## Method of moments for univariate probability distributions

Let be a random variable and are samples from a data set that is representative of .

In part B, the numerical summaries of sample sets were introduced. These include the sample mean , defined as the arithmetic mean of the sample values :

|  |  |
| --- | --- |
|  |  |

and the sample variance, defined as

|  |  |
| --- | --- |
|  |  |

In a nutshell, the method of moments consists in utilizing and as point estimators for the mean and the standard deviation of the random variable . The distribution parameters of are obtained by requiring that has mean and variance equal to and . For all probability distributions with two parameters (e.g. Normal, Lognormal, Weibull), these two conditions are sufficient to calculate parameter estimates.

In the case of the Normal distribution, the point estimates for the distribution parameters and are simply

|  |  |
| --- | --- |
|  |  |
|  |  |

In the case of the Lognormal distribution, the point estimates for the distribution parameters and are

|  |  |
| --- | --- |
|  |  |
|  |  |

Example: Estimation of timber tensile strength

We model the tensile strength of a timber specimen by a Weibull distribution with CDF

The parameters of the distribution are and . The mean and variance of the distribution are given by

We utilize the sample data in *Timber.dat*. The sample mean and sample variance are calculated as

,

.

We set and and derive the corresponding parameter values of . To this end, we note that the coefficient of variation is a function only of the shape parameter :

There is no analytical solution for , but numerical solutions are readily available (e.g. through the Matlab function *fzero*). We obtain the estimator

The second parameter is obtained from the condition

as

The resulting CDF of is shown in Figure 1, together with the empirical cumulative frequency plot.

Timber_Weibull_model.eps

Figure 1. Comparison of the estimated CDF of with the cumulative frequency diagram.

It is observed that the CDF estimate does not match well with the observed frequency of tensile strength values at the lower range. As the interest is in low values of , which determine the reliability of structures, this is unfortunate. Other distribution models, e.g. the Lognormal distribution, should be investigated.