

Spojité pravdepodobnostné rozdelenia

Číselné charakteristiky

Príklad 1:

Je daná funkcia hustoty $f(x) = 2/(x+1)$ na intervale $[0, \exp(0.5)-1]$. V Matlabe si ju môžeme definovať ako inline objekt (kvôli ďalším výpočtom pred delenie musí ísť bodka):

```
>> f=inline('2./(x+1)');
```

Ubezpečme sa, že je definovaná správne a zistíme hodnotu integrálu na danom intervale. Na numerické integrovanie využijeme príkaz quad:

```
>> format long, d=exp(0.5)-1;  
>> quad(f,0,d)
```

```
ans =  
1.000000023810609e+000
```

Očakávali sme výsledok 1 a odpoveď Matlabu nie je ďaleko od očakávaní. Nepresnosť je spôsobená tým, že quad integruje štandardne s presnosťou $1e-6$. Túto presnosť môžeme zvýšiť štvrtým parametrom:

```
>> quad(f,0,d, 1e-12)
```

```
ans =  
1.0000000000000002e+000
```

```
>> quad(f,0,d, 1e-16)
```

```
ans =  
1
```

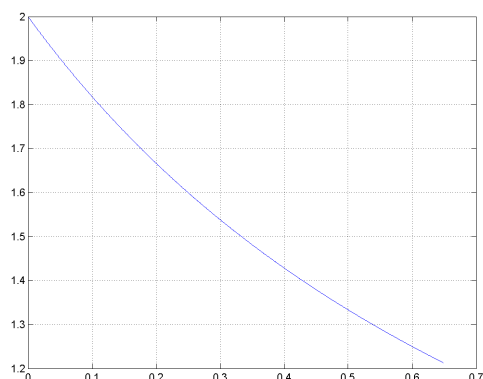
Nájďme distribučnú funkciu pravd. rozdelenia F . S integrovaním nám poradí Matlab a možnosti jeho symbolického módu. (Zeleným písmom sú uvedené časti, ktoré nebudú na testoch ani skúške vyžadované a predstavujú priestor na dobrovoľné rozširovanie si obzoru.)

```
>> s=sym('s');  
>> int(f(s))
```

```
ans =  
2*log(s+1)
```

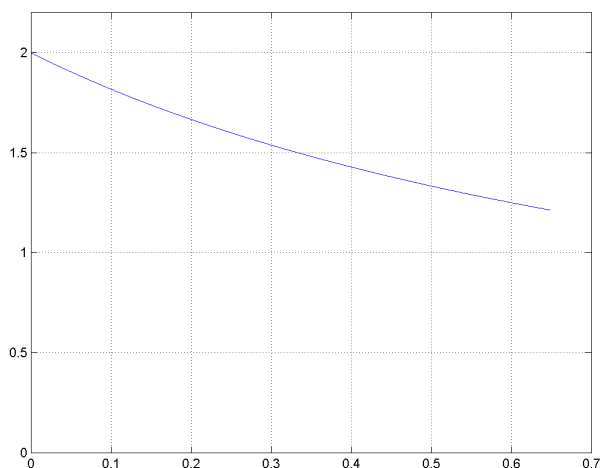
Vykreslíme teraz grafy funkcií f a F . Interval $[0, d]$ rozdelíme na 1000 častí príkazom linspace (daný interval rozdelí na n častí, zadávame však číslo $n+1$ ako počet deliacich bodov).

```
>> xd = linspace(0,d,1001);  
>> plot(xd,f(xd)), grid on
```



Nie je to celkom ono... chceme presne vidieť, ako vysoko je graf nad osou x. Upresníme teda, akú oblasť roviny má Matlab znázorniť:

```
>> plot(xd,f(xd)), axis([0, 0.7, 0, 2.2]), grid on
```



Úloha: Podobne nakreslite funkciu F (najprv si ju definujte ako inline objekt s premennou x).

Na výpočet kvantilov využijeme príkaz `fzero`, ktorým Matlab hľadá korene rovnice $f(x) = 0$. Násť kvantil q_a znamená násť x , pre ktoré platí $F(x) = a$, čo je to isté ako $F(x) - a = 0$. Nájďme najprv medián, teda 0.5-kvantil. Definujeme $Fq = F - 0.5$ a nájďme jej koreň:

```
>> Fq=inline('2*log(x+1)-0.5'); q05=fzero(Fq, 0)
```

```
q05 =  
2.840254166877414e-001
```

Príkaz `fzero` potrebuje „poradiť“, kde má začať hľadať koreň, preto sme zadali ako druhý vstup 0 (približný odhad hľadanej hodnoty). Treba sa ubezpečiť, že výsledok nie je „mimo“ prípustného intervalu.

Úloha: Nájďte kvartily funkcie F a 0.1, 0.9 a 0.95 -kvantily.

Príklad 2:

Je daná hustota $f(x,y) = 1-x+y$ pravdepodobnostného vektora (X,Y) , na ploche možných udalostí $[0,1] \times [0,1]$. Na teórii sme už zistili, ako vyzerajú f_x a f_y (hustota premenných X , Y samostatne).

Úloha 1: Zadajte f , f_x , f_y ako inline objekty. Ubezpečte sa, že f , f_x a f_y sú zadané správne.

Riešenie:

```
>> f=inline('1-x+y'); fx=inline('1.5-x'); fy=inline('0.5+y');  
>> i1=quad(fx,0,1); i2=quad(fy,0,1); i3=dblquad(f,0,1,0,1); [i1,i2,i3]
```

```
ans =  
    1    1    1
```

Na výpočet dvojitého integrálu sme použili príkaz `dblquad`, do ktorého sa vkladajú postupne hranice intervalov premennej x a y .

Úloha 2: Vypočítajte strednú hodnotu $E(X)$, $E(Y)$, rozptyl $\text{var}(X)$, $\text{var}(Y)$, kovarianciu a korelačný koeficient vektora (X,Y)

Riešenie: – *Riešte samostatne a až nakoniec si porovnajte výsledky s nasledujúcim postupom!*

```
>> fxe=inline('1.5*x-x.^2'); fye=inline('0.5*y+y.^2');  
>> Ex=quad(fxe, 0, 1); Ey=quad(fye, 0, 1); [Ex, Ey]
```

```
ans =  
    4.166666666666667e-001    5.833333333333334e-001
```

```
>> fxv=inline('1.5*x.^2-x.^3'); fyv=inline('0.5*y.^2+y.^3');  
>> varx=quad(fxv, 0, 1)-Ex^2; vary=quad(fyv, 0, 1)-Ey^2; [varx, vary]
```

```
ans =  
    7.638888888888888e-002    7.638888888888888e-002
```

```
>> fxy=inline('x.*y.*(1-x+y)'); EXY=dblquad(fxy,0,1,0,1); covxy=EXY-Ex*Ey
```

```
covxy =  
    6.9444444444444392e-003
```

```
>> cocfxy=covxy/(varx*vary)^0.5
```

```
cocfxy =  
    9.090909090909033e-002
```

Úloha 3: Na oblasti $[-1, 2] \times [-2, 4]$ je daná funkcia hustoty $f(x,y) = (2-x+y)/9$. Vypočítajte jednotlivé číselné charakteristiky pravdepodobnostného rozdelenia určeného funkciou f .

Príklad 3:

Nakoniec sa budeme venovať situácii, keď má oblasť nenulových hodnôt funkcie hustoty iný tvar ako obdĺžnik.

Nech oblasť Q je daná takto: $Q = \{(x,y) \in \mathbb{R}^2 ; x, y \geq 0 \wedge x+y \geq 2 \wedge x+y \leq 5\}$.

Funkcia hustoty bude: $f = 2/21$ pre $(x,y) \in Q$

a $f = 0$ inde.

Úloha: Načrtnite (na papier) oblasť Q .

Matlab nám nedáva priamu možnosť počítať integrál z funkcie na iných ako obdĺžnikových oblastiach. Musíme preto úlohu riešiť nepriamo – integrovať budeme na obdĺžniku $B = [0, 5] \times [0, 5]$ a funkciu f položíme rovnú nule všade na B/Q .¹ To nám v Matlabe umožňuje zapojenie boolovských výrazov do predpisu funkcie:

```
>> f=inline('(x+y<=5).*(x+y>=2)*2/21');
```

```
>> dblquad(f,0,5,0,5,1e-10)
```

```
ans =
```

```
9.999999913841786e-001
```

Na danej oblasti je hodnota integrálu 1, takže f je dobre definovaná funkcia hustoty.

Funkcie hustoty f_x a f_y (funkcia f_y je zhodná s f_x) premenných X , Y samostatne možno jednoducho zistiť z náčrtu na papieri (skúste!). Ide o „vetvičkové“ funkcie, čo opäť možno matlabu vysvetliť pomocou boolovských výrazov:

```
>> fx = inline('(2/7*(x<2)+(5-x).*(x>2)*2/21)');
```

```
>> quad(fx,0,5)
```

```
ans =
```

```
9.999974165013228e-001
```

Výpočet strednej hodnoty: (pre X aj Y rovnaká)

```
>> fxx=inline('(2/7*(x<2)+(5-x).*(x>2)*2/21).*x');
```

```
>> ex=quad(fxx,0,5,1e-11)
```

```
ex =
```

```
1.857142857162568e+000
```

Výpočet variancie: (pre X aj Y rovnaká)

```
>> fxxx=inline('(2/7*(x<2)+(5-x).*(x>2)*2/21).*x.*x');
```

```
>> vax=quad(fxxx,0,5,1e-11)
```

```
vax =
```

```
4.833333333372755e+000
```

¹ Do predpisu nedávame podmienky nezápornosti x , y . Budeme však na ne pamätať a pri výpočtoch sa budeme vyhýbať všetkým kvadrantom okrem prvého. V ďalších výpočtoch pre jednoduchosť nahradíme neostré rovnosti ostrými. Prečo si to môžeme dovoliť?

Výpočet kovariancie a korelačného koeficientu:

```
>> fxy=inline('((x+y<5).*(x+y>2)/10.5).*x.*y');
```

```
>> cox=dblquad(fxy,0,5,0,5,1e-11)
```

```
cox =  
    2.278797936574460e+000
```

```
>> cocof=cox/vax
```

```
cocof =  
    4.714754351494911e-001
```

Úloha: Pri výpočte korelačného koeficientu máme v menovateli niečo iné, ako sa uvádza v štandardnom vzorci. Prečo? (Je to naozaj iné?)