TEST MATLAB (STATISTIEK EN WISKUNDIGE DATA-ANALYSE)
(1ste zit '20-'21, reeks C)
Opleiding industrieel ingenieur

FACULTEIT INGENIEURSWETENSCHAPPEN EN ARCHITECTUUR

Naam : Richting: /20

Schrijf netjes. Vul in op de opengelaten plaatsen.
De MATLAB-code komt in de kadertjes.
Geen gsm, smartphone, rekentoestel .... Veel succes!

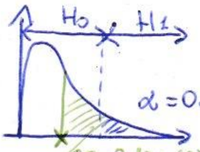
UNIVERSITEIT GENT

1. Formuleer in deze oefening telkens H_0 en H_1 en maak een schets met alle informatie indien van toepassing; vermeld de berekende waarden gegenereerd door MATLAB (geen code). t- test onafhankelijke steekproeven
De secretaresse van een huisartspraktijk vermoedt dat huisarts A meer huisbezoeken maakt dan huisarts B. Zij noteert het aantal huisbezoeken aan vijftien patiënten van huisarts A en het aantal huisbezoeken aan vijftien verschillende patiënten van huisarts B. Hieronder vind je de data die de secretaresse vergaart (aantal huisbezoeken aan patiënten per huisarts): /4

Huisarts A	8	1	8	8	3	5	6	4	1	8	3	3	6	2	4
Huisarts B	9	8	0	0	2	4	7	4	2	7	2	10	5	1	7

Kan je op basis van bovenstaande resultaten besluiten dat huisarts A meer huisbezoeken maakt dan huisarts B. Voer de test uit op 95%-niveau. (Ga de voorwaarden na voor het uitvoeren van de test in kwestie.)

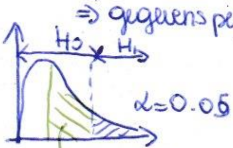
VW1: normaliteit H_0: gegevens per groep normaal verdeeld
H_1: " " " NT " "



p = { 0.7131 (A) ; 0.6670 (B) } > α

⇒ H_0 aanvaarden op 95%-niveau
⇒ gegevens per groep norm. verd. op 95% niveau

VW2: Varianties gelijk? H_0: $\sigma_A^2 = \sigma_B^2$
H_1: $\sigma_A^2 \neq \sigma_B^2$

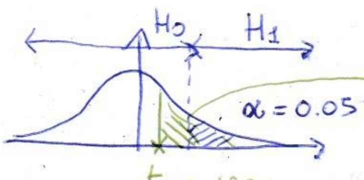


p = 0.2024 > α

⇒ H_0 aanvaarden op 95% niveau
⇒ varianties gelijk op 95%-niveau

t-test onafh. steekproeven

H_0: $\mu_A = \mu_B$
H_1: $\mu_A > \mu_B$ (eenzijdig)



t = 0.1234

p = 0.4513 > α
⇒ H_0 aanvaarden op 95%-niveau

⇒ op 95% niveau kan je nt besluiten dat huisarts A meer huisbezoeken maakt

2. Gegeven de vector $\{a_1, a_2, \cdots, a_n\}$ met $\begin{cases} a_1 = \ln(10) \\ a_n = \ln(1 + a_{n-1}) \text{ voor } n > 1 \end{cases}$

Bereken de som van de eerste tien elementen (numeriek, 4 cijfers na de komma).

$S =$ 6.7583

```
A = 1:10
A(1) = log(10)
for n = 2:10
    A(n)=log(1+A(n-1));
end
A
sum(A)
```

/3

3. Bepaal de waarde van x zodat de determinant van onderstaande matrix groter is dan nul.

$$\begin{pmatrix} x & x & 0 & 1 \\ x & x & 3 & x \\ 0 & x & 2 & x \\ 5 & x & 0 & x \end{pmatrix}$$

Antwoord als unie van deelintervallen.

$x \in$ $]-\infty, -\frac{5}{2}[\cup]0,1[$

```
syms x
A=[x x 0 1; x x 3 x
   0 x 2 x; 5 x 0 x]
det(A)
opl=solve(det(A)>0, x,
'ReturnConditions', true, 'Real', true)
opl.conditions
```

/2.5

4. Gegeven de vectoren $\vec{a} = \{1, 3(k-1), 5\}$ en $\vec{b} = \{-2, k-1, k-2\}$.
Bereken k (reëel) zodat $||\vec{a} \times \vec{b}|| = 9$.

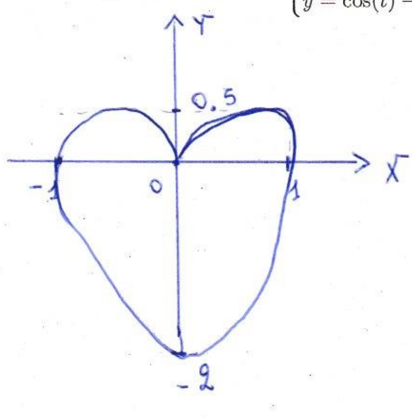
$k =$ 1

(numeriek, 4 cijfers na de komma)

```
syms k
a=[1 3*(k-1) 5]
b=[-2 k-1 k-2]
pr=cross(a,b)  n=norm(pr)
solve(n==9)
```

/2.5

5. Teken de parameterkromme $\begin{cases} x = \sin^3(t) \\ y = \cos(t) - \cos^4(t) \end{cases}$




```
t=0:0.01:2*pi;
x=sin(t).^3;
y=cos(t)-cos(t).^4;
plot(x,y)
axis equal
```

/2

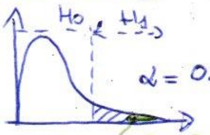
6. Formuleer in deze oefening telkens H_0 en H_1 en maak een schets met alle informatie indien van toepassing; vermeld de berekende waarden gegenereerd door MATLAB (geen code).
Bij een bank worden het aantal kredietkaarten per gezin (Y), het aantal gezinsleden (X_1), het gezinsinkomen (X_2 uitgedrukt in duizendtallen) en het aantal auto's per gezin (X_3) genoteerd. Hieronder staan deze gegevens voor 8 gezinnen. /6

Gezin	Y	X_1	X_2	X_3
1	4	2	14	1
2	6	2	16	2
3	6	4	14	2
4	7	4	17	1
5	8	5	18	3
6	7	5	21	2
7	8	6	17	1
8	10	6	25	2

Bepaal de meervoudige lineaire regressievergelijking voor Y in functie van X_1, X_2 en X_3. Kan je dit model aanvaarden als waardevol? Argumenteer ($\alpha = 0.1$). Zijn de coëfficiënten uit het gevonden regressiemodel significant verschillend van nul? Argumenteer ($\alpha = 0.1$). Bepaal een nieuwe enkelvoudige lineaire regressievergelijking ($y = \beta_0 + \beta_1 x$) voor Y in functie van de meest determinerende variabele (X_1, X_2 of X_3 en waarom?). Kan je dit model aanvaarden als waardevol? Argumenteer ($\alpha = 0.1$).

$Y = 0.2861 + 0.6346 X_1 + 0.1995 X_2 + 0.2716 X_3$

H_0: het regressiemodel is niet significant
H_1: het regressiemodel is significant



$\alpha = 0.1$

$p = 0.0294 < \alpha = 0.1$

$\Rightarrow$ H_0 verwerpen: het regressiemodel is significant met 90% betrouwbaarheid

$\beta_0 = 0.2861$: $0 \in [-3.1374, 3.7095]$ $\beta_1 = 0.6346$: $0 \notin [0.0568; 1.2124]$

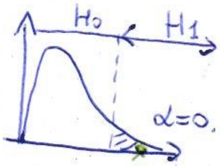
$\beta_2 = 0.1995$: $0 \in [-0.0550; 0.4541]$ $\beta_3 = 0.2716$: $0 \in [-0.7308; 1.2741]$

$\Rightarrow$ enkel β_1 is significant verschillend van nul met 90% betrouwbaarheid

Nieuwe enkelvoudige lin. regressie vgl: $y = \beta_0 + \beta_1 X_1$ (want $\beta_1 \neq 0$)

$Y = 2.8714 + 0.9714 X_1$

H_0: het regressiemodel is niet significant
H_1: het regressiemodel is significant



$\alpha = 0.1$

$P = 0.0054$

$p = 0.0054 < \alpha = 0.1 \Rightarrow H_0$ verwerpen: het regressiemodel is significant met 90% betrouwbaarheid