Homework Assignment 1

Alfred Bornefalk
Simen Söderberg

7 February 2022

Contents

1 Random Number Generation **3**
1.1 Part A: Conditional Distribution Function & Density 3
1.2 Part B: Inverse Conditional Distribution Function 3

2 Power Production of a Wind Turbine **4**
2.1 Part A: Truncated Weibull Distribution 4
2.2 Part B: Control Variate 6
2.3 Part C: Means of Importance Sampling 8
2.4 Part D: Antithetic Sampling 9
2.5 Part E: Probability Estimation 11
2.6 Part F: Average Ratio of Output 12
2.7 Part G: Capacity Factor & Availability Factor 12

3 Combined Power Production of Two Wind Turbines **13**
3.1 Part A: Expected Combined Power Generation 13
3.2 Part B: Covariance 14
3.3 Part C: Variability & Standard Deviation 15
3.4 Part D: 95 % Confidence Intervals 16

4 References **22**

1 Random Number Generation

In the first part of the home assignment, we are asked to find different expressions. f_X, F_X and F_X^-1 are assumed to be known, and $I = (a, b)$ is an interval such that $\mathbb{P}(X \in I) > 0$.

1.1 Part A: Conditional Distribution Function & Density

The expression for the conditional distribution function is obtained through

$$F_{X|X\in I}(x) = \mathbb{P}(X \leq x|X \in I) = \mathbb{P}(X \leq x|a < X < b) = \frac{\mathbb{P}(X \leq x, a < X < b)}{\mathbb{P}(a < X < b)}$$

If $x \leq a$, then $F_{X|X\in I}(x) = 0$. On the other hand, if $a < X < b$, we have

$$F_{X|X\in I}(x) = \frac{\mathbb{P}(a < X < x)}{\mathbb{P}(a < X < b)} = \frac{F_X(x) - F_X(a)}{F_X(b) - F_X(a)}$$

Lastly, for $x \geq b$, $F_{X|X\in I}(x) = 1$. Hence, we obtain

$$F_{X|X\in I}(x) = \begin{cases} 1, & \text{if } x \geq b \\ \frac{F_X(x)-F_X(a)}{F_X(b)-F_X(a)}, & \text{if } a < x < b \\ 0, & \text{if } x \leq a \end{cases} \tag{1}$$

The density function is obtained through simple derivation of the conditional density function. For $a < X < b$, we get

$$f_{X|X\in I} = \frac{d}{dx}(F_{X|X\in I}) = \frac{d}{dx}\left(\frac{F_X(x)}{F_X(b) - F_X(a)} - \frac{F_X(a)}{F_X(b) - F_X(a)}\right) = \frac{f_X(x)}{F_X(b) - F_X(a)}$$

For the other cases we derive a constant, thus the probability density function is zero. Finally, we have

$$f_{X|X\in I}(x) = \begin{cases} \frac{f_X(x)}{F_X(b)-F_X(a)}, & \text{if } a < x < b \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

For (1) and (2), it is worth noting that we do not need to be careful about the end points, i.e. changing $a < X < b$ to $a \leq X \leq b$ does not make a difference. This is attributed to the fact that X is a continuous random variable.

1.2 Part B: Inverse Conditional Distribution Function

To derive the inverse conditional distribution function we begin by noting that

$$F_X(F_X^{-1}(x)) = x \tag{3}$$

We begin by rewriting (1) for $a < x < b$

$$F_{X|X\in I}(x) = \frac{F_X(x) - F_X(a)}{F_X(b) - F_X(a)} = \frac{F_X(x) - a}{b} \tag{4}$$

Utilizing (3) and (4) we get the following

$$F_{X|X\in I}(F_X^{-1}(x)) = \frac{x-a}{b} \iff$$

$$F_{X|X\in I}^{-1}(\frac{x-a}{b}) = F_{X|X\in I}^{-1}(F_{X|X\in I}(F_X^{-1}(x))) \iff$$

$$F_{X|X\in I}^{-1}(\frac{x-a}{b}) = F_X^{-1}(x)$$

By introducing the parameter $u = \frac{x-a}{b}$ we finally get the expression

$$F_{X|X\in I}^{-1}(u) = F_X^{-1}(a+ub) \iff$$

$$F_{X|X\in I}^{-1}(u) = F_X^{-1}(F_X(a) + u(F_X(b) - F_X(a)) \tag{5}$$

When working with a monotonically increasing and right continuous $F_X(x)$ the generalized inverse of the cumulative distribution function is equal to

$$F_X^{\leftarrow}(u) = F_X^{-1}(u)$$

Applying this to our conditional cumulative distribution function enables us to sample X through the inversion sampling algorithm

draw $u \sim U(0,1)$
set $X \leftarrow F_{X|X\in I}^{\leftarrow}(u)$
return X

This way we can obtain random samples from a given distribution.

2 Power Production of a Wind Turbine

In the second section of the home assignment, we are investigating the potential of building the wind turbine D236 at a given site. The turbine will be exposed to the wind V, denoted by X in the upcoming subsections. The air density ρ at the location of the turbine is approximately 1.225 kg/m^3 and its rotor diameter d is equal to 236 m.

2.1 Part A: Truncated Weibull Distribution

In this part of the home assignment, we are comparing the 95 % confidence interval for the total amount of power [W] in the wind passing a wind turbine with a rotor diameter d [m]. The power function in our case is given by

$$P_{tot}(x) = \frac{1}{2}\rho\pi\frac{d^2}{4}x^3$$

We begin by examining the confidence interval using the standard Monte Carlo method. That is, we let $X_1, X_2, ..., X_n$ be random independent variables with density f. Here, f is the density function of the Weibull distribution. Then, by the Law of Large Numbers, as N tends to infinity,

$$\tau_N \equiv \frac{1}{N}\sum_{n=1}^{N}\phi(X_i) \to E(\phi(X_i)) \tag{6}$$

The result in (6) is the basis for the basic Monte Carlo sampler:

for $i = 1 \to N$ do
  draw $X_i \sim f$
end for
set $\tau_N \leftarrow \sum_{n=1}^{N}\frac{\phi(X_i)}{N}$
return τ_N

Through the implications of the central limit theorem (CLT), the two-sided confidence interval at level $(1-\alpha)$ is obtained by

$$I_\alpha = \tau_N \pm \lambda_{\alpha/2}\frac{\sigma(\phi)}{\sqrt{N}} \tag{7}$$

The stochastic wind speeds passing the wind turbine are modeled by the following Weibull distribution:

$$f(x) = \frac{k}{\lambda}(\frac{x}{\lambda})^{k-1}exp(-(\frac{x}{\lambda})^k), x \geq 0 \tag{8}$$

Applying the Monte Carlo Sampler with (6), (7), (8) and the file "powercurve_D236" in MATLAB with sample size $N = 10^6$ yields the 95 % confidence interval for the expected amount of power generated by the wind turbine with draws from our Weibull distribution. The results are displayed in table 1.

Table 1: 95 % confidence interval for the total amount of power, measured in W, in the wind passing a wind turbine with a rotor diameter of 236 m using the standard Monte Carlo method with draws from a Weibull distribution.

Month	Lower Bound	Upper Bound	Width
January	$9.33 \cdot 10^6$	$9.35 \cdot 10^6$	$2.35 \cdot 10^4$
February	$8.57 \cdot 10^6$	$8.59 \cdot 10^6$	$2.37 \cdot 10^4$
March	$8.03 \cdot 10^6$	$8.05 \cdot 10^6$	$2.36 \cdot 10^4$
April	$6.66 \cdot 10^6$	$6.68 \cdot 10^6$	$2.32 \cdot 10^4$
May	$6.44 \cdot 10^6$	$6.47 \cdot 10^6$	$2.29 \cdot 10^4$
June	$6.77 \cdot 10^6$	$6.80 \cdot 10^6$	$2.32 \cdot 10^4$
July	$6.45 \cdot 10^6$	$6.47 \cdot 10^6$	$2.30 \cdot 10^4$
August	$6.77 \cdot 10^6$	$6.80 \cdot 10^6$	$2.33 \cdot 10^4$
September	$7.95 \cdot 10^6$	$7.97 \cdot 10^6$	$2.36 \cdot 10^4$
October	$8.60 \cdot 10^6$	$8.63 \cdot 10^6$	$2.39 \cdot 10^4$
November	$9.34 \cdot 10^6$	$9.36 \cdot 10^6$	$2.35 \cdot 10^4$
December	$9.32 \cdot 10^6$	$9.35 \cdot 10^6$	$2.35 \cdot 10^4$

When instead using the truncated Weibull distribution with the lower limit $x = 3$ m/s and the upper limit $x = 30$ m/s, we leverage the inversion sampling algorithm and the result in (5). By weighting the power generated by the power function with the wind sampled from our inversion sampling with the probability of getting a sample in our interval $I = (3, 30)$ we get the confidence intervals on the 95 % level shown in table 2.

Table 2: 95 % confidence interval for the total amount of power, measured in W, in the wind passing a wind turbine with a rotor diameter of 236 m using the inversion method with the truncated Weibull distribution.

Month	Lower Bound	Upper Bound	Width
January	$9.32 \cdot 10^6$	$9.34 \cdot 10^6$	$2.07 \cdot 10^4$
February	$8.56 \cdot 10^6$	$8.58 \cdot 10^6$	$2.08 \cdot 10^4$
March	$8.04 \cdot 10^6$	$8.06 \cdot 10^6$	$2.06 \cdot 10^4$
April	$6.66 \cdot 10^6$	$6.68 \cdot 10^6$	$1.97 \cdot 10^4$
May	$6.45 \cdot 10^6$	$6.47 \cdot 10^6$	$1.95 \cdot 10^4$
June	$6.77 \cdot 10^6$	$6.79 \cdot 10^6$	$1.98 \cdot 10^4$
July	$6.45 \cdot 10^6$	$6.47 \cdot 10^6$	$1.95 \cdot 10^4$
August	$6.78 \cdot 10^6$	$6.80 \cdot 10^6$	$1.98 \cdot 10^4$
September	$7.94 \cdot 10^6$	$7.96 \cdot 10^6$	$2.06 \cdot 10^4$
October	$8.60 \cdot 10^6$	$8.62 \cdot 10^6$	$2.07 \cdot 10^4$
November	$9.32 \cdot 10^6$	$9.34 \cdot 10^6$	$2.07 \cdot 10^4$
December	$9.32 \cdot 10^6$	$9.34 \cdot 10^6$	$2.07 \cdot 10^4$

From table 1 and 2 we can clearly see that the width of the confidence intervals for each month is larger using the standard Monte Carlo method compared to the truncated version. This difference is explained by the fact that the truncated version never yields zero power from the power function since the input value is in the interval of [3, 30] m/s. However, for the standard Monte Carlo simulation there is a possibility of wind speeds both greater than 30 m/s and less than 3 m/s resulting in a corresponding power value of zero. This leads to a greater variance which then explains the greater width. Moreover we can see that the expected value is similar both for the standard Monte Carlo method and the truncated version. Naturally, this is a direct result of weighting the truncated Weibull distribution's expected power generation with the probability of actually getting a sample in said range.

2.2 Part B: Control Variate

For this part of the home assignment, we are asked to use the wind X as a control variate to decrease the variance, to thereafter estimate a 95 % confidence interval for the expected power in question. The real-valued random variable X satisfies the conditions for being a control variate Y since (i) $\mathbb{E}(Y) = m$ is known for each month and (ii) $\phi(X)$ and Y can be simulated at the same complexity as $\phi(X)$. This allows us to set, for some $\beta \in R$,

$$Z = \phi(X) + \beta(Y - m)$$

such that

$$\mathbb{E}(Z) = \mathbb{E}(\phi(X)) + \beta(\mathbb{E}(Y) - m) = \tau$$

Since $\mathbb{E}(Y) = m$, the expression can be simplified to

$$\mathbb{E}(Z) = \mathbb{E}(\phi(X)) = \tau \tag{9}$$

The variance of Z is defined by

$$\mathbb{V}(Z) = \mathbb{V}(\phi(X) + \beta Y) = \mathbb{V}(\phi(X)) + 2\beta\mathbb{C}(\phi(X), Y) + \beta^2 V(Y)$$

Differentiating the above, with $\phi(X)$ and Y having the covariance $\mathbb{C}(\phi(X), Y)$, yields the optimal β^*, in terms of variance reduction, through

$$0 = 2\mathbb{C}(\phi(X), Y) + 2\beta\mathbb{V}(Y) \iff \beta^* = -\frac{\mathbb{C}(\phi(X), Y)}{\mathbb{V}(Y)} \tag{10}$$

Table 3 illustrates the confidence intervals at the fixed confidence level of 95 % when applying (9) and (10) in MATLAB.

Table 3: 95 % confidence interval for the total amount of power, measured in W, in the wind passing a wind turbine with a rotor diameter of 236 m using the wind V as a control variate to reduce variance. The sample correlation r between the wind power function and the control variate is included as well.

Month	Lower Bound	Upper Bound	Width	Correlation
January	$9.33 \cdot 10^6$	$9.34 \cdot 10^6$	$1.26 \cdot 10^4$	.843
February	$8.57 \cdot 10^6$	$8.58 \cdot 10^6$	$1.12 \cdot 10^4$	.882
March	$8.03 \cdot 10^6$	$8.04 \cdot 10^6$	$1.03 \cdot 10^4$	.900
April	$6.68 \cdot 10^6$	$6.68 \cdot 10^6$	$.882 \cdot 10^4$	.924
May	$6.45 \cdot 10^6$	$6.46 \cdot 10^6$	$.855 \cdot 10^4$	.928
June	$6.78 \cdot 10^6$	$6.79 \cdot 10^6$	$.898 \cdot 10^4$	.922
July	$6.45 \cdot 10^6$	$6.46 \cdot 10^6$	$.856 \cdot 10^4$	.928
August	$6.77 \cdot 10^6$	$6.78 \cdot 10^6$	$.899 \cdot 10^4$	.922
September	$7.94 \cdot 10^6$	$7.95 \cdot 10^6$	$1.02 \cdot 10^4$	.902
October	$8.60 \cdot 10^6$	$8.61 \cdot 10^6$	$1.21 \cdot 10^4$	.865
November	$9.33 \cdot 10^6$	$9.34 \cdot 10^6$	$1.26 \cdot 10^4$	.843
December	$9.34 \cdot 10^6$	$9.36 \cdot 10^6$	$1.26 \cdot 10^4$	.844

As we can see from table 3, using the control variate method reduces the width of our confidence interval. The control variate method will yield larger variance reductions if the correlation between $\phi(X)$ and Y is close to 1. This can be observed by plugging in the optimal β^* from (10) in the variance formula for Z

$$\mathbb{V}(Z) = \mathbb{V}(\phi(X)) + 2\beta^*\mathbb{C}(\phi(X), Y) + \beta^{*2}\mathbb{V}(Y) = \mathbb{V}(\phi(X))(1 - r(\phi(X), Y))$$

Observing the values of r from table 3 it can be seen that the correlation between the power $(\phi(X))$ and the wind (Y) is strong, thereby explaining the reduction of the width. This is intuitive since the amount of power generated is a function of the wind.

2.3 Part C: Means of Importance Sampling

In part C, we investigate if we can reduce the width of the confidence interval by using the importance sampling technique. We aim to find an instrumental density $g(x)$ on X that we can simulate random variables from with the property

$$g(x) = 0 \implies \phi(x)f(x) = 0$$

We aim to find a $g(x)$ that is directly proportional to the product of $\phi(x)$ and $f(x)$; this is explained by the following equation:

$$\tau = \mathbb{E}_f(\phi(x)) = \int_X \phi(x)f(x)dx = \int_{g(x)>0} \phi(x)\frac{f(x)}{g(x)}g(x)dx = \mathbb{E}_g(\phi(X)\frac{f(X)}{g(X)})$$

If $g(x)$ is directly proportional to the product of $\phi(x)$ and $f(x)$ the equation holds, thereby giving us a "perfect" estimate of τ. This is however hard to obtain. Nevertheless, the objective is to find a $g(x)$ that makes this ratio as constant as possible. We estimate our $\tau = \mathbb{E}_g(\phi(X)\frac{f(X)}{g(X)})$ by using the standard Monte Carlo method

for $i = 1 \to N$ do
  draw $X_i \sim g$
end for
set $\tau_N \leftarrow \frac{1}{N}\sum_{n=1}^{N} \phi(X)\frac{f(X)}{g(X)}$
return τ_N

To determine our $g(x)$ we plotted $\phi(x)f(x)$ and observed that the joint distribution looked similar to a normal distribution, albeit somewhat right skewed.

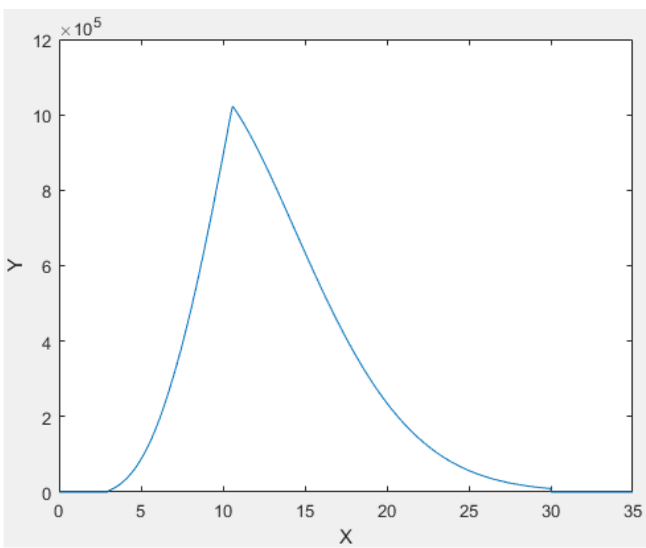

Figure 1: 2D plot of $\phi(x)f(x)$. The values in the graph are from December.

From figure 1, we decided that the normal fit was good enough for the modeling purpose, and that our instrumental density therefore should be drawn from a normal distribution. The mean for our instrumental density was chosen to be the corresponding x value to our maximum value from our joint distribution, which for all months yielded $\mu_g = 10.6$. We minimized our mean width of the confidence interval to determine the optimal sigma (down to 3 significant figures) for $g(x)$. This led us to sample with $\sigma_g = 5.08$. Generating random samples from $g(x)$ then gave the following results presented in table 4.

Table 4: 95 % confidence interval for the total amount of power, measured in W, in the wind passing a wind turbine with a rotor diameter of 236 m using importance sampling.

Month	Lower Bound	Upper Bound	Width
January	$9.33 \cdot 10^6$	$9.36 \cdot 10^6$	$2.27 \cdot 10^4$
February	$8.57 \cdot 10^6$	$8.59 \cdot 10^6$	$1.62 \cdot 10^4$
March	$8.04 \cdot 10^6$	$8.05 \cdot 10^6$	$1.46 \cdot 10^4$
April	$6.66 \cdot 10^6$	$6.68 \cdot 10^6$	$1.26 \cdot 10^4$
May	$6.45 \cdot 10^6$	$6.46 \cdot 10^6$	$1.25 \cdot 10^4$
June	$6.78 \cdot 10^6$	$6.79 \cdot 10^6$	$1.26 \cdot 10^4$
July	$6.45 \cdot 10^6$	$6.46 \cdot 10^6$	$1.25 \cdot 10^4$
August	$6.78 \cdot 10^6$	$6.79 \cdot 10^6$	$1.26 \cdot 10^4$
September	$7.94 \cdot 10^6$	$7.95 \cdot 10^6$	$1.44 \cdot 10^4$
October	$8.60 \cdot 10^6$	$8.62 \cdot 10^6$	$1.83 \cdot 10^4$
November	$9.33 \cdot 10^6$	$9.35 \cdot 10^6$	$2.25 \cdot 10^4$
December	$9.33 \cdot 10^6$	$9.35 \cdot 10^6$	$2.26 \cdot 10^4$

As we can see from table 4, the width decreased for all months compared to the standard Monte Carlo from table 1. However, comparing the width with the control variate method from table 3, we realize that the width using importance sampling is greater for all months. It is worth to note that we might not have chosen the optimal $g(x)$. That being said, our $g(x)$ is easy to simulate from and decreases the width of τ. Thus, our choice of $g(x)$ is improving the results from our original estimation from table 1.

2.4 Part D: Antithetic Sampling

In this exercise, we are to use the variance reduction technique antithetic sampling. As is usually the case, we wish to estimate $\tau = \mathbb{E}_f(\phi(x))$ by means of Monte Carlo. For simplicity, we let $V \stackrel{\text{def}}{=} \phi(x)$, so that $\tau = E(V)$. It is now assumed that we can generate another variable $\widetilde{V}$ such that (i) $E(\widetilde{V}) = \tau$, (ii) $\mathbb{V}(\widetilde{V}) = \mathbb{V}(V)$, (iii) $\widetilde{V}$ can be simulated at the same complexity as V. Then, for

$$W \stackrel{\text{def}}{=} \frac{V + \widetilde{V}}{2}$$

it holds that $\mathbb{E}(W) = \tau$ and

$$\mathbb{V}(W) = \mathbb{V}(\frac{V + \widetilde{V}}{2}) = \frac{1}{4}(\mathbb{V}(V) + 2\mathbb{C}(V, \widetilde{V}) + \mathbb{V}(\widetilde{V})) = \frac{1}{2}(\mathbb{V}(V) + \mathbb{C}(V, \widetilde{V}))$$

It is to be noted that if we can find $\widetilde{V}$ such that the antithetic variables V and $\widetilde{V}$ are negatively correlated, that is if

$$\mathbb{C}(V, \widetilde{V})) < 0 \tag{11}$$

then we will gain computational work. For this purpose, the theorem mentioned by Wiktorsson (2022, p.21) is very useful. We can apply this theorem to antithetic sampling by letting F be a distribution function and ϕ a monotone function. Then, letting $U \sim U(0,1)$, $T(u) = 1 - u$, and $\varphi(u) = \phi(F^{-1}((u))$ yields, for $V = \phi(F^{-1}((U))$ and $V = \phi(F^{-1}((1-U))$,

$$V \equiv \widetilde{V}$$

and

$$\mathbb{C}(V, \widetilde{V}) \leq 0$$

Since our case applies to the above scenario, and (11) is satisfied, we are able to use antithetic sampling in MATLAB as a means for variance reduction. The effect on the confidence intervals are exhibited in table 5.

Table 5: 95 % confidence interval for the total amount of power, measured in W, in the wind passing a wind turbine with a rotor diameter of 236 m using antithetic sampling to reduce variance.

Month	Lower Bound	Upper Bound	Width
January	$9.33 \cdot 10^6$	$9.34 \cdot 10^6$	$.090 \cdot 10^4$
February	$8.57 \cdot 10^6$	$8.58 \cdot 10^6$	$.462 \cdot 10^4$
March	$8.03 \cdot 10^6$	$8.04 \cdot 10^6$	$.291 \cdot 10^4$
April	$6.68 \cdot 10^6$	$6.68 \cdot 10^6$	$.625 \cdot 10^4$
May	$6.45 \cdot 10^6$	$6.46 \cdot 10^6$	$.696 \cdot 10^4$
June	$6.78 \cdot 10^6$	$6.79 \cdot 10^6$	$.590 \cdot 10^4$
July	$6.45 \cdot 10^6$	$6.46 \cdot 10^6$	$.696 \cdot 10^4$
August	$6.77 \cdot 10^6$	$6.78 \cdot 10^6$	$.589 \cdot 10^4$
September	$7.94 \cdot 10^6$	$7.95 \cdot 10^6$	$.283 \cdot 10^4$
October	$8.60 \cdot 10^6$	$8.61 \cdot 10^6$	$.533 \cdot 10^4$
November	$9.33 \cdot 10^6$	$9.34 \cdot 10^6$	$.898 \cdot 10^4$
December	$9.34 \cdot 10^6$	$9.36 \cdot 10^6$	$.902 \cdot 10^4$

When comparing the width of the confidence intervals in table 5 to the results from the inversion sampling, control variate and importance sampling, it is evident that antithetic sampling is the best technique to reduce the variance of our estimated power output so far. For instance, in January, the width obtained by the standard Monte Carlo method is over 25 times the size of the antithetic's.

2.5 Part E: Probability Estimation

In this part of the home assignment, the probability that the turbine delivers power, i.e. the probability that

$$P(X) > 0$$

is to be estimated. This is achieved by generating wind speeds through the standard Monte Carlo procedure, and then calculating the quotient of non-zero power outputs for the respective month. Letting A denote the event that $P(X) > 0$, the equation for the probability is as follows:

$$\mathbb{P}(A) = \frac{N_A}{N} \tag{12}$$

Implementing (11) in MATLAB yields the values displayed in table 6.

Table 6: Estimation of the probability of power generation using the standard Monte Carlo method with draws from a Weibull distribution.

Month	Probability of Power
January	.935
February	.924
March	.916
April	.878
May	.873
June	.881
July	.873
August	.880
September	.914
October	.916
November	.935
December	.935

From table 6, one can observe that the probability of power generation differs with as much as six percentage points between the highest (January, November and December) and the lowest (May and July) probabilities. Comparing this result with table 1 suggests that the higher the total expected amount of power output from the wind turbine, the higher the probability of actual power generation. Since the wind turbine produces no power outside of the interval $[3, 30]$ m/s, and the average wind speeds are higher for January, November and December than the case of May and July, it seems intuitive that wind speeds in the lower end of the spectrum are inhibiting the output more than higher ones for our renewable energy source.

2.6 Part F: Average Ratio of Output

In this part of the exercises we examine the average power coefficient, denoted by PC, and defined as

$$PC = \frac{\mathbb{E}[P(X)]}{\mathbb{E}[P_{tot}(X)]} \tag{13}$$

Where $P_{tot}(x) = \frac{1}{2}\rho\pi\frac{d^2}{4}x^3$. For our wind turbine, the rotor diameter d is 236 m and the air density ρ is 1.225 kg/m^3. Since we have $\mathbb{E}[X^m] = \Gamma(1+\frac{3}{m})\lambda^m$, we can calculate P_{tot} for each month as

$$\mathbb{E}[P_{tot}] = \frac{1}{2}\rho\pi\frac{d^2}{4}\mathbb{E}[X^3] = \frac{1}{2}\rho\pi\frac{d^2}{4}\Gamma(1+\frac{3}{k})\lambda^3$$

Obtaining the nominator in (13) is achieved with the standard Monte Carlo simulation that we have used previously, whereas the denominator is calculated with the built-in gamma function in MATLAB. The 95 % confidence intervals are presented in table 7.

Table 7: 95 % confidence interval for the average ratio of output in the wind passing a wind turbine with a rotor diameter of 236 m.

Month	Lower Bound	Upper Bound	Width
January	.163	.164	.000
February	.196	.197	.001
March	.219	.220	.001
April	.259	.260	.001
May	.269	.270	.001
June	.255	.256	.001
July	.269	.270	.001
August	.255	.256	.001
September	.223	.224	.001
October	.176	.176	.001
November	.163	.164	.000
December	.164	.164	.000

From table 7 we can observe that the average ratio of output varies between 16 and 27 %, where the months with the lower total amount of power output have the higher ratio. This can be attributed to the fact that the scaling parameter λ is lower during these months, whilst the shape parameter k remains rather constant (shifting between 1.9 and 2.0).

2.7 Part G: Capacity Factor & Availability Factor

In the last part of this section of the home assignment, we are asked to examine two important characteristics of the wind turbine, meaningly its *capacity factor* and *availability factor*. The capacity factor is defined as the ratio between the actual power output and the maximum power output. For the wind turbine in question, the value of the latter is 15 MW. The availability factor is the fraction of time that the plant actually produces output. Thus, we can simply reuse the results obtained in part E. Calculating both factors for each month and then taking the averages in MATLAB yields the following table.

Table 8: Capacity and availability factor for the site of the wind turbine.

Factor	Value
Capacity	.524
Availability	.905

From table 8, we can see that the capacity factor at the site is 52.4 % and that the availability factor is 90.5 %. Considering that wind turbines typically have a capacity factor of 20 to 40 % and an availability factor above 90 %, this does seem to be a good site to build a wind turbine. The availability factor is acceptable, whilst the capacity factor far exceeds the typical case.

3 Combined Power Production of Two Wind Turbines

In the third section of the home assignment, we are studying two wind turbines placed in the same area. The turbines will be exposed to similar winds V_1 and V_2; these are denoted by X and Y, respectively, in the forthcoming subsections.

3.1 Part A: Expected Combined Power Generation

In this exercise we calculate the expected combined power generation from two wind turbines that are located in the same area. The winds that the wind turbines are exposed by, X and Y, are both generated from the same instrumental density $g(x)$. By reusing the same instrumental density from part 2c, we draw from a normal distribution with $\sigma = 5.08$ and $\mu = 10.6$. This is possible due to the problem being one dimensional, as a consequence of the fact that the sum of the expected value of two random variables is equal to the sum of their expectations, e.g.

$$\mathbb{E}[P(X) + P(Y)] = \mathbb{E}[P(X)] + \mathbb{E}[P(Y)]$$

Calculating the combined expected power generated from our wind turbines with importance sampling in MATLAB yields the following result

$$\mathbb{E}[P(X) + P(Y)] = 15.9 \cdot 10^6$$

Hence, the expected combined power generation under the given circumstances is 15.9 MW.

3.2 Part B: Covariance

The objective in this part of the assignment is to calculate the covariance between the power generated from our wind turbines, i.e. $\mathbb{C}[(P(X), P(Y)]$. The covariance is calculated using the following formula

$$\mathbb{C}[(P(X), P(Y)] = \mathbb{E}[P(X)P(Y)] - \mathbb{E}[P(X)]\mathbb{E}[P(Y)] \tag{14}$$

From part 3a we computed the expected power generated from wind turbine 1 and 2, thus we only need to calculate $\mathbb{E}[P(V_1)P(V_2)]$ to compute the covariance. We define our target function $\phi(X, Y)$ as

$$\phi(X, Y) = P(X)P(Y)$$

To compute the total expected power from the wind turbines we compute the mean of the following equation

$$\phi(X, Y)\frac{f(X, Y)}{g(X, Y)}$$

where we draw our samples X and Y from a multivariate normal distribution $g(X, Y)$ using importance sampling in a similar way as in part 2c. To determine our mean μ and variance-covariance matrix Σ for $g(x, y)$ we plot $\phi(X, Y)f(X, Y)$ as seen in figure 2. From the plot, we observe that we have a peak for X and Y at about 10.960 and thus define μ as

$$\mu = [10.960, 10.960]$$

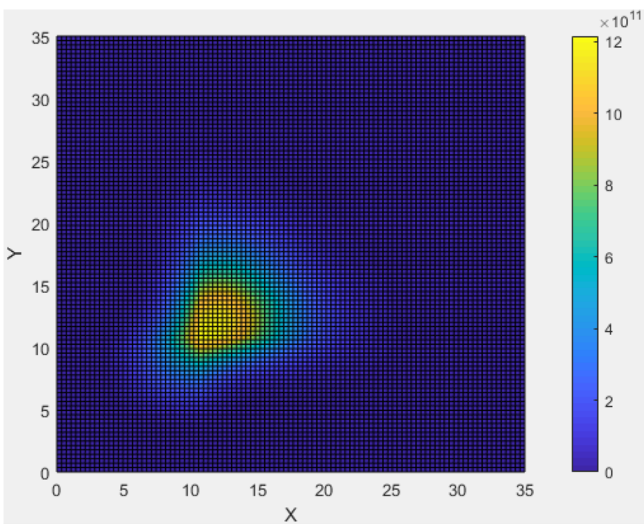

Figure 2: Surface plot of $\phi(X, Y)f(X, Y)$. The colorbar on the right shows a spectrum in which the product is the lowest for the dark blue colors, whereas the maximum value is represented by light yellow.

Determining the value of Σ is done by plotting $\phi(X,Y)\frac{f(X,Y)}{g(X,Y)}$ and testing different values of Σ. Our aim is to find a Σ for $g(X,Y)$ such that

$$g(X,Y) = 0 \rightarrow \phi(X,Y)f(X,Y) = 0$$

By repeatedly plotting the graph in figure 3 with different values for Σ, we find that a suitable variance-covariance matrix for $g(X,Y)$ is

$$\Sigma = \begin{bmatrix} 20 & 1 \\ 1 & 20 \end{bmatrix}$$

The resulting surface plot of $\phi(X,Y)\frac{f(X,Y)}{g(X,Y)}$ given our chosen μ and Σ is presented in figure 3.

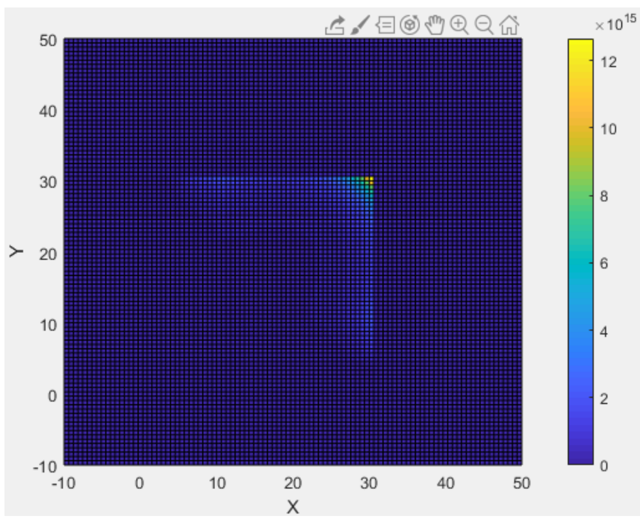

Figure 3: Surface plot of $\phi(X,Y)\frac{f(X,Y)}{g(X,Y)}$. The colorbar on the right shows a spectrum in which the product is the lowest for the dark blue colors, whereas the maximum value is represented by light yellow.

The resulting covariance between the power generated by the wind turbines can then be calculated in MATLAB, using (14). The result is as follows:

$$\mathbb{C}[P(X), P(Y)] = 1.86 \cdot 10^{13}$$

3.3 Part C: Variability & Standard Deviation

In part C the objective is to find the variability $\mathbb{V}[P(V_1) + P(V_2)]$ and the corresponding standard deviation $\mathbb{D}[P(X)+P(Y)]$. The variability is calculated with the formula

$$\mathbb{V}[P(X) + P(Y)] = \mathbb{V}[P(X)] + \mathbb{V}[P(Y)] + 2\mathbb{C}[P(X), P(Y)]$$

The covariance between the wind turbines is already calculated in part 3b. The variance for wind turbine 1 and 2, respectively, is reduced to a one dimensional problem which can be calculated as

$$\mathbb{V}[P(X)] = \mathbb{V}[\phi(X)]$$
$$\mathbb{V}[P(Y)] = \mathbb{V}[\phi(Y)]$$

where X and Y are sampled from the Weibull distribution $f(x)$ using standard Monte Carlo. Inserting the results from the variance of the wind turbines and their covariance yields the variability

$$\mathbb{V}[P(X) + P(Y)] = 1.10 \cdot 10^{14}$$

The standard deviation is then simply calculated by taking the square root of the variability:

$$\mathbb{D}[P(X) + P(Y)] = \sqrt{\mathbb{V}[P(X) + P(Y)]} = 10.5 \cdot 10^{6}$$

Hence, the standard deviation in our case is 10.5 MW.

3.4 Part D: 95 % Confidence Intervals

In the last part of the assignment the aim is to determine the 95 % confidence interval for the probability that the combined power generation of the wind turbines exceeds 15 MW, and vice versa. Hence, the objective is to compute the following two probabilities and their corresponding confidence intervals

$$\mathbb{P}[P(X) + P(Y) > 15 \cdot 10^{6}]$$
$$\mathbb{P}[P(X) + P(Y) < 15 \cdot 10^{6}]$$

We denote our target function as the sum of the power generated from the two wind turbines and set it to one if it meets the required condition above, and zero otherwise:

$$\phi_1(X, Y) = \mathbb{1}_{P(X)+P(Y)>15\cdot 10^{6}}[P(X) + P(Y)]$$
$$\phi_2(X, Y) = \mathbb{1}_{P(X)+P(Y)<15\cdot 10^{6}}[P(X) + P(Y)]$$

To compute the probabilities we use importance sampling and simulate random variables from the multivariate normal probability density function $g(x, y)$. To determine μ_1 for $g(x, y)$ in both cases we plot our target function multiplied by our joint probability density function. For the first case, when the objective is to obtain the probability that the generated total power is greater than 15 MW, we find the suitable value to be

$$\mu_1 = [9.000, 9.000]$$

The chosen μ_1 is determined by examining the surface plot of $\phi_1(X,Y)f(X,Y)$ seen in figure 4 and taking an expected value of the wind that is representative of the plot.

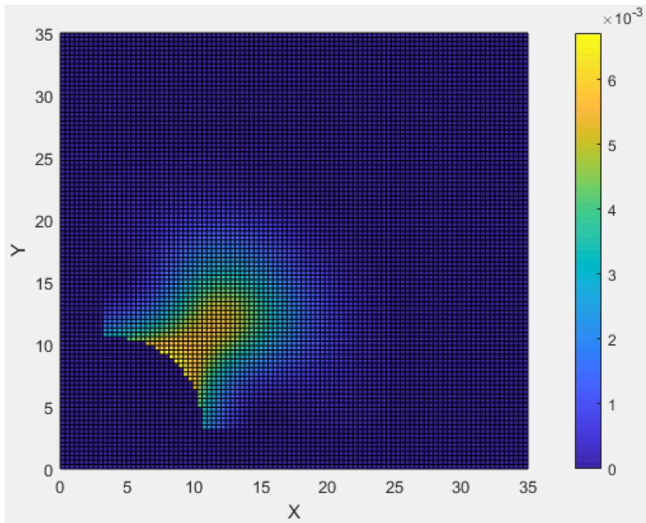

Figure 4: Surface plot of $\phi_1(X,Y)f(X,Y)$. The colorbar on the right shows a spectrum in which the product is the lowest for the dark blue colors, whereas the maximum value is represented by light yellow.

The same process is done for determining the μ_2 for $\phi_2(X,Y)$. The plot is demonstrated in figure 5 and the following values are chosen for μ_2:

$$\mu_2 = [6.000, 6.000]$$

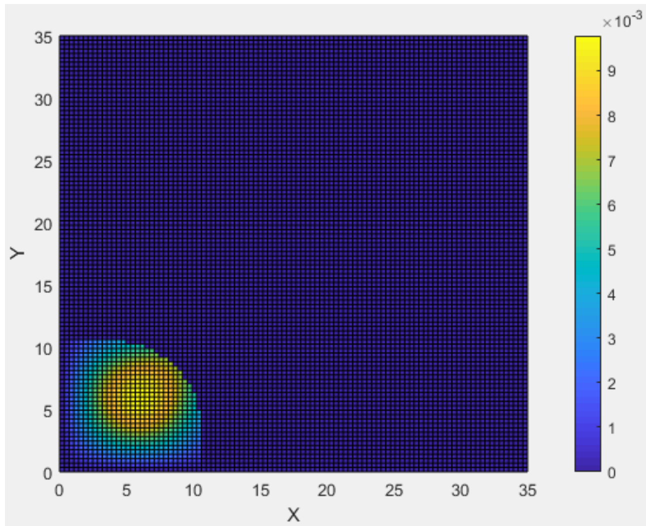

Figure 5: Surface plot of $\phi_2(X,Y)f(X,Y)$. The colorbar on the right shows a spectrum in which the product is the lowest for the dark blue colors, whereas the maximum value is represented by light yellow.

Determining the values of Σ_1 is done in the same way as in part 3b, but this time by plotting $\phi_1(X,Y)\frac{f(X,Y)}{g(X,Y)}$ and testing different values of Σ_1. By repeatedly plotting the graph in figure 6 with different values for our variance, we find that a suitable variance-covariance matrix for $g(X,Y)$ when we want to compute the probability that the power output is greater than 15 MW is

$$\Sigma_1 = \begin{bmatrix} 30 & 1 \\ 1 & 30 \end{bmatrix}$$

In figure 6, we can see that the function behaves well and does not tend to infinity for any samples drawn from $g(X,Y)$ with our chosen μ_1 and Σ_1.

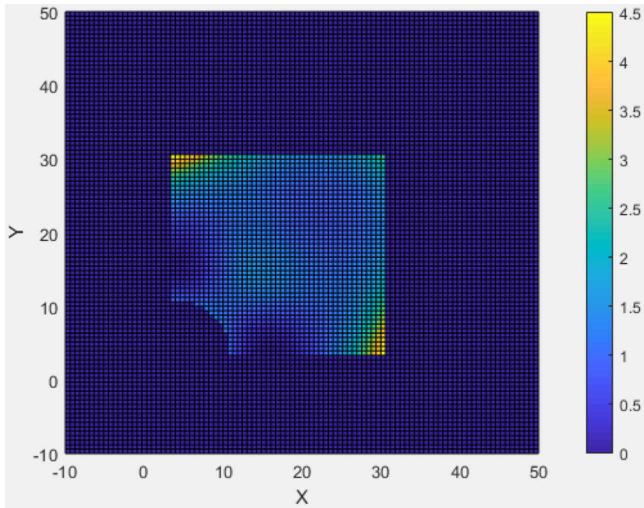

Figure 6: Surface plot of $\phi_1(X,Y)\frac{f(X,Y)}{g(X,Y)}$. The colorbar on the right shows a spectrum in which the product is the lowest for the dark blue colors, whereas the maximum value is represented by light yellow.

To determine the values for our variance-covariance matrix Σ_2 for the case when the total power generated by the wind turbines are less than 15 MW we do the same procedure as for Σ_1 and plot $\phi_2(X,Y)\frac{f(X,Y)}{g(X,Y)}$ for the different values of the variance. The plot can be found in figure 7 and the chosen Σ_2 that yields a well behaved function that does not tend to infinity is

$$\Sigma_2 = \begin{bmatrix} 60 & 1 \\ 1 & 60 \end{bmatrix}$$

Here, our chosen variance is greater. This is due to the fact that our target function $\phi_2(X,Y)$ is equal to one when the wind is outside of the power generating wind interval. To compensate for this we need a greater variance for our chosen instrumental density $g(X,Y)$.

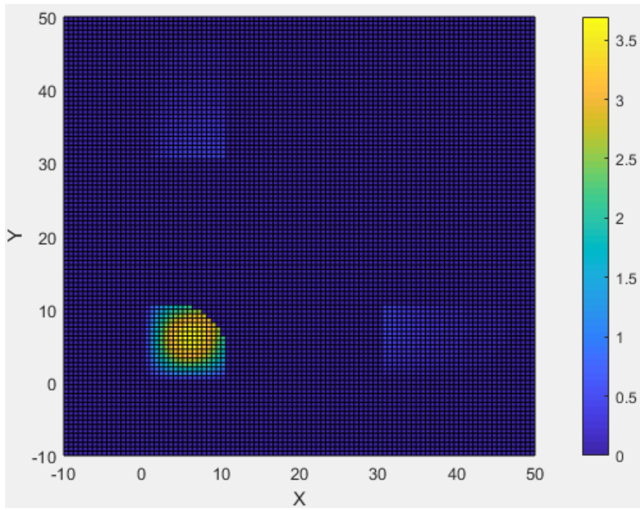

Figure 7: Surface plot of $\phi_2(X,Y)\frac{f(X,Y)}{g(X,Y)}$. The colorbar on the right shows a spectrum in which the product is the lowest for the dark blue colors, whereas the maximum value is represented by light yellow.

To calculate the probability that the combined power generated from the wind turbines are greater or less than 15 MW, we scale our target functions by the importance weight function $\omega = \frac{f(x,y)}{g(x,y)}$ and take the mean of the equations

$$\mathbb{P}[P(X) + P(Y) > 15 \cdot 10^6] = \mathbb{E}[\phi_1(X,Y)\omega]$$

$$\mathbb{P}[P(X) + P(Y) < 15 \cdot 10^6] = \mathbb{E}[\phi_2(X,Y)\omega]$$

By using the standard deviations of our calculations above, MATLAB gives us the confidence intervals in table 9.

Table 9: 95 % confidence interval for the probability that the combined power generation exceeds half of their installed capacity (15 MW), and vice versa.

Probability	Lower	Upper
Over 15 MW	.505	.507
Under 15 MW	.478	.481

A quick glimpse on the upper and lower bounds of the probabilities presented in table 9 suggests that the probabilities do not sum to one. To be more thorough, the probability that the power generated by the two wind turbines combined is either strictly greater than 15 MW or strictly less than 15 MW is

$$\mathbb{P}[P(X) + P(Y) > 15 \cdot 10^6 \cup P(X) + P(Y) < 15 \cdot 10^6] = .985$$

The suggestion is confirmed; the probabilities clearly do not sum to one. This may not be intuitive at a first glance, but is explained by the fact that the power output is exactly 15 MW if one of the wind turbines generates zero power and the other wind turbine generates maximum power. The wind turbines generates zero power when the wind is either less than 3 m/s or greater than 30 m/s, whilst they generate maximum power if the wind is between 10 to 30 m/s. This is illustrated in figure 8, where we plot the total generated power for different combinations of X and Y.

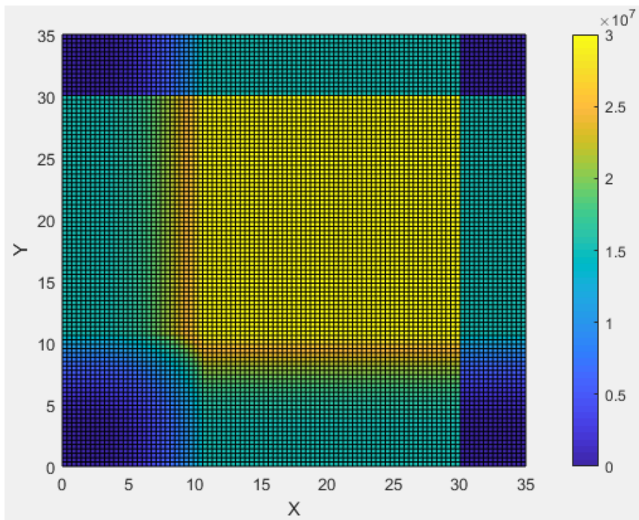

Figure 8: Surface plot of P(X)+P(Y). The colorbar on the right shows a spectrum in which the sum is the lowest for the dark blue colors, whereas the maximum value is represented by light yellow.

4 References

Wiktorsson, M 2022, *Variance reduction for MC methods*, lecture slides, Monte Carlo methods for stochastic inference FMSN50, Lund University, delivered 27 January 2022.