BAYESIAN FAIR EXPERIENCE PREMIUM

UNDER LINEAR EXPONENTIAL LIKELIHOODS

AND CONJUGATE PRIORS

By

Zia Rehman

Department of Mathematics
University of Louisville
Louisville, Kentucky

April, 2001

Introduction

In this chapter we will prove the point wise convergence of the Bayesian point estimates under conjugate priors belonging to the 'linear exponential families'. The result below demonstrates that if $E[X|q]=q$, then the Bayesian Point estimates

(i) Have a generic formula that also provides the empirical Bayes point estimate of q .

(ii) Converge point-wise to the sample mean under increasing sample sizes, as do the Maximum Likelihood estimates.

A. Bayesian Experience Premium

Theorem 5-1: For a linear exponential likelihood, having a finite mean and variance, satisfying the relationship $E[(x|q)] = q$ then using a conjugate prior (and squared error loss function) we have

$$\hat{q}_B = \frac{\sum_{i=1}^{i=n} X_i}{n+k} + \frac{k^2 TV}{(n+k)(k+1)} + \frac{ka}{(n+k)}$$

Where TV is the total distribution variance of X, n is the sample size, $a = E(q) - E[Var[(x|q)]]$.

Proof (Case I): $E[Var[(x|q)]] = E[q]$.

We assumed that $E[X|q] = q$ .From the definition of a in the Theorem, it is easy to see that

$$E(q) - E[Var[(x|q)]] = a = 0 .$$

Therefore we need to show (from the Theorem with $a = 0$) that,

$$\hat{q}_B = \frac{\sum_{i=1}^{i=n} X_i}{n+k} + \frac{k^2 TV}{(n+k)(k+1)}$$

In order to prove this relationship, we begin by noting that under a Conjugate Prior, the Bayesian Pure Premium equals the Buhlmann Credibility Estimate of the Pure Premium.

Using conventional notation, Bayesian Pure Premium is the predictive mean. That is ,

$$E[X_{n+1} \mid X = x_1, x_2 ... x_n] = \int m(q) f(q \mid x_1, x_2 ... x_n) dq$$

Where $E[X \mid q] = m(q)$ and $-\infty < q < \infty$

The above expression must then equal the Buhlmann estimate of Pure Premium.

Therefore ,

$$E[X_{n+1} \mid X = x] = \int m(q) f(q \mid x_1, x_2 ... x_n) dq = Z\overline{X} + (1-Z)E[m(q)]$$

By assumption we have we have $m(q) = q$. Also,

$$f(q \mid x_1.x_2 ... x_n) = \frac{f(q)L(q)}{\int f(q)L(q)dq}$$

Therefore

(5-1) $$\frac{\int q f(q)L(q)dq}{\int f(q)L(q)dq} = Z\overline{X} + (1-Z)E[q]$$

For $\hat{q}_B$, we find the $E[q]$ first. Use the credibility factor, Z, in the Buhlmann formula, $Z = \dfrac{n}{n+k}$ where k is given by,

(5-2) $$k = \frac{E[Var(X|q)]}{Var[(E(X|q)]} = \frac{E[q]}{Var[q]}$$

Also since Total Variance = $E[Var(X|q)] + Var[E(X|q)]$ = EVPV + VHM we have for a Poisson Likelihood,

$$\text{Total Variance } = Var(X) = E[q] + Var[q]$$

Thus,

(5-3) $$Var[q] = Var(X) - E[q]$$

So that,

(5-4) $$k = \frac{E[q]}{TV - E[q]}$$

Substituting it in the expression for Z and solving for $E[q]$ yields,

(5-5) $$E[q] = \frac{nTV[1-Z]}{n[1-Z]+Z}$$

Finally we have from (5-1) and (5-5) and the fact that the support of q is on $[0, \infty)$,

$$\frac{\int_{-\infty}^{\infty} q f(q) L(q) dq}{\int_{-\infty}^{\infty} f(q) L(q) dq} = Z\overline{X} + [1-Z] \frac{nTV[1-Z]}{n[1-Z]+Z}$$

Now the left hand side is clearly the Bayesian Posterior Mean which is equal to the Bayesian Point Estimate = $\hat{q}_B$ (under the squared error loss minimization criteria). Thus,

$$\hat{q}_B = Z\overline{X} + [1-Z] \frac{nTV[1-Z]}{n[1-Z]+Z}$$

But $Z = \dfrac{n}{n+k}$. Rearranging and plugging yields,

(5-6) $$\hat{q}_B = \frac{\sum_{i=1}^{i=n} X_i}{n+k} + \left[\frac{k^2 TV}{(n+k)(k+1)}\right]$$

Case II: $E[Var[(x|q)]] < E[q]$.

We assumed that $E[(x|q)] = q$. From the definition of a in Theorem (5-1) it follows that

$$E[Var[(x|q)]] = E[q] - a.$$

As in (5-1),

(5-7) $$\frac{\int q f(q) L(q) dq}{\int f(q) L(q) dq} = Z\overline{X} + (1-Z)E[q]$$

For $\hat{q}_B$, we find the $E[q]$ first. Use the credibility factor, Z, in the Buhlmann formula,

$Z = \frac{n}{n+k}$ where k is given by,

(5-8) $$k = \frac{E[Var(X|q)]}{Var[E(X|q)]} = \frac{E[q]-a}{Var[q]}$$

Also since Total Variance = $E[Var(X|q)] + Var[E(X|q)]$ = EVPV + VHM we have,

(5-9) Total Variance $= TV = E[q] + Var[q] - a$

Thus,

(5-10) $Var[q] = TV - E[q] + a$

So that,

(5-11) $$k = \frac{E[q]-a}{TV - E[q] + a}$$

Substituting it in the expression for Z and solving for $E[q]$ yields,

(5-12) $$E[q] = \frac{nTV[1-Z]}{n[1-Z]+Z} + \frac{a(n+Z-nZ)}{n[1-Z]+Z}$$

Finally from (5-7) we have,

$$\frac{\int_{-\infty}^{\infty} q f(q) L(q) dq}{\int_{-\infty}^{\infty} f(q) L(q) dq} = Z\overline{X} + [1-Z]\frac{nTV[1-Z]}{n[1-Z]+Z} + [1-Z]\frac{a(n+Z-nZ)}{n[1-Z]+Z}$$

Thus the left hand side is clearly the Bayesian Posterior Mean which is equal to the Bayesian Point Estimate = $\hat{q}_B$. Thus,

(5-13) $$\hat{q}_B = Z\overline{X} + [1-Z]\frac{nTV[1-Z]}{n[1-Z]+Z} + [1-Z]\frac{a(n+Z-nZ)}{n[1-Z]+Z}$$

But $Z = \frac{n}{n+k}$. Rearranging and plugging yields,

(5-14) $$\hat{q}_B = \frac{\sum_{i=1}^{i=n} X_i}{n+k} + \frac{k^2 TV}{(n+k)(k+1)} + \frac{ka}{(n+k)}$$

This completes the proof of Case II.

Case III: $E[Var[(x|q)]] > E[q]$.

We assumed that $E[(X|q)] = q$. Let the $E[Var[(x|q)]] = E[q] + b$ for some real number, $b > 0$. Then from the definition of a in Theorem (5-1) we have

$$E[Var[(x|q)]] - E(q) = b = -a$$

As in (5-1),

(5-15) $$\frac{\int q f(q) L(q) dq}{\int f(q) L(q) dq} = Z\overline{X} + (1-Z)E[q]$$

For $\hat{q}_B$, we find the $E[q]$ first. Use the credibility factor, Z, in the Buhlmann formula, $Z = \dfrac{n}{n+k}$ where k is given by,

(5-16) $$k = \frac{E[q] + b}{Var[q]}$$

Also since Total Variance = $E[Var(X|q)] + Var[E(X|q)]$ = EVPV + VHM we have,

(5-17) Total Variance $= TV = E[q] + b + Var[q]$

Thus,

(5-18) $Var[q] = TV - E[q] - b$

So that,

(5-19) $k = \dfrac{E[q] + b}{TV - E[q] - b}$

Substituting it in the expression for Z and solving for $E[q]$ yields,

(5-20) $E[q] = \dfrac{nTV[1-Z]}{n[1-Z]+Z} - \dfrac{b(n+Z-nZ)}{n[1-Z]+Z}$

Finally we have,

$$\frac{\int_{-\infty}^{\infty} q f(q) L(q) dq}{\int_{-\infty}^{\infty} f(q) L(q) dq} = Z\overline{X} + [1-Z]\,\frac{nTV[1-Z]}{n[1-Z]+Z} - [1-Z]\frac{b(n+Z-nZ)}{n[1-Z]+Z}$$

Thus the left hand side is recognized as the Bayesian Posterior Mean which is equal to the Bayesian Point Estimate = $\hat{q}_B$. Thus,

(5-21) $\hat{q}_B = Z\overline{X} + [1-Z]\,\dfrac{nTV[1-Z]}{n[1-Z]+Z} - [1-Z]\dfrac{b(n+Z-nZ)}{n[1-Z]+Z}$

But $Z = \dfrac{n}{n+k}$. Further, $b = -a$ .Rearranging and plugging yields,

$$(5\text{-}22)\quad \hat{q}_B = \frac{\sum_{i=1}^{i=n} X_i}{n+k} + \frac{k^2 TV}{(n+k)(k+1)} + \frac{ka}{(n+k)}$$

This completes the proof of Case III.

Corollary 5-2: We have for a Poisson- Gamma case,

$$(5\text{-}23)\quad \hat{q}_B = \frac{\sum_{i=1}^{i=n} X_i}{n+b} + \left[\frac{b^2 TV}{(n+b)(b+1)}\right] = \frac{a+n\overline{x}}{b+n}$$

Proof: We note that for a Gamma (a, b) prior we have $k = b$ [Reference Credibility Theory, Herzog]. Also the total distribution variance given by (5-3) above is

$$TV = E[Var(X|q)] + Var[E(X|q)] = E(q) + Var(q) = \frac{a}{b^2} + \frac{a}{b}$$

Plugging this into (5-1) with $a = 0$ (since Poisson likelihood satisfies the requirements of Case I) gives

$$\hat{q}_B = \frac{\sum_{i=1}^{i=n} X_i}{n+k} + \frac{k^2 TV}{(n+k)(k+1)} = \frac{n\bar{x}}{n+b} + \frac{b^2\left\{\frac{a}{b}+\frac{a}{b^2}\right\}}{(n+b)(b+1)} = \frac{a+n\bar{x}}{b+n}$$

Corollary 5-4: The Bayesian and the Maximum Likelihood variances are related as under (restricting ourselves to cases where $\hat{q}_{ML} = \bar{x}$ under the assumption that $E[(x|q)] = q$):

(5-24) $Var(\hat{q}_B) = Z^2 Var(\hat{q}_{ML})$.

Thus the Bayesian estimate can be regarded as a superior estimate since (5-25 implies that $Var(\hat{q}_B) \leq Var(\hat{q}_{ML})$ as $0 \leq Z \leq 1$.

Proof: We note from the formula of Theorem (5-1) that the only term containing X is the first term. The rest will therefore be zero when the variance operator is applied to them. We rewrite the formula with the observation that $Z = \frac{n}{n+k}$ as under

$$Var(\hat{q}_B) = Var\left[\frac{\sum_{i=1}^{i=n} X_i}{n+k}\right] = Var[Z\bar{x}] = Z^2 Var[\bar{x}].$$

$Var(\hat{q}_{ML}) = Var[\bar{x}]$. Thus

$Var(\hat{q}_B) = Z^2 Var(\hat{q}_{ML})$.

Since $0 \le Z \le 1$, we see that $Var(\hat{q}_B) \le Var(\hat{q}_{ML})$.

Corollary 5-5: $\hat{q}_B$ and $\hat{q}_{ML}$ converge point-wise to the same value as $n \to \infty$ for cases in which $\hat{q}_{ML} = \bar{x}$ (under the assumption that $E[(x|q)] = q$). Interestingly, the Bayesian point estimate converges to $\bar{x}$ for <u>every case</u> of the class of distributions we have considered.

We also restrict ourselves to cases where $\hat{q}_{ML} = \bar{x}$ under the assumption that $E[(x|q)] = q$.

Proof: We note that Actual TV = Constant. Also k does not depend on n. Thus in the limit,

(5-25) $$\underset{n\to\infty}{Limit}\hat{q}_B = \underset{n\to\infty}{Limit}\left[\frac{\sum_{i=1}^{i=n} X_i}{n+k}\right] + \underset{n\to\infty}{Limit}\left[\frac{k^2 TV}{(n+k)(k+1)}\right] + \left[\frac{ka}{(n+k)}\right] = \bar{x}$$

Example 11: From the foregoing Theorem, we draw the attention of the reader to the Normal-Normal case. Since Normal is a conjugate prior of the Normal Likelihood, we have using the mean of the likelihood as the parameter,

$$\underset{n\to\infty}{Limit}\hat{q}_B = \bar{x} = \underset{n\to\infty}{Limit}\hat{q}_{ML}$$

This result can be confirmed independently using the fact the posterior is Normal with mean

$$\hat{m} = \frac{ms_1^2 + s_2^2 n\bar{x}}{s_1^2 + ns_2^2}$$ [Reference Credibility Theory, Herzog, 132]

For $n \to \infty$ the $\hat{m} \to \bar{x}$ as predicted by our theorem.

C. Conclusion

Before moving on to explain the use of the Theorem as an empirical Bayesian estimate of q, we must point out the restrictions imposed by Theorem (5-1). These are as under:

(a) The underlying distribution should be linearly exponential

(b) The prior must be conjugate with a finite mean and variance

(c) $E[X|q] = q$

These three facts must be known before one attempts to use Theorem (5-1) to empirically find $\hat{q}_B$. Once known the following are the reasons of using Theorem (5-1):

(1) Since the analyst need not know the shape of the prior and the underlying distribution, he is being allowed by the Theorem to explore a <u>whole class of distributions and conjugate priors</u> with a single use of the formula. That is whether the true scenario was Poisson-Gamma or Bernoulli-Beta or Normal-Normal, the formula will adjust itself according to the empirically derived value of a (using Empirical Bayesian Methods) and yield the true value of $\hat{q}_B$. Clearly then, from a computational point of view, the utility of the formula is noteworthy.

(2) We have already established in corollary (5-4) that the Bayesian Point estimate is superior to the Maximum Likelihood estimate. That is

$$Var(\hat{q}_B) = Z^2 Var(\hat{q}_{ML}) .$$

For lower credibility situations ($0 \leq Z \leq 1$) then, it is advisable to use Theorem (5-1) to arrive at the Bayesian based estimate of Pure Premium since $Z^2 < 1$ in this situation and therefore $Var(\hat{q}_B) < Var(\hat{q}_{ML})$.

(3) As explained above, no further knowledge (beyond that assumed in (a), (b), (c) above) is needed about the underlying distribution or even the prior to find $\hat{q}_B$. One merely needs to apply simple Empirical Bayesian Procedures to find $\hat{q}_B$.

(4) It is the only formula which allows empirical Bayesian estimation of q under the assumptions outlined above.