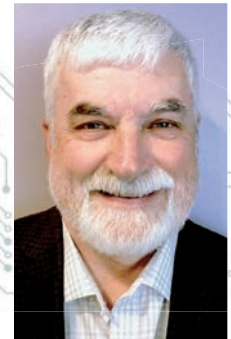




Using probability distribution and variance

to select the best overall capitalization rate of a property

By George Canning, AACI, P.App,
Canning Consultants Inc., London, ON



Introduction

This article is not about how to calculate an overall capitalization rate (OCR), but rather how to 'adjust' for an OCR range when there is a considerable spread in the 'going in' OCRs derived from a series of comparable sales. Even though the valuer has researched the sales thoroughly, often there is no reasonable explanation as to why one property has a higher or lower OCR than others.

Sometimes, the answer is clear in that a specific income property had a low OCR because the rents were coming due and there is an obvious 'upswing' in the income. In this case, the buyer is prepared to pay a premium for the property to secure those future rents. Perhaps there is a change in the economy, such as rising interest rates which, in turn, put more risk onto the investment. The buyer wants compensation for this fact by offering a price that is commensurate with a higher OCR.

Perhaps the expenses of a given property that are used to generate the OCR are not deemed reliable because there is too much variation in them. A case in point involves trailer parks and marinas. Expenses can vary immensely, even though the sizes of the marinas and campgrounds are similar. This means that the spread of the OCRs can be wide.

If there is no reasonable explanation for the differences in the OCRs of the comparables, the valuer has to choose which rate(s) are the best for the property under appraisal. To aid the valuer with this adjustment process, there are statistical tools that are designed to deal with differences. The formula that we are going to use is outlined as follows:

$$SD = \sqrt{\sum_{i=1}^n CF_i - CF)^2 P_i}$$

Table 1

COLUMN A	COLUMN B	COLUMN C	COLUMN D	COLUMN E	COLUMN D
Array of possible cap rate outcomes	The expected overall rate	Difference between col A and col B	Squaring to remove neg signs	Applying the probability of col A	Total variance of each computation col D x col E
5.0%	6.25%	-1.25%	1.56%	0.05%	0.078
5.50%	6.25%	-0.75%	0.56%	0.05%	0.028
6.0%	6.25%	-0.25%	0.06%	0.40%	0.025
6.50%	6.25%	+0.25%	0.06%	0.25%	0.016
7.0%	6.25%	+0.75%	0.56%	0.20%	0.113
7.5%	6.25%	+1.25%	1.56%	<u>0.05%</u>	<u>0.078</u>
			Total	100%	0.34 Sum of all variance

The target is SD (standard deviation), which is the distance around the mean. The CF with a subscript *i* are the OCRs from the sales. The CF is the mean of the OCRs from the sales. We have to square the difference between the list of the OCRs and the mean to eliminate any negative numbers. The P with a subscript *i* is the probability of the likelihood that each OCR of the sales is the same for the subject property. All probabilities will add up to 100%. The square root of the total sum of all variances and probabilities is then applied +/- to the CF or mean of all the OCRs to arrive at the OCR for the subject property.

If this sounds confusing, the following illustrative example will make it all clear. It is a great tool to use in appraisal reports.

Example of dealing with a wide array of OCRs

If you have a series of OCRs from sales and these rates could fit the subject property, then this is known as a **Probability Distribution**. If you have an OCR range of 5.25%, 6.1%, 6.5% 7.0%, and 7.25%, it is known as a **Discrete Probability Distribution** because the possible outcomes are limited to the specific point estimates that are already established (i.e., 5.2% to 7.25%). In other words, the OCR selected for the subject

property is a derivative of the data range. They are not continuous.

To solve for the subject property's best OCR, all that is needed is to apply variance to the problem and some probability. Variance is nothing more than the weighted average of the squared differences between each possible outcome and the expected outcome. However, we need to apply standard deviation, which is the square root of the variance, because it provides a more useable measure of dispersion.

We searched our data basis and found that the best sales had an OCR between 5.25% and 7.25%. However, we know that the OCR for the subject property is not absolute in the sense that it is not going to match perfectly with one or two of the rates from the comparable sales. However, the sales are our best evidence of an OCR for the subject property 'going into' the analysis.

We can smooth out the OCR from the sales by creating a more standard range. We created a range between 5.0% and 7.50% using a 0.50% increment. A half a point of a capitalization rate is not going to be a significant difference in the end. However, if the valuer feels that might be an issue, then the rates should be broken down using a 0.25% difference. Our new OCR range is shown in Table 1's Column A.

The average of the OCR that we created is 6.25%. It is shown in Column B. That is our only estimate of an OCR given our data set. We know that it is not perfect and we then have to 'adjust' the OCRs in the data set to modify the selected average (6.25%). We can do that by putting the data into an Excel spreadsheet and breaking down the formula for variance.

Column C is the difference between our array of OCRs and our selected 'best guess' in terms of the targeted OCR for the subject property at this point in the analysis.

Column D is the squaring of the numbers in Column C to eliminate all the negative values of the difference between the OCR range and our best targeted OCR for the subject property.

Column E is where the valuer starts to 'adjust' the 'going in rates' of 5.0% to 7.5%. In other words, what is the probability of the subject property achieving an OCR of 5.0%, 5.50%, etc? The immediate response is: "How am I supposed to know that?"

It is here that the valuer brings his or her knowledge of the marketplace and an understanding of why OCRs differ per property type. We should point out that the probability Column E tends to force the valuer to think about the relationship

of the possible choices of an OCR in Column **A**, relative to the subject property. It is an unforgiving moment, because there needs to be a balance between rationalization and probability.

Column **D** is the product of all the probabilities against the variance between the targeted or best guess average OCR and the possible number of OCRs that could fit the subject property. The actual sum of all the variances is quite meaningless, so one needs to take the square foot of that number. Doing so is just calculating the standard deviation of the total sum of the variance. It is nothing more than a yardstick to measure the distance around our targeted OCR of the subject property (6.25%). We know that the 6.25% is not the correct OCR, but it is our best guess. The decision-making process of trying to validate the OCR for the subject property through variance will lead us to a range of OCRs in which the subject property should fall.

The square root of the total variance (0.34) is 0.58%. This would be 1 standard deviation. Therefore, the OCR of the subject property is going to be 6.25% +/- 0.58% or 5.67% to 6.83%. This latter range is significantly better than the starting point.

In our example, the average of 6.25% was the best fit of the array of sales. There was no mode or median that we could consider. Depending upon the shape of the OCRs 'going into' the analysis, perhaps the mode or median would be the better choice? One could also use a weighted mean by simply doing the SUM FUNCTION in Excel. Table 2 is an example of this.

The cap rates in Table 2 are the same as those used in the original example of dealing with variance. The weight column is simply the % of each capitalization rate of the total sum, e.g., 5%/37.5. The amount of the weight is totaled to represent 100%.

The typical average of the capitalization rates is shown in blue at 6.25%. However, the weighted average in red is calculated differently by using the SUM FUNCTION in Excel. This new weighted average is 6.37%. If we use the weighted average as opposed to the mean average, the OCR of the subject property would be 5.8% to 6.9%. It is not much different from our original range. However, that could change depending upon the situation.

The advantages of using variance and probability as a source to determine an OCR of a given property is as follows:

1. It helps to remove some of the differences of the OCRs 'going into' the analysis, particularly if the market rates are wide for unknown reasons.
2. It forces the valuer to reflect upon what is the meaning behind specific capitalization rates. OCRs are more than numbers. Each one should account for compensation for a risk-free liquid investment, risk itself, the changes of illiquidity, management, and some provision for appreciation. They represent a collective consciousness of various buyers who have purchased similar, but not identical properties to that of the subject property. They were prepared to accept a given rate of return on and of that investment at a moment in time. Was the OCR a reflection of the fact that there is little real estate product on the marketplace? Perhaps the buyer felt that the OCR was reasonable because the property is near to other properties that he or she owns? Perhaps the rate was acceptable because the probability of the tenants paying the rent each month was very good and, therefore, the future income was somewhat guaranteed?

Table 2

	CAP RATES	WEIGHT	
	5	13.00	average 6.25
	5.5	15.00	weighted average 6.37
	6	16.00	
	6.5	17.00	
	7	19.00	
	7.5	20.00	
SUM	37.5	100.00	

Conclusion

Variance and probability are all about exploring your data. That is the function of all valuers: to explain the differences in real estate whether that be within the Direct Comparison Approach, GIMs, or OCRs. So why not use some statistics to help you along the way? 📊