

## Trends and Ticks

Last updated: December 2014

The figures in Family Food are sourced from The Living Costs and Food Survey run by the Office for National Statistics. One element of the survey - The Family Food Module collects detailed quantity and expenditure information on food and drink household purchases and itemised lists of food and drink eating out purchases for use by Defra.

The Office for National Statistics has overall project management and financial responsibility for the survey while Defra sponsors the specialist food data.

### Background

Trend indicators, reliability ticks and significance of change are published alongside many of the estimates from Family Food. This note describes the method of calculation used.

The aim of these quality assessments is to make estimates of averages per person per week of expenditures and quantities of food purchases easier to interpret and use. In all cases the method is approximate and based on sampling errors ignoring any other kinds of error.

### Reliability ticks

The reliability ticks come directly from the approximate standard errors of the estimates.

The standard error formula is for a ratio estimate where the ratio is the sum of purchases divided by the sum of people in the purchasing households. A version of the formula without survey weights can be found in 'Sampling Errors' by Cochran and Cox.

The formula for producing the standard error for a ledger incorporating survey weights is as follows:

Let the survey weight for a household be denoted by “w”.

Let the amount recorded in the ledger by a household be denoted by “x”.

Let the weighted total of w times x across households be denoted by T(wx).

Let the weighted total of w<sup>2</sup> times x<sup>2</sup> across households be denoted by T(w<sup>2</sup>x<sup>2</sup>).

Let the number of persons in a household be denoted by “p”.

Let the weighted total number of persons across households be denoted by T(wp)

Let the weighted total of p<sup>2</sup> across households be denoted by T(w<sup>2</sup>p<sup>2</sup>).

Let the weighted total of x times p across households be denoted by T(w<sup>2</sup>xp).

Let N be the total number of households.

Let the weighted average for the ledger, expressed per person, be  $A(wx) = T(wx)/T(wp)$ .

It is the standard error of A(wx) which is denoted by SE(A(wx)), that is required. The approximation which is accurate for large N is as follows.

$$SE(A(wx)) = A(wx) * \sqrt{ n/(n-1) * }$$

$$((T(w^2x^2)/T(wx)^2) - (2*T(w^2xp)/T(wx)T(wp)) + (T(w^2p^2)/T(wp)^2)) )$$

A feature of the approximation is that zero averages always produce zero approximate standard errors. They should not be taken to indicate that zero rates of consumption are estimated to high accuracy since these have been removed from the datasets.

The estimated standard error is converted into a relative standard error by expressing it as a percentage of the estimate itself. The ticks are derived directly from the relative standard error as given by the following ranges:

**Table 1: How to interpret ticks and crosses**

Reliability Indicator	Relative Standard Error (RSE) of the estimate
✓✓✓	0 to 2.5%
✓✓	2.5% to 5%
✓	5% to 10%
blank	10% to 20%
✗	Greater than 20%
-	Not available

As an example the estimate of the reliability flag for the average quantity of purchases of olive oil per person in 2005-06 in the UK is shown below.

**Example: Calculation of reliability flag for olive oil purchases**

Formulae	Value of expression
$T(wx)$	1365305
$T(w^2x^2)$	6906821512
$T(wp)$	58473
$T(w^2p^2)$	732504
$T(w^2xp)$	16473025
$A(wx)$	12
$T(w^2x^2)/T(wx)^2$	0.00371
$T(w^2xp)/T(wx)T(wp)$	0.00041
$T(w^2p^2)/T(wp)^2$	0.00021
$SE(A(wx))$	0.691
RSE	5.9
Reliability flag	✓

Note:

(a) The estimated quantity per week,  $A(wx)$ , involves dividing by two to convert from two weekly diary period to purchases per week. This does not affect the calculation of standard error.

(b) The factor of  $n/n-1$  is ignored because it is very close to unity.

### Significance of changes

The statistical significance of year on year changes comes from the estimated standard errors of the two annual estimates. The survey samples underlying the estimates for any two non-overlapping years can be taken as independent.

The method is to calculate the variance of the change as the sum of the variances of the two independent estimates. The calculation of the reliability flag described above already gives an estimate of the standard error of the estimate for a particular year.

Let  $SE(A_1(wx))$  be the estimate of standard error of the quantity in year 1.

Let  $SE(A_2(wx))$  be the estimate of standard error of the quantity in year 2.

Then:

$$\text{Variance}(A_1(wx)) = SE(A_1(wx))^2$$

$$\text{Variance}(A_2(wx)) = SE(A_2(wx))^2$$

$$\text{Variance}(A_2(wx) - A_1(wx)) = SE(A_1(wx))^2 + SE(A_2(wx))^2 \quad (\text{since independent})$$

$$SE(A_2(wx) - A_1(wx)) = \text{SQRT}(SE(A_1(wx))^2 + SE(A_2(wx))^2)$$

The difference between two estimates is described as statistically significant whenever the interval given by the difference  $\pm$  twice the standard error excludes zero, i.e. the confidence interval around the difference excludes the possibility of no change.

The rationale is that if the interval includes zero then there is insufficient evidence to say that there is any difference at all. If (a) the interval excludes zero, (b) the estimated standard error is accurate and (c) the estimate of the difference follows a normal distribution then as an approximation there is less than a one in twenty chance that the population difference is zero.

### Trend indicators

Trend indicators in the form of an arrow are intended to provide a guide for users of the statistics about whether there really is a short term trend. Four years is chosen as the period over which to check for presence of a statistically significant trend, since it is considered long enough to show a trend and short enough to be current.

The method is intentionally simple in order that a wide range of users will be able to understand and reproduce the calculation. It treats four annual estimates as independent measurements and examines the linear regression slope estimator. As a consequence it is sub-optimal (a better method would work directly on the raw data for all four years, taking into account the actual date of the diary.)

Bearing in mind that the x's are really constants the gradient of the linear regression slope estimator is given by:

$$\begin{aligned}
 m &= (nS_{xy} - S_x S_y) / (nS_{xx} - S_x S_x) \\
 &= (S_{xy} - (S_x/n) * S_y) * \text{constant} \\
 \alpha &= \Sigma(xy - x\bar{y}) \\
 &= \Sigma y(x - \bar{x})
 \end{aligned}$$

where the y's are the estimates of average purchases per person per year, the x's are the years.

In determining the statistical significance of the trend it makes no difference if we examine the gradient itself or the gradient times a constant. If one is significantly different from zero then so too for the other.

Because the gradient is proportional to  $\Sigma y(x - \bar{x})$  which includes x's only through  $\bar{x}$ , it makes no difference how large the x's are. It is only the deviations from their mean that are important. With four contiguous x values this always produces the same answer. The deviations are -1.5, -0.5, 0.5 and 1.5.

In the spreadsheet they are scaled down without loss of proportionality to -0.3, -0.1, 0.1 and 0.3. With this scale the expression gives the estimate of the actual gradient, but this is unnecessary for determining the significance of the trend.

We now have:

$$\begin{aligned}
 m &= \Sigma y(x - \bar{x})/5 = 0.3y_4 + 0.1y_3 - 0.1y_2 - 0.3y_1 \\
 \text{Var}(m) &= 0.09\text{var}(y_4) + 0.01\text{var}(y_3) + 0.01\text{var}(y_2) + 0.09\text{var}(y_1) \\
 &\text{(the standard error of } m = \text{the square root of the variance of } m)
 \end{aligned}$$

The standard errors of the y's are already published and thus their variances can be calculated and the expression evaluated.

Statistical significance is taken to occur when m is greater than twice its standard error. This means that there is approximately a one in twenty chance that a true zero trend would give rise to such a large observed trend. We take this as meaning that there is a one in twenty chance that there is no trend.

When significance is found the user is left to examine the four estimates and decide what it really means in terms of trend.

