ACTUARIAL RESEARCH CLEARING HOUSE
1985 VOL. 2

PARAMETRIC EMPIRICAL BAYES AND CREDIBILITY THEORY

CARL N. MORRIS

Center for Statistical Sciences
The University of Texas
Austin, Texas

The current view of empirical Bayes inference models data and parameters alike with two families of distributuions: one family for data, and one for the unknown parameters, with the latter involving wither parametric prior distributions (PEB) or nonparametric prior distributions (NPEB). The (prior) distributions of the unknown parameters is assumed to be restricted to a class Π having just one member, while frequentist statistical inference allows Π to contain all possible prior distributions. Empirical Bayes inference covers involving limited knowledge about Π, giving the statistician flexibility to develop new statistical procedures.

This paper shows credibility theory and PEB theory are intimately related, and reviews and applies recent PEB theory to the credibility situation, for the normal and other familiar distributions.

FOUNDATIONS

Empirical Bayes involves 2 families of stochastic processes, one for the data (distribution, likelihood) and one for the parameters (prior, superpopulation).

Parametric ($\alpha \in \mathcal{A}$) Empirical Bayes (PEB) Paradigm: $\Pi = \{\pi_\alpha : \alpha \in \mathcal{A}\}$ is known

	Descriptive Model	Inferential Model
Observed Data	(1) $Y \mid \theta \sim f(y \mid \theta)$ (Likelihood)	(1*) $Y \sim \bar{f}_\alpha(y)$, $\alpha \in \mathcal{A}$
Unobserved Parameters (prior)	(2) $\theta \sim g_\alpha(\theta)$, $\alpha \in \mathcal{A}$	(2*) $\theta \mid Y \sim \bar{g}_\alpha(\theta \mid y)$

Evaluations involve two integrals: $r(\alpha,t) = E_\alpha E_\theta L(\theta, t(Y))$

e.g.: $E_\alpha E_\theta(\hat{\theta}(Y) - \theta)^2$ = Bayes risk (α) + $E(\hat{\theta}(Y) - \hat{\theta}_\alpha(Y))^2$, $\hat{\theta}_\alpha$ = Bayes Estimate

Nonparametric Empirical Bayes (Robbins, $k \to \infty$), if $\alpha = \pi_1$

→ Richness of Π.

Bayes (1 prior)	PEB	NPEB	Frequency (all priors)
	(.. Restricted class Π ....)		

Table 1

Parameters to be Estimated	*Unbiased Estimate*	*Variance*	*Restricted Model Estimate*	*Concomitant Information Vector* (r dimensional)
θ_1	Y_1	V_1	$\dot\theta_1$	z_1
θ_2	Y_2	V_2	$\dot\theta_2$	z_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	
θ_k	Y_k	V_k	$\dot\theta_k$	z_k

Distributions: Descriptive Form

(1) $Y_i \mid \theta_i \sim N(\theta_i, V_i)$, $i = 1, \ldots, k$ independently

(2) $\theta_i \mid \beta, A \overset{\text{indep}}{\sim} N(z_i'\beta, A)$ $i = 1, \ldots, k$

Π: $r + 1$ dimensional

$\alpha = (A, \beta)$ unknown

Inferential Form

(1*) $Y_i \mid \beta, A \overset{\text{ind}}{\sim} N(z_i'\beta, V_i + A)$.

(2*) $\theta_i \mid Y_i, \beta, A \overset{\text{ind}}{\sim} N(\theta_i^*, V_i(1 - B_i))$

$$\theta_i^* = (1 - B_i)Y_i + B_i z_i'\beta$$

$$B_i = V_i/(V_i + A)$$

INTERVAL ESTIMATES:

$\hat\theta_i \pm z s_i$ $\qquad P(\hat\theta_i - z s_i \le \theta_i \le \hat\theta_i + z s_i) \ge 2\Phi(z) - 1$ all $\alpha = (A, \beta)$

Estimate of $\hat\theta_i$

$$\hat\theta_i = (1 - \hat B_i)Y_i + \hat B_i (z_i'\hat\beta).$$

$$\hat\beta = (Z'DZ)^{-1}Z'DY$$

$$\hat B_i = ((k - r - 2)/(k - r))\, V_i/(V_i + \hat A)$$

$$\hat A = \frac{\sum W_i\{(k/(k - r))((Y_i - z_i'\hat\beta)^2 - V_i)\}}{\sum W_i}$$

force $\hat A \ge 0$

$$D = \text{Diag}\,(W_1, \ldots, W_k) \qquad W_i = 1/(V_i + \hat A)$$

Estimate of Variability in $\hat\theta_i$

$$s_i^2 = V_i\{1 - ((k - \hat r_i)/k)\,\hat B_i\} + v_i(Y_i - z_i'\hat\beta)^2$$

$$v_i = (2/(k - r - 2))\,\hat B_i^2\,(\bar V + \hat A)/(V_i + \hat A)$$

$$\hat r_i = kW_i[Z(Z'DZ)^{-1}Z']_{ii}$$

$$\bar V = \sum W_i V_i / \sum W_i$$

MSE if $V_1 = \ldots = V_k$: "EB Minimax", but not (Frequentist) minimax.

$$E(\hat\theta_i - \theta_i)^2 = E\,E_\theta(\hat\theta_i - \theta_i)^2 = V(1 - ((k - r_i - 2)/k\; E\hat B)\;,\quad r_i \equiv k\, z_i'(Z'Z)^{-1}z_i\;,\quad \bar r = r.$$

SAS PROGRAM

```
DO I=1 TO MAXIT
WHILE ((ABS(A-A1)>EPSILON) AND ((A>0) OR (I=1)));
  A1=A;

  W=1#/(V+A);   = Wi
  WT=DIAG(W);   = D
  B=Z*INV(Z`*WT*Z)*Z`;
  MU=B*WT*Y;    = zi'β̂
  A=SUM(W#(((K#/(K-R))#((Y-MU)##2))-V))#/SUM(W);
  PRINT A;
END;
```

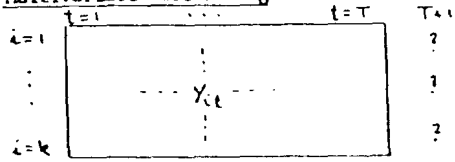
```
IF (A<0) THEN A=0;
B=((K-R-2)#/(K-R))#V#/(V+A);   = B̂i
ESTAR=((1-B)#Y)+(B#MU);        = θ̂i
VBAR=SUM(V#W)#/SUM(W);         = V̄
SMALLV=(2#/(K-R-2))#B#B#(VBAR+A)#/(V+A);   = v
VHALF=SQRT(SMALLV);
SMALLR=K#VECDIAG(B)#/(V+A);    = r̂i
SIGSTAR2=V#(1-(((K-SMALLR)#/K)#B))+(((Y-MU)##2)#SM
SIGSTAR=SQRT(SIGSTAR2);   = si
```

Exchangeable Priors:

(a) Equal Variances (Exposures)

(b) Unequal Variances

Multivariate Forecasting:



Ranking and Selecting the Best: "Ensemble Information"

i) Toxoplasmosis Rates: City 2 worse than City 1.

ii) Was Frank O'Conner better than Ruth and Cobb?

iii) Taxicab Example.

Non-Normal Data: Exchangeable data and Exchangeable priors. (NEF [illegible]F)

$$X_i|\mu_i \overset{\text{ind}}{\sim} \text{NEF}(\mu_i, V(\mu_i)/n), \quad i = 1, \dots, k \qquad \mu_i \overset{\text{ind}}{\sim} CP(\mu_0, \tau_0^2 [illegible] V(\mu_0)/(m [illegible] v_2)), \quad i = 1, \dots, k$$

$$E\mu_i|X_i = (1 - B)X_i + BE\bar{X}$$

$$B = \frac{v_2}{n + v_2}\frac{k - 1}{k} + \frac{n}{n + v_2}E\left\{\frac{V(\bar{X})}{n}\right\}\frac{k - 1}{ES}.$$

$$\hat{B} = \frac{v_2}{n + v_2}\frac{k - 1}{k} + \frac{n}{n + v_2}\hat{B}_{JS} \qquad EV(\bar{X})(k - 1)/ES \approx EV(\bar{X})(k - 3)/S.$$

$\hat{B}_{JS} = (k - 3)V(\bar{X})/nS$, the naive extension of the James Stein shrinking factor.

1. "Data Analysis Using Stein's Estimator and Its Generalizations", Bradley Efron and Carl Morris, Journal of the American Statistical Association, Vol. 70, No. 350, pp. 311-319, June 1975.

2. "Stein's Paradox in Statistics", Bradley Efron and Carl Morris, Scientific American, Vol. 236, No. 5, pp. 119-127, May 1977.

3. "Parametric Empirical Bayes Confidence Intervals", Carl N. Morris, Proceedings of the Conference on Scientific Inference, Data Analysis, and Robustness, Academic Press, 1983 p 25-50.

4. "Parametric Empirical Bayes Inference: Theory and Applications", Carl N. Morris, Journal of the American Statistical Association, Vol. 78, No. 381, March 1983 p 47-65.

5. "NEF - QVF: Statistical Theory" Ann. Statist. 1983 515-529.