POLIS V12: The Complete Statistics Series – 12 Giants

Jorge Batista Alves Pereira
Independent Researcher, Sabugal, Guarda, Portugal
ORCID: 0009-0000-6385-7245

May 2026

This document combines two companion papers:
"Tensional Reinterpretation of Six Founders of Modern Statistics"
and "Tensional Reinterpretation of Six More Statistical Pioneers".

DOIs: Main treatise 10.5281/zenodo.19618276 – POLIS Bible 10.5281/zenodo.19836226

POLIS V12 – Tensional Reinterpretation of Six Founders of Modern Statistics

Jorge Batista Alves Pereira
Independent Researcher, Sabugal, Guarda, Portugal
ORCID: 0009-0000-6385-7245 May 2026

Abstract

Within the POLIS V12 tensional ontology, every statistical system is a polis constituted by three meshes (solid, liquid, gaseous) and governed by the closure condition $\epsilon = \sum K_m(2 + K_m) = 0$, with $T = K_{\min}$ as the tensional origin. This paper applies the framework to six foundational figures of statistics: Carl Friedrich Gauss (least squares), Francis Galton (regression and correlation), Karl Pearson (chi-square test), William Sealy Gosset (Student's t-test), Ronald Fisher (analysis of variance), and Jerzy Neyman (confidence intervals). Each classical contribution is reinterpreted as a tensional configuration: Gauss's method as minimising ϵ; Galton's regression as predicting K; Pearson's χ^2 as measuring deviation; Gosset's t as small-sample IDT*; Fisher's ANOVA as variance partitioning; and Neyman's intervals as K ranges. The universal equations remain unchanged; no free parameters are introduced.

1 Introduction

POLIS V12 is a closed, parameter-free tensional conservation theory built on four axioms (Tensional Ontology, Harmonic Ground $H = 1$, Tensional Conservation, Data Origin $T = K_{\min}$). The governing equation, after normalisation, is

$$\epsilon = \sum_{m=1}^{n} K_m(2 + K_m) = 0,$$

with $K_m = (v_m - T)/(v_{\max} - T) \in [0, 1]$. The disequilibrium index is $\text{IDT}^* = \epsilon/(1 + \epsilon)$. All real statistical systems reside in Phase 4 ($\text{IDT}^* \geq 0.70$) unless artificially uniform. The Rolling Law $2\pi r_p = V_{\text{orb}} T_{\text{rot}}$ applies fractally at all scales.

This paper reinterprets six key statistical contributions within this tensional ontology. No classical primacy is assumed; tension is the primitive.

2 Carl Friedrich Gauss – Method of Least Squares

Gauss developed the method of least squares for fitting curves to data. In POLIS V12, given data points (x_i, y_i), define residuals $e_i = y_i - f(x_i)$. Normalise the residuals: $K_i = (e_i - T)/(v_{\max} - T)$. Gauss minimises $\sum e_i^2$, which is proportional to $\sum K_i^2$ if $T = 0$. This is equivalent to minimising $\epsilon = \sum K_i(2 + K_i)$ for small K_i. The least squares estimate is the tensional ground state of the data polis.

Gauss's normal distribution (error curve) is the equilibrium distribution of K when errors are independent and identically distributed. The Gaussian (bell curve) is the tensional attractor.

3 Francis Galton – Regression and Correlation

Galton studied hereditary stature, discovering regression to the mean and correlation. In POLIS V12, regression predicts the expected K_Y given K_X: $\hat{K}_Y = r \cdot K_X$, where r is the correlation coefficient. Galton's data (parents and children) showed that tall parents have children shorter (regression), i.e., $\hat{K}_{\text{child}} = b \cdot K_{\text{parent}}$ with $b < 1$. Correlation measures the linear coupling strength: $r = \text{Cov}(K_X, K_Y)/\sqrt{\text{Var}(K_X)\text{Var}(K_Y)}$.

Galton's "reversion to mediocrity" (regression toward the mean) is a tensional property: extreme values (K near 1 or 0) tend to move toward 0.5 (the mean) over generations.

4 Karl Pearson – Chi-Square Test

Pearson introduced the χ^2 goodness-of-fit test and the chi-square distribution. In POLIS V12, the χ^2 statistic measures the deviation between observed $K_{\text{obs},i}$ and expected $K_{\text{exp},i}$: $\chi^2 = \sum (K_{\text{obs},i} - K_{\text{exp},i})^2 / K_{\text{exp},i}$. Under the null hypothesis, χ^2 has a certain K distribution. A large χ^2 means that the observed K distribution is unlikely under the model (high ϵ). Pearson's test is a tensional hypothesis test.

Pearson's coefficient of skewness and kurtosis measure the shape of K distributions. He also founded the journal *Biometrika*.

5 William Sealy Gosset – Student's t-Test

Gosset (under pseudonym "Student") developed the t-test for small samples. In POLIS V12, the t-statistic is $t = (\bar{K}_Y - \bar{K}_X)/(s/\sqrt{n})$, where s is the sample standard deviation. Gosset's contribution was to derive the exact distribution of t when n is small, which is the tensional distribution of the difference between two means divided by an estimate of ϵ. The t-test is a tensional significance test.

Gosset worked at Guinness brewery, improving quality control (monitoring K of yeast).

6 Ronald Fisher – Analysis of Variance (ANOVA)

Fisher developed ANOVA to partition variance into between-groups and within-groups components. In POLIS V12, the total sum of squares $SS_{\text{total}} = \sum_{i,j} (K_{ij} - \bar{K})^2$ is partitioned into $SS_{\text{between}} + SS_{\text{within}}$. The F-statistic $F = (SS_{\text{between}}/(k-1))/(SS_{\text{within}}/(N-k))$ compares tensional variation due to treatment to residual variation. Fisher's "randomised block" design reduces ϵ by controlling for nuisance variables.

Fisher's principle of maximum likelihood estimates parameters by maximising the probability of observing the data (K). He also coined the term "null hypothesis".

7 Jerzy Neyman – Confidence Intervals

Neyman (with Egon Pearson) developed the theory of confidence intervals. In POLIS V12, a $100(1-\alpha)\%$ confidence interval for a parameter θ is an interval $[L, U]$ such that over repeated sampling, the probability that θ lies in the interval equals $1-\alpha$. This is a tensional quantification of uncertainty: the interval covers the true K with specified coverage. Neyman's approach contrasts with Fisher's fiducial inference.

Neyman also introduced the concept of "inductive behaviour" – decisions based on minimising ϵ. The Neyman-Pearson lemma gives the most powerful test for a simple hypothesis (maximising K_{power}).

8 Conclusion

The six foundational contributions to statistics are coherently reinterpreted within the POLIS V12 tensional ontology. Least squares, regression, chi-square, t-test, ANOVA, and confidence intervals all become natural consequences of the closure condition $\epsilon = \sum K_m(2+K_m) = 0$ and the fractal hierarchy of statistical polises. No free parameters are added.

Zenodo references

- Main treatise: 10.5281/zenodo.19618276
- POLIS Bible: 10.5281/zenodo.19836226

POLIS V12 – Tensional Reinterpretation of Six More Statistical Pioneers Jorge Batista Alves Pereira
Independent Researcher, Sabugal, Guarda, Portugal
ORCID: 0009-0000-6385-7245 May 2026

Abstract

This paper extends the POLIS V12 tensional reinterpretation to six additional statistical giants: Thomas Bayes (Bayes' theorem), Adolphe Quetelet (average man), Florence Nightingale (polar area diagram), Charles Spearman (factor analysis), George Box (response surface), and C. R. Rao (information geometry). Each is re-read as a tensional configuration: Bayes's theorem as updating K; Quetelet's l'homme moyen as mean K; Nightingale's coxcomb as circular K histogram; Spearman's factor as latent K; Box's principal components as K variance directions; and Rao's distance as tensional metric. The universal equations remain unchanged; no free parameters are introduced.

9 Introduction

As in the companion paper, POLIS V12 rests on four axioms. After normalisation the mother equation is

$$\epsilon = \sum_{m=1}^{n} K_m(2 + K_m) = 0,$$

with $\mathrm{IDT}^* = \epsilon/(1+\epsilon)$. All real statistical systems are in Phase 4 ($\mathrm{IDT}^* \geq 0.70$) unless artificially uniform. The Rolling Law $2\pi r_p = V_{\mathrm{orb}} T_{\mathrm{rot}}$ applies fractally.

This paper reinterprets six more foundational contributions to statistics.

10 Thomas Bayes – Bayes' Theorem

Bayes's theorem updates K values based on evidence: $P(H|E) = P(E|H)P(H)/P(E)$. In POLIS V12, the prior $K_{\mathrm{prior}} = P(H)$, the likelihood $K_{\mathrm{likelihood}} = P(E|H)$, and the posterior $K_{\mathrm{posterior}} = P(H|E)$. The theorem is a tensional update rule: $K_{\mathrm{posterior}} = (K_{\mathrm{likelihood}} \cdot K_{\mathrm{prior}})/\mathrm{normalisation}$. Bayesian inference is an iterative Phase 5 process: each new data point revises the K distribution of the hypothesis.

The "Bayesian brain" hypothesis posits that the brain performs approximate Bayesian inference (tensional prediction error minimisation).

11 Adolphe Quetelet – L'Homme Moyen (The Average Man)

Quetelet applied statistics to social phenomena, introducing the concept of the "average man" as a norm. In POLIS V12, the average man is the mean K_{mean} of a population over many variables (height, weight, chest circumference). Quetelet's body mass index (BMI

$=$ weight/height^2) is a K index of fatness. His work on crime statistics showed that crime rates are predictable (low ϵ) over large populations. Quetelet's "moral statistics" paved the way for sociology as a quantitative discipline.

He also defined the "Quetelet index" which is the basis of modern BMI.

12 Florence Nightingale – Polar Area Diagram (Coxcomb)

Nightingale used the polar area diagram to show the causes of mortality in the Crimean War. In POLIS V12, the polar diagram is a circular histogram where each wedge's angle is constant and the radius is proportional to K_{death} (number of deaths). The diagram highlighted that most deaths were due to preventable disease (unsanitary conditions). Nightingale's visualisation is a tensional mapping of ϵ per cause.

She was a pioneer of statistical graphics and of evidence-based public health.

13 Charles Spearman – Factor Analysis and General Intelligence

Spearman developed factor analysis, proposing that a general intelligence factor (g) underlies all cognitive tests. In POLIS V12, observed test scores K_i are linear combinations of a latent factor K_{general} plus specific factors $K_{\text{specific},i}$: $K_i = \lambda_i K_{\text{general}} + K_{\text{specific},i}$. Factor analysis extracts the common K that maximises shared variance. Spearman's technique is a tensional dimensionality reduction.

The two-factor theory (general + specific) has been influential in psychometrics. His work also introduced the concept of "tetrad differences" as a method to test for a single factor.

14 George Box – Response Surface Methodology

Box developed response surface methodology (RSM) for optimising processes. In POLIS V12, the response y is a function $y = f(K_{\text{inputs}})$. RSM fits a polynomial surface (often quadratic) to approximate f. The optimum (minimum or maximum K) is found by taking derivatives and solving $\nabla K = 0$ (tensional gradient zero). Box's "evolutionary operation" (EVOP) uses small perturbations to climb the K gradient.

Box also contributed to time series analysis (Box-Jenkins method) and to the design of experiments (Box-Behnken design). His famous quote: "All models are wrong, but some are useful" reflects that any K model is an approximation; only $\epsilon = 0$ is exact.

15 C. R. Rao – Information Geometry and Rao Distance

Rao introduced the concept of a Riemannian metric on statistical manifolds (Fisher information metric). In POLIS V12, a statistical model is a manifold of probability distributions, each point a K vector. The Rao distance between two distributions is a tensional divergence: $d^2 = \Delta K^T I(K) \Delta K$, where $I(K)$ is the Fisher information matrix. This metric is invariant under reparameterisation, i.e., independent of how K is scaled.

Rao's work connects statistics to differential geometry (tensional curvature). He also made contributions to linear models and multivariate analysis.

16 Conclusion

Six additional statistical pioneers are reinterpreted within the POLIS V12 tensional ontology. Bayes' theorem, the average man, polar diagrams, factor analysis, response surfaces, and information geometry all become natural consequences of the closure condition $\epsilon = \sum K_m(2 + K_m) = 0$ and the fractal hierarchy of statistical polises. No free parameters are added; the same equations that describe a physical system or a biological system also describe the logic of uncertainty.

Zenodo references

- Main treatise: 10.5281/zenodo.19618276
- POLIS Bible: 10.5281/zenodo.19836226

References for the twelve statisticians

- Gauss, C. F. (1809). *Theoria motus corporum coelestium*. Hamburg.
- Galton, F. (1886). “Regression towards Mediocrity in Hereditary Stature”. *Journal of the Anthropological Institute*, **15**, 246–263.
- Pearson, K. (1900). “On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is such that it can be Reasonably Supposed to have Arisen from Random Sampling”. *Philosophical Magazine*, **50**, 157–175.
- Gosset, W. S. (1908). “The Probable Error of a Mean”. *Biometrika*, **6**, 1–25.
- Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Edinburgh: Oliver & Boyd.
- Neyman, J. (1937). “Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability”. *Philosophical Transactions of the Royal Society A*, **236**, 333–380.
- Bayes, T. (1763). “An Essay towards Solving a Problem in the Doctrine of Chances”. *Philosophical Transactions*, **53**, 370–418.
- Quetelet, A. (1835). *Sur l’homme et le développement de ses facultés*. Brussels.
- Nightingale, F. (1858). “Notes on Matters Affecting the Health, Efficiency and Hospital Administration of the British Army”.
- Spearman, C. (1904). “General Intelligence, Objectively Determined and Measured”. *American Journal of Psychology*, **15**, 201–293.
- Box, G. E. P. & Wilson, K. B. (1951). “On the Experimental Attainment of Optimum Conditions”. *Journal of the Royal Statistical Society B*, **13**, 1–45.
- Rao, C. R. (1945). “Information and the Accuracy Attainable in the Estimation of Statistical Parameters”. *Bulletin of the Calcutta Mathematical Society*, **37**, 81–91.