Cvičení 2.

Úkol 1.: Použijte funkci clv.m pro generování 200 čísel z rozložení N(0,1). Pomocí funkce kstest.m otestujte na hladině významnosti 0,05, že vygenerovaná data se skutečně řídí rozložením N(0,1).

Návod:

Vygenerujeme n = 200 realizací z N(0,1):

```
n=200;
realizace=clv(…,…,…);
```

Vypočteme hodnoty distribuční funkce rozložení N(0,1) v bodech vygenerovaných realizací:

```
Fi=normcdf(realizace,…,…);
```

Zavoláme funkci kstest:

`[h,p,ksstat,cv]=kstest(realizace,[realizace,Fi])` (v případě, že testujeme hypotézu o rozložení N(0,1), stačí zadat jen vektor realizací)

Vstupní parametry:

`realizace` … sloupcový vektor realizací

`[realizace,Fi]` … matice nx2 obsahující vektor realizací a vektor hodnot distribuční funkce rozložení N(0,1)

Výstupní parametry:

`h` … nabývá hodnoty 0, když nulovou hypotézu nezamítáme na hladině významnosti 0,05 a hodnoty 1, když zamítáme na hladině významnosti 0,05

`p` … p-hodnota

`ksstat` … hodnota testové statistiky

`cv` … kritická hodnota

V tomto konkrétním případě jsme dostali následující výstupy:

```
h =
     0
p =
    0.6009
ksstat =
    0.0533
cv =
    0.0952
```

Závěr: Na hladině významnosti 0,05 ………………… hypotézu, že vygenerovaná data pocházejí z rozložení N(0,1), protože p-hodnota je …….., což je …………… než 0,05. Vidíme rovněž, že testová statistika nabývá hodnoty ………, což je ….. …. Než kritická hodnota ………….

Nepovinný úkol: Místo z rozložení N(0,1) generujte data z rozložení N(2,4) a pro ověření normality opět použijte funkci kstest.m.

Návod:

```
realizace=clv(…,…,…);
```

Pro výpočet hodnot distribuční funkce rozložení N(2,4) použijeme příkaz

```
Fi=normcdf(realizace,2,2);
```

Úkol 2.: Pro stejný úkol jako v bodě 1 použijte funkce clv_polynom.m, BM_transformace.m a funkci normrnd.m (je součástí statistického toolboxu).

Návod:

```
n=200;
```
Použití funkce clv_polynom.m
```
realizace=clv_polynom(…,…,…);
Fi=normcdf(realizace,…,…);
[h,p,ksstat,cv]=kstest(…,[…,…])
```

V tomto konkrétním případě jsme dostali následující výstupy:
```
h =
     1
p =
    0.0497
ksstat =
    0.0952
cv =
    0.0952
```
Závěr: Na hladině významnosti 0,05 ………………… hypotézu, že vygenerovaná data pocházejí z rozložení N(0,1), protože p-hodnota je …….., což je …………… než 0,05.

Použití funkce BM_transformace.m
```
realizace=BM_transformace(…,…,…);
Fi=normcdf(realizace,…,…);
[h,p,ksstat,cv]=kstest(…,[…,…])
```

V tomto konkrétním případě jsme dostali následující výstupy:
```
h =
     0
p =
    0.9517
ksstat =
    0.0358
cv =
    0.0952
```
Závěr: Na hladině významnosti 0,05 ………………… hypotézu, že vygenerovaná data pocházejí z rozložení N(0,1), protože p-hodnota je …….., což je …………… než 0,05.

Použití funkce normrnd.m
```
realizace=normrnd(…,…,…,…);
Fi=normcdf(realizace,…,…);
[h,p,ksstat,cv]=kstest(…,[…,…])
```

V tomto konkrétním případě jsme dostali následující výstupy:
```
h =
     0
p =
    0.5528
ksstat =
    0.0554
cv =
    0.0952
```
Závěr: Na hladině významnosti 0,05 ………………… hypotézu, že vygenerovaná data pocházejí z rozložení N(0,1), protože p-hodnota je …….., což je …………… než 0,05.

Nepovinný úkol: Místo funkce kstest.m použijte k testování normality funkci chi2gof.m a funkci adtest.m
Návod:

Funkce chi2gof.m:
Implicitně třídí data do 10 intervalů.
`[h,p] = chi2gof(realizace,'cdf',@normcdf)`
Vstupní parametry:
`realizace` … sloupcový vektor realizací
`'cdf'` … parametr, který dává funkci na vědomí, že bude použita distribuční funkce nějakého rozložení
`@normcdf` … označení distribuční funkce standardizovaného normálního rozložení
Výstupní parametry:
`h` … nabývá hodnoty 0, když nulovou hypotézu nezamítáme na hladině významnosti 0,05 a hodnoty 1, když zamítáme na hladině významnosti 0,05
`p` … p-hodnota

Funkce adtest.m:
`[h,p,adtsat,cv] = adtest(realizace)`
Vstupní parametr:
`realizace` … sloupcový vektor realizací
Výstupní parametry:
`h` … nabývá hodnoty 0, když nulovou hypotézu nezamítáme na hladině významnosti 0,05 a hodnoty 1, když zamítáme na hladině významnosti 0,05
`p` … p-hodnota
`adstat` … hodnota testové statistiky
`cv` … kritická hodnota

Úkol 3.: Pomocí funkce unifrnd.m vygenerujte 1000 čísel z rozložení Rs(0,1). Na hladině významnosti 0,05:
a) proveďte testy náhodnosti, a to test založený na bodech zvratu, test znamének diferencí a test založený na Spearmanově koeficientu;
b) proveďte testy nezávislosti, a to test založený na koeficientu autokorelace 1. až 10. řádu a Cochranův test.

Návod:
Vygenerujeme n = 1000 realizací z Rs(0,1):
```
n=1000;
x=unifrnd(…,…,…,…);
```
Zvolíme hladinu významnosti:
```
alfa=0.05;
```

Ad a)
<u>Provedení testu založeného na bodech zvratu</u>
Zavoláme funkci body_zvratu:
```
[h,p,u]=body_zvratu(…,…)
```
Vstupní parametry:
`x` … sloupcový vektor realizací
`alfa` ... hladina významnosti

Výstupní parametry:
`h` … nabývá hodnoty 0, když nulovou hypotézu nezamítáme na hladině významnosti 0,05 a hodnoty 1, když zamítáme na hladině významnosti 0,05
`p` … p-hodnota
`u` … hodnota testové statistiky

V tomto konkrétním případě jsme dostali následující výstupy:
```
h =
     0
p =
    0.3545
u =
   -0.9258
```
Závěr: Test založený na bodech zvratu na hladině významnosti 0,05 ………… hypotézu, že posloupnost vygenerovaných čísel je náhodná, protože p-hodnota je …….., což je …….. než 0,05.

Provedení testu založeného na znaménkách 1. diferencí

Zavoláme funkci znamenka_diferenci:
```
[h,p,u]=znamenka_diferenci…,…)
```
Vstupní a výstupní parametry jsou stejné jako u funkce body_zvratu.

V tomto konkrétním případě jsme dostali následující výstupy:
```
h =
     0
p =
    0.4115
u =
    0.8212
```
Závěr: Test založený na znaménkách 1. diferencí na hladině významnosti 0,05 ………….. hypotézu, že posloupnost vygenerovaných čísel je náhodná, protože p-hodnota je ………, což je ………. než 0,05.

Provedení testu založeného na Spearmanově koeficientu

Utvoříme vektor `y=[1:n]';`
Pomocí funkce tiedrank zjistíme vektor pořadí:
```
R=tiedrank(x);
```
Pomocí funkce corrcoef spočteme koeficient korelace a odpovídající p-hodnotu:
```
[rs,p]=corrcoef(y,R)
```
Rozhodnutí o nulové hypotéze učiníme na základě porovnání p-hodnoty se zvolenou hladinou významnosti alfa.
Je-li p ≤ alfa, hypotézu o pořadové nezávislosti realizací zamítáme na hladině významnosti alfa, v opačném případě nikoliv.

V tomto konkrétním případě jsme dostali následující výstupy:
```
rs =
    1.0000   -0.0225
   -0.0225    1.0000
p =
    1.0000    0.4767
    0.4767    1.0000
```
Závěr: Test založený na Spearmanově koeficientu pořadové korelace na hladině významnosti 0,05 ………….. hypotézu, že posloupnost vygenerovaných čísel je náhodná, protože p-hodnota je ………, což je ………. než 0,05.
Spearmanův koeficient nabývá hodnoty …….., což svědčí o existenci
- zanedbatelně slabé přímé pořadové závislosti
- zanedbatelně slabé nepřímé pořadové závislosti
- slabé přímé pořadové závislosti
- slabé nepřímé pořadové závislosti
- středně silné přímé pořadové závislosti
- středně silné nepřímé pořadové závislosti
- silné přímé pořadové závislosti
- silné nepřímé pořadové závislosti

Ad b)
<u>Provedení testu založeného na koeficientech autokorelace 1. až 10. řádu a Cochranova testu</u>:
1. „Ruční” výpočet + funkce Cochran.m
Vypočteme koeficient autokorelace 1. řádu a odpovídající p-hodnotu:
```
[r,p]=corrcoef(x(1:n-1),x(2:n))
```
Rozhodnutí o nulové hypotéze učiníme na základě porovnání p-hodnoty se zvolenou hladinou významnosti alfa.
Je-li p ≤ alfa, hypotézu o neexistenci autokorelace 1. řádu zamítáme na hladině významnosti alfa, v opačném případě nikoliv.
Analogicky počítáme koeficient autokorelace 2. řádu a příslušnou p-hodnotu:
```
[r,p]=corrcoef(...,...)
```
Tak postujeme dál až ke koeficientu autokorelace 10. řádu:
```
[r,p]=corrcoef(...,...)
```
Postup lze zjednodušit použitím cyklu:
```
rad=10;
for i=1:rad
    [r,p]=corrcoef(x(1:n-i),x(i+1:n))
end
```

V tomto konkrétním případě jsme dostali následující výstupy:
```
r =
    1.0000    0.0176
    0.0176    1.0000
p =
    1.0000    0.5777
    0.5777    1.0000
```
Koeficient autokorelace 1. řádu je .... a .... významný na hladině významnosti 0,05.

```
r =
    1.0000   -0.0012
   -0.0012    1.0000
p =
    1.0000    0.9688
    0.9688    1.0000
```
Koeficient autokorelace 2. řádu je .... a ....významný na hladině významnosti 0,05.

```
r =
    1.0000   -0.0016
   -0.0016    1.0000
p =
    1.0000    0.9603
    0.9603    1.0000
```
Koeficient autokorelace 3. řádu je .... a ....významný na hladině významnosti 0,05.

```
r =
    1.0000   -0.0870
   -0.0870    1.0000
p =
    1.0000    0.0060
    0.0060    1.0000
```
Koeficient autokorelace 4. řádu je .... a ....významný na hladině významnosti 0,05.

```
r =
    1.0000   -0.0044
   -0.0044    1.0000
p =
    1.0000    0.8885
    0.8885    1.0000
```
Koeficient autokorealce 5. řádu je .... a ....významný na hladině významnosti 0,05.

```
r =
    1.0000    0.0023
    0.0023    1.0000
p =
    1.0000    0.9424
    0.9424    1.0000
```
Koeficient autokorelace 6. řádu je .... a ....významný na hladině významnosti 0,05.

```
r =
    1.0000   -0.0525
   -0.0525    1.0000
p =
    1.0000    0.0982
    0.0982    1.0000
```
Koeficient autokorelace 7. řádu je .... a ....významný na hladině významnosti 0,05.

```
r =
    1.0000    0.0507
    0.0507    1.0000
p =
    1.0000    0.1104
    0.1104    1.0000
```
Koeficient autokorelace 8. řádu je .... a ....významný na hladině významnosti 0,05.

```
r =
    1.0000   -0.0244
   -0.0244    1.0000
p =
    1.0000    0.4426
    0.4426    1.0000
```
Koeficient autokorelace 9. řádu je .... a ....významný na hladině významnosti 0,05.

```
r =
    1.0000    0.0061
    0.0061    1.0000
p =
    1.0000    0.8478
    0.8478    1.0000
```
Koeficient autokorelace 10. řádu je .... a ....významný na hladině významnosti 0,05.

Cochranův test provádí funkce Cochran.m.

```
[Q,chi,h,p]=Cochran(x,rad,alfa)
```
Vstupní parametry:
`x` … sloupcový vektor realizací
`rad` … řád autokorelace
`alfa` … hladina významnosti
Výstupní parametry:
`Q` … hodnota testové statistiky
`chi` … kritická hodnota
`h` … nabývá hodnoty 0, když nulovou hypotézu nezamítáme na hladině významnosti alfa a hodnoty 1, když zamítáme na hladině významnosti alfa
`p` … p-hodnota

V tomto konkrétním případě jsme dostali následující výstupy:
```
Q =
   13.8808
chi =
   18.3070
h =
     0
p =
    0.1785
```
Závěr: Cochranův test na hladině významnosti 0,05 …………. hypotézu, že všechny koeficienty autokorelace od 1. do 10. řádu jsou nulové, protože p-hodnota je …….., což je …………… než 0,05.

<u>Použití funkce autokorelace.m</u> (oproti funkci Cochran.m poskytne ještě korelogram)

```
[T,Q,p,h]=autokorelace(x,k,alfa)
```

Funkce autokorelace.m má 3 vstupní parametry:
x ... vektor realizací
k ... maximální řád koeficientu autokorelace
alfa ... volitelný parametr, hladina významnosti pro Cochranův test, implicitně alfa=0,05
a 4 výstupní parametry:
T ... tabulka obsahující korelogram a výsledky testu hypotézy H_0: $\rho_k = 0$
Q ... hodnota testové statistiky Cochranova testu
p ... p-hodnota Cochranova testu
h ...výsledek Cochranova testu: h=0 nezamítáme hypotézu, že všechny koeficienty autokorelace jsou nula, h=1 zamítáme tuto hypotézu

```
T =
    1.0000    0.0176    0.5777         0
    2.0000   -0.0012    0.9688         0
    3.0000   -0.0016    0.9603         0
    4.0000   -0.0870    0.0060    1.0000
    5.0000   -0.0044    0.8885         0
    6.0000    0.0023    0.9424         0
    7.0000   -0.0525    0.0982         0
    8.0000    0.0507    0.1104         0
    9.0000   -0.0244    0.4426         0
   10.0000    0.0061    0.8478         0
Q =
   13.8808
p =
    0.1785
h =
     0
```