Click to prove you're human

SPSS is a widely used program forstatistical analysisinsocial science. It is also used by market researchers, health researchers, survey companies, government, education researchers, marketing organizations, data minersand others. IBM SPSS is available for free on all Classroom PCs. It can be found in the "Common Applications Folder" under "Statistical Software". All University staff and students can install SPSS on a personal computer through the Software Download Service which can be accessed in theLinks section of this resource. You can download the software, installation instructions and license codes for PC, Mac or Linux. Tutorials, the IBM Knowledge Center and other information can also be found in the Links section. Related resources | MATLAB | NVivo | MS Excel | Information Governance Last updated on 22 October 2024 [{"Business Unit":{"code":"BU048","label":"IBM Software"},"Product":{"code":"SSLVMB","label":"IBM SPSS Statistics"},"Component":"SPSS Statistics Subscription","Platform":[{"code":"PF017","label":"Mac OS"},{"code":"PF033","label":"Windows"}],"Version":"All Versions","Edition":"","Line of Business":{"code":"LOB76","label":"Data Platform"}},{"Business Unit":{"code":"BU048","label":"IBM Software"},"Product":{"code":"SS3RA7","label":"IBM SPSS Modeler"},"Component":"SPSS Modeler Subscription;Watson Studio Desktop","Platform":[{"code":"PF017","label":"Mac OS"},{"code":"PF033","label":"Windows"}],"Version":"All Versions","Edition":"","Line of Business":{"code":"LOB76","label":"Data Platform"}}] Process of using data analysisNot to be confused with Statistical interference.Part of a series onResearchResearch designEthicsProposalQuestionWritingArgumentReferencingResearch strategyInterdisciplinaryMultimethodologyQualitativeArt-basedQuantitativePhilosophical schoolsAntipositivism ConstructivismCritical rationalismEmpiricismFallibilismPositivismPostpositivismPragmatismRealismCritical realismSubtle realismMethodologyAction researchArt methodologyCritical theoryGrounded theoryHermeneuticsHistoriographyHuman subject researchNarrative inquiryPhenomenologyPragmatismScientific methodMethodsAnalysisCase studyContent analysisDescriptive statisticsDiscourse analysisEthnographyAutoethnographyExperimentField experimentSocial experimentQuasi-experimentField researchHistorical methodInferential statisticsInterviewsMappingCultural mappingPhenomenographySecondary researchBibliometricsLiterature reviewMeta-analysisScoping reviewSystematic reviewScientific modellingSimulationSurveyTools and softwareArgument technologyGIS softwareLIS softwareBibliometricsReference managementScience softwareQualitative data analysisSimulationStatisticsPhilosophy portalvteStatistical inference is the process of using data analysis to infer properties of an underlying probability distribution.[1] Inferential statistical analysis infers properties of a population, for example by testing hypotheses and deriving estimates. It is assumed that the observed data set is sampled from a larger population.Inferential statistics can be contrasted with descriptive statistics. Descriptive statistics is solely concerned with properties of the observed data, and it does not rest on the assumption that the data come from a larger population. In machine learning, the term inference is sometimes used instead to mean "make a prediction, by evaluating an already trained model";[2] in this context inferring properties of the model is referred to as training or learning (rather than inference), and using a model for prediction is referred to as inference (instead of prediction); see also predictive inference.Statistical inference makes propositions about a population, using data drawn from the population with some form of sampling. Given a hypothesis about a population, for which we wish to draw inferences, statistical inference consists of (first) selecting a statistical model of the process that generates the data and (second) deducing propositions from the model.[3]Konishi and Kitagawa state "The majority of the problems in statistical inference can be considered to be problems related to statistical modeling".[4] Relatedly, Sir David Cox has said, "How [the] translation from subject-matter problem to statistical model is done is often the most critical part of an analysis".[5]The conclusion of a statistical inference is a statistical proposition.[6] Some common forms of statistical proposition are the following:a point estimate, i.e. a particular value that best approximates some parameter of interest;an interval estimate, e.g. a confidence interval (or set estimate), i.e. an interval constructed using a dataset drawn from a population so that, under repeated sampling of such datasets, such intervals would contain the true parameter value with the probability at the stated confidence level;a credible interval, i.e. a set of values containing, for example, 95% of posterior belief;rejection of a hypothesis;[note 1]clustering or classification of data points into groups.Main articles: Statistical model and Statistical assumptionsAny statistical inference requires some assumptions. A statistical model is a set of assumptions concerning the generation of the observed data and similar data. Descriptions of statistical models usually emphasize the role of population quantities of interest, about which we wish to draw inference.[7] Descriptive statistics are typically used as a preliminary step before more formal inferences are drawn.[8]Statisticians distinguish between three levels of modeling assumptions:Fully parametric: The probability distributions describing the data-generation process are assumed to be fully described by a family of probability distributions involving only a finite number of unknown parameters.[7] For example, one may assume that the distribution of population values is truly Normal, with unknown mean and variance, and that datasets are generated by 'simple' random sampling. The family of generalized linear models is a widely used and flexible class of parametric models.Non-parametric: The assumptions made about the process generating the data are much less than in parametric statistics and may be minimal.[9] For example, every continuous probability distribution has a median, which may be estimated using the sample median or the HodgesLehmannSen estimator, which has good properties when the data arise from simple random sampling.Semi-parametric: This term typically implies assumptions 'in between' fully and non-parametric approaches. For example, one may assume that a population distribution has a finite mean. Furthermore, one may assume that the mean response level in the population depends in a truly linear manner on some covariate (a parametric assumption) but not make any parametric assumption describing the variance around that mean (i.e. about the presence or possible form of any heteroscedasticity). More generally, semi-parametric models can often be separated into 'structural' and 'random variation' components. One component is treated parametrically and the other non-parametrically. The well-known Cox model is a set of semi-parametric assumptions.[citation needed]See also: Statistical model validationThe above image shows a histogram assessing the assumption of normality, which can be illustrated through the even spread underneath the bell curve.Whatever level of assumption is made, correctly calibrated inference, in general, requires these assumptions to be correct; i.e. that the data-generating mechanisms really have been correctly specified.Incorrect assumptions of 'simple' random sampling can invalidate statistical inference.[10] More complex semi- and fully parametric assumptions are also cause for concern. For example, incorrectly assuming the Cox model can in some cases lead to faulty conclusions.[11] Incorrect assumptions of Normality in the population also invalidates some forms of regression-based inference.[12] The use of any parametric model is viewed skeptically by most experts in sampling human populations: "most sampling statisticians, when they deal with confidence intervals at all, limit themselves to statements about [estimators] based on very large samples, where the central limit theorem ensures that these [estimators] will have distributions that are nearly normal."[13] In particular, a normal distribution "would be a totally unrealistic and catastrophically unwise assumption to make if we were dealing with any kind of economic population."[13] Here, the central limit theorem states that the distribution of the sample mean "for very large samples" is approximately normally distributed, if the distribution is not heavy-tailed.Main articles: Statistical distance, Asymptotic theory (statistics), and Approximation theoryGiven the difficulty in specifying exact distributions of sample statistics, many methods have been developed for approximating these.With finite samples, approximation results measure how close a limiting distribution approaches the statistic's sample distribution: For example, with 10,000 independent samples the normal distribution approximates (to two digits of accuracy) the distribution of the sample mean for many population distributions, by the BerryEsseen theorem.[14] Yet for many practical purposes, the normal approximation provides a good approximation to the sample-mean's distribution when there are 10 (or more) independent samples, according to simulation studies and statisticians' experience.[14] Following Kolmogorov's work in the 1950s, advanced statistics uses approximation theory and functional analysis to quantify the error of approximation. In this approach, the metric geometry of probability distributions is studied; this approach quantifies approximation error with, for example, the KullbackLeibler divergence, Bregman divergence, and the Hellinger distance.[15][16][17]With indefinitely large samples, limiting results like the central limit theorem describe the sample statistic's limiting distribution if one exists. Limiting results are not statements about finite samples, and indeed are irrelevant to finite samples.[18][19][20] However, the asymptotic theory of limiting distributions is often invoked for work with finite samples. For example, limiting results are often invoked to justify the generalized method of moments and the use of generalized estimating equations, which are popular in econometrics and biostatistics. The magnitude of the difference between the limiting distribution and the true distribution (formally, the 'error' of the approximation) can be assessed using simulation.[21] The heuristic application of limiting results to finite samples is common practice in many applications, especially with low-dimensional models with log-concave likelihoods (such as with one-parameter exponential families).Main article: RandomizationSee also: Random sample and Random assignmentFor a given dataset that was produced by a randomization design, the randomization distribution of a statistic (under the null-hypothesis) is defined by evaluating the test statistic for all of the plans that could have been generated by the randomization design. In frequentist inference, the randomization allows inferences to be based on the randomization distribution rather than a subjective model, and this is important especially in survey sampling and design of experiments.[22][23] Statistical inference from randomized studies is also more straightforward than many other situations.[24][25][26] In Bayesian inference, randomization is also of importance: in survey sampling, use of sampling without replacement ensures the exchangeability of the sample with the population; in randomized experiments, randomization warrants a missing at random assumption for covariate information.[27]Objective randomization allows properly inductive procedures.[28][29][30][31][32] Many statisticians prefer randomization-based analysis of data that was generated by well-defined randomization procedures.[33] (However, it is true that in fields of science with developed theoretical knowledge and experimental control, randomized experiments may increase the costs of experimentation without improving the quality of inferences.[34][35]) Similarly, results from randomized experiments are recommended by leading statistical authorities as allowing inferences with greater reliability than do observational studies of the same phenomena.[36] However, a good observational study may be better than a bad randomized experiment.The statistical analysis of a randomized experiment may be based on the randomization scheme stated in the experimental protocol and does not need a subjective model.[37][38]However, at any time, some hypotheses cannot be tested using objective statistical models, which accurately describe randomized experiments or random samples. In some cases, such randomized studies are uneconomical or unethical.It is standard practice to refer to a statistical model, e.g., a linear or logistic models, when analyzing data from randomized experiments.[39] However, the randomization scheme guides the choice of a statistical model. It is not possible to choose an appropriate model without knowing the randomization scheme.[23] Seriously misleading results can be obtained analyzing data from randomized experiments while ignoring the experimental protocol; common mistakes include forgetting the blocking used in an experiment and confusing repeated measurements on the same experimental unit with independent replicates of the treatment applied to different experimental units.[40]Model-free techniques provide a complement to model-based methods, which employ reductionist strategies of reality-simplification. The former combine, evolve, ensemble and train algorithms dynamically adapting to the contextual affinities of a process and learning the intrinsic characteristics of the observations.[41][42]For example, model-free simple linear regression is based either on:a random design, where the pairs of observations $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ are independent and identically distributed (iid),or a deterministic design, where the variables $X_1, X_2, \ldots, X_n$ are deterministic, but the corresponding response variables $Y_1, Y_2, \ldots, Y_n$ are random and independent with a common conditional distribution, i.e., $P(Y_j \mid X_j = x) = D_x(y)$, which is independent of the index j. In either case, the model-free randomization inference for features of the common conditional distribution $D_x(\cdot)$ relies on some regularity conditions, e.g. functional smoothness. For instance, model-free randomization inference for the population feature conditional mean, $(x) = E(Y \mid X = x)$ can be consistently estimated via local averaging or local polynomial fitting, under the assumption that (x) is smooth. Also, relying on asymptotic normality or resampling, we can construct confidence intervals for the population feature, in this case, the conditional mean, (x).[43]Different schools of statistical inference have become established. These schoolsor "paradigms"are not mutually exclusive, and methods that work well under one paradigm often have attractive interpretations under other paradigms.Bandyopadhyay and Forster describe four paradigms: The classical (or frequentist) paradigm, the Bayesian paradigm, the likelihoodist paradigm, and the Akaikean-Information Criterion-based paradigm.[44]Main article: Frequentist inferenceThis paradigm calibrates the plausibility of propositions by considering (notional) repeated sampling of a population distribution to produce datasets similar to the one at hand. By considering the dataset's characteristics under repeated sampling, the frequentist properties of a statistical proposition can be quantifiedalthough in practice this quantification may be challenging.p-valueConfidence intervalNull hypothesis significance testingOne interpretation of frequentist inference (or classical inference) is that it is applicable only in terms of frequency probability; that is, in terms of repeated sampling from a population. However, the approach of Neyman[45] develops these procedures in terms of pre-experiment probabilities. That is, before undertaking an experiment, one decides on a rule for coming to a conclusion such that the probability of being correct is controlled in a suitable way: such a probability need not have a frequentist or repeated sampling interpretation. In contrast, Bayesian inference works in terms of conditional probabilities (i.e. probabilities conditional on the observed data), compared to the marginal (but conditioned on unknown parameters) probabilities used in the frequentist approach.The frequentist procedures of significance testing and confidence intervals can be constructed without regard to utility functions. However, some elements of frequentist statistics, such as statistical decision theory, do incorporate utility functions.[citation needed] In particular, frequentist developments of optimal inference (such as minimum-variance unbiased estimators, or uniformly most powerful testing) make use of loss functions, which play the role of (negative) utility functions. Loss functions need not be explicitly stated for statistical theorists to prove that a statistical procedure has an optimality property.[46] However, loss-functions are often useful for stating optimality properties: for example, median-unbiased estimators are optimal under absolute value loss functions, in that they minimize expected loss, and least squares estimators are optimal under squared error loss functions, in that they minimize expected loss.While statisticians using frequentist inference must choose for themselves the parameters of interest, and the estimators/test statistic to be used, the absence of obviously explicit utilities and prior distributions has helped frequentist procedures to become widely viewed as 'objective'.[47]See also: Bayesian inferenceThe Bayesian calculus describes degrees of belief using the 'language' of probability; beliefs are positive, integrate into one, and obey probability axioms. Bayesian inference uses the available posterior beliefs as the basis for making statistical propositions.[48] There are several different justifications for using the Bayesian approach.Credible interval for interval estimationBayes factors for model comparisonMany informal Bayesian inferences are based on "intuitively reasonable" summaries of the posterior. For example, the posterior mean, median and mode, highest posterior density intervals, and Bayes Factors can all be motivated in this way. While a user's utility function need not be stated for this sort of inference, these summaries do all depend (to some extent) on stated prior beliefs, and are generally viewed as subjective conclusions. (Methods of prior construction which do not require external input have been proposed but not yet fully developed.)Formally, Bayesian inference is calibrated with reference to an explicitly stated utility, or loss function; the 'Bayes rule' is the one which maximizes expected utility, averaged over the posterior uncertainty. Formal Bayesian inference therefore automatically provides optimal decisions in a decision theoretic sense. Given assumptions, data and utility, Bayesian inference can be made for essentially any problem, although not every statistical inference need have a Bayesian interpretation. Analyses which are not formally Bayesian can be (logically) incoherent; a feature of Bayesian procedures which use proper priors (i.e. those integrable to one) is that they are guaranteed to be coherent. Some advocates of Bayesian inference assert that inference must take place in this decision-theoretic framework, and that Bayesian inference should not conclude with the evaluation and summarization of posterior beliefs.Main article: LikelihoodismLikelihood-based inference is a paradigm used to estimate the parameters of a statistical model based on observed data. Likelihoodism approaches statistics by using the likelihood function, denoted as $L(x \mid)$, quantifies the probability of observing the given data x, assuming a specific set of parameter values θ. In likelihood-based inference, the goal is to find the set of parameter values that maximizes the likelihood function, or equivalently, maximizes the probability of observing the given data.The process of likelihood-based inference usually involves the following steps:Formulating the statistical model: A statistical model is defined based on the problem at hand, specifying the distributional assumptions and the relationship between the observed data and the unknown parameters. The model can be simple, such as a normal distribution with known variance, or complex, such as a hierarchical model with multiple levels of random effects.Constructing the likelihood function: Given the statistical model, the likelihood function is constructed by evaluating the joint probability density or mass function of the observed data as a function of the unknown parameters. This function represents the probability of observing the data for different values of the parameters.Maximizing the likelihood function: The next step is to find the set of parameter values that maximizes the likelihood function. This can be achieved using optimization techniques such as numerical optimization algorithms. The estimated parameter values, often denoted as $\bar{y}$, are the maximum likelihood estimates (MLEs).Assessing uncertainty: Once the MLEs are obtained, it is crucial to quantify the uncertainty associated with the parameter estimates. This can be done by calculating standard errors, confidence intervals, or conducting hypothesis tests based on asymptotic theory or simulation techniques such as bootstrapping.Model checking: After obtaining the parameter estimates and assessing their uncertainty, it is important to assess the adequacy of the statistical model. This involves checking the assumptions made in the model and evaluating the fit of the model to the data using goodness-of-fit tests, residual analysis, or graphical diagnostics.Inference and interpretation: Finally, based on the estimated parameters and model assessment, statistical inference can be performed. This involves drawing conclusions about the population parameters, making predictions, or testing hypotheses based on the estimated model.Main article: Akaike information criterionThis section needs expansion. You can help by adding to it. (November 2017)The Akaike information criterion (AIC) is an estimator of the relative quality of statistical models for a given set of data. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Thus, AIC provides a means for model selection.AIC is founded on information theory: it offers an estimate of the relative information lost when a given model is used to represent the process that generated the data. (In doing so, it deals with the trade-off between the goodness of fit of the model and the simplicity of the model.)Main article: Minimum description lengthThe minimum description length (MDL) principle has been developed from ideas in information theory[49] and the theory of Kolmogorov complexity.[50] The (MDL) principle selects statistical models that maximally compress the data; inference proceeds without assuming counterfactual or non-falsifiable "data-generating mechanisms" or probability models for the data, as might be done in frequentist or Bayesian approaches.However, if a "data generating mechanism" does exist in reality, then according to Shannon's source coding theorem it provides the MDL description of the data, on average and asymptotically.[51] In minimizing description length (or descriptive complexity), MDL estimation is similar to maximum likelihood estimation and maximum a posteriori estimation (using maximum-entropy Bayesian priors). However, MDL avoids assuming that the underlying probability model is known; the MDL principle can also be applied without assumptions that e.g. the data arose from independent sampling.[51][52]The MDL principle has been applied in communication-coding theory in information theory, in linear regression,[52] and in data mining.[50]The evaluation of MDL-based inferential procedures often uses techniques or criteria from computational complexity theory.[53]Main article: Fiducial inferenceFiducial inference was an approach to statistical inference based on fiducial probability, also known as a "fiducial distribution". In subsequent work, this approach has been called ill-defined, extremely limited in applicability, and even fallacious.[54][55] However this argument is the same as that which shows[56] that a so-called confidence distribution is not a valid probability distribution and, since this has not invalidated the application of confidence intervals, it does not necessarily invalidate conclusions drawn from fiducial arguments. An attempt was made to reinterpret the early work of Fisher's fiducial argument as a special case of an inference theory using upper and lower probabilities.[57]Developing ideas of Fisher and of Pitman from 1938 to 1939,[58] George A. Barnard developed "structural inference" or "pivotal inference",[59] an approach using invariant probabilities on group families. Barnard reformulated the arguments behind fiducial inference on a restricted class of models on which "fiducial" procedures would be well-defined and useful. Donald A. S. Fraser developed a general theory for structural inference[60] based on group theory and applied this to linear models.[61] The theory formulated by Fraser has close links to decision theory and Bayesian statistics and can provide optimal frequentist decision rules if they exist.[62]The topics below are usually included in the area of statistical inference.Statistical assumptionsStatistical decision theoryEstimation theoryStatistical hypothesis testingRevising opinions in statisticsDesign of experiments, the analysis of variance, and regressionSurvey samplingSummarizing statistical dataPredictive inference is an approach to statistical inference that emphasizes the prediction of future observations based on past observations.Initially, predictive inference was based on observable parameters and it was the main purpose of studying probability,[citation needed] but it fell out of favor in the 20th century due to a new parametric approach pioneered by Bruno de Finetti. The approach modeled phenomena as a physical system observed with error (e.g., celestial mechanics). De Finetti's idea of exchangeabilitythat future observations should behave like past observationscame to the attention of the English-speaking world with the 1974 translation from French of his 1937 paper,[63] and has since been propounded by such statisticians as Seymour Geisser.[64]Algorithmic inferenceInduction (philosophy)Informal inferential reasoningInformation field theoryPopulation proportionPhilosophy of statisticsPrediction intervalPredictive analyticsPredictive modellingStylometry^ According to Peirce, acceptance means that inquiry on this question ceases for the time being. In science, all scientific theories are revisable.^ Upton, G., Cook, I. (2008) Oxford Dictionary of Statistics, OUP. ISBN978-0-19-954145-4.^ "TensorFlow Lite inference". The term inference refers to the process of executing a TensorFlow Lite model on-device in order to make predictions based on input data.^ Johnson, Richard (12 March 2016). "Statistical Inference". Encyclopedia of Mathematics. Springer: The European Mathematical Society. Retrieved 26 October 2022.^ Konishi & Kitagawa (2008), p. 75.^ Cox (2006), p. 197.^ "Statistical inference - Encyclopedia of Mathematics". www.encyclopediaofmath.org. Retrieved 2019-01-23.^ a b Cox (2006) page 2^ Evans, Michael; etal. (2004). Probability and Statistics: The Science of Uncertainty. Freeman and Company. p.267. ISBN9780716747420.^ van der Vaart, A.W. (1998) Asymptotic Statistics Cambridge University Press. ISBN0-521-78450-6 (page 341)^ Kruskal 1988^ Freedman, D.A. (2008) "Survival analysis: An Epidemiological hazard?". The American Statistician (2008) 62: 110-119. (Reprinted as Chapter 11 (pages 169192) of Freedman (2010)).^ Berk, R. (2003) Regression Analysis: A Constructive Critique (Advanced Quantitative Techniques in the Social Sciences) (v. 11) Sage Publications. ISBN0-7619-2904-5^ a b Brewer, Ken (2002). Combined Survey Sampling Inference: Weighing of Basu's Elephants. Hodder Arnold. p.6. ISBN978-0340692295.^ a b Jrgen Hoffman-Jrgensen's Probability With a View Towards Statistics, Volume I. Page 399 [full citation needed]^ Le Cam (1986) [pageneeded]^ Erik Torgerson (1991) Comparison of Statistical Experiments, volume 36 of Encyclopedia of Mathematics. Cambridge University Press. [full citation needed]^ Liese, Friedrich & Mieschke, Klaus-J. (2008). Statistical Decision Theory: Estimation, Testing, and Selection. Springer. ISBN978-0-387-73193-3.^ Kolmogorov (1963, p.369): "The frequency concept, based on the notion of limiting frequency as the number of trials increases to infinity, does not contribute anything to substantiate the applicability of the results of probability theory to real practical problems where we have always to deal with a finite number of trials".^ "Indeed, limit theorems 'as n $\displaystyle n$ tends to infinity' are logically devoid of content about what happens at any particular n $\displaystyle n$. All they can do is suggest certain approaches whose performance must then be checked on the case at hand." Le Cam (1986) (page xiv)^ Pfanzagl (1994): "The crucial drawback of asymptotic theory: What we expect from asymptotic theory are results which hold approximately . . . . What asymptotic theory has to offer are limit theorems."(page ix) "What counts for applications are approximations, not limits." (page 188)^ Pfanzagl (1994): "By taking a limit theorem as being approximately true for large sample sizes, we commit an error the size of which is unknown. [. . .] Realistic information about the remaining errors may be obtained by simulations." (page ix)^ Neyman, J.(1934) "On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection", Journal of the Royal Statistical Society, 97 (4), 557625 JSTOR2342192^ a b Hinkelmann and Kempthorne(2008) [pageneeded]^ ASA Guidelines for the first course in statistics for non-statisticians. (available at the ASA website)^ David A. Freedman et alia's Statistics.^ Moore et al. (2015).^ Gelman A. et al. (2013). Bayesian Data Analysis (Chapman & Hall).^ Peirce (1877-1878)^ Peirce (1883)^ Freedman, Pisani & Purves 1978.^ David A. Freedman Statistical Models.^ Rao, C.R. (1997) Statistics and Truth: Putting Chance to Work, World Scientific. ISBN981-02-3111-3^ Peirce; Freedman; Moore et al. (2015).[citation needed]^ Box, G.E.P. and Friends (2006) Improving Almost Anything: Ideas and Essays, Revised Edition, Wiley. ISBN978-0-471-72755-2^ Cox (2006), p.196.^ ASA Guidelines for the first course in statistics for non-statisticians. (available at the ASA website)David A. Freedman et alias Statistics.Moore et al. (2015).^ Neyman, Jerzy. 1923 [1990]. "On the Application of Probability Theory to AgriculturalExperiments. Essay on Principles. Section 9." Statistical Science 5 (4): 465472. Trans. Dorota M. Dabrowska and Terence P. Speed.^ Hinkelmann & Kempthorne (2008) [pageneeded]^ Dinov, Ivo; Palanimalai, Selvam; Khare, Ashwini; Christou, Nicolas (2018). "Randomization-based statistical inference: A resampling and simulation infrastructure". Teaching Statistics. 40 (2): 6473. doi:10.1111/test.12156. PMC6155997. PMID30270947.^ Hinkelmann and Kempthorne (2008) Chapter 6.^ Dinov, Ivo; Palanimalai, Selvam; Khare, Ashwini; Christou, Nicolas (2018). "Randomization-based statistical inference: A resampling and simulation infrastructure". Teaching Statistics. 40 (2): 6473. doi:10.1111/test.12156. PMC6155997. PMID30270947.^ Tang, Ming; Gao, Chao; Goutman, Stephen; Kalinin, Alexandr; Mukherjee, Bhramar; Guan, Yuanfang; Dinov, Ivo (2019). "Model-Based and Model-Free Techniques for Amyotrophic Lateral Sclerosis Diagnostic Prediction and Patient Clustering". Neuroinformatics. 17 (3): 407421. doi:10.1007/s12021-018-9406-9. PMC6527505. PMID30460455.^ Politis, D.N. (2019). "Model-free inference in statistics: how and why". IMS Bulletin. 48.^ Bandyopadhyay & Forster (2011). See the book's Introduction (p.3) and "SectionIII: Four Paradigms of Statistics".^ Neyman, J. (1937). "Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability". Philosophical Transactions of the Royal Society of London A. 236 (767): 333380. Bibcode:1937RSPTA.236..333N. doi:10.1098/rsta.1937.0005. JSTOR91337.^ Preface to Pfanzagl.^ Little, Roderick J. (2006). "Calibrated Bayes: A Bayes/Frequentist Roadmap". The American Statistician. 60 (3): 213223. doi:10.1198/000313006X117837. ISSN0003-1305. JSTOR27643780. S2CID53505632.^ Lee, Se Yoon (2021). "Gibbs sampler and coordinate ascent variational inference: A set-theoretical review". Communications in Statistics - Theory and Methods. 51 (6): 15491568. arXiv:2008.01006. doi:10.1080/03610926.2021.1921214. S2CID220935477.^ Soofi (2000)^ a b Hansen & Yu (2001)^ a b Hansen and Yu (2001), page 747.^ a b Rissanen (1989), page 84^ Joseph F. Traub, G. W. Wasilkowski, and H. Wozniakowski. (1988) [pageneeded]^ Neyman (1956)^ Zabell (1992)^ Cox (2006) page 66^ Hampel 2003.^ Davison, page 12. [full citation needed]^ Barnard, G.A. (1995) "Pivotal Models and the Fiducial Argument", International Statistical Review, 63 (3), 309323. JSTOR1403482^ Fraser, D. A. S. (1968). The structure of inference. New York: Wiley. ISBN0-471-27548-4. OCLC440926.^ Fraser, D. A. S. (1979). Inference and linear models. London: McGraw-Hill. ISBN0-07-021910-9. OCLC3559629.^ Taraldsen, Gunnar; Lindqvist, Bo Henry (2013-02-01). "Fiducial theory and optimal inference". The Annals of Statistics. 41 (1). arXiv:1301.1717. doi:10.1214/13-AOS1083. ISSN0090-5364. S2CID88520957.^ De Finetti, Bruno (1937). "La Prvision: ses lois logiques, ses sources subjectives". Annales de l'Institut Henri Poincar. 7 (1): 168. ISSN0365-320X. Translated in De Finetti, Bruno (1992). "Foresight: Its Logical Laws, Its Subjective Sources". Breakthroughs in Statistics. Springer Series in Statistics. pp.134174. doi:10.1007/978-1-4612-0919-5_10. ISBN978-0-387-94037-3.^ Geisser, Seymour (1993) Predictive Inference: An Introduction, CRC Press. ISBN0-412-03471-9Bandyopadhyay, P. S.; Forster, M. R., eds. (2011), Philosophy of Statistics, Elsevier.Bickel, Peter J.; Doksum, Kjell A. (2001). Mathematical statistics: Basic and selected topics. Vol.1 (Second (updated printing 2007)ed.). Prentice Hall. ISBN978-0-13-850363-5. MR0443141.Cox, D. R. (2006). Principles of Statistical Inference, Cambridge University Press. ISBN0-521-68567-2.Fisher, R. A. (1955), "Statistical methods and scientific induction", Journal of the Royal Statistical Society, Series B, 17, 6978. (criticism of statistical theories of Jerzy Neyman and Abraham Wald)Freedman, D. A. (2009). Statistical Models: Theory and practice (reviseded.). Cambridge University Press. pp.xiv+442 pp. ISBN978-0-521-74385-3. MR2489600.Freedman, D. A. (2010). Statistical Models and Causal Inferences: A Dialogue with the Social Sciences (Edited by David Collier, Jasjeet Sekhon, and Philip B. Stark), Cambridge University Press.Hampel, Frank R. (February 2003). "The proper fiducial argument". Seminar fr Statistik, Eidgenssische Technische Hochschule. 114. doi:10.3929/ethz-a-004526011.Hansen, Mark H.; Yu, Bin (June 2001). "Model Selection and the Principle of Minimum Description Length: Review paper". Journal of the American Statistical Association. 96 (454): 746774. CiteSeerX10.1.1.43.6581. doi:10.1198/016214501753168398. JSTOR2670311. MR1939352. S2CID14460386. Archived from the original on 2004-11-16.Hinkelmann, Klaus; Kempthorne, Oscar (2008). Introduction to Experimental Design (Seconded.). Wiley. ISBN978-0-471-72756-9.Kolmogorov, Andrei N. (1963). "On tables of random numbers". Sankhy Ser. A. 25: 369375. MR0178484. Reprinted as Kolmogorov, Andrei N. (1998). "On tables of random numbers". Theoretical Computer Science. 207 (2): 387395. doi:10.1016/S0304-3975(98)00075-9. MR1643414.Konishi S., Kitagawa G. (2008), Information Criteria and Statistical Modeling, Springer.Kruskal, William (December 1988). "Miracles and statistics: the casual assumption of independence (ASA Presidential Address)". Journal of the American Statistical Association. 83 (404): 929940. doi:10.2307/2290117. JSTOR2290117.Le Cam, Lucian. (1986) Asymptotic Methods of Statistical Decision Theory, Springer. ISBN0-387-96307-3Moore, D. S.; McCabe, G. P.; Craig, B. A. (2015), Introduction to the Practice of Statistics, Eighth Edition, Macmillan.Neyman, Jerzy (1956). "Note on an article by Sir Ronald Fisher". Journal of the Royal Statistical Society, Series B. 18 (2): 288294. doi:10.1111/j.2517-6161.1956.tb00236.x. JSTOR2983716. (reply to Fisher 1955)Peirce, C. S. (18771878), "Illustrations of the logic of science" (series), Popular Science Monthly, vols. 1213. Relevant individual papers:(1878 March), "The Doctrine of Chances", Popular Science Monthly, v. 12, March issue, pp. 604615. Internet Archive Eprint.(1878 April), "The Probability of Induction", Popular Science Monthly, v. 12, pp. 705718. Internet Archive Eprint.(1878 June), "The Order of Nature", Popular Science Monthly, v. 13, pp. 203217.Internet Archive Eprint.(1878 August), "Deduction, Induction, and Hypothesis", Popular Science Monthly, v. 13, pp. 470482. Internet Archive Eprint.Peirce, C. S. (1883), "A Theory of probable inference", Studies in Logic, pp. 126-181, Little, Brown, and Company. (Reprinted 1983, John Benjamins Publishing Company, ISBN90-272-3271-7)Freedman, D.A; Pisani, R.; Purves, R.A. (1978). Statistics. New York: W. W. Norton & Company.Pfanzagl, Johann; with the assistance of R. Hambker (1994). Parametric Statistical Theory. Berlin: Walter de Gruyter. ISBN978-3-11-013863-4. MR1291393.Rissanen, Jorma (1989). Stochastic Complexity in Statistical Inquiry. Series in Computer Science. Vol.15. Singapore: World Scientific. ISBN978-9971-5-0859-3. MR1082556.Soofi, Ehsan S. (December 2000). "Principal information-theoretic approaches (Vignettes for the Year 2000: Theory and Methods, ed. by George Casella)". Journal of the American Statistical Association. 95 (452): 13491353. doi:10.1080/01621459.2000.10474346. JSTOR2669786. MR1825292. S2CID120143121.Traub, Joseph F.; Wasilkowski, G. W.; Wozniakowski, H. (1988). Information-Based Complexity. Academic Press. ISBN978-0-12-697545-1.Zabell, S. L. (Aug 1992). "R. A. Fisher and Fiducial Argument". Statistical Science. 7 (3): 369387. doi:10.1214/ss/1177011233. JSTOR2246073.Casella, G., Berger, R. L. (2002). Statistical Inference. Duxbury Press. ISBN0-534-24312-6Freedman, D.A. (1991). "Statistical models and shoe leather". Sociological Methodology. 21: 291313. doi:10.2307/270939. JSTOR270939.Held L., Bov D.S. (2014). Applied Statistical InferenceLikelihood and Bayes (Springer).Lenhard, Johannes (2006). "Models and Statistical Inference: the controversy between Fisher and NeymanPearson" (PDF). British Journal for the Philosophy of Science. 57: 6991. doi:10.1093/bjps/axi152. S2CID14136146.Lindley, D (1958). "Fiducial distribution and Bayes' theorem". Journal of the Royal Statistical Society, Series B. 20: 1027. doi:10.1111/j.2517-6161.1958.tb00278.x.Rahlf, Thomas (2014). "Statistical Inference", in Claude Diebolt, and Michael Haupert (eds.), "Handbook of Cliometrics (Springer Reference Series)", Berlin/Heidelberg: Springer.Reid, N.; Cox, D. R. (2014). "On Some Principles of Statistical Inference". International Statistical Review. 83 (2): 293308. doi:10.1111/insr.12067. hdl:10.1111/insr.12067. S2CID17410547.Sagitov, Serik (2022). "Statistical Inference". Wikibooks. G.A., Smith, R.L. (2005). Essentials of Statistical Inference, CUP. ISBN0-521-83971-8Wikimedia Commons has media related to Statistical inference.Wikiversity has learning resources about Statistical inferenceStatistical Inference lecture on the MIT OpenCourseWare platformStatistical Inference lecture by the National Programme on Technology Enhanced LearningAn online, Bayesian (MCMC) demo/calculator is available at causaScientiaPortal: MathematicsRetrieved from " Get authorization codes for 10 or fewer new licenses Authorization codes for new orders will be sent to primary contacts e-mail. Read details Get authorization codes for more than 10 licenses License Key Center Portal will be used to generate authorization codes. Login to portal Download perpetual and fixed term software You can download SPSS Statistics through Passport Advantage. Connect now Download concurrent license manager Concurrent license manager controls your license to run the software. Download Sentinel Open a technical support case You can open a support ticket for license-key related or technical issues. Open a ticket Download an IBM SPSS Statistics subscription You can download and activate your subscription with your IBM ID. Connect to MyIBM Resolve download issues Passport Advantage Support can help if you have trouble downloading perpetual or fixed-term software. Contact us Get support for IBM SPSS Statistics GradPack GradPack vendors provide support for IBM SPSS Statistics GradPack. Contact your vendor Download invoices Accessor download your invoices on Invoices@IBM Portal. Access invoices@IBM Update primary contact Get assistance with changing primary and technical contact information. Contact IBM eCustomer care Add seats to your subscription Add seats to your IBM SPSS Statistics Subscription through MyIBM. Connect to MyIBM Reduce seats in your subscription Contact us directly to reduce seats in your subscription. Open a ticket Cancel your subscription Cancel your IBM SPSS Statistics subscription through MyIBM. Connect to MyIBM Renew perpetual license support IBM will contact you when its time to renew your yearly support. Learn how it works Get help on expired Subscription and Support license Visit our forum for licensing support if your subscription and support has ended. Check out the forum Manage or transfer licenses See how to manage your existing devices and users. Learn more Watch demo videos Watch demos of key capabilities within IBM SPSS Statistics. Go to video library Renew support Protect your software investment through subscription and support (S&S). Learn how it works Connect with Passport Advantage Order through IBM programs, such as Passport Advantage and Passport Advantage Express. Learn more Open a licensing case Learn how to return or delete an authorization code automatically. Learn how Receive notifications Get notified when new software fixes, releases and support become available. Learn how Get SPSS Statistics GradPack support Contact your GradPack vendor to get support. Contact your vendor SPSS Statistics Community Get access to a community of experts for all your questions. Join the community Support forum Get technical support for licensing, errors, installation, trial set-up and more. Go to support forum Stack Overflow Check out IBM SPSS Statistics on Stack Overflow. Go to Stack Overflow On the IBM SPSS Statistics Support page, you will find support information related to downloading software, opening support tickets, obtaining software authorization codes, and much more. Visit the Support page The Subscription application does not require authorization codes. Use your IBMid to activate the software. The SPSS Statistics paid Subscription plan can be used offline, but it requires a user to authenticate their license at least once every two weeks. After 14 days of offline access, the application will stop functioning until an internet connection is established and you log in with your IBMid. A working internet connection is required for this step. We recommend that you download the free trial subscription to ensure compatibility prior to making purchase. When you purchase SPSS Statistics Subscription, the only information in the cloud is information required to manage your subscription account. All functions and data processed within the application are run on your machine and not in the cloud. Sorry to interruptCSS Error The IBM SPSS software platform offers advanced statistical analysis, a vast library of machine learning algorithms, text analysis, open-source extensibility, integration with big data and seamless deployment into applications. Its ease of use, flexibility and scalability make SPSS accessible to users of all skill levels. Whats more, its suitable for projects of all sizes and levels of complexity, and can help you find new opportunities,