GUESSING MODELS

Guessing models are statistical structures which aim to provide insight into the normative and optimal behaviors of people who must make choices in the apparent absence of data. Such models are related to aspects of *Bayesian* statistics* and the *Delphi method*[1], but they have a flavor and a theory of their own.

This is most easily illustrated by the anecdote of two statisticians, Bob and Mike, who engaged in a contest to guess the weights of people at a party. Bob agreed to always guess first, and on person 1 Bob guessed 142 pounds. Mike then guessed 142.1 pounds, and the subject declared Mike the winner. The contest continued; and, when final tallies were made, Bob found he had lost almost three-fourths of the time.

To model this scenario, consider a system of four p vectors

Target values:	$(\theta_1, \theta_2, \ldots, \theta_p) = \boldsymbol{\theta}$
First guess:	$(X_1, X_2, \ldots, X_p) = \mathbf{X}$
Second guesser's hunch:	$(Y_1, Y_2, \ldots, Y_p) = \mathbf{Y}$
Second guess:	$(G_1, G_2, \ldots, G_p) = \mathbf{G}$

The θ_i denote the real values to be guessed, so, for example, θ_2 would denote the weight of the second person considered by Bob and Mike. The X_i are the guesses made by the first guesser, and the Y_i represent the second guesser's best estimate of θ_i. Finally, the G_i are the guesses that are announced by the second guesser. The first problem in this theory is to determine how **G** should be determined by **X** and **Y**.

As the anecdote suggests, each player wishes to come closer to each θ_i than his opponent, so we consider

$$V(\mathbf{G}, \boldsymbol{\theta}) = \sum_{j=1}^{p} V_j(\mathbf{G}, \boldsymbol{\theta})$$

where

$$V_j(G, \theta) = \begin{cases} 1 & \text{if } |G_j - \theta_j| < |X_j - \theta_j| \\ 0 & \text{otherwise.} \end{cases}$$

With the objective of maximizing $V(\mathbf{G}, \boldsymbol{\theta})$, it is intuitive that the second guesser should always guess just a bit higher or a bit lower than the first guesser. This can be proved to be true under very general circumstances. Specifically, if $\boldsymbol{\theta}$, **X**, and **Y** are assumed to have a joint distribution that is continuous and if $\nu_i(\mathbf{X}, \mathbf{Y})$ denotes the conditional median of θ_i given **X** and **Y**, a key role is played by the strategies

$$G_i^{\epsilon} = \begin{cases} X_i + \epsilon & \text{if } X_i < \nu_i(\mathbf{X}, \mathbf{Y}) \\ X_i - \epsilon & \text{otherwise.} \end{cases}$$

These guesses $\mathbf{G}^{\epsilon}$ are called *Hotelling strategies* and the first result in the theory of guessing models is the following:

Theorem 1. The Hotelling strategies are ϵ optimal, i.e.,

$$\lim_{\epsilon \to 0} EV(\mathbf{G}^{\epsilon}, \boldsymbol{\theta}) = \sup_{\mathbf{G}} EV(\mathbf{G}, \boldsymbol{\theta}).$$

Although this result reassures intuitive feelings, it just makes the first step in telling the second guesser how to guess. Considerable ingenuity may be required to ferret out those models in which $\nu_i(\mathbf{X}, \mathbf{Y})$ can be calculated, and much of the theory of guessing strategies resides in the calculation of suitable approximations.

Consider, for example, the strategies

$$\tilde{G}_i = \begin{cases} X_i + \epsilon & \text{if } X_i < Y_i \\ X_i - \epsilon & \text{if } X_i > Y_i. \end{cases}$$

These are the "hunch" guided strategies and they model the reasonable actions of Mike in the anecdote. In some cases the hunch guided strategies are, in fact, Hotelling strategies, but even when these strategies are not ϵ optimal they have surprising power.

Theorem 2. If for each $1 \leqslant i \leqslant p$, X_i and Y_i are independent and identically distributed with a distribution that is symmetric about θ_i, then the hunch guided strategy $\tilde{G}_i$ has a $\frac{3}{4}$ probability of winning the ith contest as $\epsilon \to 0$.

The fact that one shrinks the first guess Y_i toward one's hunch X_i anticipates that the fact that other types of *shrinkers* are relevant in guessing models. In particular, when $p \geqslant 3$ and all the Y_i's are available to the second guesser, a very powerful strategy for the second guesser can be based on the *James–Stein estimator**.

NOTE

1. *Editor's note*: The Delphi method referred to above is the subject of a book [1], and is defined there as follows: "Delphi may be characterized as a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem." [1, p. 3]

(*See also* PUBLIC ADMINISTRATION, STATISTICS IN, *and* SOCIOLOGY, STATISTICS IN.)

Reference

[1] Linstone, H. A. and Turoff, M., eds. (1975). *The Delphi Method.* Addison-Wesley, Reading, Mass.

Bibliography

Hwang, J. T. and Zidek, J. V. (1981). *J. Appl. Prob.*, **19**, 321–331. (Investigates the strategy $G_i = X_i \pm \epsilon$ according as $X_i < \bar{X}$ or not.)

Pittinger, A. O. (1980). *J. Appl. Prob.*, **17**, 1133–1137. (Studies the problem of how the $\frac{3}{4}$ theorem extends to n-dimensional guessing.)

Steele, J. M. and Zidek, J. V. (1980). *J. Amer. Stat. Ass.*, **75**, 596–601. (Sets the foundation for the theory of second guessing, and provides the basis for this exposition. Contains much more information on Stein guided guessing as well as results of computer simulations.)

J. MICHAEL STEELE

GUMBEL DISTRIBUTION

The theory of statistical extremes has a short effective history. Beginning, essentially, with a paper by Dodd [10] on the distribution of the extremes (maxima and minima) of a univariate independent and identically distributed (i.i.d.) sample, the basic results regarding special properties and asymptotic behavior are contained in Fréchet [12], Fisher and Tippett [11], Gumbel [17], and von Mises [46], culminating in the fundamental paper by Gnedenko [14] (*see* EXTREME-VALUE DISTRIBUTIONS). The initial results concerning the law of large numbers* for extremes—not dealt with here—can be found in de Finetti [9]. The basic bibliography for statistical problems is still Gumbel [19]: many results and examples can be found in this fundamental reference. A large block of references can be found in Johnson and Kotz [22] and Harter [20]. A modern and essential reference for probabilistic results is Galambos [13]. Extensions and applications can be found at the end of the entry.

BASIC RESULTS

Consider a sample of k i.i.d. random variables $(Y_1, \ldots, Y_k)$ with cumulative distribution function (CDF) $F(x) = \Pr[Y \leqslant x]$. Then the CDF of $\max(Y_1, \ldots, Y_k)$ is

$$\Pr[\max(Y_1, \ldots, Y_k) \leqslant x] = \prod_1^k \Pr[X_i \leqslant x] = F^k(x);$$

in the same way,

$$\Pr[\min(Y_1, \ldots, Y_k) \leqslant x] = 1 - (1 - F(x))^k.$$

In general, to deal with samples of maxima (or minima), all obtained under the same conditions (with the same k and F), it would be necessary to know the form (and parameters) of F. But if k is large, we can try to use asymptotic distributions in statistical analysis. In fact, it can be shown that, in many cases, coefficients α_k and $\beta_k (> 0)$ exist such that $\Pr[(\max(X_1, \ldots, X_k) - \alpha_k)/\beta_k \leqslant x] = F^k(\alpha_k + \beta_k x)$ has a nondegenerate limit CDF; note that (α_k, β_k) are not uniquely defined. There are only three such limit forms (Weibull*, Gumbel, and Fréchet) which can be integrated in a condensed von Mises [46]–Jenkinson [21] form.