Rozdelenie chí-kvadrát

Hodnoty distribučnej funkcie rozdelenia chí-kvadrát môžeme získať v Matlabe pomerne jednoducho aj bez štatistického balíka. Distribučná funkcia rozdelenia chí-kvadrát so stupňom voľnosti k je

$$F(k,x)= \int_{[0,x/2]} t^{k/2-1} * e^{-t}\, dt \;/\; \int_{[0,\infty]} t^{k/2-1} * e^{-t}\, dt ,$$

čo sa zvykne stručnejšie značiť

$$F(k,x)= \boldsymbol{\gamma}(k/2,x/2) \,/\, \boldsymbol{\Gamma}(x/2) .$$

Úloha: Nájdite v skriptách alebo v poznámkach z prednášky definíciu funkcie hustoty rozdelenia chí-kvadrát a dokážte, že jej integrovaním dostaneme vyššie uvedený tvar. Na riešenie tejto úlohy treba ovládať integrovanie a osobitne metódu substitúcie.

V Matlabe môžeme realizovať výpočet hodnôt distribučnej funkcie prostredníctvom zabudovanej funkcie `gammainc` (pozor na poradie argumentov!):

F(k,x) ← `gammainc(x/2,k/2)`

Poznámka:
Pre dosť veľké *n* sa rozdelenie chí-kvadrát blíži ku Gaussovmu N(n,2n).
Úloha:
Porovnajte pre rôzne n rozdelenie chí-kvadrát a N(n,2n). Pre aké n možno hovoriť o uspokojivej aproximácii? Kreslite cez seba grafy oboch rozdelení pre rôzne n.

Intervalový odhad rozptylu

I. Náhodná 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)
```

So spoľahlivosťou 0.9 chceme určiť interval, v ktorom by sa mala nachádzať skutočná hodnota rozptylu. Na to potrebujeme 0.95 a 0.05 kvantily rozdelenia chí-kvadrát s parametrom/stupňom voľnosti k = n-1 = 14. To zrealizujeme v dvoch krokoch:

a) Najprv definujeme v Matlabe distribučnú funkciu pre chí-kvadrát stupňa 14:

```
>> F=inline('gammainc(x/2,7)')
```

Potom vyrobíme dve odvodené funkcie – miesto hľadania hodnôt F(x)=0.05 a 0.95 budeme zisťovať, kedy sú F(x)-0.05 a F(x)-0.95 rovné nule. V princípe je to tá istá úloha, rozdiel je v schopnosti Matlabu ju riešiť.

```
>> Fr=inline('gammainc(x/2,7) – 0.95')
>> Fl=inline('gammainc(x/2,7) – 0.05')
```

Kvantily nájdeme príkazom `fzero` (hľadá nulové body funkcií). Príkaz potrebuje poradiť aspoň orientačne, kde by mal hľadať riešenie. Preto najprv skusmo dosadíme zopár čísel do F a orientačne zistíme, že 0.95-kvantil by mohol byť okolo 20ky a 0.05-kvantil niekde okolo 7-ky. Teraz už môžeme spustiť hľadanie:

```
>> X95=fzero(Fr,20)
     X95 =  23.68479130484058
>> X05=fzero(Fl,7)
     X05 =   6.57063138378934
```

Môžeme pristúpiť k vyčísleniu hraníc hľadaného intervalu pre rozptyl.

```
>> HL= 14*S2/ X95, HP= 14*S2/ X05

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].

Úlohy:

1. Zopakujte výpočet pre spoľahlivosť 0.95 a 0.99, a to isté s iným výberovým vektorom x dĺžky 25.
2. Riešte opačný problém. Chceme, aby rozptyl nepresiahol hodnotu 120, tj. aby bol v intervale [?, 120]. S akou spoľahlivosťou môžeme počítať?

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
```

Úlohy:

1. Počítajte potrebné kvantily chí-kvadrátu pre k=15.
2. Nájdite hľadaný interval.
3. Zopakujte výpočet pre spoľahlivosť 0.95 a 0.99, a pre m=15 a m=22.