CREDIBLE MEANS ARE EXACT BAYESIAN FOR EXPONENTIAL FAMILIES

WILLIAM S. JEWELL
University of California, Berkeley
and Teknekron, Inc.

ABSTRACT

The credibility formula used in casualty insurance experience rating is known to be exact for certain prior-likelihood distributions, and is the minimum least-squares unbiased estimator for all others. We show that credibility is, in fact, exact for all simple exponential families where the mean is the sufficient statistic, and is also exact in an extended sense for all regular distributions with their natural conjugate priors where there is a fixed-dimensional sufficient statistic.

ACKNOWLEDGEMENT

I would like to thank Hans Bühlmann for his suggestions to improve and extend this paper.

CREDIBLE MEAN

In the usual model of nonlife insurance [3], each member of a *risk collective* is characterized by a *risk parameter* θ. Given θ, the *risk random variable*, ξ, is a random sample from a *likelihood density*, $p(x \mid \theta)$, defined over $x \varepsilon X$ (discrete or continuous). Thus, for a risk with known parameter θ, the *fair premium* is $m(\theta) = E\{\xi \mid \theta\} = \int xp(x \mid \theta)\, dx$, and the *risk variance* is $v(\theta) = V\{\xi \mid \theta\}$.

The risk parameters have a *prior density* $u(\theta)$, $\theta \varepsilon \Theta$; we assume that statistics are known from the mixed *collective density*, $p(x) = E_\theta p(x \mid \theta) = \int p(x \mid \theta)\, u(\theta)\, d\theta$, in particular, the *collective fair premium* $m = E_\theta m(\theta)$, and *collective variance* $v = E_\theta v(\theta) + V_\theta m(\theta)$.

The central problem of *experience rating* is to estimate the fair premium of an individual risk, given only collective statistics, and n years *individual experience* $\underline{x} = \{\xi_t = x_t;\ (t = 1, 2, \ldots, n)\}$, that is, to estimate $E\{\xi_{n+1} \mid \underline{x}\}$. Based on heuristic arguments, American actuaries in the 1920's proposed a *credibility formula* of form:

$$E\{\xi_{n+1} \mid \underline{x}\} \approx (1 - Z)\, m + Z \left(\frac{1}{n} \sum_{t=1}^{n} x_t\right); \quad Z = \frac{n}{n+N}; \tag{1}$$

with N determined experimentally [12].

The Bayesian formulation is straightforward. One first finds the *density of* θ *posterior to the data*:

$$u_n(\theta \mid \underline{x}) = \frac{\Pi p(x_t \mid \theta)\, u(\theta)}{\int \Pi p(x_t \mid \phi)\, u(\phi)\, d\phi} \tag{2}$$

and then:

$$E\{\xi_{n+1} \mid \underline{x}\} = \iint y p(y \mid \theta)\, u_n(\theta \mid \underline{x})\, d\theta dy. \tag{3}$$

In the 1950's, Bailey [1] and Mayerson [13] showed that (1) was, in fact, exactly the Bayesian result (3) for special prior-likelihood combinations: Beta-Binomial, Gamma-Poisson, Gamma-Exponential, Normal-Normal, and similar cases. Bühlmann [2] then showed that (1) was the minimum least-squares unbiased estimator for arbitrary families, if $N = E_\theta v(\theta)/V_\theta m(\theta)$. More detailed historical remarks can be found in [7] and [9].

We now show that the credibility formula (1) is exact for a larger class of prior-likelihood families.

Exponential Family

The Koopman-Pitman-Darmois *exponential family* of likelihoods [6] [11] is:

$$p(x \mid \theta) = \frac{a(x) \exp \sum_{i=1}^{I} \phi_i(\theta) \cdot f_i(x)}{c(\theta)}, \quad (x \in X) \tag{4}$$

where $c(\theta)$ is a normalizing factor to make $\int p(x \mid \theta)\, dx = 1$. If $\underline{x} = (x_1, x_2, \ldots, x_n)$ is a *random sample* of size n, then n and the I sums, $F_i = \sum_{t=1}^{n} f_i(x_t)$ $(i = 1, 2, \ldots, I)$, are *sufficient statistics* for θ. In fact, subject to mild regularity conditions, the exponential family is the only one which has a sufficient statistic $(F_1, F_2, \ldots, F_I; n)$ of fixed dimension for every n, given that X does not depend on θ.

Furthermore, if we pick for the family of priors the *natural conjugate prior* [4]

$$u(\theta) \propto [c(\theta)]^{-n_0} \exp \sum_{i=1}^{I} f_{0i}\phi_i(\theta), \quad (\theta \in \Theta), \tag{5}$$

then the family will be *closed under sampling*, that is, the density of θ *posterior to the data*, $u_n(\theta \mid \underline{x})$, will be of the same form as (5) with the $I + 1$ *hyperparameters* updated by: $n_0 \leftarrow n_0 + n$; $f_{0i} \leftarrow f_{0i} + \Sigma f_i(x_t)$ $(i = 1, 2, \ldots, I)$.

The Simple Exponential Family

The family of distributions for which credibility will turn out to be exact is the *single-parameter* exponential family with $f_1(x) = x$, and *natural parameterization* $\phi_1(\theta) = -\theta$, i.e.,

$$p(x \mid \theta) = \frac{a(x)\, e^{-\theta x}}{c(\theta)} \tag{6}$$

for continuous or discrete measure in the range X, determined by the nonvanishing of $a(x)$. The *sample sum or mean* is, with n, the sufficient statistic for this simple family.

The natural conjugate prior to (6) is:

$$u(\theta) = \frac{[c(\theta)]^{-n_0}\, e^{-\theta x_0}}{d(n_0, x_0)}, \tag{7}$$

defined over a *natural parameter space*, Θ, for which (6) is a density; i.e., for all values of θ for which $c(\theta)$ is finite. Restrictions on the hyperparameters (n_0, x_0) may be necessary to make (7) a density as well, i.e., to make the normalization $d(n_0, x_0)$ finite. We shall henceforth assume $n_0 > 0$; parameter updating is:

$$\begin{aligned} n_0 &\leftarrow n_0 + n \\ x_0 &\leftarrow x_0 + \sum_{t=1}^{n} x_t. \end{aligned} \tag{8}$$

We shall need certain properties of $c(\theta)$ and Θ. These follow directly from the fact that $c(\theta)$ is a transform [14] in the continuous case:

$$c(\theta) = \int_{x \in X} a(x)\, e^{-\theta x} dx. \tag{9}$$

If X is (countably) discrete, we can use the same approach by incorporating Dirac delta-functions in our definition of $a(x)$, and defining (9) over the convex hull of X.

Cases in which (9) does not exist for any value of θ, or for only one value, are uninteresting. Since Θ, if it exists, is convex, three possibilities remain; Θ has a range which is:

(a) finite;
(b) semi-infinite; or
(c) doubly infinite.

At finite ends of the range $c(\theta)$ "usually is infinite (See Note added in proof)" thus, $n_0 > 0$ insures $u(\theta) \to 0$ at these points. $u(\theta)$ must also be zero at infinite endpoints.

X cannot depend on θ, but may influence Θ. For example, if X is finite, then $\Theta = (-\infty, +\infty)$. Or, if (the convex hull of) X is $[0, \infty)$, then $\Theta = (\theta_1, +\infty)$, with θ_1 usually finite, but possibly $-\infty$ (see below). The only case in which Θ can be finite is if X is $(-\infty, \infty)$.

Also, from transform theory [14], we know that $c(\theta)$ is analytic at all interior points of Θ, and derivatives of all orders can be passed under the integral sign in (9), giving analytic functions of θ. Thus, the individual risk mean and variance are:

$$m(\theta) = \frac{-c'(\theta)}{c(\theta)} = -\frac{d}{d\theta} \ln c(\theta); \qquad (10)$$

$$v(\theta) = -\frac{dm(\theta)}{d\theta}. \qquad (11)$$

Since $v(\theta) \geq 0$ $(\theta \in \Theta)$, $m(\theta)$ must be monotone decreasing with range in (the convex hull of) X. Then, in addition to $c(\theta)$ being a positive function, infinite at finite ends of the range, we see it must be strictly convex, in fact monotone decreasing if $X = [0, \infty)$.

The Prior and Posterior Mode

By differentiation, we find:

$$\frac{du(\theta)}{d\theta} = (n_0 m(\theta) - x_0) \cdot u(\theta), \qquad (12)$$

that is, starting from its zero value at the left endpoint, $u(\theta)$ has at first zero or positive slope, then ultimately negative slope, with a unique *prior mode* $\hat{\theta}_0$ at

$$m(\hat{\theta}_0) = x_0/n_0. \qquad (13)$$

Furthermore, from (8), we see that $\hat{\theta}_n$, the mode of $u_n(\theta \mid \underline{x})$ posterior to the data, will satisfy:

$$m(\hat{\theta}_n) = \frac{x_0 + \Sigma x_t}{n_0 + n}. \qquad (14)$$

Thus, if we pick x_0/n_0 in the convex hull of X, the mean risk at the mode will always remain "in range".

However, we know that if the experience data is sampled from a risk with true parameter θ_T, that for fairly arbitrary priors, the posterior converges to a degenerate distribution at θ_T. This means that the (random) estimate $m(\hat{\theta}_n)$ converges with probability one as $n \to \infty$ to the true fair premium $m(\theta_T)$; this can also be verified by the strong law of large numbers applied to the RHS of (14). Furthermore, if $v(\theta)$ is slowly varying in the neighborhood of θ_T, then to a good approximation θ has a Normal density, with mean θ_T and variance $[(n_0 + n)\, v(\theta_T)]^{-1}$, as $n \to \infty$.

Note that $\hat{\theta}_n$ is *not* the maximum-likelihood estimator (which would be that estimate of θ gotten from (14) with $x_0 = n_0 = 0$); in modern terminology, this means our prior is *informative*, and our measurement is *imprecise*.

Credible Mean is Exact

We now show a stronger result relating to the mean risk. Integrating (12) over the natural range Θ, we get:

$$u(\theta) \mid_{\Theta} = n_0 \int_{\Theta} m(\theta)\, u(\theta)\, d\theta - x_0. \qquad (15)$$

Assuming $u(\theta)$ is zero at endpoints of Θ, the prior mean risk must be:

$$m = E_{\theta} m(\theta) = x_0/n_0. \qquad (16)$$

Furthermore, when θ is updated by (8), the mean risk posterior to the data is:

$$E\{\xi_{n+1} \mid \underline{x}\} = E_{\theta \mid \underline{x}} m(\theta) = \frac{x_0 + \Sigma x_t}{n_0 + n} = (1 - Z) \cdot m + Z\left(\frac{1}{n} \Sigma x_t\right) \qquad (17)$$

with

$$Z = \frac{n}{n + n_0}. \tag{18}$$

Thus, credibility is exact for simple exponential families!

The remaining fact that

$$n_0 = E_\theta v(\theta)/V_\theta m(\theta) = N, \tag{19}$$

can be inferred from known results [2] [5], or by forming $\frac{d^2u(\theta)}{d\theta^2}$ from (12), and integrating, to give:

$$\left.\frac{du(\theta)}{d\theta}\right|_\Theta = -n_0 E_\theta v(\theta) + n_0^2 E_\theta \{(m(\theta))^2\} - 2n_0 \cdot x_0 \cdot m + x_0^2; \tag{20}$$

then (19) follows by assuming the slope of u(θ) is zero at endpoints of Θ. Additional restrictions on the hyperparameters may be necessary to insure finiteness of the variance.

Other Sufficient Statistics

If we encounter a simple exponential family with an arbitrary sufficient statistic, a linear forecast may also exist for the *transformed* random variable. For, suppose we observe data $\underline{y} = \{\eta_t = y_t; (t = 1, \ldots, n)\}$, and know $(\Sigma f(y_t); n)$ is the sufficient statistic; then, if the sample space does not depend on θ, the likelihood must be of exponential form:

$$p(y \mid \theta) = \frac{b(y)\, e^{-\theta f(y)}}{c(\theta)}. \tag{21}$$

Assuming $f(y)$ has an inverse, we can transform the problem by setting:

$$x = f(y); a(x) = \frac{b(f^{-1}(x))}{f'(f^{-1}(x))}. \tag{22}$$

The previous result then shows that:

$$E\{f(\eta_{n+1}) \mid \underline{y}\} = \frac{n_0 E\{f(\eta)\} + \sum_{t=1}^{n} f(y_t)}{n_0 + n}. \tag{23}$$

The classic example of this is the Normal with known mean m, and unknown precision θ:

$$p(x \mid \theta) = \sqrt{\frac{\theta}{2\pi}}\, e^{-\theta(x-m)^2/2}, \quad X = (-\infty, +\infty) \tag{24}$$

for which the natural conjugate prior is known to be the Gamma, $u(\theta) = G(n_0/2 + 1; x_0; \theta)$ The variance forecast is then linear in the sample variance:

$$E\left\{\frac{(\xi_{n+1} - m)^2}{2} \mid \underline{x}\right\} = \frac{x_0 + \frac{1}{2}\Sigma\,(x_t - m)^2}{n_0 + n}. \tag{25}$$

Another example is the Pareto with range $[1, \infty)$, but unknown shape parameter θ:

$$p(x \mid \theta) = \theta x^{-(\theta+1)} \qquad X = [1, \infty). \tag{26}$$

The natural conjugate prior is $G(n_0 + 1;\ x_0;\ \theta)$, and the linear forecast is of logvariables:

$$E\{\ln(\xi_{n+1}) \mid \underline{x}\} = \frac{x_0 + \ln(\Pi\, x_t)}{n_0 + n}. \tag{27}$$

Credible Means for the General Exponential Family

It is easy to see how the argument for the simple exponential family can be extended to the general exponential family (4) with $I + 1$ sufficient statistics, if we use *natural parameterization* $\phi_i(\theta) = -\theta_i$ $(i = 1, 2, \ldots, I)$:

$$p(x \mid \underline{\theta}) = \frac{a(x)\exp\left(-\sum_{i=1}^{I} \theta_i \cdot f_i(x)\right)}{c(\underline{\theta})} \quad (x \in X). \tag{28}$$

Here the natural parameter space Θ consists of all points $\theta = [\underline{\theta}_1, \theta_2, \ldots, \theta_I]$ in R^I for which

$$c(\underline{\theta}) = \int_{x \in X} a(x)\, e^{-\Sigma\,\theta_i f_i(x)}\, dx \tag{29}$$

is finite. It is known that Θ is convex.

With this choice, the natural conjugate prior is:

$$u(\underline{\theta}) \propto [c(\underline{\theta})]^{-n_0} \exp\left(-\sum_{i=1}^{I} f_{0i}\theta_i\right) \quad (\underline{\theta} \in \Theta) \tag{30}$$

with the usual updating for the $I + 1$ hyperparameters (f_{01}, $f_{02}, \ldots, f_{0I}$; n_0. Assume $n_0 > 0$.

If we define *generalized means* and (*co*)*variances*

$$M_i(\underline{\theta}) = E_{\xi|\underline{\theta}} f_i(\xi)\,;\ C_{ij}(\underline{\theta}) = C_{\xi|\underline{\theta}}\{f_i(\xi)\,;\ f_j(\xi)\} \tag{31}$$

for all i, j (assuming they exist), then we can generalize (10) and (11):

$$M_i(\underline{\theta}) = -\frac{\partial}{\partial\theta_i}\ln c(\underline{\theta})\,; \tag{32}$$

$$C_{ij}(\underline{\theta}) = -\frac{\partial M_i(\underline{\theta})}{\partial\theta_j} = -\frac{\partial M_j(\underline{\theta})}{\partial\theta_i}. \tag{33}$$

Now assume that $u(\underline{\theta}) \equiv 0$ *everywhere on the boundary of* Θ. This seems reasonable in view of our arguments in the one-dimensional case. Then, since

$$\frac{\partial u(\underline{\theta})}{\partial\theta_i} = (n_0 \cdot M_i(\underline{\theta}) - f_{0i}) \cdot u(\underline{\theta}) \qquad (i = 1, 2, \ldots, I) \tag{34}$$

it follows by integration that for all i

$$E_{\xi} f_i(\xi) = E_{\underline{\theta}} M_i(\underline{\theta}) = \frac{f_{0i}}{n_0}, \tag{35}$$

and

$$E\{f_i(\xi_{n+1}) \mid \underline{x}\} = E_{\underline{\theta}|\underline{x}} M_i(\underline{\theta}) = \frac{f_{0i} + \sum_{t=1}^{n} f_i(x_t)}{n_0 + n}. \tag{36}$$

Note that the time constant, n_0, is the same for all components, and that f_{0i}/n_0 should be selected in the range of $f_i(x)$.

In other words, if it is known that a likelihood has a *fixed-dimensional* ($I + 1$) *sufficient statistic* and is *regular* (X does not depend on $\underline{\theta}$ plus regularity conditions), then, if the natural conjugate prior is used, it follows that there are I functions, $f_i(x)$, whose *mean values are updated by linear credibility formulae* (36):

The classic two-dimensional example is the Normal distribution with unknown mean μ and precision τ:

$$p(x \mid \mu, \tau) = \sqrt{\frac{\tau}{2\pi}}\, e^{-\tau(x-\mu)^2/2} = N(\mu, \tau^{-1}; x) \quad X = (-\infty, \infty). \tag{37}$$

The natural parameterization is obtained by substituting $\theta_1 = -\mu\tau$, $\theta_2 = \tau/2$, and $f_1(x) = x$, $f_2(x) = x^2$.

It is known that the natural conjugate prior is a two-dimensional *Normalgamma* distribution. With our choice of parameterization and hyperparameters, it has the following properties:

(a) The marginal distribution of τ is Gamma,

$$G\left(\frac{n_0 + 3}{2}\,;\ \tfrac{1}{2}\left(f_{02} - \frac{f_{01}^2}{n_0}\right); \tau\right);$$

(b) The *conditional* distribution of μ, given τ, is Normal,

$$N\left(\frac{f_{01}}{n_0}\,;\quad (n_0\tau)^{-1};\ \mu\right);$$

(c) The *marginal* distribution of μ is a generalized Student-t distribution, with $n_0 + 3$ degrees of freedom, mean f_{01}/n_0, and variance

$$(n_0 + 1)^{-1}\left(\frac{f_{02}}{n_0} - \frac{f_{01}^2}{n_0^2}\right).$$

Explicit formulae may be found in DeGroot [4], pp. 42 and 169. His notation seems to imply four hyperparameters, but there are only three independent ones.

By direct calculation we find that the collective mixed density is a generalized Student-t with $n_0 + 3$ degrees of freedom, mean f_{01}/n_0, and variance $\left(\frac{f_{02}}{n_0} - \frac{f_{01}^2}{n_0^2}\right)$. Clearly, we must pick $f_{02} \geq f_{01}^2/n_0$.

From our generalized result (36), we find:

$$E\{\xi_{n+1} \mid \underline{x}\} = \frac{f_{01} + \Sigma x_t}{n_0 + n} = (1 - Z)\left(\frac{f_{01}}{n_0}\right) + Z\left(\frac{1}{n}\Sigma x_t\right); \qquad (38)$$

$$E\{\xi_{n+1}^2 \mid \underline{x}\} = \frac{f_{02} + \Sigma x_t^2}{n_0 + n} = (1 - Z)\left(\frac{f_{02}}{n_0}\right) + Z\left(\frac{1}{n}\Sigma x_t^2\right). \qquad (39)$$

As before, $Z = n/(n + n_0)$, and

$$\frac{E_{\underline{\theta}}V\{\xi \mid \underline{\theta}\}}{V_{\underline{\theta}}E\{\xi \mid \underline{\theta}\}} = \frac{E\{\tau^{-1}\}}{V\{\mu\}} = n_0, \qquad (40)$$

but the formula for updating the total variance is more complicated:

$$V\{\xi_{n+1} \mid \underline{x}\} = (1 - Z)\left[\frac{f_{02}}{n_0} - \frac{f_{01}^2}{n_0^2}\right] + Z\left[\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{j=1}^{n} x_j\right)^2\right]$$

$$+ Z(1 - Z)\left[\frac{f_{01}}{n_0} - \frac{1}{n}\Sigma\, x_t\right]^2. \qquad (41)$$

In other words, for a Normal with unknown mean and precision, the variance forecast is *not* a linear mixture of prior variance and sample variance, but obtains predictive information from sample mean deviations as well. Credibility formulas hold only for the mean and second moment.

Special Distributions

Table 1 shows the simple exponential families for which credibility was previously known to be exact. (Notation follows [6]). Note that in many cases a transformation is necessary to get the natural parameterization. In all cases, the hyperparameters have been chosen to satisfy (16) and (19).

Likelihoods (1), (2), (4) in Table 1 can be "enriched", by noting that if the *r-fold convolution* of $p(x \mid \theta)$ is taken, $a(x)$ is replaced by $a^{r*}(x)$, $c(\theta)$ by $[c(\theta)]^r$, and the prior is of the same form, but with n_0 replaced by $n_0 r$. With this convention, the new $V_\theta m(\theta)$ will be r^2 times the value in Table 1 (with n_0 replaced by $n_0 r$), but it is still true that

$$m = x_0/n_0;\; E_\theta v(\theta) = n_0 V_\theta m(\theta);\; v = (n_0 + 1)\, V_\theta m(\theta), \qquad (42)$$

and the credible estimate (17) still holds.

The enlarged families are:

(1) Binomial $p(x \mid \pi) = \binom{r}{x}(1 - \pi)^{r-x}\pi^x \quad X = \{0, 1, 2, \ldots, r\}$

Beta-Binomial $p(x) = BB(r, x_0, rn_0 - x_0)$

(2) Negative Binomial $p(x \mid \pi) = NB(r, \pi; x)$

Beta-(Negative) Binomial $p(x) = BB(-r, x_0, -rn_0 - x_0; x)$

(4) Gamma $p(x \mid \pi) = G(r, \pi; x)$

$$\text{Shifted Pareto } p(x) = \frac{\Gamma(rn_0 + r + 1)}{x_0 \cdot \Gamma(rn_0 + 1)}\left(1 + \frac{y}{x_0}\right)^{-(rn_0+2)}. \qquad (43)$$

TABLE I: CLASSICAL SIMPLE EXPONENTIAL FAMILIES

LIKELIHOOD	PRIOR	NATURAL PARAMETERIZATION	$m(\theta)$	$V_\theta m(\theta)$	PREDICTIVE DENSITY
1. Bernoulli $p(x \mid \pi) = (1-\pi)^{1-x}\pi^x$ $X = \{0,1\}$	Beta $u(\pi) = Be(x_o, n_o - x_o; \pi)$ $\frac{x_o}{n_o} \in \Pi = [0,1]$	$\theta = \ln\left(\frac{1-\pi}{\pi}\right)$ $\Theta = (-\infty,\infty)$	$(e^\theta + 1)^{-1}$	$\frac{x_o(n_o - x_o)}{n_o^2(n_o + 1)}$	Bernoulli $p(x) = \left(1 - \frac{x_o}{n_o}\right)^{1-x}\left(\frac{x_o}{n_o}\right)^x$
2. Geometric $p(x \mid \pi) = (1-\pi)\pi^x$ $X = \{0,1,2, \ldots\}$	Beta $u(\pi) = Be(x_o, n_o + 1; \pi)$ $\frac{x_o}{n_o} \in \Pi = [0,1]$	$\theta = \ln \frac{1}{\pi}$ $\Theta = [0,\infty)$	$(e^\theta - 1)^{-1}$	$\frac{x_o}{n_o^2} \cdot \frac{n_o + x_o}{n_o - 1}$ $(n_o > 1)$	Beta-(Negative) Binomial $p(x) = BB(-1, x_o, -n_o - x_o; x)$
3. Poisson $p(x \mid \pi) = \frac{e^{-\pi}\pi^x}{x!}$ $X = \{0,1,2, \ldots\}$	Gamma $u(\pi) = G(x_o, n_o; \pi)$ $x_o \in \Pi = [0,\infty)$	$\theta = \ln \frac{1}{\pi}$ $\Theta = (-\infty,\infty)$	$e^{-\theta}$	$\frac{x_o}{n_o^2}$	Negative Binomial $p(x) = NB\left(x_o, (n_o + 1)^{-1}; x\right)$

4. **Exponential** $p(x \mid \pi) = \pi e^{-\pi x}$ $X = [0,\infty)$	Gamma $u(\pi) = G(n_o + 1, x_o; \pi)$ $x_o \in \Pi = [0,\infty)$	$\theta = \pi$ $\Theta = [0,\infty)$	θ^{-1}	$\dfrac{x_o^2}{n_o^2(n_o - 1)}$ $(n_o > 1)$	Shifted Pareto $p(x) = \left(\dfrac{n_o + 1}{x_o}\right)\left(1 + \dfrac{y}{x_o}\right)^{-(n_o+2)}$
5. Normal-Known Variance $p(x \mid \pi) = N(\pi, s^2; x)$ $X = (-\infty,\infty)$	Normal $u(\pi) = N\left(\dfrac{x_o}{n_o}; \dfrac{s^2}{n_o}; \pi\right)$ $\Pi = (-\infty,+\infty)$	$\theta = \dfrac{-\pi}{s^2}$ $\Theta = (-\infty,\infty)$	$-s^2\theta$	$\dfrac{s^2}{n_o}$	Normal $p(x) = N\left(\dfrac{x_o}{n_o}; s^2\left(1 + \dfrac{1}{n_o}\right); x\right)$

$$m = \frac{x_o}{n_o} \;;\quad E_\theta v(\theta) = n_o \cdot V_\theta m(\theta) \;;\quad v = (n_o + 1)V_\theta m(\theta)$$

$$Be(a,b;x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \qquad NB(a,b;x) = \frac{\Gamma(a+x)}{\Gamma(a)x!}(1-b)^a b^x \qquad BB(n,a,b;x) = \binom{n}{x}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{\Gamma(a+x)\Gamma(n+b-x)}{\Gamma(n+a+b)}$$

$$G(a,b;x) = \frac{b^a x^{a-1}}{\Gamma(a)} e^{-bx} \qquad N(a,b;x) = \frac{1}{\sqrt{2\pi b}} e^{-\frac{1}{2}\left(\frac{x-a}{b}\right)^2} \qquad BB(-n,a,-b;x) = \frac{\Gamma(x+n)}{\Gamma(n)x!}\frac{\Gamma(b+1)}{\Gamma(a)\Gamma(b+1-a)}\frac{\Gamma(a+x)\Gamma(n+b+1-a)}{\Gamma(n+b+1+x)}$$

Convolution in the Poisson and Normal families merely gives scale changes.

To illustrate the relationship between X and Θ, we consider two new families of exponential type. First, if

$$a(x) = \sinh x, \qquad X = [0, \infty) \tag{44}$$

we find

$$c(\theta) = (\theta^2 - 1)^{-1}; \; m(\theta) = 2\theta(\theta^2 - 1)^{-1}; \; v(\theta) = 2(\theta^2 + 1)(\theta^2 - 1)^{-2} \tag{45}$$

and

$$u(\theta) = \frac{(\theta^2 - 1)^{n_0} e^{-\theta x_0}}{d(n_0, x_0)} \quad (x_0 > 0) \tag{46}$$

for $\Theta = [1, \infty)$.

However, if

$$a(x) = \tfrac{1}{2} e^{-|x|} \qquad X = (-\infty, +\infty), \tag{47}$$

we find

$$c(\theta) = (1 - \theta^2)^{-1}; m(\theta) = -2\theta(1 - \theta^2)^{-1}; v(\theta) = 2(\theta^2 + 1)(1 - \theta^2)^{-2} \tag{48}$$

and

$$u(\theta) = \frac{(1 - \theta^2)^{n_0} e^{-\theta x_0}}{d(n_0, x_0)} \tag{49}$$

for finite range $\Theta = [-1, +1]$:

The normalizing factors $d(n_0, x_0)$ and the collective densities are modified Hankel and Bessel functions of order $n_0 + \frac{1}{2}$, and further computations are laborious. Nevertheless, credibility and the basic formulae (16)-(19) must still hold.

Bibliography

[1] Bailey, A. L., "Credibility Procedures, LaPlace's Generalization of Bayes' Rule, and the Combination of Collateral Knowledge with Observed Data", *Proceedings of the Casualty Actuarial Society*, Vol. 37, pp. 7-23, 1950; Discussion, *Proceedings of the Casualty Actuarial Society*, Vol. 37, pp. 94-115, 1950.

[2] Bühlmann, H., "Experience Rating and Credibility", *ASTIN Bulletin*, Vol. 4, Part 3, pp. 199-207, July 1967.

[3] ——, *Mathematical Methods in Risk Theory*, Springer-Verlag, New York, 1970.

[4] DeGroot, M. H., *Optimal Statistical Decisions*, McGraw-Hill, New York, 1970.

[5] Ericson, W. A., "A Note on the Posterior Mean of a Population Mean", *Journal of the Royal Statistical Society*, Series B, Vol. 31, No. 2, pp. 332-334, 1969.

[6] Ferguson, T. S., *Mathematical Statistics: A Decision Theoretic Approach*, Academic Press, New York, 1967.

[7] Hewitt, C. C., Jr., "Credibility for Severity", *Proceedings of the Casualty Actuarial Society*, Vol. 57, pp. 148-171, 1970; Discussion, *Proceedings of the Casualty Actuarial Society*, Vol. 58, pp. 25-31, 1971.

[8] Jewell, W. S., "Multi-Dimensional Credibility", ORC 73-7, Operations Research Center, University of California, Berkeley, April 1973; to appear in *Journal of the American Risk Insurance Association.*

[9] ——, "Models of Multi-Dimensional Credibility", (in preparation).

[10] ——, "The Credible Distribution", ORC 73-13, Operations Research Center, University of California, Berkeley, August 1973; *ASTIN Bulletin*, Vol. VII, Part 3.

[11] Lehmann, E. L., *Testing Statistical Hypotheses*, John Wiley and Sons, Inc., New York, 1959.

[12] Longley-Cook, L. H., "An Introduction to Credibility Theory", *Proceedings of the Casualty Actuarial Society*, Vol. 49, pp. 194-221, 1962; Available as a separate report from Casualty Actuarial Society, 200 East 42nd Street, New York, New York 10017.

[13] Mayerson, A. L., "A Bayesian View of Credibility", *Proceedings of the Casualty Actuarial Society*, Vol. 51, pp. 85-104, 1964; Discussion, *Proceedings of the Casualty Actuarial Society*, Vol. 52, pp. 121-127, 1965.

[14] Van der Pol, B. and H. Bremmer, *Operational Calculus*, 2nd Edition, University Press, Cambridge, 1955.

Note added in proof

In the original paper, it was assumed that if Θ has finite endpoints, $c(\theta) \rightarrow \infty$ at these points. A recent counterexample by R. B. Miller and A. Banerjee shows that this argument is correct only for meremorphic functions.

In the general case, additional regularity conditions must be attached to $a(x)$ to ensure that the left-hand-side of (15) is zero. For example, if X is one-sided and $\Theta = [0, \infty)$, then the class of counterexamples consists of all $a(x)$ for which $\int a(x)\, dx$ is finite, but the limit is approached more slowly than any exponential. These conditions will be discussed further in a forthcoming paper.