9.1 Case Study: Question Effects in a Survey

Often when respondents are asked for their opinion on an issue, some want to choose a middle or neutral position. The question to be asked here is "Do you think that giving buses priority at traffic signals will increase or decrease traffic congestion?" People will often state a middle position even if it is not offered to them. Whether or not it should be offered is a matter of some controversy. One school holds that where respondents fall into a middle position and it is not offered to them, they might feel constrained to choose a false response. Another argues that respondents do in fact lean to one side or the other of neutral and should be "forced" to make a decision. A sensible compromise is that if you are interested in determining *leanings*, don't suggest a middle ground, whereas if you are interested in *convictions* then do suggest a middle ground.

The data in Table 1 was collected as part of an experiment conducted by Kalton et al. [1978][1] to assess wording effects of survey questions (see also the Chapter 8 Case Study on question-wording effects given on this web site). This experiment was performed in two locations, Lancashire and London. In both locations, one sample of people was asked simply "Do you think that giving buses priority at traffic signals will increase or decrease traffic congestion?". In contrast, another roughly equal-sized group was asked "Do you think that giving buses priority at traffic signals will increase or decrease traffic congestion *or would it make no difference*?"[2] Table 1 gives the proportion of people making a positive statement (i.e. either congestion will increase or it will decrease).[3] In each case the number appearing in brackets under the proportion is the size of the sample involved. These samples are large enough for our usual large sample Normal theory to apply to proportions.

Table 1 : Bus Priority Question[a]

	Neutral option?		
	Without	With	Difference
	$\widehat{p}_1$ (n_1)	$\widehat{p}_2$ (n_2)	$\widehat{d} = \widehat{p}_1 - \widehat{p}_2$
Lancashire	0.55 (482)	0.35 (496)	0.20
London	0.50 (594)	0.38 (585)	0.12

[a]Source: Kalton et al. [1978].

[1]Kalton, G. and Schuman, H. (1982). The effect of the question on survey responses: A review. *Journal of the Royal Statistical Society, Series A*, **145**, 42–73.
[2]Our italics, to stress the difference in wording.
[3]Even when the neutral option was not suggested, 29% of respondents gave it. 18% felt unable to answer the question when there was no neutral option, which was little different from the 16% who felt unable to answer when there was a neutral option.

Let p_1 be the underlying true proportion who would give a definite opinion using the first form of the question (which doesn't suggest a neutral position), and p_2 be the underlying true proportion who would give a definite opinion using the second form of the question (which does suggest a neutral position). The difference $d = p_1 - p_2$ measures the effect of the change in the wording of the question. Estimates of the question effect are therefore given by the difference column above.

How strong is the evidence for the existence of a question effect in London? We will conduct a formal test of $H_0 : p_1 - p_2 = 0$. Since we might expect more people to make a positive response when they had to supply the neutral option themselves rather than having the question supply it, the alternative hypothesis we should be using is $H_1 : p_1 - p_2 > 0$. The test statistic has observed value

$$t_0 = \frac{\widehat{\theta} - \theta_0}{\mathrm{se}(\widehat{\theta})} = \frac{(\widehat{p}_1 - \widehat{p}_2) - 0}{\mathrm{se}(\widehat{p}_1 - \widehat{p}_2)}.$$

Here $\widehat{p}_1 = 0.50$, $n_1 = 594$, $\widehat{p}_2 = 0.38$, and $n_2 = 585$ so that $\widehat{p}_1 - \widehat{p}_2 = 0.12$. Since we are comparing proportions from two separate groups of people, we have two independent samples (cf. Fig. 8.5.1 (a) in the text). Thus, using the formula from Table 8.5.5(a)

$$\mathrm{se}(\widehat{p}_1 - \widehat{p}_2) = \sqrt{\frac{\widehat{p}_1(1 - \widehat{p}_1)}{n_1} + \frac{\widehat{p}_2(1 - \widehat{p}_2)}{n_2}} = 0.02869861$$

and $t_0 = 0.12/0.02869861 = 4.181387$. The *P-value* for this (one-sided) test, is

$$\begin{aligned} \textit{P-value} &= \mathrm{pr}(T > 4.181387), \quad \text{where } T \sim \mathrm{Normal}(0,1) \\ &= 0.00003\,. \end{aligned}$$

As the *P-value* is so very small, we have extremely strong evidence against H_0. This means that we have extremely strong evidence that a larger proportion of Londoners do make a positive statement when the possibility of making a neutral response is not suggested in the wording of the question.

How big a difference does it make? To answer this, we will compute a 95% confidence interval for the question effect $p_1 - p_2$ in London. Since $z(0.025) = 1.96$, our interval is given by

$$\widehat{p}\_\widehat{p}_2 \;\pm\; 1.96\,\mathrm{se}(\widehat{p}_1 - \widehat{p}_2) = 0.12 \;\pm\; 1.96 \times 0.0287 = [0.06, 0.18].$$

Thus, with 95% confidence, somewhere between 6% and 18% more Londoners would have taken a neutral position on this question if one is suggested to them than would have done so otherwise.

The question effect looks even stronger in Lancashire than it was in London. We will test for a difference between the question effect for the two areas. As in Case Study 8.1 on this web site, let d_{Lond} be the question effect in London (i.e. $p_1 - p_2$ for London). Our estimate of d_{Lond} is $\widehat{d}_{\text{Lond}} = 0.12$ which, as we have calculated above, has a standard error of 0.02869861. Let d_{Lanc} be the question effect in Lancashire (i.e. $p_1 - p_2$ for Lancashire). Our estimate of d_{Lanc} is $\widehat{d}_{\text{Lanc}} = 0.20$. Arguing as we did for London above, this estimate has a standard error of 0.0311794. These estimates are independent (as they come from physically independent samples) so that the standard error of the difference is given by (see Section 7.5)

$$\begin{aligned} \text{se}(\widehat{d}_{\text{Lanc}} - \widehat{d}_{\text{Lond}}) &= \sqrt{\text{se}(\widehat{d}_{\text{Lanc}})^2 + \text{se}(\widehat{d}_{\text{Lond}})^2} \\ &= \sqrt{0.0311794^2 + 0.02869861^2} = 0.04237647. \end{aligned}$$

We will test $H_0 : d_{\text{Lanc}} - d_{\text{Lond}} = 0$ (or equivalently $d_{\text{Lanc}} = d_{\text{Lond}}$). We should use the two-sided alternative $H_1 : d_{\text{Lanc}} - d_{\text{Lond}} \neq 0$ as we had no prior belief that the question effect would be larger in Lancashire. Here $\theta = d_{\text{Lanc}} - d_{\text{Lond}}$, $\theta_0 = 0$, $\widehat{\theta} = \hat{q}_{\text{Lanc}} - \hat{q}_{\text{Lond}}$ and the test statistic is

$$t_0 = \frac{\widehat{\theta} - \theta_0}{\text{se}(\widehat{\theta}\,)} = \frac{(\hat{q}_{\text{Lanc}} - \hat{q}_{\text{Lond}}) - 0}{\text{se}(\hat{q}_{\text{Lanc}} - \hat{q}_{\text{Lond}})} = \frac{0.20 - 0.12}{0.04237647} = 1.88784.$$

The *P-value* for this (two-sided) test is

$$\begin{aligned} P\text{-}value &= 2 \times \text{pr}(T > 1.88784), \qquad \text{where } T \sim \text{Normal}(0,1), \\ &= 0.059\,.\text{"} \end{aligned}$$

We therefore have some evidence of a difference in question effect between the two areas. To get an idea of the possible size of the difference between the question effect in Lancashire and London, we compute a 95% confidence interval, namely $0.20 - 0.12 \;\pm\; 1.96 \times 0.04238$, or $[-0.003, 0.163]$. The question effect is unlikely to be noticeably smaller in Lancashire than London, but could be larger by up to 16% of respondents.

We have established above that with the bus priority question, offering a neutral option made quite a difference to the proportions of people making a positive statement. However, this is not always the case. In the same study the researchers asked, “Do you think that car drivers should pay more or less for parking in city centers than they do now?” with half the respondents offered the neutral option “or should charges be kept as they are now?”. This time 42% of the sample made a positive response (i.e. pay more or pay less) when the neutral option was offered compared with 46% when is was not!

Newspapers and Market Researchers quite commonly quote percentages of those people who support a proposition *among those who state a definite opinion.* Does suggesting a neutral option make any appreciable difference to the proportion of people who thought congestion would increase among those offering a positive statement? Of the combined sample (i.e. Lancashire and London) 53% of the "without neutral option" group (570 people) made a positive response. Of these 62.3% thought congestion would increase. Of all respondents to which a neutral option was suggested 37% made a positive response (400 people). Of these 67.6% thought congestion would increase. Is there any evidence that the proportions giving a "congestion will increase" answer is different between the two groups?

Let p_1 be the population proportion for the first group (no neutral option) and p_2 for the second. We will test $H_0 : p_1 - p_2 = 0$ (no difference) versus $H_1 : p_1 - p_2 \neq 0$ (since we have no prior reason to believe that the difference will go in any particular direction). We have $n_1 = 570$, $\widehat{p}_1 = 0.623$, $n_2 = 400$, and $\widehat{p}_2 = 0.676$. Hence

$$\mathrm{se}(\widehat{p}_1 - \widehat{p}_2) = \sqrt{\frac{\widehat{p}_1(1-\widehat{p}_1)}{n_1} + \frac{\widehat{p}_2(1-\widehat{p}_2)}{n_2}} = 0.03097764,$$

and our test statistic is

$$t_0 = \frac{\widehat{\theta} - \theta_0}{\mathrm{se}(\widehat{\theta}\,)} = \frac{\widehat{p}_1 - \widehat{p}_2}{\mathrm{se}(\widehat{p}_1 - \widehat{p}_2)} = \frac{0.623 - 0.676}{0.03097764} = -1.710911.$$

The *P-value* for this (2-sided) test is

$$\begin{aligned} P\text{-}value &= 2 \times \mathrm{pr}(T > |t_0|) \qquad \text{where } T \sim \mathrm{Normal}(0,1) \\ &= 2 \times \mathrm{pr}(T > 1.711) = 0.087\,. \end{aligned}$$

We only have weak evidence of a difference. A 95% confidence interval for the true difference is given by $0.623 - 0.676 \;\pm\; 1.96 \times 0.03097764$, or $[-0.114, 0.008]$. So with 95% confidence, p_1 is somewhere between being smaller than p_2 by 11% of respondents and being bigger by 1%. Putting all this together, the difference could be of some practical importance but we have insufficient data to establish either whether it is important or in what direction the difference lies. We do know, however, that p_1 is unlikely to be appreciably bigger than p_2. ∎