



Ipsos Poll Conducted for Reuters

Awareness of Alibaba 09.18.2014

These are findings from an Ipsos poll conducted for Thomson Reuters from September 10-15, 2014. For the surveys, a sample of 1,604 Americans 18+ were interviewed online. The precision of the Reuters/Ipsos online polls is measured using a [credibility interval](#). In this case, the poll has a credibility interval of plus or minus 2.8 percentage points. For more information about credibility intervals, please see the appendix.

The data were weighted to the U.S. current population data by gender, age, education, and ethnicity. Statistical margins of error are not applicable to online polls. All sample surveys and polls may be subject to other sources of error, including, but not limited to coverage error and measurement error. Figures marked by an asterisk () indicate a percentage value of greater than zero but less than one half of one per cent. Where figures do not sum to 100, this is due to the effects of rounding. To see more information on this and other Reuters/Ipsos polls, please visit*

<http://polling.reuters.com/>.

AWARENESS OF ALIBABA

Q1. From the list below, please choose the topics or issues you have heard or read anything about in the past few days. You can choose as many or few as needed.

Chinese e-commerce website Alibaba and their upcoming stock initial public offering (IPO)	12%
---	-----

How to Calculate Bayesian Credibility Intervals

The calculation of credibility intervals assumes that Y has a binomial distribution conditioned on the parameter θ , i.e., $Y|\theta \sim \text{Bin}(n, \theta)$, where n is the size of our sample. In this setting, Y counts the number of “yes”, or “1”, observed in the sample, so that the sample mean ($\bar{y}$) is a natural estimate of the true population proportion θ . This model is often called the likelihood function, and it is a standard concept in both the Bayesian and the Classical framework. The Bayesian ¹ statistics combines both the prior distribution and the likelihood function to create a posterior distribution. The posterior distribution represents our opinion about which are the plausible values for θ adjusted after observing the sample data. In reality, the posterior distribution is one’s knowledge base updated using the latest survey information. For the prior and likelihood functions specified here, the posterior distribution is also a beta distribution ($\pi(\theta/y) \sim \beta(y+a, n-y+b)$), but with updated hyper-parameters.

Our credibility interval for ϑ is based on this posterior distribution. As mentioned above, these intervals represent our belief about which are the most plausible values for ϑ given our updated knowledge base. There are different ways to calculate these intervals based on $\pi(\theta/y)$. Since we want only one measure of precision for all variables in the survey, analogous to what is done within the Classical framework, we will compute the largest possible credibility interval for any observed sample. The worst case occurs when we assume that $a=1$ and $b=1$ and $y=n/2$. Using a simple approximation of the posterior by the normal distribution, the 95% credibility interval is given by, approximately:

$$\bar{y} \pm \frac{1}{\sqrt{n}}$$

For this poll, the Bayesian Credibility Interval was adjusted using standard weighting design effect $1+L=1.3$ to account for complex weighting²

Examples of credibility intervals for different base sizes are below. Ipsos does not publish data for base sizes (sample sizes) below 100.

Sample size	Credibility intervals
2,000	2.5
1,500	2.9
1,000	3.5
750	4.1
500	5.0
350	6.0
200	7.9
100	11.2

¹ *Bayesian Data Analysis, Second Edition, Andrew Gelman, John B. Carlin, Hal S. Stern, Donald B. Rubin, Chapman & Hall/CRC | ISBN: 158488388X | 2003*

² Kish, L. (1992). *Weighting for unequal Pi*. *Journal of Official, Statistics*, 8, 2, 183200.