

A mechanism for eliciting the mean and quantiles of a random variable

Thomas Demuyck*

July 11, 2013

Abstract

We present two Becker-DeGroot-Marschak type incentive compatible elicitation mechanisms. The first can be used to elicit an agent's belief about the mean of a random variable while the second elicits the quantiles.

JEL-classification: D81, C49

keywords: Becker-DeGroot-Marschak; elicitation; incentive compatible; mean; quantile

1 Introduction

This note deals with the truthful elicitation of an agent's belief about some features of a random variable. Although there exist a large number of different elicitation mechanisms, only two of them have the desirable property that they are incentive compatible irrespective of the agent's risk-preferences.

The first incentive compatible elicitation mechanism, proposed by Karni (2009) and Holt and Smith (2009), is based on an adaptation of the Becker-DeGroot-Marshak mechanism (BDM) (Becker, DeGroot, and Marschak, 1964). In order to elicit the subjective probability of a certain event, the mechanism compares the reported value of the agent's belief with a randomly drawn number (between zero and one). If the reported value is above this number then the agent obtains the reward if the random event takes place. If the reported value is below the random number, then the reward is given with probability equal to the value of this number. Recently, Qu (2012) extended this mechanism towards the elicitation of an entire probability distribution.

The second mechanism is the Binarized Scoring Rule (BSR) recently proposed by Hosain and Okui (2013) and Schlag and van der Weele (2013). The mechanism is based on

*Center for Economic Studies, University of Leuven, E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: thomas.demuyck@kuleuven-kortrijk.be. I would like to thank an anonymous referee for useful comments on an earlier version of this paper.

a modification of a proper scoring rule. A scoring rule measures the (ex-post) accuracy of a certain prediction. A scoring rule is proper if its expected value is minimized when the reported belief is equal to the actual belief. The usual implementation of a proper scoring rule is only incentive compatible when agents are risk neutral. The Binarized Scoring Rule modifies this method by binarizing the payoffs: the agent receives the reward if the realized prediction error (measured by the scoring rule) is smaller than an independently drawn random number. By maximizing the probability of winning the price, the agent is induced to report truthfully.

One of the main advantages of the Binarized Scoring Rule, compared to the BDM-type mechanism is its versatility. By varying the underlying proper scoring rule, the mechanism can be adjusted to elicit a various number of features of the random variable like the mean or quantiles. This versatility is lacking for the BDM-type mechanisms where it is necessary to provide a different procedure for each separate feature. In this note, we present two such procedures. The first elicits the mean of a random variable and the second elicits the quantiles.

2 Setting

We consider a setting where a principal would like to elicit the belief of an agent about a feature of a random variable. Let Θ be the random variable of interest. We assume that Θ has support on an interval $\mathcal{D} \subseteq \mathbb{R}$. The belief or knowledge of the agent about Θ is captured by the strictly increasing cumulative distribution function $F(\theta)$ ($\theta \in \mathcal{D}$). We assume that the principal can observe (at least) one realized value of Θ .

The binary lottery $\ell(p, x, y)$ ($x, y \in \mathbb{R}, p \in [0, 1]$) gives the agent an amount x with probability p and an amount y with probability $(1 - p)$. We assume that the agent has a transitive and complete preference relation \succeq over all binary lotteries which satisfies probabilistic sophistication and dominance in the sense that if $x > y$, then $\ell(p, x, y) \succeq \ell(q, x, y)$ if and only if $p \geq q$. From now on, we assume that $x > y$.

Eliciting the mean For this mechanism, we assume that the domain \mathcal{D} of the variable Θ is equal to the interval $[0, 1]$. If the support of Θ is finite, $\mathcal{D} = [\underline{\theta}, \bar{\theta}]$, we can always use the transformed variable $\frac{\Theta - \underline{\theta}}{\bar{\theta} - \underline{\theta}}$ which has support $[0, 1]$. The mean μ of the transformed variable is then related to the mean of the original variable $\tilde{\mu}$ by the transformation $\mu = \frac{\tilde{\mu} - \underline{\theta}}{\bar{\theta} - \underline{\theta}}$. The assumption that the support is bounded is crucial for obtaining an exact elicitation mechanism for the mean.¹ In the appendix, however, we show that the mechanism can also be used to elicit an ‘approximation’ of the mean in cases where the support is infinite.

¹The same actually holds for the Binarized scoring rule procedure. See Hossain and Okui (2013) for further discussion.

The mean of the variable Θ is given by:

$$\mu = \int_0^1 \theta dF(\theta).$$

Our elicitation mechanism is the following. First, the agent reports to the principal a value $m \in [0, 1]$. Next, the principal draws a random number r from a distribution with strict positive support on $[0, 1]$ (e.g. a uniform distribution). The payoff for the agent depends on values of m and r .

If $m \geq r$, then the principal waits for a realized value θ from the random variable Θ and independently draws a number z from the uniform distribution on $[0, 1]$. If $\theta > z$, the agent is given x . If $\theta \leq z$, the agent receives y .

If $m < r$, the agent is given the lottery $\ell(r, x, y)$ which gives x with probability r and y with probability $(1 - r)$.

In order to see that it is a weakly dominant strategy for the agent to truthfully report the value $m = \mu$, notice first that the payoff for the case $m \geq r$ coincides with a draw from the lottery $\ell(\mu, x, y)$. Indeed, the probability that the agent receives x is given by:

$$\begin{aligned} \int_0^1 \int_0^1 I(\theta > z) dz dF(\theta) &= \int_0^1 \left(\int_0^\theta dz \right) dF(\theta) \\ &= \int_0^1 \theta dF(\theta), \\ &= \mu \end{aligned}$$

Now let us show that it is a weakly dominant strategy for the agent to truthfully report the value $m = \mu$. First suppose that $m > \mu$. Then if $\mu \geq r$ or $r > m$, the agent receives the same lottery as when she would have reported μ . If $r \in]\mu, m]$, then the agent receives the lottery $\ell(\mu, x, y)$ but would have obtained the lottery $\ell(r, x, y)$ if she would have reported truthfully. Given that $r > \mu$, the agent is worse off by reporting m instead of μ .

If $\mu > m$ and $m \geq r$ or $r \geq \mu$, the agent receives the same lottery as if she would have reported μ . If $\mu > m$ and $r \in]m, \mu[$, then she receives the lottery $\ell(r, x, y)$ but would have received the preferred lottery $\ell(\mu, x, y)$ by reporting truthfully.

Eliciting the quantiles: Let us proceed by presenting another variation on the BDM mechanism that can be used to elicit the quantiles values of the random variable Θ . In this case, the domain \mathcal{D} of the random variable Θ is also allowed to be unbounded. Again we assume that the agent knows or has beliefs about F that is unknown to the principal. The problem of the principal is to elicit the $p\%$ -quantile θ_p :

$$F(\theta_p) = p.$$

Consider the following mechanism: The agent first reports a value $m_p \in \mathcal{D}$ to the principal. Next, the principal draws a random number r from a distribution which has

strict positive support on \mathcal{D} . Again, the payoff to the agent depends on the values of m_p and r .

If $m_p \geq r$, then the agent receives the lottery $\ell(p, x, y)$ which gives the payoff x with probability p and the payoff y with probability $(1 - p)$.

If $m_p < r$, then the principal waits for a realization θ of Θ . If $\theta \leq r$, the agent receives the payoff x . If $\theta \geq r$, the agent receives y .

First observe that when $m_p < r$, then the agent receives x with probability $F(r)$. As such, the agent actually faces the lottery $\ell(F(r), x, y)$.

Now, let us show that reporting $m_p = \theta_p$ is a weakly dominant strategy. First, if $m_p > \theta_p$ then $\theta_p \geq r$ or $r > m_p$ would have given the same lottery as reporting $m_p = \theta_p$. If $r \in]\theta_p, m_p]$, then the agent receives the lottery $\ell(p, x, y)$ but could have received the preferred lottery $\ell(F(r), x, y)$ (observe that $F(r) > F(\theta_p) = p$).

If $m_p < \theta_p$ and $r \leq m_p$ or $\theta_k \leq r$, then reporting truthfully would have given her the same lottery. However, if $m_p < \theta_p$ and $r \in]m_p, \theta_p[$, then by reporting m_p , she receives the lottery $\ell(F(r), x, y)$ while reporting truthfully would have given her the lottery $\ell(p, x, y)$. Given that $F(r) \leq F(\theta_p) = p$, the agent prefers to report θ_p .

Acknowledgements

I gratefully acknowledges the Fund for Scientific Research - Flanders (FWO-Vlaanderen) for my postdoctoral fellowship.

References

- Becker, G. M., DeGroot, M. H., Marschak, J., 1964. Measuring utility by a single-response sequential method. *Behavioral Science* 9, 226–232.
- Holt, C. A., Smith, A. M., 2009. An update on Bayesian updating. *Journal of Economic Behavior and Organization* 69, 125–134.
- Hossain, T., Okui, R., 2013. The binarized scoring rule. *The Review of Economic Studies* forthcoming.
- Karni, E., 2009. A mechanism for eliciting probabilities. *Econometrica* 77, 603–606.
- Qu, X., 2012. A mechanism for eliciting a probability distribution. *Economics Letters* 115, 399–400.
- Schlag, K. H., van der Weele, J. J., 2013. Eliciting probabilities, means, medians, variance and covariances without assuming risk neutrality. *Theoretical Economics Letters* 3, 38–42.

Appendix: Elicitation of the mean when the domain is unbounded

The elicitation mechanism for the mean used the assumption that the domain was bounded. In this appendix, we show that when the support is unbounded, then a slight variation of the algorithm can be used to elicit an approximation for the mean. We take the case where the support is equal to \mathbb{R} . A simple adaptation also allows for semi-infinite support.

In order to establish our result, we need the assumption that the tails of the distribution function are not too thick. In particular, we assume that there exists a $\delta > 0$ such that the distribution function of Θ satisfies $\lim_{\theta \rightarrow \infty} |\theta|^{2+\delta} f(\theta) = \lim_{\theta \rightarrow -\infty} |\theta|^{3+\delta} f(\theta) = 0$.²

Let $R > 0$ be a large number and consider the following mechanism. First the agent reports a value $m \in [-R, R]$. Next, the principal draws a random number r from a distribution with strict positive support on $[0, 1]$.

If $m \geq 2R(r - \frac{1}{2})$, the principal waits for a realized value of θ and independently draws a number z from a uniform distribution on $[-R, R]$. If $\theta > z$, the agent receives x . If $\theta \leq z$, the agent receives y . If $m < 2R(r - \frac{1}{2})$, the agent is given the lottery $\ell(r, x, y)$ which gives x with probability r and y with probability $(1 - r)$.

Let \tilde{p} be the probability of receiving x when $m \geq 2R(r - \frac{1}{2})$ (its exact value will be computed below). Using a similar argument as in the mechanism presented in the main text, we can show that it is a weakly dominant strategy for the agent to report $m = 2R(\tilde{p} - \frac{1}{2})$.

We show that for large values of R , this value approximates the true mean of Θ in the sense that $\lim_{R \rightarrow \infty} |m - \mu| = 0$. In order to establish our result, we first compute the probability \tilde{p} of receiving x if $m \geq 2R(r - \frac{1}{2})$.

$$\begin{aligned} \int_{-R}^R \int_{-R}^{\theta} \frac{1}{2R} dz dF(\theta) &= \int_{-R}^R \left(\frac{\theta}{2R} + \frac{1}{2} \right) dF(\theta) \\ &= \frac{\mu}{2R} - \frac{1}{2R} \left(\int_{-\infty}^{-R} \theta dF(\theta) + \int_R^{\infty} \theta dF(\theta) \right) + \frac{1}{2} \int_{-R}^R dF(\theta), \\ &= \frac{\mu}{2R} + \frac{1}{2} - \frac{1}{2R} \left(\int_{-\infty}^{-R} \theta dF(\theta) + \int_R^{\infty} \theta dF(\theta) \right) - \frac{1}{2} \int_{-\infty}^{-R} dF(\theta) - \frac{1}{2} \int_R^{\infty} dF(\theta), \\ &\equiv \tilde{p}. \end{aligned}$$

²This is the same assumption as the one imposed by Hossain and Okui (2013) to deal with situations where the scoring rule is unbounded.

This gives:

$$\begin{aligned}
|m - \mu| &= \left| 2R \left(\tilde{p} - \frac{1}{2} \right) - \mu \right|, \\
&= \left| \left(\int_{-\infty}^{-R} \theta f(\theta) d\theta + \int_R^{\infty} \theta f(\theta) d\theta \right) + R \int_{-\infty}^{-R} dF(\theta) + R \int_R^{\infty} dF(\theta) \right|, \\
&\leq \int_{-\infty}^{-R} |\theta| f(\theta) d\theta + \int_R^{\infty} |\theta| f(\theta) d\theta + R \int_{-\infty}^{-R} dF(\theta) + R \int_R^{\infty} dF(\theta).
\end{aligned}$$

Our assumption on the tails of the distribution implies that for all $\varepsilon > 0$, if θ becomes large (or small) enough, then $|\theta| f(\theta) \leq \varepsilon |\theta|^{-1-\delta}$ and $f(\theta) \leq \varepsilon |\theta|^{-2-\delta}$. As such,

$$\begin{aligned}
R \int_{-\infty}^{-R} f(\theta) d\theta + R \int_R^{\infty} f(\theta) d\theta &\leq \varepsilon R \int_{-\infty}^{-R} |\theta|^{-2-\delta} d\theta + R \int_R^{\infty} |\theta|^{-2-\delta} d\theta, \\
&= 2\varepsilon \frac{R^{-\delta}}{1-\delta}
\end{aligned}$$

This right hand side converges towards zero for R large enough. Also, for R large enough,

$$\begin{aligned}
\int_{-\infty}^{-R} |\theta| f(\theta) d\theta + \int_R^{\infty} |\theta| f(\theta) d\theta &\leq \varepsilon \left(\int_{-\infty}^{-R} |\theta|^{-1-\delta} d\theta + \int_R^{\infty} |\theta|^{-1-\delta} d\theta \right), \\
&= 2\varepsilon \frac{R^{-\delta}}{1+\delta}
\end{aligned}$$

Again, the right hand can again be made arbitrarily small by taking R large enough. As such, we see that $|m - \mu| \rightarrow 0$