Quantitative Economics: Introductory Probability

The questions in this worksheet are intended to help you become familiar with the basics of probability distributions, which you will need for the QE course. For some students this will be revision; but if you have not done probability before you should work carefully through the chapters indicated below of Thomas, or Wonnacott and Wonnacott, or a similar textbook.

References:
Lecture slides for lectures 1-4.

James H. Stock and Mark W. Watson, *Introduction to Econometrics*, Brief Edition, or complete Edition. Chapter 1 (introduction to QE); Chapter 2, Sections 2.1-2.4.

R. L. Thomas, *Using Statistics in Economics*, Prerequisites, and Chapters 1, 2

T. H. Wonnacott and R. J. Wonnacott, *Introductory Statistics for Business and Economics*, Chapters 1-5

Sheldon Ross, *A First Course in Probability*, Chapers 1-7

. . . and there are many other good introductions to probability and statistics.

Questions:

1. An urn contains 1 black ball, 2 white balls and 7 red balls. You draw one ball at random. A random variable is defined:

$$X = \begin{cases} 30 & \text{if black} \\ 10 & \text{if white} \\ 0 & \text{if red} \end{cases}$$

   (a) Write down the probability mass function and the cumulative distribution function, and sketch them.

   (b) What is (i) $\mathrm{P}(X > 0)$ (ii) $\mathrm{P}(X \leq 20)$?

   (c) Find the mean and variance of X.

2. The probability density of the uniform distribution, $U(a, b)$, is: $f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$

   Find, by integrating:

   (a) the cumulative distribution function

   (b) the mean;

   (c) the variance;
   Hints: Find $\mathrm{E}(X^2)$ first. Note that $b^3 - a^3 = (b - a)(a^2 + ab + b^2)$.

   and draw the cdf.

3. The *Quantiles* of a distribution are obtained by inverting the distribution function at regular intervals:

   - The *quartiles* are the points x_1, x_2 and x_3 such that: $F_X(x_1) = 0.25$, $\mathrm{F}_X(x_2) = 0.5$, $\mathrm{F}_X(x_3) = 0.75$ and the second quartile x_2 is the median.
   - The 94th *percentile* of the distribution of X is the value of x such that $F_X(x) = 0.94$
   - Similarly the *7th decile* is the point x such that $F_X(x) = 0.7$
   - and *quintiles* ...etc.

   (a) What are the quartiles of $N(0,1)$?

   (b) Find the 1^{st} decile and the 78^{th} percentile of $U(0,2)$.

4. Suppose that students' marks on the economics prelims paper are normally distributed with mean 61 and standard deviation 9.5.

   (a) What is the probability that a student scores (i) less than 50? (ii) 70 or more?

   (b) What score is exceeded by only 10% of students?

   (c) Find the median, and the upper and lower quartiles.

   (d) What proportion of students have scores within 5 marks of the mean?

5. Suppose that the probability of rain on a randomly-chosen day in August is 0.1. Let Y be the number of rainy Saturdays in August. Under what conditions can we assume that Y has a binomial distribution, $Y \sim B(n, p)$, where $n = 4$ and $p = 0.1$? Find the probability mass function for $B(4, 0.1)$.[1] Verify by direct calculation that (under the required conditions) the expected number of rainy Saturdays is equal to np. Is the distribution positively or negatively skewed?

6. If X and Y are random variables, and a, b and c are constants:[2]

$$\mathrm{E}(a + bX + cY) = a + b\mathrm{E}(X) + c\mathrm{E}(Y)$$

   By applying this result, show that for random variables X and Y:

   (a) $\quad \mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$

   (b) $\quad \mathrm{Var}(a + bX + cY) = b^2\mathrm{Var}(X) + 2bc\mathrm{Cov}(X,Y) + c^2\mathrm{Var}(Y)$

7. The following table shows the joint probability distribution of age and labour market status:

	Employed	Unemployed	Out of Labour Force
Young (16-25)	0.108	0.005	0.037
Old (26-60)	0.493	0.085	0.272

[1] Either look up the formula for binomial probabilities – e.g. in Thomas, or Wikipedia – or use Excel.

[2] This may seem obvious; if you want to prove it formally, follow the steps in the lecture slides for $\mathrm{E}(a + bX) = a + b\mathrm{E}(X)$, but use the joint distribution of X and Y.

Define two random variables: $A = 0$ if Young; 1 if old; $L = 0$ if employed; 1 if unemployed; 2 if out of labour force.

(a) Find the marginal distributions of age and labour market status, and the distributions of labour market status conditional on age.

(b) Are the two variables independent? How do you know?

(c) Find the mean $E(L)$ and the conditional means $\mathrm{E}(L|A = 0)$ and $\mathrm{E}(L|A = 1)$. Hence verify the law of iterated expectations:

$$\mathrm{E}(L) = E(L|A = 0)\mathrm{P}(A = 0) + E(L|A = 1)\mathrm{P}(A = 1)$$

8. (From *Stock and Watson.*) Suppose that the annual returns on stocks and bonds are R_s and R_b respectively, where R_s is random with mean 0.08 (8%) and standard deviation 0.07, and R_b is random with mean 0.05 and standard deviation 0.04. The correlation between R_b and R_s is 0.25. You have a sum of money to invest, and you put a fraction w into stocks and $(1 - w)$ into bonds, to obtain return $R = wR_s + (1 - w)R_b$. What value of w would you choose to maximise the expected return? What value of w would you choose if you wanted to minimise the variance of the return? What is the expected return on this portfolio?

   Hint: You will need to use the results from question 6.

9. Starting from the result 6a, together with the definition that X and Y are independent if and only if:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

   show (either for discrete or for continuous random variables) that if two random variables are independent, their covariance is zero.

10. Suppose that Z has a Standard Normal distribution, and define another random variable by $Y = Z^2$. Using the fact that the normal distribution has skewness 0 and kurtosis 3, find the mean and variance of Y. (Note: the distribution of Y is called the $\chi^2(1)$ distribution.)