

## Labour Force Survey – Estimating and reporting uncertainty

---

In the [Labour Market Report](#), the headline estimate of unemployment, employment and economic inactivity are presented alongside a **confidence interval** and a statement of whether changes over the quarter or year are **statistically significant**.

This short guide explains the two concepts and how you can use the information to better understand estimated changes in labour market status. Section 2 provides information on how confidence intervals are calculated for the Labour Market Report and contains information for more technical users.

### 1. Confidence Intervals for Survey Estimates

The Labour Force Survey (LFS) is a sample survey. It provides *estimates* of population values (the number and rate of people who are unemployed, employed or economically inactive in NI). The difference between an estimate from a sample and the (unknown) population value is the result of “sampling errors” and “non-sampling errors”.

- Sampling errors are due to only using a sample and not the entire population, therefore estimates from samples may not equal the unknown population value. The larger the sample, the smaller the sampling error. Confidence intervals are one way of providing information on sampling error.
- Non-sampling errors are due to aspects of surveys other than the sampling, such as households being unreachable or refusing to participate in a survey, imperfections in the sampling frame, issues in the questionnaire design, inaccurate answers from respondents, processing and analysis errors. These types of errors would be present in the statistics even if the entire population was surveyed and can be difficult to quantify using the sample data alone. Steps are taken to ensure some of these errors are mitigated for the LFS, such as strict quality control measures during data processing and analysis, continuous improvement of the sampling frame such as removal of ineligible addresses, extensive testing of the questionnaire etc.

We provide information on sampling error through **confidence intervals**. The difference between the true population value and the estimate based on a sample is called the sampling error, and its variability is measured by the standard error of the estimate. The standard error can be used to calculate the confidence interval around the estimate. Confidence intervals are a standard way of expressing the statistical accuracy of survey based estimates. Typically 90%, 95% or 99% confidence intervals are used. The Labour Market Report includes **95%** confidence intervals around headline estimates. If a large number of random samples were taken, the 95% confidence intervals of the resulting estimates would contain the population value 95% of the time. The estimate and the 95% confidence interval are presented as follows:

**Estimate +/- confidence interval.**

Confidence intervals calculated at the 95% level are 1.96 standard errors above and below the sample estimate. Therefore, the larger the sampling error, the larger the corresponding confidence interval.

Tables 1 & 3 in the monthly Labour Market Report (LMR) provide estimates of levels, rates and changes in headline labour market statistics. The LMR also provides confidence intervals around estimates and changes over time.

## Examples

---

The following examples are intended to help users understand how to use the information on sampling variability provided in the LMR.

### Example 1

The seasonally adjusted estimate for the number of employed people aged 16 years and over for NI in [March to May 2019](#) was 871,000; with a confidence interval of +/- 19,000

This means that the best estimate of number of employed people was 871,000. The confidence interval shows the precision of the estimate. At +/- 19,000 the confidence interval is relatively small, meaning the estimate is considered precise. In 95 out of 100 samples the true population value would lie within the 95% confidence interval, in this case between 852,000 and 890,000.

The seasonally adjusted estimate for the number of unemployed people aged 16 years and over for NI in [March to May 2019](#) was 28,000; with a confidence interval of +/- 6,000

This means that the best estimate of number of unemployed people was 28,000. The confidence interval shows the precision of the estimate. At +/- 6,000 the confidence interval indicates that the estimate of unemployment is acceptably precise. It is less precise than the estimate of employment. In 95 out of 100 samples the true population value would lie within the 95% confidence interval, in this case between 22,000 and 34,000.

**Table 1: Northern Ireland Labour Market Summary with sampling variability**

<b>March-May 2019</b>	<b>Estimate</b> (confidence interval)	<b>Change over quarter</b> (confidence interval)	<b>Change over year</b> (confidence interval)
Unemployment <sup>1</sup>	<b>28,000</b> (+/- 6,000)	<b>1,000</b> (+/-5,000)	<b>-2,000</b> (+/-9,000)
Employment <sup>2</sup>	<b>871,000</b> (+/-19,000)	<b>11,000</b> (+/-16,000)	<b>21,000</b> (+/-28,000)
Economically inactive <sup>2</sup>	<b>570,000</b> (+/-19,000)	<b>-11,000</b> (+/-16,000)	<b>-13,000</b> (+/-27,000)
Unemployment rate <sup>1</sup>	<b>3.1%</b> (+/-0.7)	<b>0.1pps</b> (+/-0.6pps)	<b>-0.3pps</b> (+/-1.0pps)
Employment rate <sup>2</sup>	<b>71.7%</b> (+/-1.6)	<b>-0.5pps</b> (+/-1.2pps)	<b>1.7pps</b> (+/-2.3pps)
Economic inactivity rate <sup>2</sup>	<b>25.9%</b> (+/-1.5)	<b>-0.6pps</b> (+/-1.2pps)	<b>-1.5pps</b> (+/-2.2pps)

LFS data are seasonally adjusted.

<sup>1</sup> People aged 16 and over. Unemployment rate = total unemployed as a proportion of the economically active.

<sup>2</sup> Levels for all persons aged 16 and over, rates for working age (16-64).

## Example 2

The seasonally adjusted estimate for the employment rate of people aged 16-64 NI in [March to May 2019](#) was 71.7%; with a confidence interval of +/- 1.6 percentage points (pps).

This means that the best estimate for the employment rate was 71.7%. The confidence interval shows the precision of the estimate. At +/- 1.6pps the confidence interval is relatively small, meaning the estimate is considered precise. In 95 out of 100 samples the true population value would lie within the 95% confidence interval, in this case between 70.2% and 73.3%.

The seasonally adjusted estimate for the unemployment rate of people aged 16 years and over for NI in [March to May 2019](#) was 3.1%; with a confidence interval of +/-0.7pps.

This means that the best estimate for the unemployment rate was 3.1%. The confidence interval shows the precision of the estimate. At +/-0.7pps the confidence interval indicates that the estimate of unemployment is acceptably precise. It is less precise than the estimate for the employment rate. In 95 out of 100 samples the true population value would lie within the 95% confidence interval, in this case between 2.4% and 3.8%.

## 2. Statistical Significance

**Statistical significance** helps us distinguish what observed changes or relationships in data we should pay attention to, and which are likely to be down to chance alone.

### Example

The third bullet point in the Key Points section of the [January to March 2019](#) publication reads:

*The LFS indicated that the NI unemployment rate (16+) decreased over the quarter (0.9pps) and the year (0.2pps) to 2.9% in January-March 2019. The quarterly change was statistically significant, and is likely to reflect real change. The NI unemployment rate was below the UK rate (3.8%), the Republic of Ireland rate (5.6%) and the EU rate (6.5%).*

In the example above, the quarterly change was statistically significant. By this we mean that the size of the quarterly change was large enough that it was unlikely to have resulted only from the variable nature of samples.

To determine whether the quarterly change was statistically significant, we first calculated the difference between the two quarters (see table below circled green), then the confidence interval around this difference (circled in blue). As the change over the quarter was greater than the sampling variability of the change, the change was determined to be statistically significant.

**Table 1: Northern Ireland Labour Market Summary with sampling variability**

<b>January-March 2019</b>	<b>Estimate</b> (confidence interval)	<b>Change over quarter</b> (confidence interval)	<b>Change over year</b> (confidence interval)
Unemployment <sup>1</sup>	<b>26,000</b> (+/- 6,000)	<b>-8,000</b> (+/-7,000)	<b>-1,000</b> (+/-9,000)
Employment <sup>2</sup>	<b>860,000</b> (+/-19,000)	<b>7,000</b> (+/-15,000)	<b>14,000</b> (+/-28,000)
Economically inactive <sup>2</sup>	<b>582,000</b> (+/-18,000)	<b>3,000</b> (+/-13,000)	<b>-7,000</b> (+/-26,000)
Unemployment rate <sup>1</sup>	<b>2.9%</b> (+/-0.7)	<b>-0.9pps</b> (+/-0.8pps)	<b>-0.2pps</b> (+/-1.0pps)
Employment rate <sup>2</sup>	<b>71.3%</b> (+/-1.4)	<b>1.0pps</b> (+/-1.1pps)	<b>1.5pps</b> (+/-2.0pps)
Economic inactivity rate <sup>2</sup>	<b>26.5%</b> (+/-1.4)	<b>-0.2pps</b> (+/-1.0pps)	<b>-1.4pps</b> (+/-2.0pps)

LFS data are seasonally adjusted.

<sup>1</sup> People aged 16 and over. Unemployment rate = total unemployed as a proportion of the economically active.

<sup>2</sup> Levels for all persons aged 16 and over, rates for working age (16-64).

**For more technical users**

The confidence interval around the quarterly change in the unemployment rate in January to March was +/- 0.8pps.

This was calculated by calculating the standard error of the change using the following formula and multiplying by 1.96, the bound for a 95% confidence interval:

$$\text{Confidence interval} = 1.96 * \text{standard error}$$

$$\text{SE(prop)C} = \text{SQRT}((V1+V2)*(1-0.8c))$$

Where

$$V = \text{Variance} = (\text{DEFT} * (\text{SQRT}(p(1-p)/n)))^2$$

**SE** is standard error

**Prop C** is the difference between the two quarter's unemployment rates

**SQRT** is the square root

**V1** is the variance around the previous quarter's unemployment rate

**V2** is the variance around the current quarter's unemployment rate

**c** is the correlation coefficient (in this case 0.59)

**0.8** accounts for the 80% sample overlap between quarters

**DEFT** is the Design Factor (in this case 1.13)

**p** is the unemployment rate (the percentage of economically active people who are unemployed, weighted)

**n** is the unweighted base (in this case the number of unweighted economically active persons)

Therefore, in the January to March quarter the confidence interval around the change in the unemployment rate was calculated as follows –

$$1.96 * \text{standard error}$$

$$1.96 * \text{SQRT}((0.0000179 + 0.0000115) * (1 - 0.8 * 0.59))$$

Where

$$V1 = 0.0000179 = (1.13 * (\text{SQRT}(3.76\% * (1 - 3.76\%) / 2582)))^2$$

$$V2 = 0.0000115 = (1.13 * (\text{SQRT}(2.86\% * (1 - 2.86\%) / 3082)))^2$$