No part of the candidate's evidence in this exemplar material may be presented in an external assessment for the purpose of gaining an NZQA qualification or award.

SUPERVISOR'S USE ONLY



+

91584







Mana Tohu Mātauranga o Aotearoa New Zealand Qualifications Authority

Level 3 Mathematics and Statistics (Statistics) 2024

91584 Evaluate statistically based reports

Credits: Four

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

Check that the National Student Number (NSN) on your admission slip is the same as the number at the top of this page.

You should attempt ALL the questions in this booklet.

Pull out Resource Booklet 91584R from the centre of this booklet.

Make sure that you have the Formulae and Tables Booklet L3–STATF.

Show ALL working.

If you need more room for any answer, use the extra space provided at the back of this booklet.

Check that this booklet has pages 2–12 in the correct order and that none of these pages is blank.

Do not write in any cross-hatched area (1/1/2). This area will be cut off when the booklet is marked.

YOU MUST HAND THIS BOOKLET TO THE SUPERVISOR AT THE END OF THE EXAMINATION.





© New Zealand Qualifications Authority, 2024. All rights reserved.

No part of this publication may be reproduced by any means without the prior permission of the New Zealand Qualifications Authority.

QUESTION ONE: UK ADULTS DON'T KNOW THE LOCATION OF THEIR BODY PARTS, NEW RESEARCH SHOWS

Refer to Resource One in the resource booklet to answer the following question parts.

(a) The Pall Mall survey media release did not include any information about the survey questions used to locate body parts, but HOW the survey questions were asked is important for correct interpretation of the results of the survey.

Give an example of how the responders' ability to locate body parts could have been measured in this survey.

An example of how the steppinder could have been measured is if they were asked, " can you locate blank body part" and were asked to do so by pointing to the location on an unlabelled diagram of the body I each body part is missing it's laber and the responder must click lit online) or point to lif in person) the body part they think they are naming. This increases accuracy of data results as it means the respondent can be made precise with their answers cg.

(b) Construct and interpret a confidence interval for the proportion of UK men and women who can confidently state where their rectum is located.

4 55% MARGIN OF ERMOR :

$Mat = \pm \int_{-}^{1}$	W
$M \circ E = \pm \sqrt{\frac{1}{2000}}$	that
M0F = 0.02236067977	men
MOE = 2.2%	state
CONFIDENCE INTERVAL	locat
(I= 65 % ± 2.2%	and
ct = 52.8% - 57.2%	size
	of err

> This confidence interval tells
us the can be pretty sure
that the proportion of UK
men and women the
smagt who can confidently
state where their rectum is
located is between 52.8%
and 57.2%. It the sample
size was larger, the margin
/ ot error would be smaller and
we could be more contident that
this data is accurately represent
titive of what is occurring back
atistice) 91584 2024 in the population.

03286

Mathematics and Statistics (Statistics) 91584, 2024

(c) One component to consider when evaluating statistical survey reports is who funded the study.

Explain whether there is a potential issue with the source of funding for this study.

There is a potential user with the source of tunding for this study as it was funded by a to private healthcare clininc, "Pall Mall Medical." This means it is possible they want to influencing people to join their clinic. We see this when they state at the end, "Pall Mall Medical is one of the leading Private Healthcare a cosmetic surgery provider in the VK, offering outstanding services to selt-paying patients a patients with private malthcare insurance." The type of language utical here eg. "outstanding' "leading", alongside the survey's statistics, is likely going to be influential to the reader, leading to sele this, we need to take into account hav the data has been displayed to vis and if it has been done in certain way in attempt to influence A different media report from the same study used the following headline: vis to think a certain

(d) A different media report from the same study used the following headline: *vs to think a certain* "Butt of the Joke: Survey finds people easily locate their rear, but the hunt for their 'private parts' *k* 13 is tricky!"

Is the headline correct with the implication that a higher proportion of UK men and women can confidently locate their rectum than their reproductive organs?

Construct ONE confidence interval and interpret this interval as part of your answer.

557.	52% * (meaning it is possible there is no difference between proportions or even that reproductive organs prop-
redum = 55%	This confidence interval diplays that we
reproductive organs = 52%	can be pretty sure the proportion of UK
% diff = 3%	men & women in this survey who could locate
Margin of Error:	their rectum is between 1.5% lower
$MoE = \pm \sqrt{\frac{1}{2000}} \times 2$	and 7.5% higher than the proportion
MOE = 0.04472135955	who could confidently locate their reproductive
 MOE = 4.5%	organs. As the contidence interval contains
confidence interval:	zero, there is not sufficient evidence to
CI = % dift ± MOE	support this claim. that a higher proportion
CI = 3% ± 4.6%	of VK nen & woman can confidently locate
(I = -1.5% - 7.5%) Mathematics and Statist	tics (Statistics) 91584, 2024 03286

(e) The Pall Mall survey was carried out by OnePoll, a UK online research agency. OnePoll describes their 'panel' as a community of UK adults who have registered to take part in market research and opinion polls. When panellists join OnePoll, they agree to participate in at least three surveys per year for which they get paid the equivalent of between 40c and \$NZ1. Members of the panel can claim payment when their total has reached the equivalent of approximately \$NZ50.

OnePoll sent the Pall Mall survey to a sample from their panel who fit the target population.

Discuss whether the OnePoll sampling method is likely to generate a representative sample of all UK men and women.

Support your answer with statistical reasoning, including clearly identifying the target population, the sampling frame, and at least one potential non-sampling error.

The target population is all UK men and woman. the sampling Frame is all One foil members/parelists. This means it may not be representitive of the whole target population of all UK men & women because of bias. People a part of One Poll may be a certain type of perion eg. more opinionated etc, meaning it is not a representitue sam of the population as people who don't want to share there opinions are left out of the sample transe & thus, the sample * It states here that "Ore poll sent the ... survey to a sample from their panel who Fit the population." Because we don't know the size of the sample frane, we cannot be sure that this was done through a random sampling method (eg. maybe in order to pick a group who fit they had to hand pick participants , maning they're turning away others). This possibility of non-random sampling being used this means the sample that may have had errors and so might not generate a truly representive sample of all UK men and women.

* This makes results inaccurate a thus, not representitive.

QUESTION TWO: SOCKS-OVER-SHOES PROVEN AS MEANS OF REDUCING WINTER FALLS

Refer to Resource Two in the resource booklet to answer the following question parts.

(a) The study reported results on three outcome variables including: self-rated slipperiness, observer-rated slipperiness, and the time it took to descend the slope.

5

Discuss why the researchers chose to use self-rated slipperiness as their primary outcome instead of the time it took to descend the slope.

This is likely are to the fact that the speed of descent doesn't neccessarily correlate with spipperiness, the for example, a person might find running (walking tast with bigger strides down a hill their natural or pretented method to combat spipperiness but another person might find slower walking (smaller stride their prettored method. (If the researchers then view speed of descent, this would inaccurately results & because both find to would it slippering but the fast time of the first preson suggest otherwise.)

(b) Identify the explanatory and response variables for this study.

Explanatory variable: Whether or not shoes had a pair of socks over them (when a participant was walking down a hill.)

Response variable: Primary outcome : Selt-rated Slipperiness and (But there was also abserver rated slipperiness and speed of descent).

- (c) It was not possible to blind the participants or the outcome assessors to the treatment allocation in this study.
 - (i) Describe why blinding was not possible in this situation.

Blinding Was not possible in this situation because it was obvious who was reciping treatment, and who was not; This is not something that you (sock-meaning (normal show people) (normal can add to their shoes without them knowing, it is easily visible.

Mathematics and Statistics (Statistics) 91584, 2024

(ii) Discuss TWO measures that the researchers used in this study to minimise the impact of not blinding the participants and, for each, describe how these measures may have helped minimise bias in this study.

Measure one: Participants * assessing were simply told that researchers were interested in assessing the performance of different types of Footwear and different types of socks worn over the top. This minimises bias because then the participants will respond honestly to what they felt with or would socks. If they were told that the test was to see if socks were grip, they might be influenced to try convince themselves they are getting a certain result eg. Measure two: Spess supperivers, when it they didn't know the study's purpose they would have likely said something different eg. that they didn't notice a difference.

J Participants footwear was photographed for later reference to encourage participants and assessed to think that characteristics of footnear were important. This minimises bias as it leads their minds away from the true difference between the two groups (sock a no sock). This would further incline them to answer honestly about socks and socks only but rather about this footwear as a whole.

(d) Recruiters for the study administered a baseline questionnaire to participants, which included details about potential confounding variables, shown in Table 2(a).

Discuss TWO of these variables and, for each, describe how they may have been confounding, and their potential effect on the findings of this study.

Variable one: Amount of women. It is important that this confounding variable is controlled because otherwise the data inaccurately, for example, the treatment all group if anter are group had women and the other all men, you might get

inaccurate results which display that socks may have not made as significant of a difference as expected but really, that is due to this confounding variable of them all being women meaning their footwear may have been harder to walk in

in general eg. high heels. This makes it hard for never was to identify causation swhy the Variable two: Evariable is contourneling) and so must be controlled.

Age Range. Age is another important confounding variable to control as it can significantly impact an individuals walking speed, Eq. if the treatment group was all people over 50 and the treatment make it appear as though the socks much a bigger difference to walking speed than they really did (because older people naturally walk slover, & younger people naturally walk and with more stability

4) If both of these were not controlled, it would be difficult to identify causation and thus, inferences made about the whole population might be inaccurate as data is not representitive. Mathematics and Statistics (Statistics) 91584, 2024

03286

NY HAVY

QUESTION THREE: 50% OF MEN SURVEYED THINK THEY COULD LAND A PASSENGER PLANE – EXPERTS DISAGREE

Refer to Resource Three in the resource booklet to answer the following question parts.

(a) The *New Zealand Herald* headline for Resource Three (a) is: "50% of men surveyed think they could land a passenger plane – experts disagree."

Explain how evidence from this report has been used to generate this headline.

The report added up the somewhat confident men & the very confident men to create the statistic in this head ine: 26% somewhat confident, 20% very confident. #(20+26 = 46%). They then rounded this up (likely because 50% will grab peoples attention more than 46% would).

(b) Identify one of the survey percentages in the YouGov study, and explain why it would not be appropriate to use the rule of thumb for the margin of error to construct an approximate 95% confidence interval for the population proportion, using this survey percentage.

opproxi It would not be appropriate to create an 95% confidence interval For the population proportion of Female who Feel very confident. that they could "sately land a passenger plane ... relying" on air traffic control " because this statistic was File. 7% st sits outside of the 30 -70% range that is eligible to be weed for rule of thumb nethod. It you did the presented rule of thimb method for this 7%, the margin of error would be over estimated and would thus be too wide, meaning it is thus not an appropriate method.

Question Three continues on the next page.

Mathematics and Statistics (Statistics) 91584, 2024

8 Resource Three (a) and Resource Three (b) are from two different studies. (c) Discuss the main differences between the designs of these two studies. Support your answer with statistical reasoning, including clearly justifying the study designs, the types of inferences (claims) that can be made, and the assumptions needed to do so. **Resource Three (a): YouGov study** Study design: Observational. - Observing a group with out controlling, manipulating anything about what is going on back in the population A causal claim can not be made because there may Inferences: be lurking variables impacting data results. However, assumptions can be made that, it backed 4 1 confounding up by further evidence, could be accurate claims. Assumptions: (source) You would need to assume consider where the study came From eg. online poll, interview etc, how many people nere sampled and how the sample was chosen (hopefully by random sampling), bias a inaccuracies etc to be able to use this data accurately to make claims Resource Three (b): University of Waikato study

Study design: Experimental. - Had a control group and had a treatment group in which something was changed I manipulated.

Inferences: A causal claim to an be made be cause it is an experimental study w/ a control * treatment group.

Assumptions: A causal claim can be made assuming that the sample was random sampling i that it is representitive of the population. Additionally, lurking variables such as gender, age, IR etc. must be controlled (taken into account for an accurate causal claim to be made. furthermore, we a causal chaim can only be made assuming chances of bias' were minimized as much as possible - study states part, was 'volunteers' for viewing the video, hinting that its possible self - selection bias may be prefert.

03286

イント・イ

(d) In Resource Three (b), the study researchers concluded that "We found watching the video inflated people's confidence that they could land a plane."

Using evidence from Figure 4, write TWO comparative comments that support the study researchers' conclusion.

Commentione: For both questions, the confidence level was wider than 'no video' at the 100% on the graph. This displays it is

possible the claim is correct and confidence was increated

I move people in this range to the for 'video' than for 'no video'.

Comment two: The for both graphs, the median is higher the For 'video' than for 'nor video' eg." sits around #35 or 4 for 'video' and around 30 for 'no video'. This supports the claim

as it deplays a difference between medians of at least 5, meaning it is possible klaim is correct. More evidence may be needed to confirm this though, such as, or confidence interval comparing the two. In the study from Resource Three (b), participants either watched a video or not, then were told: graups.

"Now we're going to ask you a few questions. Don't try to analyse and puzzle things out – just go with your gut feel or hunch. Respond as quickly as possible within a couple of seconds. Remember this is an emergency situation."

Participants answered the following questions in this order:

(e)

- Q1: "How confident are you that you would be able to land the plane without dying?" (0 = not at all confident, 100 = very confident)
- Q2: "How confident are you that you would be able to successfully land the plane as well as a pilot could?" (0 = not at all confident, 100 = very confident).

In a second repeated study (with new participants), the researchers randomised the order of these two questions, with approximately half of the participants asked the questions in the order above, and half asked in the opposite order.

Explain why the researchers asked the questions in different orders. In the second study, The researchers asked the questions in different orders in order to minimize bigs, so the participants don't find a pattern in what they're being asked. They could then identify whether the order of the questions has any impact on a participants answers (by comparing to the first trial we original participants.) eq. if Q2 asked first, maybe they the participants would provide a different answer to the question compared to if they were asked it second, following the 'without dying' version # -> Researchers can now identify whether this occurred or not to Mathematics and Statistics (Statistics) 91584, 2024 they can control it. 03286

Excellence

Subject: Statistics

Standard: 91584

Total score: 21

Q	Grade score	Marker commentary
		1(a) - A sensible suggestion – "pointing to the location on an unlabelled diagram"
		1(b) – Correctly calculates confidence interval and correctly interprets the confidence interval with correct context "We can be pretty sure" and "UK men and women"
		1(c) – Identifies the company that funded the study is "private" and describes one way the survey can be used to company's advantage "influencing people to join clinic". Describes how this issue could impact the study "type of language used"
		1(d) – Correctly calculates confidence interval:
		Correct interpretation of the confidence interval with the following features:
		 Doubt – "we can be pretty sure".
		 Population – "UK men and women".
One	E7	 Upper and lower limits of confidence interval described correctly – "1.5% lower and 7.5% higher".
		 Variables identified –" locate rectum" and "locate reproductive organs".
		Correct Judgement of claim with the following features:
		Confidence interval contains zero.
		Evidence not sufficient to make the claim.
		 Claim statement – "there is not sufficient evidence that a higher proportion of UK men and women can confidently locate their rectum than their reproductive organs".
		1(e) – target population identified. Explanation of the likely representativeness of the sample discussed "People a part of OnePoll may be a certain type – e.g., more opinionated, meaning not representative" And mentions uncertainty around the sampling of the sample frame. There is no explanation of how/why the issue identified (opinionated people) affects the finding of the study.

		2(a) – Does not identify or discuss self-rated slipperiness or
	E7	wearing of socks to determine slipperiness as the study goal.
		2(b) – Correctly identifies the explanatory and response variables in context.
		2(c) – (i) and (ii) are marked together. Candidates must correctly answer part (i) before any credit for responses in part (ii) are considered.
		 (i) – candidate mentions that blinding is not possible and gives a clear, valid reason why in this study.
Two		(ii) For both Measure One and Measure Two, the candidate has identified a valid measure used in the study to minimise the impact of not blinding, and has described how it may have helped minimise the impact of not blinding in this study (for both measures).
		2(d) – The candidate has identified and described two potential confounding variables. However, the candidate response for Variable one does not discuss HOW the confounding variable may interact with the response variable(s). With the Variable two response they have discussed how " <i>might make it appear the socks make a bigger difference to slipperiness rating/walking speed</i> ".
		To gain the highest grade for part d, the candidate needed to discuss how either confounding variable listed in Variable one or Variable two is not well controlled by the study design, using the group percentages from Table 2a as evidence.
	E7	1(a) - A sensible suggestion – "pointing to the location on an unlabelled diagram".
		1(b) – Correctly calculates confidence interval and correctly interprets the confidence interval with correct context "We can be pretty sure" and "UK men and women".
Three		1(c) – Identifies the company that funded the study is "private" and describes one way the survey can be used to company's advantage "influencing people to join clinic". Describes how this issue could impact the study "type of language used".
		1(d) - Correctly calculates confidence interval.
		Correct interpretation of the confidence interval with the following features.
		 Doubt – "we can be pretty sure".
		 Population – "UK men and women".
		 Upper and lower limits of confidence interval described correctly - "1.5% lower and 7.5% higher".

 Variables identified –" locate rectum" and "locate reproductive organs".
Correct Judgement of claim with the following features:
Confidence interval contains zero.
• Evidence not sufficient to make the claim.
 Claim statement – "there is not sufficient evidence that a higher proportion of UK men and women can confidently locate their rectum than their reproductive organs".
1(e) – target population identified. Explanation of the likely representativeness of the sample discussed "People a part of OnePoll may be a certain type – e.g., more opinionated, meaning not representative" And mentions uncertainty around the sampling of the sample frame. There is no explanation of how/why the issue identified (opinionated people) affects the finding of the study.