Příklady ze Statistiky

Regrese ——————————————

```
Příklad 1
V továrně byla sledována závislost celkových
nákladů "n" (v tis. Kč.) na produkci "p". Byly
zaznamenány následující údaje
p = [532 297 378 121 519 613 592 497];
n = [ 48  32  42  27  45  51  53  48];
a) Pomocí lineární regrese odhadněte náklady pro produkci
   1000 výrobků.
b) Pomocí lineární regrese odhadněte, pro jakou produkci budou
   náklady rovny 100 000 Kč.

Výsledek
 param  = 0.0535136, 19.510024
 yp  = 73.023636
 xp  = 1504.1028


Příklad 3
Do vodní nádrže unikla jedovatá látka. Byl aplikován
neutralizační prostředek a v okamžicích "xi" byla měřena
zbytková koncentrace škodliviny "yi" v nádrži. Zjištěné údaje
jsou (čas v minutách, koncentrace v promile)
xi = [5 12 20 26 29 38 65 126];
yi = [19 17 18 17 17 15 14 7];
Vypočtěte korelační koeficient pro lineární regresi, posudte
její vhodnost a odhadněte, kdy bude koncentrace škodliviny
nulová.

Výsledek
 Korelační koeficient
 r  = -0.9832531
Parametry
 p  = -0.0948295, 19.305033
Nulová koncentrace bude v x0
 x0  = 203.57626


Příklad 5
Na sledovaném procesu byla naměřena data
xi = [5 12 20 26 29 38 40 45];
yi = [19 17 12 1 27 35 44 76];
Pro tato data provedte polynomiální regresi "k"-tého stupně a
exponenciální regresi. Pomocí chyby predikce určete, která
z regresí je vhodnější.
```

```
k = 3;

Výsledek
 SEp  = 0.0365208
 SEe  = 0.1113473
 Polynomiální je lepší.
```

```
Příklad 6
Na sledovaném procesu byla naměřena data
x1i = [15 12 11  9  9  8  5  3]';
x2i = [3   9  5 11 28 14 32 58]';
yi = [9 7 22 12 27 31 44 36]';
Pro tato data provedte vícenásobnou lineární regresi
a testujte její vhodnost pomocí testu na bělost reziduí.

Výsledek
SE  = 0.2799241
Regrese se jeví jako vhodná.
```

Parametrické testy ——————————————-

```
Příklad 7
Za předpokladu, že výška dětí ve věku 10 let má normální
rozdělení s rozptylem "sig2", určete interval al-I, ve kterém
bude ležet neznámá střední hodnota výšky dětí, jestliže
z výběru o rozsahu "n" byl vypočten průměr "mx".
Interval určete jako a) oboustranný,  b) pravostranný.
sig2 = 38;
n    = 12;
mx   = 127.3;
al   = .01;

Výsledek
 CI_B  = 122.71628   131.88372
 CI_L  = 123.16023   Inf
 CI_R  = -Inf   131.43977
```

```
Příklad 8
Za předpokladu, že výška dětí ve věku 10 let má normální
rozdělení určete interval al-I, ve kterém bude ležet neznámá
střední hodnota výšky dětí, jestliže z výběru o rozsahu 'n'
byl vypočten průměr 'mx' a rozptyl 's2x'.
Interval určete jako a) oboustranný,  b) levostranný.
n    = 12;
mx   = 127.3;
```

```
s2x  = 38;
al   = .01;
```

Výsledek

```
 CI_B  = 121.77318   132.82682
 CI_L  = 122.46314   Inf
 CI_R  = -Inf   132.13686
```

Příklad 9

Pro zjištění přesnosti metody pro stanovení obsahu manganu v oceli byla provedena nezávislá měření vzorků. Chceme stanovit hranici, pro níž platí, že pravděpodobnost hodnoty skutečného rozptylu metody větší než tato hranice bude jen 5%.
Změřené hodnoty 'x' jsou

```
x = [4.3 2.9 5.1 3.3 2.7 4.8 3.6];
```

Výsledek

Hranice je 3.209

Příklad 10

Na magistrále v úseku s doporučenou rychlostí 80 km/h jsme kontrolovali rychlost vozidel a získali údaje 'xi' a 'ni' (hodnoty a četnosti). Určete oboustranný 'al'-interval pro rozptyl rychlostí, kterými řidiči v tomto úseku jezdí

```
xi = [60 70 75 80 85 90 110];
ni = [ 3 27 36 29 25 31   8];
al = .05;
```

Výsledek

```
  CI  = 79.89   124.34
```

Příklad 11

Na magistrále v úseku s doporučenou rychlostí 80 km/h jsme kontrolovali rychlost vozidel. Výběrem byly zjištěny rychlosti vozidel 'x'. Určete oboustranný 'al'-IS pro podíl řidičů, kteří překračují doporučenou rychlost o 'r' km/h.

```
x = [78 86 65 92 83 92 85 66 42 82 ...
99 92 75 81 66 76 89 76 97 76 ...
75 56 76 78 96 77 86 79 86 93];
al= .05;
r = 3;
```

Výsledek

```
  CI_obou  = 0.2246955   0.5753045
  CI_levo  = 0.   0.5471202
```

Příklad 12
Ze souboru ocelových nosníků stejné nominální délky jsme provedli náhodný výběr s délkami 'x'. Výrobce zaručuje rozptyl výrobků na hodnotě 'sig2'. Na hladině významnosti 'al' testujte tvrzení: Nominální délka nosníků je 'd'.

```
x    = [6.2 7.5 6.9 8.9 6.4 7.1];
d    = 6.5;
sig2 = .8;
al   = .05;
```

Výsledek

```
 pv  = 0.0678892
```

Příklad 13
Ze souboru ocelových nosníků stejné nominální délky jsme provedli náhodný výběr 'n' nosníků a vypočetli průměr 'mx' a směrodatnou odchylku 'sx'. Na hladině významnosti 'al' testujte tvrzení: Nominální délka nosníků je 'd'.

```
n  = 50;
mx = 5.77;
sx = .8;
d  = 6;
al = .05;
```

Výsledek

```
 pv  = 0.1534113
```

Příklad 15
Přesnost nastavení obráběcího stroje se zjistí z rozptylu průměrů vyráběných součástek. Stroj je nutné znovu nastavit, jestliže jeho rozptyl je větší než 'sig2'. Provedli jsme výběr a zjistili hodnoty průměrů 'nx' - četnosti, 'mx' - průměry. Je třeba stroj znovu nastavit? Testujte na hladině 'al'.

```
nx   = [ 5  12  32  11   8   3];
mx   = [95 100 105 110 115 120];
sig2 = 28;
al   = .05;
```

Výsledek

```
 pv  = 0.066
```

Příklad 16
Pevnost materiálu se ověřuje pomocí metod A a B. Tentýž materiál byl podroben testování oběma metodami. Výsledky testů jsou 'xA' a 'xB'. Na hladině významnosti 'al' testujte

shodu obou testů. Rozptyly lze považžovat za shodné.
```
xA = [20.1 19.6 20.0 19.9 20.1];
xB = [20.9 20.1 20.6 20.5 20.7 20.5];
al = .05;
```

Výsledek
```
 pv  = 0.0024367
```

Příklad 17
Testujeme sjetí pneumatik na nevé a pravé straně u předních kol automobilů. U náhodně vybraných automobilů byly změřeny vždy oba údaje s výsledky 'xL' a 'xP'. Testujte na hladině 'al'.
```
xL = [1.8 1.0 2.2 0.9 1.5];
xP = [1.5 1.1 2.0 1.1 1.4];
al = .05;
```

Výsledek
```
pv  = 0.5528894
```

Příklad 18
Na magistrále v úseku s doporučenou rychlostí 80 km/h jsme kontrolovali rychlost vozidel. Výběrem byly zjištěny rychlosti vozidel 'x'. Na hladině významnosti 'al' testujte hypotézu:
Podil řídičů, kteří překračují doporučenou rychlost o 'r' km/h. není větší než 'P' procent.
```
x = [78 86 65 92 83 92 85 66 42 82 ...
99 92 75 81 66 76 89 76 97 76 ...
75 56 76 78 96 77 86 79 86 93];
al= .05;
r = 3;
P = 20;
```

Výsledek
```
pv  = 0.0030849
```

Příklad 19
Na magistrále v úseku s doporučenou rychlostí 80 km/h jsme kontrolovali rychlost vozidel směrem do Prahy a ven z Prahy. Výběrem jsme získali rychlosti do Prahy a z Prahy. Rychlosti jsme roztřídili a zanamenali jako četnosti 'nD' a 'nZ' a hodnoty 'rD' a 'rZ' (D je do Prahy a Z je z Prahy).
Na hladině 'al' testujte hypotézu: Z Prahy jezdí auta rychleji.
```
nD = [ 5 11 17 65 98 73 79  63   3];
rD = [65 70 75 80 85 90 95 100 110];
nZ = [ 8 22 13 71 48 64 89  24   5];
```

```
rZ = [65 70 75 80 85 90 95 100 110];
al = .01;
```

Výsledek

```
 pv  = 0.0195199
```

Příklad 20

Byla provedena namátková kontrola seřízení předních světel automobilů. U každého automobilu byla kontrolována obě světla s výsledky v centimetrech (+ nad a - pod optimální úrovní). Získané hodnoty jsou 'xP' (pravé) a 'xL' (levé reflektory). Na hladině 'al' testujte tvrzení, že jednotlivé automobily mají oba reflektory seřízeny stejně.

```
xP = [-3   5  16  9 -8 -2  23  5  -6 -3];
xL = [-5 -12  22 -3 -9  1  -1  2 -13 -5];
al = .1;
```

Výsledek

```
pv  = 0.0749262
```

Příklad 21

Byla provedena namátková kontrola seřízení předních světel automobilů. U každého automobilu byla kontrolována obě světla s výsledky v centimetrech (+ nad a - pod optimální úrovní). Získané hodnoty jsou 'xP' (pravé) a 'xL' (levé reflektory). Na hladině 'al' testujte tvrzení, že levé reflektory jsou u automobilů méně sklopeny než pravé.

```
xP = [-3   5  16  9 -8 -2  23  5  -6 -3];
xL = [-5 -12  22 -3 -9  1  -1  2 -13 -5];
al = .1;
```

Výsledek

```
 pv  = 0.0374631
```

Příklad 22

Na křižovatce jsme po určitou dobu zaznamenávali počty automobilů jedoucích rovně (R) a odbočujících doleva (L) a doprava (P). Zjištěné hodnoty byly 'xR', 'xL' a 'xP'. Na hladině 'al' testujte tvrzení, že podíl automobilů jedoucích rovně je stejný jako těch, kteří odbočují.

```
xR = 62;
xL = 39;
xP = 46;
al = .1;
```

Výsledek

pv = 0.060902

Příklad 23
Na křižovatce jsme opakovaně zaznamenávali počty
automobilů jedoucích rovně (R) a odbočujících doleva (L)
a doprava (P). Zjištěné hodnoty jsou 'xR', 'xL' a 'xP'. Na
hladině 'al' testujte tvrzení, že průměrné množství automobilů
odbočujících je větší než těch, kteří jedou rovně.
xR = [82 78 92 83 99 97];
xL = [29 42 34 38 45 34];
xP = [31 44 36 54 31 24];
al = .05;

Výsledek
 pv = 1.130D-69

Anova ———————————

Příklad 24
Sledujeme tři stroje. Náhodně zjištujeme jejich hodinové
produkce "x1", "x2" a "x3". Je prvdivé tvrzení, že na hladině
významnosti "al" jsou průměrné produkce všech tří strojů
shodné?
al = 0.05;
x1 = [53 55 49 58 52 61 56 55];
x2 = [49 56 52 45 51 56 44 51];
x3 = [52 53 52 54 55 53 53 52];

Výsledek
  pv = 0.0541908

Příklad 25
V měsíci bezpečnosti sledujeme počet nehod na pěti pražských
křižovatkách. Výsledky jsou v následující tabulce

rok:	1999	2000	2001	2002	2003
počet_1	3	5	2	1	3
počet_2	6	2	5	3	4
počet_3	3	2	1	1	2
počet_4	4	1	1	2	2
počet_5	4	2	5	5	6

Lze na hladině významnosti "al" tvrdit, že průměrný počet
nehod je na všech křižovatkách stejný?
al = 0.01;

```
Výsledek
Jednofaktorová anova
pv  = 0.0207414

Dvoufaktorová anova
Shoda v křižovatkách  0.0195312

Shoda v čase  0.2749788
```

Neparametrické testy ———————————

```
Příklad 26
Na křižovatce jsme v různých intervalech zaznamenávali počty
projíždějících automobilů. Změřená hodnoty jsou "di" - delka
intervalu pozorování (min) a "xi" - počet pozorovaných
automobilů
d = [15 10 20 35 10 50];
x = [71 56 98 121 44 271];
Na hladině "al" testujte tvrzení, že automobily projíždějí
rovnoměrně.
al = .05;

Výsledek
 pv  = 0.001917


Příklad 27
Následující tabulka udává četnosti drobných nehod ve velké
továrně, pozorované během jednoho dne
-------------------------------------
doba   8-10h. 10-12h. 12-13h. 13-17h.
počet   2       7       1       16
-------------------------------------
Na hladině "al" testujte tvrzení, že nehody jsou rozloženy
rovnoměrně.
al = .05;

Výsledek
 pv  = 0.1286455


Příklad 28
Zjištovala se závislost mezi barvou očí a vlasů u mužů. Ve
výběru dotázaných mužů byly zjištěny následující četnosti
----------------------------------
oči \ vlasy   světlé  hnědé  tmavé
modré           90      75     55
šedé            96     136     88
```

```
hnědé          108     135    119
---------------------------------
```
Na hladině "al" testujte tvrzení, že barva očí a vlasů je
navzájem nezávislá.
al = .05;

Výsledek
pv = 0.0167358

Příklad 29
Na dvou strojích se pravidelně střídají dva operátoři.
Výrobky, které se na strojích vyrobí projdou kontrolou a každý
vadný je označen podle stroje (S) a operátora (O). Byly
zjištěny následující údaje
```
---------------------------------------------------
stroj     1 2 1 1 2 2 2 1 2 1 1 1 2 1 2 2 2 1 2 1 2
operátor  2 2 1 2 1 1 2 2 2 1 2 2 1 2 1 2 2 2 1 1 2
---------------------------------------------------
```
Na hladině "al" testujte tvrzení, že operátoři a stjoje jsou
při výrobě zmetků nezávislí.
al = .05;

Výsledek
pv = 0.4663938

Příklad 30
Dva doktoři doporučují léčit rýmu dvěma různými způsoby.
Zjištěné výsledky (doba léčení) jsou 'x1' and 'x2'.
Testujte shodu léčebných metod.
x1=[5 8 7 8 4 5 5 6 9 3 5 8 6];
x2=[3 4 9 5 4 9 9 8];

Výsledek
pv = 1.

Příklad 31
Osm sportovců z jistého sportovního klubu podstoupilo test výkonnosti.
Každý z nich hodil oštěpem a poté byl podroben intenzivnímu tréningu.
Druhý den házel podruhé. Obě délky hodu byly zaznamenány jako 'x1' a 'x2'.
Trenér očekává, že jeden den trénigu je málo na zvýšení výkonnosti.
Testujte na hladině 'al'.

x1=[68 81 69 72 66 91 98 89 75 68];
x2=[79 62 70 75 68 81 85 94 71 62];

Výsledek

```
pv = 1.
```

Příklad 32
Testujte jestli jeleni a myši mají stejně dlouhé obě přední nohy.
Změřená data jsou 'x1' (levé) a 'x2' (pravé).

```
x1=[135 123 3.1 2.5 98 124 131 3.4 2.8 128];
x2=[136 121 2.9 2.6 101 121 130 3.5 2.9 126];
```

Výsledek
pv = 1.

Příklad 33
Tři kontroloři hodnotí úroveň pěti stýnků s rychlým občerstvením.
Každý kontrolor ohodnotí každý stánek. Výskledky jsou v tabulce 'Tab'.
Řádky odpovídají kontrolorům, sloupce stánkům. Hodnocení je 1 až 10
(10 je nejlepší). Testujte, zda stánky mají stejnou kvalitu.

```
Tab=[10 8 3 9 7
8  7 5 9 10
8  9 5 7  6];
```

Výsledek
pv = 0.1545873

Příklad 34
Továrna vyrábí výrobky, které by měly mít stejnou váhu.
Pro výrobu jsou použity čtyři stroje. Od každého stroje
jsme odebrali několik výrobků s vahami 'x1' až 'x4'.
Testujte, zda výhy výrobků z jednotlivých strojů jsou shodné.

```
x1=[39.4 34.8 35.6 35.1 35.8];
x2=[34.4 34.2 35.1 31.1 32.5 33.8];
x3=[30.2 35.1 34.2 36.3 30.8 35.6 35.2];
x4=[39.1 34.3 38.6 34.5 36.4 36.1];
```

Výsledek
pv = 0.0335338

Example 35
Označme X délku (cm) ryby určitého druhu. Změřili jsme
vzorek X z náhodně vybraných ryb. Lze prohlásit,
že medián délky těchto ryb se významně neliší od
hodnoty X0 cm?
X=[5.0 3.9 5.2 5.5 2.8 6.1 6.4 2.6 1.7 4.3]
X0=3.7

Výsledek
Málo dat!

Example 36
Totéž jako v 35, ale s delším vzorkem.
Označme X délku (cm) ryby určitého druhu. Změřili jsme
vzorek X z náhodně vybraných ryb. Lze prohlásit,
že medián délky těchto ryb se významně neliší od
hodnoty X0 cm?
X=[5.0 3.9 5.2 5.5 2.8 6.1 6.4 2.6 1.7 4.3 3.5 3.2];
X0=3.7;

Výsledek
pv=1

Poznámka:
1. Zkuste měnit hodnotu X0.
2. Zkuste přidat další data z množiny
5.6 6.8 3.4 8.4 6.9 7.5 9.2 9.4 8.9 7.6