Efterår2016Reeksamen

```
clearvars
clc
close all
```

Opgave 1: Stokastiske Variable

En kontinuert stokastisk variabel X har følgende tæthedsfunktion (pdf):

$$f_X(x) = \begin{cases} A \cdot x + B & -2 \leq x \leq 3 \\ 0 & ellers \end{cases}$$

1) En skitse af $f_X(x)$ er som på figur 1, hvor det bemærkes at $f_X(3) = 0$.
For hvilken værdi af $f_X(-2) = k$, er $f_X(x)$ en gyldig tæthedsfunktion?
Begrund dit svar.

```
clearvars
```

For at fx(x) er en gyldig tæthedsfunktion gælder følgende:

$$\int_{-2}^{3} fx(x)dx = 1$$

Areal skal under kurven være 1, derved bruger vi formlen fra en retvinklet trekant.

$$Areal = 1 = \frac{1}{2} * k * 5 \rightarrow \quad k = \frac{2}{5} \quad = \quad 0.4$$

2) Vis at fordelingsfunktionen (cdf) $F_X(x)$ for X er givet ved:

$$F_X(x) = \begin{cases} 0 & x < -2 \\ \frac{A}{2} \cdot x^2 + B \cdot x + C & -2 \leq x \leq 3 \\ 1 & 3 < x \end{cases}$$

Antag at $A = -\frac{2}{25}$ og $B = \frac{6}{25}$ og $C = \frac{16}{25}$.

$$Fx(x) = \int_{nedre\ værdi}^{x} x\, dx + tidl.\ led$$

- From pdf to cdf: $F_X(x) = \int_{-\infty}^{x} f_X(u)\, du = Pr(X \leq x)$
- From cdf to pdf: $f_X(x) = \frac{dF_X(x)}{dx}$

$$Fx(x) := \frac{a}{2} \cdot x^2 + b \cdot x + c$$

$$\frac{d}{dx} Fx(x) \rightarrow b + a \cdot x$$

3) Brug $f_X(x)$ til at finde forventningsværdien $E[X]$ og variansen ${\sigma_X}^2$ for X.

Antag at $A = -\frac{2}{25}$ og $B = \frac{6}{25}$.

Ved
forventningsværien gælder følgende:

$$\boldsymbol{E|x| = \int_{-\infty}^{\infty} x * fx(x) dx}$$

```
clearvars
syms x
A=-2/25; B = 6 /25;
fx(x) = piecewise( x < -2, 0,...
                    -2 <= x <= 3, A*x+B,...
                    x > 3 ,0)
```

fx(x) =

$$\begin{cases} 0 & \text{if } x < -2 \\ \frac{6}{25} - \frac{2x}{25} & \text{if } x \in [-2, 3] \\ 0 & \text{if } 3 < x \end{cases}$$

```
ex =  int(x*fx(x), -Inf, Inf)
```

ex =

$$-\frac{1}{3}$$

```
ex2 = int(x.^2*fx(x), -Inf, Inf)
```

ex2 =

$$\frac{3}{2}$$

$$var(x) = E\left|x^2\right| - E|x|^2$$

```
variance =  ex2-ex^2
```

variance =

$$\frac{25}{18}$$

4) Hvad er sandsynligheden $\Pr(X < 0)$? Begrund dit svar.

Antag at $A = -\frac{2}{25}$ og $B = \frac{6}{25}$ og $C = \frac{16}{25}$.

Følgende gælder:

$$Pr(x < 0) = Fx(0), \qquad Pr(x < 0{,}4) = \frac{A}{2} * x^2 + B * x + C$$

```
C = 16/25
```

```
C = 0.6400
```

```
Fx(x) = piecewise( x < -2, 0,...
            -2 <= x <= 3, A/2*x^2+B*x+C,...
            x > 3 ,1)
```

```
Fx(x) =
```

$$\begin{cases} 0 & \text{if } x < -2 \\ -\frac{x^2}{25} + \frac{6\,x}{25} + \frac{16}{25} & \text{if } x \in [-2, 3] \\ 1 & \text{if } 3 < x \end{cases}$$

```
Pr1 = Fx(0)
```

```
Pr1 =
```

$$\frac{16}{25}$$

Opgave 2: Stokastiske Processer

En diskret stokastisk process er givet for den n´te sample ved:

$$X(n) = W(n) + 0{,}7$$

hvor $W(n)$ er i.i.d. fordelte efter:

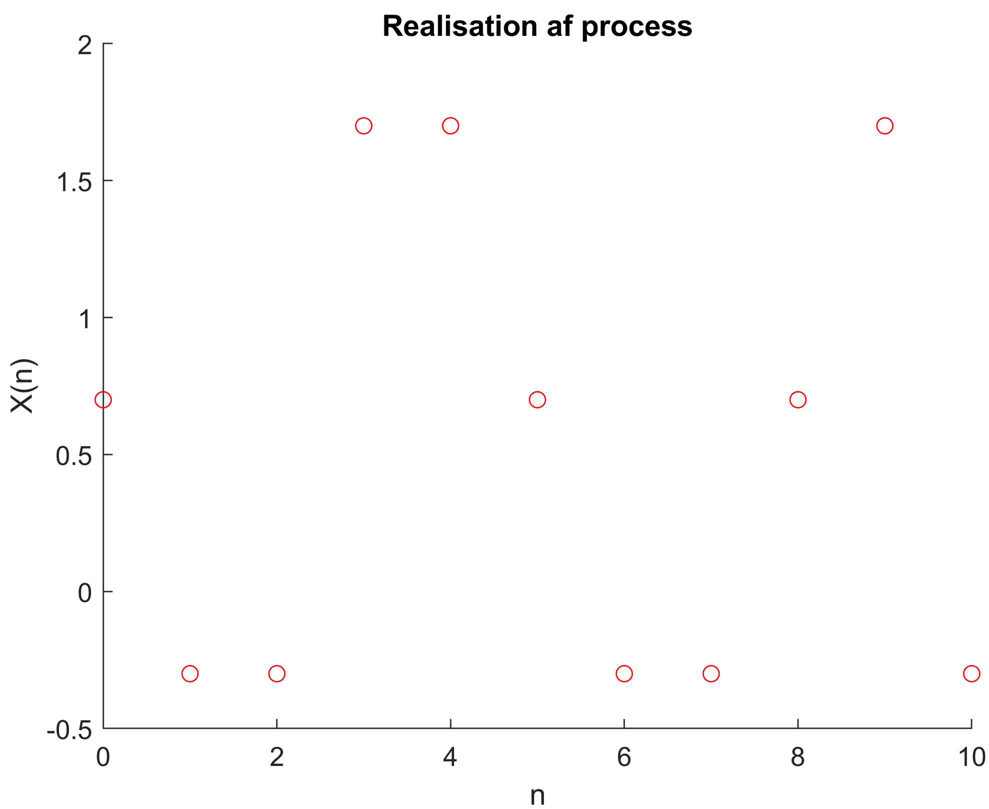
$w(n)$	-1	0	1
$f_{W(n)}(w(n))$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

1) Skitsér 11 samples fra 0 – 10 af én realisation af $X(n)$. Angiv hvorledes realisationen er fremkommet, brug en tilfældighedsgenerator, f.eks. unidrnd() i matlab.

$\mathcal{N}(\mu,\sigma^2)$
Mean value: μ
Variance: σ^2

```
x = 0:1:10;
samples = length(x);
```

```
w = unidrnd(3, 1, samples) - 2 + 0.7;
figure(1);
hold on
xlabel("n"); ylabel("X(n)"); title("Realisation af process");
scatter(x, w, 'or')
```



2) Bestem ensemble middelværdien og ensemble variansen for processen $X(n)$.

$$\boldsymbol{E}|\boldsymbol{x}| = \sum_{x} w * fw(w) + process\ værdi$$

```
x = [-1 0 1]
```

```
x = 1×3
    -1     0     1
```

```
fx= [1/3 1/3 1/3]
```

```
fx = 1×3
    0.3333    0.3333    0.3333
```

```
ex = sum(x.*fx)+0.7
```

```
ex = 0.7000
```

```
ex2= sum(x.^2.*fx)
```

```
ex2 = 0.6667
```

```
%%%% OBS %%%% Variance giver ikke rigtige svar
```

Ved variansen gælder følgende:

$$\sigma_x^2 = E|x^2| - E|x|^2$$

$$\sigma_x^2 = E|x^2| - E|x|^2$$

$$\boldsymbol{E|x^2|} = \sum_x w^2 * fw(w) + 0{,}7 \to (-1)^2 * \frac{1}{3} + 0^2 * \frac{1}{3} + 1^2 * \frac{1}{3} = \frac{2}{3}$$

$$\sigma_x^2 = E|x^2| - E|x|^2 = \frac{2}{3}$$

```
variance2 = ex2 - ex^2
```

```
variance2 = 0.1767
```

3) Opstil formlen til at bestemme autokorrelationsfunktionen for $X(n)$.

$$\mathrm{Rxx}(\tau) = \sum x(n)x(n-\tau) * \mathrm{fx}(x)$$

hvor $f(x) = \mathrm{fw}(w + 0.7)$

4) Angiv om processen $X(n)$ er WSS (stationær i den brede forstand), og om den er ergodisk. Begrund dine svar.

Da processen er uafhængig af tiden, og at middelværdien og variansen ikke ændrer sig over tid, så kan vi konstanter at processen er stationær.

Processen er egodisk da pmf kan bestemmes udfra en realisation

Opgave 3: Sandsynlighedsregning

I et tognet angives forsinkelser til at kunne skyldes blade på skinnerne, signalfejl eller personalemangel. Hændelserne er uafhængige og er ikke disjunkte.

Hvis der opstår en forsinkelse, vil der være blade på skinnerne $\frac{1}{4}$ af gangene, der vil være signalfejl $\frac{1}{2}$ af gangene, og der vil være personalemangel $\frac{1}{4}$ af gangene.

Haendelse A : Blade på skinnerne

Haendelse B: Signal fejl

Haendelse C: Personmangel

1) Skitsér Venn diagrammet (hændelsesdiagram for udfaldsrummet) for de tre begivenheder, når blade på skinnerne er hændelse A, signalfejl er hændelse B, og personalemangel er hændelse C.

2) Find sandsynligheden $\Pr(A \cap B)$. Dvs. sandsynligheden for, at der både er blade på skinnerne og signalfejl.

```
Pr_A = 1/4;
Pr_B = 1/2;
Pr_C = 1/4;

Pr_A_B = Pr_A*Pr_B
```

```
Pr_A_B = 0.1250
```

3) Find sandsynligheden $\Pr(A \cup B)$. Dvs. sandsynligheden for, at der enten er blade på skinnerne eller signalfejl.

Union: $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$

```
Pr_AUB = Pr_A + Pr_B - Pr_A_B
```

```
Pr_AUB = 0.6250
```

4) Find sandsynligheden $\Pr(A \cup B \cup C)$. Dvs. sandsynligheden for, at en forsinkelse på toget skyldes blade på skinnerne, signalfejl eller personalemangel.

Da A, B og C er uafhængige så tages der Pr(A U B U C)

```
Pr_A + Pr_B + Pr_C - Pr_A_B - (Pr_B*Pr_C) - (Pr_A*Pr_C) + (Pr_A*Pr_B*Pr_C)
```

```
ans = 0.7188
```

Opgave 4: Statistik

I et studie af en anerkendt metode til vægttab undersøges 10 patienter før behandling, og ét år efter behandling.

Patient nr.	Vægt før (kg)	Vægt efter (kg)
1	140	130
2	138	121
3	110	127
4	154	101
5	125	92
6	169	170
7	142	143
8	162	170
9	131	134
10	122	85

1) Opstil en NULL og Alternativ hypotese for at bestemme, om behandlingen har ændret patienternes vægt.

Null hypothesis. Brugt som testen af ingen forskel.

$H_0 = \mu 1 = \mu 2$

Alternanive hypotese

$H_1 = \mu 1 \neq \mu 2$

2) Bør testen, der udføres, være parret eller uparret? Begrund dit svar.

Da det er samme subjekt for før og efter, så bør det være en parret.

3) Estimér middelforskellen af patienternes vægt før og efter behandling.

```
patient = 1:1:10;
foer = [ 140 138 110 154 125 169 142 162 131 122];
efter = [130 121 127 101 92 170 143 170 134 85];

middelforskel = sum(efter-foer)/length(patient)
```

```
middelforskel = -12
```

4) Estimér variansen for forskellen af patienternes vægt før og efter behandling.

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

```
s2 = 1/(length(patient)-1)*(sum((efter-foer-middelforskel).^2))
```

```
s2 = 508.8889
```

5) Anvend en parret t-test til hypotesetest af din hypotese. Kan NULL hypotesen afvises med et signifikansniveau på 0,05? Begrund dit svar.

$$t = \frac{\bar{d}}{\sqrt{s^2/n}}$$

```
t = middelforskel/sqrt(s2/length(patient))
```

```
t = -1.6822
```

$$2 \cdot |1 - \Phi(|z|)$$

```
pval = 2*(1-tcdf(abs(t),(length(patient)-1)))
```

```
pval = 0.1268
```

Da pval er større end 0.05, så fejler vi med at afvise hypotese 0.

6) Opstil og find 95% konfidensintervallet for vægtforskellen før og efter behandlingen. Angiv hvilken formel, der er brugt.

Lower bound:

$$\delta_- = (\bar{x}_1 - \bar{x}_2) - t_0 \cdot s\sqrt{\frac{1}{n_1}}$$

Upper bound:

$$\delta_+ = (\bar{x}_1 - \bar{x}_2) + t_0 \cdot s\sqrt{\frac{1}{n_1}}$$

```
t0=tinv(0.975,(length(patient)-1))
```

```
t0 = 2.2622
```

```
koefMinus = middelforskel-t0*sqrt(s2)*sqrt((1/length(patient)))
```

```
koefMinus = -28.1374
```

```
koefPositiv = middelforskel + t0*sqrt(s2)*sqrt((1/length(patient)))
```

```
koefPositiv = 4.1374
```