

Question 1

Jin Barai

(a)

The range is **location invariant**. As per the definition, an attribute is location invariant if

$$a(y_1 + b, \dots, y_N + b) = a(y_1, \dots, y_N)$$

Since, the difference between the maximum value and the minimum value in the population would remain the same even if we add the same value to all the units in the population.

(b)

The range is **scale equivariant**. As per the definition, an attribute is scale equivariant if

$$a(y_1 * m, \dots, y_N * m) = m * a(y_1, \dots, y_N)$$

Clearly, by multiplying all the units in the population with a certain number m, the difference between the range would increase by m.

(c)

The range is **neither**. As per the definition, an attribute is location-scale equivariant if

$$a(m * y_1 + b, \dots, m * y_N + b) = (m * a(y_1, \dots, y_N)) + b$$

As per the definition, an attribute is location-scale invariant if

$$a(m * y_1 + b, \dots, m * y_N + b) = a(y_1, \dots, y_N)$$

This does not hold true for range, and hence we conclude that it is neither location-scale variant nor location-scale invariant.

(d)

The range is **replication invariant**. If the population duplicates k-1 times (such that there are k copies). As per the definition, an attribute is replication invariant if

$$a(P^k) = a(P)$$

And evidently the maximum and the minimum in the population don't change, and hence the difference doesn't change either. So we it is replication invariant.

(e)

Given

$$P = \{y_1, y_2, \dots, y_{N-1}\}$$

For any variate y ,

$$P^* = \{y_1, y_2, \dots, y_{N-1}, y\}$$

We know that Sensitivity curve can be given by,

$$SC(y) = N(a(P^*) - a(P))$$

Thus

$$a(P^*) = \begin{cases} y_{(N-1)} - y & y \leq y_1 \\ y_{(N-1)} - y_1 & 0 \leq y \leq y_{(N-1)} \\ y - y_1 & y_{(N-1)} \leq y \end{cases}$$
$$SC(y) = \begin{cases} N(y_1 - y) & y \leq y_1 \\ 0 & 0 \leq y \leq y_{(N-1)} \\ N(y - y_{(N-1)}) & y_{(N-1)} \leq y \end{cases}$$

(f)

```
set.seed(341)
pop <- rnorm(10000)

sc = function(y.pop, y, attr, ...) {
  N <- length(y.pop) + 1
  sapply(y, function(y.new) {
    N * (attr(c(y.new, y.pop), ...) - attr(y.pop, ...))
  })
}

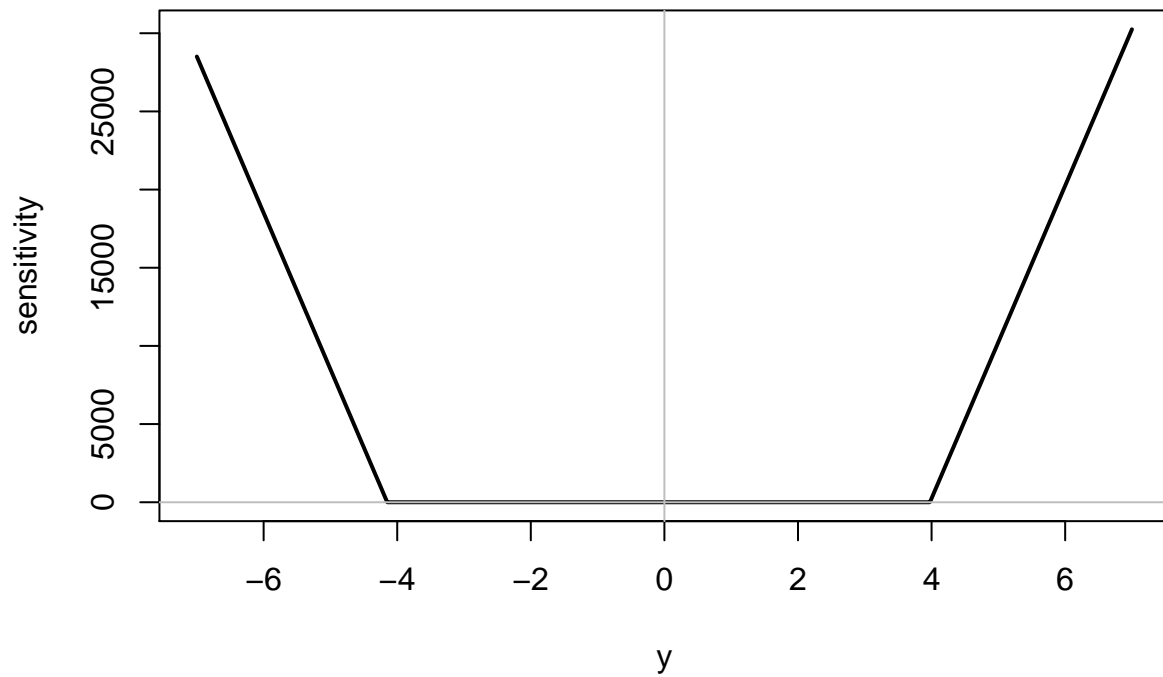
spread <- function (pop) {
  (max(pop) - min(pop))
}

y <- seq(-7, 7, by = 0.01)

plot(y,
      sc(y.pop = pop, y = y, attr = spread),
      type = "l",
      lwd = 2,
      main = "Sensitivity curve for Spread",
      ylab = "sensitivity")

abline(h = 0, v = 0, col = "grey")
```

Sensitivity curve for Spread



(g)

Good Thing: It is location invariant hence adding any values to the units in the population will still give us the same answer.

Bad Thing: If you observe the Sensitivity Curve, you can clearly see that for extreme observations the spread value will completely change.