Monte Carlo analysis of Wind Turbine

Nils Hallerfelt

February 2023

1 Random Number Generation

Let X be a random variable on $\mathbb{R}$ with density f_X and invertible distribution function F_X . Here f_X, F_X , and the inverse F_X^{-1} are assumed to be known. Let $I = (a, b)$ be an interval such that $P(X \in I) > 0$.

Given this, and a suitable random number generator from the Uniform distribution $U(0, 1)$ we can generate random numbers from f_X:

Method 1. *Random Number Generation, Inversion Method.*

1. *Generate N samples $U_1, \ldots, U_N$ from $U(0, 1)$.*
2. *For each sample U_i, set $X_i = F_X^{-1}(U_i)$*

Theorem 1. *Each sample X_i generated from method 1 has distribution F_X.*

1.1 The cdf $F_{X|X\in I}$ and pdf $f_{X|X\in I}$

To find the conditional cdf $F_{X|X\in I}$, we can first rewrite it as a probability and use Bayes theorem:

$$F_{X|X\in I} = \mathbb{P}(X < x|X \in I) = \mathbb{P}(X < x|a \leq X \leq b) = \frac{\mathbb{P}(X < x, a \leq X \leq b)}{\mathbb{P}(a \leq X \leq b)} \tag{1}$$

Here we can see three cases, $x < a$, $x > b$, and $x \in I$. For $x < a$ we get that $\mathbb{P}(X < x, a \leq X \leq b) = 0$, and for $x \geq b$ we get that $\mathbb{P}(X < x, a \leq X \leq b) = 1$. For $x \in I$ we can rewrite 1:

$$\frac{\mathbb{P}(X < x, a \leq X \leq b)}{\mathbb{P}(a \leq X \leq b)} = \frac{F_X(x) - F_X(a)}{F_X(b) - F_X(a)} \tag{2}$$

So in conclusion, we get that:

$$F_{X|X\in I} = \begin{cases} x > b & 1 \\ x \in I & \dfrac{F_X(x) - F_X(a)}{F_X(b) - F_X(a)} \\ x < a & 0 \end{cases} \tag{3}$$

In order to the obtain the pdf $f_{X|X\in I}$, we simply take the derivative of $F_{X|X\in I}$ w.r.t. x.

$$f_{X|X\in I} = \begin{cases} x \in I & \dfrac{f_X(x)}{F_X(b) - F_X(a)} \\ \text{otherwise} & 0 \end{cases} \tag{4}$$

1.2 The inverse cdf: $F^{-1}_{X|X\in I}$

Here we first observe $F_{X|X\in I}(x) = p \in [0,1]$, and that $F_X^{-1}(F_X(x)) = x$. We can rewrite in 3:

$$\frac{F_X(x) - F_X(a)}{F_X(b) - F_X(a)} \implies (F_X(b) - F_X(a))F_{X|X\in I}(x) + F_X(a) = F_X(x) \tag{5}$$

By applying F_X^{-1} on both sides, we get that:

$$F_X^{-1}\big((F_X(b) - F_X(a))p + F_X(a)\big) = x = F^{-1}_{X|X\in I}(p) \quad p \in (0,1) \tag{6}$$

$$F^{-1}_{X|X\in I}(p) = \begin{cases} p \in [0,1] & x \\ \text{otherwise} & \text{not defined} \end{cases} \tag{7}$$

Using this we can generate random samples from X, conditioned on that $X \in I$. That is to say, we can can generate random samples from the distribution $f_{X|X\in I}$. This can be done using the inverse method:

Method 2. *Random Number Generation, Truncated Inversion Method:*

1. *Generate N random samples $U_1, \ldots, U_N$ from $U(0,1)$, using a suitable random number generator.*
2. *Use the inverse cdf $F^{-1}_{X|X\in I}$ on each sample: $X_i = F^{-1}_{X|X\in I}(U_i)$.*

Theorem 2. *Each sample X_i generated from method 2 have distribution $F_{X|X\in I}$.*

2 Power Production of a Wind Turbine

In this section we are going to examine a single wind turbine. The total amount of power [W] in the wind passing a wind turbine is given by:

$$P_{tot}(v) = \frac{1}{2}\rho\pi\frac{d^2}{4}v^3 \tag{8}$$

where d [m] is the rotor diameter, $\rho[\mathrm{kg/m^3}$ is the air density, and v [m/s] is the wind speed. However, as 8 gives the theoretical power, this is not realized in practise. For the example turbine that we will work with (Vestas V164 9.5 MW turbine), the actual power produced $P(v)$ follows the power curve in figure [*power curve*]. Here we have a cut-in wind speed of $3.5\mathrm{m/s}$ and a cut-off wind speed of $25\mathrm{m/s}$. The Vestas V164 9.5 MW turbine has a motor diameter $d = 164$.

The power produced by a turbine depends on the winds. Here historical meteorological data has been used to estimate the winds in a given area for each month of the year. From the meteorological data it can be seen that the winds for each month can be modelled by Weibull distributions (one for each month). A Weibull distribution has density function:

$$f(v) = \frac{k}{\lambda}\big(\frac{v}{\lambda}\big)^{k-1}\exp\left(-\big(\frac{v}{\lambda}\big)^k\right) \quad v \geq 0. \tag{9}$$

As we can see from the pdf, the Weibull distribution is parameterized by shape (k) and scale (λ). The shape parameter determines the shape of the distribution and can take on any positive value. If k is smaller than 1, the distribution is said to have a "fat tail," meaning that the probability of extreme events is higher than in the case of a distribution with a larger k. If k is larger than 1, the distribution is more peaked and has a faster decay of tail probabilities. The scale parameter determines the location and spread of the distribution and must be positive. It represents the scale at which the failure rate begins to increase. If λ is larger, the distribution is shifted to the right and spreads out, while if λ is smaller, the distribution is shifted to the left and becomes more concentrated.

The expected value of V can be calculated analytically:

$$\mathbb{E}[V^m] = \Gamma\big(1 + \frac{m}{k})\lambda^m \tag{10}$$

where Γ is the gamma-function.

For this specific case, the scale and shape parameters are estimated to be according to table 1.

month	1	2	3	4	5	6	7	8	9	10	11	12
λ	10.6	9.7	9.2	8.0	7.8	8.1	7.8	8.1	9.1	9.9	10.6	10.6
k	2.0	2.0	2.0	1.9	1.9	1.9	1.9	1.9	2.0	1.9	2.0	2.0

Table 1: The scale and shape parameters for each month

We will now investigate the potential of building a wind turbine at the site in question. The is firstly done by using different Monte Carlo sampling techniques to construct approximate 95% confidence intervals (CI) of the expected power output for each month.

We will from now on denote the approximated expected power output by $\hat{\mu}_{method}$, where the subscript describes the method that has been used. So for example, the approximated expected power output using the standard Monte Carlo will be denoted as $\hat{\mu}_{MC}$.

2.1 Standard MC and Truncated MC

We start by constructing a 95% CI using standard Monte Carlo approximation of the expected power output for each month, i.e. $\hat{\mu}_{MC}^{(i)}$ for $i \in [1, 12]$. The true expected value for the generated power can be written as:

$$\mathbb{E}_{f_V}\big[P(V)\big] = \int_V P(v) f_V(v) dv. \tag{11}$$

So the task lies in approximating the integral in 11. We can here use *the law of large numbers* to motivate the standard Monte Carlo method. First, take N random samples $V_1, \dots, V_N$ with density f_V. By the law of large numbers we can observe that

$$\frac{1}{N}\sum_{i}^{N} P(V_i) \to \int_V P(v) f_V(v) dv, \quad \textit{as } N \to \infty. \tag{12}$$

So in order to approximate 11 we can use the following method:

Method 3. *Standard Monte Carlo Estimation*

1. *Generate N random samples with density f_V.*

2. *Calculate* $\hat{\mu}_{MC} = \dfrac{1}{N}\sum_i^N P(V_i)$.

To to get an estimate of the 95% CI we can use the *central limit theorem* which implies that $\sqrt{N}(\hat{\mu} - \mathbb{E}_{f_V}\left[P(V)\right] \to \mathbf{N}(0, \sigma^2(P))$. And thus an approximate CI can be constructed as:

$$\left(\hat{\mu} - \lambda_p \sqrt{\frac{\sigma(P)^2}{N}}, \hat{\mu} + \lambda_p \sqrt{\frac{\sigma^2(P)}{N}}\right) \tag{13}$$

where λ_p is the p-quantile of the standard normal distribution. Generally $\sigma^2(P)$ is not known, however it can be estimated as:

$$v\hat{a}r(P) = \frac{1}{N-1}\sum_i^N \big(P(V_i) - \frac{1}{N}\sum_l P(V_l)\big)^2 \tag{14}$$

where $v\hat{a}r(P)$ is the sample variance of P (it is assumed that the samples $V_1, \ldots, V_N$ are not the whole population but a set of random samples from a larger population).

We are now ready to construct a 95% CI for each month of the year using standard Monte Carlo. With method 3, using $N = 10^5$ for each month, and method 1 in step 1 for random number generation, we compute $\hat{\mu}_{MC}^{(m)}$ for $m \in [1, 12]$. The variances are computed for each month using eq 14, and CI:s are constructed for each month using eq 13 with $\lambda_p = 1.96$. The results can be seen in Table 2.

month	upper bound	$\hat{\mu}_{MC}$	lower bound
1	4.5438e+06	4.6155e+06	4.6873e+06
2	4.0907e+06	4.1605e+06	4.2303e+06
3	3.7662e+06	3.8342e+06	3.9023e+06
4	2.9063e+06	2.9688e+06	3.0313e+06
5	2.8549e+06	2.9166e+06	2.9784e+06
6	3.0172e+06	3.0808e+06	3.1443e+06
7	2.8331e+06	2.8952e+06	2.9574e+06
8	3.0175e+06	3.0812e+06	3.145e+06
9	3.6911e+06	3.7593e+06	3.8275e+06
10	4.2129e+06	4.2844e+06	4.3559e+06
11	4.5521e+06	4.6242e+06	4.6962e+06
12	4.5496e+06	4.6214e+06	4.6932e+06

Table 2: Estimated confidence bounds using the standard Monte Carlo method, together with the estimated expected value.

We can also construct a 95% CI for each month using a truncated version of the standard Monte Carlo. A truncated version of Monte Carlo simulation refers to a Monte Carlo simulation where the random sampling is limited to a certain range or interval. This is often done when the system being modeled has known constraints or boundaries. Truncated Monte Carlo simulation allows us to focus the simulation on the region of interest and can reduces the computational time required to obtain reliable results. Remember that there was a cut-in wind speed of 3.5 $\mathrm{m/s}$ and a cut-off wind speed of 25 $\mathrm{m/s}$. So, we can focus the simulation on this area. We can use method 3 with slight modifications. In step 1, we use method 2 to sample from the region of interest. In step 2, we need to also multiply by a normalization factor $F_V(b) - F_V(a)$:

$$\hat{\mu}_{TR} = \frac{1}{N}\sum_i^N P(V_i) \cdot (F_V(b) - F_V(a)). \tag{15}$$

where F_V is the cdf of the original distribution. If we do not multiply the estimate of $\mathbb{E}[P(V)|a < V < b]$ by the normalization factor $F(b) - F(a)$, the estimate will not have the correct scale since it will only represent the expected value of $P(V)$ over the interval (a, b) conditioned on the event that V is in that interval. In other words, it will only represent the expected value of $P(V)$ for the portion of the distribution that falls within the interval. The normalization factor accounts for the fraction of the entire distribution that falls within the interval, and by multiplying the estimate by this factor,we obtain an estimate of the expected value of $P(V)$ over the entire distribution.

We also need to calculate the variance. We can do this using eq 14, where we also need to multiply it by $(F_V(b) - F_V(a))^2$. Multiplying the variance by the normalizing factor $(F_V(b) - F_V(a))^2$ ensures that the variance of the Monte Carlo estimate reflects the variance of $P(V)$ over the entire distribution, and not just over the portion that falls within the interval.

$$v\hat{a}r(\hat{\mu}_{TR}) = var(P(V))(F_V(b) - F_V(a))^2 \tag{16}$$

where V is sampled on the interval $I = (a, b)$. We can now construct CI:s for each month using eq 13. The results are presented in Table 3.

month	upper bound	$\hat{\mu}_{TR}$	lower bound
1	4.5665e+06	4.6274e+06	4.6884e+06
2	4.1019e+06	4.1608e+06	4.2196e+06
3	3.7779e+06	3.8349e+06	3.8919e+06
4	2.9261e+06	2.9765e+06	3.0268e+06
5	2.8582e+06	2.9074e+06	2.9567e+06
6	3.0284e+06	3.0796e+06	3.1309e+06
7	2.8352e+06	2.8849e+06	2.9345e+06
8	3.0294e+06	3.0807e+06	3.1321e+06
9	3.6971e+06	3.7542e+06	3.8112e+06
10	4.2152e+06	4.2744e+06	4.3337e+06
11	4.5733e+06	4.6345e+06	4.6956e+06
12	4.5662e+06	4.6271e+06	4.6881e+06

Table 3: Estimated confidence bounds using the truncated version of standard Monte Carlo method, together with the estimated expected value.

2.2 Using the Wind V as a Control Variate

We will now construct 95% CI:s for the expected generated power, using wind V as a control variate to reduce the variance. The control variate strategy is a variance reduction technique where one makes use of some correlated estimator which is known. In our case, we want to estimate eq 11. We can intuitively see that $P(V)$ is correlated to V, as higher values of V generally results in higher values for $P(V)$ ($P(V)$ is monotonously increasing over the interval $(3.5, 14)$). So, V could potentially be used as a control variate. Now we calculate the standard Monte Carlo approximations, $\hat{\mu}_{MC}$ and $\hat{V}_{MC}$. Note that $\hat{\mu}_{MC}$ and $\hat{V}_{MC}$ will be correlated when $corr(\hat{\mu}_{MC}, \hat{V}_{MC}) \neq 0$, and that a positive correlation means that an unusually high outcome of $\hat{V}_{MC}$ tends to correlate to an unusually high outcome for $\hat{\mu}_{MC}$. Thus $\hat{\mu}_{MC}$ should be corrected downwards. The opposite is true for a negative correlation. (Course book). This motivates the control variate estimator:

$$\hat{\mu}_{CV} = \hat{\mu}_{MC} + \alpha(\hat{V}_{MC} - \mathbb{E}_{f_V}[V]) \tag{17}$$

As we can see in eq 17 a constant α should be chosen to minimize the variance of $\hat{\mu}_{CV}$:

$$var(\hat{\mu}_{CV}) = var(\hat{\mu}_{MC}) + \alpha^2 var(\hat{V}_{MC}) + 2\alpha cov(\hat{\mu}_{CV}, \hat{V}_{MC})$$
$$\Longrightarrow$$
$$\min_{\alpha} var(\hat{\mu}_{CV}) = var(\hat{\mu}_{CV}) - \frac{cov(\hat{\mu}_{CV}, \hat{V}_{MC})^2}{var(\hat{V}_{MC})}$$

when

$$\alpha = -\frac{cov(\hat{\mu}_{CV}, \hat{V}_{MC})}{var(\hat{V}_{MC})} \tag{18}$$

We estimate the sample variance of $\hat{\mu}_{MC}$ and $\hat{V}_{MC}$ using eq 14, and the sample covariance can be estimated by:

$$c\hat{o}v(\hat{\mu}_{CV}, \hat{V}_{MC}) = \frac{1}{N-1}\sum_{i}^{N}(P(V_i) - \hat{\mu}_{MC})(V_i - \hat{V}_{MC}) \tag{19}$$

We thus also have an estimate for α, and $\hat{\mu}_{CV}$ can be calculated. An estimate of the variance can now also be calculated:

$$v\hat{a}r(\hat{\mu}_{CV}) = v\hat{a}r(\hat{\mu}_{CV}) - \frac{c\hat{o}v(\hat{\mu}_{CV}, \hat{V}_{MC})^2}{v\hat{a}r(\hat{V}_{MC})}.$$

Once again, we can now calculate CI:s using eq 13. The results are can be seen in Table 4.

month	upper bound	$\hat{\mu}_{CV}$	lower bound
1	4.6261e+06	4.6593e+06	4.6925e+06
2	4.1274e+06	4.1537e+06	4.18e+06
3	3.8154e+06	3.8385e+06	3.8616e+06
4	2.9995e+06	3.0191e+06	3.0386e+06
5	2.8469e+06	2.8657e+06	2.8845e+06
6	3.066e+06	3.0855e+06	3.105e+06
7	2.8673e+06	2.8866e+06	2.9059e+06
8	3.0558e+06	3.0766e+06	3.0975e+06
9	3.7508e+06	3.7737e+06	3.7965e+06
10	4.189e+06	4.2196e+06	4.2503e+06
11	4.6058e+06	4.6395e+06	4.6733e+06
12	4.6206e+06	4.6527e+06	4.6847e+06

Table 4: Estimated confidence bounds using the the wind V as a control variate with the Monte Carlo method, together with the estimated expected value.

2.3 Importance Sampling

In this section we will use importance sampling as variance reduction technique. The importance sampling technique uses the fact that eq 11 can be written in an alternative form:

$$\mathbb{E}_{f_V}\big[P(V)\big] = \int_V P(v)f_V(v)dv = \int_{g(v)>0} P(v)\frac{f_V(v)}{g(v)}g(v)dv = \mathbb{E}_g\big[P(V)\omega(V)\big] \tag{20}$$

where ω are the *importance ratios* $\frac{f_V(V)}{g(V)}$. We can write this as a function:

$$\omega : \{v \in V : g(v) > 0\} \ni v \to \frac{f_V(v)}{g(v)}$$

The (possibly) reduced variance induced by importance sampling comes by causing events with lower probability to happen more often in the simulation by using the envelope (g). In other words, we oversample portions of the target distribution (f_V), and use the importance ratios to correct for the oversampling. The natural question is how to choose the envelope g.

As V is sampled from g, $g(V)$ will often be larger then $f(X)$. Still, it must be so that $\mathbb{E}[f(V)/g(V)] = 1$. Since $f(V)/g(V)$ will more often take values in between 0 and 1, values for when $f(V)/g(V)$ is not between 0 and 1 can be large. The expected variance for $f(V)/g(V)$ is then expected to be large. To counter this, we want $f(V)/g(V)$ to be large only when $P(V)$ is small, so that the variance of $P(V)f(V)/g(V)$ is small. We can translate this to a key requirement for an effective envelope g, that $P(V)f(V)/g(V)$ is close to constant in the support of g.

We can also use *effective sampling size* (ESS) to measure the efficiency of the envelope. The ESS in is a measure of how well the importance weights account for the variability of the samples and provides a measure of the efficiency of the estimation procedure. A low ESS indicates that the importance weights are not effective in capturing the target distribution, leading to a biased or imprecise estimation, while a high ESS suggests that the estimation procedure is more efficient and reliable. [Computational Statistics] presents the following equation to calculate the effective sample size:

$$\hat{N}(g, f) = \frac{N}{1 + v\hat{a}r(\omega(V))} \tag{21}$$

With these things in mind we have tried to find a suitable envelope g. First, we plotted the function $P(V)f_V(V)$. The criteria that $P(V)f(V)/g(V)$ is close to constant in the support of g implies that g should be proportional to $P(V)f(V)$ in the support. See figure 1 As we can see, the shape of $P(V)f(V)$ looks somewhat like a normal distribution or a gamma distribution. So, for g to be proportional, its shape should also look somewhat like a normal distribution. For simplicity we have chosen a normal distribution, also since we did not find a suitable gamma distribution. Because of the restrictions on P, we had a hard time using eq 21 as a measure of the effectiveness of the envelope.

To find suitable parameters for our normal distribution, we chose the estimated $L2$-norm of the gradient as our measure:

$$\min_{\mu,\sigma} ||\nabla P(v)\frac{f_V(v)}{g(v;\mu,\sigma)}||, \quad v \in (0, 30). \tag{22}$$

Then we simply looped over the different combinations of $\mu, \sigma \in (0, 20)$ with a step-size of 0.1 and picked the combination which minimized eq 22. We found $\mu = 10.7$, $\sigma = 4.3$ as the best result.

Here, a maybe more suitable approach would have been to use the Kullback-Leibler Divergence as a measure, however we do not include this here.

Now that the g has been decided, we proceeded with the actual importance sampling.

Method 4. *Monte Carlo Importance Sampling*

1. *Take N random samples $V_1, \ldots, V_N$ from g.*
2. *Calculate the weight ratios $\omega(V_i) = \dfrac{f_V(V_i)}{g(V_i}$.*
3. *Calculate the weighted mean: $\hat{\mu}_{IS} = \dfrac{1}{N}\sum_i^N P(V_i)\omega(V_i)$*

We can now construct CI:s in the usual manner. The results can be seen in Table 5.

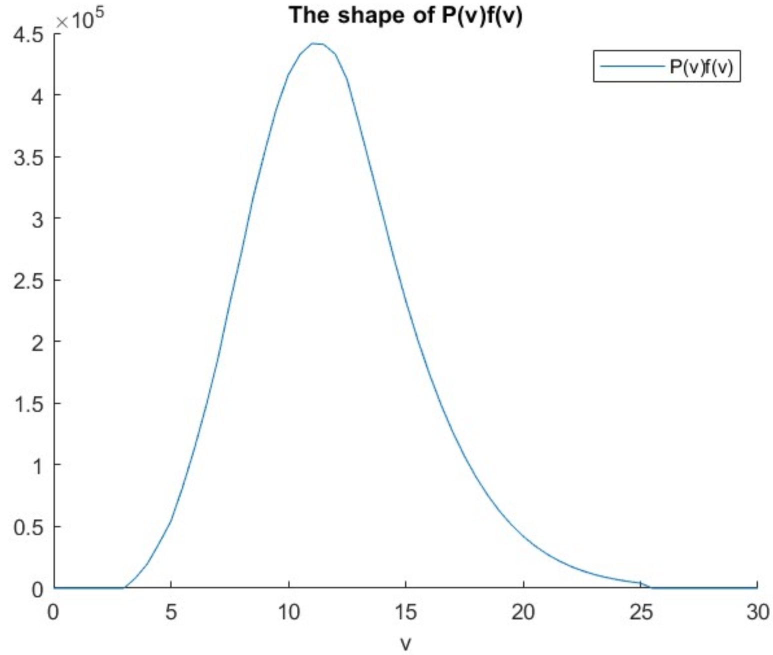

Figure 1: This shows the shape of the $P(V)f_V(V)$ in an area that should be similar to the support of the envelope. The x-axis represents the winds.

2.4 Antithetic Sampling

Antithetic Sampling is a variance reduction technique used in Monte Carlo simulation that involves generating two random variables that are negatively correlated to reduce the variance of the estimator. The basic idea behind antithetic sampling is that if two random variables are negatively correlated, their average will have a smaller variance than the variance of a single random variable. Furthermore, if both estimators are unbiased we have that:

$$\hat{\mu}_{AS} = \frac{\hat{\mu}_1 + \hat{\mu}_2}{2} \tag{23}$$

and

$$var(\hat{\mu}_{AS}) = \frac{1}{4}(var(\hat{\mu}_1) + var(\hat{\mu}_2) + \frac{1}{2}cov(\hat{\mu}_1, \hat{\mu}_2) \tag{24}$$

Now assume we have an unbiased estimate $\hat{\mu}_1$. We want to construct an identically distributed estimate $\hat{\mu}_2$. We can use the following theorem:

Theorem 3. *Let $V = \rho(U)$, where $\rho : R \to R$ is a monotone function. Moreover, assume that there exists a non-increasing transform $T : \mathbb{R} \to \mathbb{R}$ such that $U = T(U)$. Then $V = \rho(U)$ and $\tilde{V} = \rho(T(U))$ are identically distributed. Furthermore:*

$$cov(V, \tilde{V}) = cov(\rho(U), \rho(T(U))) \leq 0$$

Then letting $U \sim \mathbf{U}(0, 1)$, noting that $\rho(U) = P(F_V^{-1}(U))$ is monotone, $T = 1 - U$ a non-decreasing transform, and with $V = \rho(U)$, $\tilde{V} = \rho(T(U))$, we get that $V \equiv \tilde{V}$ and $cov(V, \tilde{V}) \leq 0$

So, we can use method 1 two times on the same set of uniform samples and then estimate each with standard Monte Carlo (method 3), and then use eq 23. We can calculate the variance in accordance with eq 24, and construct CI:s in the usual manner. The results are shown in table 6.

month	upper bound	$\hat{\mu}_{IS}$	lower bound
1	4.5869e+06	4.6386e+06	4.6903e+06
2	4.1185e+06	4.1512e+06	4.1839e+06
3	3.7945e+06	3.8211e+06	3.8478e+06
4	2.9835e+06	3.0044e+06	3.0252e+06
5	2.8573e+06	2.8778e+06	2.8983e+06
6	3.0569e+06	3.0776e+06	3.0984e+06
7	2.8292e+06	2.8499e+06	2.8706e+06
8	3.0567e+06	3.0776e+06	3.0984e+06
9	3.6984e+06	3.7248e+06	3.7512e+06
10	4.2111e+06	4.2505e+06	4.2899e+06
11	4.6052e+06	4.6569e+06	4.7086e+06
12	4.5903e+06	4.6403e+06	4.6902e+06

Table 5: Estimated confidence bounds using g as an envelope in the importance sampling with the Monte Carlo method, together with the estimated expected value.

month	upper bound	$\hat{\mu}_{AS}$	lower bound
1	4.6427e+06	4.6527e+06	4.6628e+06
2	4.1334e+06	4.1458e+06	4.1583e+06
3	3.8132e+06	3.8287e+06	3.8442e+06
4	2.9909e+06	3.013e+06	3.035e+06
5	2.8374e+06	2.8602e+06	2.883e+06
6	3.0623e+06	3.084e+06	3.1057e+06
7	2.8721e+06	2.8951e+06	2.918e+06
8	3.0609e+06	3.0828e+06	3.1046e+06
9	3.7639e+06	3.7801e+06	3.7962e+06
10	4.2109e+06	4.2238e+06	4.2366e+06
11	4.6499e+06	4.6598e+06	4.6697e+06
12	4.6508e+06	4.6601e+06	4.6695e+06

Table 6: Estimated confidence bounds using antithetic sampling as variance reduction technique with the Monte Carlo method, together with the estimated expected value.

We can compare each of the methods. This is done in Figure 2, that shows how wide the confidence bounds for each method for each month get. As we can see from both the presented tables and Figure 2, the standard Monte Carlo performs worst in terms of variance. This is of course expected as no measures has been taken to reduce it. The truncated version has a somewhat reduced variance compared to the standard Monte Carlo as it focuses more on the region of interest. The IS has less variance, however it is inconsistent in that the variance gets larger for stronger winds (more generated power), and this is a disadvantage. This probably is a result from the fact that the same envelope was used for all months, and not fitted for each month. The AS method gives very low variance with the same number of samples, it should however be noted that its somewhat more computationally heavy, as we transform the uniform samples into two Weibull samples two times using the inverse on U and $1 - U$. The same goes for mean and variance calculations.

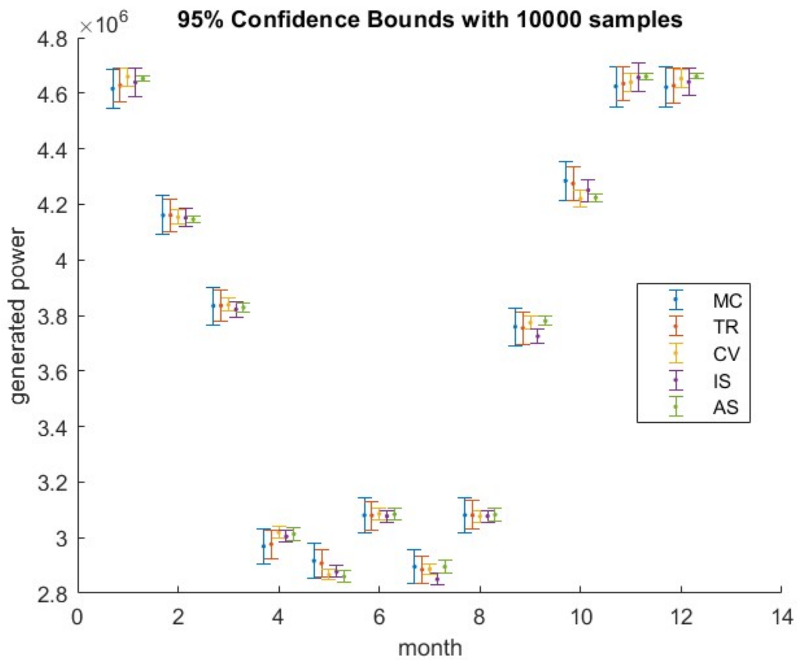

Figure 2: All estimated bounds: MC = standard Monte Carlo, TR = Truncated standard Monte Carlo, CV = Monte Carlo with control variate, IS = Monte Carlo with importance sampling using g as an envelope, AS = Monte Carlo with antithetic sampling. The upper end of each bar represents the upper confidence, and the lower end represents the lower confidence. The x-axis is the month, and the y-axis represents the generated power for this month.

2.5 Probability that the turbine produces power

We can calculate the probability that the turbine produces power as:

$$\mathbb{P}(P(V) > 0) = \mathbb{P}(V > 25) - \mathbb{P}(V < 3.5) = F_V(25) - F_V(3.5) \tag{25}$$

which can be applied to for each month. The results can be seen in Table 7

month	1	2	3	4	5	6	7	8	9	10	11	12
		0.88	0.86	0.81	0.80	0.82	0.80	0.82	0.86	0.87	0.89	0.89

Table 7: Estimated probabilities that the turbine produces power for each month of the year.

2.6 Average Power Coefficient

We also create an approximate 95% confidence interval for the average ratio of actual wind turbine output to total wind power (average power coefficient) , i.e.:

$$\frac{\mathbb{E}[P(V)]}{\mathbb{E}[P_{tot}(V)]}. \tag{26}$$

The denominator can be calculated analytically using eq 10 multiplied by the constant of P_tot, with $m = 3$:

$$\mathbb{E}[P_{tot}(V)] = \frac{1}{2}\rho\pi\frac{d^2}{4}\Gamma\big(1 + \frac{m}{k}\big)\lambda^m. \tag{27}$$

We then use standard monte carlo to estimate $\mathbb{E}[P(V)]$ and construct CI:s. The results can be seen in Table 8.

month	upper bound	$\hat{\mu}_{AS}$	lower bound
1	0.22326	0.22678	0.23029
2	0.25954	0.26399	0.26845
3	0.27689	0.28194	0.28699
4	0.31924	0.32598	0.33272
5	0.32368	0.33074	0.33781
6	0.31236	0.31895	0.32555
7	0.32495	0.332	0.33905
8	0.30989	0.31643	0.32298
9	0.28581	0.29102	0.29624
10	0.23355	0.23758	0.2416
11	0.22687	0.23039	0.23391
12	0.22319	0.2267	0.2302

Table 8: Estimated confidence bounds for the average power coefficient, together with the estimated expected value of the ratio.

2.7 Capacity Factor and Availability Factor

We can calculate the capacity factor as $C = \dfrac{\hat{\mu}_{MC}}{9.5e6}$ and the availability factor as $A = F_V(25) - F_V(3.5)$ for each month and then take the average. This results in

$$\hat{A} = 0.85 \rightarrow 85\% \quad \hat{C} = 0.39 \rightarrow 39\% \tag{28}$$

To be interested in building a turbine $\hat{A}$ should be a least 90% and $\hat{C}$ should be at least 20%. As we can see it is suitable from a capacity point of view, the availability is however too low. One could make an argument that the high capacity could make up for the availability, for example by calculating the the average power generated per time unit or the total produced power. However, the low availability could be a problem not in the sense of total generated power, but in the stability and selling prices for power, as the price could be high when the turbine produces and low when its not available. In conclusion, from this analysis a wind turbine should not be built. However, further analysis could change this conclusion.

3 Combined production of Two Wind Turbines

3.1 Expected Power of the two Turbines

We want to estimate the combined produced power from the two turbines. This can written the following way:

$$\mathbb{E}_{f_{v1,v2}}[P(V_1)+P(V_2)] = \int_{V_1}\int_{V_2}(P(v_1)+P(V_2))f(v_1,v_2)dv_1dv_2 \tag{29}$$

$$= \int_{V_1}\int_{V_2} P(v_1)f(v_1,v_2)dv_1dv_2 + \int_{V_1}\int_{V_2} P(v_2)f(v_1,v_2)dv_1dv_2 \tag{30}$$

$$= \int_{V_1} P(v_1)\int_{V_2} f(v_1,v_2)dv_2dv_1 + \int_{V_2} P(v_2)\int_{V_1} f(v_1,v_2)dv_1dv_2 \tag{31}$$

$$= \int_{V_1} P(v_1)f(v_1)dv_1 + \int_{V_2} P(v_2)f(v_2)dv_2 \tag{32}$$

$$= \mathbb{E}_{f_{v_1}}[P(V_1)] + \mathbb{E}_{f_{v_2}}[P(V_2)] \tag{33}$$

$$= 2\mathbb{E}_{f_{v_1}}[P(V_1)] = 2\mathbb{E}_{f_{v_2}}[P(V_2)] \tag{34}$$

Where in step 30 we marginalize over v_1 and v_2 respectively. We can thus estimate using method 4, using the envelope as before, and simply multiply by 2:

$$2\hat{\mu}_{IS} = 7.7093e6$$

3.2 Estimating the Covariance between the turbines

For two jointly distributed real-valued random variables V_1 and V_2 we can write the covariance as:

$$cov(P(V_1), P(V-2)) = \mathbb{E}_{f_{(v_1,v_2)}}[P(V_1)P(V_2)] - \mathbb{E}_{f_{v_1}}[P(V_1)]\mathbb{E}_{f_{v_2}}[P(V_2)]$$
$$= \mathbb{E}_{f_{(v_1,v_2)}}[P(V_1)P(V_2)] - \mathbb{E}_{f_{v_1}}[P(V_1)]^2$$
$$= \mathbb{E}_{f_{(v_1,v_2)}}[P(V_1)P(V_2)] - \mathbb{E}_{f_{v_2}}[P(V_2)]^2$$

So we can calculate the second term easily with importance sampling in one dimension. The first term, however, needs to be calculated using multivariate importance sampling.

The univariate importance sampling easily extends to the multivariate case., with the implication that it might be harder to find the envelope $g(v_1, v_2)$. For our case, we choose to extend our previously found envelope. We chose to do this because the shape and scale of the marginal distributions both are $\lambda = 9.13$ and $k = 1.96$, the same as the parameters we used to find the univariate g which then should extend nicely because of symmetry.

$$g(v_1, v_2) = g((v_1, v_2); \mu, \sigma) \tag{35}$$

We can now proceed as before with the importance sampling, now using random samples from $g(v_1, v_2)$. The weights are calculated as:

$$\omega(V_{1,i}, V_{2,i}) = \frac{f_{V_1,V_2}(V_{1,i}, V_{2,i})}{g(V_{1,i}, V_{2,i}); \mu, \sigma)} \tag{36}$$

getting that

$$\mathbb{E}_{f_{(v_1,v_2)}}[P(V_1)P(V_2)] = \mathbb{E}_{g_{(V_1,V_2)}}[P(V_1)P(V_2)\omega(V_1, V_2)] \tag{37}$$

so

$$\frac{1}{N}\sum_{i}^{N} P(V_{1,i})P(V_{2,i})\omega(V_{1,i},V_{2,i}) = 2.0791e+13 \tag{38}$$

The covariance can now be calculated, getting:

$$c\hat{o}v(P(V_1),P(V_2)) = \frac{1}{N}\sum_{i}^{N} P(V_{1,i})P(V_{2,i})\omega(V_{1,i},V_{2,i}) - \hat{\mu}_{IS}^2 = 6.5502e+12$$

3.3 The Variability $var[P(V_1)+P(V_2)]$

To calculate the variability, we can use:

$$var[P(V_1)+P(V_2)] = var[P(V_1)] + var[P(V_2)] + 2cov(P(V_1),P(V_2)) \tag{39}$$

where the variance of $P(V_1)$ and $P(V_2)$ should be equal because of symmetry. We get the estimated sample variance

$$v\hat{a}r[P(V_1)+P(V_2)] = 3.6766e+13 \tag{40}$$

3.4 $\mathbb{P}[P(V_1)+P(V_2) > 9.5\text{MW}]$ and $\mathbb{P}[P(V_1)+P(V_2) < 9.5\text{MW}]$

We now want to estimate the probability that the combined power generated from the two turbines is greater then half of their installed capacity. We can do this using an indicator function and importance sampling as a variance reduction technique.

$$\mathbb{P}[P(V_1)+P(V_2) > 9.5\text{MW}] = \int\int_{P(V_1)+P(V_2)>9.5\text{MW}} f(v_1,v_2)dv_1dv_2 \tag{41}$$

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbf{1}_{P(V_1)+P(V_2)>9.5\text{MW}} f(v_1,v_2)dv_1dv_2 \tag{42}$$

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbf{1}_{P(V_1)+P(V_2)>9.5\text{MW}} \frac{f(v_1,v_2)}{g(v_1,v_2)} g(v_1,v_2)dv_1dv_2 \tag{43}$$

$$= \mathbb{E}_g[\mathbf{1}_{P(V_1)+P(V_2)>9.5\text{MW}}\omega(V_1,V_2)] = 0.3767 \tag{44}$$

To get $\mathbb{P}[P(V_1)+P(V_2) < 9.5\text{MW}]$ we simply flip the inequality in the indicator function: $\mathbb{P}[P(V_1)+P(V_2) < 9.5\text{MW}] = 0.5952$. CI:s have also been constructed for both probabilities, see Table 9 and Table 10.

	upper bound	$\mathbb{E}$	lower bound
$\mathbb{E}[\mathbf{Pr}(P(v_1)+P(v_2)) > 9.5MW]$	0.3608	0.3776	0.3944
$\mathbb{E}[\mathbf{Pr}(P(v_1)+P(v_2)) < 9.5MW]$	0.4493	0.5572	0.6652

Table 9: Estimated confidence bounds for the probability that the combined power generated by the two turbines is greater (respectively less) than half of their installed capacity

It would be better to find another envelope here, one that samples such that the event we are examining happens more often, i.e where the indicator function is not 0