1. Data on house prices in major Australian cities are collected.
   (a) What are some of the descriptive statistics that should be computed for this data set?
   (b) State one graph that should be produced for this data. What are two key features that you would expect to see in the graph?

2. Given that $P(A \cup B) = 0.7$, $P(B) = 0.4$ and $P(A \cap B) = 0.3$, find $P(B \mid A)$. Are $A$ and $B$ independent? Justify your answer.

3. A computer assembly business receives DVD burners from two suppliers. DVDA supplies 70% of the parts, but on average 3% of their parts are defective. DVDStar provides the remaining 30% of the parts, and on average 5% of their parts are defective. In order to maximise worker utility and efficiency, the company requires that the overall proportion of defective items does not exceed 3.5%.
   (a) Does the company meet its target proportion of defective items? Justify your answer.
   (b) If the company wants to further reduce the overall proportion of defective items it receives, which supplier should be asked to decrease its proportion of defectives? Give reasons for your answer.

4. Random variables $X$ and $Y$ are normally distributed with means 10 and 14 and standard deviations 2 and 4 respectively. In addition, the covariance between $X$ and $Y$ is -4. Let $W = 2X - 4Y$. What is the distribution of $W$?

5. A consultant is hired to collect data on the perceptions of consumers to a new health product. The consultant sends ten of her staff each to one of ten major shopping centres in Perth for a week (Monday to Friday during office hours), and each of them surveys passers by at random regarding their perception of the new product.
   (a) What the possible sources of bias in this survey?
   (b) Suggest improvements to the data collection process.

6. A sample of size 25 is taken from a population that is normally distributed with mean $\mu = 20$ and standard deviation $\sigma = 4$. Let $\overline{X}$ denote the sample mean.
   (a) State the distribution of the sample mean $\overline{X}$.
   (b) Evaluate $P(\overline{X} > 20)$.
   (c) Without computing the probabilities, state which is greater, $P(\overline{X} > 22)$ or $P(\overline{X} < 17)$? Justify your answer.
7. Weekly postage expenses for a large corporation have a mean of $800 and a standard deviation of $150. Additionally, the weekly postage expenses for “adjacent” weeks are not independent, and have a covariance of -1,000 dollars. Assume that weekly postage expenses are normally distributed. The Business Manager takes the total expenses for three successive weeks.

(a) Find the distribution of the total postage expense for this period.

(b) The Manager has budgeted $1,000 per week for postage expenses. What is the probability that the total expenses for the three weeks exceed the budget amount?

8. An insurance company knows that in the entire population of home owners, the mean loss from fire is $750 with a standard deviation of $600. The distribution of losses is strongly right skewed — many policies have a $0 losses, but a few have large losses. If the company sells 10,000, what is the probability that the average loss is greater than $765?

9. A researcher takes the salaries (in $100,000) of a random sample of five CEOs of large corporations and computes the following statistic:

\[ \bar{X}' = \frac{1}{15} \sum_{i=1}^{5} X_i, \]

where \(X_1, X_2, \ldots, X_5\) are the salaries. Let \(\mu\) denote the mean income of all CEOs in Australia, and \(\sigma\) the standard deviation.

(a) Find the mean and variance of \(\bar{X}'\).

(b) How does your answer compare with the mean and standard deviation of the sample mean of the five salaries?

(c) Assume that the salaries of CEOs is normally distributed with a standard deviation of 2. Find \(P(\bar{X}' > \bar{X})\).

10. A random sample of size \(n\) is selected from a population with standard deviation 10. How large a sample is required if \(P(\mid \bar{X} - \mu \mid > 1) < 0.05\)?

11. (June 2006 Examination) A pizza parlour estimates that the time (hours) it takes to deliver a pizza from when it is ordered is a continuous random variable \(X\) with probability density function

\[ f_X (x) = \begin{cases} \frac{4}{3} - \frac{2x}{3} & \text{if } 0 < x < 1, \\ 0, & \text{otherwise}. \end{cases} \]

What proportion of pizzas are delivered within half an hour of being ordered? (4 marks)

12. (June 2006 Examination) An investor estimates the joint probability distribution of the returns for a $1,000 investment for each of two stocks, \(X\) and \(Y\). From this distribution she computes the following summaries:

\[ E(X) = 100, E(Y) = 80, E(XY) = 7800, \text{Var}(X) = 100, \text{Var}(Y) = 400. \]

(a) Which stock should the investor prefer and why? (2 marks)

(b) Compute the covariance between \(X\) and \(Y\). (1 mark)
(c) If the yield for stock $X$ increases this year compared to last year, will the yield for stock $Y$ be likely to decrease or increase? Give a reason for your answer. (2 marks)

13. (2006 June Examination) Suppose the income of executives is normally distributed with the incomes of men and women independent of each other. The means and standard deviations (in $10,000) respectively are 20, 2.5 for men and 15, 2.0 for women. Let $F$ and $M$ denote the incomes of male and female executives respectively.

(a) What is the probability that a randomly chosen female executive earns more than a randomly chosen male executive? (4 marks)

(b) Find the probability that an executive couple has a combined income of more than $400,000. (3 marks)

(c) Ten executive couples are selected at random. What is the probability that at least one of the couples has a combined income of more than $400,000? (3 marks)

(d) A researcher selects 400 executive couples at random. Use a suitable approximation to estimate the probability that at most 20 of them have a combined income of more than $400,000. (5 marks)

©Nazim Khan, 2007