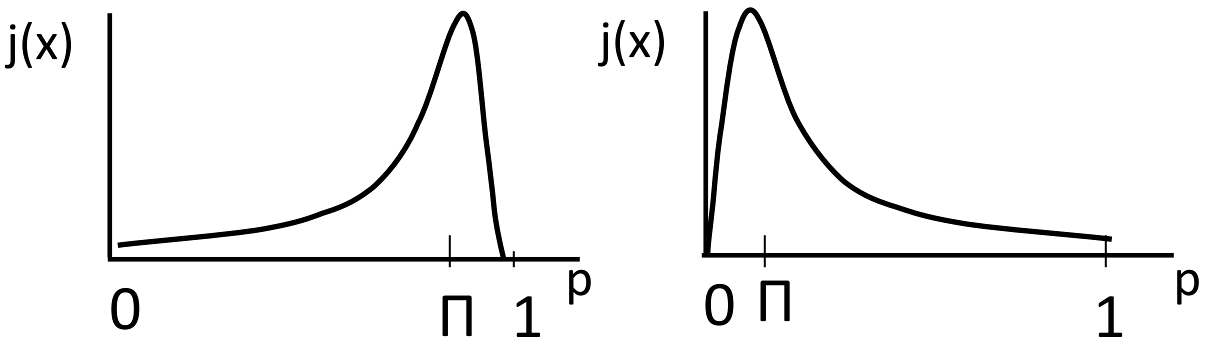
- Jaká je P, že většina jedinců (tedy minimálně 3) má vyšší hladinu cholesterolu ? ~ Tzn. výběr alespoň obecně odpovídá zkoumané populaci ?
  - P(X > 3) = P(3) + P(4) + P (5) = 0,346 + 0,259 + 0,078 = 68 %

Odhad parametru Π binomického rozdělení

- Při vícenásobném odhadu se parametr Π chová jako normálně rozložen



U malých nebo velkých hodnot p (Π) je však předpoklad normality omezen



Odhad parametru Π binomického rozložení

$$\pi \approx \hat{p}\,; \quad \hat{p} = r/n$$

1) Bodový

$$\hat{p}; \quad s_{\bar{p}}^2 = \frac{\hat{p}(1-\hat{p})}{n-1}$$

2) Intervalový – aproximace

$$\hat{p} - Z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}} \le \pi \le \hat{p} + Z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}}$$

$$\pi: \quad \hat{p} \pm Z_{1-\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n-1}}$$

Odhad parametru Π binomického rozložení: příklad I

X: % jedinců s daným znakem

n = 100 jedinců

r = 60; $\hat{p} = 0{,}6$

$s_{\hat{p}} = 0{,}049$

Interval spolehlivosti : 95 %

$Z_{0{,}975} = 1{,}96$

$$0{,}6 - 1{,}96 \cdot 0{,}049 \leq \pi \leq 0{,}6 + 1{,}96 \cdot 0{,}049$$

$$\boxed{0{,}504 \leq \pi \leq 0{,}697}$$

$$\Downarrow$$

$$P(0{,}504 \leq \pi \leq 0{,}697) \geq 0{,}95$$

Odhad parametru Π binomického rozložení

- Intervalový odhad bez aproximací na normální rozložení

$$L_1 = \frac{r}{r + (n - r + 1) \cdot F_{\alpha/2}^{(\nu_1;\nu_2)}}$$

spodní limit intervalu

$$\nu_1 = 2(n - r + 1); \quad \nu_2 = 2r$$

$$L_2 = \frac{(r+1) \cdot F_{\alpha/2}^{(\nu_1';\nu_2')}}{n - r + (r+1) \cdot F_{\alpha/2}^{(\nu_1';\nu_2')}}$$

horní limit intervalu

$$\nu_1' = 2(r+1) = \nu_2 + 2$$
$$\nu_2' = 2(n-r) = \nu_1 - 2$$

$$P(L_1 \leq \pi \leq L_2) \geq 1 - \alpha$$

Odhad parametru Π binomického rozložení: příklad II

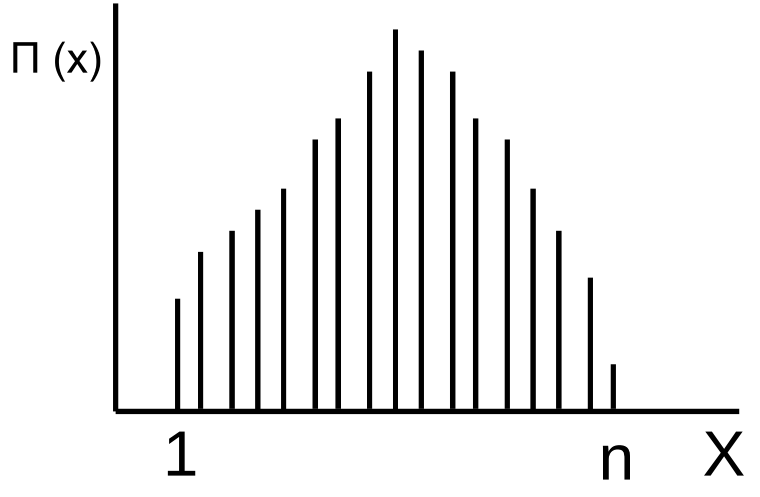
Náhodný vzorek n = 200 jedinců.
Zjištěno pouze r = 4 jedinci bez určitého znaku.

$$\hat{p} = 4/200 = \underline{\underline{0{,}02}}$$

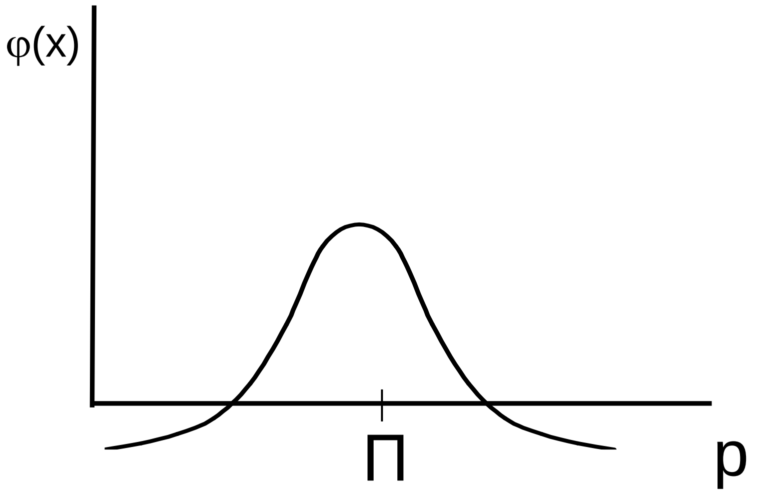
95% interval spolehlivosti = ?

Spodní hranice	Horní hranice
$\nu_1 = 2(n-r+1) = 2(200-4+1) = 394$ $\nu_2 = 2r = 2\cdot 4 = 8$ $F^{(394;8)}_{1-\alpha/2} = \underline{\underline{3{,}67}}$ $L_1 = \dfrac{4}{4+(200-4+1)\cdot 3{,}67} = \underline{\underline{0{,}0055}}$	$\nu'_1 = 2(r+1) = 10$ $\nu'_2 = 2(n-r) = 2(200-4) = 392$ $F^{(10;392)}_{1-\alpha/2} = \underline{\underline{2{,}08}}$ $L_2 = \dfrac{(4+1)\cdot 2{,}08}{200-4+(4+1)\cdot 2{,}08} = \underline{\underline{0{,}051}}$

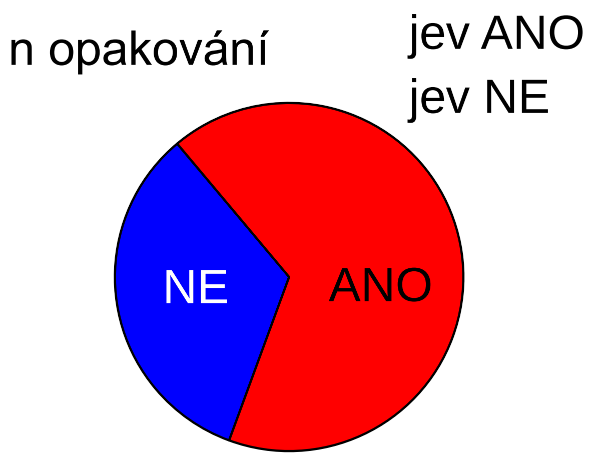
Binomické rozložení v datech: vizualizace



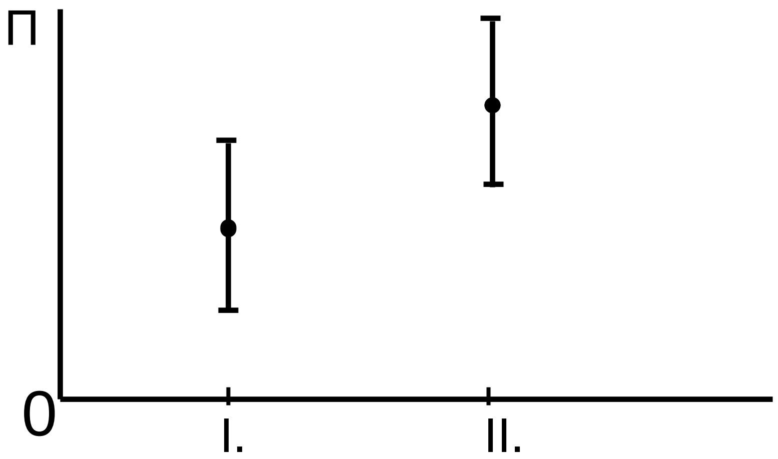
Pravděpodobnost výskytu hodnot X



Modelové rozložení odhadovaného parametru



Binární podstata původních hodnot



Interval spolehlivosti pro Π

Statistické testování binomických dat

I. Liší se odhad p od předpokládané hodnoty P ?

II. Liší se dva nebo více odhadů p ?

- závislé odhady -

- nezávislé odhady -

III. Je výskyt kategorií dvou jevů nezávislý ?

IV. Hodnocení relativního rizika z výskytu určitého jevu v rámci skupiny lidí

Jednovýběrový binomický test

H_0	*H_A*	*Testová statistika*	*Interval spolehlivosti*
p ≤ Π	**p > Π**	**z**	$z > z_{1-\alpha}$
p ≥ Π	**p < Π**	**z**	$z < z_{\alpha}$
p = Π	**p ≠ Π**	**z**	½z½ $> z_{1-\alpha/2}$

$$Z = \frac{n \cdot \hat{p} - n \cdot \pi}{\sqrt{n \cdot \hat{p}(1-\hat{p})}} \cong \frac{\left| n \cdot \hat{p} - n \cdot \pi \right| - 0{,}5}{\sqrt{n \cdot \hat{p}(1-\hat{p})}}$$

Korekce na kontinuitu

H_0	*H_A*	*Testová statistika*	*Interval spolehlivosti*
p ≤ Π	**p > Π**	$L_1 = \frac{(r+1)F_{\alpha,v_1',v_2'}}{n-r+(r+1)F_{\alpha,v_1',v_2'}}$	$p = r/n > L_1$
p ≥ Π	**p < Π**	$L_2 = \frac{r}{r+(n-r+1)F_{\alpha,v_1',v_2'}}$	$p < L_2$
p = Π	**p ≠ Π**	$L_1; L_2 \ (F_{\alpha/2}; F_{1-\alpha/2})$	$p < L_2 \ \vee \ p > L_1$

Test π ? p: Příklad 1

- Stromy s pozměněným tvarem koruny
  - n = 9 000 jedinců
  - r = 2 250 změněných jedinců
- Jak je pravděpodobná změna u až 1/3 jedinců?

$$Z = \frac{n \cdot p - n \cdot \pi}{\sqrt{p(1-p) \cdot n}} = \frac{2250 - 3000}{\sqrt{0{,}25 \cdot 0{,}75 \cdot 9000}} = \underline{\underline{-18{,}26}}$$

$\alpha = 5\ \%$; $Z_{1-\alpha/2} = 1{,}96$; $Z_{1-\alpha} = 1{,}645$

$Z > Z_{1-\alpha/2}$ ………zamítáme H_0: $p < 0{,}01$

- 95 % Interval spolehlivosti … p: (0,241; 0,258)

Test π ? p: Příklad 2

- Pravděpodobnost narození chlapce je asi 1/2.
- Máte zhodnotit výsledky průzkumu populace, která žije v silně poškozeném životním prostředí.
- Průzkum se týká 1000 náhodně vybraných rodin a zjištěný podíl narozených chlapců je 0.41.
  Jaké jsou vaše závěry o této populaci (zda se rodí stejný podíl chlapců jako v běžné populaci?)
  Jak se váš odhad zpřesní, když použijete vzorek n = 10 000 rodin při zachování odhadu p = 0.41?

Test π ? p: Příklad 2

- Použijeme jednovýběrový binomický test s nulovou hypotézou H0: p=π, hladina významnosti α=0,05
- Testová statistika: a příslušný kvantil:
- Protože nulovou hypotézu ???
- Interval spolehlivosti:
- Pokud použijeme n=10 000, bude int. spolehlivosti ???:

Test π ? p: Příklad 2

- Použijeme jednovýběrový binomický test s nulovou hypotézou H0: p=π, hladina významnosti α=0,05
- Testová statistika: $Z = \frac{n \cdot \hat{p} - n \cdot \pi}{\sqrt{n \cdot \hat{p}(1-\hat{p})}} = \frac{1000 \cdot 0{,}41 - 1000 \cdot 0{,}5}{\sqrt{1000 \cdot 0{,}41 \cdot 0{,}59}} = -5{,}79$ a příslušný kvantil: $Z_{1-\frac{\alpha}{2}} = Z_{0{,}975} = 1{,}96$

- Protože $|Z| > Z_{0{,}975}$ nulovou hypotézu zamítáme. Chlapci se ve zkoumavé populaci nerodí s pravděpodobností 0,5.
- Interval spolehlivosti: $\pi: \ \hat{p} \pm Z_{1-\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n-1}} = 0{,}4 \pm Z_{0{,}975} \cdot 0{,}046 = 0{,}41 \pm 1{,}96 \cdot 0{,}016 = 0{,}41 \pm 0{,}03$

- Pokud použijeme n=10 000, bude int. spolehlivosti užší:

$$\pi: \ \hat{p} \pm Z_{1-\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n-1}} = 0{,}41 \pm 1{,}96 \cdot 0{,}005 = 0{,}41 \pm 0{,}01$$

Test π ? p: Příklad 3

- Příklad testu bez aproximace na normální rozložení

12 jedinců bylo zkoumáno pro výskyt určitého znaku,
10 jedinců znak nemělo

? Jak hodně se tento výsledek liší od výsledku 6 - 6: tedy od situace, kdy polovina jedinců znak má?

a) Využití distribuční funkce

r	0	1	2	3	4	5	6	7	8	9	10	11	12
P(r)	0,00024	0,00293	0,01611	0,05371	0,12085	0,19335	0,22559	0,19336	0,12085	0,05371	0,01611	0,00293	0,00024

P (r ≥ 10) = 0,01611 + 0,00393 + 0,00024 = 0,01928

H_0: p = 0,5 je tedy značně nepravděpodobná

b) Pozorované $\hat{p} = {}^{10}\!/_{12} = 0{,}833$ překročilo horní limit 95 % intervalu spolehlivosti pro p:

$$p = 0{,}5: \; L_2 = \frac{(6+1)\cdot 2{,}64}{12-6+(6+1)\cdot 2{,}64} = \underline{\underline{0{,}755}}$$

Dvouvýběrový binomický test (p1 ? p2)

$$Z = \frac{\left|\hat{p}_1 - \hat{p}_2\right|}{\sqrt{\frac{\overline{p}\left(1-\overline{p}\right)}{n_1} + \frac{\overline{p}\left(1-\overline{p}\right)}{n_2}}}$$

$$\overline{p} = \frac{n_1 \cdot \overline{p}_1 + n_2 \cdot \overline{p}_2}{n_1 + n_2}$$

$$\left(\hat{p}_1 - \hat{p}_2\right) \pm Z_{\left(1-\alpha/2\right)} \cdot \sqrt{\frac{\overline{p}\left(1-\overline{p}\right)}{n_1} + \frac{\overline{p}\left(1-\overline{p}\right)}{n_2}}$$

Dvouvýběrový binomický test (p1 ? p2)

- Tento příklad je ukázkou testování rozdílů mezi dvěma binomickými populacemi (tedy srovnání dvou odhadů parametru p).
- Celkem 49 pokusných myší bylo použito k testování toxického preparátu během dvouměsíční kultivace. Následující tabulka obsahuje původní data zároveň s testem nulové hypotézy: Podíl přežívajících jedinců je u zasažené populace stejný.

	Alive	Dead	Total	Proportion alive	Proportion dead
Treated	15	9	24	$\hat{p}_1 = 0{,}625$	$\hat{q}_1 = 0{,}375$
Not Treated	10	15	25	$\hat{p}_2 = 0{,}400$	$\hat{q}_2 = 0{,}600$
Total	25	24	49	$\hat{p} = 0{,}510$	$\hat{q} = 0{,}490$

$$Z = \frac{0{,}625 - 0{,}400}{\sqrt{\frac{(0{,}510)\,(0{,}490)}{24} + \frac{(0{,}510)\,(0{,}490)}{25}}} = \frac{0{,}225}{\sqrt{0{,}010413 + 0{,}009996}} = 1{,}573$$

Z0,05(2) = t0,05(2) = 1,96

Nezamítáme H_0: 0,10 < P < 0,20

S korekcí na kontinuitu:

$$Z = \frac{\frac{15-0{,}5}{24} - \frac{10+0{,}5}{25}}{0{,}143} = \frac{0{,}604 - 0{,}420}{0{,}143} = 1{,}287$$

Z0,05(2) = t0,05(2) = 1,96

Nezamítáme H_0: 0,10 < P < 0,20