

**Statistical Methods in Valuation Analysis: Data Sets and Samples  
 (Part Four of a Six-Part Series)**

This fourth installment of the six-part *Health Capital Topics* series on the application of statistical methods by valuation analysts will provide a brief overview of datasets, samples, and their utilization in various valuation approaches and methodologies. As discussed in Part One of this series, entitled, “*Review of Principles and Applications*,”<sup>1</sup> a strong understanding of commonly used statistical methods is useful to a valuation analyst in creating, defending, and/or critiquing valuation reports, including when engaging in forecasting and benchmarking functions. Further, any improper use of statistical methods by a valuation analyst may lead to erroneous inferences by the analyst or their client that may significantly affect the calculated value indication. It is imperative that valuation analysts possess a working understanding of the collection, structure, and patterns associated with data given the frequent reliance of valuation analysis on the statistical applications built from data sets and samples.

Variables used as inputs in a valuation analysis may be inherently *random*, i.e., the variable takes on a numerical value previously unknown to the valuation analyst until the value is observed through some measuring process.<sup>2</sup> Populations, and the samples drawn from them, may exhibit certain patterns, referred to as *statistical distributions*, some of which occur regularly in statistics and have special names (e.g., *binomial* or *normal* distributions) with known properties that may provide a valuation analyst with convenient analytical tools for understanding the data.<sup>3</sup>

*Random variables* and their distributions are typically classified into one of two general categories: (1) *discrete* (i.e., data that can only take particular values within an interval); or, (2) *continuous* (i.e., data that is able to take on all possible values within an interval).<sup>4</sup>

Discrete distributions have been developed to describe observed patterns in certain types of *discrete random variables*, including: (1) *binomial* distribution; (2) *hypergeometric* distribution; (3) *negative binomial* distribution; and, (4) *Poisson* distribution.<sup>5</sup> The *Binomial* distribution may be used to estimate the probability of a “*success*,” or the phenomena for which the researcher is looking, over any number of fixed trials, e.g., the probability of observing eight or more heads if a coin is tossed ten times.<sup>6</sup> The *hypergeometric* distribution is related to the *binomial* distribution, but gives the

probability for the *number* of successes in a given sample, e.g. the probability of observing exactly eight heads in ten coin tosses.<sup>7</sup> The *negative binomial* distribution fixes the number of successes desired and informs the researcher as to the necessary number of trials to acquire the preferred number of successes, e.g., the number of required coin tosses to observe eight heads.<sup>8</sup> The *Poisson* distribution uses a parameter, measured as a rate per unit time or unit area, which may be used to find the probability that an event will happen within a specified time interval, e.g., if the average number of people visiting an exhibit is ten people per day, the probability that 15 will visit tomorrow.<sup>9</sup>

Similarly, *continuous distributions* have been developed to explain the observed patterns in *continuous random variables*, including: (1) *normal* distribution; (2) *lognormal* distribution; and, (3) *beta* distribution, among others.<sup>10</sup> The *normal* distribution (bell-shaped curve) is symmetrically centered on its mean,<sup>11</sup> and is considered “*the most important [distribution] in all of probability and statistics.*”<sup>12</sup> The *normal* distribution gives the cumulative probability for a range of values leading up to (or away from) a selected data point, typically located on either side of the mean for any real valued number, e.g., half of observations are less than the mean observation.<sup>13</sup> Transforming the *normal* distribution under a logarithmic operator for a random variable greater than zero yields the *lognormal* distribution, which models data where the percentage change is *normally* distributed.<sup>14</sup> The *Beta* distribution is a two *parameter* distribution, whose parameter values allow for a variation in the shape of the curve, and generates probabilities for a random variable that is bounded from above and below over any interval, e.g., perhaps effective for modeling data that is not symmetric but proportionally bounded between zero and one (i.e., the *standard beta*).<sup>15</sup>

A valuation analyst’s understanding of statistical distributions has direct application to valuation. One example of a *discrete random variable* is the *condition factor* assigned to assets in the valuation of *furniture, fixtures, and equipment* (FF&E). When conducting a site visit, an appraiser observes the condition of each item, and may assign a number to a piece of FF&E (e.g., one (1) for “*good*” through seven (7) for “*scrap*,” with various conditions represented between these bounds).<sup>16</sup> If the condition of certain pieces of FF&E were not

captured during the site visit, a valuation analyst may utilize the observed attributes about the condition of the collection of assets to make inferences about the missing data, i.e., to impute the most likely condition value for the missing data. However, the ability to use statistical techniques to make such inferences is dependent on the analyst's assumptions related to the likely distribution of the condition factor. For instance, if a sufficiently large amount of the collected data which is from FF&E is heavily weighted towards one end of the *condition factor* spectrum, any missing pieces may be inferred to exhibit a similar level of quality.

A second example of a *discrete random variable* is the procedure volume of a healthcare provider, since providers are not likely to provide half of a procedure, the observed volume will be a discrete number, i.e., an integer. The comparison of a specific hospital's procedure volume to a broader market may be built from statistically inferred performance metrics for that hospital, e.g., utilizing the *Poisson* distribution to find the probability of a certain number of procedures within a given month, to determine how to better schedule staff or medical equipment to increase productivity and make efficient use of equipment or staff.

One example of a *continuous random variable* a healthcare valuation analyst will likely encounter is *practice expense*. It is reasonable to expect expenses to range from zero (i.e., no expenses) to some measurably, higher amount (e.g., a number with eight decimal places, in millions of dollars, or both). A valuation analyst may use *practice expenses* for a statistical comparison of a specific practice's expenses with other comparable entities in an effort to determine profitability. This comparison is predicated on an assumption regarding the distribution of possible practice expenses, e.g., an analyst might assume a *lognormal* distribution.

Additionally, a valuation analyst may use the assumption of a *normal* distribution to investigate how a specific entity's net profit compares to the net profits of a sample of market comparables, or the *lognormal* distribution to investigate volatility in "*dollars per unit productivity*" across states with differing income levels. In either case, understanding whether the data is *discrete* or *continuous*, and understanding the implications of the assumed distribution, will lead to more robust inferences required for competent valuation analysis.

Often, the size of a population renders it impossible or impractical to study in its entirety. Under these circumstances, *sampling* is the preferred method to develop a representation and draw inferences regarding a population.<sup>17</sup> Recall, from Part Two of this series, entitled, "*Descriptive Statistics*,"<sup>18</sup> that a relationship exists between sample size, the *Central Limit Theorem*, and the *Law of Large Numbers*. Given a random sample from a population, as the number of observations within the sample increases, the calculated sample mean approaches the value of the population mean, and as samples are repeated, the distribution of the sample means approximates a *normal* distribution.<sup>19</sup> Conversely,

small sample sizes may vary greatly from the true population mean and distribution.

Looking ahead in Part Five of this series, entitled, "*Regression Analyses*," an additional application of these properties will be used to describe how a valuation analyst may estimate any unknown parameter (through regression analysis) and determine if this estimate meets the requirements necessary to convey meaningful information about a sample.

As noted in Part Two of this series, entitled, "*Descriptive Statistics*,"<sup>20</sup> certain unwanted properties (e.g., non-representative samples and sample bias) may plague a sample while hindering a valuation analyst's ability to accurately describe a population. In light of these properties, a valuation analyst should be familiar with the role that methodologies associated with the collection of sample data from a population play in judging the data's quality. Two important fundamental assumptions of sample data are: (1) *independence* of the data; and, (2) whether the data are *identically distributed observations*.<sup>21</sup> *Independence* within data requires that one observation cannot influence another, i.e., the occurrence of an observation does not influence the measurement of another observation.<sup>22</sup> *Identically distributed* data arise from adherence to an exact collection methodology when creating a sample, ensuring that the measurement and scope of the information are shared by all observations.<sup>23</sup> It may be important for a valuation analyst to make assumptions about the collected data in order to sufficiently generalize the data to gather meaningful interpretations of the population through other statistical tests, such as data collected by industry surveys.

Many such statistical tests are built upon an assumption of *normality*, or that the data is *normally* distributed, which if wrong, may lead to "*inaccurate inferential statements*."<sup>24</sup> A valuation analyst may wish to test whether a sample is approximately *normally* distributed, suggesting that the sample's properties are closely related to the properties associated with a *normal* distribution. First, there are simple "*quick checks*" that may automatically disqualify a sample, e.g., the sample must be continuous, and the sample may take on all sample values in the support for a *normal* distribution (e.g., a variable restricted to non-negativity cannot be *normal*).

Next, the sample may be subject to tests of normality that are more rigorous. An example of a *normality test* is the *Chi-Squared Goodness of Fit Test* that uses *maximum likelihood tests* (a method that identifies the necessary parameter values required to make test true) to investigate the probability that the sample is *normally distributed*.<sup>25</sup> A second *test for normality* is the *Jarque-Bera Test* that estimates the sample mean and variance to compare the sample to a *normal* distribution through hypothesis testing.<sup>26</sup> A third *test for normality* is the *Shapiro-Wilk Test* that utilizes *order statistics* (a sub-sample of the data through independent observations, e.g., the collection of the smallest values) and their means and variance to compare the sample to a *normal* distribution through hypothesis testing.<sup>27</sup>

Occasionally upon examination of a sample, certain members of the sample may appear to differ significantly from the remainder of the data, potentially arising from data collection errors or unexpected characteristics of the data. These points, referred to as *outliers*, warrant special attention to obviate any deleterious effects that may arise from their inclusion in the analysis.<sup>28</sup> A valuation analyst who suspects that there may be outliers in the data may opt to use a statistical test (e.g., Grubb’s test, discussed below)<sup>29</sup> to assist in the identification of outliers.<sup>30</sup> This process is somewhat subjective, and relies on the ability of a valuation analyst to appropriately identify outliers within the data and develop forecasts that adjust for the outlier effects.<sup>31</sup> Depending on the nature of the outlier, it may or may not be wise to exclude the data point. Outliers arising from data collection errors should always be excluded, but other outliers may highlight important factors affecting the random variable of interest.

In valuation analysis, since historical data may be limited to the two or three years prior to the valuation date, the valuation analyst should carefully consider whether to remove an outlier from a data set.<sup>32</sup> With such small sample sizes, rigorously identifying outliers may be difficult, and the analyst should seek, as a rule, to conserve as much relevant information as possible.<sup>33</sup> A

statistical test by Dr. Frank E. Grubb recommends certain criteria for identifying outliers through hypothesis testing based upon sample means and standard deviations, and where the process is adjusted to reflect the number of suspected outliers by the valuation analyst.<sup>34</sup> An alternative heuristic for identifying an outlier uses *Chebyshev’s inequality*, proven by Russian mathematician Pafnuty Chebyshev.<sup>35</sup> The inequality can be applied to any sample, regardless of distribution, which states that the probability of the distance between a data point and the sample mean is greater than a certain standard deviation (one, two, ... standard deviations) is at most one over the number of standard deviations squared.<sup>36</sup> A valuation analyst can conclude that, for any sample, regardless of distribution, the probability of a data point being three standard deviations away from the mean is, at most, one over the square of three, or approximately 11 percent, and may be an outlier.

Understanding the structure of data, its distribution, and its effects of sampling is important for valuation analysts in making inferences about data that statistics alone cannot answer. The fifth installment of this six-part series will shift from an overview of general statistics to a focus on *regression analysis*, its uses in the valuation analysis, and potential pitfalls or mistakes in its interpretation.

1 “Statistical Methods in Valuation Analysis: Review of Principles and Applications (Part One of a Six-Part Series)” Health Capital Topics, Vol. 9, No. 7 (July 2016).  
 2 “Probability and Statistics for Engineering and the Sciences” By Jay L. Devore, Australia: Thompson Brooks/Cole, 2004, p. 98.  
 3 *Ibid*, p. 104.  
 4 *Ibid*, p. 100.  
 5 *Ibid*, p. 120-138.  
 6 *Ibid*, p. 120-122.  
 7 *Ibid*, p. 128-129.  
 8 *Ibid*, p. 132-133.  
 9 *Ibid*, p. 135.  
 10 *Ibid*, p. 160-186.  
 11 *Ibid*, p. 161-162.  
 12 *Ibid*, p. 161.  
 13 *Ibid*, p. 162.  
 14 *Ibid*, p. 184.  
 15 *Ibid*, p. 185.  
 16 “Healthcare Valuation: Financial Appraisal of Enterprise, Assets, and Services.” By Robert James Cimasi, MHA, ASA, FRICS, MCBA, CVA, CM&AA, Volume 2, Hoboken, NJ: John Wiley and Sons, 2014, p. 748-750.  
 17 Devore, 2004, p. 3.  
 18 “Statistical Methods in Valuation Analysis: Descriptive Statistics (Part Two of a Six-Part Series)” Health Capital Topics, Vol. 9, No. 8 (August 2016).  
 19 Devore, 2004, p. 239.

20 Health Capital Topics, Statistical Methods in Valuation Analysis: Descriptive Statistics (Part Two of a Six-Part Series), August 2016.  
 21 “Common Errors in Statistics (and How to Avoid Them)” By Phillip I. Good and James W. Harden, 2nd ed., Hoboken, NJ: John Wiley and Sons, 2006, p. 36.  
 22 Devore, 2004, p. 86-87.  
 23 Good and Harden, 2006, p. 37-38.  
 24 “A Test for Normality of Observations and Regression Residuals” By Carlos M. Jarque and Anil K. Bera, International Statistical Review, Vol. 55, No. 2 (August 1987), p. 164.  
 25 Devore, 2004, p. 649-652.  
 26 Jarque and Bera, August 1987, p. 164.  
 27 “An Analysis of Variance Test for Normality (Complete Samples)” By S.S. Shapiro and M.B. Wilk, Biometrika, Vol. 52, No. 3/4 (December 1965), p. 591-593.  
 28 Devore, 2004, 2004, p. 30.  
 29 “Procedures for Detecting Outlying Observations in Samples” By Frank E. Grubbs, Technometrics, Vol. 11, No. 1 (February 1969).  
 30 Cimasi, 2014, p. 59.  
 31 *Ibid*.  
 32 *Ibid*.  
 33 *Ibid*.  
 34 Grubbs, February 1969, p. 2-3.  
 35 Devore, 2004, p. 119.  
 36 *Ibid*.



(800)FYI - VALU

Providing Solutions  
in the Era of  
Healthcare Reform

Founded in 1993, HCC is a  
nationally recognized healthcare  
economic financial consulting firm

- [HCC Home](#)
- [Firm Profile](#)
- [HCC Services](#)
- [HCC Experts](#)
- [Clients & Projects](#)
- [HCC News](#)
- [Upcoming Events](#)
- [Contact Us](#)
- [Email Us](#)

## HCC Services

- [Valuation Consulting](#)
- [Commercial Reasonableness Opinions](#)
- [Commercial Payer Reimbursement Benchmarking](#)
- [Litigation Support & Expert Witness](#)
- [Financial Feasibility Analysis & Modeling](#)
- [Intermediary Services](#)
- [Certificate of Need](#)
- [ACO Value Metrics & Capital Formation](#)
- [Strategic Consulting](#)
- [Industry Research Services](#)



**Robert James Cimasi**, MHA, ASA, FRICS, MCBA, CVA, CM&AA, serves as Chief Executive Officer of **HEALTH CAPITAL CONSULTANTS (HCC)**, a nationally recognized healthcare financial and economic consulting firm headquartered in St. Louis, MO, serving clients in 49 states since 1993. Mr. Cimasi has over thirty years of experience in serving clients, with a professional focus on the financial and economic aspects of healthcare service sector entities including: valuation consulting and capital formation services; healthcare industry transactions including joint ventures, mergers, acquisitions, and divestitures; litigation support & expert testimony; and, certificate-of-need and other regulatory and policy planning consulting.

Mr. Cimasi holds a Master in Health Administration from the University of Maryland, as well as several professional designations: Accredited Senior Appraiser (ASA – American Society of Appraisers); Fellow Royal Institution of Chartered Surveyors (FRICS – Royal Institution of Chartered Surveyors); Master Certified Business Appraiser (MCBA – Institute of Business Appraisers); Certified Valuation Analyst (CVA – National Association of Certified Valuators and Analysts); and, Certified Merger & Acquisition Advisor (CM&AA – Alliance of Merger & Acquisition Advisors). He has served as an expert witness on cases in numerous courts, and has provided testimony before federal and state legislative committees. He is a nationally known speaker on healthcare industry topics, and is the author of several books, the latest of which include: *“The Adviser’s Guide to Healthcare – 2nd Edition”* [2015 – AICPA]; *“Healthcare Valuation: The Financial Appraisal of Enterprises, Assets, and Services”* [2014 – John Wiley & Sons]; *“Accountable Care Organizations: Value Metrics and Capital Formation”* [2013 – Taylor & Francis, a division of CRC Press]; and, *“The U.S. Healthcare Certificate of Need Sourcebook”* [2005 – Beard Books].

Mr. Cimasi is the author of numerous additional chapters in anthologies; books, and legal treatises; published articles in peer reviewed and industry trade journals; research papers and case studies; and, is often quoted by healthcare industry press. In 2006, Mr. Cimasi was honored with the prestigious *“Shannon Pratt Award in Business Valuation”* conferred by the Institute of Business Appraisers. Mr. Cimasi serves on the Editorial Board of the Business Appraisals Practice of the Institute of Business Appraisers, of which he is a member of the College of Fellows. In 2011, he was named a Fellow of the Royal Institution of Chartered Surveyors (RICS). In 2016, Mr. Cimasi was named a *“Pioneer of the Profession”* as part of the recognition of the *National Association of Certified Valuators and Analysts (NACVA) “Industry Titans”* awards, which distinguishes those whom have had the greatest impact on the valuation profession.



**Todd A. Zigrang**, MBA, MHA, ASA, FACHE, is the President of **HEALTH CAPITAL CONSULTANTS (HCC)**, where he focuses on the areas of valuation and financial analysis for hospitals, physician practices, and other healthcare enterprises. Mr. Zigrang has over 20 years of experience providing valuation, financial, transaction and strategic advisory services nationwide in over 1,000 transactions and joint ventures. Mr. Zigrang is also considered an expert in the field of healthcare compensation for physicians, executives and other professionals.

Mr. Zigrang is the co-author of *“The Adviser’s Guide to Healthcare – 2nd Edition”* [2015 – AICPA], numerous chapters in legal treatises and anthologies, and peer-reviewed and industry articles such as: *The Accountant’s Business Manual* (AICPA); *Valuing Professional Practices and Licenses* (Aspen Publishers); *Valuation Strategies: Business Appraisal Practice*; and, *NACVA QuickRead*. In addition to his contributions as an author, Mr. Zigrang has served as faculty before professional and trade associations such as the American Society of Appraisers (ASA); the National Association of Certified Valuators and Analysts (NACVA); Physician Hospitals of America (PHA); the Institute of Business Appraisers (IBA); the Healthcare Financial Management Association (HFMA); and, the CPA Leadership Institute.

Mr. Zigrang holds a Master of Science in Health Administration (MHA) and a Master of Business Administration (MBA) from the University of Missouri at Columbia. He is a Fellow of the American College of Healthcare Executives (FACHE) and holds the Accredited Senior Appraiser (ASA) designation from the American Society of Appraisers, where he has served as President of the St. Louis Chapter, and is current Chair of the ASA Healthcare Special Interest Group (HSIG).



**John R. Chwarzinski**, MSF, MAE, is Senior Vice President of **HEALTH CAPITAL CONSULTANTS (HCC)**. Mr. Chwarzinski’s areas of expertise include advanced statistical analysis, econometric modeling, as well as, economic and financial analysis. Mr. Chwarzinski is the co-author of peer-reviewed and industry articles published in *Business Valuation Review* and *NACVA QuickRead*, and he has spoken before the Virginia Medical Group Management Association (VMGMA) and the Midwest Accountable Care Organization Expo.

Mr. Chwarzinski holds a Master’s Degree in Economics from the University of Missouri – St. Louis, as well as, a Master’s Degree in Finance from the John M. Olin School of Business at Washington University in St. Louis. He is a member of the St. Louis Chapter of the American Society of Appraisers, as well as a candidate for the Accredited Senior Appraiser designation from the American Society of Appraisers.



**Jessica L. Bailey-Wheaton**, Esq., is Vice President and General Counsel of **HEALTH CAPITAL CONSULTANTS (HCC)**, where she conducts project management and consulting services related to the impact of both federal and state regulations on healthcare exempt organization transactions and provides research services necessary to support certified opinions of value related to the Fair Market Value and Commercial Reasonableness of transactions related to healthcare enterprises, assets, and services. Ms. Bailey is a member of the Missouri and Illinois Bars and holds a J.D., with a concentration in Health Law, from Saint Louis University School of Law, where she served as Fall Managing Editor for the *Journal of Health Law & Policy*.



**Kenneth J. Farris**, Esq., is an Associate at **HEALTH CAPITAL CONSULTANTS (HCC)**, where he provides research services necessary to support certified opinions of value related to the Fair Market Value and Commercial Reasonableness of transactions related to healthcare enterprises, assets, and services, and tracks impact of federal and state regulations on healthcare exempt organization transactions. Mr. Farris is a member of the Missouri Bar and holds a J.D. from Saint Louis University School of Law, where he served as the 2014-2015 Footnotes Managing Editor for the *Journal of Health Law & Policy*.