Ratio Statistics

This procedure provides a variety of descriptive statistics for the ratio of two variables.

Notations

The following notation is used throughout this chapter unless otherwise stated:

n	Number of observations
A_i	Numerator of the *I*-th ratio (*i* = 1, …, *n*). This is usually the appraisal roll value.
S_i	Denominator of the *i*-th ratio (*i* = 1, …, *n*). This is usually the sale price.
R_i	The *i*-th ratio (*i* = 1, …, *n*). Often called the appraisal ratio.
f_i	Case weight associated with the *i*-th ratio (*i* = 1, …, *n*).

Data

This procedure requires for *i* = 1, …, *n* that:

- $A_i > 0$,
- $S_i > 0$,
- $f_i > 0$, and
- w_i is a whole number. If the SPSS Weight variable contains fractional values, then only the integral parts are used.

A case is considered valid if it satisfies all four requirements above. This procedure will use only valid cases in computing the requested statistics.

Ratio Statistics

Ratio

$$R_i = \frac{A_i}{S_i}, \quad i = 1, \ldots, n$$

Minimum

The smallest ratio and is denoted by $R_{\min}$.

Maximum

The largest ratio and is denoted by $R_{\max}$.

Range

The difference between the largest and the smallest ratios. It is equal to $R_{\max} - R_{\min}$.

Median

The middle number of the sorted ratios if n is odd. The mean (average) of the two middle ratios if the n is even. The median is denoted as $\tilde{R}$.

Average Absolute Deviation (AAD)

$$AAD = \sum_{i=1}^{n} f_i \left| R_i - \tilde{R} \right| \Big/ \sum_{i=1}^{n} f_i$$

Coefficient of Dispersion (COD)

$$COD = 100\% \times \frac{AAD}{\tilde{R}}$$

Coefficient of Concentration (COC)

Given a percentage 100% × g, the coefficient of concentration is the percentage of ratios falling within the interval $\left[(1-g)\tilde{R},\ \ (1+g)\tilde{R}\right]$. The higher this coefficient, the better uniformity.

Mean

$$\overline{A/S} = \overline{R} = \sum_{i=1}^{n} f_i R_i \Big/ \sum_{i=1}^{n} f_i$$

Standard Deviation (SD)

$$s = \sqrt{\frac{1}{(F-1)} \sum_{i=1}^{n} f_i \left(R_i - \overline{R}\right)^2}$$

where $F = \sum_{i=1}^{n} f_i$.

Coefficient of Variation (COV)

$$COV = 100\% \times \frac{s}{\overline{R}}$$

Weighted Mean

$$\overline{A/S} = \frac{\sum_{i=1}^{n} f_i A_i}{\sum_{i=1}^{n} f_i S_i} = \frac{\sum_{i=1}^{n} f_i S_i R_i}{\sum_{i=1}^{n} f_i S_i}$$

This is the weighted mean of the ratios weighted by the sales prices in addition to the usual case weights.

Price Related Differential (a.k.a. Index of Regressivity)

$$PRD = \frac{\overline{A/S}}{\overline{A}/\overline{S}}$$

This is quotient by dividing the Mean by the Weighted Mean.

Property appraisals sometimes result in unequal tax burden between high-value and low-value properties in the same property group. Appraisals are considered *regressive* if high-value properties are under-appraised relative to low-value properties. On the contrary, appraisals are considered *progressive* if high-value properties are relatively over-appraised. The price related differential is a measure for measuring assessment regressivity or progressivity. Hence the price related differential is also known as the index of regressivity.

Recall that the [unweighted] mean weights the ratios equally, whereas the weighted mean high-value properties are under-appraised, thus pulling the weighted mean below the mean. On the other hand, if the PRD is less than 1, high-value properties are relatively over-appraised, pulling the weighted mean above the mean.

Confidence Interval for the Median

Distribution Free

Given the confidence level $100\% \times (1-\alpha)$, the confidence interval for the median is an interval $(R_{[r]}, R_{[n-r+1]})$ such that

$$1-\alpha = 1 - 2\, I_{0.5}(n-r+1, r) = \frac{1}{2^n} \sum_{k=r}^{n-r} \binom{n}{k},$$

where $R_{[k]}$ is the $100\% \times k/n$ quantile, and $I_{0.5}(n-r+1, r)$ is the incomplete Beta function.

An equivalent formula is

$$\frac{\alpha}{2} = I_{0.5}(n-r+1, r) = \frac{1}{2^n}\sum_{k=0}^{r-1}\binom{n}{k}.$$

Since the rightmost term is the cumulative Binomial distribution and it is discrete, r is solved as the largest value such that

$$\frac{\alpha}{2} \le \frac{1}{2^n}\sum_{k=0}^{r-1}\binom{n}{k}.$$

Thus the confidence interval has coverage probability of at least $1-\alpha$.

Normal Distribution

Assuming the ratios follow a normal distribution, a two-sided $100\% \times (1-\alpha)$ confidence interval for the median of a normal distribution is

$$\left(\overline{R} + g_{(\alpha/2;0.5,d)} \times s,\quad \overline{R} + g_{(1-\alpha/2;0.5,d)} \times s\right)$$

where $g_{(\gamma;p,d)}$ are values defined in Table 1 of Odeh and Owen (1980).

The value $g_{(\gamma;p,d)}$ is, in fact, the solution to the following equations:

$$\Pr\left(T_d \le g\sqrt{n} \mid \delta = K_p\sqrt{n}\right) = \gamma$$

with T_d follows a noncentral Student *t*-distribution where d is degrees of freedom associated with the standard deviation s, δ is noncentrality parameter, γ is the probability, n is the sample size, and K_p is the upper p percentile point of a standard normal distribution.

Confidence Interval for the Mean

The normal distribution is used to approximate the distribution of the ratios. The $100\% \times (1-\alpha)$ confidence interval for the mean is:

$$\overline{R} \pm t_{\alpha/2;F-1} \times s \times \sqrt{\sum_{i=1}^{n} f_i^{\,2}} \Big/ F$$

where $t_{\alpha/2;F-1}$ is the upper $\alpha/2$ percentage point of the t distribution with $F-1$ degrees of freedom, and where $F = \sum_{i=1}^{n} f_i$.

Confidence Interval for the Weighted Mean

Using the Delta method, variance of the weighted mean is approximated as

$$\operatorname{var}\left(\frac{\overline{A}}{\overline{S}}\right) \approx \frac{\operatorname{var}(\overline{A})}{\overline{S}^2} - \frac{2\overline{A}\operatorname{cov}(\overline{A},\overline{S})}{\overline{S}^3} + \frac{\overline{A}^2 \operatorname{var}(\overline{S})}{\overline{S}^4}.$$

where

$$\operatorname{var}(\overline{A}) = \frac{1}{(F-1)}\sum_{i=1}^{n} f_i \left(A_i - \overline{A}\right)^2 \times \sum_{i=1}^{n} f_i^2 \Big/ F^2 \text{ ,}$$

$$\operatorname{var}(\overline{S}) = \frac{1}{(F-1)}\sum_{i=1}^{n} f_i \left(S_i - \overline{S}\right)^2 \times \sum_{i=1}^{n} f_i^2 \Big/ F^2 \text{ , and}$$

$$\operatorname{cov}(\overline{A},\overline{S}) = \frac{1}{(F-1)}\sum_{i=1}^{n} f_i \left(A_i - \overline{A}\right)\left(S_i - \overline{S}\right) \times \sum_{i=1}^{n} f_i^2 \Big/ F^2 \text{ .}$$

References

International Association of Assessing Officers (1990). *Property Appraisal and Assessment Administration*. International Association of Assessing Officers: Chicago, Illinois.

Odeh, Robert E., and Owen, D. B. (1980). *Tables for Normal Tolerance Limits, Sampling Plans, and Screening*. Marcel Dekker: New York.