mean parameter in the conjugate normal model with the variance known is:

$$\mu \sim \mathcal{N}\left[\left(\frac{m}{s^2}+\frac{n\bar{x}}{\sigma_0^2}\right)\Big/\left(\frac{1}{s^2}+\frac{n}{\sigma_0^2}\right), \left(\frac{1}{s^2}+\frac{n}{\sigma_0^2}\right)^{-1}\right]. \tag{4.15}$$

If an improper prior for μ is specified as $\mathcal{N}(m = 0, s^2 = \infty)$, then the resulting posterior is clearly proper $\mathcal{N}(\bar{x}, \sigma_0^2/n)$. The compromise suggested is to make the prior variance (s^2) very large so that the prior is normal and proper but very spread out. This is often considered a "conservative" choice of prior since the relative probability structure is quite flat. These *diffuse priors* or *vague priors* are quite popular in hierarchical models where there are many regression-style parameters of only moderate interest.

The best rationale for improper priors, provided that there is appreciable substantive motivation for assigning them, is that if the model is set up so that the likelihood dominates the prior to such an extent that the posterior is still proper, then their use is not harmful. However, the only way to get an improper posterior is to specify an improper prior, and this decision rests entirely with the researcher.

4.5 Informative Prior Distributions

This section introduces several forms of informed (or informative) priors. In this chapter we looked closely at conjugate prior specifications, which are generally informed (although specifying ∞ for various parameters can dilute such qualities). Informative priors are those that deliberately insert information that researchers have at hand. On one level this seems like a reasonable and reasoned approach since previous scientific knowledge should play a role in statistical inference. The key concern that some readers, reviewers, and editors harbor is that the author is deliberately manipulating prior information to obtain a desired posterior result. Therefore there are two important requirements to any written research using informative priors: overt declaration of prior specifications, and detailed sensitivity analysis to show the effect of these priors relative to uninformed types. The latter requirement is the subject of Chapter 6, and the former requirement is discussed periodically in this section.

So where do informative priors come from? Generally there is an abundance of previous work in the social and behavioral sciences that can guide the researcher, including her own. So generally, these priors are derived from:

- ▷ previous studies, published work,
- ▷ researcher intuition,
- ▷ interviewing substantive experts,
- ▷ convenience through conjugacy,
- ▷ nonparametrics and other data derived sources,

which can obviously be overlapping definitions. Prior information from previous studies need not be in agreement. One fruitful strategy is to construct prior specifications from competing intellectual strains in order to contrast the resulting posteriors and say something informed about the relative strength of each. The last item on this list can be productive if the data used are distinct from that at hand to be used to construct the likelihood functions. There is considerable controversy, otherwise, about "double-use" of the data.

4.5.1 Power Priors

Ibrahim and Chen (2000a) introduce an informed prior that explicitly uses data from previous studies (also discussed by Ibrahim, Chen, and Sinha [2003] as well as Chen and Ibrahim [2006]). Their idea is to weight data from earlier work as input for the prior used in the current model. Define $\mathbf{x}_0$ as these older data and $\mathbf{x}$ as the current data. Their primary application is to clinical trials for AIDS drugs where a considerable amount of previous data exist. In a social science context, there are many settings where previous research informs extant model specifications. Our interest centers on the unknown parameter θ, which is studied in both periods. Specify a regular prior for θ, $p(\theta)$ that would have been used un-modified if the previous data were not included. This can be a diffuse prior if desired, although it will become informed through this process.

An elementary power prior is created by updating the regular prior with a likelihood function from the previous data in a very simple manner, which is scaled by a value $a_0 \in [0{:}1]$:

$$p(\theta|\mathbf{x}_0, a_0) \propto p(\theta)[L(\theta|\mathbf{x}_0)]^{a_0}. \tag{4.16}$$

It is important to remember that this is still a *prior* form and the regular process follows wherein the posterior is obtained by conditioning this distribution on the data through the likelihood function based on the current data:

$$\pi(\theta|\mathbf{x}, \mathbf{x}_0, a_0) \propto p(\theta|\mathbf{x}_0, a_0)L(\theta|\mathbf{x}). \tag{4.17}$$

The parameter a_0 scales our confidence in the similarity or applicability of the previous data for current inferences. If it is close to zero then we do not particularly value the older observations or studies, and if it is close to one then we believe strongly in the ties to the current data. So lower values favor the regular prior specification, $p(\theta)$, and the choice of this parameter can be very influential. Note that we would not want this value to exceed one, since that would be equivalent to valuing older over newer data.

To reduce the influence of a single choice for a_0, we specify a mixture of these priors using a specified distribution for this parameter, $p(a_0|.)$. Thus (4.16) becomes

$$\begin{aligned} p(\theta|\mathbf{x}_0) &= \int_0^1 p(\theta|\mathbf{x}_0, a_0)p(a_0|\cdot)da_0 \\ &= \int_0^1 p(\theta)[L(\theta|\mathbf{x}_0)]^{a_0} p(a_0|\cdot)da_0. \end{aligned} \tag{4.18}$$

Here the parameterization for $p(a_0|\cdot)$ is left vague since its parametric form remains undefined. Chen and Ibrahim recommend a beta distribution as the "natural" choice, but point out that truncated forms of the normal or gamma work as well. The mixture specification has the effect of inducing heavier tails in the marginal distribution of θ and thus represents a more conservative choice of prior.

4.5.2 Elicited Priors

A completely different class of priors is derived not from real or desired mathematical properties, but from previous human knowledge on the subject of investigation. These elicited priors are discussed in detail in Gill and Walker (2005), with a detailed application to attitudes towards the judicial system in Nicaragua. Typically the source for elicited priors is from subject area experts with little or no concern for the statistical aspects of the project. These include physicians, policy-makers, theoretical economists, and qualitative researchers in various fields. However, there is no reason that politicians, study participants, outside experts, or opinion leaders in general could not be used as a source for informative priors as well.

The bulk of the published work on elicited priors is on the Bayesian analysis of clinical trials. In these settings, it is typical to elicit qualitative priors from the clinicians as a means of incorporating local expertise into the calculations of posteriors and trial stopping points (Freedman and Spiegelhalter 1983, Kadane 1986, Spiegelhalter, Abrams, and Myles 2004, Chapter 5). There is also a small literature on elicitation of priors for variable selection (Garthwaite and Dickey 1988, 1992; Ibrahim and Chen 2000b). Here we will concentrate on the more basic task of using elicitation to specify a particular parametric form for the prior. It is relatively common to use conjugate priors or mixtures of conjugate priors for this task so as to remove additional complications. However, this is certainly not a mathematical or theoretical restriction.

Although an overwhelming proportion of the studies employing elicited priors are in the medical and biological sciences, the methodology is ideal for a wide range of social science applications. In virtually every field and subfield of the various disciplines there are practicing "experts" whose opinions can be directly or indirectly elicited. Furthermore, the fact that the social and behavioral sciences are focused on varying aspects of human behavior means that describing current knowledge and thinking about some specific behavioral phenomenon probabilistically is a more realistic way to incorporate disparate judgments. Restated, uncertain and divided opinion is better summarized in probabilistic language than with deterministic alternatives.

The central challenge here is how to translate expert knowledge into a specific probability statement. This process ranges from informal assignments to detailed elicitation plans and even regression analysis across multiple experts (Johnson and Albert 1999, Chapter 5). Spetzler and Staël von Holstein (1975) outline three general steps in the process:

1. **Deterministic Phase.** The problem is codified and operationalized into specific variables and definitions.

2. **Probabilistic Phase.** Experts are interviewed and tested in order to assign probabilistic values to specific outcomes.

3. **Informational Phase.** The assigned probabilities are tested for inconsistencies and completeness is verified.

The deterministic phase includes specifying the explanatory variables and possibly their assumed parametric role in the model (Steffey 1992), determining data sources and data collection processes fitted to this methodology (Garthwaite and Dickey 1992), determining how many experts to query and where to find these experts (Carlin, *et al.* 1993), and finally judging their contributions (Hogarth 1975). Some of this work is far from trivial: experts might need to be trained prior to elicitation (Winkler 1967), variable selection can be influenced by the difficulty of elicitation (Garthwaite and Dickey 1992), and cost projections can be difficult.

The informational phase is somewhat mechanical and it includes testing elicitation responses for consistency, calibrating responses with known data, and perhaps weighting expert opinions. Determining consistency is an important requirement and experts differ in their familiarity with the details of the project at hand. Less experienced respondents tend to show more inconsistencies (especially with continuous rather than discrete choices), and more experienced respondents as well as normative experts show high levels of consistency (Hogarth 1975, Winkler 1967). By consistency it is meant that answers do not contradict each other, for instance, the subset of an event having a higher probability than the event itself. Calibration generally involves comparison of results after the rest of the analysis and can be a safety check for future work as well as a confirmation of the reliability of the experts (Seidenfeld 1985). Sometimes these checks are further complicated when the subject is a rapidly changing area and the experts' earlier statements can quickly become outdated (Carlin, *et al.* 1995). Leamer (1992) also gives a diagnostic approach that helps categorize elicitations into blunt responses removing the necessity of further inquiry.

By far the most challenging is the probabilistic phase and this has consumed the bulk of the literature. For instance, the experts can be asked fixed value ("P-methods") and/or fixed probability ("V-methods") questions where specific estimates of the probability or relative likelihood of events are queried (Spetzler and Staël von Holstein 1975, p.347). In addition, the experts can be asked these questions directly with regard to a cumulative density function (CDF), or indirectly by way of physical devices or hypothetical constructions. A more challenging, but perhaps informative, approach is to ask open-ended questions and code the response. In all of these cases, it is important to clarify to experts that they are giving probability estimates rather than utility assessments (Kadane and Winkler 1988). The concern is that these experts will otherwise express their normative ideals about outcomes, and preferred outcomes will be given unrealistic probabilities.

In general it is not feasible to ask subject-matter experts to make determinations about coefficient estimates or about moments for specified PDFs and PMFs convenient to the statistician. So the most common strategy, dating back to the seminal paper of Kadane *et al.* (1980), is to query these experts about outcome variable quantiles for given (hypothetical) levels of explanatory variables. For example, in one study an emergency room physician is asked about survival probabilities of patients with specified injury type, injury severity score, trauma score, age, and type of injury (Bedrick, Christensen, and Johnson 1997). The idea is then to take these quantiles and solve for the parameters of an assumed distributional form for the (often conjugate) prior.

One particularly simple application is in the case of a binomial outcome. For psychological reasons, it appears to be easier to elicit hypothetical binary outcomes. Using the beta conjugate prior several authors have suggested algorithms for elicitation (Chaloner and Duncan 1983, 1987; Gavasakar 1988). The basic process is to hypothesize a fixed set of Bernoulli trials, ask the expert to give a most likely number of successes given the particular scenario and reasonable bounds on the uncertainty, work these values backward into the beta-binomial PMF to get the beta parameters, and finally show the expert the posterior implications of these values. If they are found to be unreasonable, then adjustments are made and the process repeats itself.

4.5.2.1 The Community of Elicited Priors

The priors that are elicited from experts can have a variety of characterizations. Kass and Greenhouse (1989) coined the phrase "community of priors" to describe the range of attitudes that equally qualified experts may have about the same phenomenon. These can be categorized as well:

- ▷ **Clinical Priors.** These are priors elicited from substantive experts who are taking part in the research project. This is often done because these individuals are easily captured for interviews and are motivated by a direct stake in the outcome.

- ▷ **Skeptical Priors.** These are priors built with the assumption that the hypothesized effect does not actually exist and are usually operationalized with a zero mean. Skeptical priors can be created because of actual skepticism or because overcoming such a prior provides stronger evidence: "...set up as representing an adversary who will need to be disillusioned by the data ..." (Spiegelhalter *et al.* 1994, Spiegelhalter, Abrams, and Myles 2004).

- ▷ **Enthusiastic Priors.** These are obviously the opposite of the skeptical prior. The priors are built around the positions of partisan experts or advocates and generally assume the existence of the hypothesized effect. For comparative purposes, enthusiastic priors can be specified with the same variance, but different mean, as corresponding skeptical priors.

▷ **Reference Priors.** Such priors are occasionally produced from expert sources, but they are somewhat misguided because the purpose of elicitation is to glean information that can be described formally.

The priors are restated from Spiegelhalter *et al.* (1994) in order to be less focused on the application to clinical trials. The key point is to understand the differing perspectives of experts. One approach is to contrast the posterior results obtained from divergent prior perspectives, including a formalized process of overcoming adversarial prior specifications in favor of priors more sympathetic to research questions through additional sampling (Lindley and Singpurwalla 1991), randomization strategies (Kass and Greenhouse 1989), scoring rules (Savage 1971), or other means.

4.5.2.2 Simple Elicitation Using Linear Regression

An analyst asks an expert for predictions on an expected outcome for some interval-measured event of interest. The V-method question asked is: what would be an expected low value as a 0.25 quantile (labeled $x_{0.25}$) and an expected high value as a 0.75 quantile (labeled $x_{0.75}$)? These two supplied quantile values, $x_{0.25}$ and $x_{0.75}$, correspond to normal z-scores $z_{0.25} = -0.6745$ and $z_{0.75} = 0.6745$, which specify the shape of a normal PDF since there are two equations and two unknowns:

$$z_{0.25} = \frac{x_{0.25} - \alpha}{\beta} \qquad\qquad z_{0.75} = \frac{x_{0.75} - \alpha}{\beta}. \tag{4.19}$$

Here α and β are the mean and standard deviation parameters of the normal PDF:

$$f(x|\alpha, \beta) = (2\pi\beta^2)^{-\frac{1}{2}} \exp\left[-\frac{1}{2\beta^2}(x - \alpha)^2\right]. \tag{4.20}$$

This notation for the parameters of a normal is different, but the reason shall soon become apparent. When we solve for α and β in (4.19), we have a fully defined prior distribution in (4.20) and the elicitation is complete.

Of course, one expert is typically not enough to produce robust prior forms, so now query experts $1, 2, \ldots, J$. This produces an over-specified series of equations since there are $J \times 2$ equations and only two unknowns (Spiegelhalter *et al.* [1994], for instance, use $J = 10$). It is necessary to assume that these experts are *exchangeable* meaning that they all provide equal quality elicitations.

Secondly, given the cost of interviewing, we are likely to ask each expert for more than just these two quantiles. Each assessor is asked to give five quantile values at $m = [0.01, 0.25, 0.5, 0.75, 0.99]$ corresponding to standard normal points z_m. At this point, (4.19) can be re-expressed for the quantile level m given by assessor j: $x_{jm} = \alpha + \beta z_{jm}$, and the total amount of expert-elicited information constitutes the following over-specification

($J \times 5$ equations and 2 unknowns) of a normal distribution:

$$
\begin{array}{llll}
x_{11} = \alpha + \beta z_{11} & x_{21} = \alpha + \beta z_{21} \ldots & x_{(J-1)1} = \alpha + \beta z_{(J-1)1} & x_{J1} = \alpha + \beta z_{J1} \\
x_{12} = \alpha + \beta z_{12} & x_{22} = \alpha + \beta z_{22} \ldots & x_{(J-1)2} = \alpha + \beta z_{(J-1)2} & x_{J2} = \alpha + \beta z_{J2} \\
x_{13} = \alpha + \beta z_{13} & x_{23} = \alpha + \beta z_{23} \ldots & x_{(J-1)3} = \alpha + \beta z_{(J-1)3} & x_{J3} = \alpha + \beta z_{J3} \\
x_{14} = \alpha + \beta z_{14} & x_{24} = \alpha + \beta z_{24} \ldots & x_{(J-1)4} = \alpha + \beta z_{(J-1)4} & x_{J4} = \alpha + \beta z_{J4} \\
x_{15} = \alpha + \beta z_{15} & x_{25} = \alpha + \beta z_{25} \ldots & x_{(J-1)5} = \alpha + \beta z_{(J-1)5} & x_{J5} = \alpha + \beta z_{J5}
\end{array}
$$

The approach suggested by this setup is to run a simple linear regression to estimate α as the intercept and β as the slope. There are two issues to worry about. One must check for logical inconsistencies in consistent quantile values for each assessor (see Lindley, Tversky, and Brown [1979] for a discussion of problems). Also, it is critical to apply necessary mathematical constraints such as ensuring that the estimated coefficient for β remains positive (if substantively required) since the basic linear model imposes no such restriction (Raiffa and Schlaifer 1961).

■ **Example 4.1: Eliciting Expected Campaign Spending.** We are interested in eliciting a prior distribution for expected campaign contributions received by major-party candidates in an impending U.S. Senate election in order to specify an encompassing Bayesian model. Elicitation replaces data here since the election has not yet taken place. Eight campaign experts are queried for quantiles at levels $m = [0.1, 0.5, 0.9]$, and they provide the following values reflecting the national range of expected total intake by Senate candidates (in thousands):

$$
\begin{array}{lll}
x_{11} = 400 & x_{12} = 2500 & x_{13} = 4000 \\
x_{21} = 150 & x_{22} = 1000 & x_{23} = 2500 \\
x_{31} = 300 & x_{32} = 900 & x_{33} = 1800 \\
x_{41} = 250 & x_{42} = 1200 & x_{43} = 2000 \\
x_{51} = 450 & x_{52} = 1800 & x_{53} = 3000 \\
x_{61} = 100 & x_{62} = 1000 & x_{63} = 2500 \\
x_{71} = 500 & x_{72} = 2100 & x_{73} = 4200 \\
x_{81} = 300 & x_{82} = 1200 & x_{83} = 2000
\end{array}
$$

where x_{83} is expert 8's third quantile. None of the experts have supplied quantile values out of logical order, so these results are consistent. Using these "data" we regress x on z to obtain the intercept and slope values:

```
x <- c( 400, 2500, 4000,  150, 1000, 2500,  300,  900,
       1800,  250, 1200, 2000,  450, 1800, 3000,  100,
       1000, 2500,  500, 2100, 4200,  300, 1200 ,2000)
z <- qnorm(rep(c(0.1,0.5,0.9),8))
summary(lm(x~z))
```

which returns:

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1506.3      127.0  11.858 4.99e-11
qnorm(y)       953.4      121.4   7.854 8.00e-08
```

Thus the normal prior mean is $\hat{\alpha} = 1506.3$ and the normal prior standard deviation is $\hat{\beta} = 953.4$.

4.5.2.3 Variance Components Elicitation

A problem with direct quantile elicitation is that assessors often misjudge the probability of unusual values because it is more difficult to visualize and estimate tail behavior than to estimate means or medians. Hora *et al.* (1992) found that when non-technical assessors are asked to estimate spread by providing high probability coverage intervals such as at 99%, then they tend to perceive this as near-certainty coverage and overstate the bounds. But this finding is not universal: in other settings people tend to think of rare events in the tails of distributions as more likely than they really are (an effect exploited by casinos and lotteries). Accordingly, O'Hagan (1998) improves elicited estimates of spread by separately requiring assessors to consider two types of uncertainty:

▷ uncertainty about an estimate relative to an assumed known summary statistic

▷ secondary uncertainty of this summary only.

The elicitee first gives a (modal) point estimate for the explanatory variable coefficient, τ, and is then asked "given your recent estimate of τ, what is the middle 50% probability interval around τ?" The elicitees must understand that this is the interval that contains the middle half of the expected values. So this (V-method) specifies a density estimate centered at the assessors modal point, and if the form of the distribution is assumed or known, then the exact value for the variance can be calculated under a distributional assumption (normal, Student's-t, or a log-normal if a right-skewed interval is required).

O'Hagan (1998) prefers asking for the middle 66% of the density, which he calls the "two-to-one interval" since the middle coverage is twice that of the combined tails. Now if a normal prior is used then this interval quickly yields a value for the standard deviation since it covers approximately two of them. It should actually be multiplied by $\frac{68}{66}$ to be exactly correct but analysts often do not worry about the difference since measurement error is almost certainly greater than the difference. Now the researcher calculates the implied variance from this and shows the assessor credible intervals at familiar $(1-\alpha)$-levels. If these are deemed by the elicitee to be too large or too small, then the process is repeated.

We want to elicit prior distributions for τ_i across n cases, with unknown total $T = \sum_{i=1}^{n} \tau_i$. The assessor first provides point estimates for each case: $x_1, x_1, \ldots, x_n$, so that the estimated total is given by $x_T = \sum_{i=1}^{n} x_i$. These are useful values, but it is still necessary to get a measure of uncertainty in order to produce a variance for the full elicited prior distribution.

The individual deviance of the ith estimate from its true value can be rewritten algebraically:

$$\tau_i - x_i = \left(\tau_i - \frac{x_i}{x_T}T\right) + \frac{x_i}{x_T}\left(T - x_T\right). \tag{4.21}$$

The first quantity on the right-hand side of (4.21) is the deviance of τ_i from an estimate that would be provided if we knew T for certain:

$$E[\tau_i|T] = \frac{x_i}{x_T}T \tag{4.22}$$

(which can be considered as between-case deviance). Now the second quantity on the right-hand side of (4.21) is the weighted deviation of T, i.e., uncertainty about the true total. The expected value form (4.22) helps us obtain the variance of τ_i:

$$\begin{aligned}
\text{Var}(\tau_i) &= E\left[\text{Var}(\tau_i|T)\right] + \text{Var}\left(E[\tau_i|T]\right) \\
&= E\left[\tau_i - \frac{x_i}{x_T}T\right]^2 + \left(E\left[E[\tau_i|T]^2\right] - \left(E\left[E[\tau_i|T]\right]\right)^2\right) \\
&= E\left[\tau_i - \frac{x_i}{x_T}T\right]^2 + \left(\frac{x_i}{x_T}\right)^2 \text{Var}(T),
\end{aligned} \tag{4.23}$$

which shows the general form of the two variance components. A better form for elicitation is achieved by dividing both sides of this equation by x_i^2:

$$\text{Var}\left(\frac{\tau_i}{x_i}\right) = \text{Var}\left(\frac{\tau_i}{x_i} - \frac{T}{x_T}\right) + \text{Var}\left(\frac{T}{x_T}\right). \tag{4.24}$$

At this point elicitees are queried about the middle spread around these two quantities individually. *First*, they give an estimate of middle spread around $\frac{T}{x_T}$, assuming accuracy of the sum x_T as an estimate of T. *Then*, they give the middle spread around each $\frac{\tau_i}{x_i}$ assuming that $\frac{T}{x_T} = 1$. This means that there is no second component in the variance to consider at this moment. Once the individual means and variances are elicited, these $\frac{x_i}{x_T}$ values are put into the assumed distribution defined over $[0\!:\!1]$ (they are proportions) to form the complete prior distribution specification. Two common distributional forms are the normal CDF and the beta distribution. In the case where $\frac{\tau_i}{T}$ is from a beta distribution, we can solve for the parameters with the beta distribution mean and variance: $E\left[\frac{\tau_i}{T}\right] = \frac{\alpha}{\alpha+\beta}$, $\text{Var}\left(\frac{\tau_i}{T}\right) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

■ **Example 4.2: Minority Political Participation.** This example is from Gill and Walker (2005). An expert on minority electoral participation is asked to estimate upcoming Hispanic turnout for n precincts in a given district: $\tau_1, \tau_2, \ldots, \tau_n$, with total

Hispanic turnout in the district equal to T. She first gives estimates $x_1, x_2, \ldots, x_n$ for each precinct, which give a district turnout estimate of T by summing, x_T. This result does not yet give the variance information necessary to build a prior using an assumed normal distribution.

The expert is now asked to provide the two-to-one interval for $\frac{T}{x_T}$, giving $[0.7\!:\!1.3]$: they believe that the summed estimate of Hispanic turnout is correct to plus or minus 30% with probability 0.66 (from the two-to-one interval). To confirm the expert's certainty about this, the value $\sigma_T = 0.3$ is plugged into the normal CDF at levels to give credible interval summaries:

$$50\%\, CI = [\Phi_{\mu=1,\sigma=0.3}(0.25)\!:\!\Phi_{\mu=1,\sigma=0.3}(0.75)] = [0.798\!:\!1.202]$$

$$99\%\, CI = [\Phi_{\mu=1,\sigma=0.3}(0.005)\!:\!\Phi_{\mu=1,\sigma=0.3}(0.995)] = [0.227\!:\!1.773].$$

These are then displayed to the elicitee, and if she agrees that these are reasonable summaries then the variance is $\sigma_T^2 = (0.3)^2 = 0.09$ and there is no need to iterate here. The expert is now asked to repeat this process for each of the x_i estimates under the assumption that $x_T = T$ (the estimate of the total above is correct). This temporary fixing of x_T means that the right-hand-side of (4.24) reduces to the variance of $\frac{\tau_i}{x_i}$ and the expert can do the same interval process as was done with $\frac{T}{x_T}$ for each of the n precincts.

Suppose that two-to-one interval for the estimate of Hispanic turnout at the first precinct ($x_1 = 0.2$) is given as $[0.5\!:\!1.5]$: she believes the estimate to be correct to plus or minus 50% with probability 0.66. This gives a variance of $\sigma_1 = (0.5)^2 = 0.25$, and we will note that the subsequent 50% and 99% credible interval summaries are approved by the expert. So the total elicited variance for the first precinct is given by (4.24) where x_1^2 is moved back to the right-hand-side: $\text{Var}(\tau_1) = x_1^2(0.25 + 0.09) = 0.0136$.

4.5.2.4 Predictive Modal Elicitation

If the outcome variable of interest is distributed Bernoulli or binomial, then it is usually straightforward to query experts *directly* for prior probabilities. For psychological reasons probabilities of binary or summed binary outcomes are more intuitively easy to visualize (Cosmides and Tooby 1996). Using the natural (conjugate) choice of a beta conjugate prior Chaloner and Duncan (1983, 1987) produce the predictive modal (PM) elicitation algorithm (see also Gavasakar [1988] for a second application). First fix a hypothetical total number of Bernoulli trials, and then ask the elicitee to specify the most likely number of successes as well as reasonable bounds on the uncertainty. These values are then worked backward into the beta-binomial parametric setup (Chapter 1) to get the beta prior distribution parameters. The elicitee is now shown the implications of their stipulated values on the shape of the beta prior on a computer terminal or with pre-prepared flip-charts. If the