ME591 - Assignment 4

Herman Kolstad Jakobsen
20196493

October 15, 2019

ME591 Assignment 4 Herman K. Jakobsen
20196493

Problem 1a)

The power spectral density function is defined as

$$G_{xx}(f) = \begin{cases} \frac{1}{100} f, & 0 \leq f \leq 200 \\ 2, & 200 \leq f \leq 300 \\ -\frac{1}{100} f + 5, & 300 \leq f \leq 400 \\ 1, & 400 \leq f \leq 500 \\ -\frac{1}{100} f + 6, & 500 \leq f \leq 600 \\ 0, & \text{otherwise} \end{cases}$$

The mean square is given as

$$\psi_x^2 = R_{xx}(0) = \int_{-\infty}^{\infty} S_{xx}(f)\, df = \int_0^{\infty} G_{xx}(f)\, df$$

$$= \frac{1}{100}\int_0^{200} f\, df + 2\int_{200}^{300} df + \int_{300}^{400} -\frac{1}{100} f + 5\, df + \int_{400}^{500} df$$

$$+ \int_{500}^{600} -\frac{1}{100} f + 6\, df$$

$$= \frac{1}{100}\left[\frac{1}{2} f^2\right]_0^{200} + 2\left[f\right]_{200}^{300} + \left[-\frac{1}{200} f^2 + 5f\right]_{300}^{400} + \left[f\right]_{400}^{500}$$

$$+ \left[-\frac{1}{200} f^2 + 6f\right]_{500}^{600}$$

$$= \frac{1}{200}\left(200^2\right) + 2(300-200) + \left(-\frac{400^2}{200} + 5\cdot 400 - \left(-\frac{300^2}{200} + 5\cdot 300\right)\right)$$

$$+ (500-400) + \left(-\frac{600^2}{200} + 6\cdot 600 - \left(-\frac{500^2}{200} + 6\cdot 500\right)\right)$$

$= 200 + 200 - 800 + 2000 + 450 - 1500 + 100$

$- 1500 + 3600 + 1250 - 3000$

$\underline{\underline{\psi_x^2 = 700}}$

7b)

Without pre-filtering, aliasing will occur. It will not be possible to distinguish between power contributions from frequency f or frequency $2nf_c \pm f$, where f_c is the Nyquist cut-off frequency defined as

$$f_c = \frac{f_s}{2} = \frac{600}{2} = \underline{300 \text{ Hz}}$$

Since $G_{xx}(f) = 0$ for $f > 600\text{Hz}$ we will only have $n=1$. So, for instance,

$0 \le f \le 100$ Hz will be aliased with power from $500 \le f \le 600$ Hz and $600 \le f \le 700$ Hz

Zero contribution

$100 \le f \le 200$ Hz -11- $400 \le f \le 500$ Hz

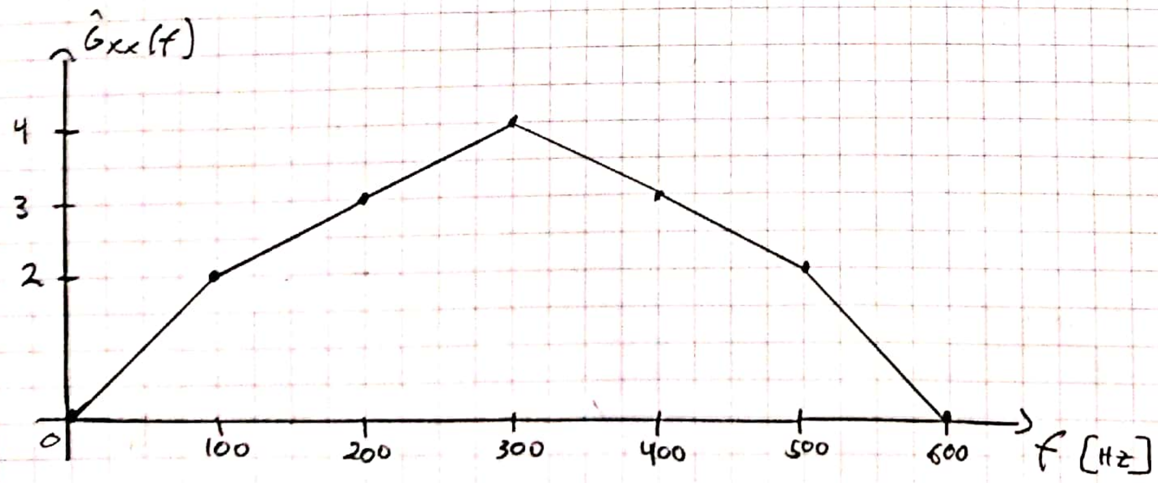
$200 \le f \le 300$ Hz -11- $300 \le f \le 400$ Hz

$300 \le f \le 400$ Hz -11- $200 \le f \le 300$ Hz

$400 \le f \le 500$ Hz -11- $100 \le f \le 200$ Hz

$500 \le f \le 600$ Hz -11- $0 \le f \le 100$ Hz

The estimate $\hat{G}_{xx}(f)$ is then



1c)

To prevent aliasing during the AD conversion, the signal should be low-pass filtered before it is fed into the ADC. The low-pass filter should remove information in the signal that might exist at frequencies above the Nyquist cut-off frequency. Hence all information above

$f_c = 300$ Hz

should be removed, ideally. The estimate $\hat{G}_{xx}(f)$ is then

Now

$0 \leq f \leq 100$ Hz will get power contribution from $500 \leq f \leq 600$ Hz but this will result in zero contribution. Likewise for the other frequencies.

The mean square is then

$$\psi_x^2 = \int_0^\infty G_{xx}(f)\, df$$

$$= \int_0^{200} \frac{1}{100} f \, df + \int_{200}^{300} 2 \, df$$

$$= 200 + 200$$

$$\psi_x^2 = 400$$

Problem 2a)

Given that

$$A = \sum_{i=1}^{N-1} A_i = \sum_{i=1}^{N-1} \left[\sum_{j=i+1}^{N} h_{ij} \right], \quad h_{ij} = \begin{cases} 1 & \text{if } x_i > x_j \text{ for } i<j \\ 0 & \text{otherwise} \end{cases}$$

So for our case

$$A = \sum_{i=1}^{19} \left[\sum_{j=i+1}^{20} h_{ij} \right]$$

Practically, this means that every A_i is the number of how many of the consecutive samples are greater than the current sample.

Applying this, we end up with

$A_1 = 0$ $A_{11} = 9$

$A_2 = 5$ $A_{12} = 3$

$A_3 = 7$ $A_{13} = 0$

$A_4 = 6$ $A_{14} = 5$

$A_5 = 8$ $A_{15} = 1$

$A_6 = 12$ $A_{16} = 0$

$A_7 = 4$ $A_{17} = 0$

$A_8 = 0$ $A_{18} = 0$

$A_9 = 4$ $A_{19} = 0$

$A_{10} = 4$

So

$$A = \sum_{i=1}^{19} A_i = \underline{68}$$

For $\alpha = 0{,}05$ level of significance we have

$$A_{20;\,0{,}975} < A_{test} < A_{20;\,0{,}025}$$

$$64 < A_{test} < 125$$

We see that

$$64 < A = 68 < 125$$

Therefore, the hypothesis of stationarity is failed to be rejected at the 5% level of significance.

For $\alpha = 0{,}1$ we have

$$A_{20;\,0{,}95} < A_{test} < A_{20;\,0{,}05}$$

$$69 < A_{test} < 120$$

We see that

$$A = 68 < 69 < A_{test} < 120$$

Hence, the hypothesis of stationarity is rejected at the 10% level of significance.

2b)

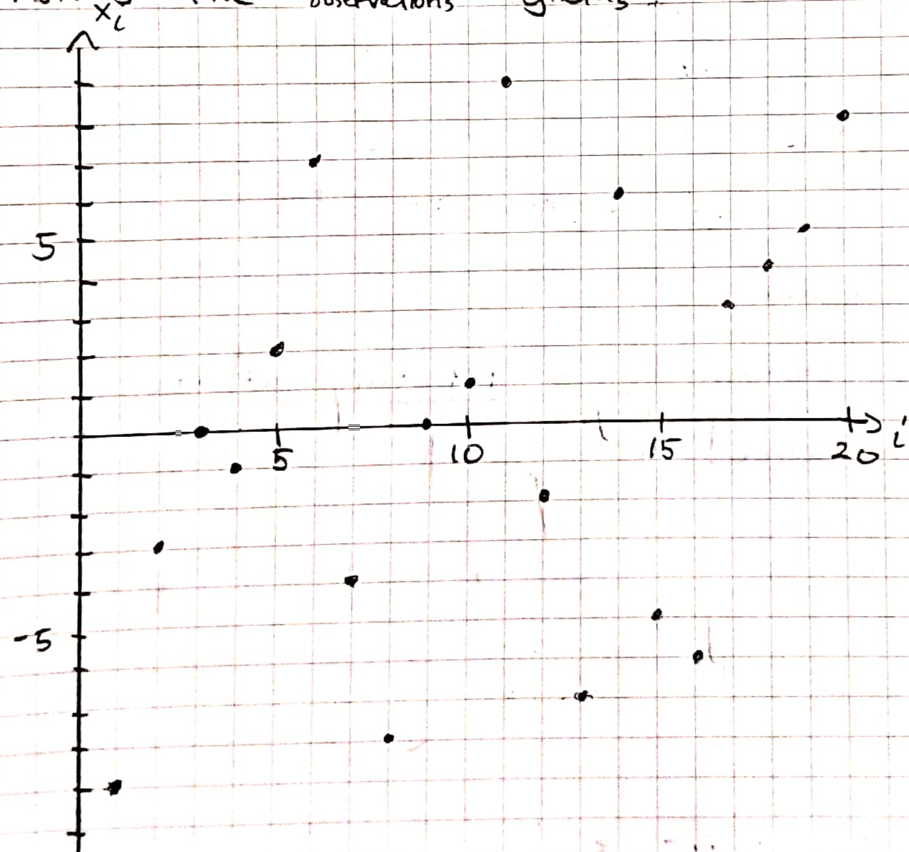
The hypothesis of stationarity is failed to be rejected when there is 5% chance of making a Type I error.

However, the hypothesis of stationarity is rejected when there is a 10% chance of making a Type I error.

Based on this, it may look like that the random variable is stationary and hence no trend is embedded in the variable.

The probability of getting these observations, given stationarity, lies between 5~10%.

Plotting the observations yields



Looking at the plot, it may be argued that a small monotonic trend is embedded in the variable. Hence, the random variable is non-stationary, which is in contrast to the conclusion of the hypothesis test with level of significance equal to 5%.

The observations are slightly increasing in value?

Problem 3a)

Defining a new random variable Z, which is the sum of the normal distributed variable that represents amount of water.

Hence

$$X \sim \mathcal{N}(50 \cdot \mu, 50 \cdot \sigma^2) = \mathcal{N}(50 \cdot 2, 50 \cdot 0{,}7^2) = \mathcal{N}(100, 24{,}5)$$

Note that it is assumed that the summed variables are <u>independent</u>

The total amount of water that will be consumed on the trip is then given by

$$X \sim \mathcal{N}(100, 24{,}5)$$

So

$$P(\text{run out of water}) = 1 - P(\text{not run out of water})$$

Will standardize the normal distributed variable X

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 100}{\sqrt{24{,}5}} \sim \mathcal{N}(0,1)$$

Then

$$P(\text{not run out of water}) = P(X \leq 110)$$
$$= P\left(Z \leq \frac{110 - 100}{\sqrt{24{,}5}}\right)$$
$$= P(Z \leq 2{,}02)$$
$$= \underline{\underline{0{,}9783}}$$

Ultimately

$$P(\text{run out of water}) = 1 - 0{,}9783 = 0{,}0217 = \underline{\underline{2{,}17\ \%}}$$

If we were to remove the assumption of independent variables, we could calculate

$$P\left(\begin{array}{c}\text{average water use}\\ \text{per man}\end{array} > 2{,}2\ L\right)$$

instead

3b)

Since the population (not injected rats) variance is unknown, we must use the t-distribution. Hence, our test statistic is

$$T_0 = \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t_{n-1}$$

Defining the hypothesis

H_0: The drug has no effect

H_1: The drug has an effect

This can be expressed as

$H_0: \mu = 1,2$

$H_1: \mu \neq 1,2$

Hence, our test statistic is (assuming H_0 is true)

$$T_0 = \frac{1,05 - 1,2}{\frac{0,5}{\sqrt{100}}} = \underline{-3}$$

Having $DoF = n-1 = 99$ and $\alpha = 0,05$; the table gives the
(95% confidence interval)
following value

$T = 1,984$

We see that

$T_0 > |T| \Rightarrow T_0 = -3 < -T = -1,984$

Thus, we reject the null-hypothesis.
With a 5% level of significance, the drug has an effect.

Problem 4a)

The degree of freedom is defined as

$$n = K - 3$$

where K is the number of intervals. Thus

$$n = 10-3 = 7$$

4b)

Interval No.	Interval	P Probability	$F = NP$ Expected frequency	f Observed frequency	$\frac{(F-f)^2}{F}$ Discrepancy
1	0-9	0,0617	6,17	9	1,2980
2	10-19	0,0673	6,73	4	1,1074
3	20-29	0,1064	10,64	14	1,0610
4	30-39	0,1424	14,24	10	1,2625
5	40-49	0,1615	16,15	17	0,0447
6	50-59	0,1552	15,52	12	0,7983
7	60-69	0,1264	12,64	11	0,2128
8	70-79	0,0872	8,72	11	0,5961
9	80-89	0,051	5,1	8	1,6490
10	90-100	0,0411	4,11	4	0,0029

$$\chi^2 = 8,03$$

- The probability is calculated from a normal distribution (assuming the variable is normal distributed).

e.g. Interval 1: $P(x \leq 10)$

Interval 2: $P(10 \leq x \leq 20)$

⋮

Interval 10: $P(x \geq 90)$

With the level of significance at 5%, we have

$$\chi^2_{7;0,05} = 14,07$$

Since

$$X^2 = 8,03 < \chi^2_{7;0,05} = 14,07$$

we fail to reject the hypothesis of normal distribution.

Hence, we can conclude that the random variable X is normal distributed.