# Pei-Shan Yen

: Linkedln My-Website : pyen2@uic.edu : 608-335-7436

## Summary

Self-motivated, enthusiastic with a positive attitude. 15 years of combined professional with 12 years of volunteering experience in statistical consulting, data analysis, and marketing research. Highly competent in problem-solving, communication skills, strategic planning, and interdisciplinary collaboration.

## Skills

**Software:** R (7 yrs), R Shiny (1 yr), R Markdown (3 yrs), SAS (5 yrs), SPSS (12 yrs), Stata (1 yr), Minitab (2 yrs), SQL (1 yr), Latex (3 yrs), Unix/Linux (0.5 yr)

**Statistics:** Longitudinal Data Analysis, Statistical Inference and Modeling, Data Visualization, Marketing Research, Survey Methodology, Data Science

## Education

**University of Illinois at Chicago (UIC)** **Chicago, USA**
*Ph.D. Candidate in Biostatistics* (GPA: 4.00) 2018 - Present

- *Research*: Longitudinal Data Analysis, Causal Mediation Model, Neuroimaging Studies
- *Honor*: 2021 Paul Levy Memorial Scholarship

**National Chengchi University** **Taipei, Taiwan**
*M.S. in Statistics* (GPA: 3.86) 2010

- *Research*: An International Study of 2010 Population Census Methods
- *Honor*: Invited Paper to the Journal of Population Studies

**University of Taipei** **Taipei, Taiwan**
*B.S. in Mathematics* (GPA: 3.55) 2005

- *Honor*: Ranked top 3% in the national mathematics undergraduate student competition

## Doctoral Research

**Longitudinal High-Dimensional Mediation Model for Multimodal Neuroimaging Data**
*Advisor: Dr. Dulal Bhaumik*

- Developing one of the first penalized longitudinal mediation models for high-dimensional data.
  1. Modeling: Construct the depiction of the pathway between therapeutic treatments and neurobehavior outcome with two-modality high-dimensional human brain connectivity mediators. Brain structural and functional connectivities are measured by MRI technique.
  2. Feature Selection: Employ the high-dimensional sparsity and non-smoothness penalty functions incorporating temporal and spatial associations to select significant brain connectivity. Bootstrapping approach is proposed to verify the significance of connectivity.
  3. Prioritize False Discovery Rate (FDR): Present the multiple hypothesis testing procedure by taking into account the order of importance of brain connectivity.
  4. Model Validation: Evaluate and validate the proposed theoretical model using both statistical simulations and real-world data for the patients suffering from mental disorders (Coding: R).

## Professional Experiences

**Graduate Research Assistant, Biostatistics Core** Jul 2019 - Present
*Center for Clinical and Translational Science, UIC, USA*

- Participated in 15 medical research projects with physicians and researchers from UIC, VA medical centers, and Rush University.
  1. Assisted 4 principal statisticians to provide consultation.
  2. Performed statistical data analysis and tabulated customized results (Coding: SAS, R).
  3. Supported the preparation of 2 grant proposals, 4 manuscripts, and 2 conference abstracts.
- Developing an R Shiny app to visualize all the datasets of the Research Electronic Data Capture database (REDCap). The functions include clinical trial recruitment tracking, descriptive statistical tabulated results, and interactive plots. Preparing the user manual and, upon completion, upload the R Shiny app source code to GitHub (Coding: R Shiny/R Markdown/JSON).

**Graduate Teaching Assistant, School of Public Health** Aug 2018 - May 2019
*Division of Epidemiology and Biostatistics, UIC, USA*

- Tutored the Graduate Courses "Biostatistics II" and "Analytic and Research Methods".

**Senior Research Data Scientist, Acquisition and Loyalty Division** Apr 2014 - Apr 2016
*Taiwan Mobile, Taiwan ($2^{nd}$ largest telecom company in Taiwan)*

- Designed and conducted marketing analytic projects to locate and prioritize customer churn behavior. Proposed strategy options. Monitored marketing plans and sales performance.
- Led the data science team (a group of 8 data engineers & data scientists) in projects involving 7,000,000 users to predict churn behavior.
  1. Constructed a proprietary machine learning model to predict the risk of churn behavior (annual churn rate reduced by 3%).
  2. Proposed 3 natural language processing projects to prevent daily churn behavior by monitoring real-time user phone, text, and Internet surfing behavior (daily churn rate reduced by 28%).
  3. Analyzed the behavior of 800,000 long-term iPhone users, executed mystery shopper surveys, and designed promotional plans for new iPhones sales.

**Senior Data Analyst, Energy Research Center** Nov 2011 - Feb 2014
*NCKU Research and Development Foundation, Taiwan*
(*The leading institute to perform solar-heating subsidy program sponsored by Taiwan government*)

- Led 7 large-scale field surveys, constructed cost-benefit analysis models, and analyzed the payback period with a team of 12 engineers in the poultry & butchery industry, hospitality industry, and schools in rural areas. Successfully motivated 5 large-scale system installations.
- Analyzed 130,000 solar heater user behavior in Taiwan. Evaluated lifetime of solar systems with cost-benefit analysis for the subsidy policy adjustments.
- Prepared 2 grant proposals, 3 annual reports, and 8 manuscripts for solar heater development.

**Senior Data Analyst, 104 Web Survey Center** May 2010 - Apr 2011
*104 Group, Taiwan* (*The largest human resource company in Taiwan*)

- Proposed new research plans, executed 15 online survey research projects involving 2,000,000 customers and composed reports. Collaborated with clients including Philips (Healthcare), Shiseido (Cosmetics), Kanebo (Cosmetics), and Carat (Global Media & Marketing Agency).
- Constructed job classification models to design vocational training courses for the Ministry of Labor, Taiwan.

**Data Analyst, Market Survey Research Center** Jun 2005 - Oct 2010
*National Chengchi University, Taiwan* (*Best College of Commerce in Taiwan*)

- **Graduate Research Assistant** Sep 2007 - Oct 2010
  1. Provided 3 projects involving statistical consulting service and used the GeneSpring software to perform human genetic clustering for the Institute of Biomedical Sciences, Academia Sinica.
  2. Conducted user behavior and pricing studies to help design an innovative mobile device "ORBI" for interactive exhibitor information, map guide and the exchange of digital business cards. The project was sponsored by IBM and HTC.
  3. Led a team of 6 software engineers in project to redesign the computer-assisted telephone surveying system and added visualization and tracking functions.
- **Data Analyst** Jun 2005 - Aug 2007
  1. Led a team of 26 interviewers and conducted 29 quantitative and qualitative research projects. Hosted 12 focus group interviews. Performed 6 central location tests, 19 telephone surveys, 3 field surveys, 2 mail surveys, and 2 online surveys.
  2. Prepared research proposals, conducted analyses, composed reports, and offered feasible solutions to various government agencies, industries, and research institutes (Coding: SPSS).

## Selected Publications

**Book**

1. C.S. Yue and **P.S. Yen**. "Big data: The data wisdom and knowledge-based economics." *Tsang Hai Publishing*, Taiwan, 2016.

**Journal Publications**

1. D.K. Bhaumik, Y.B. Song, **P.S. Yen**, and O.A. Ajilore. "Power Analysis for High Dimensional Neuroimaging Studies." *Psychiatric Annals* (under review).
2. I. Imayama, A. Gupta, **P.S. Yen**, Y.F. Chen, B. Keenan, R. R. Townsend, J. A. Chirinos, F. M. Weaver, D. W. Carley, S. T. Kuna, and B. Prasad. "Racial Differences in Cardiovascular Response to Positive Airway Pressure Treatment." *Journal of Clinical Sleep Medicine* (under review).
3. C.C. Li, A.K. Matthews, **P.S. Yen**, Y.F. Chen, and X.Q. Dong. "Intimate partner violence and Its Relationship with Psychological Distress Among Older Asian Americans: Results from the California Health Interview Survey." *Asian Journal of Psychiatry*, 2021.
4. C.C. Li, A.K. Matthews, **P.S. Yen**, Y.F. Chen, and X.Q. Dong. "The Influence of Perceived Discrimination in Healthcare Settings on Psychological Distress Among a Diverse Sample of Older Asian Americans." *Aging & Mental Health*, 2021.
5. C.W. Kuo, **P.S. Yen**, W.C. Chang and K.C. Chang. "The design and optical analysis of compound parabolic collector." *Procedia Engineering*, 2014.
6. **P.S. Yen** and C.S. Yue. "An international study of 2010 population census methods." *Journal of Population Studies*, 2010.

**Conference Proceedings**

1. J. Li, K. Qin, P.B. Doshi, X.D. Zhu, P. Kotini-Shah, M. Granado, H. Kim, **P.S. Yen**, R. Gordon, H.E. Wang, and T. Hoek. "Novel Blood Biomarker Taurine and Glutamate for Early Predicting Outcomes After Out-of-hospital Cardiac Arrest." *AHA Resuscitation Science Symposium*, USA, 2020.
2. Y.F. Wang, K. Schachtschneider, Y.H. Huang, R. Gaba, P. Nguyen, **P.S. Yen**, M. Sverdlov, G. Guzman, and R. Lokken. "Infiltration of T-cell subpopulations in hepatocellular carcinoma is associated with tumor size, BCLC stage, and transplant-free survival." *Society of Interventional Radiology Conference*, USA, 2020.
3. **P.S. Yen** and C.W. Kuo. "Policy for solar water heaters in Taiwan: An International Perspective." *Grand Renewable Energy*, Tokyo, Japan, 2014.
4. **P.S. Yen** and C.S. Yue. "An International Study of 2010 Population Census Methods." *Taiwan Demography Association Annual Meeting*, Taipei, Taiwan, 2010.
5. **P.S. Yen** and C.S. Yue. "Feedback on Taiwan population registered system: A lesson from international experience using census techniques." *Population Policy Symposium in Taipei City*, Taipei, Taiwan, 2010.

**Invited Talks**

1. **P.S. Yen** and S. Konda. "Clustering Depression Data." *Department of Epidemiology and Biostatistics, University of Illinois at Chicago*, Nov, 2019.
2. **P.S. Yen** and K.M. Chung. "Evaluation of solar water heater utilization status in Taiwan rural areas." *2013 Energy Conservation & Carbon Reduction*, IC Broadcasting Company, Hsinchu, Taiwan, May, 2013.

## Professional Organizations

**Student Member** Jul 2017 - Present
*American Statistical Association (ASA)*

**Student Member** Oct 2018 - Present
*Eastern North American Region International Biometric Society (ENAR)*

GRADUATE COURSES

- (*PhD level*) Generalized Linear Model, Linear Model, Multivariate Analysis, Advanced Statistical Inference, Large Sample Theory, Bayesian Methods, Missing Data, Computational Statistics.
- (*MS level*) Clinical Trial (2022 Spring), Experiment Design, Health Analytics, Longitudinal Data Analysis, Survival Analysis, Categorical Data Analysis, Epidemiology, Demography, Sampling Methods, Statistical Consulting, Marketing Management, Product Management

SELECTED DOCTORAL COURSE PROJECTS

- *Bayesian Statistics*: Using Bayesian Ordinal Regression Model to Identify Suicide Risk Factors
- *Missing Data*: A Review of EM algorithm (EM, MCEM, SAEM)
- *Computational Statistics*: Developed R package `MixPoiRayExp` including an EM Algorithm for Finite Mixture Models
- *Generalized Linear Model*: A Review of Complete and Quasi-Complete Separation in Logistic Regression
- *Independent Study for Linear Model*: Two-Component Mix Model and Gamma Mixed Model for Estimating the Unknown True Concentration of Ground Water
- *Linear Model*: Linear Mixed-Effects Model for Detecting Disrupted Connectivities in Heterogeneous Data
- *Health Analytics*: Using Machine Learning Classification Methods and SMOTE Resampling Skills to Predict Disease Risk

VOLUNTEER EXPERIENCES

**President and Organizer of the Millennials Sector** May 2021 - Present
*American Beiyinu Alumnae Foundation, Walnut, CA, USA*

- Redesigned database and questionnaire to collect alumnae information to build networks. The number of alumnae members recruited increased from 328 to 1,397 within 5 months.
- Led and mentored a team of 45 in multiple projects in the analysis of the alumnae member profile, organization structure, marketing strategies, and the organization of alumnae activities, including career consultation, monthly workshops, and fund-raising activities.

**Data Analyst** Oct 2009 - Present
*Asian Mental Health Association (AMHA), Taipei, Taiwan*
(*The leading Taiwanese organization that provides crisis response and mental health service*)

- Resigned PTSD questionnaires for COVID-19 first responders and conducted statistical analysis for the evaluation of crisis response training courses.
- Provided data collection and study design ideas for AMHA, the Ministry of Health and Welfare of Taiwan, and the Red Cross Society of Taiwan. Executed 18 crisis response training for the disaster victims, first responders, the Army of the Republic of China, and various NGOs.
- Assisted the director of AMHA in the proposal and execution of the 2009 Typhoon Morakot Project to reduce the suicide rates of 247 victims. Develop a model to rank the suicidal tendency of victims and demonstrated statistical applications for 82 NGO directors. The results contributed to the initiation of a national-wide emergency plan.

**Freelance Statistician & Volunteering Statistician** Jul 2006 - Present
*Self-Employed, Taipei, Taiwan*

- Provided statistics tutoring and consulting sessions for a total of 47 clients from various professional fields, including physicians, nurses, pharmaceutical scientists, researchers, students, and provided pro-bono statistics sessions for 7 underprivileged students in Taiwan & US.

**Data Analyst** Oct 2016 - Feb 2018
*American Red Cross – Wisconsin Region, Madison, WI, USA*

- Wrote weekly statistical reports for the emergency response vehicle management.

**Data Analyst** Jul 2013 - Jun 2014
*Taipei Lifeline Association, Taipei, Taiwan*

- Used 2011-2014 phone call records from national suicide prevention lifeline to perform data analysis for identifying suicide high-risk group.