Approximating the marginal effect of discrete regressors in logit models

Vince Daly

pproximating the marginal effect of discrete regressors in logit models

Abstract

Logit models are non-linear in their explanatory variables. Derivatives with respect to the explanatory variables therefore only approximate the response to discrete changes in regressor values, yet have gained some support within the literature. This note investigates such an approximation strategy and delineates the conditions under which it does or does not give satisfactory results. The conclusion is that it is a risky strategy and that therefore the correct alternative should be employed.

Key words: Logit Models, Approximation Error

JEL classification number: C25

Introduction

Where models are non-linear in their explanatory variables, differentiation with respect to an explanatory variable will only approximate the modelled response to a discrete change in that variable's value. Nevertheless, use of this approximation has been noted within contemporary reputable textbook literature as being capable of accuracy for the particular case of binary choice models - Greene (2000: p817). Intuitively, this is an unexpected property for such models to possess and thus a puzzle that does not seem to have been yet addressed[1]. The puzzle is of more than academic interest since the rate of response of dependent variable to explanatory variable may, in applied work, be used to assess practical questions where accuracy is a serious concern. Examples could include determining whether or not a binary policy instrument is sufficiently influential or deciding the size – and thus the expense, of a sufficiently large change in a continuous instrument. This present note investigates the conditions required for such an approximation strategy to deliver acceptable accuracy, concluding that these are too restrictive to make the strategy advisable. The note proceeds by first briefly introducing logit models, then deriving an expression for correctly calculating the response to a discrete change in a regressor. The proportionate error that follows from using a derivative-based approximation to this correct calculation is then derived and its behaviour is analysed. A concluding section summarises the implications of the analysis.

Logit Models of Success Probability

Binary response models restrict the dependent variable, Y, to taking one of two values, which may be coded $y_i = 0$ and $y_i = 1$, at each observation point, $i = 1,2,\cdots N$. Simply as a matter of notation, the outcome $y_i = 1$ can be denoted "success". The logit model is a particular example of binary response modelling, so named because it models the "logit", i.e. the logarithm of the odds ratio, as a linear regression. With x_i' being a $1\times k$ row of observed regressor values at the i^th^

[1] On the basis of a search of all abstracts held in the EconLit database, using "logit" and "approximat*" as keywords.

observation point, and β a $k \times 1$ column of parameters, a logit model may be presented as follows.

$$Ln\left(\frac{p_i}{1-p_i}\right) = x_i'\beta$$
$$p_i = \Pr[y_i = 1 \mid x_i'] = E[y_i \mid x_i'] \qquad (1.$$

Inverting the logit transformation and introducing a scalar disturbance, u_i, gives the implied model for y_i:

$$y_i = p_i + u_i = \frac{e^{x_i'\beta}}{1+e^{x_i'\beta}} + u_i$$

This non-linear model is typically estimated by numerical Maximum Likelihood, to give

$$p = F\left(x'\hat{\beta}\right) = \frac{e^{x'\hat{\beta}}}{1+e^{x'\hat{\beta}}} \qquad (2.$$

If the row x' is specified to be one of the observed x_i' then (2) provides a fitted value, $\hat{y}_i = \hat{p}_i$. More generally, the estimated model may be used to simulate $E[y \mid x]$ for any instance of x', observed or hypothetical. Because the function $F(z) = e^z / (1+e^z)$ is the cumulative density function for the logistic distribution the fitted or simulated values will fall within the range (0,1), making them acceptable as estimates of the "conditional success probability", $\Pr[y = 1 \mid x']$.

A natural focus in the application of the estimated model is to consider its implications for the influence of the regressors upon success probability. For continuous regressors the point rate of response can be obtained by differentiation.[2]

[2] Appendix 1 provides support for this and other assertions made within the main text.

$$\frac{\partial p}{\partial x_j} = p(1-p)\hat{\beta}_j \tag{3.}$$

On occasion it may be more useful to evaluate how success probability responds to a discrete change in a regressor value, i.e. $\Delta p / \Delta x_j$ - for example when modelling the consequences of changes in the value of a policy instrument. Moreover, some regressors – for example 0 / 1 dummies, are intrinsically discrete, making instantaneous rates of response theoretically irrelevant. The non-linear nature of the model prompts an expectation that the marginal effects of discrete changes may not be well-approximated by the point rates of response.

However, Greene (2000: p817) suggests that evaluating the derivative $\partial p / \partial x_j$ in fact provides an approximation to the discrete change $\Delta p / \Delta x_j$ that is "often surprisingly accurate" for binary choice models and substantiates this assertion with an example. The purpose of this present note is to examine more closely the merits of such an approximation strategy. Greene makes the case for binary response models generally, offering a fitted probit[3] model as a supporting example. This note investigates the approximation strategy in the context of logit models, because of their relative analytical tractability, maing some comments upon Probit models in the consclucing section. Greene focuses on a particular instance of x', namely the case where regressors are equal to their sample averages. The analysis presented here is not restricted to that case but is applicable to it.

Calculation of Marginal Effects

This section provides an expression for the response of an estimated success probability to a discrete change in the value of a single regressor, which may be specified without loss of generality as the k^{th}. To simplify notation, let the fixed contribution of the remaining regressors to the fitted / simulated logit be

[3] Where Logit models use the cumulative density function (cdf) of the Logistic distribution, Probit models use the cdf of the standard Normal distribution.

$A = \sum_{j=1}^{j=k-1} \hat{\beta}_j x_j$. Let the k^{th} regressor have estimated coefficient b and consider a change in the value of this regressor from $x_k = x$ to $x_k = x + \Delta x$.

The consequence of the shift in value of the k^{th} regressor is a shift in the estimated success probability. Let the estimated success probabilities before and after the change of regressor value be, respectively

$$ {}_1p = p = \frac{e^{A+bx}}{1+e^{A+bx}} \quad , \quad {}_2p = p + \Delta p = \frac{e^{A+bx+b\Delta x}}{1+e^{A+bx+b\Delta x}} $$

Using the continuous derivative to approximate Δp gives

$$ {}^a\Delta p = \frac{\partial p}{\partial x_j} \times \Delta x = p\left(1-p\right)b\ \Delta x \approx \Delta p $$

Whereas the actual discrete change in the estimated probability of success is obtained by the subtraction $\Delta p = {}_2p - {}_1p$ and, following some algebra, may be expressed as

$$ \Delta p = \left(1-p\right)p\,\Phi \quad , \quad \Phi = \frac{\left(e^{b\Delta x}-1\right)}{1+p\left(e^{b\Delta x}-1\right)} \qquad (4. $$

Correct calculation of the effect of a discrete change in a regressor value, whether by computation of Φ, or simply by the subtraction ${}_2p - {}_1p$, requires only small effort. This suggests that the effort should only be avoided if the returns are even smaller – for example if the approximation error is negligible. This possibility is now investigated.

Assessing Approximation Error

The cases $p = 0$, $p = 1$, $b = 0$ and $\Delta x = 0$ may be excluded from consideration since here ${}^a\Delta p = \Delta p = 0$. Before proceeding to a detailed analysis, it is worth noting that, for given values of p and b, $\Phi(\Delta x) \neq -\Phi(-\Delta x)$. Thus the symmetry implicit in the approximation $\Delta p/\Delta x \approx (\partial p/\partial x) = p(1-p)b$ is always qualitatively incorrect.

A more complete assessment of the approximation strategy can be based upon the proportionate approximation error, which is $\alpha = \dfrac{{}^{a}\Delta p - \Delta p}{\Delta p} = \dfrac{b\Delta x - \Phi}{\Phi}$ and can be algebraically re-arranged[4] as

$$\alpha = \delta + pb\Delta x \quad , \quad \delta = \frac{b\Delta x - \left(e^{b\Delta x} - 1\right)}{\left(e^{b\Delta x} - 1\right)}$$

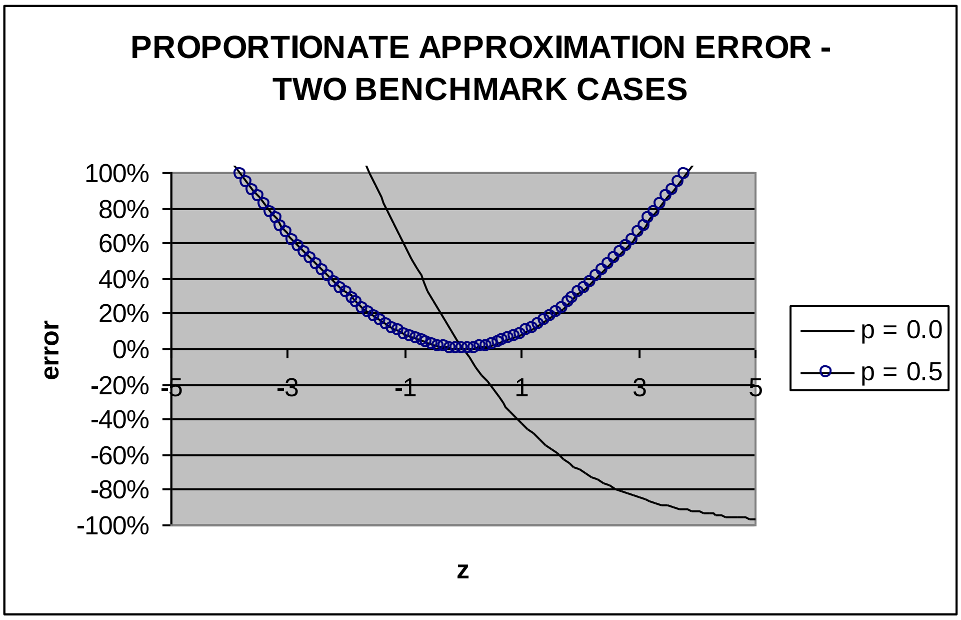
The first term of the RHS, δ, is recognisably the proportionate error that would result from using $b\Delta x$ as an approximation for $\left(e^{b\Delta x} - 1\right)$ and α differs from this by an amount, $pb\Delta x$, whose size and sign are in general unrestricted.

Now, a first-order McLaurin expansion has $e^{z} - 1 \approx z$ so that, when $|b\Delta x|$ is sufficiently small as to make this first-order approximation satisfactory, then $\Phi \approx b\Delta x/(1 + pb\Delta x)$ and $\alpha \approx pb\Delta x$. This confirms that the approximation noted by Greene (2000) will indeed have small error whenever $b\Delta x$ is sufficiently close to zero. However, there are no generally applicable *a priori* limits to the absolute size of $b\Delta x$ and the analysis should therefore proceed without the assumption that $e^{b\Delta x} - 1 \approx b\Delta x$.

The object of analysis is α - the proportionate error that results from approximating $\Delta p/\Delta x_j$ by $\partial p/\partial x_j$. Consider α as a function of $z \equiv b\Delta x$, and indexed by the value of p - now treated as a parameter. Setting $p = 0$ provides a benchmark case in which $\alpha_0(z) = \delta(z)$ is identical to the proportionate approximation error in

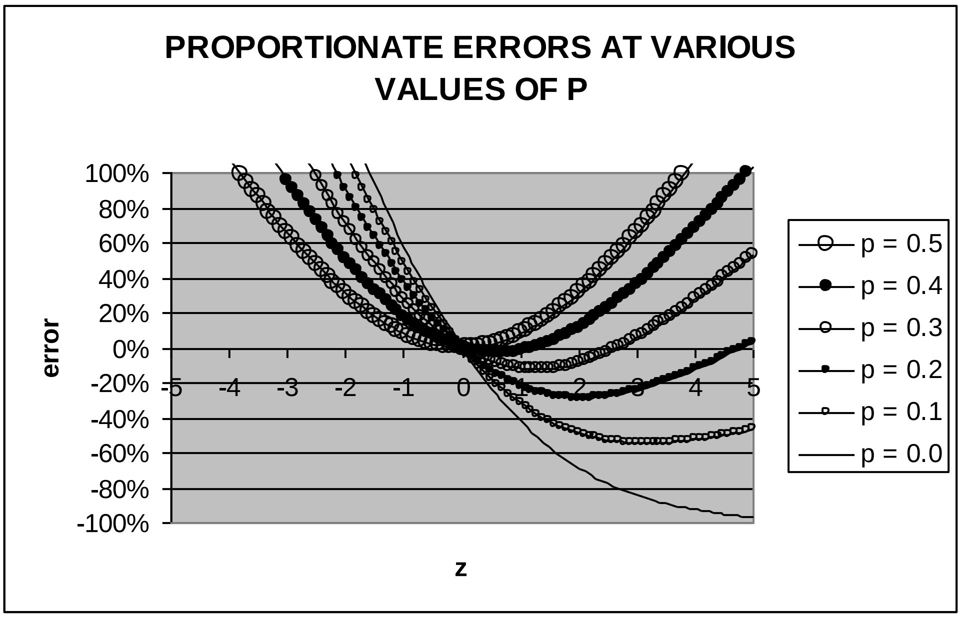
[4] The detailed derivation is in Appendix. 1.

$e^z - 1 \approx z$, which approaches -1 asymptotically as $z \to \infty$ but increases without bound as $z \to -\infty$. Since $p = 0$ is not achievable with finite regressor values, this benchmark should be read as a limiting case – approached as $p \to 0$. Setting $p = 0.5$ provides a second useful benchmark: $\alpha_{0.5}(z) = \delta(z) + 0.5z$, which is symmetric around zero, and non-negative. The following graph shows both benchmark cases.



The two benchmark cases indicate that p, the value of success probability before the change of regressor value, has a fundamental influence upon the approximation error. For $p = 0.5$, the approximation over-states the size of the response to a discrete change in a regressor value, whether this change is one that increases success probability - $z\,(\equiv b\Delta x) > 0$, or decreases success probability - $z\,(\equiv b\Delta x) < 0$. The size of this approximation error increases without bound as $|z| \to \infty$. The second benchmark, $\alpha_0(z)$, indicates that for values of p approaching $p = 0$ the error is asymmetric: the approximation over-states the size of response when $b\Delta x < 0$ and understates it when $b\Delta x > 0$. The proportionate size of error increases without bound as $z \to -\infty$ but is bounded below by -1 as $z \to +\infty$.

For any $p \neq 0.5$, consider $\alpha_p(z) = \alpha_{0.5}(z) + (p - 0.5)z$. This has the appearance of a (distorted) rotation of $\alpha_{0.5}(z)$ around $z = 0$ - clockwise for $p < 0.5$ and anti-clockwise for $p > 0.5$. Instances of $p < 0.5$ are mirror images of the corresponding $p > 0.5$ - i.e. $\alpha_{0.5-\theta}(-z) = \alpha_{0.5+\theta}(z)$, $0 \le \theta \le 0.5$. Hence an investigation of one or the other will suffice. The analysis therefore proceeds for the case of clockwise rotation, illustrated by the following graph.



It is convenient to deal separately with positive and negative values of $z\,(\equiv b\Delta x)$. In the case of clockwise rotation, i.e. for $p < 0.5$, it can be seen that, at any value of z along the negative half-axis, the approximation error worsens relative to the case of $p = 0.5$. The further the departure of p from 0.5, the greater the error that results at any given $b\Delta x < 0$ when approximating $\Delta p / \Delta x_j$ by $\partial p / \partial x_j$. For any given value of p the size of error increases with the size of $b\Delta x$. An investigator who employs the approximation $\Delta p / \Delta x_j \approx \Delta x_j \times \partial p / \partial x_j$ when $b\Delta x < 0$ introduces an error whose size increase with the size of the coefficient, b, and with the size of the change in regressor value, Δx, and also with the size of $(p - 0.5)$.

For the other half-axis, $z\,(\equiv b\Delta x) > 0$, the consequences of the clockwise rotation of $\alpha_p(z)$ are less straightforward; $\alpha_p(z)$ decreases algebraically relative to $\alpha_{0.5}(z)$ but this algebraic reduction does not necessarily imply a reduction in the absolute size of approximation error. The slope of $\alpha_p(0)$ is steeper for smaller values of $p < 0.5$, causing the approximation errors close to the origin to be larger in size for smaller values of p. As z increases, however $\alpha_p(z)$ approaches a (negative) turning point. Thereafter its slope is positive, initially decreasing the size of negative error but eventually causing the error to become positive and thereafter to increase in size. The smaller the value of p, the less steep this positive slope and therefore the more extended the range of values of z for which the absolute size of $\alpha_p(z)$ is less than at its turning point.

Any arbitrary criterion for acceptable approximation error, such as (say) $\alpha_p(b\Delta x) < \pm 20\%$, is represented in the preceding graph by a horizontal band centred on the horizontal axis. The arguments of the previous paragraphs lead to the conclusion that

(i) on the half-axis $z\,(\equiv b\Delta x) < 0$, this criterion is only met when z occupies a range close to the origin, whose width narrows as p falls below $p = 0.5$;

(ii) on the half-axis $z\,(\equiv b\Delta x) > 0$, this criterion is also met when *z* occupies a range close to the origin, whose width narrows as p falls below $p = 0.5$, and is additionally satisfied for a second range of values of z whose width increases as p approaches $p = 0$. For values of p sufficiently close to $p = 0.5$, these two ranges will be contiguous[5].

Cases of $p > 0.5$ are a mirror image of the above.

We are now in a position to understand why evaluating the derivative $\partial p/\partial x_j$ provides an approximation to the discrete change $\Delta p/\Delta x_j$ that has been described as "often surprisingly accurate". It is because

- the approximation (not surprisingly) achieves whatever accuracy might be required when $b\Delta x$ occupies a range sufficiently close to zero

[5] This is because the minimum point of the proportionate error function then falls within the error tolerance band.

- when $p < 0.5$ it will also be acceptably accurate for some additional finite range within $b\Delta x > 0$, though not within $b\Delta x < 0$
- when $p > 0.5$ it will also be acceptably accurate for some additional finite range within $b\Delta x < 0$, though not within $b\Delta x > 0$

The end-points of these ranges are the roots of $\alpha_p(z) \pm \alpha^* = 0$, where α^* is the target level of accuracy. The roots can be obtained by numerical solution[6]; the following table gives the detail for the case where the target level of accuracy for the approximation is set at $\pm 5\%$.

Table 1

p	Range of $b\Delta x$ for which accuracy is within $\pm 5\%$.
0.1	Between -0.12 and 0.13 and also between 9.49 and 10.50
0.2	Between -0.16 and 0.18 and also between 4.50 and 5.09
0.3	Between -0.23 and 0.28 and also between 2.33 and 2.96
0.4	Between -0.38 and 1.64
0.5	Between -0.78 and 0.78
0.6	Between -1.64 and 0.38
0.7	Between -2.96 and -2.33 and also between -0.28 and 0.23
0.8	Between -5.09 and -4.50 and also between -0.18 and 0.16
0.9	Between -10.50 and -9.49 and also between -0.13 and 0.12

Table 1 confirms that there are combinations of p, b and Δx for which the strategy of employing point rates of response to approximate discrete changes can achieve a given target level of accuracy. However, these combinations do not dominate within the range of the table, which is $-10.50 < b\Delta x < +10.50$, and all combinations outside of this range fail to achieve the target accuracy level. Moreover, whenever the accuracy criterion is met for some $b\Delta x = z^*$, there is no guarantee that it is also met for $b\Delta x = -z^*$.

[6] See Appendix 2

It remains to be considered whether the approximation strategy has any improved chances of success when the regressor being considered for variation is itself binary, i.e. a $0/1$ dummy variable, and all other regressors are set equal to their sample means. This case deserves attention because it is one that has been used in applied work to characterise the influence of a binary regressor upon success probability and because it is described by Greene (2000: p817) as one in which the point rate of response "generally produce[s] a reasonable approximation to the change in probability". The analysis presented above has established that the extent of approximation error depends only upon the sign and size of $z\left(\equiv b\Delta x\right)$ and the success probability, p, associated with the given regressor values. Investigation of the case highlighted by Greene may therefore be limited to a consideration of whether or not it constrains the values of $b\Delta x$ and / or p.

Firstly, note that, despite the non-linearity of a Logit model, its Maximum Likelihood residuals are orthogonal to its regressors (Cameron and Trivedi, 2005: p469), with the implication that these residuals average to zero if the model includes an intercept. In this case, the sample average of fitted success probabilities, which, in the notation of equation (2), is $\bar{p} = \left(1/N\right)\sum p_i$, $p_i = F\left(x_i'\hat{\beta}\right)$, is equal to the observed sample frequency of successes: $\bar{y} = \left(1/N\right)\sum y_i$. Limiting attention to applications in which this sample frequency is in the range $0.3 < p < 0.7$ then permits appeal to Amemiya (1981), who notes (p1488) that a linear approximation to the logistic cumulative density function (cdf) "works well" for probabilities in this range. The implication of this approximate linearity is that setting regressors to their sample average values will produce an estimated probability approximately equal to the average of the fitted probabilities, i.e. $F\left(\bar{x}'\hat{\beta}\right) \approx \bar{p} = \bar{y}$, and therefore also in the quoted range. As the sample frequency of success moves away from the quoted range into the lower or upper tail then, because of the increased curvature of the cdf, $F\left(\bar{x}'\hat{\beta}\right)$ moves further into the tail than does $\bar{p}$, so that it is no longer possible to claim $F\left(\bar{x}'\hat{\beta}\right) \approx \bar{y}$. In summary, if the sample frequency of success is in the range $0.3 < p < 0.7$ then the success probability associated with regressors being at their sample means will probably also be in this range if the model employs an intercept. This places some limitations upon the behaviour of the approximation error, but not sufficient to lead to any useful conclusions without also considering the size of $b\Delta x$.

Since the change in a $0/1$ dummy variable is ± 1, then the quantity $z\left(\equiv b\Delta x\right)$ becomes in this case $\pm b$. Note from equation (1) that $\pm b$ is the change in the logarithm of the odds ratio that is occasioned by the regressor switching between its possible values of zero and unity. If it could be speculatively accepted that binary regressors do not commonly induce dramatic changes in the odds ratio in favour of success, this would imply a bound on the commonly encountered values of $|b|$. For example, an assertion that binary regressors are rarely capable of changing an odds ratio more dramatically than from (say) 50/50 to 75/25, which is a change of logit from 0 to 1.1, would imply that b rarely exceeds 1.1 in modulus. Table 1 then suggests that if, across a range of applications employing regressors at their sample average values, b varies uniformly within $1.1 < b < 1.1$ and $\bar{y}$ varies uniformly within $0.3 < \bar{y} < 0.7$, the chances of approximation accuracy being within $\pm 5\%$ in any one application are of the order of 50%. Widening the permitted range of variation for b and $\bar{y}$ would reduce the average accuracy of the approximation strategy.

Thus it is possible to put the case that "When applied at the point of regressor sample means, for logit models incorporating an intercept, and estimated on data for which success frequency is not particularly high or low, and in which the binary regressor being considered does not have a dramatically large influence, then the approximation strategy has a fair chance of meeting accuracy criteria that are not too stringent." This does not seem a sufficient basis upon which to argue in favour of employing the approximation as a matter of course, even for the limited circumstances highlighted by Greene (2000: p817).

Conclusion

In Logit models the link between regressors and regressand is highly non-linear. Greene (2000, p817) notes that, nevertheless, the strategy of using continuous derivatives to approximate the response of a binary regressand to a discrete change in a regressor value is may be acceptably accurate, particularly when regressors are at their sample average values. This note has demonstrated that, for logit models at least, such accuracy is fortuitous, being contingent in a complex manner upon the sign and size of the discrete change to be considered, the sign and size of the coefficient attaching to this regressor and the modelled value of the regressand prior to the discrete change. The approximation strategy should therefore be considered a risky one and exact calculation of the marginal effect of discrete regressors should be preferred to it. There has been no explicit analysis of Probit models above but Amemiya (1981: p1487) characterises them as approximately a linear transformation of Logit models for $0.3 < p < 0.7$, which suggests that similar conclusions may be expected for Probit models within this range of success probabilities. For probabilities outside of this range, as with Logit models, the increased non-linearity is sufficient to put in doubt the usefulness of linear approximations to the impact of discrete marginal changes in regressor values.

References

Amemiya, T. (1981) “Quantitative Response Models: A Survey” *Journal of Economic Literature* vol XIX (December 1981) pp 1483 – 1536

Cameron, A.C. and P.K. Trivedi (2005) *Microeconometrics: Methods and Applications* CambridgeU.P.

Greene, W.H. (2000) *Econometric Analysis* 4th edn. Prentice Hall

Appendix 1

This appendix provides support for assertions made within the text.

(page 4)

To establish that $\partial p / \partial x_j = p(1-p)\hat{\beta}_j$, begin with the fitted logit model,

$p = F\left(x'\hat{\beta}\right) = e^{x'\hat{\beta}} / \left(1 + e^{x'\hat{\beta}}\right)$ and apply the chain rule with $g = e^{x'\hat{\beta}}$ to get

$$\frac{\partial p}{\partial x_j} = \frac{\partial}{\partial g}\left(\frac{g}{1+g}\right) \times \frac{\partial g}{\partial x_j} \quad = \quad \frac{1}{(1+g)^2} \times g\hat{\beta}_j .$$

Recalling that $p = g/(1+g)$, and therefore $(1-p) = 1/(1+g)$, gives the required result.

(page 6)

To establish the actual discrete change in the estimated probability, begin with

$$p = \frac{e^{A+bx}}{1+e^{A+bx}} \quad , \quad p + \Delta p = \frac{e^{A+bx+b\Delta x}}{1+e^{A+bx+b\Delta x}}$$

Hence $\Delta p = \dfrac{e^{A+bx+b\Delta x}\left(1+e^{A+bx}\right) - e^{A+bx}\left(1+e^{A+bx+b\Delta x}\right)}{\left(1+e^{A+bx+b\Delta x}\right)\left(1+e^{A+bx}\right)}$, which

simplifies to $\Delta p = \dfrac{e^{A+bx+b\Delta x} - e^{A+bx}}{\left(1+e^{A+bx+b\Delta x}\right)\left(1+e^{A+bx}\right)}$

This may be recast as $\Delta p = \dfrac{(p+\Delta p)}{e^{A+bx+b\Delta x}} \times \dfrac{p}{e^{A+bx}}\left(e^{A+bx+b\Delta x} - e^{A+bx}\right)$ and then,

using $p = \dfrac{g}{1+g} \Rightarrow g = \dfrac{p}{1-p} \Rightarrow \dfrac{1}{g} = \dfrac{1-p}{p}$, as

$$\Delta p = (1-p-\Delta p)(1-p)e^{A+bx}\left(e^{b\Delta x} - 1\right).$$

Now, using $p = \frac{g}{1+g} \Rightarrow (1-p)g = p$, leads to $\Delta p = (1-p-\Delta p)p\left(e^{b\Delta x}-1\right)$ and thus to

$$\Delta p = (1-p)p\,\Phi \quad , \quad \Phi = \frac{\left(e^{b\Delta x}-1\right)}{1+p\left(e^{b\Delta x}-1\right)}$$

(page 7)

The proportionate approximation error is

$\alpha = \frac{{}^{a}\Delta p - \Delta p}{\Delta p} = \frac{(1-p)pb\Delta x - (1-p)p\Phi}{(1-p)p\Phi} = \frac{b\Delta x - \Phi}{\Phi} = \frac{b\Delta x}{\Phi} - 1$. Substitution of the detail for Φ gives $\alpha = b\Delta x \frac{1+p\left(e^{b\Delta x}-1\right)}{\left(e^{b\Delta x}-1\right)} - 1$, which can be re-arranged as $\alpha = \frac{b\Delta x + pb\Delta x\left(e^{b\Delta x}-1\right) - \left(e^{b\Delta x}-1\right)}{\left(e^{b\Delta x}-1\right)}$, i.e. $\alpha = \frac{b\Delta x - \left(e^{b\Delta x}-1\right)}{\left(e^{b\Delta x}-1\right)} + pb\Delta x$, which is presented in the text as $\alpha = \delta + pb\Delta x \quad , \quad \delta = \frac{b\Delta x - \left(e^{b\Delta x}-1\right)}{\left(e^{b\Delta x}-1\right)}$ and later as $\alpha(z) = \delta + pz \quad , \quad \delta(z) = \frac{z-\left(e^{z}-1\right)}{\left(e^{z}-1\right)} \quad , \quad z = b\Delta x$.

(page 8)

To show the symmetry of $\alpha_{0.5}(z)$, first recall that the text defines

$\alpha(z) = \delta(z) + pz \quad , \quad \delta = \frac{z-\left(e^{z}-1\right)}{\left(e^{z}-1\right)}$. The asymmetry in $\delta(z)$ is detailed as

$\delta(-z) = \frac{-z-\left(e^{-z}-1\right)}{\left(e^{-z}-1\right)} = \frac{-ze^{z}-\left(1-e^{z}\right)}{\left(1-e^{z}\right)} = \frac{ze^{z}-\left(e^{z}-1\right)}{\left(e^{z}-1\right)}$, which leads to

$$\delta(-z) = \frac{z-\left(e^{z}-1\right)+\left(ze^{z}-z\right)}{\left(e^{z}-1\right)} = \delta(z) + z$$

Consequently, we can show that $\alpha_p(z)$ is generally asymmetric:

$$\alpha_p(-z) = \delta_p(-z) - pz = \delta(z) + z - pz = \alpha_p(z) + (1-2p)z$$

But from this it follows that in the particular case of p=0.5 we have

$\alpha_p(-z) = \alpha_p(z)$.

(page 8)

To show the non-negativity of $\alpha_{0.5}(z)$, note that the symmetry of $\alpha_{0.5}(z)$ requires that non-negativity only be proved for $z > 0$ or for $z < 0$. Also note that $\delta(0) = 0$ and $\delta'(0) = -0.5$ so a demonstration that $\delta''(0) > 0$ will establish that $\delta(z) > -0.5z$, which proves the non-negativity of $\alpha_{0.5}(z)$.

By differentiation we have

$$\delta'(z) = \frac{\partial \delta}{\partial z} = \frac{(e^z - 1) - ze^z}{(e^z - 1)^2} \Rightarrow \delta''(z) = \frac{(e^z - 1)^2(-ze^z - 2e^z) + 2ze^{2z}(e^z - 1)}{(e^z - 1)^4}$$

Hence, $\delta''(z) > 0 \Leftarrow 2ze^{2z}(e^z - 1) > (e^z - 1)^2(2 + z)e^z$

Proceed initially with the assumption $z > 0$:

$$\delta''(z) > 0 \Leftarrow 2ze^z > (e^z - 1)(2 + z)$$

$$\delta''(z) > 0 \Leftarrow 2z(e^z - 1) + 2z > (e^z - 1)(2 + z)$$

$$\delta''(z) > 0 \Leftarrow z(e^z - 1) + 2z > 2(e^z - 1)$$

$$\delta''(z) > 0 \Leftarrow 2z > (2 - z)(e^z - 1)$$

Now,

$$2(e^z - 1) = 2\left\{z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \cdots\right\}$$

$$z(e^z - 1) = \left\{0 + \frac{z^2}{1!} + \frac{z^3}{2!} + \frac{z^4}{3!} + \cdots\right\}$$

So $(2 - z)(e^z - 1) = 2z + \left\{z^2\left(\frac{2}{2!} - \frac{1}{1!}\right) + z^3\left(\frac{2}{3!} - \frac{1}{2!}\right) + z^4\left(\frac{2}{4!} - \frac{1}{3!}\right) + \cdots\right\}$

i.e. $\left(2-z\right)\left(e^{z}-1\right)=2z+\left\{z^{2}\left(\frac{2-2}{2!}\right)+z^{3}\left(\frac{2-3}{3!}\right)+z^{4}\left(\frac{2-4}{4!}\right)+\cdots\right\}$

Hence, for $z>0$, $\left(2-z\right)\left(e^{z}-1\right)<2z \quad \Rightarrow \quad \delta''(z)>0$

Since $\delta(0)=0$, $\delta'(0)=-0.5$ and $z>0 \Rightarrow \delta''(z)>0$ then
$z>0 \Rightarrow \delta(z)>-0.5z$

When $\delta(z)>-0.5z$ then $\alpha_{0.5}(z)=\delta(z)+0.5z>0$.

This proves that $\alpha_{0.5}(z)$ is non-negative for $z>0$ and the previously established symmetry of $\alpha_{0.5}(z)$ then means that it is also non-negative for $z<0$. For completeness, note that $\alpha_{0.5}(0)=0$ because $\delta(0)=0$.

To confirm that, therefore, $\alpha_{0.5}(z)$ always "over-states" the (size of the) response, first express the non-negativity as $\frac{{}^{a}\Delta p-\Delta p}{\Delta p}>0$ when . $z\neq 0$. To confirm that this implies an over-statement of size, argue as follows.

- For $b\Delta x>0$, ${}^{a}\Delta p>0$ and $\Delta p>0$ so $\frac{{}^{a}\Delta p-\Delta p}{\Delta p}>0 \Rightarrow {}^{a}\Delta p>\Delta p$, i.e. $\left|{}^{a}\Delta p\right|>\left|\Delta p\right|$
- For $b\Delta x<0$, ${}^{a}\Delta p<0$ and $\Delta p<0$ so $\frac{{}^{a}\Delta p-\Delta p}{\Delta p}>0 \Rightarrow {}^{a}\Delta p<\Delta p$, i.e. $\left|{}^{a}\Delta p\right|>\left|\Delta p\right|$

(page 8)

Asymptotic behaviour of $\alpha(z)$:

As $z \to \infty$, e^z dominates z, so $\delta(z) = \frac{z}{(e^z - 1)} - 1 \to -1$ and hence $\alpha(z) = \delta(z) + pz \to -1 + pz$.

As $z \to -\infty$, $e^z \to 0$, so $\delta(z) = \frac{z}{(e^z - 1)} - 1 \to -z - 1$ and $\alpha(z) = \delta(z) + pz \to -1 - (1 - p)z$.

(page 9)

To show that clockwise rotations are "mirror images" of anti-clockwise rotations, i.e. $\alpha_{0.5-\theta}(-z) = \alpha_{0.5+\theta}(z)$, $0 \le \theta \le 0.5$, first establish

$$\alpha_p(z) = \delta(z) + pz = \delta(z) + 0.5z + (p - 0.5)z = \alpha_{0.5}(z) + (p - 0.5)z$$

Now use this to show that $\alpha_{0.5-\theta}(-z) = \alpha_{0.5+\theta}(z)$:

$$\begin{aligned}
&\alpha_{0.5-\theta}(-z) = \alpha_{0.5}(-z) + (0.5 - \theta - 0.5)(-z) \\
&= \alpha_{0.5}(z) + (\theta)(z) \\
&= \alpha_{0.5}(z) + (0.5 + \theta - 0.5)z \\
&= \alpha_{0.5+\theta}(z)
\end{aligned}$$

ADDITIONAL LY

a) confirming that $sign(\Phi)=sign(b\Delta x)$

z > 0:

$$\Phi=\frac{\left(e^{z}-1\right)}{1+p\left(e^{z}-1\right)} \quad , \quad z=b\Delta x$$

$$z>0\Rightarrow e^{z}>1\Rightarrow\Phi>0$$

z < 0:

$$\Phi=\left(e^{z}-1\right)/\left[(1-p)+pe^{z}\right]$$

$$0<p<1,0<e^{z}<1\Rightarrow\Phi<0$$

b) slope of $\delta(z)$, $\alpha(z)$

For $z\neq 0$

$$\delta(z)=\frac{z}{\left(e^{z}-1\right)}-1\Rightarrow\delta'(z)=\frac{\partial\delta}{\partial z}=\frac{\left(e^{z}-1\right)-ze^{z}}{\left(e^{z}-1\right)^{2}}$$ and

$$\alpha'(z)=\delta'(z)+p$$

As $z\rightarrow 0$

$$\delta'(z)=\frac{e^{z}-1-ze^{z}}{\left(e^{z}-1\right)^{2}}=\frac{A}{B}$$

We use $e^{z}=1+\frac{z}{1!}+\frac{z^{2}}{2!}+\frac{z^{3}}{3!}+\cdots$ and $ze^{z}=z+\frac{z^{2}}{1!}+\frac{z^{3}}{2!}+\frac{z^{4}}{3!}+\cdots$ to get $A=e^{z}-1-ze^{z}=z^{2}\left\{\frac{1}{2!}-\frac{1}{1!}\right\}+z^{3}\left\{\frac{1}{3!}-\frac{1}{2!}\right\}+z^{4}\left\{\frac{1}{4!}-\frac{1}{3!}\right\}+\cdots$. i.e.

$A = z^2\left[\left\{\frac{1}{2!}-\frac{1}{1!}\right\}+z^3\left\{\frac{1}{3!}-\frac{1}{2!}\right\}+z^4\left\{\frac{1}{4!}-\frac{1}{3!}\right\}+\cdots\right]$. We write this as So as $A = z^2\left[-\frac{1}{2}+A_0\right]$, where $A_0 \to 0$ as $z \to 0$,

We use $e^z - 1 = \frac{z}{1!}+\frac{z^2}{2!}+\frac{z^3}{3!}+\cdots = z[1+B_0]$, where $B_0 = \frac{z}{2!}+\frac{z^2}{3!}+\frac{z^3}{4!}+\cdots \to 0$ as $z \to 0$. To get $B = \left(e^z - 1\right) = z^2[1+B_0]^2$

Hence, $\delta'(z) = \frac{A}{B} = \frac{z^2[-1/2+A_0]}{z^2[1+B_0]^2} = \frac{-1/2+A_0}{[1+B_0]^2}$ and as $z \to 0$: $A_0 \to 0$, $B_0 \to 0$, $\delta'(z) \to -0.5 = \delta'(0)$.

Consequently, $\alpha'(z) = \delta'(z) + p \to p - 0.5$ as $z \to 0$; in particular $\alpha'_{0.5}(z) \to 0$.

Appendix 2

This appendix provides the roots of $\alpha_p(z) \pm \alpha^* = 0$ for several combinations of success probability (p) and accuracy criterion ($\pm\alpha^*$). The roots have been obtained numerically by a first-order Newton-Raphson iteration implemented in Excel.

target accuracy = +/- 5% (used to provide Table 1 in the text)

p	+0.05	-0.05	-0.05	+0.05
0.1	-0.1219	0.1284	9.4928	10.4971
0.2	-0.1596	0.1752	4.4966	5.0926
0.3	-0.2283	0.2834	2.3290	2.9613
0.4	-0.3800			1.6354
0.5	-0.7785			0.7785
0.6	-1.6354			0.3800
0.7	-2.9613	-2.3290	-0.2834	0.2283
0.8	-5.0926	-4.4966	-0.1752	0.1596
0.9	-10.4971	-9.4928	-0.1284	0.1219

target accuracy = +/- 10%

p	+0.1	-0.1	-0.1	+0.1
0.1	-0.2382	0.2646	8.9888	10.9982
0.2	-0.3072	0.3716	4.1737	5.3750
0.3	-0.4250	0.7059	1.8479	3.2200
0.4	-0.6502			1.9316
0.5	-1.1065			1.1065
0.6	-1.9316			0.6502
0.7	-3.2200	-1.8479	-0.7059	0.4250
0.8	-5.3750	-4.1737	-0.3716	0.3072
0.9	-10.9982	-8.9888	-0.2646	0.2382

target accuracy = +/- 20%

p	+0.2	-0.2	-0.2	+0.2
0.1	-0.4567	0.5665	7.9725	11.9993
0.2	-0.5752	0.8782	3.4230	5.9203
0.3	-0.7610			3.6834
0.4	-1.0677			2.4023
0.5	-1.5806			1.5806
0.6	-2.4023			1.0677
0.7	-3.6834			0.7610
0.8	-5.9203	-3.4230	-0.8782	0.5752
0.9	-11.9993	-7.9725	-0.5665	0.4567