

PASW (SPSS)/Excel Workshop 2 – S2, 2010 – Answers

In Assignment 2 of STATS 10x you will need to use Excel or PASW (SPSS) to perform some calculations, that is, finding Normal probabilities and Inverse Normal values in **Question 3**.

You may also like to use Excel to check you've done your 'by hand' calculations correctly for your confidence interval in **Question 6**.

Instructions from your assignment sheet read:

Question guide

- Question 3 will require use of Excel, PASW or a graphics calculator for calculating probabilities from Normal distributions. Do not hand in any computer output for these questions. Use Excel, PASW or a graphics calculator to find the solutions. **DO NOT USE TABLES.**

Question 3. [9 marks] [Chapter 6]

Reminder: When calculating Normal probabilities, use PASW (SPSS), Excel or a graphics calculator. **Do not use tables.** Report any probabilities to **4 decimal places**.

On the following pages are some questions from the **Worked Examples** which you can find on Cecil.

Question 6 [Chapter 6] (helpful for Question 3, Assignment 2)

A medical trial was conducted to investigate whether a new drug extended the life of a patient who had lung cancer. 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, 25, 26, 27, 29,
36, 38, 39, 39, 40, 41, 41, 43, 44, 44, 45, 46, 46, 49, 50, 50, 54, 54, 59.

Sample mean \approx 31.1 months and sample standard deviation \approx 16.0 months.

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.

- (i) What is the probability that a patient survives for no more than one year?
- (ii) What percentage of patients survive for at least two years?
- (iii) What proportion of patients will survive between one year and two years?
- (iv) What is the highest survival time that 80% of patient survival times exceed?
- (v) What is the lower quartile of the survival times?
- (vi) Calculate the central 80% of survival times.

Question 6 Solutions

(a) Let X be the survival time (in months) for a patient on the drug.

$$X \sim \text{Normal} (\mu = \underline{31.1}, \sigma = \underline{16.0})$$

(i) $\Pr(X \leq \underline{12}) = \underline{0.1163}$ (4dp)

(ii) $\Pr(X \geq \underline{24}) = 1 - \Pr(X \leq \underline{24})$
 $= 1 - \underline{0.3286}$
 $= \underline{67.14\%}$ (2dp)

(iii) $\Pr(\underline{12} \leq X \leq \underline{24})$
 $= \Pr(X \leq \underline{24}) - \Pr(X \leq \underline{12})$
 $= \underline{0.3286} - \underline{0.1163}$
 $= \underline{0.2123}$ (4dp)

(iv) We want to calculate x such that $\Pr(X \geq x) = \underline{0.8}$

$$\Rightarrow \Pr(X \leq x) = \underline{0.2}$$

$$\Rightarrow x = \underline{17.6}$$
 (1dp)

The highest survival time that 80% of patient survival times exceed is 17.6 months.

(v) We want to calculate x such that $\Pr(X \leq x) = \underline{0.25}$

$$\Rightarrow x = \underline{20.3}$$
 (1dp)

The lower quartile of the survival times is 20.3 months.

(vi) For the central 80% of survival times: $\Pr(x_L \leq X \leq x_U) = 0.80$

$$\Rightarrow \Pr(X \leq x_L) = \underline{0.1}$$

$$\Rightarrow x_L = \underline{10.6}$$
 (1dp)

$$\Pr(x_L \leq X \leq x_U) = 0.80$$

$$\Rightarrow \Pr(X \leq x_U) = \underline{0.9}$$

$$\Rightarrow x_U = \underline{51.6}$$
 (1dp)

The central 80% of survival times fall between 10.6 months and 51.6 months.

Useful places to look for help by assignment question

Assignment question number	Worked Examples question number	Lecture Workbook page number
Q1	Q5	
Q2	Q4	
Q3	Q6 & 7	
Q4	Q8	
Q5	-	Use Part Time Work examples (Ch 8, pgs 1 & 2)
Q6	Q10	

Also, don't forget where else you can get assignment help! They are:

- Your **lecturer's office hours!** See Cecil for details – if they don't suit you, email or call them to book a time.
- **Statistics Assistance Area** – ask a tutor or your neighbour
- **Statistics Computer Lab** – ask a lab demonstrator or your neighbour
- The STATS 10x **forum**: www.stat.auckland.ac.nz/forum/10x

Question 10 [Chapter 8] (similar to Question 6, Assignment 2)

In April 1996 the New Zealand Consumers' Institute conducted a survey on home computer use. 7400 subscribers to Consumer Magazine were randomly selected and sent a survey form. Of those surveyed, 2730 had a computer for personal use at home. The respondents who had a home computer were given

a list of computer activities and were asked to indicate all of those that they engaged in. They were also asked to indicate the number of hours per week that they used their computer. The Consumers' Institute used the results to draw conclusions about subscribers **who own a home computer**. The results of the survey are given in the two tables below.

Computer Activities		Computer Use (hours per week)	
Word-processing	2621	Not used	27
Games	1502	Used for less than 2 hours	328
Spreadsheets	819	Over 2 and up to 7 hours (incl)	764
Accounting	655	Over 7 and up to 14 hours (incl)	710
Databases	437	Over 14 and up to 21 hours (incl)	546
Internet	328	Over 21 and up to 28 hours (incl)	109
Drawing	300	Over 28 and up to 35 hours (incl)	136
Desktop publishing	246	Over 35 and up to 44 hours (incl)	55
Fax / answering machine	82	Over 44 hours	55
		Total	2730

(c) Identify the sampling situation as (a) *Two independent samples*, (b) *Single sample, several response categories* or (c) *Single sample, two or more Yes/No items* in the following cases:

(i) Consider the results from the New Zealand Consumers' Institute survey on home computer use above. We want to compare the proportion of respondents who use their home computer for spreadsheets with the proportion of respondents who use their home computer for accounting.

Situation (c): One sample, two or more Yes/No items.

(ii) Consider the results from the New Zealand Consumers' Institute survey on home computer use above. We want to compare the proportion of respondents who use their home computer for over 7 and up to 14 hours per week and the proportion of respondents who use their home computer for over 14 and up to 21 hours per week.

Situation (b): One sample, several response categories.

(iii) As part of a nationwide telephone survey in 1998, data was collected on people who use their home computers for Internet activities. We want to compare the proportion of respondents (in this survey) who use their home computer for the Internet to the proportion of respondents from the Consumers' Institute survey who use their home computer for the Internet.

Situation (a): Two independent samples.