Detailed solutions, making sure that you describe your algorithms in text, and not only as code, must be submitted **before Tuesday 12 Feb, 13:00:00**. You are strongly encouraged to work in groups of two.

Report submissions are accepted in PDF format only.

Also submit an email with your MATLAB-files (or implementation in other language), with a file named `proj1.m` that can be used to run your analysis (Remember to submit **all** of the files you use to the create your solution). Email all files (including the report PDF) to `fms091@matstat.lu.se`. Set the subject line to:
"Project 1 by "STILID1" and "STILID2" "

Discussion between groups is permitted, as long as your report reflects your own work.

Random number generation

1. The solution to the following problem may be useful in the next section. 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 $\mathsf{I} = (a, b)$ be an interval such that $\mathbb{P}(X \in \mathsf{I}) > 0$.

   (a) Find the conditional distribution function $F_{X|X\in\mathsf{I}}(x) = \mathbb{P}(X \leq x | X \in \mathsf{I})$ and density $f_{X|X\in\mathsf{I}}(x)$ of X given that $X \in \mathsf{I}$.

   (b) Find the inverse $F^{-1}_{X|X\in\mathsf{I}}$. How can this be used for simulating X conditionally on $X \in \mathsf{I}$?

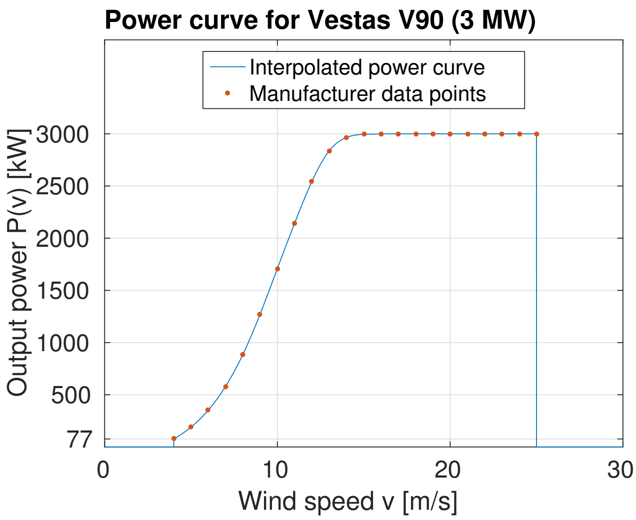
Power production of a wind turbine

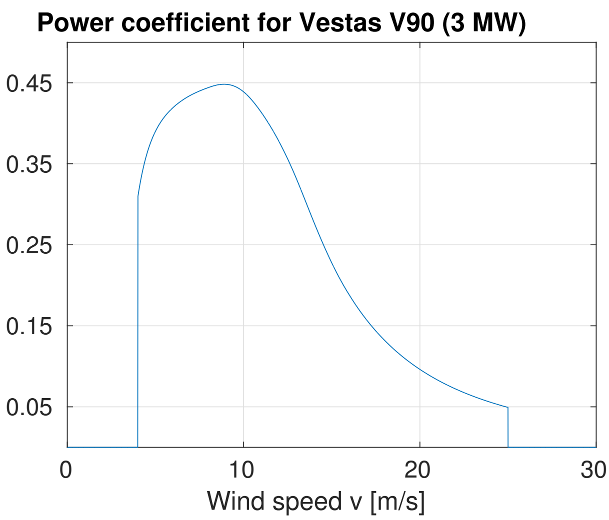
2. Renewable energy sources becomes more and more important for the generation of electrical power throughout the world. In this assignment we are therefore going to take a closer look at wind power. The total amount of power $[W]$ in the wind passing a wind turbine with rotor diameter d is given by

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

   where ρ $[kg/m^3]$ is the air density (about 1.225 kg/m^3 at sea level), d in m is the rotor diameter of the wind turbine and v $[m/s]$ is the wind speed. However a wind turbine cannot completely stop the air flow, to work in practise. It turns out that is optimal to reduce the wind speed by two thirds to get out the optimal power. This is known as the Betz-limit (after Albert Betz who was a German physicist) and gives that we can utilise at maximum $16/27 \approx 59\%$ of the total wind power (see Ragheb and Ragheb (2011) for a detailed derivation). The Betz limit is derived under idealised assumptions and real wind turbines can usually utilise at most about $40 - 45\%$ of the total power. This fraction of the total power that a wind turbine can utilise is called the *power coefficient*.

   The amount $P(v)$ of electrical power produced by an actual wind turbine at a given wind speed v $[m/s]$ is described by the *power curve* of the turbine. For a Vestas V90 3 MW turbine we have a power curve and a power coefficient (measured at air density $1.225 kg/m^3$) as given on next the page:





This behaviour is typical for real wind turbines. We have here a cut-in wind speed of 4.0 m/s and a cut-off wind speed of 25 m/s. Outside the interval $[4, 25]$ m/s the wind turbine produces no power. The wind turbine Vestas V90 has a rotor diameter of $90m$ and a tower height of $80m$. The file `powercurve_V90.mat` contains a function object P that gives the output of the wind turbine. Load the object with `load powercurve_V90`. Use it as a regular function, e.g. `P(6.5)` will return the output power in (W) for the wind speed 6.5 m/s. The function P does accept vector valued input of wind speeds.

Of course the wind speed at a wind turbine will depend on the location of the turbine and vary during the year. Meteorological records can be used to estimate the distribution of winds in a given area. We plan to build a wind turbine at a site in northern Europe; from records we see that such stochastic wind speeds V can be modelled by a Weibull distribution

$$f(v) = \frac{k}{\lambda}\left(\frac{v}{\lambda}\right)^{k-1} \exp\left(-\left(\frac{v}{\lambda}\right)^{k}\right), \quad v \geq 0, \tag{2}$$

where k and λ varies between months as follows

Par/Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
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

This will give an yearly average expectation of the wind speed slightly above 8.1 m/s, which is an reasonable approximation of the yearly average wind speed in the southern part of Öresund (between Malmö and Saltholm, Lat: 55.67°, Lon: 12.86°) at the height $80m$ (remember that the centre of our rotor is located at $80m$). Data is taken from `http://www.windmap.se` and `https://opendata-download-metobs.smhi.se/explore/?parameter=2#`. In MATLAB, Weibull-distributed random numbers can be generated using `wblrnd`; in addition, the probability density and cumulative distribution functions of the Weibull distribution can be evaluated using `wblpdf` and `wblcdf`.

We now wish to investigate the potential of building a wind turbine at the site in question. So solve the following for each month.

(a) Create an approximate 95% confidence interval for the expected amount of power generated by the wind turbine using draws from a Weibull distribution. Compare the width of the confidence interval using standard Monte Carlo and the truncated version considered in problem 1. Remember to properly adjust for the conditioning when computing the estimate using the truncated Weibull distribution.

(b) Compare the above result to an approximate 95% confidence interval created by means of importance sampling based on some convenient instrumental distribution.

(c) The power curve $P(v)$ is monotonously increasing over the interval $(4, 16)$ and constant over $(16, 25)$. Use this for reducing the variance of the estimator in (a) via antithetic sampling. Construct a new 95% confidence interval using the antithetic estimator and compare it to the ones obtained in (a) and (b). Which approach would you recommend and why?

(d) Calculate/estimate the probability that the turbine delivers power, $\mathbb{P}(P(V) > 0)$.

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

$$\frac{\mathbb{E}P(V)}{\mathbb{E}P_{\text{tot}}(V)},$$

where P_{tot} is given by formula (1). Note that the denominator can be calculated explicitly without simulation, do this also.

(f) Two important characteristics of power plants are the *capacity factor*, or the ratio of the actual output over a time period and the maximum possible output during that time (3 MW times the time span for the Vestas 90 3 MW); and the *availability factor*, or the amount of time that electricity is produced during a given period divided by the length of the period (you can here re-use the result from (d)). Wind turbines typically have a capacity factor of 20–40% and an availability greater than 90%. Does this seem like a good site to build a wind turbine? Look here at the averages over all months, i.e. first calculate the measures for each month and then take averages.

Combined power production of two wind turbines

3. Two wind turbines placed in the same area will be exposed to similar winds V_1 and V_2, which we here model using a bivariate Weibull distribution (given by a generalized Farlie-Gumbel-Morgenstern copula, see Bairamov and Kotz and Bekci (2001) Equation (4)) for $v_1 \geq 0, v_2 \geq 0$. This gives that the joint cumulative distribution function of v_1 and v_2 is given by

$$F(v_1, v_2) = F(v_1)F(v_2)\left[1 + \alpha\left(1 - F(v_1)^p\right)^q\left(1 - F(v_2)^p\right)^q\right],$$

where we have chosen $\alpha = 0.638,\ p = 3$ and $q = 1.5$. We thus have the joint density function:

$$f(v_1, v_2) = f(v_1)f(v_2)\left[1 + \alpha\left(1 - F(v_1)^p\right)^{q-1}\left(1 - F(v_2)^p\right)^{q-1}\left(F(v_1)^p(1+pq) - 1\right)\left(F(v_2)^p(1+pq) - 1\right)\right],$$

where f and F are again the probability density and cumulative distribution functions of the univariate Weibull distribution.

Note that the marginal distributions of V_1 and V_2 are Weibull. To simplify the setup we only use one set of parameters $k = 1.96$ and $\lambda = 9.13$. This corresponds approximately to the average yearly behaviour. A more realistic model would be to also include a dependence on the wind direction in α but that is too complicated to include in this project.

Now use importance sampling (based on some convenient instrumental distribution) to estimate the following:

(a) The expected amount of combined power generated by both turbines, i.e. $\mathbb{E}(P(V_1) + P(V_2))$. This actually reduces to a one dimensional problem, why is this true?

(b) The covariance $\mathbb{C}(P(V_1), P(V_2))$.

(c) The variability $\mathbb{V}(P(V_1) + P(V_2))$ in the amount of combined power generated by both turbines as well as the standard deviation $\mathbb{D}(P(V_1) + P(V_2))$.

(d) Find an approximate 95% confidence interval for the probability $\mathbb{P}(P(V_1) + P(V_2) > 3.0MW)$ that the combined power generated by the two turbines is greater than half of their installed capacity and for $\mathbb{P}(P(V_1) + P(V_2) < 3.0MW)$. (use importance sampling or other variance reduction techniques for **both** probabilities) Do the probabilities sum to one or not? Why?

Good luck!

References

Bairamov, I., Kotz, S., and Bekci, M. (2001). New generalized Farlie-Gumbel-Morgenstern distributions and concomitants of order statistics. Journal of Applied Statistics, 28(5), 521-536. Available at `http://ludwig.lub.lu.se/login?url=http://www.tandfonline.com/doi/pdf/10.1080/02664760120047861?needAccess=true`

Magdi Ragheb and Adam M. Ragheb (2011) Wind Turbines Theory - The Betz Equation and Optimal Rotor Tip Speed Ratio, in *Fundamental and Advanced Topics in Wind Power*, (eds) Rupp Carriveau. Available at `http://www.intechopen.com/download/pdf/16242`.