ChatGPT

Core Statistical Methods

1. **Yusuf A. Sani (2017)** – *“Comparative Analysis of Random Number Generators Using Monte Carlo Algorithm.”* B.Sc. Thesis, Yobe State Univ. – Describes statistical tests for validating randomness, using chi-square goodness-of-fit to check uniformity of outcomes and chi-square contingency tests (independence tests) to detect non-random patterns in sequences [1]. *Open-access PDF:* RG full text.

2. **Justine Lequesne & Philippe Regnault (2018)** – *“Goodness-of-fit tests based on Kullback-Leibler divergence.”* arXiv:1806.07244 [stat.ME] – Proposes goodness-of-fit testing methods that use Shannon entropy and Kullback–Leibler divergence to quantify deviations of an empirical distribution from an expected distribution [2]. These information-theoretic tests can complement traditional tests in assessing the fairness (uniform randomness) of outcomes. *Open access:* DOI:10.48550/arXiv.1806.07244.

3. **David Abuhanna (2015)** – *“Roulette: More than just a Chance.”* Honors Thesis, Univ. of Nevada, Las Vegas – Uses computer-logged data from a physical roulette wheel to perform statistical analysis of fairness. The study found a significant statistical association between the ball’s release position and the landing pocket, implying the wheel was not generating outcomes with perfect independence or uniformity [3]. *Open access:* UNLV Digital Scholarship.

4. **Michael Small & Chi Kong Tse (2012)** – *“Predicting the outcome of roulette.” Chaos* **22**(3): 033150 – Demonstrates that even a well-calibrated casino-grade roulette wheel can exhibit slight physical biases detectable via statistical analysis. Using high-speed camera data and physics modeling, the authors identified systematic, statistically significant outcome biases (e.g. due to a small tilt) deviating from the uniform random expectation [4]. *Open access:* DOI:10.1063/1.4753920 (preprint available on arXiv).

Power-Analysis and Sample-Size Guidance

1. **Abel Rodríguez & Bruno Mendes (2018)** – *Probability, Decisions and Games: A Gentle Introduction using R.* Wiley – Provides a quantitative illustration of how many roulette spins are needed to detect a small bias. Through simulations and Chebyshev’s inequality, it shows that a subtle bias (on the order of a few percent in pocket probability) may **require hundreds of thousands of spins** to confirm with high confidence [5] [6]. (For example, even ~5,000 spins might not reveal a ~5% bias, and on the order of 5×10^5 spins could be needed for reliable detection.) *Open-access PDF:* Chapter 3 excerpt.

2. **John F. Newsom (2021)** – *“Sample Size Issues for Categorical Analyses and Logistic Regression.”* (Lecture notes, Portland State Univ.) – Gives practical power-analysis guidelines for chi-square tests. It notes that detecting a **small effect size** difference (Cohen’s $w \approx 0.1$, comparable to a few-percent deviation in probabilities) with high power (~0.8) may require on the order of **n ≈ 800** observations [7]. This underscores that very subtle roulette biases demand large sample sizes to achieve statistically significant results. *Open access:* PSU archive PDF.

3. **NIST (2010)** – *NIST SP 800-22 Rev.1a: A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications.* – Although focused on random bit generators, this official guide underscores sample size in randomness testing. It recommends sequence lengths on the order of **10^3–10^7** for many tests, since only with large samples can subtle biases be detected and asymptotic test statistics hold [8]. This implies that to robustly test a physical randomizer like a roulette wheel, one must collect thousands or millions of spins. *Open access:* NIST PDF.

Industry or Regulatory Documents

1. **Gaming Laboratories International (2016)** – *GLI-11: Standards for Gaming Devices in Casinos, Version 3.0.* – Widely adopted technical standard that mandates rigorous statistical testing of casino RNG devices (including electro-mechanical devices). It specifies that outcomes must pass goodness-of-fit tests for uniform distribution (e.g. a chi-square “Total Distribution” test) and that sequential results must pass tests for independence (e.g. runs tests, serial correlation tests) [9]. The standard requires these tests to be evaluated at a high confidence level (99%), with sufficiently large sample sizes to reliably catch any deviation beyond acceptable randomness criteria [10] [11]. *Open access:* GLI-11 PDF.

2. **eCOGRA (2013)** – *Generally Accepted Practices (eGAP) – Fair Gaming Requirements.* – Industry accreditation criteria for online gambling fairness. The eGAP standards require that all game outcomes be **statistically random**, which is verified by ensuring the RNG outputs are **uniformly distributed** across the outcome range and **independent** from event to event [12]. Certified operators must test their roulette results (and other games) both before launch and periodically thereafter to meet these randomness standards. *Open access:* eCOGRA Fairness (BOS Report, 2013).

3. **National Institute of Standards and Technology (NIST, US) – SP 800-22 Test Suite (2010)** – Serves as a de facto reference for randomness testing in regulated industries. Many gaming regulators and labs reference tests like those in NIST SP800-22 for assessing physical randomizers. For example, Gaming regulators often incorporate the NIST battery (frequency test, runs test, serial correlation, entropy tests, etc.) as acceptable criteria for randomness certification [13] [14]. The NIST tests define pass-fail thresholds (typically at $\alpha=0.01$ or 0.05) which roulette outcome data must satisfy to be considered fair. *Open access:* NIST SP800-22 Rev.1a PDF.

4. **Gaming Laboratories Intl. – Technical Bulletins (various)** – *Example:* **“Technical Specifications for RNG Testing”** (GLI, 2020) outlines how manufacturers must submit hardware/software for statistical evaluation. Such documents emphasize collecting very large samples of RNG output (sometimes millions of data points) and applying batteries of tests (chi-square, runs, autocorrelation, etc.) to ensure compliance [15] [16]. These industry guidelines directly translate to practices like long-term camera-based logging of roulette spins to gather enough data for statistical fairness audits. *Open access:* GLI Tech Specs (RNG Testing).

---

[1] (PDF) Comparative Analysis of Random Number Generators Using Monte Carlo Algorithm
https://www.researchgate.net/publication/391903530_Comparative_Analysis_of_Random_Number_Generators_Using_Monte_Carlo_Algorithm

2 Package vsgoftest for R: goodness-of-fit tests based on Kullback-Leibler divergence
https://arxiv.org/pdf/1806.07244

3 "Roulette: More than just a Chance" by David Abuhanna
https://digitalscholarship.unlv.edu/honors_theses/27/

4 [1204.6412] Predicting the outcome of roulette
https://arxiv.org/abs/1204.6412

5 6 programmer-books.com
https://www.programmer-books.com/wp-content/uploads/2018/08/probability-decisions-games.pdf

7 web.pdx.edu
https://web.pdx.edu/~newsomj/cdaclass/ho_sample%20size.pdf

8 NIST SP 800-22, A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications
https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-22r1a.pdf

9 10 11 13 14 GLI-11 V3.0
https://gaminglabs.com/wp-content/uploads/2018/09/GLI-11-Gaming-Devices-V3-0.pdf

12 eCogra - Operator Template
https://www.bos.nu/wp-content/uploads/2020/01/Benchmarking-Study_Bos_14Nov2013.pdf

15 16 Technical Specifications for RNG Testing - GLI
https://gaminglabs.com/getting-started/technical-specifications-for-rng-testing/