Bayesian Point Estimation

- General Approach:

For a vector of observations, $\mathbf{X} = X_1, X_2, ..., X_n$, a Bayes decision rule, $d^*(\mathbf{X})$, is any function d that minimizes the expected risk $E_\theta\left[E_{\mathbf{X}|\theta}\left[L(d,\theta)|\theta\right]\right]$ over all d, where $L(d,\theta)$ is the loss function that describes the relative penalty to be paid for using the estimate d when the actual value of the parameter is θ.

Note that

$$E_\theta\left[E_{\mathbf{X}|\theta}\left[L(d,\theta)|\theta\right]\right] = \int_\Theta \int_{\mathbf{X}} \left[L(d,\theta) f(\mathbf{X}|\theta) d\mathbf{X}\right] \pi(\theta) d\theta .$$

If $L(d,\theta)$ is non-negative, then it can be shown that the value of d that minimizes this expression is equivalent to the value of d that minimizes

$$E_{\theta|\mathbf{X}}\left[L(d,\theta)|\mathbf{X}\right] = \int_\Theta L(d,\theta)\pi(\theta|\mathbf{X}) d\theta ,$$

where $\pi(\theta|\mathbf{X}) = \dfrac{f(\mathbf{X}|\theta)\pi(\theta)}{\int_\Theta f(\mathbf{X}|\theta)\pi(\theta) d\theta}$.

- Commonly Used Loss Functions:

The three most commonly used loss functions (and their associated Bayes decisions) are:

(1) Squared-error loss, $L(d,\theta) = (d-\theta)^2$; this yields the Bayes decision $d^*(\mathbf{X}) = E_{\theta|\mathbf{X}}[\theta|\mathbf{X}]$, the posterior mean.

(2) Absolute loss, $L(d,\theta) = |d-\theta|$; this yields the Bayes decision $d^*(\mathbf{X})$ = the posterior median.

(3) "All-or-nothing" loss, $L(d,\theta) = \begin{cases} 0 \text{ if } d = \theta \\ 1 \text{ otherwise} \end{cases}$; this yields the Bayes decision $d^*(\mathbf{X})$ = the posterior mode.

- Admissibility:

  A Bayes decision rule $d^*(\mathbf{X})$ is <u>admissible</u> if and only if

$$E_{\mathbf{X}|\theta}\left[L(d^*,\theta)|\theta\right] \le E_{\mathbf{X}|\theta}\left[L(d,\theta)|\theta\right]$$

  for all parameter values θ, for all competing decision rules d.

  Result: If a Bayes decision rule is unique, then it is admissible.

Bayesian Interval Estimates

- For a given probability, $1-\alpha$ (where α often equals .10, .05, or .01), find values a_1 and a_2 such that

$$\Pr\left\{a_1 \le \theta \le a_2\right\} = 1-\alpha ,$$

  using the posterior distribution of θ given the vector of observations $\mathbf{X}$.

  Note that the values a_1 and a_2 will generally not be unique.

Bayesian Credibility

- Let

$$X_i|\theta \sim f\left(x|\mu(\theta),\sigma^2(\theta)\right)$$

  denote the i^{th} observation of a claim count (frequency), claim amount (severity), or total claim cost (pure premium), where:

  $\mu(\theta) = E_{X|\theta}\left[X_i|\theta\right]$, the hypothetical mean (HM), and
  $\sigma^2(\theta) = Var_{X|\theta}\left[X_i|\theta\right]$, the process variance (PV).

- Let

$$\theta \sim \pi(\theta)$$

  denote a "state of nature" parameter, where:

  $E_\theta\left[\mu(\theta)\right] =$ the prior mean (H),
  $Var_\theta\left[\mu(\theta)\right] =$ the variance of the hypothetical means (VHM),
  $E_\theta\left[\sigma^2(\theta)\right] =$ the expected value of the process variance ($EVPV$), and
  $R = \overline{X}$, a recent observation.

- The Bayesian posterior mean,

$$E_{\theta|X}\left[\mu(\theta)|X_1, X_2, \ldots, X_n\right]$$

  minimizes the expected loss when the loss function is quadratic.

  The posterior mean is sometimes called the Bayesian credibility estimator.