Displaying Uncertainty With Shading

Christopher H. JACKSON

A new technique is presented for illustrating several probability distributions on the same axes. The *density strip* is a shaded monochrome strip whose darkness at a point is proportional to the probability density of the quantity at that point. These are ideal for comparing distributions arising from parameter estimation, such as posterior distributions from Bayesian multiple regression or meta-analysis. Such distributions are more commonly illustrated as a point and line representing point and interval estimates. This may give the false perception that all points within the line are equally likely, and that points outside the line are impossible. The density strip represents the entire distribution in one dimension, giving a fuller description of the uncertainty surrounding the quantity. The strips fade gradually to white in the tails of a typical distribution, in contrast with line plots and strips whose thickness is proportional to the density, which terminate at a clear limit. This discourages casual significance testing based on comparing an arbitrary point in the tail of the distribution to a threshold. Shaded strips can also be generalized to shaded regions, which illustrate the uncertainty surrounding a continuously varying unknown quantity, such as a survival curve or a forecast from a time series.

KEY WORDS: Color; Distribution; Forecasting; Density; Meta-analysis; Multiple regression.

1. DISPLAYING UNCERTAINTY IN ONE DIMENSION

Suppose we want to illustrate quantities which are subject to uncertainty, such as parameter estimates or predictions from a model. A probability distribution has been obtained to represent the uncertainty surrounding each quantity, for example, a Bayesian posterior or predictive distribution. The distribution may be known analytically, or estimated from a bootstrap or Markov chain Monte Carlo sample. A distribution is usually illustrated by plotting its probability density. For example, Figure 1 displays the density of a standard normal distribution, a skewed distribution (standard exponential), and a bimodal distribution (mixture of two normals). However, to compare several distributions side by side, one-dimensional representations may be clearer. Most commonly, these consist of a point estimate and a line covering a symmetric 95% probability region (Figure 1(a)). The coefficients of multiple regression models are commonly presented in this way (see Figure 2, top left). Similar illustrations are used in meta-analysis for comparing estimated effects from different studies (the *forest plot*, see Figure 3, top left) or in clinical trials for displaying subgroup-specific treatment effects or interactions.

Christopher H. Jackson is Research Associate, MRC Biostatistics Unit, Institute of Public Health, Forvie Site, Robinson Way, Cambridge, U.K. (E-mail: *chris.jackson@mrc-bsu.cam.ac.uk*). The author thanks David Spiegelhalter, Nikos Demiris, Simon Thompson, Ian White, Daniela DeAngelis and the editor for helpful ideas and encouragement.

1.1 Illustrating the Whole Distribution

This common presentation of confidence or credible intervals has been criticized for giving the perception that the data support all points within the interval equally, which is usually false. Louis and Zeger (2008) suggested that when presenting results in text, interval estimates are indicated in smaller type than the point estimate, for example: $0.98_{(0.91,1.04)}$. For graphical presentation, several ways have been proposed to convey more information about the shape of the underlying distribution. For example, the interquartile range can be shown as a box (Figure 1(b)), in the manner of the boxplot (Tukey 1977) commonly used for summarizing data. This may be elaborated to show additional pairs of quantiles with additional boxes or tick marks. This principle was extended by Esty and Banfield (2003) to illustrate the entire empirical distribution of data as a *box-percentile plot*. This is a strip whose width at a point is proportional to the probability of a more extreme point (Figure 1(c), median and interquartile range indicated by vertical lines). This can illustrate an entire distribution in one dimension. However, its meaning is not always intuitively clear, for example, it does not indicate the peaks of the bimodal distribution in Figure 1.

This problem can be addressed by defining the width or thickness of the strip as proportional to the *probability density*. Lee and Tu (1997) described variable-width plots for summarizing data using a density estimate placed next to its reflection, and Hintze and Nelson (1998) described the similar *violin plot* which superimposes such a graph on a box plot. The same principle can be used to illustrate parameter estimates, for example, Spiegelhalter (1999) presented *eyeball plots* of Bayesian posterior distributions, and Barrowman and Myers (2003) proposed the similar *raindrop plot* for illustrating likelihood-based confi-

 DOI: 10.1198/000313008X370843

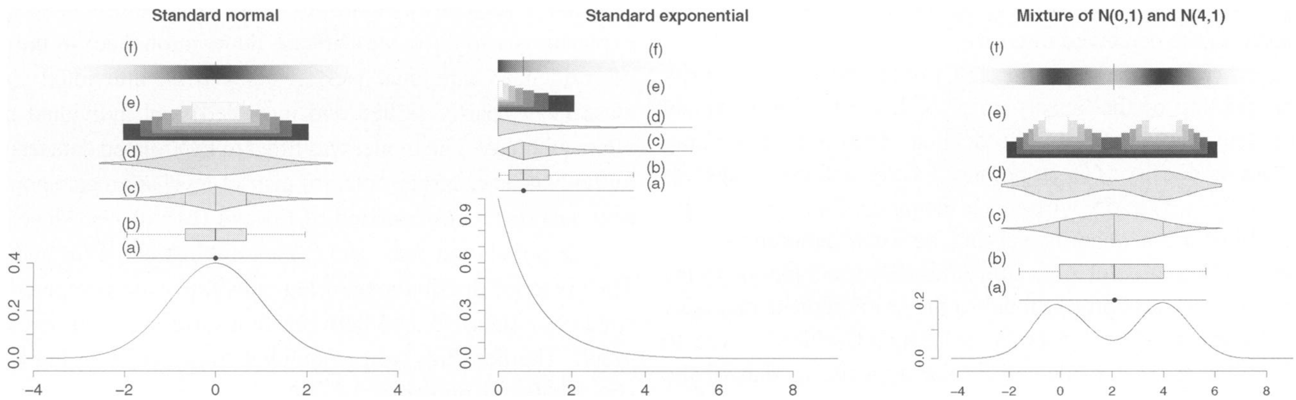

Figure 1. Density of a standard normal distribution, standard exponential distribution and a bimodal normal mixture, with six compact representations of each distribution.

dence regions for parameters. Figure 1(d) shows an illustration of this kind for three distributions. Since it bulges in regions of higher probability density, it can represent the two modes of the mixture distribution. In the tails of the distribution, its width decays to the width of a printed line. These approaches enable a substantially greater *data density* (Tufte 1983) than point-and-line drawings, conveying more information per unit of space and unit of ink.

1.2 Casual Significance Testing

Confidence or credible intervals for parameters also suggest a duality between points inside and outside the interval. Regression coefficients are often classified as "significant" according to whether their 95% interval estimates exclude the null effect. The casual use of hypothesis tests based on arbitrary thresholds is frequently criticized (Gelman and Stern 2006), particularly in medical research (Sterne et al. 2001). Replacing point-and-line illustrations of interval estimates with varying-width strips removes the perception that all points in the interval have equal probability. However, even these must terminate with a line in the tail of the distribution, where the density decreases towards but never reaches zero. The end of a pixel-width line is easily visible on even the highest-resolution of current displays. If used for summarizing the empirical distribution of data, it is reasonable for the strip to terminate at the sample minimum and maximum (as in the violin plots of Hintze and Nelson 1998). But if used for illustrating a distribution of a parameter estimate with unbounded support, an arbitrary end point must be chosen. In Figure 1(c,d), the strips terminate at the 0.5% and 99.5% quantiles of the distribution.

2. DENSITY STRIPS

An alternative to varying width for illustrating probability density is by intensity of color or shading. Cohen and Cohen (2006) combined shading with occlusion to illustrate empirical distributions of data as *sectioned density plots*. The estimated density is categorized into ranges, and regions of higher density are displayed as offset rectangles with lighter shading. This technique is shown for the three known distributions in Figure 1(e).

A smooth and more space-efficient alternative is the *density strip*. This is defined as a thin horizontal rectangle which is darkest at the point of highest probability density, white at points of zero density, and shaded with darkness proportional to the density (Figure 1(f)). Unlike varying-width strips and sectioned density plots, these represent a two-dimensional density plot using one physical dimension, with no loss of information. The second "virtual" dimension is defined by shading. The second physical dimension, the thickness of the strip, contains no information. Thus, large numbers of distributions with widely varying maximum densities may be compared using efficiently stacked strips, each occupying the same amount of vertical space. As described in Section 4, infinitesimally thin strips which merge into each other may even be used to illustrate the uncertainty around a continuously changing quantity. Density strips do not terminate at a clear limit as the density reduces to zero. Since they fade gradually to white, they illustrate that any choice of threshold in the tail, beyond which points are deemed "extreme," is arbitrary.

2.1 Definition of the Shading

In computer graphics, colors are conventionally represented using the red-green-blue (RGB) model, which leads to an intuitive definition for the shading of the density strip. In modern displays, $2^8 = 256$ distinct levels of each of red, green, and blue are available to color each pixel. The intensities of red, green, and blue are indicated on a decimal scale from 0 to 255, or commonly to programmers, in hexadecimal notation as `00` to `FF`. Black is represented by (0,0,0) or `000000`, and white is (255,255,255) or `FFFFFF`. Shades of gray have equal levels of red, green, and blue: a gray "halfway" between black and white is represented by (128, 128, 128) or `808080`. There is a nonlinear relationship between perceived and actual color intensity, and visual displays perform *gamma correction* to transform this into a linear relationship (Poynton 1993, 1998), so that the RGB color value specified by the programmer becomes directly proportional to the perceived intensity.

Therefore, the darkness, or 255 minus the RGB intensity of the shading of the density strip, is defined as proportional to the density. Specifically, the decimal definition of the gray level for a density $f()$ at point x is the nearest integer to $(1-f(x)/f(x_0))\times 255$, where x_0 is the mode. *Choropleth maps*, which illustrate a quantity varying between geographical areas by shading or color, also conventionally use a monochrome scale with darkness proportional to the value represented (see, e.g., Brewer 1994). This color ordering of white at zero to black at the maximum density is the opposite to that of the sectioned density plots presented by Cohen and Cohen (2006) (e.g., Figure 1(e)). Density strips do not exploit perspective, and since most printed graphics are displayed on white backgrounds, white suggests absence of information.

Alternatively, if color display or printing is available, the decimal color at x can be chosen as $p \times (c_R, c_G, c_B) + (1-p) \times (255, 255, 255)$, where $p = f(x)/f(x_0)$ varies between 0 and 1, and (c_R, c_G, c_B) is some dark color chosen for the maximum density. Thus the color intensity varies with the density—variations in intensity can better represent numeric quantities than variations in color hue or saturation (Tufte 1983; Cleveland 1993; Brewer 1994). If desired, the proportional mapping of density to darkness could be replaced by another monotone transformation. In particular, replacing p by p^γ ($\gamma > 0$) modifies the "gamma correction," for example, setting $\gamma < 1$ will darken the tails of the distribution, and $\gamma > 1$ will shorten the black area around the peak.

2.2 Examples

Figure 1(f) shows density strips for the three standard distributions. For the normal distribution, the strip is indistinguishable from black over about half the interquartile range. Beyond about three standard deviations from the mean, at which the probability of a more extreme value is about 0.001, the shade becomes indistinguishable from white. For the exponential distribution with mean 1, the darkest area of the strip is close to the mode of zero, in contrast to the interquartile range in (b) and (c) which surrounds the median of 0.69. The shading correctly represents the twin peaks of the mixture distribution. While density strips are intended to illustrate that parameter uncertainty is smoothly varying, if particular points of interest in the distribution are still required, the strip can be supplemented with tick-marks. In Figure 1(f), the medians are indicated as thin ticks, but the conventional pairs of outer quantiles are omitted.

3. APPLICATIONS OF DENSITY STRIPS

3.1 Multiple Regression

In Figure 2, a published graph of point-and-interval estimates (Jackson et al. 2008) is redrawn using density strips and varying-width strips. This illustrated a Bayesian multiple regression model for the risk of hospital admission for cardiovascular disease in London, involving four socio-demographic explanatory variables: the Carstairs deprivation index of the area of residence, individual lack of car access, individual social class IV/V (partly skilled and unskilled) and individual non-white ethnicity. The model was fitted to five related datasets: (a) individual-level survey data, (b) district-level aggregate population data, (c) a combination of (a) and (b), (d) ward-level aggregate population data, and (e) a combination of (a) and (d). The posterior distributions of log odds ratios are compared between five datasets, and between four covariates, on the same graph. The densities were computed from MCMC samples of size 10,000, using kernel density estimation.

Shaded as in Figure 2 (bottom left) the strips tend to emphasize the less precise estimates with the widest credible intervals. To prevent this, the plot can be adapted as in Figure 2 (bottom right) by multiplying the shading level for each strip by its density divided by the maximum density over all strips. This ensures that the most precise estimate appears black, and the shading at any x-value on the same axes is proportional to the probability density of that value. A legend can then be included on the figure to indicate the densities represented by the color range. This method of shading density strips ensures that the total amount of "black ink" in each strip on the same figure is equal. Analogously, the varying-width strips (top right) are scaled so that the area of each strip is identical, just as when comparing standard probability density plots, the area under each curve should be 1.

Compared to the traditional line drawings of posterior means and 95% credible intervals (top left), the density strips give a better impression of how well the data support different ranges of parameter values, particularly for the parameters with wide credible intervals. Compared to the varying-width strips, the shaded density strips discriminate better between densities in the tail of the posterior distribution. In particular, a dichotomy is not drawn between "significant" (e.g. non-white, (a)) and "non-significant" coefficients (non-white, (b)).

3.2 Meta-Analysis

In meta-analysis, estimates of a quantity obtained from different studies are compared, and a pooled summary estimate is usually calculated as a weighted average of these. In the forest plot commonly used to illustrate meta-analyses (Lewis and Clarke 2001) the size of the symbol plotted at the point estimate is usually varied in proportion to the weight given to the study in the combined estimate. This highlights the highest-weighted studies with the tightest interval estimates, and draws the eye away from the smallest studies with the widest interval estimates. The weight is related to the size of the study, and in a fixed-effects meta-analysis is usually defined as the inverse of the variance of the point estimate. Figure 3 (top left) shows a forest plot of this type from a meta-analysis by Veenstra et al. (1998) of trials of the effectiveness of coating venous catheters with silver sulfadiazine for preventing catheter infection. (This

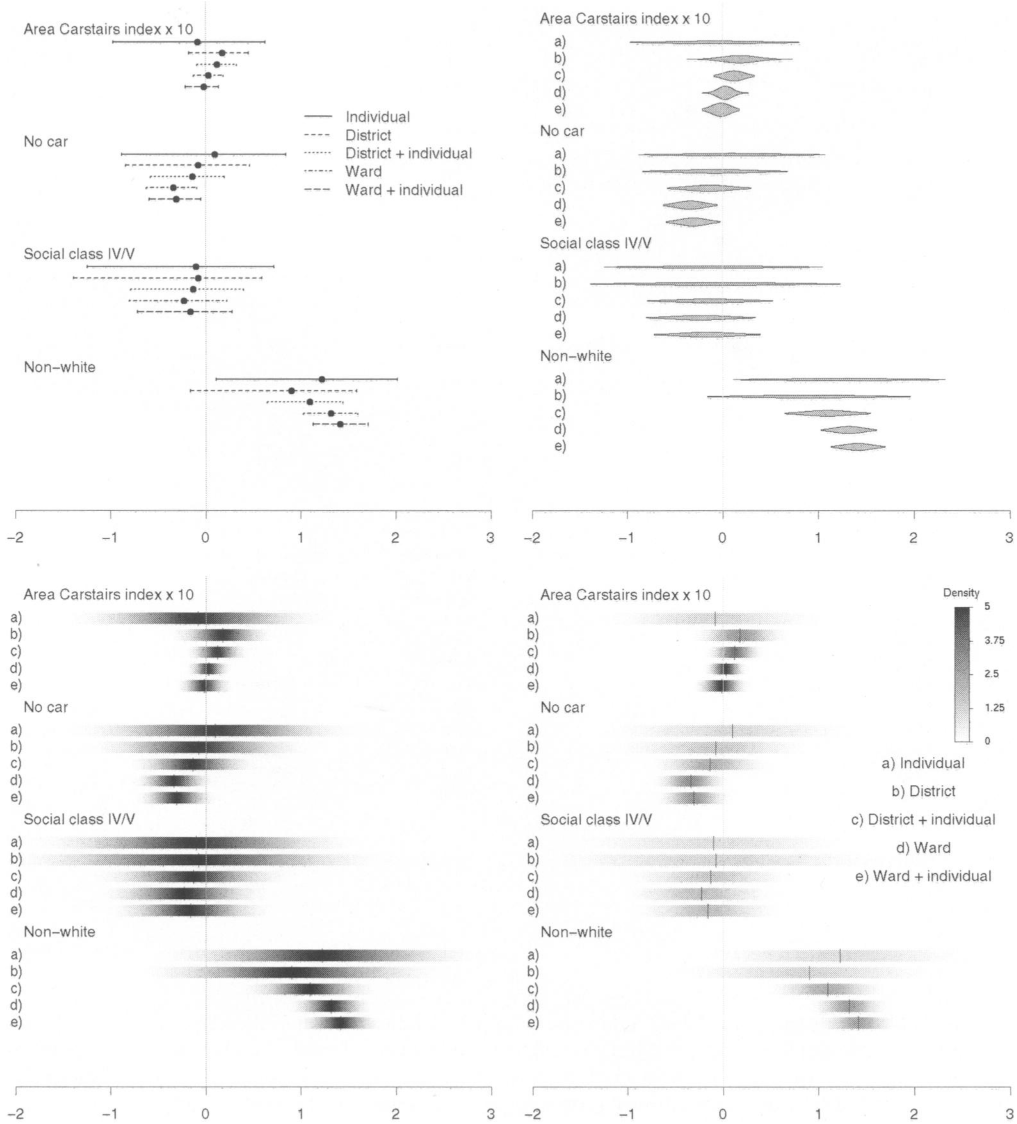

Figure 2. Estimates from a regression with four covariates fitted to five related datasets (Jackson et al. 2008). Posterior mean log odds ratios of hospital admission for cardiovascular disease in London with 95% credible intervals alone (top left), varying-width strips terminating at the 95% credible limits (top right), density strips shaded black at each mode (bottom left), and density strips with shading scaled by the density relative to the maximum density over all strips (bottom right).

differs from the published graph in that here the x-axis is not drawn on the log scale, to illustrate skewed distributions.) The forest plot is redrawn using density strips (bottom) and varying width strips (top right). The densities are calculated assuming a log-normal "fiducial" distribution for each odds ratio. Illustrating the entire distribution of each odds ratio shows the skewness more clearly than the standard forest plot. The tapering shading on the right of the density strips discourages classifying studies as showing a "significant effect" if their 95% confidence limits exclude an odds ratio of 1.

The shading of the density strips is adjusted to indicate the study weighting in a similar manner to forest plots. The shading in Figure 3 (bottom left) is defined by multiplying the darkness for each study by its weight divided by the weight of the top-weighted study. Thus, the point estimate from the top-weighted study appears black, and the point estimates from the smallest studies are light gray. The distribution of the fixed-effects pooled estimate is indicated at the bottom with an arbitrarily thicker density strip, black at its mode. The width of the varying-width strips (top right) is scaled in the same way. Scaling the shading of each density strip by its maximum density relative to the maximum density over all strips, the same method as in Figure 2 (bottom right), produces a qualitatively similar graph (Figure 3, bottom right). There is less variation in the shading, since the density at the mode for study i is proportional to the inverse standard error $1/\hat{\sigma}_i$, whereas the study weight is proportional to $1/\hat{\sigma}_i^2$. In making the estimates appear more similar between studies, this is closer in spirit to Bayesian random-effects meta-analysis (Spiegelhalter et al. 2004), in which study-specific estimates are shrunk towards a pooled summary.

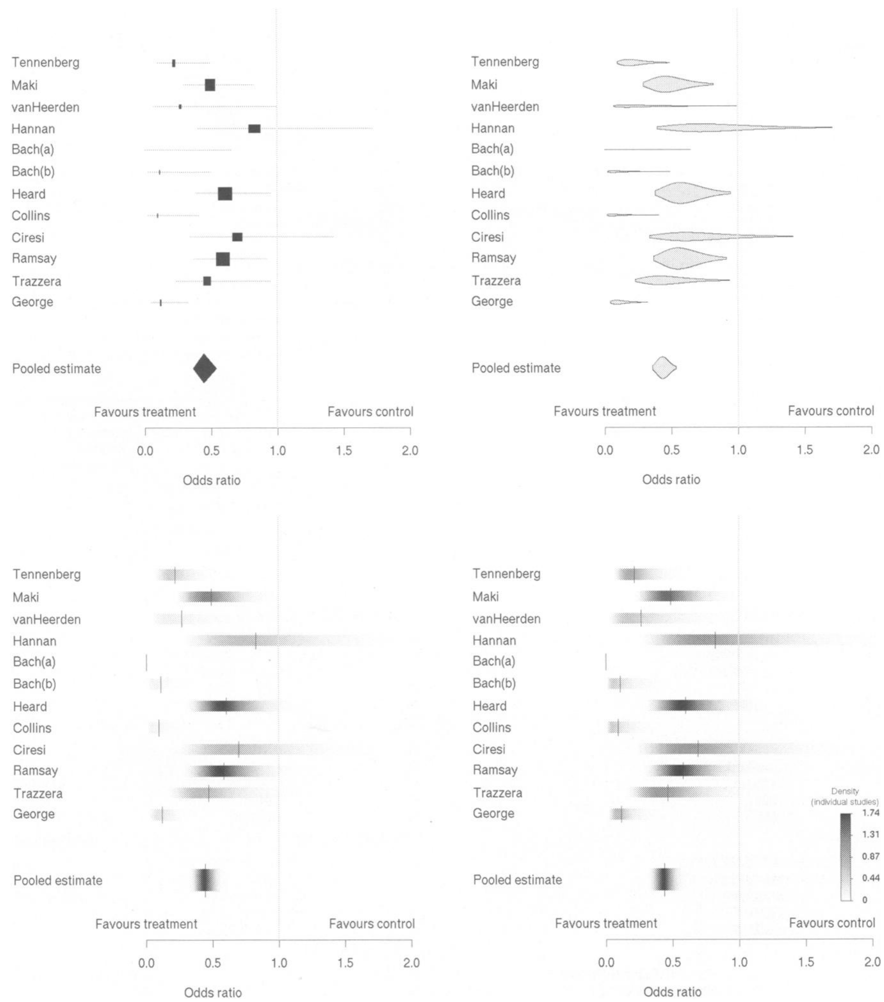

Figure 3. Meta-analysis data of Veenstra et al. (1998): odds ratios from 12 studies of bacterial catheter colonisation in coated compared to noncoated catheters. Top left: traditional forest plot of estimates and 95% confidence intervals with box area at the point estimate proportional to the study weight (inverse variance). Top right: varying-width strips scaled proportionally to the study weight. Bottom left: density strips with ticks at the point estimate and darkness at the point estimate proportional to the study weight. Bottom right: density strips with darkness at the point estimate equal to the density relative to the maximum density over all strips.

In both of these definitions of the density strip for meta-analysis, the largest studies stand out, since the vicinity of their point estimates appears black. Interpreting the graph as a whole, the darker regions highlight the parameter values which are supported more by the data as a whole, either because they are obtained from a more powerful study, or because they are closer to the mean estimate from a particular study. Thus the darkest regions are those which contribute most to the summary estimate, indicated at the bottom of the forest plot. Alternatively, the width of the strips could be varied, which can be interpreted in a similar way—areas with more "black ink" contribute more to the summary estimate.

4. REPRESENTING CONTINUOUSLY VARYING UNCERTAINTY

4.1 Survival Analysis

The shading principle can be extended to illustrate distributions which vary continuously, for example through time. Figure 4 shows a Kaplan–Meier estimate of survival in patients with acute myelogenous leukemia (Miller 1981). Point estimates and pointwise confidence limits are plotted at a series of times. Shading can be used to show the uncertainty surrounding the time-varying survival probability in greater detail. For each distinct estimated survival probability, a "fiducial" density is calculated at a series of ordinates, using Greenwood standard errors and a normal approximation to log(−log(survival)). Contours of constant density are then calculated and filled with a constant shade, using the interpolation algorithm described by Cleveland (1993) as implemented in the `filled.contour()` function in R (R Development Core Team 2008). The resulting graph is a two-dimensional generalization of the density strip, which can capture the smoothly varying uncertainty surrounding the survival estimate. Such graphs achieve a very high "data density" (like the galaxy density map commended by Tufte (1983, p. 154)), since every point within the two-dimensional shaded area has a distinct probability density represented by a distinct color.

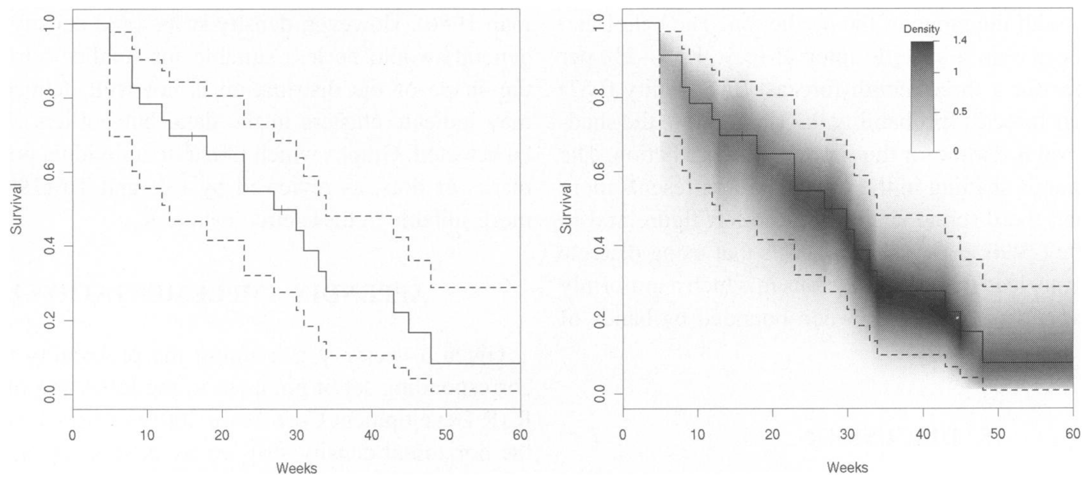

Figure 4. Kaplan–Meier estimates of survival for patients with acute myelogenous leukemia, with 95% confidence limits (left) and also uncertainty represented by shading proportional to density (right), calculated using Greenwood standard errors and a normal approximation to log(−log(survival)).

4.2 Forecasting

Shading also can help to illustrate the uncertainty of forecasts from time series, when the forecast consists of an entire distribution instead of just a point or interval prediction (Tay and Wallis 2000). Figure 5 illustrates forecasts of consumer price inflation in the UK (Bank of England 2007). A forecast based on a normal distribution is computed for every quarter between the second quarter of 2007 and the third quarter of 2010, and the density is calculated at a series of 80 ordinates on the y-axis. Densities between forecast points or between ordinates are interpolated. The illustration currently used by the Bank of England for such forecasts is the *fan chart* (Wallis 1999), so called since the uncertainty increases with the forecast range, resulting in a fan shape. Fan charts use a series of shaded bands, where the darkest band is centered at the region of highest density, and each successive pair of bands moving outwards contains the same percentage of the distribution. The fan chart of these data (Bank of England 2007, pages 8, 40) looks very similar to Figure 5 (left). However, Figure 5 (left) is based on the *density* instead of the *cumulative* density. The shading is black at the mode of each forecast distribution, and thereafter the darkness is proportional to the density within each time point. As with the "box-percentile plots" (Figure 1(c)), plotting the cumulative density instead of the density leads to similar characterizations of unimodal distributions, but multimodal distributions are not represented well.

In Figure 5 (right), instead of coloring all point forecasts black, the shading is black at the highest density over all fore-

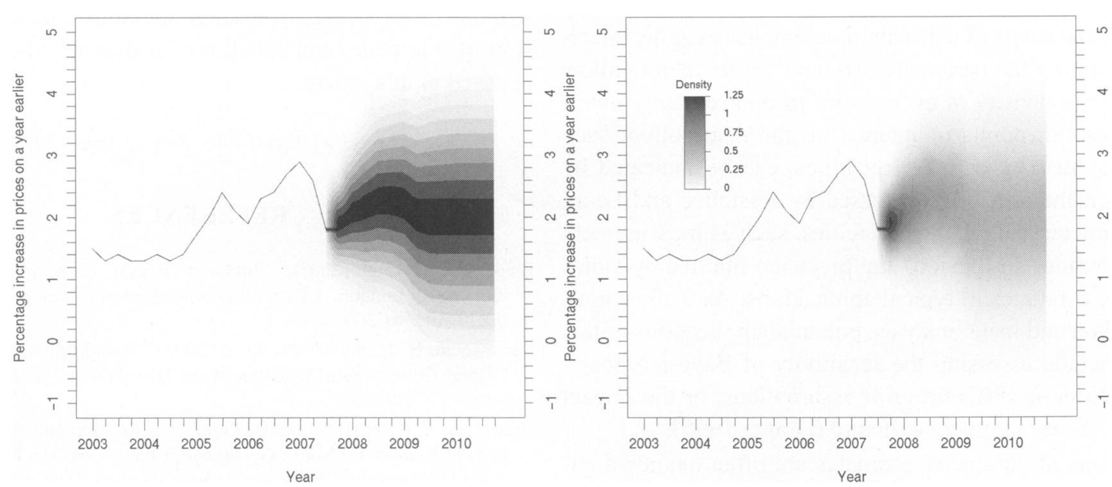

Figure 5. UK consumer price inflation forecasted from the second quarter of 2007. Left: discrete shading with intensity proportional to the posterior density of the forecasted value within each time; right: continuous shading with intensity proportional to the density over all times.

casts, and the darkness is proportional to the density over all forecasts. This definition emphasizes that the uncertainty increases rapidly with the range of the prediction. The belief that inflation will be within a specific interval (say, 1.5%–2% per year) is stronger for a three-month forecast (probability 0.52) than a three-year forecast (probability 0.23), therefore the shading of this interval is darker for the shorter-term prediction. The virtually continuous shading in the right figure represents more information than the discrete shading in the left figure and in fan charts. Tufte (1990, p. 92) also suggests that using discrete bands of color may lead to optical illusions in which a uniformly colored band appears nonuniform when bounded by bands of similar colors.

5. DISCUSSION

Shading is a flexible technique for illustrating uncertainty. Density strips give a compact one-dimensional illustration of distributions arising from parameter estimation. Darker shading indicates regions of values with greater probability density. These improve on traditional line drawings of point-and-interval estimates, accurately indicating that a posterior density is smoothly varying, instead of suggesting that all points within the interval are equally well-supported by the data. If used to present regression coefficients, their property of fading gradually from gray to white discourages casually categorizing effects as "significant" if their interval estimates exclude a null effect. When comparing several distributions on the same figure, the shading can be varied so that the more precise estimates appear darker and each strip contains the same amount of "ink." This clearly indicates the support of the data for each point. Alternatively, in meta-analysis, the shading can be varied with the study weight. The shading principle can be extended to represent the uncertainty surrounding a continuously varying quantity, such as a forecast or survival curve, with a two-dimensional shaded region.

Although perceiving fine differences in color can be more difficult than perceiving differences in position, length, area, and volume (Cleveland 1985), the purpose of the density strip is to indicate the shape of a distribution, emphasizing the uncertainty surrounding the parameter estimate, rather than to allow the value of the density at every point to be accurately determined. If desired, points of interest on the distribution, such as the mean, median, or other quantiles, can be indicated by tick-marks on the strip. As discussed by Kastellec and Leoni (2007), parameter and data uncertainties, such as measurement error, may "render an illusion" the precision implied by tables of parameter estimates in typical applications. As well as multiple regression and meta-analysis, potential applications of this technique include assessing the sensitivity of Bayesian posterior distributions to different prior assumptions, or the impact of uncertainties about model structure (Draper 1995).

Distributions of parameter estimates are often unknown analytically, and estimated from a sample, using a method such as MCMC or bootstrapping. This suggests that as well as illustrating parameter uncertainty, density strips could also be used for illustrating the estimated sampling distributions of observed data, using a method such as kernel density estimation (Silverman 1986). However, density strips (and density estimation in general) would be less suitable for smaller samples in which the shape of the distribution is not well defined. Dark areas may indicate clusters in the data, but outliers are unlikely to be detected. Graphs which illustrate individual points with tick-marks or dots, as reviewed by Lee and Tu (1997), would be more suitable in these circumstances.

APPENDIX: IMPLEMENTATION IN R

Given a vector `y` containing the probability density of the corresponding set of points in `x`, the following function for the R (R Development Core Team 2008) statistical software draws the horizontal density strip on an existing graph whose x-axis contains the range of `x`. The strip is of width `width`, centered on *y*-position `at`. The mode is shaded as black by default, or by a proportion of black defined by `scale`, if supplied. Tick-marks are drawn at x-position `ticks`, if supplied.

This is implemented by drawing a row of tall, thin rectangles, each with no border line, filled with different shades. For efficiency, contiguous rectangles of the same shade are merged to form a larger rectangle.

```
denstrip <- function(x, y, at, width, scale=1,
    ticks=NULL) {
 y <- y / max(y) * scale
 n <- length(x)
 cols <- gray(1 - y[1:(n-1)])
 first.col <- c(TRUE, cols[2:(n-1)]
    != cols[1:(n-2)])
 next.col <- c(first.col,TRUE); next.col[1] <- FALSE
 rect(xleft=x[-n][first.col], xright=x[next.col],
       ybottom=at-width/2, ytop=at+width/2,
       border=NA, col = cols[first.col])
 if (!is.null(ticks)) {
    segments(ticks, at-width*0.75, ticks,
       at+width*0.75)
 }
}
```

The R package `denstrip`, available from *http://CRAN.R-project.org/package=denstrip*, includes a more flexible version of this code, and functions for drawing the other graphs featured in this article.

[Received August 2008. Revised August 2008.]

REFERENCES

Bank of England (2007), "Inflation Report (November)," Threadneedle Street, London. URL: *http://www.bankofengland.co.uk/publications/inflationreport/2007.htm*.

Barrowman, N. J., and Myers, R. A. (2003), "Raindrop Plots: A New Way to Display Collections of Likelihoods and Distributions," *The American Statistician*, 57, 268–274.

Brewer, C. A. (1994), "Color Use Guidelines for Mapping and Visualization," in *Visualization in Modern Cartography*, eds. A. M. MacEachren and D. R. F. Taylor, Tarrytown, NY: Elsevier Science, chap. 7, pp. 123–147.

Cleveland, W. S. (1985), *The Elements of Graphing Data*, Monterey, CA: Wadsworth.

——— (1993), *Visualizing Data*, Summit, NJ: Hobart Press.

Cohen, D. J., and Cohen, J. (2006), "The Sectioned Density Plot," *The American Statistician*, 60, 167–174.

Draper, D. (1995), "Assessment and Propagation of Model Uncertainty" (with discussion), *Journal of the Royal Statistical Society*, Series B, 57, 45–97.

Esty, W. W., and Banfield, J. D. (2003), "The Box-Percentile Plot," *Journal of Statistical Software*, 8.

Gelman, A., and Stern, H. (2006), "The Difference Between "Significant" and "Not Significant" is not Itself Statistically Significant," *The American Statistician*, 60, 328–331.

Hintze, J. L., and Nelson, R. D. (1998), "Violin Plots: A Box Plot–Density Trace Synergism," *The American Statistician*, 52, 181–184.

Jackson, C. H., Best, N. G., and Richardson, S. (2008), "Hierarchical Related Regression for Combining Aggregate and Individual Data in Studies of Socio-economic Disease Risk Factors," *Journal of the Royal Statistical Society*, Series A, 171, 159–178.

Kastellec, J., and Leoni, E. (2007), "Using Graphs Instead of Tables in Political Science," *Perspectives on Politics*, 5, 755–771.

Lee, J., and Tu, N. (1997), "A Versatile One-Dimensional Distribution Plot: The BLiP Plot," *The American Statistician*, 51, 353–358.

Lewis, S., and Clarke, M. (2001), "Forest Plots: Trying to see the Wood and the Trees," *British Medical Journal*, 322, 1479–1480.

Louis, T. and Zeger, S. (2008), "Effective communication of standard errors and confidence intervals," *Biostatistics*, published online 10 June 2008, doi 10.1093/biostatistics/kxn014.

Miller, R. G. (1981), *Survival Analysis*, New York: Wiley.

Poynton, C. (1993), ""Gamma" and its Disguises: The Nonlinear Mappings of Intensity in Perception, CRTs, Film and Video," *SMPTE Journal*, 102, 1099–1108.

——— (1998), "The Rehabilitation of Gamma," in *Proceedings of SPIE 3299*, eds. B. E. Rogowitz and T. N. Pappas, SPIE, Bellingham, WA, pp. 232–249.

R Development Core Team (2008), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Silverman, B. (1986), *Density Estimation for Statistics and Data Analysis*, New York: Chapman and Hall.

Spiegelhalter, D. J. (1999), "Surgical Audit: Lessons from Nightingale and Codman," *Journal of the Royal Statistical Society*, Series A, 126, 45–58.

Spiegelhalter, D. J., Abrams, K. R., and Myles, J. P. (2004), *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*, Chichester, U.K.: Wiley.

Sterne, J. A. C., Davey-Smith, G., and Cox, D. R. (2001), "Sifting the Evidence—What's Wrong with Significance Tests?" *British Medical Journal*, 322, 226–231.

Tay, A. S., and Wallis, K. F. (2000), "Density Forecasting: A Survey," *Journal of Forecasting*, 19, 235–254.

Tufte, E. (1983), *The Visual Display of Quantitative Information*, Cheshire, CT: Graphics Press.

——— (1990), *Envisioning Information*, Cheshire, CT: Graphics Press.

Tukey, J. W. (1977), *Exploratory Data Analysis*, Reading, MA: Addison-Wesley.

Veenstra, D., Saint, S., Saha, S., Lumley, T., and Sullivan, S. (1998), "Efficacy of Antiseptic Impregnated Central Venous Catheters in Preventing Nosocomial Infections: A Meta-analysis," *Journal of the American Medical Association*, 281, 261–267.

Wallis, K. P. (1999), "Asymmetric Density Forecasts of Inflation and the Bank of England's Fan Chart," *National Institute Economic Review*, 167, 106–112.