Year 13 Statistics EAS Workbook

Robert Lakeland & Carl Nugent

NuLake Ltd
Innovative Publisher of Mathematics Texts

Contents

Statistical Reports 3.12

This achievement standard involves evaluating statistically based reports.

Achievement	Achievement with Merit	Achievement with Excellence
• Evaluate statistically based reports.	• Evaluate statistically based reports, with justification.	• Evaluate statistically based reports, with statistical insight.

- This achievement standard is derived from Level 8 of The New Zealand Curriculum and is related to the achievement objectives
  - Evaluate a wide range of statistically based reports, including surveys and polls, experiments, and observational studies:
    - critiquing causal-relationship claims
    - interpreting margins of error.
- Evaluate statistically based reports involves identifying and commenting on key features in reports relevant to any conclusions made in those reports.
- Evaluate statistically based reports, with justification involves supporting the comments made by referring to statistical evidence and processes described in reports, relevant to conclusions made in those reports.
- Evaluate statistically based reports, with statistical insight involves integrating statistical and contextual information to assess the quality of reports with respect to conclusions made in those reports.
- Evaluating statistical reports requires familiarity with:
  - the statistical enquiry cycle
  - principles of experimental design
  - surveys and polls, including potential sources of bias
  - interpreting statistical inferences
  - interpreting a wide variety of statistical tables and graphs
  - analysing a wide variety of statistical situations
  - critiquing causal-relationship claims
  - interpreting margins of error.

Acknowledgements

The authors wish to acknowledge the inspiration and help we have received from the CensusAtSchool team and in particular Chris Wild and Matt Reagan, Statistics Department of the University of Auckland. The CensusAtSchool site is a must visit for students and teachers of statistics (www.censusatschool.org.nz and new.censusatschool.org.nz).

The iNZightVIT project is led by Professor Chris Wild from the Department of Statistics at the University of Auckland. He comes up with the big (and small) ideas that make iNZight intuitive and easy to learn.

The development of iNZight itself has been shared amongst many statistics students from the university, who have worked part-time on making various parts of iNZight.

Methods of Sampling

Methods of Sampling

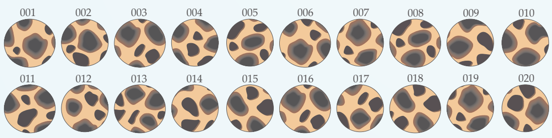
Researchers will use many different methods of collecting samples. The reason they will choose different ones depends upon time, cost and convenience. Sometimes the news media will select a method which will involve the subjects in the process and will increase the news content of the final report.

When evaluating statistical reports a major consideration is the non-sampling errors and how the method of sampling has a major bearing on the proportion of non-sampling errors. These methods are arranged from high credibility to low credibility.

Random Sample

A random sample is the selection of an unbiased sample from a population. This requires every element of the population to be numbered so that an algorithm or method can be devised to select a sample where each element has the same probability of being selected. If a sample of fifty is required then fifty different random numbers are generated in the range 1 to the number in the population. The fifty selected elements of the population are then inspected, questioned or observed to collect the information required.

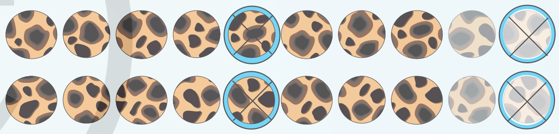
A random sample when correctly administered introduces no non-sampling errors and so is the preferred approach. Because it requires every element of the population to be numbered it is sometimes difficult and occasionally impossible. If we wanted to collect a sample of teenagers in Palmerston North it would be very difficult to get a list of all teenagers in Palmerston North. Similarly a sample of leaves on a tree would be impossible to number without chopping down the tree and removing every leaf.



Every element of the population must be numbered.

Systematic Sample

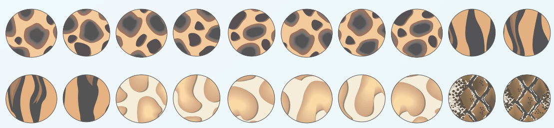
In a systematic sample every *n*th member of the population is selected until the desired sample size is achieved. For a systematic sample the population needs to be ordered in some way (e.g. in alphabetical order, seating order or the order they are made). For example, with our tree problem we could sample every 7th leaf along a branch until we have our sample size. It is often easier to systematically sample a population, but if the population has a natural pattern it is important that in selecting every *n*th member you do not select elements which are in some way not typical of the population. If a biscuit making machine produced biscuits in batches of 40 and you selected every 20th biscuit you risk selecting a repeating pattern, especially if the machine regularly made some biscuits a different size. If you selected every 27th biscuit then you are not in sync with the natural sequence and your sample is likely to be representative.



Be careful of selecting within a natural pattern. Here the biscuits at the end are under cooked but are over represented in the sample.

Stratified Sample

In an attempt to get a sample to better reflect the population, a research organisation sometimes extracts a stratified sample. For example, if you wanted to make sure that there was a minority group represented in the sample in the correct proportion. In a stratified sample, the research organisation identifies the significant characteristics for the population and the researcher randomly selects sufficient people in each category to reflect the proportions in the population. Stratified samples have increased credibility by making sure all groups in the population are represented in proportion. For example, if we knew the population of Palmerston North was 71% Pakeha, 14% Maori, 6% Polynesian and 9% Asian then we could select a random or systematic sample and when our sample was 71% Pakeha we would reject any further Pakeha and continue drawing a sample so each subset of the population was selected in proportion.



The sample must have the same proportions as the population.

Cluster Sampling

In a cluster sample a whole group or batch is selected to represent the population. Hopefully the

Achievement/Merit – Comment on each piece of research reported in the media.

These pieces of research are based on actual research but have been severely abridged for this exercise so no acknowledgement has been given as they may misrepresent the original research.

For each piece of reported research

a) Identify it as an observational study, experiment or randomised experiment.

b) Identify the explanatory variable.

c) Identify the response variable.

d) Identify the placebo if given.

e) If it is an experiment, is it single blind or double blind?

f) Did the research show a causal relationship?

g) Was there a possible confounding variable? If so identify it.

30. To test the effectiveness of nicotine patches on the cessation of smoking, researchers recruited 240 smokers. The volunteers were in good health and had a history of smoking at least 20 cigarettes per day for the past year, and were motivated to quit.

Volunteers were randomly assigned to receive either 22 mg nicotine patches or patches containing no active ingredients for 8 weeks. Neither the participants nor the nurses taking the measurements knew who received the nicotine patches.

After the 8-week period of patch use, almost half of the group wearing nicotine patches had quit smoking, whereas only one-fifth of the other group had.

31. A simple random sample of 120 college students, 60 females and 60 males, were recruited to find the relationship between the student's height and their average parents' height.

Each student was asked to list their own height and the heights of their mother and father. The heights of their parents were averaged to calculate a new variable called Parent Average Height.

Researchers were able to show that there was a relationship between the student height and that of their parents.

Two Independent Groups cont...

If we look at the MoE for the two groups, male and female (using the rule of thumb) we get:

MoE males $= \pm \dfrac{1}{\sqrt{235}} = 6.5\%$

MoE females $= \pm \dfrac{1}{\sqrt{265}} = 6.1\%$

$1.5 \times$ Average MoE $= 1.5 \times \left(\dfrac{6.5+6.1}{2}\right) = \pm 9.5\%$

The value we obtained by bootstrapping was 8.25% (approximately half the bootstrap interval, $^{-}0.6\%$ to 15.9%) which is between 1 x Average MoE (6.3%) and 2 x Average MoE (12.6%), so the rule of thumb of 1.5 x Average MoE for two independent groups is a good approximation.

Calculating the confidence interval by hand using the rule of thumb we have MoE males = 6.5%, MoE females = 6.1%, 1.5 x Average MoE = ± 9.5% and difference between females and males is

$47.2\%(\frac{125}{265}) - 39.6\%(\frac{93}{235}) = 7.6\%$, so the confidence interval for the difference between females and males who travel to school by car is 7.6% ± 9.5% = $^{-}1.9\%$ to 17.1%.

From the $^{-}1.9\%$ to 17.1% confidence interval we can infer that the percentage of females that travel to school by car is between 1.9 percentage points lower and 17.1 percentage points higher than the percentage of males that travel to school by car.

Because the $^{-}1.9\%$ to 17.1% confidence interval includes the value 0 we cannot however claim that the percentage of females that drive to school by car is greater than that of males.

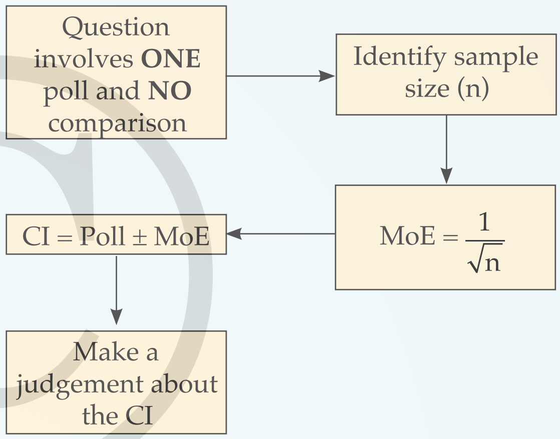
Be aware if the poll percentages, that is, the percentage of males and females that travel to school by car were outside the range 30% to 70% then the margin of error would be actually smaller than that given by the rule of thumb formula and the corresponding confidence interval would be smaller.

If you are dealing with poll percentages outside the 30% to 70% range you can still use the rule of thumb, but you should mention that the actual margin of error is smaller and that the corresponding confidence interval would be also smaller.

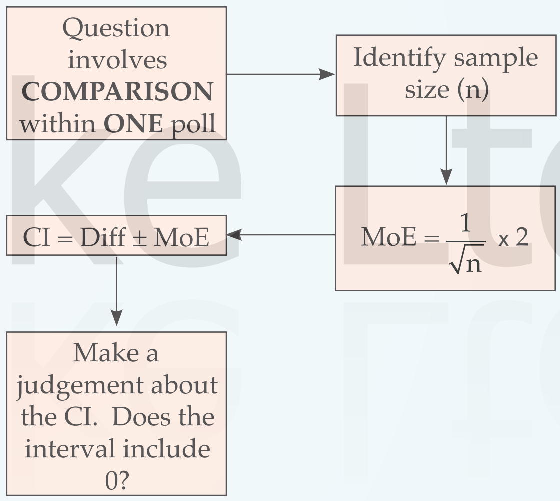
To summarise the approach you should use when asked to form a confidence interval study the diagrams on the right and adopt the following procedure.

1. Identify the type of question you are being asked.
2. Identify whether you are dealing with one poll and no comparison, a comparison within a single poll or two polls and a comparison.
3. Identify the sample size(s).
4. Calculate the appropriate margin of error using the applicable rule of thumb.
5. Calculate the applicable confidence interval.
6. Interpret the confidence interval and make an appropriate judgement in relation to the claim.

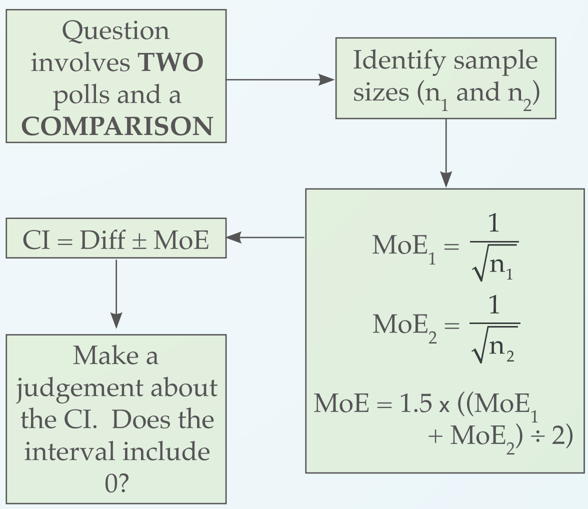
ONE POLL AND NO COMPARISON



ONE POLL AND A COMPARISON



TWO POLLS AND A COMPARISON



Probability 3.13

This achievement standard involves applying probability concepts in solving problems.

Achievement	Achievement with Merit	Achievement with Excellence
• Apply probability concepts in solving problems.	• Apply probability concepts, using relational thinking, in solving problems.	• Apply probability concepts, using extended abstract thinking, in solving problems.

- This achievement standard is derived from Level 8 of The New Zealand Curriculum and is related to the achievement objective:
  - Investigate situations that involve elements of chance
    - calculating probabilities of independent, combined, and conditional events.

- Apply probability concepts in solving problems involves:
  - selecting and using methods
  - demonstrating knowledge of concepts and terms
  - communicating using appropriate representations.
- Relational thinking involves one or more of:
  - selecting and carrying out a logical sequence of steps
  - connecting different concepts or representations
  - demonstrating understanding of concepts

  and also relating findings to a context, or communicating thinking using appropriate statements.
- Extended abstract thinking involves one or more of:
  - devising a strategy to investigate or solve a problem
  - identifying relevant concepts in context
  - developing a chain of logical reasoning
  - making a generalisation

  and also where appropriate, using contextual knowledge to reflect on the answer.
- Problems are situations that provide opportunities to apply knowledge or understanding of mathematical concepts and methods. Situations will be set in real-life or statistical contexts.
- Methods include a selection from those related to:
  - true probability versus model estimates versus experimental estimates
  - randomness
  - independence
  - mutually exclusive events
  - conditional probabilities
  - probability distribution tables and graphs
  - two way tables
  - probability trees
  - Venn diagrams.

Conditional Probability

Conditional Probability

Often when dealing with probability problems we do not use all the sample space for the experiment, but only a restricted part of it. We may ask about the probability of an event when we know a predetermined condition has been met.

For example, in question 27 we asked about the probability of a person having a disease given that they had tested positive. We restricted the sample space to just the people testing positive. We are able to calculate the probability by ignoring all results that did not test positive.

Informal Conditional Probability

If we are given information that restricts the acceptable answers of an experiment or problem, then we are limited to just this restricted sample space in calculating the probability. Consider a fitness test involving 100 students, 50 boys and 50 girls. Their results are given in this table.

Students 100	Boys 60	Girls 40
Pass	35	15
Fail	25	25

50 of the 100 students pass the test so we could say the probability of passing is 0.5. If we now change the question and ask about the probability that a boy passes the test we have restricted the sample space to the 60 boys. The probability of passing the test when we know it is a boy is 35 out of 60 or 0.583.

Again if we know that a person passed the test we restrict our sample space to the 50 students that passed. Now if we ask what is the probability that a person who passed is a boy we get 35 out of 50 or 0.7.

When we restrict the sample space to a given condition we can solve a conditional probability problem informally (without a formula).

Formal Conditional Probability

To indicate conditional probability we introduce a new notation namely, P(A | B) which is read as the **'probability of A given B'**. There is a formula that enables us to calculate a conditional probability without thinking of the restricted sample space. It is defined as

$$P(A \mid B) = P(A \text{ and } B) \text{ divided by } P(B)$$

or

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}$$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

The conditional probability formula

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

is given to you in the supplied formula sheet in the exam. You need to remember the symbol

$$P(A \mid B) = P(A \text{ given } B \text{ is true})$$

and $P(A \cap B)$ **= P(A and B) are true.**

Example

With just the 26 black cards from a pack of cards, you randomly select one card. What is the probability that you have a 'king' given that you have a picture card? The sample space is 2S, 3S, 4S, 5S, 6S, 7S, 8S, 9S, 10S, JS, QS, KS, AS, 2C, 3C, 4C, 5C, 6C, 7C, 8C, 9C, 10C, JC, QC, KC, and AC.
An ace is a picture card.

Informal approach with restricted sample space. The sample space where one is a picture card is

JS, QS, KS, AS, JC, QC, KC, and AC.

Of these eight results, two are kings.

Therefore the probability that you have a 'king' (K) given that you have a picture card (P) is

$$P(K \mid P) = \frac{2}{8}$$
$$= \frac{1}{4}$$

Formal approach with the formula.

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}$$

Being a king AND a picture card occurs 2 out of 26 times.

$$P(K \text{ and } P) = \frac{2}{26}\left(\frac{1}{13}\right)$$

Picture cards occurs 8 out of 26 times.

$$P(P) = \frac{8}{26}\left(\frac{4}{13}\right)$$

$$P(K \mid P) = \frac{\frac{2}{26}}{\frac{8}{26}}$$
$$= \frac{1}{4}$$

Example

In September 2012 New Zealand switched from analogue to digital television. One year earlier a Colmar Brunton poll was published which gave the following information on people's readiness for digital television.

Are you ready for digital TV?	NZ European	Other	Totals
Yes ready	4045	1673	5718
Not ready	828	390	1218
Totals	4873	2063	6936

a) What is the probability that a randomly selected person is ready for digital TV?

b) Given that a person is ready for digital TV, what is the probability that they are NZ European?

c) Is being 'NZ European' and being 'ready' for digital TV independent of each other?

a) P(ready) = $\frac{5718}{6936}$ (0.8244)

b) Number ready is first row = 5718

P(NZ Euro. | ready) = $\frac{4045}{5718}$ (0.7074)

c) P(NZ European) = $\frac{4873}{6936}$ (0.7023)

P(ready) = $\frac{5718}{6936}$ (0.8244)

P(NZ Euro. AND ready) = $\frac{4045}{6936}$ (0.5832)

P(NZ Euro.) x P(ready) = 0.5792

These are very close but not equal so ethnicity and readiness for digital are substantially independent but not totally. Ethnicity has some effect on whether a person is ready for digital TV.

Achievement – Answer the following probability problems.

80. The council conducted a survey on Saturday morning on the use of the skate board park. They got the following results.

Use of skate board park	Male	Female
Participant	21	7
Spectator	8	12

a) What is the probability that a randomly selected person was a spectator?

b) What is the probability that a random participant was male?

c) What is the probability that a randomly selected female was a spectator?

81. A school conducted a survey of students who were absent from the period one class and got the following results.

Students absent	Junior (Y10 and Y9)	Senior (Y11 to Y13)
Absent no excuse	11	0
Excused	23	32

a) What is the probability that a randomly selected absent student was a junior?

b) What is the probability that a student was absent without an excuse and was a junior?

c) Explain why being a senior and absent without excuse are mutually exclusive.

Probability Distributions 3.14

This achievement standard involves applying probability distributions in solving problems.

Achievement	Achievement with Merit	Achievement with Excellence
• Apply probability distributions in solving problems.	• Apply probability distributions, using relational thinking, in solving problems.	• Apply probability distributions, using extended abstract thinking, in solving problems.

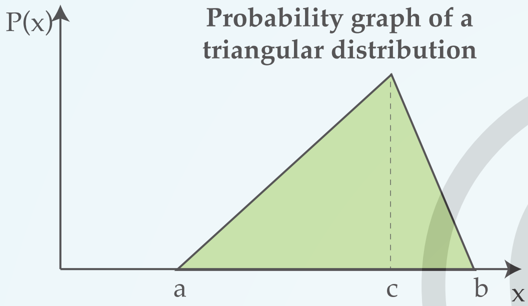
- This achievement standard is derived from Level 8 of The New Zealand Curriculum and is related to the achievement objectives:
  - Investigate situations that involve elements of chance:
    - calculating and interpreting expected values and standard deviations of discrete random variables.
    - applying distributions such as the Poisson, binomial, and normal.
- Apply probability distributions in solving problems involves:
  - selecting and using methods
  - demonstrating knowledge of concepts and terms
  - communicating using appropriate representations.
- Relational thinking involves one or more of:
  - selecting and carrying out a logical sequence of steps
  - connecting different concepts or representations
  - demonstrating understanding of concepts

  and also relating findings to a context, or communicating thinking using appropriate statements.
- Extended abstract thinking involves one or more of:
  - devising a strategy to investigate or solve a problem
  - identifying relevant concepts in context
  - developing a chain of logical reasoning
  - making a generalisation

  and also, where appropriate, using contextual knowledge to reflect on the answer.
- Problems are situations that provide opportunities to apply knowledge or understanding of mathematical concepts and methods. Situations will be set in real-life or statistical contexts.
- Methods include a selection from those related to:
  - discrete and continuous probability distributions
  - mean and standard deviation of random variables
  - distribution of true probabilities versus distribution of model estimates of probabilities versus distribution of experimental estimates of probabilities.

Triangular Probability Distribution

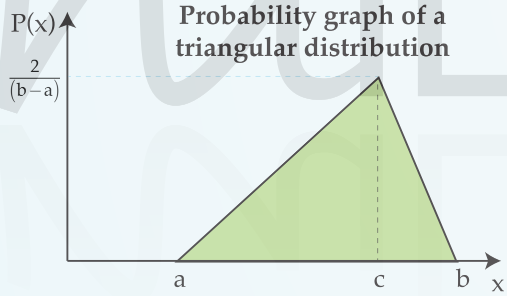
If we only know the minimum, maximum and mode of a distribution we assume the distribution is triangular. For example, if we are expecting a bus to arrive between a and b but typically it arrives at c then we assume the distribution is triangular.



We know the area of the triangle must be one so we can calculate the vertical height. Let the height be h and from the area of a triangle we have

$$\frac{1}{2}(b-a)h = 1$$

$$h = \frac{2}{(b-a)}$$



The height of the line from (a, 0) to (c, h) enables us to calculate the probability from a to c.

$$f(x) = \frac{2(x-a)}{(b-a)(c-a)}$$

The height of the line from (c, h) to (b, 0) enables us to calculate the probability from c to b.

$$f(x) = \frac{2(b-x)}{(b-a)(b-c)}$$

Extra Notes

We use a Triangular Distribution when we have

- **a distribution that is approximately Triangular OR**
- **we only have the Maximum, Minimum and Modal x values.**

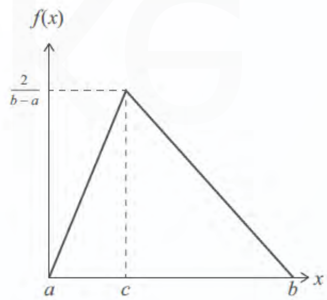
Fortunately you do not have to remember these equations or the graph as both are given to you in the examination tables booklet. See the clippings below.

Get a copy of the tables booklet and inspect it prior to any examination.

Triangular Distribution

The probability density function, $f(x)$, for a triangular distribution is defined as:

$$f(x) = \begin{cases} 0, & x < a \\ \dfrac{2(x-a)}{(b-a)(c-a)}, & a \le x \le c \\ \dfrac{2(b-x)}{(b-a)(b-c)}, & c \le x \le b \\ 0, & x > b \end{cases}$$



Area of a triangle $= \frac{1}{2}$base $\times$ height

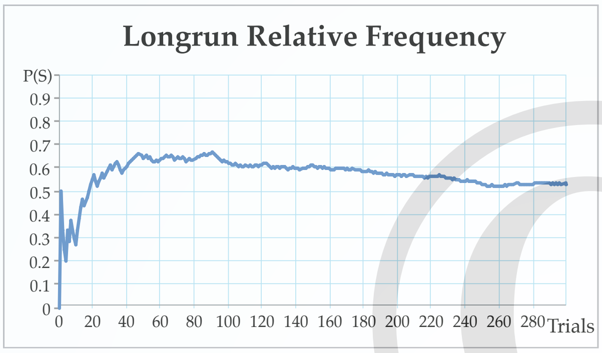
Continuous Uniform Distribution

The probability density function, $f(x)$, for a continuous uniform distribution is defined as:

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & \text{for } a \le x \le b \\ 0, & \text{elsewhere} \end{cases}$$

Example

Data has been collected of a basketball player shooting sets of three shots from the 'free throw' line. After 150 throws the probability of success was 0.613. After 300 throws the probability was 0.53.



Study the data and with justification

a) Identify the best model to find the probability of him getting two or three shots in out of a set of three.

b) Calculate the probability from your model of two or three shots scoring in a set of three shots.

c) Identify potential problems with your model explaining and justifying any conclusions.

a) Best model appears to be Binomial as

- the number of trials is **fixed**.
- each trial is **in or out**.
- each trial must be **independent** of the others.
- the **probability** of success at each trial must be the **same**.

These last two conditions are suspect but Binomial still seems the best model with $\pi = 0.613$

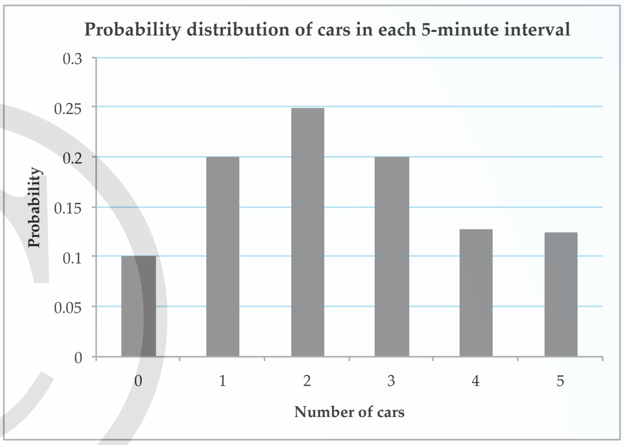
b) With $\pi = 0.613$

$$\begin{aligned} P(X \geq 2) &= 1 - 0.333 \\ &= 0.667 \end{aligned}$$

c) Is a subsequent shot independent of the shot before it? Does a sequence of misses affect the next shot? Also over time, the probability may increase (skill increase) or decrease (tiredness). Still you would expect the probability for the first 100 shots to be the best predictor of success.

Example

The manager of a fast food outlet recorded the number of cars in the drive-through lane every 5-minutes for the peak dinner period and plotted these on a probability distribution graph. The drive-through can have a maximum of five cars.



Study the graph and use it to answer the questions.

a) Identify the best model to find the probability of there being 'n' cars in a specific time period. Justify your choice of the model. What is the weakness of your model?

b) Calculate the mean and variance of the number of cars per 5-minutes and use this to check your model against your experimental results.

c) On the next night the manager randomly samples 5-minute slots. He inspected six 5-minute slots and three of these have no cars. Select a model to calculate the probability of three out of six having no cars and calculate the probability.

a) Best model appears to be a Poisson model as the parameter is a rate.

- cars should arrive at **random**.
- each trial must be **independent** of the others.
- cars **cannot** arrive **simultaneously**
- the **rate** will depend upon the time interval.

The drive-through is limited to five cars and this will distort the data.

b) Mean = 2.425 / 5-minutes Variance 2.244.
As variance is similar to mean it confirms the best model is Poisson.
$P(X = 0) = 0.088$, $P(X = 1) = 0.21$, $P(X = 2) = 0.26$, $P(X = 3) = 0.21$, $P(X = 4) = 0.13$, $P(X = 5) = 0.06$
The distribution is similar but $P(X = 5)$ is reduced.

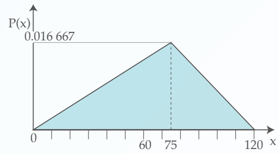
c) Binomial as probability should be constant ($\pi = 0.1$), trials are fixed ($n = 6$) and should be independent.
$P(X = 3) = 0.01458$

Page 144 Q 44. cont...

44. d) $P(10 < X < 20) = 0.6667$
$P(20 < X < 22) = 0.16$
$P(10 < X < 22) = 0.8267$

Page 145

45. a)



b) Area triangle = 0.025
$P(0 < X < 15) = 0.025$

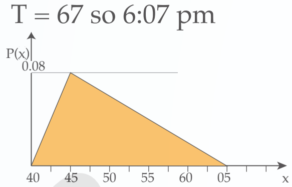
c) Area of trap. = 0.225
Area of triangle = 0.375
$P(60 < X < 120) = 0.6$

d) Area of trap. 1 = 0.225
Area of trap. 2 = 0.2083
$P(60 < X < 90) = 0.4333$

e) Probability of arriving after 6:15 is 0.375. Should be 0.5 for median.

f) Find triangle with area 0.5
$0.5 \times T \times 0.000222T = 0.5$
T = 67 so 6:07 pm

46. a)



b) $P(X > 45) = 0.5 \times 20 \times 0.08$
$= 0.8$

c) $P(X < 60) = 1 - P(X > 60)$
$= 0.95$

d) Probability of arriving after 9:45 is 0.8. Should be 0.5 for median.

e) Let T = time to 10:05
Find triangle with area 0.5
$0.5 \times T \times 0.004T = 0.5$
T = 15.8 minutes
Median = 9:49.2

Page 148

47. a) $P(0 < Z < 1.452) = 0.4268$ (7)
b) $P(Z > 1.452) = 0.0732$ (3)
c) $P(Z < 1.452) = 0.9268$ (7)
d) $P(^-1.452 < Z < 1.452)$
$= 0.8536$ (5)

Page 149

48. a) $P(Z < {}^-0.973) = 0.1652$ (3)
b) $P(^-0.895 < Z < 1.059)$
$= 0.3147 + 0.3552$
$= 0.6699$
c) $P(0.652 < Z < 2.074)$
$= 0.4810 - 0.2428$
$= 0.2382$
d) $P(^-1.953 < Z < {}^-1.049)$
$= 0.1217$

Page 152

49. a) $P(X > 108) = 0.0548$
b) $P(X < 95) = 0.1587$
c) $P(98 < X < 107) = 0.5746$
d) 159

50. a) $P(X > 4250) = 0.0345$
b) $P(2600 < X < 4000)$
$= 0.7950(1)$
c) Number outside range
= 205

Page 153

51. a) $P(X < 24) = 0.0913$ (2)
b) $P(X > 20) = 0.9962$
c) $P(25 < X < 30)$
$= 0.5890$ (89)
d) 61 days

52. a) $P(X > 30) = 0.0159$
b) $P(18.5 < X < 25)$
$= 0.6423$ (1)
c) $P(X < 17.5 \text{ or } X > 30)$
$= 0.0790$ (1)
d) 0.07898×650
= 51 or 52 boys

Page 154

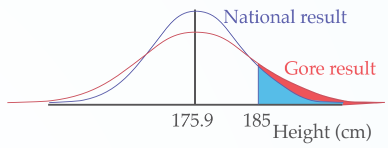
53. a) $P(X > 5) = 0.1587$
$[P(X > 5)]^3 = 0.0040$
b) $P(X < 3) = 0.0668$
$[P(X < 3)]^3 = 0.0003$

54. a) $P(X < 3) = 0.141\ 99$
$P(Y < 3) = 0.308\ 54$
Both = 0.0438
b) $P(X < 2) = 0.037\ 07$
$P(Y < 2) = 0.040\ 06$
Either = 0.0756
c) $P(X > 5) = 0.3605$
$P(Y < 2) = 0.0400$
$P(X > 5)$ and $P(Y < 2)$
$= 0.0144$

Page 155

55. a) $P(X < 0.1667) = 1.6 \times 10^{-6}$
(any answer that rounds to 4 dp).
Only one call so it does not call into question the parameters.
b) $P(\text{three} > 15) = (0.2602)^3$
$= 0.0176$
c) Time to answer each call is independent of any other call and the parameters are constant throughout the day. This is unlikely as they will change at times of high demand.

56. a) $P(X > 185) = 0.2222$
b) $P(X < 165) = 0.1798$
$P(\text{three} < 165) = 0.0058$
c) Not reasonable as each of the friends may feel more comfortable with someone close to their own height therefore not random.
d) Standard deviation of 17 yo male heights in Gore must be larger (16.1) as it has more results further from the mean.



Page 159

57. 73.3 m

58. Distinction = 63 or better
Merit = 55 to 62

59. Lower = 117
Upper = 131

Page 160

60. a) Mean = 17.4 mm
b) Reject > 20.5 mm

61. a) Std. Dev. = 18.5 kg
b) $P(X > 70.0) = 0.1587$

Page 163

62. a) $P(X > 65.5) = 0.0846$
b) $P(X < 57.5) = 0.2660$
c) $P(57.5 < X < 61.5) = 0.3802$
d) 0.0846×500
= 42.3
= 42 or 43 sacks

63. a) $P(X < 1.45) = 0.1908$
b) $P(X < 1.45)^2 = 0.0364$
c) $P(X < 1.55)^4 = 0.1743$

Page 164

64. a) $P(X > 5.025) = 0.3951$
b) $P(X > 4.975) = 0.4250$
c) $P(4.975 < X < 5.025)$
$= 0.0299$
d) Lower quartile = 4.4135 kg
So salmon recorded as 4.40 kg and lighter.

65. a) $P(X > 30.5) = 0.0401$
b) $P(X > 28.5) = 0.0861$
2.14 times or just over twice as likely.