1. The probability that a given character is miscopied when I send an email is 0.001, independently of all other characters.

   (a) If I send an email of 2000 characters, state

      (i) the exact distribution,

      (ii) a suitable approximate distribution,

   for the number, X, of miscopied characters. Use the approximate distribution to find $P(X = 0)$ and $P(X > 2)$.

   (b) If I send a second email, consisting of 3000 characters and independent of the first, state corresponding approximate distributions

      (i) for the number, Y, of miscopied characters in the second email,

      (ii) for the total number, Z, of miscopied characters in the two emails combined.

   Use this distribution of Z to find $P(Z = 4)$ and then use the approximate distributions of X, Y and Z to find the conditional probability $P(X = 2|Z = 4)$.

   (c) In the course of a week I send 50 emails, all independent and consisting of 100000 characters in total. State

      (i) the exact distribution,

      (ii) a suitable approximation,

   for the total number, W, of miscopied characters. Use the approximate distribution to find $P(W > 115)$.

2. The random variable X has the binomial distribution with probability mass function

$$p(x) = \frac{6!}{x!(6-x)!}\left(\frac{2}{5}\right)^x\left(\frac{3}{5}\right)^{6-x},$$

   where $x = \{0, 1, 2, 3 \ldots, 6\}$.

   (a) Derive E(X) and Var(X).

   (b) Find $P(X \leq 1)$

      (a) exactly,

      (b) by using a Normal approximation with a suitable continuity correction.

   (c) State circumstances under which a Normal approximation to the binomial distribution might be useful, and comment on your results.

   (d) Let $\bar{X}$ be the mean of a random sample of size 400 taken from the distribution of X. Calculate $Var(X)$, and use a Normal approximation to the distribution of $\bar{X}$ to find $P(2.35 < \bar{X} \leq 2.45)$. State with a reason whether or not you would expect your answer to be a good approximation to the exact probability.

3. The random variable X has the binomial distribution with probability mass function

$$\Pr(X = x) = \binom{2}{x}p^x(1-p)^{2-x}$$

   where $x = \{0, 1, 2\}$ and $0 < p < 1$.

(a) Write down $E(X)$, $Var(X)$ and $P(X = 2)$ in terms of the parameter $p$. Also find $P(X = 0|X < 2)$ and $P(X = 1|X < 2)$, simplifying your answers as far as possible.

(b) Let $Y = X_1 + X_2 + \ldots + X_{100}$ be the sum of 100 independent random variables, each distributed as X.

(i) Explain why Y has the $B(200, p)$ distribution.

(ii) Use a suitable approximation to find $P(Y > 140)$ when $p = 2/3$.

(iii) Use a suitable approximation to find $P(Y > 2)$ when $p = 0.02$.

(iv) Use a suitable approximation to find $P(Y \leq 197)$ when $p = 0.98$.

4. XYZ airline operates a baggage weight allowance of 25 kg per passenger. Check-in records show that the actual weight, W kg, of a randomly chosen passenger's baggage can reasonably be assumed to be Normally distributed with mean 24 and variance 1.

(a) Find $P(W > 25)$.

(b) A passenger with baggage weighing more than 25 kg is charged 5 for each kg by which the weight of his or her baggage exceeds 25 kg, all fractions of a kg being rounded up to the next whole number.

(i) If $C$ denotes the excess baggage charge in  for a randomly chosen passenger, find the probabilities $P(C = 0)$, $P(C = 5)$ and $P(C = 10)$.

(ii) Given that $P(C = 15) = 0.0013$ approximately and that $P(C > 15)$ is negligible, find $E(C)$ and $Var(C)$.

(iii) Assuming that 100 000 passengers independently fly with XYZ in a year, write down the mean and variance of CT, the total excess baggage costs paid to XYZ in a year. Use a Normal approximation to find the value of CT which is exceeded with probability 0.05.

(iv) Comment briefly on the assumptions made in your calculations in part (iii).

5. A blended wine is intended to comprise two parts of Sauvignon to one part of Merlot. The amounts dispensed to make up a nominal 75cl bottle of this wine are X cl of Sauvignon and Y cl of Merlot, where X and Y are assumed to be independent Normally distributed random variables with respective means 52 and 26 cl and respective variances 1 and 0.5625.

(a) Find the probability that the actual volume of wine dispensed into a bottle is less than the nominal volume.

(b) Find the distribution of X  2.2Y, and use this distribution to find the probability that the ratio of Sauvignon to Merlot is greater than 2.2. By a similar method, find the probability that this ratio is less than 1.8. Hence state to three decimal places the probability that the ratio of Sauvignon to Merlot in a randomly chosen bottle differs from 2 to 1 by more than 10

(c) Based on your final answer to part (b), and assuming that bottles are filled independently, write down

(i) the exact distribution,

(ii) a suitable approximation,

for the number of bottles in a thousand in which the ratio of Sauvignon to Merlot differs from 2 to 1 by more than 10%. Hence find the approximate probability that there are 10 or more bottles in a consignment of 1000 in which the ratio of Sauvignon to Merlot differs from 2 to 1 by more than 10%, giving your answer to three decimal places.

6. A coin has probability $p$ of showing heads and probability $1p$ of showing tails when it is tossed, independently each time. Let X be the random variable denoting the number of times the coin shows heads when it is tossed n times.

   (a) Show that

   $$\Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

   making clear all the steps of your reasoning. Under what conditions can the distribution of X be approximated by a Normal distribution?

   (b) A student uses the Normal approximation to approximate $P(X \leq 3)$ when $n = 20$ and $p = 0.2$. Calculate the answer he should obtain, use tables of the exact distribution of X to compute the percentage error in the answer, and comment briefly.

   (c) For integer $x \geq 1$, let N be the random variable denoting the number of tosses of the coin needed to obtain $x$ heads. Show from first principles that

   $$\Pr(N = n) = \binom{n-1}{x-1} p^x (1-p)^{n-x}$$

   where $n = x, x+1, x+2, \ldots$.

   (d) Evaluate this probability for the case p = 0.2, x = 3 and n = 20, and compare your result with the exact $P(X = 3)$ for the binomial distribution with the same values of $p$, $x$ and $n$.