

The Statistics Online Computational Resource



It's Online, Therefore It Exists!

Ivo D. Dinov
University of Michigan

www.SOCR.umich.edu

www.SOCR.ucla.edu

The SOCR Resource



Core SOCR Resources



Data

http://wiki.socr.umich.edu/index.php/SOCR_Data



Concepts & Methods

http://wiki.socr.umich.edu/index.php/SOCR_Books



Virtual Demos

<http://www.SOCR.umich.edu>



Tools & Activities

wiki.socr.umich.edu/index.php/SOCR_EduMaterials



Infrastructure

<http://socr.umich.edu/html/Navigators.html>
http://wiki.socr.umich.edu/index.php/SOCR_News
http://wiki.socr.umich.edu/index.php/SOCR_Servers
http://socr.ucla.edu/docs/SOCR_Documentation.html

SOCR is as a Cloud Resource

Resource Type & Usage	Description	URLs
Data & Web-services Users: 80,000 Post 12 months: 15,000	Research-derived, simulated, transactional and clinical data archives. Dashboard for matching multi-source socioeconomic and medical datasets, big data analytics, graphical data exploration and discovery	http://wiki.socr.umich.edu/index.php/SOCR_Data http://wiki.stat.ucla.edu/socr/index.php/SOCR_Data http://socr.umich.edu/HTML5/Dashboard/
Computational Infrastructure Users: 2.5 million Post 12 months: 600,000	Comprehensive collection of web-tools for demonstrating probability, statistics, mathematics and engineering concepts. These include probability calculations, statistics analysis tools, data modeling and visualization, virtual games, simulations and experiments	http://socr.umich.edu http://socr.umich.edu/HTML5/Tables/ http://github.com/SOCR http://github.com/SOCR http://SOCR.appspotcode.com http://SOCRcloud.attlassian.net/browse/SOCR-7 http://www.fadnotes.org
Analytical Tools Users: 30,000 Post 12 mo's: 8,000	Modern HTML5 resources for exploratory analytics, data discovery, simulation, and visualization	http://socr.umich.edu/HTML5/ http://socr.umich.edu/HTML5/BrainViewer/
Learning Resources Users: 2,000,000 Post 12 months: 1,200,000	Community-built, open-access and multilingual resources blending information technology, scientific techniques and modern pedagogical concepts	http://wiki.socr.umich.edu/index.php/EBook http://wiki.stat.ucla.edu/socr/index.php/EBook http://wiki.socr.umich.edu/index.php/SMMS http://socr.umich.edu/people/dinov/2MMS_Courses.html http://wiki.socr.umich.edu http://wiki.stat.ucla.edu/socr

FEATURES

- 80+ Distributions
- Graphs
- PDFs
- CDFs
- Inverse CDFs
- Random Sampling
- Moments
- Web Interfaces
- Interactive
 - Probability-values
 - Critical-values

SOCR Distributions

Bernoulli Distribution; Beta Distribution; Beta (Generalized) Distribution; Binomial Distribution; Birthday Distribution; Cauchy Distribution; Chi-Square Distribution; Circle Distribution; Continuous Uniform Distribution; Die Distribution; Discrete Arcsine Distribution; Discrete Uniform Distribution; Erlang Distribution; Error Distribution; Exponential Distribution; Fisher's F Distribution; Fisher-Tippett Distribution; Gamma Distribution; General Cauchy Distribution; Geometric Distribution; Gilberts Distribution; Gumbel Distribution; Half-Normal Distribution; Hypergeometric Distribution; Laplace Distribution; Logarithmic Distribution; Logistic Distribution; Log-Normal Distribution; Matching Distribution; Maxwell Distribution; Mixture Distribution; Negative-Binomial Distribution; Normal Distribution; Pareto Distribution; Point-Mass Distribution; Poisson Distribution; Poker-Dice Distribution; Power-Function Distribution; Rayleigh Distribution; Student's T Distribution; Student's T Non-Central Distribution; Triangle Distribution; Von Mises Distribution; WalkMax Distribution; WalkPosition Distribution; Weibull Distribution;

www.socr.umich.edu/html/dist
<http://goo.gl/Chwfw4>
www.Distributome.org/V3/calc

SOCR Experiments

FEATURES

- 65+ Experiments
- Simulations
- Summary Stats
- Models vs. Observed data
- GUIs
- Web Interface
- Based on VLPs

Ballot Experiment Ball and Urn Experiment Bertrand Experiment Beta Coin Experiment Beta Estimate Experiment Binomial Coin Experiment Binomial Timeline Experiment Birthday Experiment Bivariate Normal Experiment Bivariate Uniform Experiment Buffon's Coin Experiment Buffon's Needle Experiment CardExperiment Chi Square Dice Experiment Chuck A Luck Experiment Coin Die Experiment Coin Sample Experiment Confidence Interval Experiment Coupon Collector Experiment Craps Experiment Dice Experiment Dice Sample Experiment Die Coin Experiment Finite Order Statistic Experiment Fire Experiment Galton Board Experiment Game Gamma Estimate Experiment Gamma Experiment Markov Chain Experiment Match Experiment Mean Estimate Experiment Mean Test Experiment Mixture Model EM Experiment Monty Hall Experiment Negative Binomial Experiment Normal Estimate Experiment Order Statistics Experiment Pareto Estimate Experiment Problem of Points Experiment Two-Dimensional Poisson Experiment Poisson Experiment Two-Type Poisson Experiment Poker Dice Experiment Poker Experiment Probability Plot Experiment Proportion Estimate Experiment Proportion Test Experiment Quantile Jupyter Random Variable Experiment Random Walk Experiment Red and Black Experiment Roulette Experiment Sample Mean Experiment Sign Test Experiment Spinner Experiment Triangle Experiment Uniform Estimate Experiment Variance Estimate Experiment Variance Test Experiment Voter Experiment

www.socr.umich.edu/html/exp
<http://goo.gl/9Oup92>
www.Distributome.org/V3/exp

SOCR Analyses

FEATURES

- Param+NonParam
- Graphs
- Summary Stats
- R Interface
- GUIs
- Web Interface

ANOVA - One Way
ANOVA - Two Way
 χ^2 Model Goodness-of-Fit Test
Multiple Regression Analysis
One Sample T Test
Simple Regression Analysis
Two Independent Sample T Test
Two Independent Sample Wilcoxon Rank Sum Test
Two Paired Sample Sign-Test
Two Paired Sample Signed-Rank Test (Wilcoxon)
Two Paired Sample T Test
...

www.socr.umich.edu/html/ana
<http://goo.gl/NclSrW>

RESULT:
Sample size=19
INDEPENDENT = Group
DEPENDENT = Dependent

DF Model = 2
DF Error = 16
DF Corrected Total = 18
RSS MODEL = 48035.94956140351
RSS ERROR = 2714.2083333333335
RSS TOTAL = 47745.15789473685
MSS MODEL = 22515.474761904762
MSS ERROR = 169.63020833333334
F-VALUE = 132.72658257916632
P-VALUE = 1.090714185084433E-10

SOCR Modeler

FEATURES

- Distribution Model Fitting
- Fourier and Wavelet Data Modeling
- Random Number Generator (any SOCR distribution)
- Graphs
- GUIs
- Web Interface

BetaFit_Modeler
ExponentialFit_Modeler
FourierFit_Modeler
GammaFit_Modeler
MixedFit_Modeler
NormalFit_Modeler
PoissonFit_Modeler
WaveletFit_Modeler

www.socr.umich.edu/html/mod
<http://goo.gl/qdtpNN>

SOCR Charts

FEATURES

- 60+ Dynamic Interactive Graphs
- Summary Stats
- GUIs
- Web Interface
- Based on JFreeCharts
- Motion-Charts

www.socr.umich.edu/html/cha
<http://goo.gl/tuv09D>

Prob & Stats EBook

- Preface
- Chapter I: Introduction to Statistics
- Chapter II: Describing, Exploring, and Comparing Data
- Chapter III: Probability
- Chapter IV: Probability Distributions
- Chapter V: Normal Probability Distribution
- Chapter VI: Relations Between Distributions
- Chapter VII: Point and Interval Estimates
- Chapter VIII: Hypothesis Testing
- Chapter IX: Inferences From Two Samples
- Chapter X: Correlation and Regression
- Chapter XI: Analysis of Variance (ANOVA)
- Chapter XII: Non-Parametric Inference
- Chapter XIII: Multinomial Experiments and Contingency Tables
- Additional EBook Chapters

<http://wiki.socr.umich.edu/index.php/EBook>

Features

- it is community-built
- open-access (dev & use)
- blends concepts with IT
- multi-lingual

Scientific Methods for Health Analytics EBook

Chapter I: Fundamentals
Chapter II: Applied Inference
Chapter III: Linear Modeling
Chapter IV: Special Topics

- Scientific Visualization
- PCOR/CER methods Heterogeneity of Treatment Effects
- Big-Data/Big-Science
- Missing data
- Genotype-Environment-Phenotype associations
- Medical imaging
- Data Networks
- Adaptive Clinical Trials
- Databases/registries
- Meta-analysis
- Causality/Causal Inference, SEM
- Classification methods
- Time-series analysis
- Scientific Validation
- Geographic Information Systems (GIS)
- Rasch measurement model/analysis
- MCMC sampling for Bayesian inference
- Network Analysis

<http://wiki.socr.umich.edu/index.php/SMHS>

Features

- it is community-built
- open-access (dev & use)
- blends concepts with IT
- multi-lingual

SOCR Usage

- Course instructors and teachers
 - Class notes, interactive tools & activities useful for student motivation & concept demonstration
- Students and trainees
 - SOCR class notes, analyses, computational and graphing tools for learning & practicing
- Model developers, software engineers & researchers
 - Light-weight plug-in oriented SOCR computational libraries and infrastructure useful in their algorithm designs and research efforts.

SOCR = Statistical Computing Technology-based Instruction Open-Source Project

File Size: 100KB

Default: Deutsch Español Français Italiano Português 日本語

한국어 中文 繁體中文 Pycckий Nederlands Ελληνικά

SOCR Usage

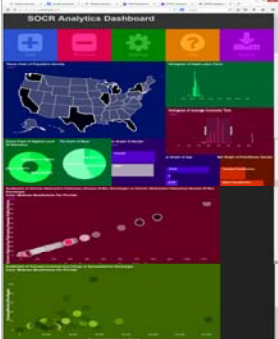
SOCR Visitor Log: Last 200 users (10:07:13, 3 Apr 2015)



www.socr.umich.edu/html/SOCR_UserGoogleMap.html

- > 8.5M active SOCR users (1 count/user/day)
- National & International users
- 18% Quarterly ↑

SOCR Dashboard

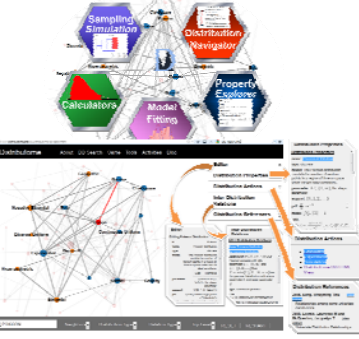


<http://socr.umich.edu/HTML5/Dashboard>

- Web-service combining and integrating multi-source socioeconomic and medical datasets
- Big data analytic processing
- Interface for exploratory navigation, manipulation and visualization
- Adding/removing of visual queries and interactive exploration of multivariate associations
- Powerful HTML5 technology enabling mobile on-demand computing

Husain, et al., J Big Data, 2015

Probability Distributome



<http://Distributome.org>

- Computational & graphical infrastructure
- Enables the discovery, exploration and application of diverse spectra of probability distributions.
- Human & machine interfaces for traversal, search, and navigation of all common probability distributions
- Distribution modeling, applications, investigation of inter-distribution relations, analytical representation & computational use
- Open-source, community-built, and web-based
- HTML5 & Web2.0 standards

Dinov, et al., CoSt, 2015

SOCR Instruction

• SOCR is Tested in:

- Lower Div., Upper Div. and Graduate Courses
- Major, Minor and Service Courses
- Small (~10) & Large Classes (100)
- 20 Instructors
- > 20 courses
- > 2,500 students

• Validated on Variety of Study Designs
Pooling results from 3 studies, SOCR effect $p < 0.00098$
(PMCID: PMC2740633)

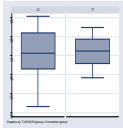
• Studies of student knowledge acquisition, learning-styles, attitudes towards Prob & Stats and quantitative performance

www.socr.umich.edu/html/SOCR_Publications.html

wiki.socr.umich.edu/index.php/SOCR_Courses

Demographics	Stat13 section1 Control	Stat13 section2 Treatment
Freshmen	24	7
Sophomores	18	14
Juniors	16	38
Seniors	23	29
Graduates	2	0
Total	83	88

Group	High	Low	Mean	Mean	Standard Deviation	Statistics
Midterm	Control	100	55	55.5	10	$\chi^2 = 1.57$ $p(100) = 0.969$
Treatment	100	48	48	48	10	$\chi^2 = 7.74$ $p(100) = 0.008$
Final	Control	100	42	40	10	$\chi^2 = 7.74$ $p(100) = 0.008$
Treatment	100	35	37	35.5	12	$\chi^2 = 4.89$ $p(100) = 0.028$
Overall Performance	Control	20-90	53.6	56.82	84.97	9.1
Treatment	20-90	42-72	58.26	55.48	9.9	9.1



SOCR Acknowledgments

Funded by
NSF 0442992, 0716055, 1023115
NIH U54 P20 NR015331, U54 EB020406,
P50 NS091856, P30 DK089503

Team
I Dinov, N Christou, J Sanchez, R Gould, S, Husain, B Athey,
J Cui, A Che, A Toga, K Siegrist, D Pearl, A Kalinin, ...



http://socr.umich.edu/html/SOCR_Acknowledgments.html
http://socr.umich.edu/html/SOCR_Publications.html
www.SOCR.umich.edu
www.StatisticsResource.org





SOCR Demos

- Functors & Distributions
- Experiments
- Analyses
- Charts
- Modeler
- Data

www.SOCR.umich.edu