22.4UC2 Using the Statistics Toolbox

MATLAB provides a Statistics Toolbox containing many useful tools. The MATLAB installations here at Heriot-Watt have the Statistics Toolbox included. The topics covered by the Toolbox include the following:

- Probability distributions
- Descriptive statistics
- Hypothesis testing
- Multivariate statistics
- Statistical plots

For example, the Toolbox provides 20 probability distributions. For each distribution there are five associated functions, namely the probability density function (`pdf`), the cumulative distribution function (`cdf`), the inverse of the cumulative distribution function, a random number generator, and the mean and variance as a function of the parameters. Details can be found in the online documentation (for example at

`http://www.ma.hw.ac.uk/matlab/help/helpdesk.html`).

Below are a couple of examples on how to use this package.

(a). After typing

```
x = [-3:0.1:3];
f = normpdf(x,0,1);
```

the variable `f` will contain the density of the normal pdf with parameters $\mu = 0$ and $\sigma = 1$. After typing

```
p = normcdf(x,0,1);
```

the variable `p` will hold the probabilities associated with the standard normal pdf.

(b). The command

```
normrnd(0,1)
```

will create a random number 'distributed' according to the standard normal distribution. The command

```
normrnd(0,1,10,3)
```

will create 30 random numbers distributed according to the standard normal distribution, printed in 10 rows with 3 columns each.

Here is a list of some of the probability distributions. Their parameters have the obvious meaning.

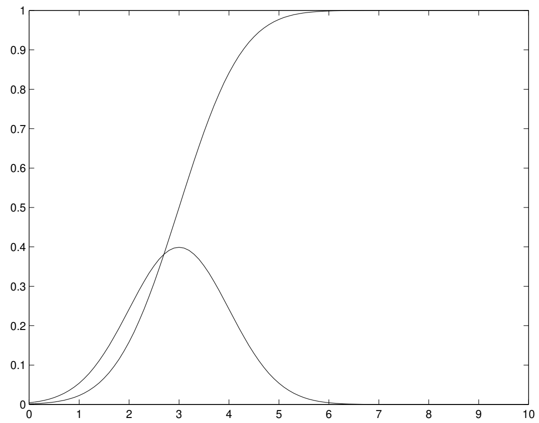
- Binomial distribution: `binopdf(x,n,p);`
- Chi-square distribution: chi2pdf(x,ν);

- Exponential distribution: `exppdf(x,`μ`);`
- Gamma distribution: `gampdf(x,`α`,`β`);`
- Normal distribution: `normpdf(x,`μ`,`σ`);`
- Poisson distribution: `poisspdf(x,`λ`);`
- Student's t-distribution: `tpdf(x,`ν`);`
- Weibull distribution: `weibpdf(x,`$\frac{1}{\beta}$`,`$\alpha+1$`);`

You can plot these functions using the usual commands. For example

```
x = 0:0.1:10;
y = normpdf(x,3,1);
z = normcdf(x,3,1);
plot(x,y,x,z)
```

will plot both the normal density function and its cumulative distribution (for $\mu = 3$ and $\sigma = 1$):



Given a sample (vector) x, the commands

```
geomean(x)
harmmean(x)
mean(x)
median(x)
```

will produce the geometric mean, harmonic mean, the arithmetic average, and the 50th percentile of the data stored in x. The commands

```
range(x)
var(x)
std(x)
```

will produce the range (difference between maximum and minimum values), the variance, and the standard deviation of the data.