STATS 101/108 WORKSHOP

TEST PREP 2: CHAPTERS 4 AND 6

WEDNESDAY 8 SEPTEMBER, 2010

Students **MUST REGISTER** for all workshops with
The Student Learning Centre, 3rd Floor, Information Commons

Student Learning Centre

Topics we teach and can provide advice on include:

- ✓ Essay writing
- ✓ Computer skills
- ✓ Reading and notetaking
- ✓ Memory and concentration
- ✓ Report writing
- ✓ Test and examination skills
- ✓ Thesis and dissertation writing
- ✓ Tutorial skills
- ✓ Research skills
- ✓ Time and stress management
- ✓ Mathematics
- ✓ **Statistics**
- ✓ Oral presentation and seminar skills
- ✓ Language learning
- ✓ Specific learning disabilities
- ✓ Motivation and goal setting
- ✓ Survival skills (in the University system)

Programmes within SLC include:

- Te Puni Wananga
  Maori university tutors committed to enhancing Maori students' success

- Fale Pasifika
  Pacific Island tutors committed to enhancing success for Pacific Island students

- Students with impairments
  Learning assessments are available for students with specific learning disabilities; academic assistance is available for these students and those with mental health impairments.
  If you have any special learning requirements, please feel free to discuss this with Leila in person or via email.

- Academic English Conversation Groups
  Improve your academic English; develop communication skills including critical/creative thinking and clear expression of ideas and opinions. Weekly class held at the SLC on Thursdays, 3-5pm (during semester)

Statistical help available at the SLC

The Student Learning Centre (SLC) offers help for STATS 101/108 by offering:

- one-on-one tutoring help, and
- a number of workshops

One-on-one help

The SLC employs tutors specifically to help students with one-on-one assistance for STATS 101/108. One-on-one tutoring must be booked at SLC reception on the third floor of the Information Commons in person or by calling 373-7599 X 88850. Enquire at the SLC reception for available times.

Note: SLC tutors are not allowed to help students complete their assignments.

SLC STATS 101/108 Workshops

Any questions regarding STATS 101/108 workshops should be forwarded to:

Leila Boyle
SLC Statistics Co-ordinator
l.boyle@auckland.ac.nz

Workshops are run in a relaxed environment, typically set at a pace for those students that find the Statistics Department's tutorials too fast. All workshops allow plenty of time for questions. In fact, this is encouraged ☺

1) Saturday Workshops

These five 3-hour workshops are held on Saturdays throughout the semester to help students with different sections of the course.

2) Computer Workshops: Excel / PASW (SPSS)

These three computer-based workshops introduce students to the skills needed for Excel and PASW (SPSS) use in STATS 101/108 assignments.

3) Pre-test Workshops

These three workshops will cover the <u>basics</u> that you need for the test.

4) Pre-exam Workshops

These six workshops will cover the <u>basics</u> that you need for the exam.

Note: All workshops concentrate on questions reviewing the basic concepts, rather than questions on finer details. They are designed to assist students to achieve a pass; they are not designed to cover all material.

The timetable for these workshops is available with this handout. Currently the SLC website is still partly down so online enrolments are not available until further notice. Please enrol in each of your preferred classes at the Student Learning Centre by:

- **Going to the SLC in person**
- **Emailing slc@auckland.ac.nz with your name, ID number and the workshop/s you wish to attend.**
- **Calling the SLC reception on 373-7599 ext. 88850 and enrol over the phone. Make sure you know which workshop/s you want to enrol in and have your ID number handy.**

Useful Websites

- SLC webpage: **www.slc.auckland.ac.nz** (The SLC website currently has all functionality except online enrolment! Download an undergraduate brochure and enrol in workshops in person or by emailing/phoning the SLC Reception as per above instructions).
- Cecil: **https://cecil.auckland.ac.nz/**
- **Leila's website for STATS 101/108 SLC workshop handouts & information: www.stat.auckland.ac.nz/~leila**

Revision Notes

Chapter 4 - Probability

Look at blue pages for extra test/exam questions for practice

- A **probability** is a number between 0 and 1 that quantifies uncertainty.
- There are two main sources of probabilities that we will deal with.

  1. Probabilities using a model – some models that may involve equally likely outcomes are *tossing a coin* and *rolling a die*
  2. Probabilities from data

- A **random experiment** is an experiment where the outcome cannot be predicted.
- A **sample space** is the collection of all possible outcomes.
- An **event** is a collection of outcomes. An event **occurs** if any outcome making up that event occurs.
- If all **outcomes** are **equally likely:** $\text{pr}(A) = \dfrac{\text{no. of outcomes in } A}{\text{total no. of outcomes}}$
- The **complement** of an event A, denoted $\overline{A}$, occurs if A does not occur. A, and $\overline{A}$ are **mutually exclusive** events, ie they CANNOT occur at the same time.
- General probability rules:

  1. $\text{pr}(S) = 1$
  2. $\text{pr}(\overline{A}) = 1 - \text{pr}(A)$

- **Statistical Independence** – two events (A & B) are statistically independent if knowing whether B has occurred gives no new information about the chances of A occurring.

  i.e. $\text{pr}(A|B) = \text{pr}(A)$ and $\text{pr}(A \text{ and } B) = \text{pr}(A) \times \text{pr}(B)$

- **Two Types of Test/Exam Questions**

  1. Given a table of numbers/proportions, find the probability:
  - ☞ Easier question/s (can be between 1 and 3 of this type).
  - ☞ *May* want to convert the table into table of probabilities first.

2. Given a short story with proportions, percentages and/or counts about two factors (qualitative variables), find the probability:

- ☞ Harder question/s (can be 1 or 2 of this type).
- ☞ Need to *interpret* the story first, and then construct a table.
- ☞ Use the table to find 1 or 2 probabilities.
- ☞ Steps to constructing a table:

  Step 1: highlight numbers

  Step 2: highlight factors

  Step 3: define factor levels

  Step 4: label table

  Step 5: enter appropriate table total

  Step 6: enter row/column totals from story

  Step 7: enter cell numbers from story

  Step 8: enter remaining numbers by +/-

- **Four Types of Probability Calculations**

1. Probability of <u>AN EVENT</u> (basic/simple)

- ☞ pr(*A*) → pr (an event)

2. Probability of an event <u>AND</u> another event:

- ☞ pr(*A* and *B*) → pr(one event and another event)
- ☞ *Finding* pr(*A*) *<u>and</u>* pr(*B*) (intersection)

3. Probability of an event <u>OR</u> another event:

- ☞ pr(*A* or *B*) → pr(one event or another event)
- ☞ *<u>Add</u>* pr(*A*) *to* pr(*B*), then *<u>subtract</u>* pr(*A* and *B*)

} Use ***GRAND TOTAL (TABLE TOTAL)***

4. CONDITIONAL Probability:

☞ Harder to detect but will usually have one of the key words:

- "Given that…"
- "Of those…"
- "Among those…"

} Use **ROW TOTAL/S** *OR* **COLUMN TOTAL/S**

- Look for language that restricts you to part of the table instead of the whole table.
- pr (one event | another event) = $\text{pr}(A|B) = \dfrac{\text{pr}(A \text{ and } B)}{\text{pr}(B)}$

The following table, taken from *Facts New Zealand*, 1993, is used in **Examples 1** to **4**. It shows household income according to family type (A = Solo parent with one or more children, B = Couple with one or more children, C = Couple with no children, D = Single person, E = Other).

Household Income	**Family Type** A	B	C	D	E	**Total**
	No. of households (× 1000)					
0-9999	9	13	5	57	2	86
10,000 - 19,999	43	26	61	92	8	230
20,000 - 39,999	36	107	93	63	40	339
40,000 - 74,999	13	153	87	17	47	317
75,000 +	1	61	25	3	19	109
Total	102	360	271	232	116	1081

Example 1: If one of the 1,081,000 families is chosen at random, the probability that the total family income is less than $20,000 is:

pr (   ) = ____________________

= (4dp)

Example 2: Given that the family is of type A [Solo parent with child(ren)], the probability that the total family income is less than $20,000 is:

pr (   ) = ____________________

= (4dp)

Example 3: The probability that a randomly chosen family's total income is less than $20,000 or the family is of type A is:

pr (   ) = ____________________

= (4dp)

Example 4: The probability that a randomly chosen family is type A and their total income is less than $20,000 is:

pr (   ) = ____________________

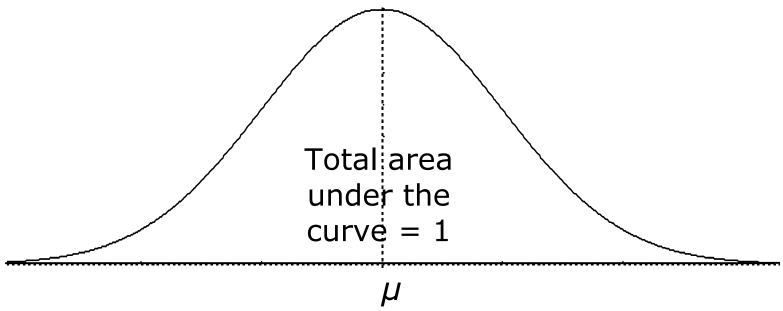
= (4dp)

Chapter 6 – Continuous Random Variables

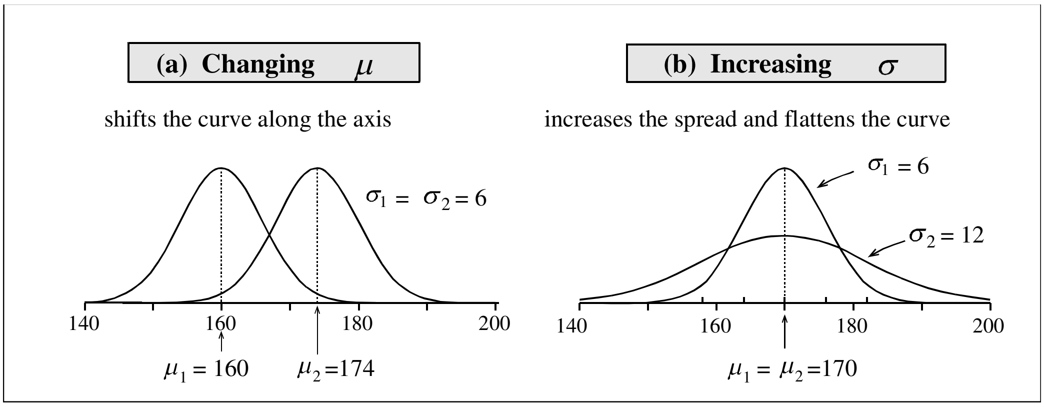
Look at blue pages for good notes and test/exam questions for practice

- A density curve is the probability distribution of a continuous random variable.
- There are no gaps between the values that a continuous random variable can take and therefore, when we calculate probabilities for a continuous random variable it does not matter whether **interval endpoints** are included or excluded

Normal Distribution



- The Normal Distribution has a probability density function curve, which is smooth, **bell-shaped**, and **symmetric**.
- The shape of the curve is solely determined by the parameters μ (mean) and σ (standard deviation).



- The Normal distribution is important because it:
  - fits a lot of data particularly well
  - can be used to approximate other distributions
  - is very important in statistical inference
- If X is a continuous random variable from a Normal distribution then:
  - $E(X) = \mu$ and $sd(X) = \sigma$
  - Probability distribution function of X is written: $X \sim \text{Normal}(\mu, \sigma)$

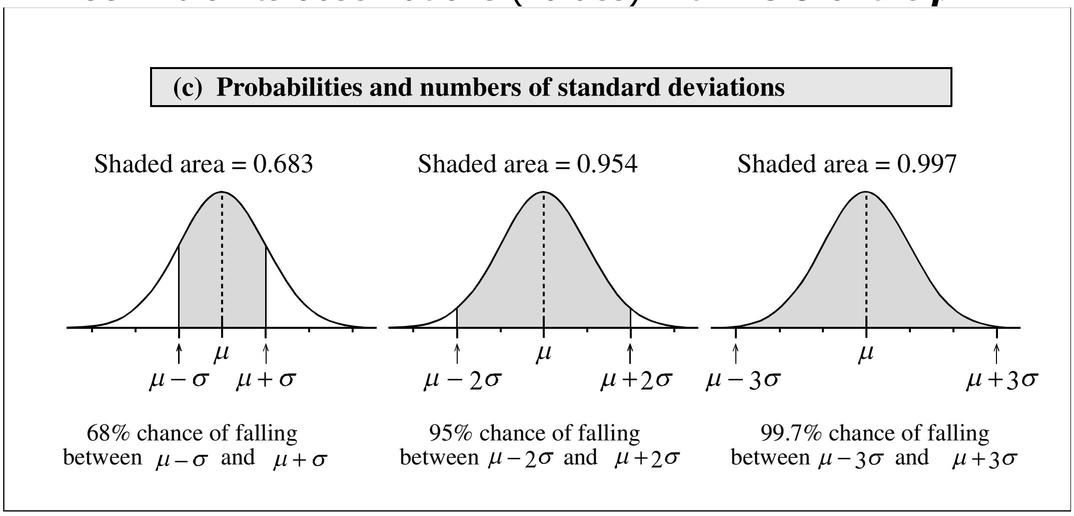
Chapter 6 test/exam questions

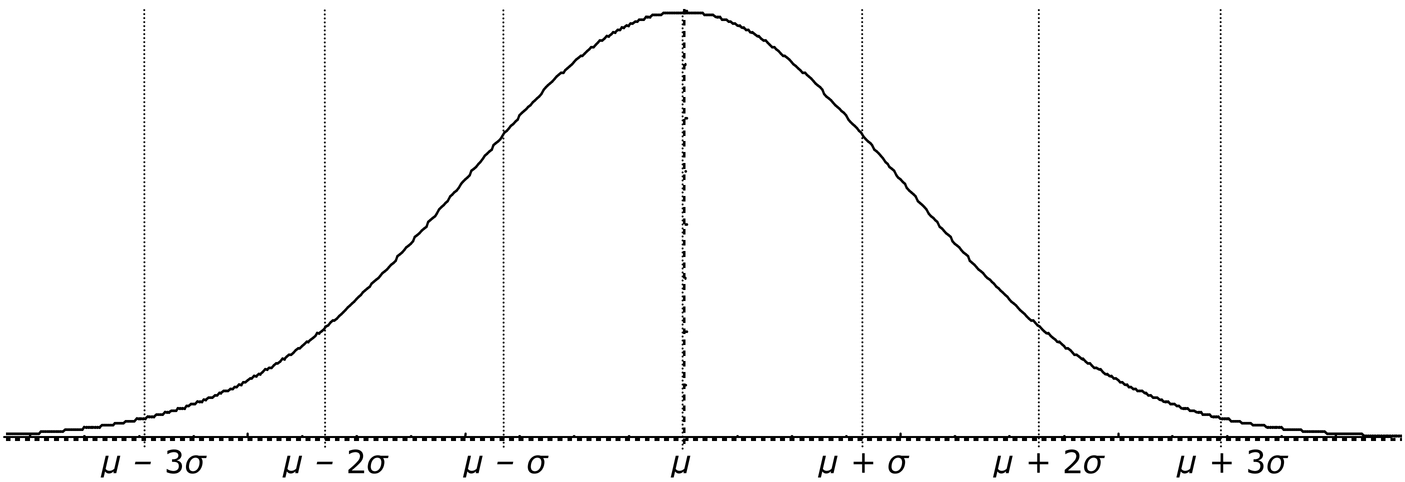
When doing Chapter 6 problems, it is sensible to draw a Normal curve and then mark on it what is known and what is unknown. There are **<u>three</u> (3)** types of Chapter 6 test/exam questions:

1. ***True/False (Normal) Chapter 6 problem***

There will be five statements, each about one or the other or both of two different Normal distributions. Use the <u>68-95-99.7% rule</u> or <u>*z*-scores</u> to determine whether four of the statements are true or false. The fifth statement will probably be comparing the means (centres/averages) and standard deviations (spread/variability) of the two distributions.

- **<u>68-95-99.7% rule</u>**: A population with a Normal distribution has:
  - ✓ 68% of its observations (values) within 1 $\boldsymbol{\sigma}$ of the $\boldsymbol{\mu}$
  - ✓ 95% of its observations (values) within 2 $\boldsymbol{\sigma}$ of the $\boldsymbol{\mu}$
  - ✓ 99.7% of its observations (values) within 3 $\boldsymbol{\sigma}$ of the $\boldsymbol{\mu}$



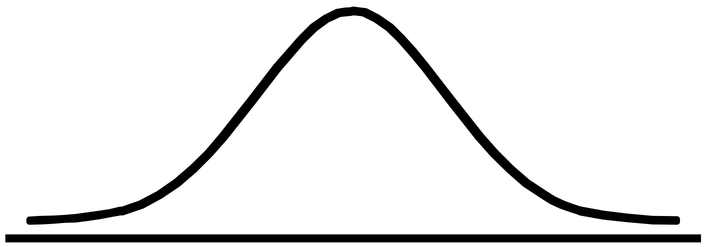


The **<u>z-score</u>**, $z = \dfrac{x - \mu}{\sigma}$, is a standardised number. It represents the number of standard deviations, σ, the value of x is away from the mean, μ. We can use z-scores to compare two or more different Normal distributions.

2. *Normal probability problem, i.e. find a probability associated with a number*

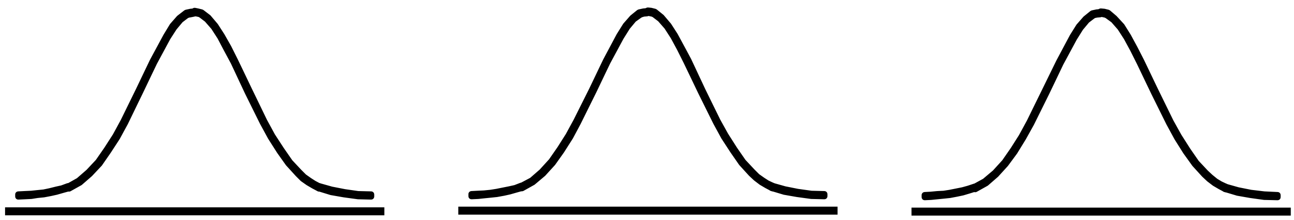
When finding a probability, shade the desired area under the curve and then devise a way to obtain it using lower tail probabilities which is all the computer can give. There are three types of Normal probability problems:

- *Find a lower tail probability (area)*
  The computer can find/give the answer directly.



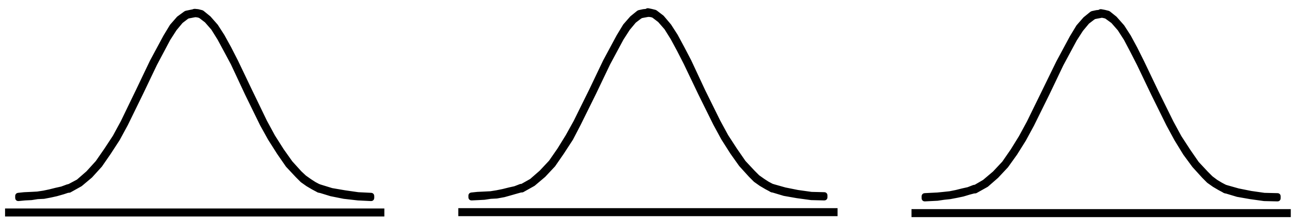
pr($X \leq x$)

- *Find an upper tail probability (area)*
  The computer cannot find/give the answer directly so subtract the lower tail from 1.



pr($X \geq x$) = 1 – pr($X \leq x$)

- *Find a probability (area) between two numbers*
  The computer cannot find/give the answer directly so subtract the lower tail beneath the smaller number from the lower tail beneath the larger number.

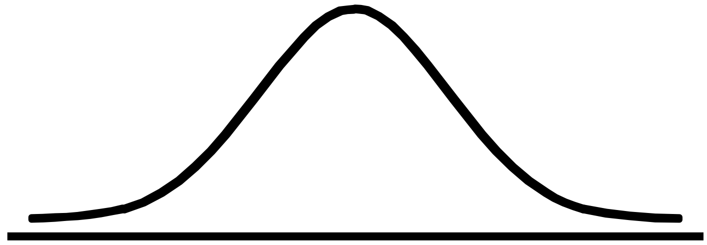


pr($a \leq X \leq b$) = pr($X \leq b$) – pr($X \leq a$)

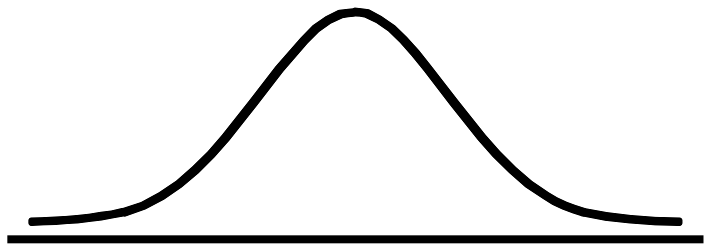
3. *Inverse Normal problem, i.e. find a number associated with a probability*

This type of problem occurs when we know the probability (e.g. the highest 10% in the class) and we need to find out the number associated with it, *x* (e.g. the mark). There are three types of inverse Normal problems:

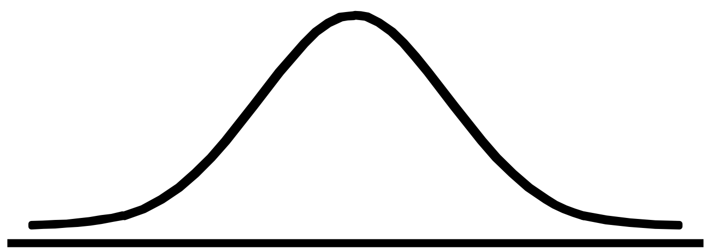
- *Given a lower tail probability, find the number associated with it*
  The computer can find/give the answer directly.



- *Given an upper tail probability, find the number associated with it*
  The computer cannot find/give the answer directly so subtract the upper tail probability from 1 & use the lower tail probability to find the answer.



- *Given a central area/probability, find the two numbers associated with it (the lower limit and the upper limit)*
  The computer can give the two limits as long as you use the lower tails/areas beneath each of them.



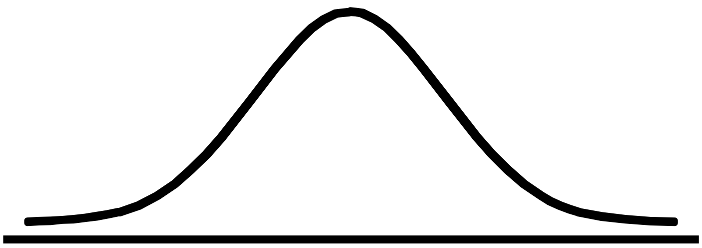
Examples 5 to **10** are about the following information.

The systolic blood pressure of 18-year-old women is Normally distributed with a mean of 120mm Hg and a standard deviation of 12mm Hg.

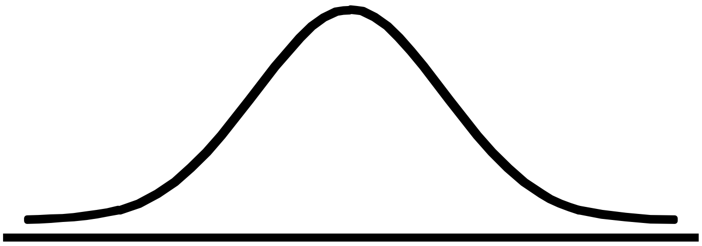
Normal with mean = 120.000 and standard deviation = 12.0000

x	Pr(X ≤ x)	Pr(X ≤ x)	x
110	0.202	0.050	100
120	0.500	0.250	112
125	0.662	0.350	115
130	0.798	0.500	120
150	0.994	0.650	125
		0.750	128
		0.950	140

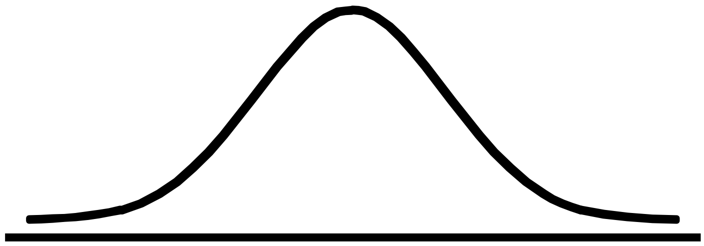
Example 5: The proportion of 18-year-old women with a systolic blood pressure between 130 and 150 is: [The table given is for a Normal(120, 12) distribution].



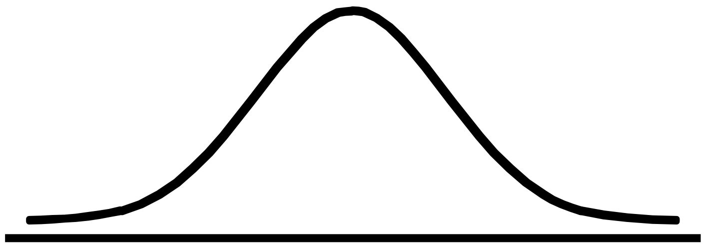
Example 6: The interquartile range of systolic blood pressure of 18-year-old women is:



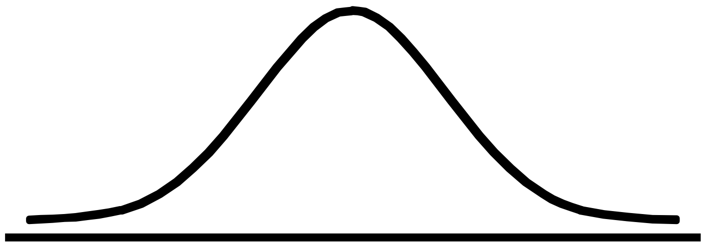
Example 7: 35% of the population of 18-year-old women would have a systolic blood pressure that is no more than:



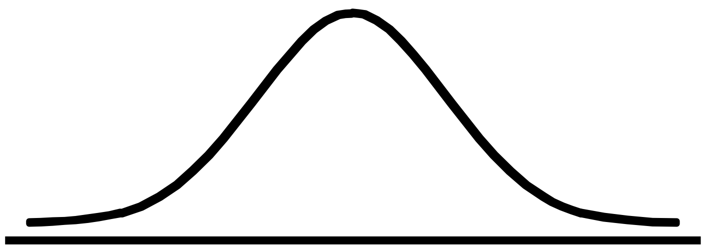
Example 8: The probability that an 18-year-old women's systolic blood pressure exceeds 110 is:



Example 9: The probability that an 18-year-old women with a systolic blood pressure of at most 125 is:



Example 10: The systolic blood pressure that 5% of the population of 18-year-old women would exceed is:



Chapter 4 and 6 – Questions

Questions 1 and **2** refer to the following information.

In the U.S. and in Europe, the presence of air bags in an automobile has become a key factor in deciding whether to purchase a particular model of automobile. A random sample of 93 automobiles was cross-classified by their size and by the level of air bag installation. Below is the cross-classification of the two variables, TYPE-OF-CAR and AIR-BAGS, for the 93 automobiles.

	AIR-BAGS			
TYPE-OF-CAR	**None in the car**	**Driver's side only**	**Driver's & passenger's side**	**Total**
Small	16	5	0	21
Sporty	3	8	3	14
Compact	5	9	2	16
Mid-size	4	11	7	22
Large	0	7	4	11
Van	6	3	0	9
Total	34	43	16	93

1. The proportion of cars in this sample that are small and have no air bags installed is:
   (1) 0.118
   (2) 0.471
   (3) 0.762
   (4) 0.366
   (5) 0.172

2. In this sample, the proportion of sporty cars having no airbags is:
   (1) 0.088
   (2) 0.214
   (3) 0.118
   (4) 0.366
   (5) 0.032

Questions 3 to **5** are about the following information.

A recent nation-wide telephone survey was conducted on behalf of the Alcohol and Public Health Research Unit (Wylie, 1996) in order to assess drinking patterns and alcohol related issues in New Zealand. A random sample of 4,232 people aged between 14 to 65 years was selected using telephone numbers from throughout New Zealand. The response rate was 76%. The gender/age composition of the 'top 10%' of drinkers (those who had consumed the most alcohol in the previous 12 months) is given in the table below.

	Male	Female	Total
14-17 years	11	4	15
18-19 years	25	11	36
20-24 years	92	21	113
25-29 years	42	7	49
30-39 years	56	11	67
40-49 years	32	3	35
50-65 years	35	3	38
Total	293	60	353

Table 7 : Gender/Age Composition of the 'top 10%' of Drinkers.

3. The proportion of the 'top 10%' of drinkers sampled who are under 20 years of age is:
   (1) 0.3729
   (2) 0.1445
   (3) 0.3201
   (4) 0.4646
   (5) 0.0425

4. The percentage of the 'top 10%' of drinkers sampled who are males and are in the 20-24 year age group is:
   (1) 26.1%
   (2) 32.0%
   (3) 81.4%
   (4) 31.4%
   (5) 43.7%

5. Which one of the following statements is false for the above table?
   (1) This two-way table of counts is an appropriate tool to use to explore the relationship between gender and age composition among the 'top 10%' of drinkers.
   (2) The gender classification in this table enables a comparison to be made between the number of males and females in the 'top 10%' of drinkers.
   (3) The gender of the respondent is a qualitative variable.
   (4) The age group in which a respondent is classified is a qualitative variable.
   (5) The age group classification in this table allows the number of the 'top 10%' of drinkers in their forties to be compared with the number in their fifties.

Questions 6 and **7** refer to the following information.

In recent years in Great Britain there has been a substantial decline in drink-driving and in the number of alcohol-related deaths on the roads. However, drink-driving remains a serious problem. The following table classifies 15730 casualties from road crashes in Great Britain in 1995 in which at least one of the drivers or riders involved was over the legal blood alcohol limit for driving.

Table: Great Britain Road Crash Casualties.

	Children	Adults under 60 years	Adults over 60 years	**Totals**
Pedestrians	120	510	70	700
Cyclists	50	120	10	180
Motor cyclists	20	860	20	900
Car drivers				
▪ Over limit	10	5050	150	5210
▪ Under limit	0	2330	190	2520
Car passengers	680	4600	240	5520
Other	30	640	30	700
Totals	910	14110	710	15730

6. The proportion of casualties who were adult pedestrians is:
   (1) 0.037
   (2) 0.829
   (3) 0.032
   (4) 0.942
   (5) 0.729

7. The proportion of cyclist casualties who were children is:
   (1) 0.003
   (2) 0.055
   (3) 0.278
   (4) 0.011
   (5) 0.198

8. In a STATS 10x/191 assignment, students were asked 'Do you have access to a computer at home?'. In total 922 students responded to this question, 305 of whom were **108** students, and 143 of whom were **191** students. A total of 708 students responded 'Yes', while 116 of the **101** students and 69 of the **108** students responded 'No'. The probability of a randomly chosen student responding 'Yes', given they are in **191**, is

   (1) 0.161

   (2) 0.124

   (3) 0.797

   (4) not possible to calculate.

   (5) 0.767

9. In a *Listener* (Feb. 4, 1995) article, it was reported from a survey of people 15 years and over that 83% approved of abortion if the mother's health was at risk. Thirty-six percent of people surveyed were in the age group 25-39 years. Of these, 85% approved of abortion if the mother's health was at risk.

   What is the probability that a person selected at random from the age group 15 and over approved of abortion if the mother's health was at risk or the person was in the 25-39 age group?

   (1) 0.340
   (2) 0.884
   (3) 0.485
   (4) 0.299
   (5) 0.891

10. Auditors developing systems to check the accuracy of regular tax returns for such taxes as GST, look at the changes in a firm's returns between tax periods. If the change is greater than some threshold, the firm's return is subjected to a rigorous audit. Such systems designed to detect cases of tax evasion must face the problem of false positives, that is, that the system indicates that the return is suspicious when, in fact, the change represents a real alteration in business conditions rather than tax evasions.

    Let E be the event that the firm is really attempting to evade tax, and T be the event that the system indicates possible tax evasion. Experience indicates that the incidence of tax evasion is 1 in 100 firms, while 90% of all cases of tax evasion are detected. Of those firms that are not really attempting to evade tax, the system indicates that 5% are possible tax evaders. The probability that a firm has actually evaded tax given that the system indicates evasion is:

    (1) 0.009
    (2) 0.050
    (3) 0.154
    (4) 0.900
    (5) 1.000

11. The probability of having a positive ELISA test given that you have HIV is 0.95. The probability of having a positive ELISA test given that you don't have HIV is 0.05. The probability of having HIV is 0.004. The probability of having HIV given a positive ELISA test is:

    (1) 0.076
    (2) 0.886
    (3) 0.071
    (4) 0.05
    (5) 0.95

12. A steel manufacturer produces pipes with a diameter that is approximately Normally distributed, with a mean of 10 cm and a standard deviation of 0.1 cm. Pipes with a diameter falling within the interval from 9.8 cm to 10.2 cm are acceptable, outside these limits they are rejected.

```
Normal with mean = 10.0000 and standard deviation = 0.100000

         x     P(X<=x)
    9.8000      0.0228
    9.9000      0.1587
   10.0000      0.5000
   10.1000      0.8413
   10.2000      0.9772
```

The proportion of acceptable pipes is approximately (use the computer output given above):

(1) 0.9544
(2) 0.9772
(3) 0.0228
(4) 1.0000
(5) 0.6826

13. If Z ~ Normal (μ_Z = 0, σ_Z = 1) then pr($-0.63 \le Z \le 0.86$) is (use the computer output given below):

(1) 0.541
(2) 0.459
(3) 0.230
(4) 0.931
(5) 0.431

```
Normal with mean = 0 and standard deviation = 1.00000

         x   P( X <= x)
   -0.8600       0.1949
   -0.6300       0.2643
    0.0000       0.5000
    0.6300       0.7357
    0.8600       0.8051
```

14. Which one of the following is not a feature of the Normal density curve:

(1) The curve is a symmetric, bell shaped and centred at μ.
(2) The mean is the same as the median.
(3) Roughly 50% of values lie within 1 standard deviation of the mean.
(4) The standard deviation σ governs the spread.
(5) Roughly 95% of values lie within 2 standard deviations of the mean.

Questions 15 and **16** are about the following information.

A research project investigated the effect of a drug on low-density lipoproteins (*LDL*) in quail. The quail were randomly divided into two groups – Group A and Group B. Group A was fed a diet mixed with the drug and Group B was fed the same diet but without the drug. At the conclusion of the project, the plasma *LDL* levels for Group A were found to be approximately Normally distributed with a mean of 62.2 and a standard deviation of 17.1. The corresponding plasma *LDL* levels for Group B were approximately Normally distributed with a mean of 67.2 and a standard deviation of 14.0.

Group A Normal(62.2, 17.1)	
x	$\Pr(X \le x)$
52	0.275
59	0.426
63	0.519
66	0.588
69	0.655
72	0.717
75	0.773
78	0.822
81	0.864

Group B Normal(67.2, 14.0)	
x	$\Pr(X \le x)$
34.5	0.010
44.2	0.050
49.3	0.100
51.6	0.133
61.7	0.347
67.2	0.500
85.2	0.900
94.6	0.975
99.6	0.990

15. The proportion of quail in **Group A** which have a plasma *LDL* level of at least 66.0 is about:

(1) 0.275
(2) 0.519
(3) 0.588
(4) 0.412
(5) 0.481

16. Approximately 10% of the quail in **Group B** had a plasma *LDL* level of less than:

(1) 34.5
(2) 61.7
(3) 49.3
(4) 85.2
(5) 44.2

Questions 17 to **19** are about the following information.

On most airlines, the maximum allowable weight for checked baggage is 20kg per economy class passenger. Excess baggage is defined to be any baggage over the 20kg limit. Airline staff know that the weight, X, of economy class passengers' checked baggage is approximately Normally distributed with mean μ_X = 15kg and standard deviation σ_X = 3kg.

The table given is for a Normal(15, 3) distribution. Use it to answer questions 22 and 23.

x	$Pr(X \leq x)$	x	$Pr(X \leq x)$
8.000	0.010	16.250	0.662
10.006	0.048	17.900	0.833
11.150	0.100	18.840	0.900
12.102	0.167	20.000	0.952
13.680	0.330	22.000	0.990
15.000	0.500	24.000	0.999

17. The probability that a randomly selected economy class passenger has an excess baggage is approximately:

(1) 0.048
(2) 0.500
(3) 0.167
(4) 0.952
(5) 0.000

18. Approximately 90% of economy class passengers have checked baggage weighing less than:

(1) 8.0kg
(2) 18.8kg
(3) 11.2kg
(4) 22.0kg
(5) 24.0kg

19. Let $\overline{X}$ be the mean checked baggage weight for a randomly selected group of 20 economy class passengers. The mean, $\mu_{\overline{X}}$, and standard deviation, $\sigma_{\overline{X}}$, of $\overline{X}$ are:

(1) $\mu_{\overline{X}}$ = 15kg, $\sigma_{\overline{X}}$ = 0.67kg
(2) $\mu_{\overline{X}}$ = 15kg, $\sigma_{\overline{X}}$ = 0.15kg
(3) $\mu_{\overline{X}}$ = 300kg, $\sigma_{\overline{X}}$ = 0.67kg
(4) $\mu_{\overline{X}}$ = 300kg, $\sigma_{\overline{X}}$ = 13.42kg
(5) $\mu_{\overline{X}}$ = 15kg, $\sigma_{\overline{X}}$ = 3kg

Questions 20 to **23** are about the following information.

A medical trial was conducted to investigate whether a new drug extended the life of a patient who had lung cancer. Assume that the survival time (in months) for patients on this drug is Normally distributed with a mean of 31.1 months and a standard deviation of 16.0 months.

X~Normal(31.1, 16.0)

x	pr(X<=x)	x	pr(X<=x)	x	pr(X<=x)
1	0.029969	9	0.083601	21	0.263938
2	0.034475	10	0.093626	22	0.284763
3	0.039523	11	0.104513	23	0.306341
4	0.045156	12	0.116288	24	0.328612

20. What is the probability that a patient survives no more than one year?

(1) 0.030
(2) 0.104
(3) 0.116
(4) 0.884
(5) 0.784

21. What is the probability that a patient survives at least 1 year but no more than 2 years?

(1) 0.4448
(2) 0.2241
(3) 0.0045
(4) 0.6314
(5) 0.2123

X~Normal(31.1, 16.0)

pr(X<=x)	x	pr(X<=x)	x
0.05	4.8	0.8	44.6
0.1	10.6	0.85	47.7
0.15	14.5	0.9	51.6
0.2	17.6	0.95	57.4

22. The length of time that only the longest surviving 15% of patients exceed is:

(1) 57.4 months
(2) 10.6 months
(3) 14.5 months
(4) 47.7 months
(5) 17.6 months

23. The central 60% of survival times are:

(1) 4.8 to 57.4 months
(2) 0 to 44.6 months
(3) 10.6 to 51.6 months
(4) 17.6 to 44.6 months
(5) 14.5 to 47.7 months

24. The z-score is:

(1) The number of standard deviations X is away from the mean.

(2) How many standard deviations the mean is away from X.

(3) $\frac{\mu - X}{\sigma}$

(4) positive if X is below the mean and negative if X is above the mean.

(5) $\frac{X - \mu}{\sqrt{\sigma}}$

25. If $x = 17$ is an observation from a random variable X, where X is distributed as Normal($\mu_X = 8$, $\sigma_X = 2.40$), calculate the z-score for X.

(1) −37.5

(2) 3.75

(3) 0.375

(4) −3.75

(5) 6

26. For the above z-score the best interpretation is:

(1) The random variable X is 3.75 standard deviations below the mean.

(2) The random variable X is -3.75 standard deviations above the mean.

(3) The random variable X is 3.75 standard deviations above the mean.

(4) The random variable X is -3.75 standard deviations below the mean.

(5) The random variable X is 3.75 standard deviations either side of the mean.

ANSWERS

1.	**(5)**	2.	**(2)**	3.	**(2)**	4.	**(1)**	5.	**(5)**	6.	**(1)**
7.	**(3)**	8.	**(3)**	9.	**(2)**	10.	**(3)**	11.	**(3)**	12.	**(1)**
13.	**(1)**	14.	**(3)**	15.	**(4)**	16.	**(3)**	17.	**(1)**	18.	**(2)**
19.	**(1)**	20.	**(3)**	21.	**(5)**	22.	**(4)**	23.	**(4)**	24.	**(1)**
25.	**(2)**	26.	**(3)**								