Expected value of more precise/context-relevant calibration target data

Introduction

Calibration target data is one of the key elements of any calibration process. However, like any data, calibration targets are prone to precision and/or validity issues.

The Standard error (SE) is a measure of the precision of a statistical estimate. Specifically, it quantifies the variability or imprecision of a sample statistic, such as the mean or proportion, compared to the true population parameter. We quantify the Expected Value of Perfect Information (EVPI) as the value of reducing uncertainty in decision-making. The EVPI represents the maximum amount a decision-maker would be willing to pay for perfect information that eliminates all uncertainty associated with a decision, reducing the SE to zero.

Bias, on the other hand, refers to the systematic or consistent error in measurements or estimates. It can lead to inaccuracies in the results, and it can result from flaws in study design, data collection methods, or other systematic errors.

Theoretical derivations

The problem

Prior

$$x \sim N(\mu, \Sigma)$$

where

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

The probability density function (PDF) for a multivariate normal distribution with a mean vector μ and a covariance matrix Σ is given by:

$$f(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)$$

Measurement

$$y = L^T x + \varepsilon$$

where

- L is a matrix that transforms the vector x into another vector (or scalar, depending on the context). The transformation is defined by the structure and values of L. L^T represents the transposition of the matrix L. Transposing a matrix involves flipping it over its diagonal, turning the matrix's rows into columns and vice versa.

In the current example: NB $= x_1$ and Mortality $= x_1 + x_2$. If $y =$ NB, $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $L = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $L^T = \begin{bmatrix} 1 & 0 \end{bmatrix}$.

- ε is the measurement error $\varepsilon \sim N(0, s_N^2)$.

The likelihood is given by:

$$L(x, s_N^2; y) = \frac{1}{\sqrt{2\pi s_N^2}} \exp\left(-\frac{(y - L^T x)^2}{2 s_N^2}\right)$$

$$L(x, s_N^2 | y_1, y_2, \ldots, y_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi s_N^2}} \exp\left(-\frac{(y_i - L^T x)^2}{2 s_N^2}\right)$$

Decision [Mark ??]

The decision is to use intervention if $x_1 > 0$ and keep the status quo otherwise.

The incremental net benefit is [Mark ??]

$$NB = y = x_1$$
$$INB(x) = x_1 . \theta(x_1)$$

where

$$\theta(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Reduced normals

To calculate the distribution of x conditional on observing measurement y, $P(x|y)$, note

$$P(x \& y) \propto \exp[-Q(x_1, x_2)/2]$$

Given the function

$$Q = \frac{(L^T x - y)^2}{s_N^2} + (x - \mu)^T \Sigma^{-1} (x - \mu)$$

and the assertion that

$$Q = (x - a)^T M^{-1} (x - a)$$

our goal is to derive expressions for M and a.

Expanding Q

First, we expand both parts of Q:

$$\frac{(L^T x - y)^2}{s_N^2} = \frac{(L^T x)^T (L^T x) - 2y L^T x + y^2}{s_N^2}$$

$$(x - \mu)^T \Sigma^{-1}(x - \mu) = x^T \Sigma^{-1} x - x^T \Sigma^{-1} \mu - \mu^T \Sigma^{-1} x + \mu^T \Sigma^{-1} \mu$$

Combining and Rearranging

Next, we combine these expansions and rearrange terms to form a quadratic expression in x:

$$Q = \frac{(L^T x)^T (L^T x)}{s_N^2} - \frac{2y L^T x}{s_N^2} + \frac{y^2}{s_N^2} + x^T \Sigma^{-1} x - 2x^T \Sigma^{-1} \mu + \mu^T \Sigma^{-1} \mu$$

The goal is to rewrite Q in the form of $(x - a)^T M^{-1}(x - a)$. This form is a general quadratic expression which can be expanded as:

$$(x - a)^T M^{-1}(x - a) = x^T M^{-1} x - x^T M^{-1} a - a^T M^{-1} x + a^T M^{-1} a$$

Comparing Coefficients

By comparing coefficients from the expanded form of Q and the general quadratic form, we can identify M^{-1} and a:

$$\text{For the quadratic term in } x(x^T M^{-1} x):$$

$$\frac{LL^T}{s_N^2} + \Sigma^{-1} = M^{-1}$$

$$\therefore \ M^{-1} = \Sigma^{-1} + \frac{LL^T}{s_N^2}$$

$$\text{For the linear term in } x(x^T M^{-1} a):$$

$$\Sigma^{-1} \mu + \frac{Ly}{s_N^2} = M^{-1} a$$

$$\text{Multiplying both sides by } M \text{ gives:}$$

$$\therefore \ a = M \left(\Sigma^{-1} \mu + \frac{Ly}{s_N^2} \right)$$

Posterior hyperparameters

So, the posterior distribution of the NB is:

$$P(x|y) = N(x; a, M)$$

$$where:$$

$$M^{-1} = \Sigma^{-1} + \frac{LL^T}{s_N^2}$$

$$a = M\left(\Sigma^{-1}\mu + \frac{Ly}{s_N^2}\right)$$

A useful integral}

Using integration by parts:

$$\int_0^\infty dx\, x.N(x; m, s^2) = s^2.N(0, m, s^2) + m.\left[1 - \Phi\left(-\frac{m}{s}\right)\right] \stackrel{\text{def}}{=} f(m, s)$$

where $N(x; m, s^2)$ is the density of a normal distribution and $\Phi(x)$ is the cdf of the standard normal distribution, and the last equation serves to define the function f.

EVPI:

The Standard error (SE) measures the precision of a statistical estimate. Specifically, it quantifies the variability or imprecision of a sample statistic, such as the mean or proportion, compared to the true population parameter.

In a set of competing options, D, the optimal choice for decision-makers like NICE is that associated with the highest "utility". The net benefit (NB) is a utility function given by the following notation for each choice or option d.

$$NB_d(\theta)$$

$\theta = (\theta_1, \dots, \theta_n)$ is the model input parameters representing real-world quantities. Since observed data (parameters) may not be known with certainty, we are uncertain about the value(s) of θ and, therefore, uncertain about the NB. Because of the uncertainty associated with the NB, the optimal choice is determined based on maximising the *expected NB*.

$$\mathbb{E}\{NB_d(\theta)\}$$

Which for the set of options, D, becomes:

$$\max_{d \in D} \mathbb{E}\{NB_d(\theta)\}$$

We quantify the Expected Value of Perfect Information (EVPI) as the value of reducing uncertainty in decision-making. The EVPI represents the maximum amount a decision-maker would be willing to pay for perfect information that eliminates all uncertainty associated with their decision, reducing the SE to zero. The EVPI is given by:

$$EVPI = \mathbb{E}\{\max_{d \in D} NB_d(\theta)\} - \max_{d \in D} \mathbb{E}\{NB_d(\theta)\}$$

In the current example we have one choice d_1 where the $NB = y = x_1$ and the NB posterior distribution, and its moments, is given by:

$$P(x|y) = \quad N(x; a, M)$$

$$where:$$

$$M^{-1} = \quad \Sigma^{-1} + \frac{LL^T}{s_N^2}$$

$$a = \quad M\left(\Sigma^{-1}\mu + \frac{Ly}{s_N^2}\right)$$

To estimate the EVPI, we assume a second choice d_2 where the $NB = 0$. Hence the EVPI becomes:

$$\begin{aligned} EVPI \quad &= \mathbb{E}\{\max(P(x|y), 0)\} - \max(\mathbb{E}\{P(x|y)\}, 0) \\ &= \mathbb{E}\{\max(P(x|y), 0)\} - \max(a, 0) \end{aligned}$$

Methods

Inputs

We use a simple decision model with two inputs, x_1 and x_2.

```
# Define model inputs:
m1 <- 0
m2 <- 190
mu <- c(m1, m2)

n_observed <- 100
L <- matrix(rep(c(2, 0), n_observed), nrow=n_observed, byrow=TRUE) #
Transformation matrix L for n observations
s_N_squared <- 2 # Variance of the measurement error, assuming identical for
all observations
y <- rnorm(n_observed, 0, sqrt(s_N_squared)) # Example: n observed values
generated around 2 with noise

v1 <- 1000
v2 <- 100
rho <- 0
Sigma <- matrix(c(v1, rho * v1 * v2, rho * v1 * v2, v2), 2, 2)

# Calculate the inverse of Sigma
Sigma_inv <- solve(Sigma)

set.seed(1)
n <- 1e6
x <- MASS::mvrnorm(n, mu, Sigma)
x1 <- x[, 1]
x2 <- x[, 2]
```

Getting the hyperparameters
```
# Calculation of M^-1 and a requires adjustment for multiple observations
# Adjust L and y to match the dimensions
```

```r
M_inv <- Sigma_inv + (t(L) %*% L) / s_N_squared
M <- solve(M_inv)

a <- M %*% ((Sigma_inv %*% mu) + ((t(L) %*% y) / s_N_squared))

# Print the results
print("Hyperparameter a:")

[1] "Hyperparameter a:"

print(a)

            [,1]
[1,]    0.02356392
[2,] 190.00000000

print("Covariance matrix M:")

[1] "Covariance matrix M:"

print(M)

            [,1] [,2]
[1,] 0.004999975    0
[2,] 0.000000000  100

# Sample from the posterior
hyper_params_samples <- MASS::mvrnorm(n, a, M)[,1]

# Get the EVPI:
hyper_params_evpi <- mean(pmax(hyper_params_samples, 0)) -
pmax(mean(hyper_params_samples), 0)

print("Hyper parameters EVPI:")

[1] "Hyper parameters EVPI:"

print(hyper_params_evpi)

[1] 0.01798918
```