Cvičení z NSTP097

Pravděpodobnostní rozdělení, limitní věty

IX.

Úloha 1: Hustoty a distribuční funkce

R má zabudované funkce pro výpočet hustot, distribučních funkcí a kvantilových funkcí pro řadu běžných rozdělení. Taktéž umí generovat náhodné výběry z těchto rozdělení.

Zkusme si nakreslit pár hustot a distribučních funkcí. Například pro exponenciální rozdělení Exp(λ): `dexp(x,lambda)` počítá hustotu v bodě `x` nebo ve vektoru bodů `x`. Podobně `pexp(x,lambda)` počítá distribuční funkci v bodě `x` nebo ve vektoru bodů `x`.

Nakresleme si hustotu a distribuční funkci rozdělení Exp(4). Rozumné rozpětí pro osu x grafu bude od 0 do 2. Vytvoříme vektor 500 bodů od 0 do 2, v nichž budeme hustotu a distribuční funkci počítat:

```
x=seq(0,2,length=500)
```

Nyní spočítáme hustotu Exp(4) v těchto bodech:

```
hust=dexp(x,4)
```

a nakreslíme obrázek bodů `hust` proti `x`:

```
plot(x,hust,type="l")
```

[Pozor: `"l"` je malé písmeno ℓ, nikoli číslice 1]

Nyní zkuste nakreslit distribuční funkci Exp(4) pomocí funkce `pexp()`.

Nakonec vygenerujeme náhodný výběr z Exp(4). To dělá funkce `rexp()`. Řekněme, že chceme vygenerovat 25 nezávislých veličin:

```
nvyb=rexp(25,4)
```

První argument je počet požadovaných veličin, druhý argument je parametr exponenciálního rozdělení. Výsledek si v R vypíšeme zadáním jména objektu, který chceme vidět: `nvyb`. Spočtěte ještě průměr z vašeho náhodného výběru: `mean(nvyb)`. Jaká je střední hodnota rozdělení Exp(4)?

Zopakujte tento postup pro další rozdělení: N(0,1), N(-2,4), Γ(2,0.5), Γ(0.5,0.1), C(0,1). Hustoty, distribuční funkce, a náhodné výběry dostaneme takto:

N(μ,σ^2)	Hustota:	`dnorm(x,mi,sigma)` (Pozor! `sigma` je $\sqrt{\sigma^2}$!)
	D.f:	`pnorm(x,mi,sigma)`
	Náh. výběr:	`rnorm(n,mi,sigma)`
$\Gamma(a,p)$	Hustota:	`dgamma(x,rate=a,shape=p)`
	D.f:	`pgamma(x,rate=a,shape=p)`
	Náh. výběr:	`rgamma(n,rate=a,shape=p)`
C(a,b)	Hustota:	`dcauchy(x,a,b)`
	D.f:	`pcauchy(x,a,b)`
	Náh. výběr:	`rcauchy(n,a,b)`

Poznali byste od oka výběr z Cauchyova rozdělení od výběru z normálního rozdělení?

Porovnejme si hustoty normálního a Cauchyova rozdělení. Uděláme to tak, že si obojí spočítáme a nakreslíme do jednoho obrázku:

```
x=seq(-5,5,length=500)
hustc=dcauchy(x,0,1)
hustn=dnorm(x,0,1)
plot(x,hustc,type="l",ylim=1.2*range(c(hustc,hustn)),ylab="Hustota")
lines(x,hustn,lty=2)
legend("topleft",lty=c(1,2),legend=c("C(0,1)","N(0,1)"),inset=0.05)
```

Nakreslete si stejným způsobem obrázek porovnávající hustotu exponenciálního rozdělení se střední hodnotou 2 s gama rozdělením se střední hodnotou 2 a parametrem $p = 0.5$. [Musíte změnit rozsah bodů `x`, v nichž se hustoty počítají a kreslí.]

Úloha 2: Průměry ze vzrůstajícího počtu pozorování

Podívejme se, jak hodnoty průměru závisí na tom, kolik pozorování do průměru zahrneme. Nejprve si vezměme 50 výběrů o $n = 25$ pozorováních z rozdělení $\Gamma(0.5{,}2)$. Pozorování si uspořádáme do matice o 50 řádcích a 25 sloupcích:

```
n=25
nvyb.50=rgamma(50*n,rate=0.5,shape=2)
nvyb.50=matrix(nvyb.50,ncol=n)
```

Spočtěme si 50 průměrů z 50 výběrů (řádkové průměry matice `nvyb.50`):

```
prumery=rowMeans(nvyb.50)
```

Vypište si průměry. Spočtěte odhad rozptylu průměrů (`var(prumery)`). Jaký je skutečný rozptyl průměrů z 25 pozorování s rozdělením $\Gamma(0.5{,}2)$? Udělejte si histogram průměrů:

```
hist(prumery,main=paste("Pocet pozorovani n =",n),xlim=c(2,6))
```

Zopakujte si úlohu pro rozsahy výběrů $n = 250$ a $n = 2500$ z téhož rozdělení. Jak se chovají průměry při rostoucím počtu pozorování? Která věta teoreticky zdůvodňuje výsledky, které vidíte?