4B.1  (a) According to a national survey, 10.4% of the population of 18–24-year-olds in Australia are left-handed.
   i. In a class of 160 students, how many would you expect to be left-handed?
   ii. A survey result shows that there are actually 26 left-handed students in the class. What is the probability that in a class of 160 students, at least 26 of them are left-handed? What have you assumed?
(b) In a double-blind trial, 15 individual received treatment $T$ and 15 received a placebo $T'$. The follow-up testing indicated that objective improvement was exhibited by 8 of the individuals who received treatment $T$, but only 1 of the $T'$ individuals showed improvement.

Let $X$ denote the number of treated patients who showed improvement—so that we have observed $X = 8$ in this trial. Find $\Pr(X \geq 8)$ under the assumption that the treatment has no effect.

4B.2  (a) The expected number of deaths due to bladder cancer for all workers in a tyre plant over a 20-year period, based on national mortality rate, is 1.8. Suppose 6 deaths due to bladder cancer were observed over the period among the tyre workers. How unusual is this event? i.e. evaluate $\Pr(X \geq 6)$ assuming the national rate is applicable.

(b) A standard test for gout is based on the serum uric acid level. The serum uric acid level, $L$ mg/100L is approximately Normally distributed: with mean 5.0 and standard deviation 0.8 among healthy individuals; and with mean 8.5 and standard deviation 1.2 among individuals with gout.

Suppose we diagnose people as having gout if their serum uric acid level is greater than 6.50 mg/100L.
   i. Find the sensitivity of this test.
   ii. Find the specificity of this test.

Review problems

4B.3  (a) The Binomial, Hypergeometric and Poisson distributions are each concerned with the number of “events” that occur in a specified “time”. Specify what the “events” and the “time” means in each case.

(b) The following table indicates the approximations relating Hg, Bi, Pn and N distributions:

<table>
<thead>
<tr>
<th>sampling indept trials</th>
<th>Hg</th>
<th>Bi</th>
<th>Pn</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i(A)$</td>
<td>$i(B)$</td>
<td>$i(C)$</td>
<td>$i(D)$</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
</tbody>
</table>

Specify conditions $(A), (B), (C), (D), (E)$ for each of the indicated approximations to be reasonable (in the form of “something large” and/or “something else small”).

(c) If $X$ is integer-valued and $X^*$ the normal approximation, then $\Pr(4 < X \leq 10) \approx \Pr(3.5 < X^* < 9.5)$ or $\Pr(3.5 < X^* < 10.5)$ or $\Pr(4.5 < X^* < 9.5)$ or $\Pr(4.5 < X^* < 10.5)$?

(d) $X_1$ and $X_2$ are independent; $X_1 \overset{d}{=} N(6, 1^2)$ and $X_2 \overset{d}{=} N(4, 2^2)$.

Specify the distribution of each of $\frac{1}{2}(X_1 + X_2)$, $X_1 - X_2$, $2X_1 + X_2$ and $3X_1 - 2X_2$. 
Tutorial problems

4B.4 The quoted figure for the 5-year mortality rate for a particular form of leukaemia is 80%. In the hospital where you are a resident interested in neoplasm research, of the last five cases with this form of leukaemia, four are cured and one died. Do you feel you should check to see if some new procedure was used on the patients or that they were special in some other way, or pass off the cures to chance?

4B.5 (a) Among ten individuals, five are classified as $A$ (and five as $A'$); also, five are classified as $B$ (and five as $B'$). [For example: $A$ = female and $B$ = caucasian.] This means that the following tables are possible:

\[
\begin{array}{cccccccc}
0 & 5 & 1 & 4 & 2 & 3 & 3 & 2 & 4 & 1 & 5 & 0 \\
5 & 0 & 1 & 4 & 3 & 2 & 2 & 3 & 1 & 4 & 0 & 5 \\
\end{array}
\]

Assuming that the classifications $A$ and $B$ are independent, specify the probabilities associated with each of these possibilities.

(b) Of ten cancer patients, five are treated with drug $U$ and five are given a placebo (at random, of course). After one year, 1/5 of the individuals treated with $U$ have a recurrence of the cancer, while 4/5 of those given the placebo have a recurrence. Let $Z$ denote the number of patients treated with $U$ having a recurrence after one year, find $\Pr(Z \leq 1)$, assuming that $U$ has no effect.

4B.6 A survey of 100 individuals yielded the following table (in which most of the entries are missing):

<table>
<thead>
<tr>
<th>$B$</th>
<th>$B'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>$Z$</td>
</tr>
<tr>
<td>$A'$</td>
<td>$Z$</td>
</tr>
<tr>
<td>30</td>
<td>70</td>
</tr>
</tbody>
</table>

For example: $A$ = right-thumb over, $B$ = right-arm over.

i. What do you “expect” to get if $A$ and $B$ are independent? What is $E(Z)$ assuming independence?

ii. What range of values is it reasonable to expect for $Z$ assuming independence of $A$ and $B$? i.e. find $\text{sd}(Z)$ and compute $E(Z) \pm 2\text{sd}(Z)$.

iii. What would you think if $Z = 20$? (What are the other entries if $Z = 20$?)

4B.7 The mean number of patients admitted to the emergency room of a certain hospital is 2 per hour. What is the probability that, in an 8-hour shift, less than 10 patients were admitted to the emergency room?

4B.8 For $X \sim N(\mu = 50, \sigma^2 = 10^2)$, find

i. $\Pr(X \leq 47)$;

ii. $\Pr(X \geq 64)$;

iii. $\Pr(47 < X \leq 64)$;

iv. $c$, such that $\Pr(X \geq c) = 0.95$;

v. $c$, such that $\Pr(X < c) = 0.025$.

4B.9 Use suitable approximations to compute the following:

i. If $X \sim \text{Bi}(100, 0.8)$, find $\Pr(75 \leq X \leq 85)$.

ii. If $X \sim \text{Bi}(500, 0.01)$, find $\Pr(X \leq 3)$.

iii. If $X \sim \text{Pu}(25)$, find $\Pr(X \geq 31)$. 
**Binomial**

4B.10 Use the Statistical Tables to find the following probabilities:

i. \( \Pr(X \leq 3) \) for \( X \sim \text{Bi}(10, 0.2) \).

ii. \( \Pr(3 < X \leq 7) \) for \( X \sim \text{Bi}(15, 0.8) \).

iii. \( \Pr(1 \leq X \leq 3) \) for \( X \sim \text{Bi}(6, 0.25) \).

iv. \( \Pr(X \geq 16) \) for \( X \sim \text{Bi}(20, 0.75) \).

v. \( \Pr(3 \leq X \leq 6) \) for \( X \sim \text{Bi}(12, 0.4) \).

Check your answers using MINITAB.

4B.11 The differential is a standard measurement made during a blood test. It consists of classifying white blood-cells into five categories, one of which is neutrophils. The usual practice is to look at 100 randomly selected cells under a microscope and count the number of cells in each category. For a normal adult, the probability of a neutrophil is 0.6. Find the probability that, in a normal adult, between 50 and 75 inclusive of the 100 white blood cells will be neutrophils?

4B.12 Suppose 12 patients have been exposed to the HIV virus. If 30% of patients who have been exposed to the virus develop AIDS within 3 years, what is the probability that at most 4 of the 12 patients will develop AIDS within 3 years?

**Hypergeometric**

4B.13 In each of the following, define, if possible, a random variable having a hypergeometric distribution giving the appropriate values of \( n, R, N \). Further, write down an expression for the probability of the event which has occurred, but do not attempt to evaluate.

i. In Melbourne in September, \( \frac{2}{3} \) days are fine. This year there were 25 fine days.

ii. A box contains five red, three white and two black marbles. Two marbles are drawn out and they are both white.

iii. A hand of thirteen cards is dealt from a pack of 52, and it contains two aces.

iv. A shipment of 1000 items, 5% of which are defective, was submitted to a sampling inspection involving the taking of a random sample of 100 items from the shipment. The sample contained three defectives.

v. A committee of three was selected at random from a meeting consisting of ten men and five women, and two women were selected.

vi. Three digits (0-9) are chosen at random, and two of them are even.

vii. A distributor sells items in packages of 100 of which on average 10% are defective. A consumer tests a package of these items and finds two defective. The number of defective items is found to be \( k \).

viii. Suppose that 1000 fish netted in a lake are tagged and released. Several weeks later a new catch of 1000 fish is made, and in it is found 100 that have been tagged.

ix. To compare two drugs A and B, five patients are given drug A and five drug B. If five of the patients were cured, and five were not, find the probability that four of those were cured with drug A, assuming the drugs are equally effective.

4B.14 A random sample of three items is selected without replacement from a batch of ten items which contains four defectives. Find the probability that there is at most one defective item in the sample.

4B.15 Industrial quality control programs often include inspection of incoming materials from suppliers. If parts are purchased in large lots, a typical plan is to select 20 parts at random from a lot and inspect them. A lot is judged acceptable if one or fewer defective parts are found among those inspected. Otherwise, the lot is rejected and returned to the supplier. Find the probability of accepting lots that have 5% defective.
**Poisson**

4B.16 Given that $X \overset{d}{=} \text{Pn}(2)$, find $\Pr(X \leq 2)$, using the formulae and the Statistical Tables.

4B.17 Under microscopic investigation, on the average five particular micro-organisms are found on a $1 \text{ cm}^2$ untreated specimen. One such specimen was chemically treated. If it is assumed that the treatment was ineffective and if the Poisson process is applicable, what is the probability of finding fewer than three organisms?

4B.18 The number of cases of tetanus reported in a single month has a Poisson distribution with mean 4.5. What is the probability that there are at least 35 cases in a six-month period?

4B.19 Suppose the number of reported cases of hay fever in a town (caused by airborne dust and pollen) over a single calendar year is modelled with a Poisson distribution with mean 20.
   i. In the space of 3 months, 5 cases have been reported. Is this unusual?
   ii. Do you think it is appropriate to model the number of reported cases over a year by a Poisson distribution? Briefly give a reason why or why not.

**Normal**

4B.20 For $X \overset{d}{=} N(\mu = 10, \sigma^2 = 4^2)$, find
   i. $\Pr(X \leq 7)$;
   ii. $\Pr(X \geq 14)$;
   iii. $\Pr(12 < X \leq 17)$;
   iv. $c$, such that $\Pr(X \geq c) = 0.95$;
   v. $c$, such that $\Pr(X < c) = 0.025$.

4B.21 A new university subject has assessment that places a 30% weighting on a mid-semester test and 70% weighting on an end-of-year exam. Let $X$ and $Y$ be random variables describing the marks (out of 100) on the test and exam, respectively. We assume they are independent. You are told they are modelled using Normal distributions. In particular $X \overset{d}{=} N(\mu = 65, \sigma^2 = 20^2)$ and $Y \overset{d}{=} N(\mu = 75, \sigma^2 = 15^2)$.
   i. Write down an expression describing a student’s final mark
   ii. Specify the distribution of the final mark.
   iii. Calculate the probability that a student’s final mark will be enough to earn first-class honours (gain over 80)?
   iv. Give a reason why the true mark distributions cannot be Normal.
   v. Is it fair to assume $X$ and $Y$ are independent? State why or why not in one sentence.

4B.22 A medical trial was conducted to investigate whether a new drug extended the life of a patient with lung cancer.

Assume that the survival time (in months) for patients on the drug is Normally distributed with a mean of 30 and a standard deviation of 15. Calculate
   i. the probability that a patient survives for no more than one year;
   ii. the proportion of patients who are expected to survive for between one year and two years;
   iii. the largest number of months such that at least 80% of the patients are expected to survive;
   iv. the expected quartiles of the survival times.

The survival times (in months) for 38 cancer patients who were treated with the drug are as follows:

| 1  | 1  | 5  | 9  | 10 | 13 | 14 | 17 | 18 | 18 | 19 | 21 | 22 | 25 | 25 | 26 | 26 | 27 | 29 |
| 36| 38| 39| 39| 40| 41| 41| 43| 44| 44| 45| 46| 46| 49| 50| 50| 54| 54| 54| 59|

The sample mean is 31.1 months and the sample standard deviation is 16.0 months. Is there any reason to question the validity of the assumption that $T \overset{d}{=} N(\mu=30, \sigma=15)$?
4B.23 (a) Given $X \sim \text{Bi}(n=300, p=0.015)$, use a suitable approximation to find $\Pr(X \leq 4)$.
(b) Given $X \sim \text{Bi}(n=300, p=0.15)$, use a suitable approximation to find $\Pr(X \leq 40)$.

4B.24 Suppose $X \sim N(\mu=23, \sigma^2=3^2)$, $Y \sim N(\mu=30, \sigma^2=4^2)$, and $X$ & $Y$ are independent. Find $\Pr(X > Y)$. 
Mental Statistics  [NO CALCULATORS!]  [Time allowed: 3 minutes]

You should work these out using Statistical Tables, brain cells and pencil only. This is of the sort of table work it will be expected that you can do without difficulty in the final examination. Give your answers correct to three decimal places.

1. \( \Pr(X = 6) \), where \( X \sim \text{Bi}(10, 0.4) \).  \( \Pr(X = 6) = \) _________

2. \( \Pr(X = 6) \), where \( X \sim \text{Bi}(12, 0.8) \).  \( \Pr(X = 6) = \) _________

3. \( \Pr(X \leq 2) \), where \( X \sim \text{Bi}(15, 0.3) \).  \( \Pr(X \leq 2) = \) _________

4. \( \Pr(X \geq 16) \), where \( X \sim \text{Bi}(20, 0.6) \).  \( \Pr(X \geq 16) = \) _________

5. \( \Pr(X = 6) \), where \( X \sim \text{Pn}(45) \).  \( \Pr(X = 6) = \) _________

6. \( \Pr(2 \leq X \leq 6) \), where \( X \sim \text{Pn}(64) \).  \( \Pr(2 \leq X \leq 6) = \) _________

7. \( \Pr(X \geq 15) \), where \( X \sim \text{Pn}(75) \).  \( \Pr(X \geq 15) = \) _________

8. \( \Pr(Z > 1.234) \), where \( Z \sim \text{N}(0, 1) \).  \( \Pr(Z > 8) = \) _________

9. \( \Pr(X < 10) \), where \( X \sim \text{N}(10, 4^2) \).  \( \Pr(X < 10) = \) _________

10. \( \Pr(X < 15) \), where \( X \sim \text{N}(10, 4^2) \).  \( \Pr(X < 15) = \) _________

11. \( \Pr(X > 8) \), where \( X \sim \text{N}(10, 4^2) \).  \( \Pr(X > 8) = \) _________

12. \( c_{0.8}(Z) \), where \( Z \sim \text{N}(0, 1) \).  \( c_{0.8}(Z) = \) _________

13. \( c_{0.5}(X) \), where \( X \sim \text{N}(10, 4^2) \).  \( c_{0.5}(X) = \) _________

14. \( c_{0.25}(X) \), where \( X \sim \text{N}(10, 4^2) \).  \( c_{0.25}(X) = \) _________

15. \( c_{0.999}(X) \), where \( X \sim \text{N}(10, 4^2) \).  \( c_{0.999}(X) = \) _________