Pak. J. Statist.
2011 Vol. 27(4), 543-554

THE TAIL FUNCTIONS APPROACH TO CONFIDENCE ESTIMATION IN SURVEY SAMPLING

Borek Puza
School of Finance and Applied Statistics, College of Business and Economics
Australian National University, Canberra, ACT 0200, Australia
Email: borek.puza@anu.edu.au

ABSTRACT

This paper shows how the method of tail functions (Puza and O'Neill, 2005) can be applied in survey sampling so as to produce improved confidence intervals for a finite population quantity of interest. The method works best in the presence of prior information and involves choosing a tail function that leads to the interval which is optimal in some sense, such as minimising prior expected length. The method is illustrated by application to inference on the finite population mean, and tested via a Monte Carlo study.

KEYWORDS

Survey sampling; confidence interval; tail function; optimisation; prior expected length.

1. INTRODUCTION

In survey sampling (Cochran, 1977) one often requires a confidence interval (CI) for the quantity of interest, for example the finite population mean. Using the central limit theorem and other results, it is often possible to construct an approximate CI based on a normal approximation to the distribution of the relevant point estimator, for example the sample mean. If the sample size is sufficiently large, the approximate CI will have a frequentist coverage probability that is close to the desired confidence level (e.g. 95%).

When prior information is available, this may be utilised in several different ways. First, one may adopt the Bayesian framework of Ericson (1988), whereby the finite population is considered as arising from a super-population whose parameters (for example, the mean and variance of the underlying normal distribution) are assigned a joint prior distribution that reflects one's prior beliefs. Secondly, one may formulate a Bayesian model with a prior distribution placed directly on the finite population quantity of interest (for example, the finite population mean). A problem with these two approaches is that when either is used to a construct a credible interval for the quantity, such as the highest posterior density region, that interval will fail to have, even approximately, the desired frequentist coverage probability for all possible values of the quantity. We will return to this point in Section 5.

Thirdly, one may apply the tail functions (TFs) approach of Puza and O'Neill (2005). This approach, which is non-Bayesian, involves generalising the 'ordinary' CI for the quantity of interest using a class of TFs, and searching for the TF in that class which leads to the optimal CI in some sense, such as minimising the prior expected length of the CI. In contrast to Bayesian methods, the TFs approach is guaranteed to yield a CI which has the desired frequentist coverage probability for all possible values of the inferential target quantity.

The theory of TFs is reviewed in Section 2 and applied in Section 3 so as to define an alternative CI for the finite population mean. Section 4 provides results of a Monte Carlo study used to assess the proposed CI, and Section 5 contains a summary and discussion.

2. THE TAIL FUNCTIONS APPROACH TO CONFIDENCE ESTIMATION

The general theory of tail functions may be outlined as follows. Consider a data vector or scalar D whose distribution depends on a scalar parameter θ for which we wish to construct a $1-\alpha$ CI. Suppose that there exists a scalar pivotal quantity $Q = g(D,\theta)$ (a function of D and θ whose distribution does not depend on θ) with distribution function $F_Q(q)$, so that $F_Q(Q) \sim U(0,1)$. Next, let $\tau(\theta)$ be any non-decreasing function over the support of θ with a range in the interval [0,1]. Then

$$1-\alpha = P(\alpha\tau(\theta) \le F_Q(g(D,\theta)) \le 1-\alpha+\alpha\tau(\theta)) = P(L \le \theta \le U) \tag{1}$$

where $L = L(D)$ and $U = U(D)$ are the solutions in θ of $F_Q(g(D,\theta)) = 1-\alpha+\alpha\tau(\theta)$ and $\alpha\tau(\theta) = F_Q(g(D,\theta))$, respectively. Assuming that these solutions exist, they can either be expressed analytically or, when that is not possible, be obtained using an iterative search procedure, such as the Newton Raphson algorithm. Since (1) is true for all possible values of θ, $[L,U]$ is an exact $1-\alpha$ CI for θ. In this context we call $\tau(\theta)$ the *tail function* (TF).

Example 1: Inference on the normal mean

Suppose that $X \sim N(\mu,\sigma^2)$, where σ^2 is known and we desire a $1-\alpha$ CI for μ. To this end, define the pivotal quantity $Z = (X-\mu)/\sigma \sim N(0,1)$, and choose a TF of the form

$$\tau(\mu) = \begin{cases} \delta, & \mu < \eta \\ \gamma, & \mu \ge \eta \end{cases} \tag{2}$$

where $\eta \in \Re$, $\delta \in [0,1]$ and $\gamma \ge \delta$ are tuning constants. (In terms of the above general theory, $D = X$, $\theta = \mu$, $Q = Z = g(D,\theta) = (X-\mu)/\sigma$, and $F_Q(Q) = \Phi(Z) \sim U(0,1)$. Here, $\Phi(\cdot)$ denotes the standard normal distribution function). Equation (1) now becomes

$$1-\alpha = P\left(\alpha\tau(\mu) \le \Phi\left(\frac{X-\mu}{\sigma}\right) \le 1-\alpha+\alpha\tau(\mu)\right) = P(L \le \mu \le U) \tag{3}$$

where $L = L(X)$ and $U = U(X)$ are to be determined. In this case, these functions can be obtained easily, without the need for iterative techniques. First observe from (3) that

$$1-\alpha = P\left(\alpha\tau(\mu) \le \Phi\left(\frac{X-\mu}{\sigma}\right) \le 1-\alpha+\alpha\tau(\mu)\right) = P(A \le X \le B) \tag{4}$$

where:

$$\begin{aligned} A = A(\mu) = \mu + \sigma\Phi^{-1}(\alpha\tau(\mu)) &= \begin{cases} \mu + \sigma\Phi^{-1}(\alpha\delta), & -\infty < \mu < \eta \\ \mu + \sigma\Phi^{-1}(\alpha\gamma), & \eta \le \mu < \infty \end{cases} \\ B = B(\mu) = \mu + \sigma\Phi^{-1}(1-\alpha+\alpha\tau(\mu)) &= \begin{cases} \mu + \sigma\Phi^{-1}(1-\alpha(1-\delta)), & -\infty < \mu < \eta \\ \mu + \sigma\Phi^{-1}(1-\alpha(1-\gamma)), & \eta \le \mu < \infty \end{cases} \end{aligned} \tag{5}$$

We see that the functions $x = A(\mu)$ and $x = B(\mu)$ are non-decreasing and made up of straight lines. Inverting these functions, respectively, leads to the required CI $[L,U]$, where:

$$\begin{aligned} L = L(x) &= \begin{cases} x - \sigma z_{\alpha(1-\delta)}, & -\infty < x < \eta + \sigma z_{\alpha(1-\delta)} \\ \eta, & \eta + \sigma z_{\alpha(1-\delta)} \le x \le \eta + \sigma z_{\alpha(1-\gamma)} \\ x - \sigma z_{\alpha(1-\gamma)}, & \eta + \sigma z_{\alpha(1-\gamma)} < x < \infty \end{cases} \\ U = U(x) &= \begin{cases} x + \sigma z_{\alpha\delta}, & -\infty < x < \eta - \sigma z_{\alpha\delta} \\ \eta, & \eta - \sigma z_{\alpha\delta} \le x \le \eta - \sigma z_{\alpha\gamma} \\ x + \sigma z_{\alpha\gamma}, & \eta - \sigma z_{\alpha\gamma} < x < \infty \end{cases} \end{aligned} \tag{6}$$

and where $z_p = \Phi^{-1}(1-p)$ denotes the upper p-quantile of the standard normal distribution.

When $\gamma = \delta$ we call the tail function in Equation (2) (and the associated CI) 'constant'. The constant TF $\tau(\mu) = 0.5, -\infty < \mu < \infty$, may also be called 'ordinary' because it leads to the 'ordinary' CI for a normal mean, $[x \pm z_{\alpha/2}\sigma]$. Two other constant TFs are $\tau(\mu)$ = 0 and $\tau(\mu)$ = 1, and these lead to the well-known one-sided CIs $(-\infty, x + z_\alpha\sigma]$ and $[x - z_\alpha\sigma, \infty)$, respectively.

When $\gamma = 1-\delta$, we say the tail function in Equation (2) (hereafter called "TF (2)") is symmetric. Otherwise, we say it is asymmetric. (CIs corresponding to symmetric or asymmetric TFs will likewise be called symmetric or asymmetric). Symmetric CIs defined by $\delta = 0$ may also be called Pratt CIs, after Pratt (1961) who derived them using a different argument based on the inversion of hypothesis tests.

Figure 1 shows TF (2) for various values of η, δ and γ, and Figure 2 displays the associated CIs, as given by (6). (Note that by equality of (3) and (4), the lines in Figure 2 could also be drawn using (5)). It will be observed that the ordinary CI has a constant width, whereas each of the other CIs is shortest for values of x in a neighbourhood of η

and widest when x is far from η. Also, the width of a symmetric CI is itself symmetric about η and has no local minima or maxima. These facts will be made use of later when it comes to choosing the best TF for a particular application.

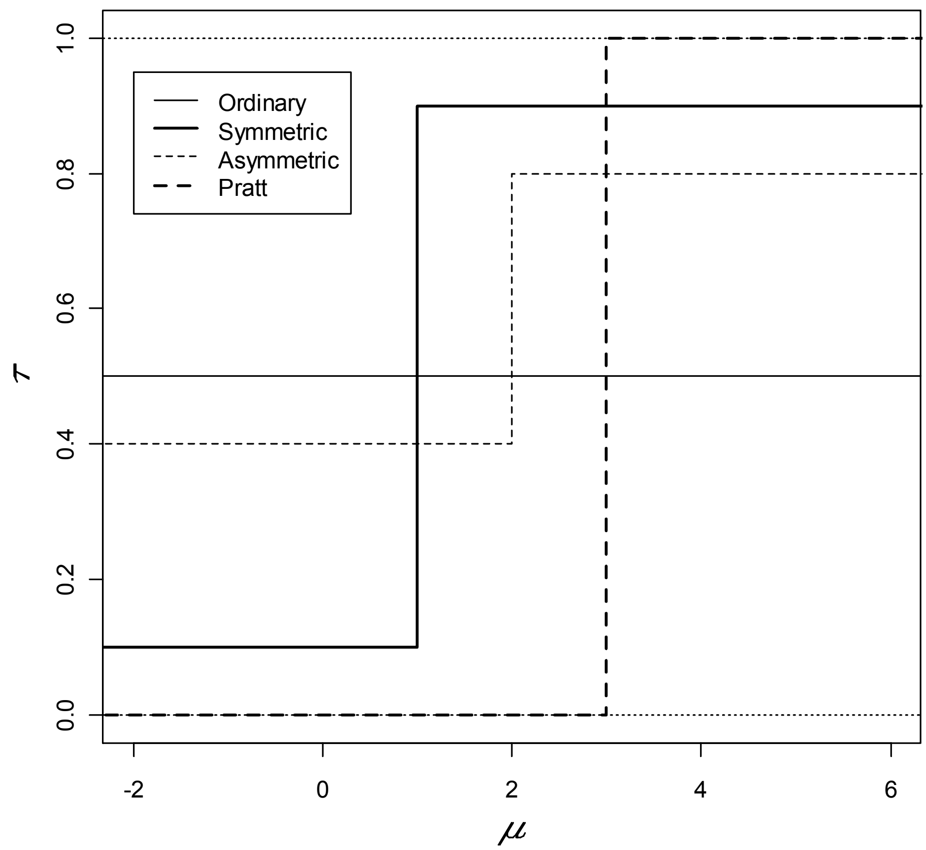

Fig. 1: Several tail functions: Ordinary ($\delta = \gamma = 0.5$); Symmetric ($\gamma = 1 - \delta$) with $\delta = 0.1$, $\eta = 1$; Asymmetric ($\gamma \neq 1 - \delta$), with $\delta = 0.4$, $\gamma = 0.8$, $\eta = 2$ Pratt ($\delta = 0 = 1 - \gamma$) with $\eta = 3$.

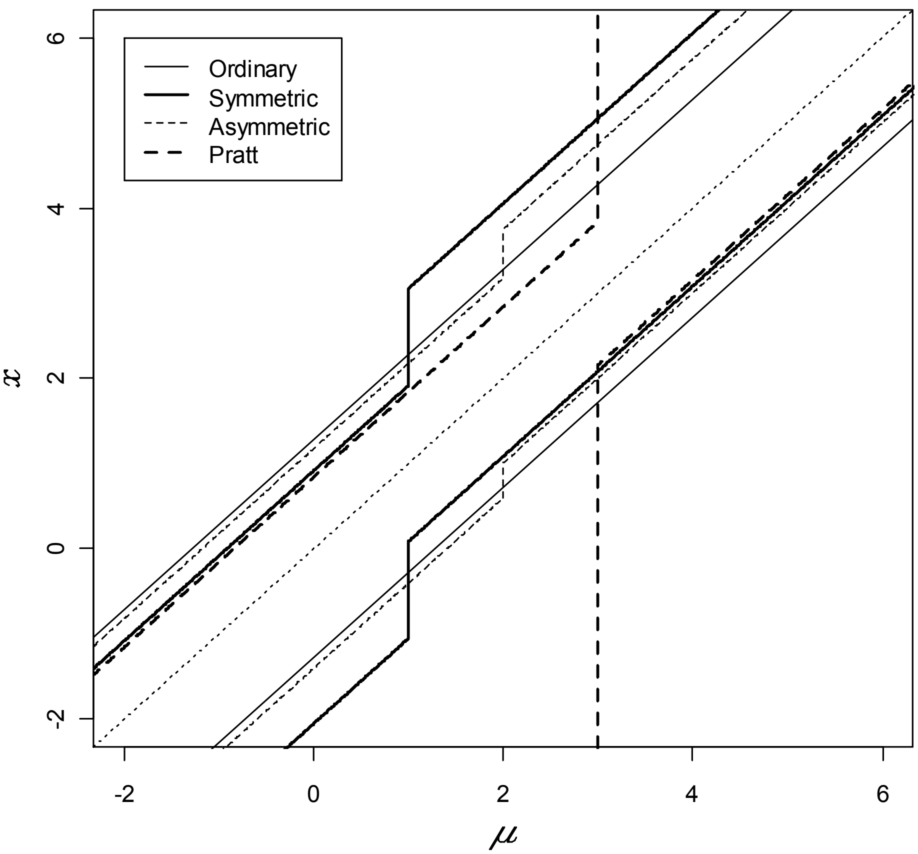

Fig. 2: Several 80% confidence intervals, all with $\sigma = 1$: Ordinary ($\delta = \gamma = 0.5$); Symmetric ($\gamma = 1 - \delta$) with $\delta = 0.1$, $\eta = 1$; Asymmetric ($\gamma \neq 1 - \delta$), with $\delta = 0.4$, $\gamma = 0.8$, $\eta = 2$; Pratt ($\delta = 0 = 1 - \gamma$) with $\eta = 3$.

3. A SIMPLE APPLICATION TO SURVEY SAMPLING

We will now show how the theory of tail functions in the last section may be applied in the context of survey sampling. Recall that under simple random sampling without replacement (SRSWOR) of n units from N, the sample mean, $\bar{y} = (y_1 + ... + y_n)/n$, has expected value equal to the finite population mean, $\bar{Y} = (Y_1 + ... + Y_N)/N$, and variance $h^2 S^2$, where $h^2 = n^{-1} - N^{-1}$ and $S^2 = (N-10)^{-1} \sum_{i=1}^{N} (Y_i - \bar{Y})^2$ (the finite population variance). Assuming that n is 'large', the Central Limit Theorem and other results imply that

$$\frac{\bar{y} - \bar{Y}}{hS} \sim N(0,1) \text{ (approximately)}$$

which then leads to the approximate $1-\alpha$ CI for $\bar{Y}$ given by $[\bar{y} \pm z_{\alpha/2} hS]$. In practice, S will be unknown and replaced in this formula by s, where $s^2 = (n-1)^{-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$ (the sample variance), so that the final reported CI becomes $[\bar{y} \pm z_{\alpha/2} hs]$ (Cochran, 1977, Section 2.8).

We may now view this example of survey sampling inference in the context of Section 2, with $x = \bar{y}$, $\mu = \bar{Y}$, $\sigma = hS$ (or $\sigma = hs$) and TF $\tau(\bar{Y}) = 1/2, -\infty < \bar{Y} < \infty$. The question now arises as to whether it might be possible to construct a better CI based on a different TF.

To address this question, suppose that prior information is available (before sampling) in the form of a distribution $\bar{Y} \sim N(a, b^2)$ (where $b < \infty$) and a point estimate c of S. Because the prior on $\bar{Y}$ is unimodal and symmetric about a, we may, on the basis of patterns in Figure 2, hypothesize that the best choice of η in TF (2) is a, and the optimal tail function is symmetric (with $\gamma = 1-\delta$). There remains the question of how δ should be specified.

One strategy is to temporarily fix $s = S = c$ and search for the value of δ that minimises the *prior expected length* (PEL) of the CI, meaning the function of δ given by

$$PEL_{\delta} = EW(\bar{y}) = \int_{-\infty}^{\infty} W(\bar{y}) f(\bar{y}) d\bar{y}$$

where $W(\bar{y}) = U(\bar{y}) - L(\bar{y})$ is the width (or length) of the CI (which depends on δ), and $f(\bar{y})$ is the prior predictive density of the sample mean (all defined on the basis that S is known and equal to c). Once the optimal value of δ has been found, this value can be used to calculate CI (6), with $x = \bar{y}$, $\mu = \bar{Y}$, $\sigma = hs$ (where s is the *sample* variance) and TF (2) with $\eta = a$ and $\gamma = 1-\delta$. This inferential strategy will now be illustrated by way of an example.

Example 2: Inference based on a pilot survey

Consider a pilot survey, whereby a SRSWOR of size m = 50 has been taken from a finite population of size M = 1050. Suppose that $\bar{y}'$ = 50.249 and s' = 4.7388 are the sample mean and sample standard deviation of the pilot sample values, $y_1', ..., y_m'$. Next, denote the remaining $N = M - m$ = 1000 population values by $Y_1, ..., Y_N$, and let $\bar{Y}$ and S be their sample mean and sample standard deviation. Finally, consider the survey proper, whereby a future sample of size n = 200 will be taken from the N units, again via SRSWOR, and let $\bar{y}$ and s denote the sample mean and sample standard deviation of the sample values, $y_1, ..., y_n$.

In this setting, it is reasonable to write $S \cong c = s'$ and $\bar{Y} \sim N(a, b^2)$ (approximately), where $a = \bar{y}'$ = 50.249 and $b^2 = (m^{-1} - M^{-1})s'^2$ = 0.42773. We now choose tail function (2) with $\gamma = 1 - \delta$ and consider the problem of specifying a suitable value of δ.

For any particular value of δ (e.g. 0.3) the *conditional expected length* (CEL) of the associated 80% CI for $\bar{Y}$ may be defined as

$$K_\delta(\bar{Y}) = E\{W(\bar{y}) \mid \bar{Y}\} = \int_{-\infty}^{\infty} W(\bar{y}) f(\bar{y} \mid \bar{Y}) d\bar{y}$$

where $W(\bar{y}) = U(\bar{y}) - L(\bar{y})$ (which depends on δ), $f(\bar{y} \mid \bar{Y}) = \phi((\bar{y} - \bar{Y})/\sigma)/\sigma$, $\sigma = hs'$ = 0.29971, $\eta = a$ = 50.249 and α = 0.2. (Here, $h^2 = n^{-1} - N^{-1}$ = 0.004 and $\phi(\cdot)$ denotes the standard normal density function). The CEL is shown in Figure 3 for each δ = 0, 0.1, 0.2, 0.3, 0.4 and 0.5, respectively. Note that when $\delta = 0.5$ the CEL is a constant equal to the length of the ordinary 80% CI, namely $2z_{\alpha/2}\sigma$ = 0.7682.

For each δ it is now possible to calculate the corresponding prior expected length via

$$PEL_\delta = EW(\bar{y}) = EE\{W(\bar{y}) \mid \bar{Y}\} = EK_\delta(\bar{Y}) = \int_{-\infty}^{\infty} K_\delta(\bar{Y}) f(\bar{Y}) d\bar{Y},$$

where $f(\bar{Y}) = \phi((\bar{Y} - a)/b)/b$. Figure 4 shows the PEL, with points and numerical values at δ = 0, 0.1, 0.2, 0.3, 0.4, 0.5. To two decimal places, the optimal value of δ is 0.20, corresponding to a PEL of 0.7251. This minimum value of the PEL is 5.6% smaller than 0.7682, the PEL of the ordinary CI defined by δ = 0.5.

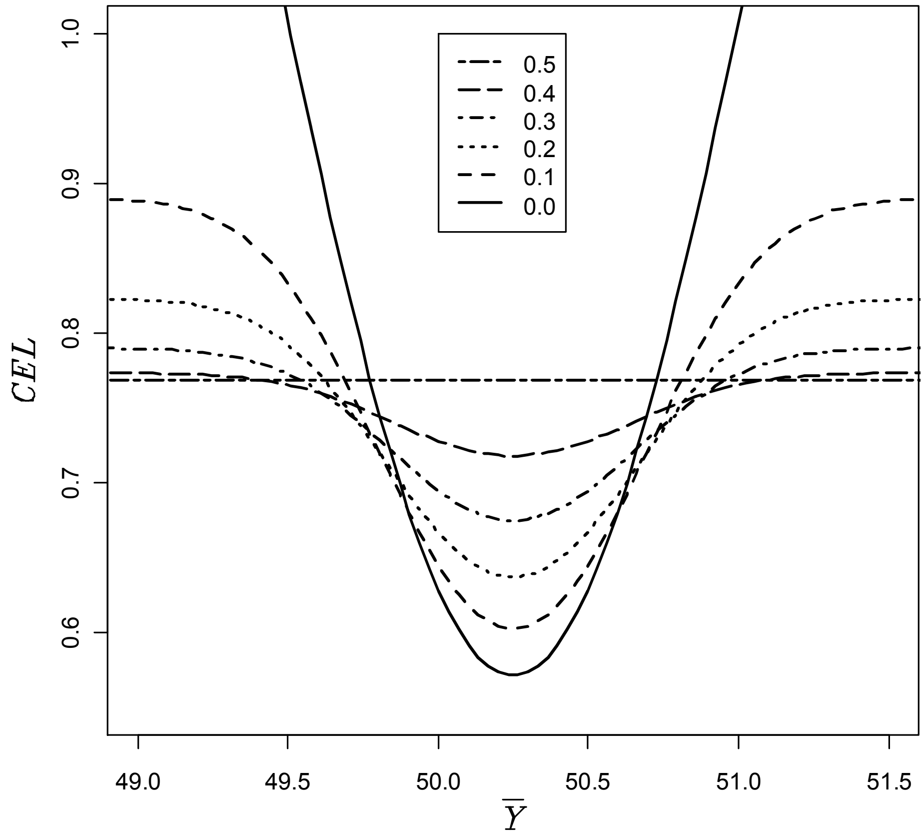

Fig. 3: The conditional expected length of a confidence interval for $\delta = 0, 0.1, \ldots, 0.5$.

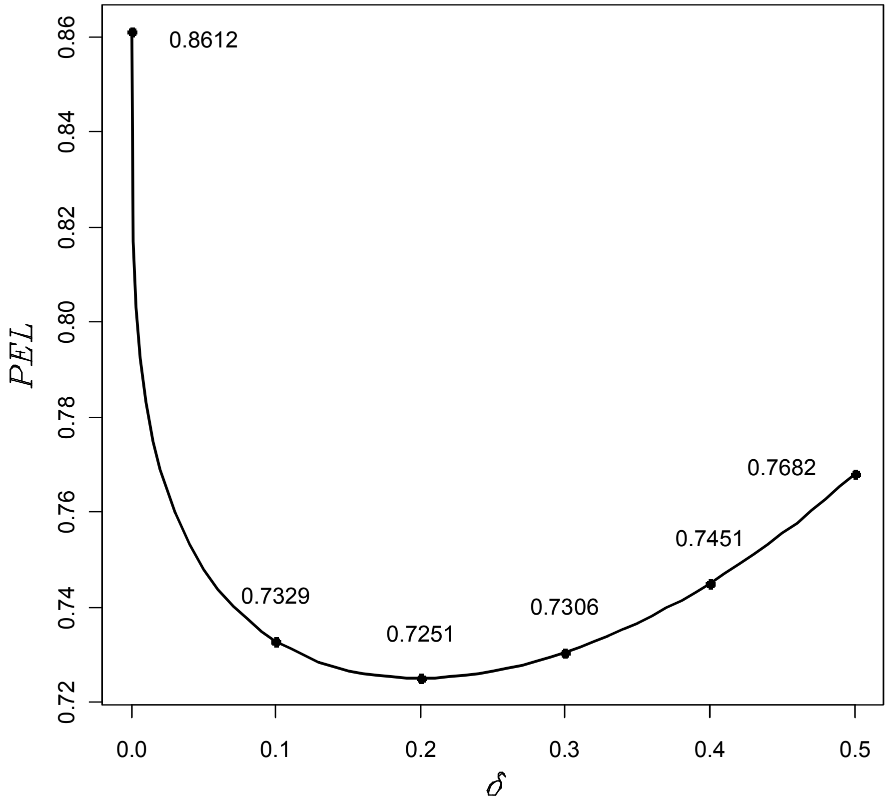

Fig. 4: The prior expected length of a confidence interval, with values indicated at 0, 0.1, 0.2, 0.3, 0.4, 0.5. To two decimals, the optimal value of δ is 0.20.

4. A MONTE CARLO STUDY

In Example 2, it was estimated that using tail function (2) with $\eta = \bar{y}'$ = 50.249, $\delta = 0.2$ and $\gamma = 1-\delta$ = 0.8, would decrease the PEL of the ordinary 80% CI (defined by $\delta = 0.5$) by 5.6%. However, this figure was arrived at using a number of approximations and assumptions. (These will be discussed further in Section 5). The following reports a Monte Carlo study that was conducted to address this concern about the accuracy of the 5.6% figure.

First, let it be known that the finite population of size 1050 had originally been generated as a random sample from the normal distribution with mean 50 and standard deviation 5, and that the 1000 population values which remained after removal of the pilot sample had a mean of $\bar{Y}$ = 49.623 (the inferential target) and a sample standard deviation of S = 5.0049. (Note that many other distributions could have been used to generate the finite population, without substantially altering the results of this Monte Carlo study, since the normality of the sample mean follows generally by the Central Limit Theorem when n is sufficiently 'large').

The first step in the Monte Carlo experiment was to take a sample of size n = 200 from the N = 1000 finite population values via SRSWOR and calculate the sample mean, $\bar{y}$ = 49.936, and the sample standard deviation, s = 5.3429. Then two 80% CIs were constructed: the ordinary CI, $[\bar{y} \pm z_{\alpha/2} hs]$ = [49.503, 50.369], and CI (6), $[L(x), U(x)]$ = [49.600, 50.272]. These intervals were obtained using the specifications $\alpha = 0.2$, $x = \bar{y}$, $\eta = \bar{y}'$ = 50.249, $\delta = 0.2$, $\gamma = 1 - \delta$ = 0.8, $\sigma = hs$ and $h^2 = n^{-1} - N^{-1} = 0.004$.

In this case, the inferential target $\bar{Y}$ = 49.623 lay in both the ordinary and alternative CIs. Also, the lengths of these CIs were 0.8661 and 0.6721, respectively; thus the alternative CI was 0.1940 units, or 22.4%, shorter than the ordinary CI.

This first step of the experiment was repeated independently another 9999 times (after replacing the 200 sampled values each time), leading to a total of 10000 ordinary CIs and 10000 corresponding alternative CIs, each pair of intervals being based on a different sample of size 200 from the same finite population of 1000 values. In each case the calculations were identical to those above except for different values of $\bar{y}$ and s being used.

It was found that $\bar{Y}$ was contained in 79.61% of the 10000 ordinary 80% CIs and in 79.84% of the 10000 alternative 80% CIs. These results are consistent with the coverage of both CIs being approximately 80%. (A 95% CI for the true coverage probability of the ordinary 80% CI is $[0.7984 \pm 1.96\sqrt{0.7984(1-0.7984)/10000}]$ = [0.7882, 0.8040]. Likewise, a 95% CI for the true coverage probability of the alternative 80% CI is [0.7905, 0.8063]).

Also, the average length of the 10000 ordinary CIs was 0.8107, and the average length of the 10000 alternative CIs was 0.8013. Thus the alternative CI appears on average to be about 0.009424 units, or 1.162%, shorter than the ordinary CI. (The sample mean and sample standard deviation of the 10000 differences between the lengths of the two CIs are 0.00924 and 0.09609. Therefore a 95% CI for the mean difference between the two lengths is $[0.00924 \pm 1.96 \times 0.09609/\sqrt{10000}]$ = [0.007541, 0.011308], which is entirely above zero. Likewise, the mean percentage reduction in length, of which the 22.4% figure mentioned above is an example, may be estimated as 1.09%, with 95% CI [0.8685, 1.3305]).

From these findings it appears that the alternative 80% CI with δ = 0.2 is on average definitely shorter than the standard 80% CI (defined by δ = 0.5). However, the improvement appears to be only about 1%, rather than about 5% as indicated by the analysis in Example 2.

One reason for this is that said analysis uses the approximation $s = S = s'$ = 4.7388 (the sample standard deviation of the 50 pilot sample values), whereas s (the sample standard deviation of the 200 sample values yet to be observed) is unknown and variable over repeated sampling. The Monte Carlo study takes this variability of s properly into account and provides an unbiased estimate of the improvement to be had by taking δ as 0.2 relative to 0.5. It should be pointed out that, regardless of the error in the 5.6% figure, both δ = 0.2 and δ = 0.5 lead to CIs with a frequentist coverage probability very close to the nominal 80%.

5. SUMMARY AND DISCUSSION

This paper has described the method of tail functions of Puza and O'Neill (2005) and explored some of its potential applications in the field of survey sampling. An example was used to show how the method can improve confidence estimation of the finite population mean when prior information is available, such as from a pilot survey, in situations where it is desireable to construct a confidence interval which is as short as possible on average.

It may be argued that Bayesian methods provide a more natural vehicle for incorporating prior information, and would, moreover, provide greater reductions in prior expected interval length. However, it is well known that a Bayesian credible interval will fail to have the desired frequentist coverage probability if the target parameter happens to take on a value that is highly inconsistent with its prior distribution. An example of this phenomenon may be found in Puza and O'Neill (2005, Section 5; 2006, Example 2). The tail functions approach provides an alternative to the Bayesian approach, one that provides more modest reductions in interval length on average but preserves and guarantees frequentist coverage probabilities.

As a caveat, it should be pointed out that when the method of tail functions is used to reduce the prior expected length of a confidence interval, this typically comes at a cost, this being the risk that the final interval will in fact be *wider* than the ordinary interval. This risk can clearly be seen in Figure 2, and also in Figure 3 where the conditional expected length of the alternative interval defined by $\delta = 0.2$ clearly exceeds the constant length of the ordinary confidence interval defined by $\delta = 0.5$ if $\overline{Y}$ happens to be 49 (for instance).

The smallness of the improvement reported in Section 4 (a reduction in prior expected length of about 1%) may be seen as a reason not to bother with the tail functions approach. However, greater improvements can certainly be achieved using other tail functions not considered here, and in situations involving other inferential targets and sampling/estimation schemes. In Examples 1 and 2, tail function (2) was chosen mainly for illustrative purposes, because it leads to a simple CI (6) which is expressible in closed form and is readily verified with no need for iterative search techniques such as the Newton Raphson algorithm. A TF which requires such a search technique - and would most likely provide improvements over TF (2) above - is given in Puza and O'Neill (2005, Equation 11 and Figure 1).

The analysis in Example 2 (leading to an estimated maximum reduction in PEL of 5.6% at $\delta = 0.2$) could be refined if there existed a joint prior density for $\overline{Y}$ and *S*, $f(\overline{Y}, S)$ (rather than just $f(\overline{Y})$ and a point prior estimate of *S*), and if one could specify a joint density for $\overline{y}$ and *s* given $\overline{Y}$ and *S*, $f(\overline{y}, s \mid \overline{Y}, S)$ (rather than just $f(\overline{y} \mid \overline{Y})$ with *S* fixed at s'). In that case, it would make sense to redefine the conditional expected length of the alternative CI as

$$K_\delta(\overline{Y}, S) = E\{W(\overline{y}, s) \mid \overline{Y}, S\} = \int_{\overline{y}-\infty}^{\infty} \int_{s=0}^{\infty} W(\overline{y}, s) f(\overline{y}, s \mid \overline{Y}, S) d\overline{y} ds$$

where $W(\bar{y},s)$ is the width of the CI as before, but now also as a function of s (rather than with s fixed at s'). The CEL could then be calculated for a range of values of $\bar{Y}$ and S and depicted in a series of graphs like Figure 3, one for each selected value of S, or alternatively as a surface in a single three dimensional figure. The prior expected length of the alternative CI could then be calculated - possibly with the aid of Monte Carlo methods, as in Puza and O'Neill (2005, Section 5) - for each specified value of δ, according to

$$PEL_\delta = EW(\bar{y},s) = EE\{W(\bar{y},s)\,|\,\bar{Y},S\} = EK_\delta(\bar{Y},S) = \int_{\bar{Y}=-\infty}^{\infty}\int_{S=0}^{\infty} K_\delta(\bar{Y},S)f(\bar{Y},S)d\bar{Y}dS$$

The final result would be a graph similar to Figure 4, leading to a different optimal value of δ and a different minimum PEL. We leave this more sophisticated analysis - and others involving more than one tuning constant (not just δ) - as an avenue for future research.

In this paper we have focussed on confidence estimation of the finite population mean $\bar{Y}$ based on the sample mean $\bar{y}$ under simple random sampling without replacement. However, tail functions can also be applied in contexts involving other inferential targets (e.g. finite population totals, ratios and proportions) and other estimators (e.g. systematic sample means, stratified estimators, the Horvitz-Thompson estimator). In many situations an estimator has a normal distribution centred at the inferential target (approximately), and the theory of Section 3 can be applied with only slight modifications. We leave this as a topic for future research. Other avenues for further research include the application of TFs in situations involving priors on superpopulation *parameters* rather than on finite population quantities as here (see Ericson, 1988), and in situations involving constraints on the inferential targets (see Puza and O'Neill, 2008, and Mandelkern, 2002).

REFERENCES

1. Pratt, J.W. (1961). Length of confidence intervals. *J. Amer. Statist. Assoc.*, 56, 549-567.
2. Cochran, W.G. (1977). *Sampling Techniques, 3rd Edition*. New York: John Wiley.
3. Ericson, W.A. (1988). Bayesian inference in finite populations. In *Handbook of Statistics, Volume 6* (editors P.R. Krishnaiah and C.R. Rao), 213-247. Elsevier Science Publishers B. V.
4. Mandelkern, M. (2002). Setting confidence intervals for bounded parameters. *Statistical Science*, 17, 149-172.
5. Puza, B. and O'Neill, T. (2005). *Interval Estimation via Tail Functions*. Working Paper STAT0005WP, College of Business and Economics, Australian National University, Canberra. Website: http://cbe.anu.edu.au/research/papers/pdf/STAT0005WP.pdf
6. Puza, B. and O'Neill, T. (2006). Interval estimation *via* tail functions. *The Can. J. Statist.*, 34(2), 299-310.
7. Puza, B. and O'Neill, T. (2008). Optimal constrained confidence estimation via tail functions. *The Mathematical Scientist*, 33, 134-140.