Konstrukce intervalu spolehlivosti pro průměr (či rozdíl průměrů)

a) průměr – Terapeut zkouší efektivitu nového přístupu k terapii nevhodného chování u dětí. Vybere si malý reprezentativní vzorek dětí s určitým druhem nevhodného chování (např. závažné narušování výuky) a týdenním pozorováním u nich stanoví frekvenci nevhodného chování. Poté proběhne terapie a pak opět týdenním pozorováním stanoví frekvenci nevhodného chování. Nakonec odečtením zjistí rozdíl mezi frekvencí před a po terapii.

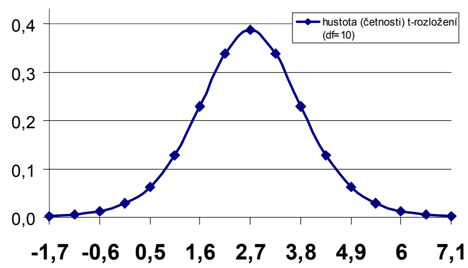
Před	po	rozdíl (před – po)
11	8	3
6	6	0
15	18	-3
22	14	8
8	7	1
9	10	-1
18	15	3
4	0	4
10	5	5
11	4	7

	N	min	max	*m*	s_m	*s*
VAR00001	10	-3	8	2,7	1,1	3,5

I. rozhodnout se, jakou pravděpodobnost chyby α jsme ochotni akceptovat: 5%, 1% ...

II. uvědomit si zda znám či neznám populační rozptyl – normální rozl. nebo *t*?

III. najít směrodatnou chybu průměru: $s / \sqrt{n}$



IV. najít *z* či *t* hodnotu odpovídající příslušnému kvantilu patřičného rozložení (2,5., 0,5.) – u *t*-rozl. si uvědomit *df*

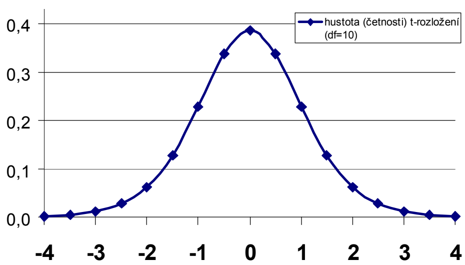
=NORM.S.INV(α/2) norm.s.inv(0,025)=-1,96

5%: norm.s.inv(0,025)=-1,96

1%: norm.s.inv(0,005)=-2,58

=T.INV(α/2; *df*)[1] např. t.inv(0,025;9)=-2,26

...nebo v tabulkách



V. Interval: výběrový průměr ± (kvantil × směrodatná chyba)

[1] Ve starších verzích Excelu a jiných tabulkových kalkulátorech (Open/LibreOffice, Gnumeric) se funkce jmenuje TINV (bez tečky) a funguje mírně odlišně. Zadaný percentil dělí dvěma, takže pro 95% interval zadáváme TINV(0,05;df) a její výsledek je vždy kladný. Jinak řečeno TINV(α;df)=T.INV(1–(α /2);df).

b) korelace – tatáž data, před= věk, počet dnů strávených v nemocnici
r = 0,8

I. Rozhodnout se, jakou pravděpodobnost chyby α jsme ochotni akceptovat:
5%, 1% ...

II. Výběrové rozložení korelace neznáme. Když se ale korelační koef. urč. způsobem transformuje, pak výběrové rozložení této transformované statistiky známe – jde o **normální** rozložení se $\boldsymbol{s_0=1/\sqrt{(n-3)}}$. Jde o Fisherovu Z-transformaci: $Z = 0{,}5 \ln((1+r)/(1-r))$ (v Excelu to počítá funkce FISHER(r))

fisher(0,8)=1,10

III. Spočítat směrodatnou chybu transformované korelace: $\boldsymbol{s_0=1/\sqrt{(n-3)}}$

IV. Najít příslušný kvantil normálního rozložení (2,5., 0,5.)

=NORM.S.INV(α/2)

5%: norm.s.inv(0,025)=-1,96

1%: norm.s.inv(0,005)=-2,58

V. Interval: výběrová korelace ± (kvantil × směrodatná chyba)

VI. Interval máme sestrojený v Z-transformovaných hodnotách. Musíme tedy ještě jeho meze transformovat zpět na koeficient r. K tomu slouží v Excelu FISHERINV.

fisherinv(0,36) = 0,35

fisherinv(1,84) = 0,95

C) Interval spolehlivosti a test hypotézy o relativních četnostech

p má přibližně normální rozložení s průměrem π a $\sigma_p = \sqrt{\dfrac{(1 - n/N).p.(1-p)}{n}}$

První činitel v čitateli zohledňuje, jak velkou část populace máme ve vzorku. Je-li populace vzhledem k vzorku obrovská(nekonečná), nemusíme ho používat. Pak σ_p=odmocnina(p(1-p)/n)

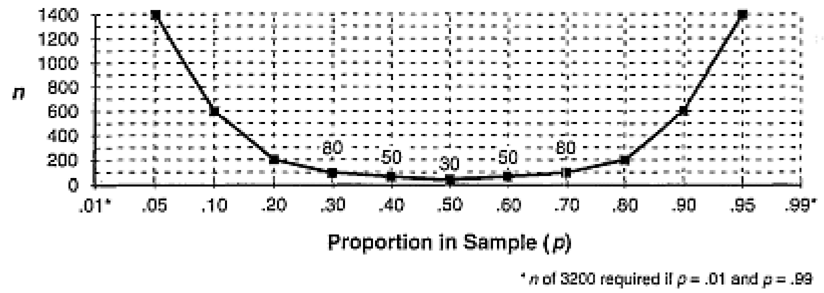


FIGURE 13.4 Minimum Sample Size (n) Needed for Use of the Normal Approximation When Setting Confidence Intervals for the Proportion (π) in the Population.

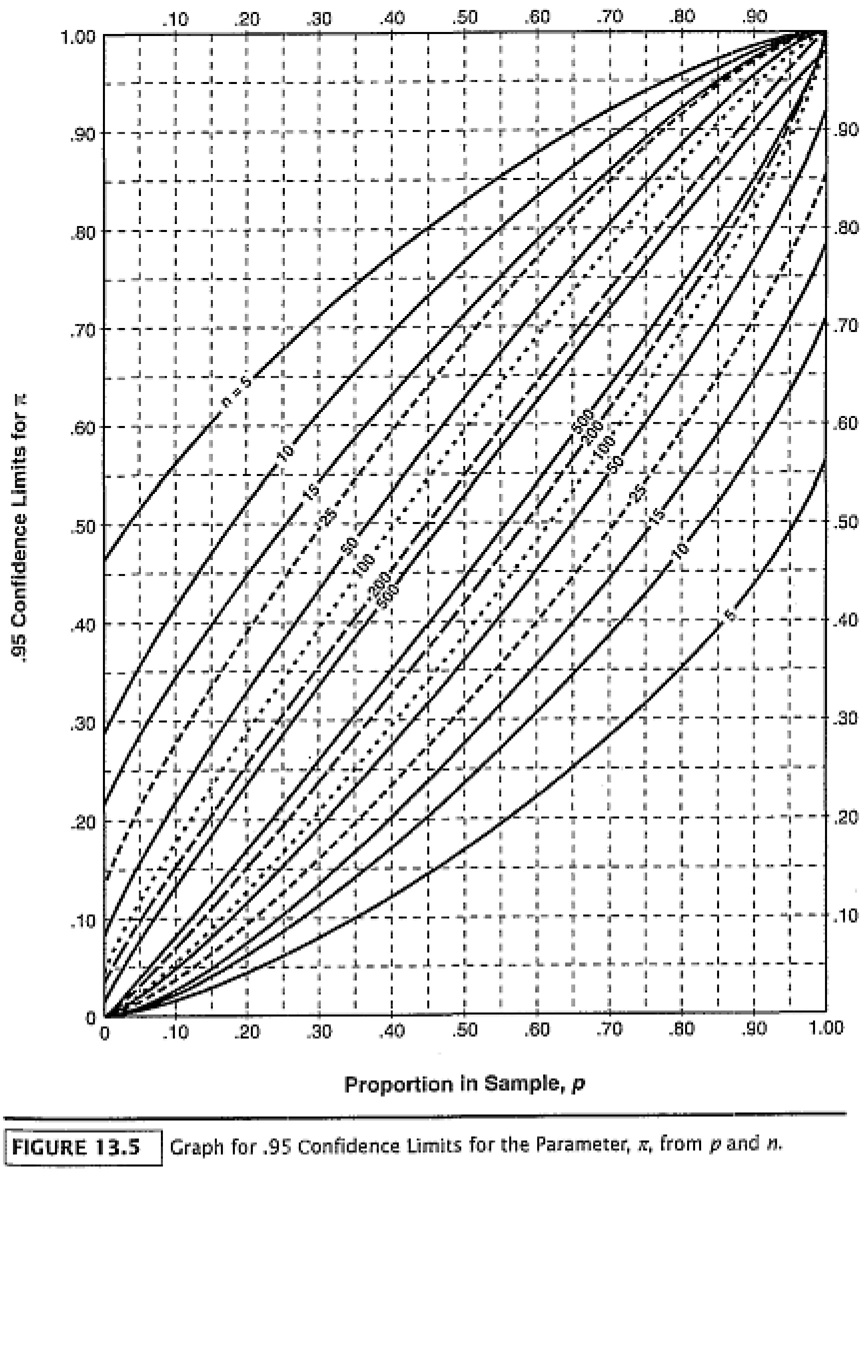


FIGURE 13.5 Graph for .95 Confidence Limits for the Parameter, π, from p and n.