Dr Isabella Gollini

Probability and Statistics

Examples 4

1. Consider the *triangular distribution* on the interval $(0, \theta)$, that is, the distribution with p.d.f.

$$f(x) = \begin{cases} 2x/\theta^2 & (0 < x < \theta) \\ 0 & \text{otherwise} \end{cases}$$

(i) Find expressions in terms of θ for the mean and variance of this distribution.

(ii) Find an expression for the c.d.f. $F(x)$ of this distribution.

2. Given $X \sim N(20, 4)$, (so that $\mu = 20$ and $\sigma^2 = 4$) find

(i) $\Pr(X < 22)$

(ii) $\Pr(X < 19)$

(iii) $\Pr(X > 16)$

(iv) $\Pr(17 < X < 22)$

In each case use the table of the standard normal distribution function to evaluate the probabilities and use `R` to check the results.

3. The weights, in grams, of a certain variety of plum are normally distributed with mean 20 and variance 16.

Find the probability that the mean of a random sample of 36 plums is within 1 gram of the population mean.

4. Let the r.v. X represent the birth weight in grams of a baby born at a certain hospital. Assume that the mean of X is 3320 and the standard deviation 660. Let $\bar{X}$ be the sample mean of a random sample of 225 babies born at the hospital.

Stating any assumptions that you make or theoretical results that you make use of, find the probability that $\bar{X}$ lies between 3223 and 3407.

5. For a continuous random variable X the *lower quartile* q_1 and the *upper quartile* q_3 of its distribution are defined to be the values such that

$$\Pr(X < q_1) = 0.25$$

and

$$\Pr(X > q_3) = 0.25.$$

The interquartile range is $q_3 - q_1$.

For the special case of a r.v. Z with the standard normal distribution,

(i) Find q_1 and q_3 hence the interquartile range.

(ii) Evaluate $\Pr(|Z| > 4q_3)$.

(iii) Discuss briefly how these results relate to the length of the whiskers in a box-and-whisker plot as described in Lecture 1.

6. 25 samples, drawn at regular intervals of time, where each sample is of size 4, have been taken of titanium buttons coming off a production line. For each button a measurement of hardness (in DPN units) is recorded. Previous experience has shown that if the process is in control then the hardness of the buttons is normally distributed with mean 127 and standard deviation 3.4.

The data are stored in the data set `Hardness.csv`. It is composed of 100 rows and 2 column. The first column identifies the sample, and the second column contains the values of the measurements.

Using `R`, plot the data on an appropriate $\bar{x}$-chart and comment on any noteworthy features of the plot.

[Hint: Load the data and create the objects: `Xbar`, `mean`, `sd`, `n`, and `plot_title`, then you can copy and paste the code in the file `ControlCharts.txt` available on Moodle (That is the same code of Chapter 7, page 9.)].