Intervalový odhad rozptylu

I. Štatistická veličina X s rozdelením N(m,s^2) sa realizovala v n=15 pokusoch takto:

```
x = rand(1,15)*30;
```

Stredná hodnota ani rozptyl nie sú známe. Odhadneme rozptyl v (n-1) štatistike:

```
>> S2 = var(x)
                        S2 =    74.15253660045266
```

So spoľahlivosťou 0.9 chceme určiť interval, v ktorom by sa mala nachádzať skutočná hodnota rozptylu.

Na to budeme potrebovať 0.95 a 0.05 kvantily rozdelenia chí-kvadrát s parametrom k = n-1 = 14. Definujme najprv pomocnú nenormovanú funkciu fx:

```
>> fx=inline('x.^6.*exp(-x/2)');
```

Skúsme dosadiť väčšie číslo:

```
>> quad(fx,0,290,1e-11)
                        ans =      92160
```

Integrál od 0 po číslo 290 a rôzne iné čísla zhruba od 100 vyššie vychádza približne 92160. Pri takomto odhade kvantilov chi-kvadratu si treba dávať pozor na to, že od istého momentu narazíme na hranice presnosti matlabu a začneme dostávať nezmyselné výsledky. Integrál od 0 do 490 už napr. vychádza úplne scestne. Odklon od konvergencie znamená vstup na územie, kde sa snaha o presnosť stane svojou vlastnou obeťou.

Pre náročnejších nemožno neuviesť aj presné zistenie hodnoty integrálu. Integrujme fx v symbolickom móde:

```
>> t=sym('t'); y=int(fx(t))

y =  -2*t^6*exp(-1/2*t)-24*t^5*exp(-1/2*t)-240*t^4*exp(-1/2*t)-1920*t^3*exp(-1/2*t)-
     1520*t^2*exp(-1/2*t)-46080*t*exp(-1/2*t)-92160*exp(-1/2*t)
```

Pozorný pohľad na výpis ukazuje, že nájdená primitívna funkcia y v nekonečne konverguje k nule. To znamená, že náš hľadaný integrál je y(∞)-y(0) = -y(0). Stačí teda zistiť hodnotu v nule (dosadzovanie do symbolických výrazov je trochu nezvyklé):

```
>> t=0; ck= - subs(y)
        ck =  92160
```

Teraz môžeme definovať hustotu ch-kvadrátu fxi pre k=14:

```
>> fxi=inline('x.^6.*exp(-x/2)/92160');
```

****
Kvantily môžeme hľadať osvedčenou bisekciou. Skusmo zistíme, že

```
>> quad(fxi,0,25)
ans =    9.654326810988620e-001

>> quad(fxi,0,20)
ans =    8.698585751386978e-001
```

a teda 0.95-kvantil môžeme hľadať na intervale [20, 25]:

```
>> a=20;b=25; for k=1:50, s=(a+b)/2; if (quad(fxi,0,s,1e-13)<0.95) a=s; else b=s; end,end, a, b

a =    2.368479130484054e+001
b =    2.368479130484054e+001
```

Takže Xi14_095= 23.68479130484054.
Podobne zistíme, že

```
>> quad(fxi,0,6)
ans =     3.350852951126015e-002

>> quad(fxi,0,7)
ans =    6.528807342311321e-002
```

a teda 0,05-kvantil budeme hľadať medzi 6 a 7:

```
>> a=6;b=7; for k=1:50, s=(a+b)/2; if (quad(fxi,0,s,1e-13)<0.05) a=s; else b=s; end,end, a, b

a =    6.570631383789316e+000
b =    6.570631383789317e+000
```

Takže zhruba Xi14_005 = 6.5706313837893165e+000.

****

Skúsme urobiť ten istý výpočet s použitím primitívnej funkcie, ktorú sme už podaktorí (tí náročnejší) našli. Definujme si primitívnu funkciu ako inline objekt

```
>> Fx=inline('-1/46080*x^6*exp(-1/2*x)-1/3840*x^5*exp(-1/2*x)-1/384*x^4*exp(-1/2*x)-1/48*x^3*exp(-1/2*x)-1/8*x^2*exp(-1/2*x)-1/2*x*exp(-1/2*x)-exp(-1/2*x)+1')
```

!!! Prečo je na konci doplnené +1 ???

Počítajme obidva kvantily:

```
>> Fx95=inline('-1/46080*x^6*exp(-1/2*x)-1/3840*x^5*exp(-1/2*x)-1/384*x^4*exp(-1/2*x)-1/48*x^3*exp(-1/2*x)-1/8*x^2*exp(-1/2*x)-1/2*x*exp(-1/2*x)-exp(-1/2*x)+1-0.95')

>> fzero(Fx95,0)
ans =    2.368479130484058e+001
```

```
>>      Fx05=inline('-1/46080*x^6*exp(-1/2*x)-1/3840*x^5*exp(-1/2*x)-1/384*x^4*exp(-1/2*x)-1/48*x^3*exp(-1/2*x)-1/8*x^2*exp(-1/2*x)-1/2*x*exp(-1/2*x)-exp(-1/2*x)+1-0.05')

>> fzero(Fx05,0)
                      ans =    6.570631383789346e+000
```

Až na posledné jedno či dve miesta výsledok zodpovedá výpočtu bisekciou, dá sa však oprávnene predpokladať, že tento priamy výpočet je presnejší.

****

Môžeme pristúpiť k vyčísleniu hraníc hľadaného intervalu pre rozptyl.

```
>> HL= 14*S2/ Xi14_095, HP= 14*S2/ Xi14_005

HL =    4.383132203970737e+001
HP =    1.579963001723772e+002
```

So spoľahlivosťou 0.9 sa neznámy rozptyl s^2 veličiny X nachádza na intervale [43.831322, 157.9963]. Blesková kontrola správnosti – obsahuje tento interval odhadnutú hodnotu S2?

Úlohy:

Zopakujte výpočet pre spoľahlivosť 0.95 a 0.99, prípadne s iným výberovým vektorom x.

II. Predpokladajme teraz, že stredná hodnota m veličiny X je známa a m = 20 (z nameraných hodnôt vychádza odhad 17.82445246758215).

Odhadneme S v n-štatistike pri známom m:

```
>> m=20; Sm2=(x-m)*(x-m)'/15
Sm2 =    73.94204122623188
```

Počítajme potrebné kvantily Chí-kvadrátu pre k=15 (normovaciu konštantu zistíme podobne ako v predošlom prípade):

```
>> fxi=inline('x.^6.5.*exp(-x/2)/338733.2118922602');
```

0,95-kvantil je medzi 24 a 25 (zistíme skusmo) a preto:

```
>> a=24;b=25; for k=1:50, s=(a+b)/2; if (quad(fxi,0,s,1e-13)<0.95) a=s; else b=s; end,end,a,b

a =    2.499579013972802e+001
b =    2.499579013972802e+001
```

Takže Chi15_95=24.99579013972802.

0.05-kvantil budeme hľadať medzi 7 a 8:

```
>> a=7;b=8; for k=1:50, s=(a+b)/2; if (quad(fxi,0,s,1e-13)<0.05) a=s; else b=s; end,end, a, b

a =    7.260943927669224
b =    7.260943927669225
```

Chi15_05 = 7.2609439276692245.

Úloha pre náročnejších: Zopakujte výpočet oboch kvantilov priamo z primitívnej funkcie (ktorú musíte nájsť).

Hranice intervalu sú:

```
>> HL= 15*Sm2/Chi15_95 , HP= 15*Sm2/ Chi15_05

HL =    44.37268155747418
HP =    152.7529534935414
```

a teda hľadaný rozptyl sa nachádza v intervale [44.37, 152.75] so spoľahlivosťou 0.9.

Úlohy:

Zopakujte výpočet pre spoľahlivosť 0.95 a 0.99, a pre m=15 a m=22.