Learning outcomes: Chapter 5

1. You should understand the concept of a *probability density function*.

2. You should understand why the *Normal distribution* is widely used in many real–life applications.

3. You should be able to sketch the shape of the probability density function for a Normal distribution, given it's mean μ and standard deviation σ, and you should know how this distribution changes for different values of μ and σ.

4. You should understand the conventional notation for a Normal distribution with mean μ and variance σ^2, that is, $X \sim N(\mu, \sigma^2)$.

5. Given that $Z \sim N(0, 1)$, you should be able to use statistical tables to obtain probabilities of the type $P(Z < z)$. You should also be able to calculate "greater than" probabilities, i.e. $P(Z > z)$, and the probability that Z lies between two values, i.e. $P(z_1 < Z < z_2)$.

6. Given that $X \sim N(\mu, \sigma^2)$, you should be able to operate the transformation

$$Z = \frac{X - \mu}{\sigma},$$

and then use statistical tables, to find probabilities such as $P(X < x)$, $P(X > x)$ and $P(x_1 < X < x_2)$.

7. You should know when it might be appropriate to use a Uniform $U(a, b)$ distribution. Given $X \sim U(a, b)$, you should also be able to:

– Find probabilities using the formula

$$P(X < x) = \frac{x - a}{b - a};$$

– Calculate $E[X] = \frac{a + b}{2}$;

– Calculate $Var(X) = \frac{(b - a)^2}{12}$.

8. Given that X follows an exponential distribution with rate λ, i.e. $X \sim \exp(\lambda)$, you should be able to:

– Find probabilities using the formula

$$P(X < x) = 1 - e^{-\lambda x};$$

– Calculate $E[X] = 1/\lambda$;

– Calculate $Var(X) = 1/\lambda^2$.

Prize question*

A car hire firm has a large fleet of cars for hire by the day and it is found that the fleet suffers breakdowns at the rate of 21 per week.

(a) Suggest a suitable distribution for X: the number of breakdowns per year.

(b) Using an approximation, find the probability that in any year more than 1,150 breakdowns occur.

(c) State the conditions under which your approximation in (b) is appropriate.

(d) Explain what is meant by the term *overdispersion*. Why might this be an issue when using the model you proposed in part (a)? How might this be overcome?

*** To be handed in to me by no later than the lecture on Friday 10th March**