CURRICULUM VITAE OF DAVID OLIVE

I. PROFESSIONAL AFFILIATION AND CONTACT INFORMATION

A. Present University Department or Unit: Mathematics

B. Official Address:

School of Mathematical & Statistical Sciences
Neckers Building, Mailcode 4408
1245 Lincoln Drive, Southern Illinois University
Carbondale, IL 62901-4408

II. EDUCATION

Ph.D. Statistics, University of Minnesota, January 1997 - September 1998.
M.S. Mathematics, Northern Illinois University, January 1988 - August 1989.
B.S. Electrical Engineering, University of Illinois, August 1984 - January 1988.

III. PROFESSIONAL EXPERIENCE

- Professor of Mathematics: Southern Illinois University at Carbondale, July 2014 - present.
- Associate Professor of Mathematics: Southern Illinois University at Carbondale, July 2004 - June 2014.
- Assistant Professor of Mathematics: Southern Illinois University at Carbondale, August 1999 - June 2004.
- Visiting Assistant Professor of Statistics: University of Minnesota, September 1998 - June 1999.

IV. RESEARCH AND CREATIVE ACTIVITY

A. Interests and Specialties: Prediction Regions, Regression, Robust Statistics, Statistical Inference, Statistical Learning, Applied Probability.

B. Current Projects:

- i) Online Math 586 course notes Olive (2021a) *Prediction and Statistical Learning* gives a) a mixture distribution estimator for variable selection methods like forward selection for multiple linear regression, GLMs, and some survival regression models, b) asymptotically optimal prediction intervals to compare methods like ridge regression, lasso, and forward selection, c) bootstrap confidence regions, including confidence intervals, for hypothesis testing after model or variable selection, and d) plots for visualizing models in the background of the data. See (http://parker.ad.siu.edu/Olive/slearnbk.htm). e) Variable selection material needs to be updated with Rathnayake and Olive (2021), which give large sample theory for many variable selection estimators.
- ii) The online text Olive (2021b) *Robust Statistics* is a major revision of the online (2008) Math 583 course notes *Applied Robust Statistics*. See (http://parker.ad.siu.edu/Olive/robbook.html).

- iii) Course notes for Math 584: Olive (2021c) *Theory for Linear Models*: see (http://parker.ad.siu.edu/Olive/linmodbk.htm).

- iv) Online course notes for Math 473: Olive (2021d) *Survival Analysis*: see (http://parker.ad.siu.edu/Olive/survbk.htm).

- v) Rajapaksha, K.W.G.D.H. and Olive, D.J. (2021), "Wald Type Tests with the Wrong Dispersion Matrix" is at (http://parker.ad.siu.edu/Olive/ppwaldtype.pdf)

- vi) Haile, M.G., and Olive, D.J. (2021a), "Bootstrapping ARMA Time Series Models after Model Selection" is at (http://parker.ad.siu.edu/Olive/pptsboot.pdf).

- vii) Haile, M.G. and Olive, D.J. (2021b), "Prediction Intervals and Regions for Some Time Series, Random Walks, and Renewal Processes" is at (http://parker.ad.siu.edu/Olive/pptspi.pdf).

- viii) Jin, Y. and Olive, D.J. (2021), "Large Sample Theory for Some Ridge-Type Regression Estimators," is at (http://parker.ad.siu.edu/Olive/ppridgetype.pdf).

In preparation:

ix) "Data Splitting for Inference After Model Selection," likely with Lingling Zhang.

Unsolicited offers to consider *Applied Robust Statistics* for publication have been received from Cambridge University Press, Wiley, Springer, Jones & Bartlett Learning, and Chapman & Hall/CRC.

According www.google.com, I have had one of the highest ranked websites for the key words **"statistics preprints"** since 2004 (ranked number one for about 6.5 years).

C. Grants Applied for: NSF CAREER 2003 (placed in "fund if possible"),
NSA (2005: both reviews suggested funding the grant, but withdrew application after NSF 2005 resulted in a grant, 2009, 2010, 2011),
NSF DMS (1999, 2000, 2001, 2003, 2004, 2005, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2019).

D. Grants Received:

NSF-DMS-0202922 (*Applied Robust Statistics*, $43,561, single PI, 2002-2004).

NSF-DMS-0600933 (*Robust Statistics*, $89,162, single PI, 2006-2009). About 5 NSF-DMS Statistics grants a year are awarded to a single PI in departments (like the SIU Math department) that do not offer a graduate degree in Statistics, but unlike SIU, usually their university does have a Statistics or Biostatistics department.

E. Honors and Awards: None.

F. Papers and Presentations at Professional Meetings:

1. "Visualizing 1D Regression" at ICORS 2003: Antwerp, Belgium (invited).

2. "Response Plots for Experimental Design" at ASA Quality and Productivity Conference, June 2008, Madison, Wisconsin.

G. Other:

i) I have developed a big website on Statistics.

ii) Invited talk "Response Plots and Related Plots for Regression" at the University of Iowa Department of Biostatistics, March 2010.

V. PUBLICATIONS AND CREATIVE WORKS

23 papers, 2 papers to appear, 3 books, 1 paper in a book of conference proceedings, (27 publications, 2 papers to appear)

A. Books:

Olive, D.J. (2014), *Statistical Theory and Inference,* Springer, New York, NY. The Springer eBook is available on SpringerLink, Springer's online platform, (http://dx.doi.org/10.1007/978-3-319-04972-4).

Olive, D.J. (2017a), *Linear Regression*, Springer, New York, NY. The Springer eBook is available on SpringerLink, Springer's online platform, (http://dx.doi.org/10.1007/978-3-319-55252-1).

Olive, D.J. (2017b), *Robust Multivariate Analysis*, Springer, New York, NY. The Springer eBook is available on SpringerLink, Springer's online platform, (https://link.springer.com/book/10.1007%2F978-3-319-68253-2).

B. Articles in Professional Journals:

1. Hawkins, D.M., and Olive, D.J. (1999a), "Improved Feasible Solution Algorithms for High Breakdown Estimation," *Computational Statistics & Data Analysis,* 30, 1-11.

2. Hawkins, D.M., and Olive, D. (1999b), "Applications and Algorithms for Least Trimmed Sum of Absolute Deviations Regression," *Computational Statistics & Data Analysis,* 32, 119-134.

3. Olive, D., and Hawkins, D.M. (1999c), "Comment on 'Regression Depth' by P.J. Rousseeuw and M. Hubert," *Journal of the American Statistical Association,* 94, 416-417. (Invited comment to one of four "Theory and Methods" discussion papers published by JASA in 1999.)

4. Cook, R.D., and Olive, D.J. (2001), "A Note on Visualizing Response Transformations in Regression," *Technometrics,* 43, 443-449.

5. Olive, D.J. (2001), "High Breakdown Analogs of the Trimmed Mean," *Statistics & Probability Letters,* 51, 87-92.

6. Hawkins, D.M., and Olive, D.J. (2002), "Inconsistency of Resampling Algorithms for High Breakdown Regression Estimators and a New Algorithm," *Journal of the American Statistical Association,* (with discussion), 97, 136-148. (One of four "Theory and Methods" discussion papers published by JASA in 2002.)

7. Hawkins, D.M., and Olive, D.J. (2002), "Rejoinder to discussion of 'Inconsistency of Resampling Algorithms for High Breakdown Regression Estimators and a New Algorithm'," *Journal of the American Statistical Association,* 97, 156-159.

8. Olive, D.J. (2002), "Applications of Robust Distances for Regression," *Technometrics,* 44, 64-71.

9. Olive, D.J., and Hawkins, D.M. (2003), "Robust Regression with High Coverage," *Statistics & Probability Letters*, 63, 259-266.

10. Olive, D.J. (2004), "A Resistant Estimator of Multivariate Location and Dispersion," *Computational Statistics & Data Analysis,* 46, 99-102.

11. Olive, D.J., and Hawkins, D.M. (2005), "Variable Selection for 1D Regression Models," *Technometrics*, 47, 43-50.

12. Olive, D.J. (2005), "Two Simple Resistant Regression Estimators," *Computational Statistics & Data Analysis,* 49, 809-819.

13. Olive, D.J. (2007), "Prediction Intervals for Regression Models," *Computational Statistics & Data Analysis,* 51, 3115-3122. (3rd "Hottest CSDA Article" for January-March 2007.)

14. Olive, D.J., and Hawkins, D.M. (2007), "Behavior of Elemental Sets in Regression," *Statistics & Probability Letters,* 77, 621-624.

15. Chang, J., and Olive, D.J. (2010), "OLS for 1D Regression Models," *Communications in Statistics: Theory and Methods*, 39, 1869-1882.

16. Zhang, J., Olive, D.J., and Ye, P. (2012), "Robust Covariance Matrix Estimation with Canonical Correlation Analysis," *International Journal of Statistics and Probability*, 1, 119-136.

17. Olive, D.J. (2013a), "Plots for Generalized Additive Models," *Communications in Statistics: Theory and Methods*, 42, 2610-2628.

18. Olive, D.J. (2013b), "Asymptotically Optimal Regression Prediction Intervals and Prediction Regions for Multivariate Data," *International Journal of Statistics and Probability*, 2, 90-100.

19. Olive, D.J., Pelawa Watagoda, L.C.R., and Rupasinghe Arachchige Don, H.S. (2015), "Visualizing and Testing the Multivariate Linear Regression Model," *International Journal of Statistics and Probability*, 4, 126-137.

20. Olive, D.J. (2018), "Applications of Hyperellipsoidal Prediction Regions," *Statistical Papers*, 59, 913-931.

21. Rupasinghe Arachchige Don, H.S., and Olive, D.J. (2019), "Bootstrapping Analogs of the One Way MANOVA Test," *Communications in Statistics: Theory and Methods*, 4, 5546-5558.

22. Pelawa Watagoda, L.C.R., and Olive, D.J. (2021), "Bootstrapping Multiple Linear Regression after Variable Selection," *Statistical Papers*, 62, 681-700.

23. Pelawa Watagoda, L.C.R., and Olive, D.J. (2021), "Comparing Six Shrinkage Estimators with Large Sample Theory and Asymptotically Optimal Prediction Intervals," *Statistical Papers*, 62, 2407-2431. (2020 Web of Science early access)

24. Olive, D.J., Rathnayake, R.C., and Haile, M.G. (2021), "Prediction Intervals for GLMs, GAMs, and Some Survival Regression Models," *Communications in Statistics: Theory and Methods*, to appear. (2021 Web of Science early access)

25. Rathnayake, R.C., and Olive, D.J. (2021), "Bootstrapping Some GLM and Survival Regression Variable Selection Estimators," *Communications in Statistics: Theory and Methods*, to appear. (2021 Web of Science early access)

C. Creative Contributions:

1) *Variable Selection*: i) Found the large sample theory for many variable selection estimators, including multiple linear regression, many GLMs, some time series models, and some survival regression models. For some of these variable selection estimators, bootstrap hypothesis testing has been developed. A bootstrap technique to get rid of selection bias is to fit the variable selection estimator $\hat{\boldsymbol{\beta}}^*_{VS}$ to a bootstrap sample to find a submodel, then draw another bootstrap sample and fit the same submodel to get the bootstrap estimator $\hat{\boldsymbol{\beta}}^*_{MIX}$. Here the number of predictors p is fixed, and the sample size $n \to \infty$. See Haile and Olive (2021a), Pelawa Watagoda and Olive (2019, 2020) and Rathnayake and Olive (2021).

ii) Found prediction intervals that can be used after variable selection even if $p > n$ provided the estimator is consistent. The prediction intervals work for the model $Y = m(\boldsymbol{x}) + e$ where the iid error distribution is unknown, and for parametric regression models such as GLMs, GAMs, and some survival regression models. Prediction intervals after model selection for some time series models were also developed. See Haile and Olive (2021b), Olive (2017ab, 2018, 2021ab), Pelawa Watagoda and Olive (2020), and Olive, Rathnayake, and Haile (2021).

iii) Let model I_{min} minimize a criterion C such as AIC or C_p. Instead of looking at models I with $C(I) - C(I_{min}) \leq a$, look at models with no more predictors than I_{min} such that $C(I) - C(I_{min}) \leq a$ where $a = 7$ for AIC and 4 for C_p. Fast methods of variable selection using software originally meant for multiple linear regression were extended to a much larger class of regression models. Also when using C_p, if model I contains k predictors including a constant, use the screen $C_p(I) \leq \min(2k, p)$ instead of $C_p(I) \leq k$. See Olive and Hawkins (2005) and Olive (2017a).

iv) Working to improve data splitting: randomly divide the data set into two half sets. On the first half set fit the model selection method, e.g. forward selection or lasso, to get the a predictors. On the second half set, use the standard regression inference from regressing the response on the predictors found from the first half set. This method can be inefficient if $n \geq 10p$, but is useful for sparse models if $n \leq 5p$, if the probability that the model underfits goes to zero, and if $n \geq 10a$. See Olive (2021ab).

2) Developed the 1D regression model where the response variable Y is independent of the vector of predictor variables $\boldsymbol{x}$ given some real valued function $h(\boldsymbol{x})$, written $Y \perp\!\!\!\perp \boldsymbol{x} | h(\boldsymbol{x})$ where $\boldsymbol{x}$ is a $p \times 1$ vector and n is the sample size. Generalized linear models (GLMs), generalized additive models (GAMs), and some survival regression models are 1D regression models, as is the additive error regression model $Y = h(\boldsymbol{x}) + e$ which includes multiple linear regression, nonlinear regression, and nonparametric regression.

See Chang and Olive (2010), Olive and Hawkins (2005), and Olive (2010, 2013a, 2017ab, 2021abcd).

3) *Statistical Learning:* For i) and ii), consider 1D regression models where $h(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{\beta}$ is a linear combination of the predictors. Such models include multiple linear regression, GLMs, and survival regression models. A shrinkage variable selection estimator applies the classical estimator, such as a GLM, to the variables selected by the shrinkage estimator.

i) Let p be fixed. See 1) i). The bootstrap method in 1) i) is useful for hypothesis testing.

ii) Let p be fixed. Showed shrinkage variable selection estimators, such as OLS after least angle regression, lasso variable selection, and elastic net variable selection, are $\sqrt{n}$ consistent under much milder conditions than the corresponding shrinkage methods: the shrinkage variable selection estimators tend to be $\sqrt{n}$ consistent when the corresponding shrinkage estimator is consistent.

iii) Developed prediction intervals for many 1D regression models including GLMs, GAMs, additive error regression, and multiple linear regression. See 1) iii). The prediction interval (PI) is useful if d is the number of variables used by the regression estimator and $n >> d$. The PI can be used for models selected with 10–fold cross validation or variable selection.

iv) Response and transformation plots have been developed to visualize the model. Points outside the PI can be plotted if n is huge, to eliminate a black band about the estimated mean function in the plots. See Olive, Rathnayake, and Haile (2021), and Olive (2021a) *Prediction and Statistical Learning*.

v) Developed a method for multivariate outlier detection when the number p of predictor variables is not necessarily less than the sample size n. See Olive (2017b, 2021abc).

vi) For multiple linear regression where p is fixed, developed simpler proofs for large sample theory for elastic net, ridge regression, and lasso, as well as other regularized estimators. See Pelawa Watagoda and Olive (2020), and Olive (2021ab).

vii) See 1), 2) and 3).

4) *Prediction*: a) Residuals and the shorth estimator can be used to make asymptotically optimal prediction intervals for future response Y_f for regression models of the form $Y = m(\boldsymbol{x})+e$ even when the iid error distribution is unknown. These prediction intervals can be visualized with a response plot. b) To predict a future multivariate observation $\boldsymbol{x}_f$ without knowledge of the distribution of $\boldsymbol{x}$, replace the classical region $\{\boldsymbol{x}|D^2_{\boldsymbol{x}} \leq \chi^2_{p,1-\alpha}\}$ by $\{\boldsymbol{x}|D^2_{\boldsymbol{x}} \leq D^2_{(U_n)}\}$, where $U_n \geq [n(1-\alpha)]$, but $U_n/[n(1-\alpha)] \to 1$ as $n \to \infty$. These prediction regions can be visualized with a DD plot. c) For multivariate linear regression with m response variables $\boldsymbol{y}$, replace the classical region $\{\boldsymbol{y}|D^2_{\boldsymbol{y}} \leq \chi^2_{m,1-\alpha}\}$ by $\{\boldsymbol{y}|D^2_{\boldsymbol{y}} \leq D^2_{(U_n)}\}$ where $U_n \geq [n(1-\alpha)]$. These prediction regions can be visualized with the DD plot of the residual vectors. d) Prediction intervals for time series have been developed. e) Prediction intervals after model or variable selection, that can be useful even if $p > n$, have been developed. f) One of the prediction regions can be applied to a bootstrap sample to get a confidence region useful for bootstrap hypothesis testing.

See Olive (2007, 2013b, 2018, 2021abcd), Olive (2017b) *Robust Multivariate Analysis*, Olive, Rathnayake, and Haile (2021), and Pelawa Watagoda and Olive (2019, 2020).

Competing prediction intervals and regions are either parametric and tend to be too short when the parametric distribution is misspecified, or, like some bootstrap methods, tend to take too long to compute.

5) *Data Visualization:* Brillinger, Li, Cook, and Weisberg note that if the response Y is independent of the predictors $\boldsymbol{x}$ given $\boldsymbol{x}^T\boldsymbol{\beta}$, then a response plot of $\boldsymbol{x}^T\hat{\boldsymbol{\beta}}$ versus Y is useful. Extending this work for a 1D regression model, if Y is independent of $\boldsymbol{x}$ given some real valued function $h(\boldsymbol{x})$, then a response plot of $\hat{h}(\boldsymbol{x})$ versus Y can be used to visualize the regression model, including additive error regression, multiple linear regression, GLMs, and GAMs. Plots for survival analysis, response transformations, detecting outliers, detecting overdispersion, goodness of fit, lack of fit, and for comparing models were also developed. Several of these techniques are effective for multivariate regression and MANOVA. Many of the plots are useful if $p > n$. Response plots should have impact similar to that of residual plots in the 1960s. See Olive (2013a) and Olive (2017ab, 2021abc).

5) *Bootstrapping Hypothesis Tests and Confidence Regions:* The percentile confidence interval is also a prediction interval for a future value of the bootstrap statistic. The percentile method can be generalized by applying the Olive (2013b) prediction region to the bootstrap sample of vector valued statistics. See Haile and Olive (2021a), Pelawa Watagoda and Olive (2019), Olive (2017ab, 2018, 2021abcd), and Rathnayake and Olive (2021). This technique is useful for bootstrapping variable selection models. See 1) and 2).

6) *Practical Robust Multivariate Location and Dispersion Estimators:* The first practical highly outlier resistant FCH, RFCH, and RMVN estimators of multivariate location and dispersion backed by theory were developed, extending Olive (2004). These estimators can be used to make DD plots for outlier detection and to determine if the distribution is multivariate normal or from some other elliptically contoured distribution. See Olive (2017b, 2021ab) and Zhang, Olive, and Ye (2012). Also developed a method for multivariate outlier detection when the number p of predictor variables is not necessarily less than the sample size n. See Olive (2017b, 2021abc).

7) *Practical High Breakdown Robust Regression:* Developed the first practical multiple linear regression estimator (HBREG) shown to be both high breakdown and asymptotically efficient. This estimator has been extended to a robust multivariate linear regression estimator ($m > 1$ response variables). See Olive (2017b).

8) *Theory for the Multivariate "High Breakdown Robust Estimators" Actually Used:* Hawkins and Olive (2002) showed that the Rousseeuw Yohai paradigm of using a fixed number of elemental sets to approximate high breakdown regression and multivariate location and dispersion estimators results in inconsistent zero breakdown estimators. This paper gave the first theory for multivariate "high breakdown robust estimators" actually used. Also see Olive (2017b, 2021ab).

9) Use the DD plot as a diagnostic for multivariate analysis. For example, if the covariance matrix is hypothesized to be $\boldsymbol{\Sigma}_0$, plot $D_i(\overline{\boldsymbol{x}}, \boldsymbol{S})$ versus $D_i(\overline{\boldsymbol{x}}, \boldsymbol{\Sigma}_0)$. See Olive (2002, 2017b).

10) *Statistical theory:* For a large class of exponential families, developed two limit theorems, found the minimal mean square error estimators, and found the power functions of UMP tests. Developed ten new random variables along with corresponding inference.

Showed inference for the half normal, Pareto, and two parameter Power distributions is simple. Showed the shorth can estimate highest density intervals which can be used for bootstrap and Bayesian tests of hypotheses. See Olive (2014). Similarly, the Olive (2013b) prediction region can be used to estimate a Bayesian credible region.

11) *Miscellaneous Results:* Extended OLS theory for multiple linear regression inference to a much larger class of 1D regression models. See Chang and Olive (2010). Developed many robust point estimators, a few robust confidence intervals, and an outlier resistant prediction interval for the location model. See Olive (2008) *Applied Robust Statistics*. Showed resampling methods such as the bootstrap and permutation tests should have the number of resamples $B \to \infty$ rather than using B fixed. See Olive (2014). Found a useful F approximation for the Hotelling Lawley test statistic. See Olive (2017b) and Olive, Pelawa Watagoda, and Rupasinghe Arachchige Don (2015). Developed large sample theory and bootstrap methods for one way MANOVA type estimators where groups do not need a common covariance matrix. See Rupasinghe Arachchige Don and Olive (2019).

12) Online Course notes Math 584: *Theory for Linear Models*, Math 586 (583): *Prediction and Statistical Learning*, Math 583: *Applied Robust Statistics* and *Robust Statistics*, Math 580: *A Course in Statistical Theory*, Math 473: *Survival Analysis*, and Math 484: *Multiple Linear and 1D Regression* are fairly good.

The online Course notes for Math 585: *Robust Multivariate Analysis* are not so good, but these notes are an early version of the book Olive (2017b). The Math 484 and 580 notes are early versions of the books Olive (2014, 2017a).

13) Developed a large website for Statistics.

D. Chapters in Professional Books: None.

E. Popular and Creative Writing: None.

F. Book Reviews:

25 Book Reviews: a) 24 for *Technometrics*: 2003 45: 362-363; 2004 46: 492; 2005 47: 233-234; 2007 49: 496; 2008 50: 236; 2009 51: 218-218, 221-222, 224, 343-344; 2010 52: 139, 261, 464-465; 53: 214, 221; 2012 54: 201, 327-328, 443; 2013 55: 249, 378, 570-1; 2014 56: 265, 551; 2017 59: 131-132; 2020 62: 288.

b) 1 for the *Journal of the American Statistical Association*: 2010 105: 1631.

G. Other:

a) Published Abstract and Paper in Conference Proceedings:

- "Asymptotically Optimal Complexity Based Model Selection," with Andrew Barron and Yuhong Yang, Abstract in *The Institute of Mathematical Statistics Bulletin,* 1992, 21, 162.

- Olive, D.J. (2004b), "Visualizing 1D Regression," in *Theory and Applications of Recent Robust Methods*, eds. Hubert, M., Pison, G., Struyf, A., and Van Aelst, S., Birkhäuser, Basel, Switzerland, 221-233.

b) Thirteen Unpublished Manuscripts Cited (more than 100 times) by Other Authors. (The top four were revised and published, the fifth and sixth were incorporated in the text

Olive (2014) *Statistical Theory and Inference*, while the last six have been incorporated in the text Olive (2017b) *Robust Multivariate Analysis*.)

- Zhang, Z., and Olive, D.J. (2009), "Applications of a Robust Dispersion Estimator."
- Chang, J., and Olive, D.J. (2007), "Resistant Dimension Reduction."
- Olive, D.J. (2008a), *A Course in Statistical Theory* (earlier version of Olive (2014) *Statistical Theory and Inference*).
- Olive, D.J. (2010), *Multiple Linear and 1D Regression* (earlier version of Olive (2017a) *Linear Regression*).
- Abuhassan, H., and Olive, D.J. (2008), "Inference for the Pareto, Half Normal and Related Distributions."
- Olive, D.J. (2009), "The Number of Samples for Resampling Algorithms."
- Olive, D.J. (2006), "Robust Estimators for Transformed Location Scale Families."
- Olive, D.J. (2005), "A Simple Confidence Interval for the Median."
- Olive, D.J. (2008b, online since 2002), *Applied Robust Statistics* (may try to publish a revision Olive (2021b) *Robust Statistics*).
- Olive, D.J., and Hawkins, D.M. (2007), "Robustifying Robust Estimators."
- Olive, D.J., and Hawkins, D.M. (2008), "High Breakdown Multivariate Estimators."
- Olive, D.J., and Hawkins, D.M. (2010), "Robust Multivariate Location and Dispersion."
- Olive, D.J., and Hawkins, D.M. (2011), "Practical High Breakdown Regression."

c) Online Course Notes Available from my Personal Website:

- Versions of Olive, D.J. (2008), *Applied Robust Statistics* (588 pages) were used for Math 583 in Fall 2004, and for Math 490 in Fall 2011. A version of Olive, D.J. (2021b) *Robust Statistics* was used for Math 583 in Fall 2020.
- Versions of Olive, D.J. (2008), *A Course in Statistical Theory* (388 pages) were used for Math 580 in Spring 2008, Spring 2012, and Spring 2014.
- Half of an earlier version of Olive, D.J. (2010), *Multiple Linear and 1D Regression* (641 pages) was used for Math 484 in Fall 2009. An early version of Olive (2017a) was used in Fall 2016.
- An earlier version of Olive, D.J. (2013c), *Robust Multivariate Analysis* (393 pages) was used for Math 583 in Fall 2012.

- An early version of Olive (2021a) *Prediction and Statistical Learning* was used for Math 583 in Fall 2017, and can be used for Math 586.

- An versions of Olive (2021c) *Theory for Linear Models* were used for Math 584 in Spring 2019 and 2021.

Also, the following documents were used as supplemental reading.

1) Olive, D.J. (2008, ch. 16) *Multiple Linear and 1D Regression* for Math 473 Reliability and Survival Analysis.

2) Olive, D.J. (2008, ch. 5-9) *Multiple Linear and 1D Regression* for Math 583 Experimental Design. Also see Olive (2017a, ch. 5-9).

3) Olive, D.J. (2008, ch. 10-12) *Multiple Linear and 1D Regression* for Math 485 Categorical Data Analysis. Also see Olive (2017a, ch. 13).

VI. TEACHING EXPERIENCE

A. Teaching Interests and Specialties:
Probability and Statistics.

B. Teaching and Training Grants: None.

C. Teaching Awards and Honors: None.

D. Current Graduate Faculty Status: Regular Graduate Faculty Member.

E. Number of Master's and Ph.D. Committees on which I have Served: 92.

F. Names of Students who have completed Master's Theses and Doctoral Dissertations under my Direction:

26 MS Students: 1) Samer Salem (2001) *Visualizing 1 Dimensional Structure*, 2) Flaviu Hodis (2003) *Variable Selection in Regression*, 3) Mussie Tesfamicael (2003) *Visualizing Binary Regression*, 4) Rajan Lamichhane (2005) *Some Simple Tools for Assessing Regression Models*, 5) Mohan Pant (2006) *Poisson Confidence Intervals and Diagnostics for Poisson Regression*, 6) Jessica Bayer (2007) *Binomial Confidence Intervals and Diagnostics for Binomial Regression*, 7) Elmer Agudelo Rodriguez (2007) *Regression and ANOVA under Heterogeneity*, 8) Jenna Haenggi (2009) *Plots for the Design and Analysis of Experiments*, 9) Joshua Powers (2011) *Plots and Prediction Intervals for Generalized Additive Models*, 10) Lasanthi C.R. Pelawa Watagoda (2013) *Plots and Testing for Multivariate Linear Regression*, 11) Hasthika S. Rupasinghe Arachchige Don (2013) *Robust Multivariate Linear Regression*, 12) Sarah Wiley (2014) *Time Series Analysis and Prediction Intervals*, 13) Ayed Alanzi (2014) *Robust Principal Component Analysis*, 14) Chathurangi H.K. Pathiravasan (2015) *Bootstrapping Hypothesis Tests*, 15) Luke Yang (2016) *Confidence Intervals for the Survival Function*, 16) Mona Almaghshi (2016) *Prediction Intervals after Forward Selection*, 17) Craig Bartelsmeyer (2017) *Prediction Intervals for Lasso and Relaxed Lasso Using d Variables*, 18) Mulubrhan Haile (2017) *Prediction Intervals after Forward Selection Using EBIC*, 19) Kosman W.G. Dimuthu H. Rajapaksha (2017) *Prediction Intervals after Forward Selection Using d Variables*, 20) Sung-ho Kim (2017) *Prediction Intervals for Partial Least Squares and Principal Component Regression Using d Variables*, 21) Chamila K. Ranaweera (2018) *Kriging and*

Nonparametric Regression, 22) Handong Wang (2018) *Outlier Detection for High Dimensional Data*, 23) Charles Murphy (2018) *Bootstrapping Forward Selection with BIC*, 24) Drew Imhoff (2018) *Bootstrapping Forward Selection with C_p*, 25) Mashael Alshammari (2019) *Bootstrap Confidence Intervals for β_i Using Forward Selection with C_p*, 26) Lingling Zhang (2019) *Prediction Intervals for Scaled Shrinkage Estimators*, 27) Daina McKinney (2021) *Some t-Type Confidence Intervals*

7 PhD Students: 1) Jing Chang (2006) *Resistant Dimension Reduction,* 2) Hassan Abuhassan (2007) *Some Transformed Distributions*, 3) Jiangfeng Zhang (2011) *Applications of a Robust Dispersion Estimator*, 4) Hasthika S. Rupasinghe Arachchige Don (2017) *Bootstrapping Analogs of the One Way MANOVA Test*, 5) Lasanthi C.R. Pelawa Watagoda (2017) *Inference after Variable Selection*, 6) Rasanji C. Rathnayake (2019) *Inference for Some GLMs and Survival Regression Models after Variable Selection*, 7) Kosman W.G.D.H. Rajapaksha (2021) *Wald Type Tests with the Wrong Dispersion Matrix*.

G. Other:

a) Twenty One Reading Classes (3 hours unless noted): 1) Spring 2002 (Applied Regression), 2) Spring 2004 (Theory of Linear Models: 2 students), 3) Spring 2005 (Regression Graphics: 2 students), 4) Fall 2005 (Large Sample Theory: 2 students), 5) Summer 2007 (Regression Graphics), 6) Fall 2007 (Robust Statistics), 7) Spring 2008 (Large Sample Theory), 8) Spring 2010 (Actuarial Mathematics, 2 hours), 9) Fall 2011 (Robust Statistics), 10) Spring 2013 (Robust Statistics: 4 students), 11) Spring 2015 (Large Sample Theory: 3 students), 12) Summer 2015 (Actuarial Mathematics), 13) Fall 2015 (bootstrap), 14) Spring 2017 (Survival Analysis, 1 hour of Math 495), 15) Spring 2017 (Math 400 as Math 495), 16) Fall 2019 (Large Sample Theory, 6 students), 17) Fall 2019 (Statistical Learning), 18) Spring 2020 (Applied Statistics 1 hour), 19) Fall 2020 (Math 495). 20) Spring 2021 (Statistical Learning). 21) Summer 2021 (Statistical Learning, 2 hours of Math 495)

b) PhD Oral Prelims: 49.

c) Contributed questions for PhD written preliminary exams for 20 math students and for 7 students outside the math department.

d) I currently have 3 Ph.D. students and 1 MS students.

e) I have helped write, proctor and grade 33 Math 580 Statistical Inference PhD qualifying exams since 2000 and 1 Math 584 Linear Models qualifying exam since 2021.

f) I introduced ARC, R and Compaq Fortran to the computer lab, and gave a presentation on ARC in Spring 2004.

g) With the Statistics group, proposed the eleven courses Math 474: Time Series, and Math 490: Special Topics in Mathematics, Math 584: Linear Models, Math 585: Multivariate Analysis, Math 586: Statistical Computing and Learning, Math 486: Statistical Computing, and proposed the undergraduate major in Actuarial Mathematics with courses Math 400: Interest Theory and Financial Derivatives, Math 401: Life Contingencies I, Math 402: Life Contingencies II, Math 403: Loss Models I, and Math 404: Loss Models II. (Worked on an undergraduate data science degree, an undergraduate Statistics degree, an undergraduate Economics–Statistics degree, and a tuition paying Master's Statistics degree.)

h) I got Math 484 and Math 474 approved for the Validation by Educational Experience: VEE-Applied Statistical Methods for the Society of Actuaries and the Casualty Actuarial Society.

i) Fall 2015: Took over two of Dr. Budzban's students for undergraduate research: Sports Data Analysis Project.

VII. UNIVERSITY EXPERIENCE

1. Department Committees:

Undergraduate Programs Committee (2004,2006,2008,2010,2011,2012,2013,2014, 2015, 2016, 2017, 2018, spring 2019)

Personnel Committee (2015,2016,2017)

Statistics Hiring Committee (2014)

Productivity Metric Committee (2013)

Mathematics Promotion Report Committee (2013)

Promotion and Tenure (2006,2008,2009,2011,2012,2016,2017,2020)

Subcommittee on Graduate Programs (2011)

Graduate Programs Committee (2005,2009,2018,2019,2020,2021)

Math Field Day (5 times)

2. College and University Committees and Councils:

COS Promotion and Tenure (2016)

Institute for Statistics Task Force # 3 (2013).

C. Other: i) Undergraduate Advisor 2006-2007. ii) Summer 2016–: **Undergraduate advisor for Mathematics majors.** iii) Math fellowships since 2019. Awarded: 2019 1) Doctoral Fellowship, 2) DRA, 2020 3) Doctoral Fellowship.

VIII. PROFESSIONAL SERVICE

A. Membership in Professional Organizations: American Statistical Association.

B. Offices Held and Honors Awarded in Professional Associations: None.

C. Consultantships: (No paid consulting at SIU, but I did 3 projects in Spring 1992 at the University of Illinois, and worked at Northern Illinois University's consulting lab June 1994 to August 1995.) Helped students on regression for Elizabeth Spector's high school project on insulation, Joanne Numrich's MS on conjugated linoleic acid in beef, and Alyass Hosin's PhD on the effect of coal waste additives on cement strength. See Olive (2017a: problems 3.14, 3.15, and 3.18; and Olive 2017b: Example 14.12).

Helped students and former students for papers including i) Rupasinghe Arachchige Don, H.S., and Pelawa Watagoda, L.C.R. (2018), "Bootstrapping Analogs of the Two Sample Hotelling's T^2 Test," *Communications in Statistics: Theory and Methods*, 47, 2172-2182. ii) Fan, R. (2016), "A Squared Correlation Coefficient of the Correlation Matrix," unpublished manuscript at (http://parker.ad.siu.edu/Olive/sfan.pdf). iii) Wijekularathna, D.K., Yi, H., and Roka, A. (2019), "Test Based on Kurtosis for Multivariate Normality," *Austrian Journal of Statistics*, 48, 1-10.

D. Evaluation of Manuscripts for Journals and Book Publishers and of Grant Proposals for Agencies:

I have refereed papers for *American J. of Mathematical and Management Sciences, American Statistician, Applicable Analysis and Discrete Mathematics, Annals of Statistics, Applied Mathematics and Computation, Austrian Journal of Statistics, Behavior Research Methods, Biometrika, British J. of Mathematics & Computer Science, Canadian J. of Statistics, Communications in Statistics (Sim, and Theory and Methods), Computational Statistics & Data Analysis, Electronic Journal of Applied Statistical Analysis, Entropy, ICORS, Information Sciences, Int. J. on Artificial Intelligence Tools, International Statistical Review, J. of the American Statistical Association, J. of Applied Statistics, J. of Computational and Graphical Statistics, J. of Machine Learning Research, J. of Statistical Computation and Simulation, J. of Statistical Planning and Inference, Mathematical and Computational Applications, Mathematics, Mathematics and Computers in Simulation, Methodology and Computing in Applied Probability, Metron, Multivariate Behavioral Research, Pakistan Journal of Statistics, Pakistan Journal of Statistics and Operation Research, PLOS ONE, Psychometrika, ScienceAsia, Sensors, South African Statistical Journal, Statistica Sinica, Statistical Science, Statistics and Computing, Sustainability, Technometrics,* and *Test.*

Number of referee jobs for journals by year: 1998 2, 2000 5, 2001 1, 2002 5, 2003 3, 2004 2, 2005 5, 2006 14, 2007 10, 2008 5, 2009 4, 2010 4, 2011 11, 2012 5, 2013 9, 2014 4, 2015 11, 2016 9, 2017 2, 2018 7, 2019 8, 2020 7, 2021 5

Evaluation of Book Proposals:

2001: 1 for Marcel-Dekker. 2005: 1 for Wiley. 2014: 1 for World Scientific, 2016: 1 for Wiley. 2017: 1 for Wiley. 2019: 1 for Chapman & Hall/CRC.

Evaluation of Grant Proposals:

2009: 1 for NSF

E. Papers and Presentations at Professional Meetings (other than those listed under "Research and Creative Activity"): None.

F. Other

Advisory Committee

University of Minnesota School of Statistics Advisory Committee (2015-2017): *of statistical leaders outside the School to collaborate with us in advising, supporting and promoting the School.* See (http://www.stat.umn.edu/people/advisory.html).

Math Fellowships

In 2019-2021, I handled the Department of Mathematics applications for the Doctoral Fellowships, Master's fellowships, and Dissertation Research Awards. In 2019, Math got both a dissertation research assistantship DRA and a doctoral fellowship. In 2020, Math got a doctoral fellowship. In 2021, Math got 5 DRAs.

Foreign Examiner of Ph.D. Thesis

2011: 1 for Student in Pakistan

Proctor for SOA/CAS Actuarial Exam

2011: twice

Supervisor for the (Actuarial) FAP Final Assessment end-of-module evaluation

2012 for 1 student

External Reviewer for Promotion to Associate or Full Professor

I have done this at least 3 times for SIU grads and at least 3 times for nonSIU grads including professors at Abdul Wali Khan University (Pakisatan), DePaul University, Northern Arizona University, and Universiti Utara Malaysia.

Math 282 Proficiency Exam

2015: for 2 students

IX. COMMUNITY SERVICE

One presentation (with Dr. Bhattacharya) for high school students.