The Simple Theory of Informal Rules

Nugzar Nachkebia
Westmount School, Hawkes Bay Campus, Hastings, New Zealand
nugzar.nachkebia@westmount.school.nz

Marina Alexander
Woodford House, Havelock North, Hastings, New Zealand
Marina.alexander@woodford.school.nz

Keywords: Statistics education; Informal statistical inference; Probability; Sampling variation

Abstract

The latest changes in the New Zealand high school mathematics curriculum include basic study of informal and formal statistical inference. Statistics education researchers have made a great effort to deliver these changes to the New Zealand high schools. However, informal statistical inference concepts currently adopted in New Zealand high schools are visually oriented and rule based, but not closely linked to the context. The aim of this paper is to investigate the relationships between the key probability and informal statistical inference concepts. Based on the simplified classical formal statistical inference we derive and critically evaluate the rules for the informal statistical inference currently adopted in New Zealand high schools.

1. Introduction

Modern technology enables students to easily learn the calculations related to the statistical inference as some user friendly interactive software like TinkerPlots (Konold and Miller 2005), are available now. However, research indicates that students typically have difficulties in understanding the statistical reasoning and statistical concepts. See for example, Batanero (2000); Saldahna and Thompson (2002); Chance et al. (2004); Castro-Sotos, et al. (2007); Liu &Thompson (2009); Harradine et al. (2011); Pratt et al. (2008).

Research indicates also that some high school mathematics teachers have the similar misconceptions and difficulties in the understanding of the statistical inference as their students do. See for example, Vallecillos (1999); Lui & Thompson (2009); Harradine et al. (2011).

A pivotal point in statistics education was in 1999 when SRTL (Statistical Reasoning, Thinking and Literacy) series of biannual international research forums were organized. To overcome the difficulties in the understanding of statistical inference, statistics education researchers at the SRTL – 4 forum came to a consensus that students should be provided with the opportunity of learning the informal statistical inference (Ben-Zvi et al., 2007).

As the importance of the statistical inference is increasingly recognized worldwide, its concept has become part of the high school curriculum in many countries. The latest changes of the New Zealand high school mathematics curriculum include the basic study of the statistical inference. For example, year 11 and 12 students (15 – 16 years old) learn sampling methods, sample distributions, sampling variability and informal inference; year 13 students (17 years old) learn confidence intervals and formal statistical inference (NZ Ministry of education, 2010).
The limited research, available about New Zealand teachers' understanding of statistical inference, indicates that some high school mathematics teachers could share the same difficulties in understanding the ideas of the statistical inference as their colleagues around the globe. See for example Pfannkuch (2006) and Wild (2006).

Wild et al. (2010, 2011) have made a great effort to deliver ideas of informal statistical inference to New Zealand high schools. The concepts developed by Wild and his colleagues represent an innovative approach that has given invaluable support to teachers and their students in understanding the key concepts of the sampling variability while rightfully keeping the focus on the statistical enquiry cycle. The visual approach developed by Wild et al. (2010, 2011) (www.censusatschool.org.nz/2009/informal-inference/WPRH/) enables students to develop their understanding of sampling variability and make inference about populations from the single samples. As Wild et al. (2011) mentioned, the support around the concepts of the informal statistical inference has been primarily targeted to the mainstream of students only – 'We want to arrive at conceptions of statistical inference that are accessible to the bulk of students and not merely an intellectual *élite*' (p.251, Wild et al., 2011). As a result of the target audience mentioned above, the informal inference rules developed by Wild et al. (2010, 2011) are empirically derived by simulating normal data, but not closely linked to the context. However, many high school students have expressed a desire for a deeper understanding of the connections between the informal inference rules and their existing understanding of the key probability and statistical concepts (P. Doyle, personal communication, August 1, 2012; New Zealand Association of Mathematics Teachers, Hawkes Bay Branch meeting, personal communication, November 11, 2012). We still have some students, who prefer to know the basis for a rule rather than applying it without understanding.

Based on the simplified classical formal statistical inference in this article we unpack and explain the theory behind the informal statistical inference rules currently implemented in the NZ mathematics curriculum.

2. Basis for the rules of informal inference

The main concept of classical statistical inference is based on the fact, that if the values of a variable of a population have normal distribution, the mean values of all possible random samples of the population can be modeled very well by the normal distribution. The mean value of a single random sample (μ_{sample}) varies from sample to sample and the standard deviation of the sample means is inversely proportional to the square root of the sample size (Kendall and Stuart, 1967).

It is impossible to figure out the exact value of the population mean if we have a single random sample. Instead, we can give the interval estimate of values for the population mean (μ).

In this paper we use the 95% confidence interval for the population mean – i.e. we use the property of the normal distribution that 95% of the data lie about within the 2 standard deviations of the mean (McGill et al., 1978):

$$\mu_{sample} - 2\frac{\sigma}{\sqrt{n}} < \mu < \mu_{sample} + 2\frac{\sigma}{\sqrt{n}} \quad \text{(Inequality 1)}$$

When analyzing the statistical information visually, box plots have been commonly used since they were introduced by Tukey (1970). To find the relationship between the standard deviation of a sample (S) and its interquartile range (IQR) for a normally distributed data, we use inverse standard normal distribution (Fig.1), which leads to:

$$IQR = 1.35S \qquad \text{(Equation 2)}$$

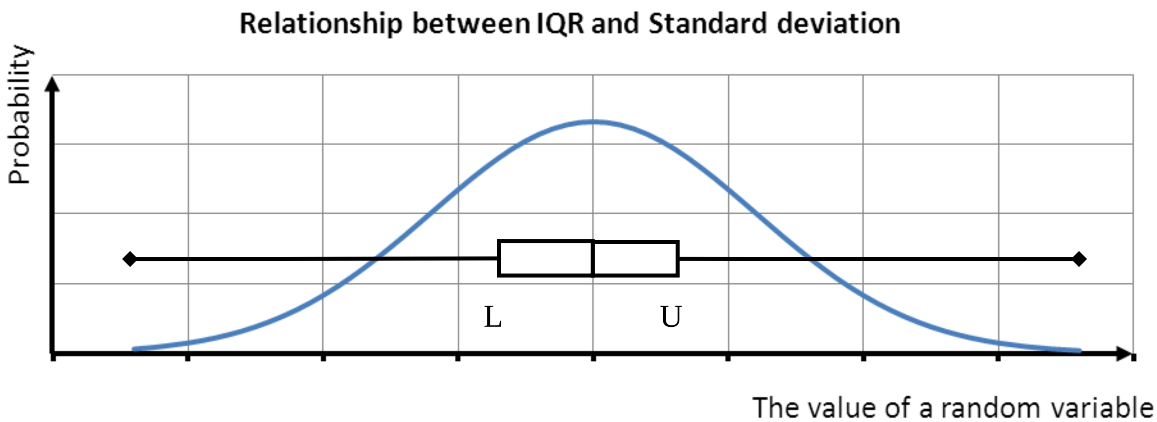


Figure 1. For a normally distributed data there is a simple relationship between the standard deviation (S) and the interquartile range of a sample (IQR): IQR = 1.35 S, where IQR = UQ – LQ.

The inequality 1 and the equation 2 give the 95% confidence interval for a population mean

$$\mu_{sample} - 1.5\,\frac{IQR}{\sqrt{n}} < \mu < \mu_{sample} + 1.5\,\frac{IQR}{\sqrt{n}} \qquad \text{(Inequality 3)}$$

The Inequality (3) is the basis for deriving informal rules of the statistical inference currently adopted in New Zealand high schools.

3. Basis of making a call about two populations

At a high school level we teach rules to make a call about two populations. By applying these rules students can make a call whether the population B tends to be bigger than the population A. However, we do not teach them - what does it mean that the population B tends to be bigger than the population A (or vice versa). For some reason, we have missed giving them the most important piece of information for making a call – the definition. We think that in teaching informal statistical inference, the starting point should be the introduction of the probabilistic definition for making a call, the definition for comparison of two groups. For example, the classical question: ‘do Year 11 boys tend to be taller than Year 11 girls in NZ?' involves in general three different types of comparisons:

1) Comparison of a particular Y11 boy's height with a particular Y11 girl's height. E.g. comparison of Adam's and Emily's heights;
2) Comparison of random samples of Y11 boys and Y11 girls. E.g. comparison of two random samples of 30 boys and 40 girls selected from the data base of www.censusatschool.co.nz;
3) Comparison of two populations. E.g. comparison of the heights of all Year 11 boys and Year 11 girls from the database of www.censusatschool.co.nz.

While the first comparison is trivial and easily comprehendible for students, the last two need precise mathematical / statistical definition.

Definition: If two groups A and B have finite number of elements, the group B tends to be bigger than the group A if the probability of the difference between their elements being positive is more than a half:

$$P(B_m - A_n > 0) > 0.5 \qquad \text{(Inequality 4)}$$

where, B_m and A_n represent the values of m^{th} and n^{th} elements of the corresponding groups. While this definition is obviously inconvenient for practical use, it gives us a probabilistic basis to derive all the commonly used informal rules (milestone tests) for making a call about two populations. Below we use this definition for comparison of two populations.

4. Comparison of two populations

When comparing two populations, we assume that the data in both populations are normally distributed. The fictional examples of the distributions of two populations and their differences are presented on the Figure 2 and 3 below. We use a letter B to represent group of boys and a letter G – group of girls.

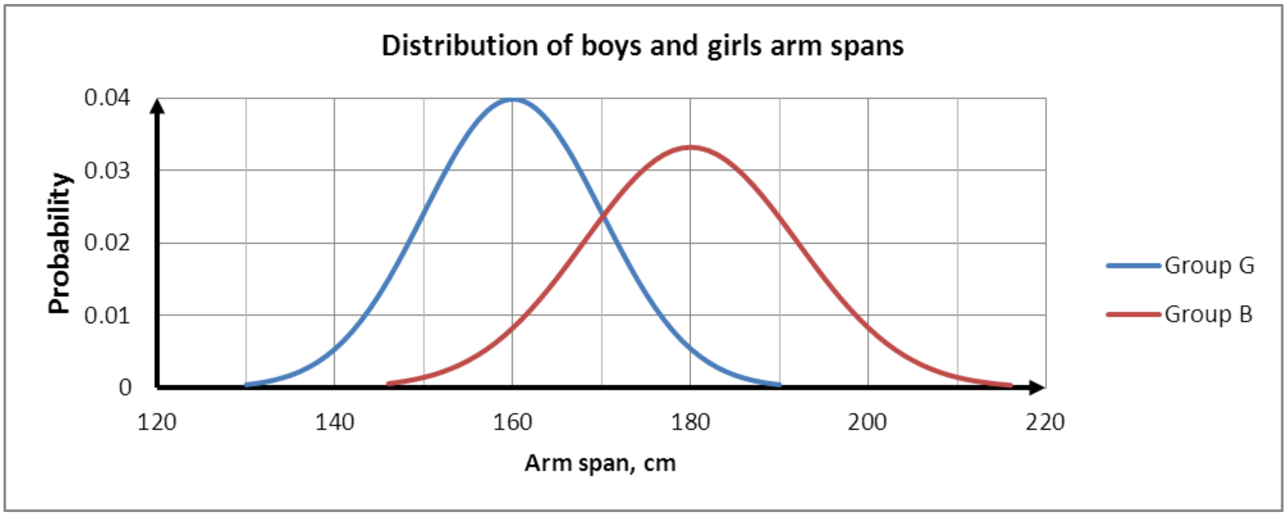


Figure 2. Fictional distribution of two random variables in the two populations. $\mu_B = 180$ cm, $\sigma_B = 12$ cm and $\mu_G = 160$ cm, $\sigma_G = 10$ cm.

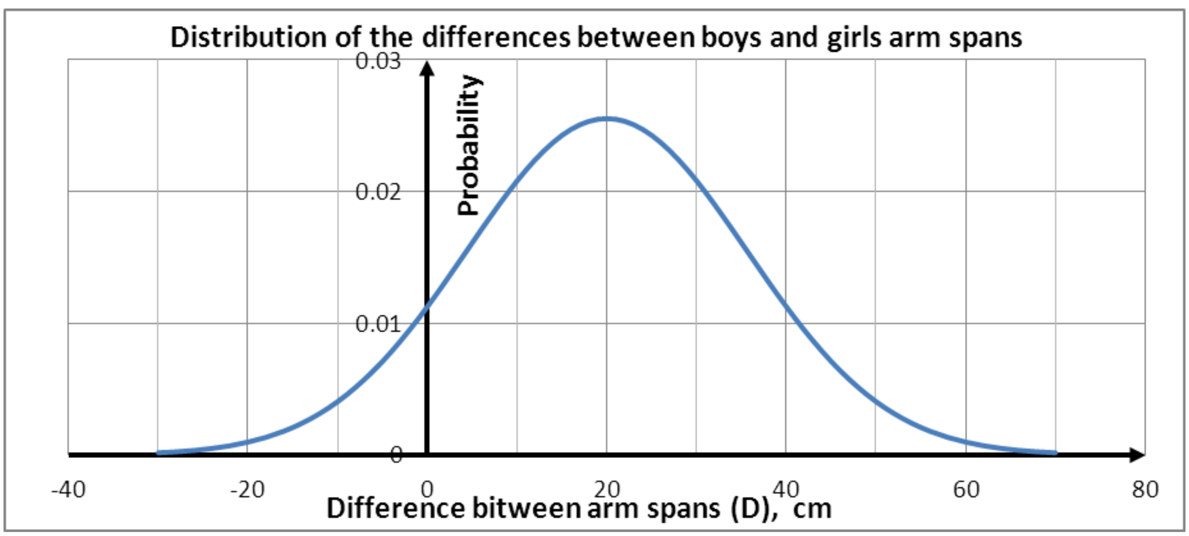


Figure 3. Distribution of the differences between the elements of the two independent populations. $\mu_D = \mu_B - \mu_G = 20$ cm and standard deviation $\sigma_D = \sqrt{(\sigma_B^2 + \sigma_G^2)} = 15.62$cm. As $\mu_D = 20 > 0$, $P(D > 0) > 0.5$, meaning that boys population tends to have longer arm span than the girls population.

In general, if two normally distributed populations A and B have the mean values μ_A and μ_B and the mean value of the new population D (where D = B – A) is positive ($\mu_D = \mu_B - \mu_A > 0$), then the population B tends to have greater values than the population A as:

$$P(B_m - A_n > 0) = P(0 < B_m - A_n < \mu_D) + 0.5 > 0.5 \qquad \text{(see Fig. 3)}$$

Application 1: If μ_A and μ_B are the mean values of the normally distributed populations A and B, the population B tends to have greater values than the population A if the mean value of the population B is greater than the mean value of the population A, regardless of the spread of the populations

$$\mu_B > \mu_A \qquad \text{(Inequality 5)}$$

Application 2: To make a call about two normally distributed independent populations all that is needed is to estimate the mean values of the populations using their samples. Note that this is valid for the normally distributed data. It also may be true for some other types of distributions. However, in general, the difference between the means/medians of two populations alone is not enough to make a call.

5. Comparison of two populations using single samples

Application of the interval estimate for population means (inequality 3) for the population A and B gives us the 95% confidence interval for each population mean

$$\mu_{sample\ B} - 1.5\ \frac{IQRB}{\sqrt{n}} < \mu_{population\ B} < \mu_{sample\ B} + 1.5\ \frac{IQRB}{\sqrt{n}} \qquad \text{(Inequality 6)}$$

$$\mu_{sample\ A} - 1.5\ \frac{IQRA}{\sqrt{n}} < \mu_{population\ A} < \mu_{sample\ A} + 1.5\ \frac{IQRA}{\sqrt{n}} \qquad \text{(Inequality 7)}$$

To make a call that “the population B tends to be bigger than the population A” the lower margin of the 95% confidence interval of the population B must be greater than the upper margin of the population A:

$$\mu_{sample\ B} - 1.5\ \frac{IQRB}{\sqrt{n}} > \mu_{sample\ A} + 1.5\ \frac{IQRA}{\sqrt{n}} \qquad \text{(Inequality 8)}$$

The inequality (8) leads us to the rule for making a call – if the sample sizes for two populations are the same (n), the population B tends to be bigger than the population A, if

$$\mu_{sample\ B} - \mu_{sample\ A} > 1.5\ \frac{IQRB + IQRA}{\sqrt{n}} \qquad \text{(Inequality 9)}$$

In general, if the sample size for the population A is n and for the population B is m, the population B tends to be bigger than the population A, if

$$\mu_{sample\ B} - \mu_{sample\ A} > 1.5\ \frac{IQRB}{\sqrt{m}} + 1.5\ \frac{IQRA}{\sqrt{n}} \qquad \text{(Inequality 10)}$$

here $\mu_{\text{sample A}}$ and $\mu_{\text{sample B}}$ are the mean values of each sample, n and m are sample sizes, IQRA and IQRB are the interquartile ranges of the corresponding samples.

In the next section, based on the inequality 9 we derive and critically evaluate the rules for the informal statistical inference currently adopted in New Zealand high schools.

6. Explaining the rules of the informal inference - "How to make the call"

To derive the rules for the informal statistical inference we consider symmetrical box plots where the median is in the middle between the values of the lower quartile (LQ) and upper quartile (UQ). We also look at the simple relationships between LQ, UQ, IQR (inter quartile range) and median (μ):

$$IQR = UQ - LQ \qquad \text{(Equation 11)}$$

$$\mu = LQ + \frac{IQR}{2} \qquad \text{(Equation 12)}$$

$$\mu = UQ - \frac{IQR}{2} \qquad \text{(Equation 13)}$$

6.1. Rule 1 – boxes do not overlap (at any level of NZ curriculum)

If there is no overlap between the boxes of the box and whisker plots of two samples (Figure 4), we can make a call that there is a difference in the population values regardless of the sample sizes (Wild et al., 2011, p.260). Below we show that this rule is true for the sample sizes ≥ 9 only.

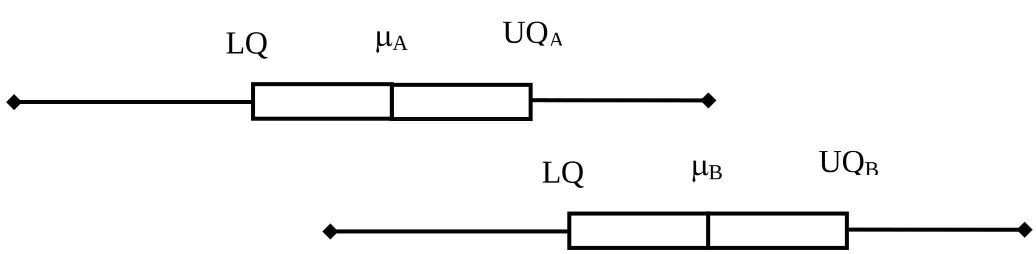


Figure 4. If there is no overlap between the boxes of the box and whisker plots of two samples, we can make a call that there is a difference in the population values if the sample sizes are at least 9.

In this case we can see that

$$LQ_B > UQ_A \qquad \text{(Inequality 14)}$$

which leads us to

$$\mu_B - \mu_A > \frac{IQRB + IQRA}{2} \qquad \text{(Inequality 15)}$$

The inequality (9) and (15) result in the sample sizes of at least 9:

$$\frac{IQRB + IQRA}{2} \geq 1.5\ \frac{IQRB + IQRA}{\sqrt{n}} \quad \rightarrow \quad n \geq 9. \qquad \text{(Inequality 16)}$$

6.2. Rule 2 – Boxes overlap, but neither median lies inside the other box (age 14, level 5 of NZ curriculum)

If the boxes of samples overlap, but neither median lies inside the other box (Figure 5), we can make a call that the population B tends to be bigger than the population A (Wild et al., 2011, p.260).

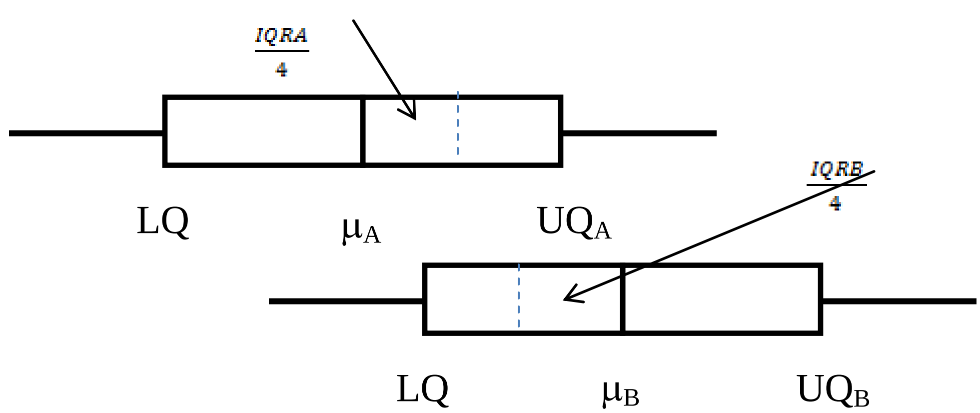


Figure 5. If the boxes of two samples overlap, but neither median lies inside the other box, the sample sizes must be at least 36 to make a call that there is a difference in the population values.

We examine this rule below and find out that this is true for the reasonable sample sizes only (n≥ 36). As we can see from the diagram above:

$$\mu_B - \mu_A > \frac{IQRB + IQRA}{4} \qquad \text{(Inequality 17)}$$

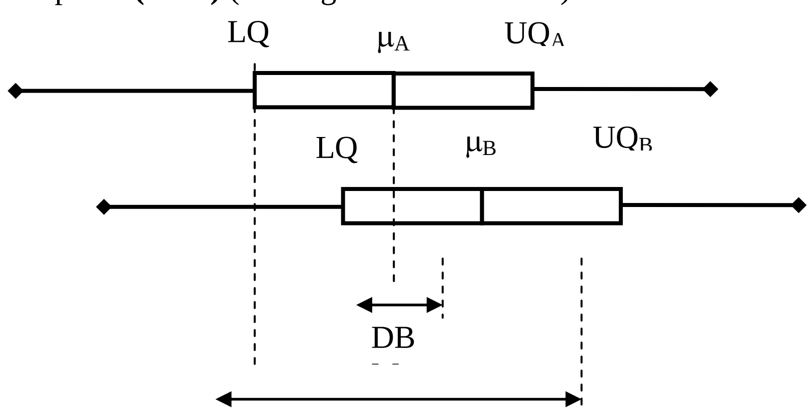
To make a call about two populations, the condition (9) should be met, which gives us:

$$\frac{IQRB + IQRA}{4} \geq 1.5 \frac{IQRB + IQRA}{\sqrt{n}} \quad \rightarrow \quad n \geq 36 \qquad \text{(Inequality 18)}$$

Conclusion: if the boxes of two samples overlap, but neither median lies inside the other box, the sample sizes must be at least 36 to make a call that there is a difference in the population values.
In this case Wild et al. (2011, p.260) recommend to restrict sample sizes between 20 and 40. However, we cannot see any reason why the sample size cannot be more than 40.

6.3. Rule 3 - The boxes overlap and each median lies inside the other box (level 5 and 6 of NZ curriculum)

No claim of any kind can be made at the level 5 (age 14) of the curriculum. At the level 6 (age 15) of the curriculum we look at the difference between medians **(DBM)** and the overall visible spread **(OVS)** (See Figure 6 and Table 1)



DBM = μ_B - OV (Equation 19)
OVS = $UQ_B - LQ_A$ (Equation 20)

Figure 6. Handy diagram for making a call, commonly used in New Zealand high schools at level 6 (age 15), Wild et al. (2011, p.260).

To derive the rule 3 above empirically developed by Wild et al. (2011, p.260), we used equations / inequalities (9), (11 – 13), (19), and (20). We can make the claim "population B tends to be bigger than the population A" (or vice-versa), if:

$$DBM > \frac{30VS}{\sqrt{n}+3} \qquad \text{(Inequality 21)}$$

where n is the sample size in each group.

In the table below we provide the comparison of results obtained empirically by Wild et al. (2011, p.260) and derived theoretically in the present paper for the rule 3.

Table 1. Comparison of results obtained empirically by Wild et al. (2011, p.260) and derived theoretically in the present paper (Inequality 21)

Sample size	**Make a call that B tends to be greater than A (or vice-versa) back in population if:**	
	Wild et al. (2011, p. 260)	**Present article – inequality (21)**
30	DBM > 33% of OVS	DBM > 35% of OVS
50	Not available	DBM > 30% of OVS
100	DBM > 20% of OVS	DBM > 23% of OVS
1000	DBM > 10% of OVS	DBM > 9% of OVS

6.4. Rule 4 - The boxes overlap and each median lies inside the other box (level 7 of NZ curriculum, age 16)

At the level 7 of the curriculum we make a call that B tends to be greater than A (or vice-versa) back in population if the margins of interval estimates of the population means do not overlap (Wild et al., 2011, p.260). In this article this rule is given by the inequalities (9) or (10).

In this section we derived and explained informal rules of the statistical inference implemented in the NZ high school curriculum. However, we should point out the limitations of these informal rules.

7. Limitations of the Informal Rules of the Statistical Inference

Limitation 1: The informal inference rules developed by Wild et al. (2010, 2011) are empirically derived by simulating normal data. Our calculations also are based on the assumption that the data in both populations can be modeled by the normal distribution. Therefore the informal statistical rules currently adopted in New Zealand high schools are limited to normally distributed data.

Limitation 2: Our calculations are based on the assumption that the standard deviations of the sample and the population are equal. As all key statistical parameters of the different samples and their populations are different, we do not have any solid reason to suggest, that their standard deviations are the same. On one hand, the standard deviation of a sample can be less than the standard deviation of a population, as the range of the sample is highly likely to be smaller than the range of the population. On the other hand, the standard deviation of a sample increases if the shape of the distribution of the sample is not symmetrical. The range of the sample and the shape of the distribution concurrently affect the value of the standard deviation.

A generated data set of normally distributed Paua shell weights was trialed at Woodford House for the internally assessed achievement standard AS 2.9 (Use statistical methods to make an inference, level 6 of New Zealand curriculum, age 16). Simple random sampling method was used for selecting samples of sizes from 30 to 50. The samples selected by students indicate that about 50% of the samples have the standard deviation less than the population standard deviation. This means, that making a call about two populations has not 5% but more than 5% chance to be incorrect.

Limitation 3: For symmetrically distributed data the values of the mean and median are equal. However, even for a sample selected from a symmetrical population the shape of the distribution usually is non-symmetrical and the values of the mean and the median are different. As the median value is not affected by extreme values of the data, preference should be given to the median values. Therefore, in the expressions (9) and (10) $\mu_{\text{sample A}}$ and $\mu_{\text{sample B}}$ should be read as the medians of the appropriate samples.

Limitation 4: If all the reasons for limitations 1 to 3 do not exist, i.e. the data in the population and its sample both are normally distributed and the standard deviation of the population and sample are the same, our conclusion about the two populations still has **5% chance to be incorrect** as we use 95% confidence interval for estimation of the population means.

8. Summary

8.1. When a comparison of two populations is based on the comparison of their samples, we can make a call that the population B tends to have bigger values than the population A if:

$$\mu_{\text{sample B}} - \mu_{\text{sample A}} > 1.5\,\frac{IQRB + IQRA}{\sqrt{n}}$$

where $\mu_{\text{sample A}}$ and $\mu_{\text{sample B}}$ are the median values and n is the size of each sample, IQRA and IQRB are the interquartile ranges of the corresponding samples.

8.2. If the sample sizes are different, population B tends to have bigger values than the population A if:

$$\mu_{\text{sample B}} - \mu_{\text{sample A}} > 1.5\,\frac{IQRB}{\sqrt{m}} + 1.5\,\frac{IQRA}{\sqrt{n}}$$

where n and m are the sample sizes of the population A and B respectively.

8.3. In terms of DBM (difference between medians) and OVS (overall visible spread) population B tends to have bigger values than the population A (or vice versa) if:

$$DBM > \frac{3OVS}{\sqrt{n}+3}$$

where n is the size of each sample. This formula can be used for the level 6 and 7 of the NZ curriculum (ages 15 and 16).

8.4. Any call about two populations using single samples has at least 5% chance to be incorrect as we use 95% confidence interval for the population mean estimate.

8.5 There is no significant difference between informal inference rules empirically developed by Wild et al. (2010, 2011) and the informal rules derived in this paper using the formal classical inference techniques. As the informal inference rules developed by Wild et al. (2011, p.260) are empirical rules, they are limited to the particular sample sizes, while the rules provided in this paper can be applied for any sample sizes.

We acknowledge the landmark significance of Wild's and his colleagues work to deliver the ideas of informal statistical inference to New Zealand high schools. We consider the present paper as a supplement to the work done by Wild et al. (2010, 2011) and Pfannkuch et al. (2010).

8.6. We did not use real data to check the informal rules currently adopted at NZ high schools. We do not argue against the importance of using real data in teaching statistics.

However, we believe that without knowledge of basis of informal inference rules and how to derive these rules, we cannot estimate the limits of their applicability. Otherwise we may develop the wrong impression that these rules can be applied to every real life populations. In fact, these informal inference rules are not valid for all types of distributions. As both the informal rules developed by Wild et al. (2010, 2011) and the rules presented in this paper are based on normal distribution, they may be applicable for some populations only. These rules should be considered as guidelines only for real data and introductory part to distribution free statistical inference.

10. Acknowledgements

The authors are grateful to Robert McDonald for editing the original manuscript. We appreciate very much the valuable comments from Philip Doyle and Farida Kachapova. At the same time, these acknowledgements should not be seen as implying that they are in the complete agreement with what we are presenting in this article.

References

Batanero, C. (2000), "Controversies around significance tests", *Mathematical Thinking and Learning, 2*(1-2), 75-98.

Ben-Zvi, D., Gil, E., and Apel, N. (2007), "What is hidden beyond the data? Helping young students to reason and argue about some wider universe", In J. Ainley & D. Pratt (Eds.), *Proceedings of the Fifth International Research Forum on Statistical Reasoning, Thinking and Literacy: Reasoning about Statistical Inference: Innovative Ways of Connecting Chance and Data*. The University of Warwick, (pp. 1-26). 11–17 November 2007.

Castro-Sotos, A. E., Vanhoof, S., Noortgate, W. and Onghena, P. (2007), "Students' misconceptions of statistical inference: A review of the empirical evidence from research on statistics education", *Educational Research Review*, 2, 98–113.

Chance, B., delMas, R. C., and Garfield, J. (2004), "Reasoning about sampling distributions", In D. Ben-Zvi & J. Garfield (Eds.). *The challenge of developing statistical literacy, reasoning and thinking* (pp. 295-323). Amsterdam: Kluwer.

Harradine, A., Batanero,C. and Rossman, A. (2011), "Students and Teachers' Knowledge of Sampling and Inference", pp. 1-15. Published in C. Batanero, C., Burrill,G. & Reading (Eds.), C., *Teaching Statistics in School-Mathematics-Challenges for Teaching and Teacher Education: A Joint ICMI/IASE Study* (pp. 235- 246), DOI 10.1007/978- 94-007-1131-0, Springer Science+Business Media B.V. 2011. Available at http://www.ugr.es/~batanero/ARTICULOS/Chap_Harradine.pdf

Kendall, M. G., and Stuart, A. (1967), "*The Advanced Theory of Statistics*", Vol. 1, 2nd ed. Ch. 14, New York: Hafner Publishing Co.

Konold, C., and Miller, C.D. (2005), "TinkerPlots: Dynamic Data Exploration", (Version 1.0) Emeryville, CA: Key Curriculum Press. Available at http://www.keypress.com/tinkerplots

Liu, Y., and Thompson, P. W. (2009), "Mathematics teachers' understandings of protohypothesis Testing", *Pedagogies*, 4 (2), 126-138. Available at http://www.patthompson.net/PDFversions/2009LiuThompsHypTesting.pdf

McGill, R., Tukey, J. W., Larsen, W. A. (1978), "Variation of Box plots", *The American Statistician*, Vol. 32, No. 1.pp. 12-16. Available at http://amstat.tandfonline.com/doi/pdf/10.1080/00031305.1978.10479236

Ministry of education, (2010), Teaching and Learning Guide for Mathematics and Statistics. Available at http://seniorsecondary.tki.org.nz

Pfannkuch, M. (2006), "Comparing box plot distributions: A teacher's reasoning", *Statistics Education Research Journal*, 5(2), 27-45. Available at http://www.stat.auckland.ac.nz/~iase/serj/SERJ5(2)_Pfannkuch.pdf

Pfannkuch, M., Regan, M., Wild, C. J. and Horton, N. (2010), "Telling data stories: essential dialogues for comparative reasoning", *J. Statist. Educ.*, 18, no. 1. Available at http://www.amstat.org/publi cations/jse/v18n1/pfannkuch.pdf

Pratt, D., Johnston-Wilder, P., Ainley, J., and Mason, J. (2008), "Local and global thinking in statistical inference", *Statistics Education Research Journal*, 7(2), 107-129. Available at http://www.stat.auckland.ac.nz/~iase/serj/SERJ7(2)_Pratt.htm

Saldanha. L., and Thompson, P. (2002), "Conceptions of sample and their relationship to statistical inference", *Educational Studies in Mathematics*, 51, 257-270. Available at http://link.springer.com/article/10.1023%2FA%3A1023692604014?LI=true

Saldanha. L., and Thompson, P. (2007), "Exploring connections between sampling distributions and statistical inference: an analysis of students' engagement and thinking in the context of instruction involving repeated sampling", *International Electronic Journal of Mathematics* Education, 3, 270-297. Available at http://www.iejme.com/032007/d9.pdf

Tukey, J. W. (1970), "*Exploratory Data Analysis*" (Limited Preliminary Edition), Vol. 1, Ch. 5, Reading, Mass: Addison-Wesley Publishing Co.

Vallecillos, A. (1999), "Some empirical evidence on learning difficulties about testing Hypotheses", *Proceedings of the International Statistical Institute* 52nd Session. Helsinki: International Statistical Institute. Available at http://www.stat.auckland.ac.nz/~iase/publications/5/vall0682.pdf

Wild, C. J. (2006), "The concept of distribution", *Statistics Education Research Journal*, 5(2), 10-26. Available at http://www.stat.auckland.ac.nz/~iase/serj/SERJ5(2)_Wild.pdf

Wild, C. J., Pfannkuch, M., Regan, M. and Horton, N. J. (2011), "Towards more accessible conceptions of statistical inference", *J. R. Statist. Soc. A* (2011) 174, Part 2, pp. 247–295. Available at http://onlinelibrary.wiley.com/doi/10.1111/j.1467-985X.2010.00678.x/pdf

Wild, C. J., Pfannkuch, M., Regan, M. and Horton, N. J. (2010), "Inferential reasoning: learning to "make a call" in theory". In *Proc. 8th Int. Conf. Teaching Statistics* (ed. C. Reading). The Hague: International Statistical Institute. Available at http://www.stat.auckland.ac.nz/~iase/publications/icots8/ICOTS8_8B1_WILD.pdf

Authors contact information:

Nugzar Nachkebia
Mathematics teacher
Westmount School, Hawkes Bay Campus,
605 Williams Street
Hastings, New Zealand
nugzar.nachkebia@westmount.school.nz
Tel: +6468770273 (Home)
+64211326002 (Mob)

Marina Alexander
HOD Mathematics
Woodford House
Iona road, Havelock North
Hastings, New Zealand
Marina.alexander@woodford.school.nz
Tel: +64211173004 (Mob)