Theorems for Exchangeable Binary Random Variables with Applications to Voting Theory

Serguei Kaniovski

March 26, 2013

Overview

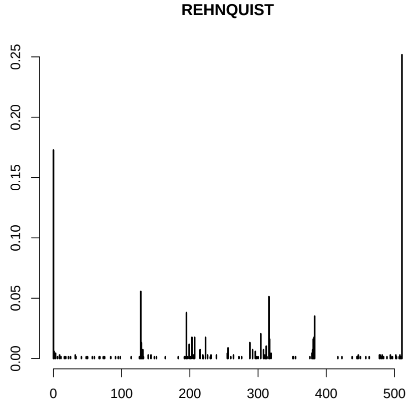
1. Examples of expertise and power computations based on a non-trivial probability distribution over the set of all voting outcomes
2. The property of exchangeability as a stochastic model of a representative agent (voter)
3. Known parameterizations of the joint probability distribution of n correlated binary random variables
4. The probability of at least k successes in n correlated binary trials. The generalized binomial distribution
5. The bounds on this probability when the higher-order correlations are unknown
6. Application to the Condorcet Jury Theorem and voting power in the sense of Penrose - Banzhaf
7. Concluding remarks

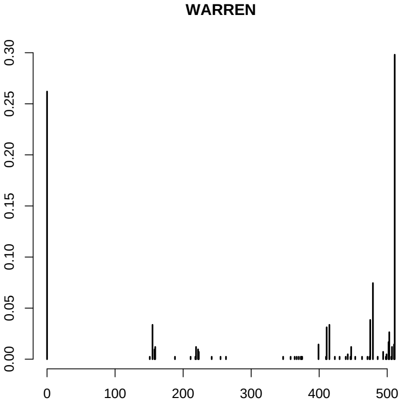
Examples of voting under simple majority rule ($n = 3$)

$\mathbf{v} = (v_1, v_2, v_3)$			$p = 0.5$, $c = 0$	$p = 0.75$, $c = 0$	$p = 0.75$, $c = 0.2$	$p_1 = 0.75$, $p_{2,3} = 0.6$, $c = 0.2$
1	1	1	0.125	0.422	0.506	0.357
1	1	0	0.125	0.141	0.094	0.136
1	0	1	0.125	0.141	0.094	0.136
1	0	0	0.125	0.047	0.056	0.122
0	1	1	0.125	0.141	0.094	0.051
0	1	0	0.125	0.047	0.056	0.056
0	0	1	0.125	0.047	0.056	0.056
0	0	0	0.125	0.016	0.044	0.086
Condorcet probability			0.5	0.844	0.788	0.679
Banzhaf probability 1			0.5	0.376	0.3	0.384
Banzhaf probability 2			0.5	0.376	0.3	0.365
Banzhaf probability 3			0.5	0.376	0.3	0.365

Computing the probability of a correct verdict, or the voting power as the probability of casting a decisive vote, requires a joint probability distribution on the set of all voting profiles $\mathbf{v} \in \mathbb{R}^{2^n}$. The influence of voting weights and decision rule is separate from that of the distribution. Exchangeability leads to a representative agent model, in which the independence assumption is relaxed

Voting in the U.S. Supreme Court





Empirical evidence overwhelmingly refutes the assumption of independent votes required in the classic versions of the Condorcet Jury Theorem and the Banzhaf measure of voting power

The joint probability distribution of n binary r.v.

The Bahadur parametrization

$Z_i = (V_i - p_i)/\sqrt{p_i(1-p_i)}$ for all $i = 1, 2, \ldots, n$, $p_i = p$

$c_{i,j} = E(Z_i Z_j)$ for all $1 \leq i < j \leq n$, $c_{i,j} = c$

$c_{i,j,k} = E(Z_i Z_j Z_k)$ for all $1 \leq i < j < k \leq n$, $c_{i,j,k} = c_3$

$\ldots$

$c_{1,2,\ldots,n} = E(Z_1 Z_2 \ldots Z_n)$, $c_{1,2,\ldots,n} = c_n$

$$\pi_{\mathbf{v}} = \bar{\pi}_{\mathbf{v}} \left(1 + \sum_{i<j} c_{i,j} z_i z_j + \sum_{i<j<k} c_{i,j,k} z_i z_j z_k + \cdots + c_{1,2,\ldots,n} z_1 z_2 \ldots z_n \right)$$

where $\bar{\pi}_{\mathbf{v}} = \prod_{i=1}^{n} p_i^{v_i} (1-p_i)^{(1-v_i)}$ is the probability under the independence

The George - Bowman parametrization for exchangeable binary r.v.

$$\pi_i = \sum_{j=0}^{i} (-1)^j C_i^j \lambda_{n-i+j}, \text{ where } \lambda_i = P(X_1 = 1, X_2 = 1, \ldots, X_i = 1), \quad \lambda_0 = 1$$

The generalized binomial distribution

This probability finds wide application in reliability and decision theory

The probability of at least k successes in n correlated binary trials

$$P_n^k(\mathbf{p},\mathbf{C}) = P_n^k(\mathbf{p},\mathbf{I}) + \prod_{h=1}^{n} p_h \sum_{1\le i<j\le n} c_{i,j}\alpha_i\alpha_j A_{i,j}^k(\boldsymbol{\alpha}), \quad \text{where}$$

$$A_{i,j}^k(\boldsymbol{\alpha}) = \begin{cases} 0, & k=0 \\ \displaystyle\sum_{\substack{i_s\neq i,\, i_s\neq j \\ 1\le i_1<\cdots<i_{n-k}\le n}} \alpha_{i_1}^2\ldots\alpha_{i_{n-k}}^2 - \sum_{\substack{j_s\neq i,\, j_s\neq j \\ 1\le j_1<\cdots<j_{n-k-1}\le n}} \alpha_{j_1}^2\ldots\alpha_{j_{n-k-1}}^2, & k=1,\ldots,n-1 \\ 1, & k=n \end{cases}$$

Here $\mathbf{p}$ is the vector of marginal probabilities, $\boldsymbol{\alpha}$ such that $\alpha_i = \sqrt{(1-p_i)/p_i}$, $\mathbf{C} = (c_{ij})$ the $n\times n$ correlation matrix, $\mathbf{I}$ the $n\times n$ identity matrix

In the following we will use a simpler formula, in which the r.v. are exchangeable and all higher-order correlations vanish. In this case the distribution is completely defined by n, p and the second-order correlation coefficient c

3-parameter generalized binomial distribution

For an odd n, $c_3 = c_4 = \cdots = c_n = 0$ and $(p, c) \in \mathcal{B}_n$ (Bahadur set)

$$P_n^k(p, 0) = \sum_{t=k}^{n} C_n^t p^t (1-p)^{n-t} = I_p(k, n-k+1)$$

$$P_n^k(p, c) = I_p(k, n-k+1) + 0.5c(n-1)\left(\frac{k-1}{n-1} - p\right)\frac{\partial I_p(k, n-k+1)}{\partial p}$$

where $I_x(a, b)$ is the regularized incomplete beta function

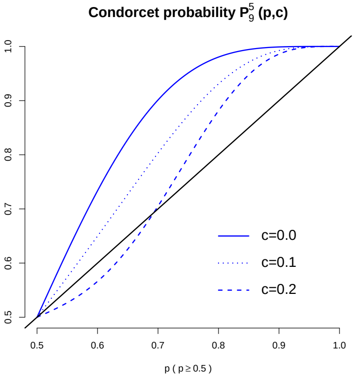
Bounds on $P_{n,p}^k$ for given n and p can be found by linear programming. Di Cecco provides bounds for given n, p and c such that $(p, c) \in \mathcal{B}_n$

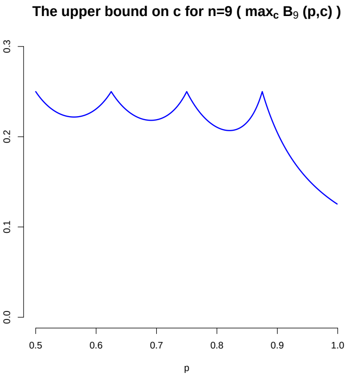
Bounds on $P_{n,p}^k$ when all correlation coefficients are unknown

$$\max\left\{\frac{np-k+1}{n-k+1}, 0\right\} \le P_{n,p}^k(c, c_3, \ldots, c_n) \le \min\left\{\frac{np}{k}, 1\right\}$$

The Condorcet probability and the Bahadur set

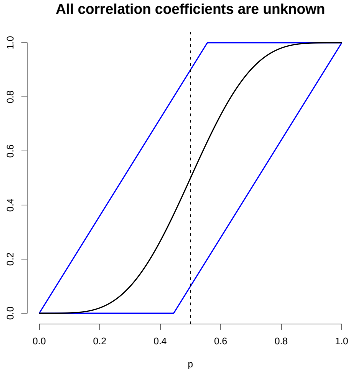
The set $\mathcal{B}_n$ contains all admissible values of c for given n and p, provided $c_3 = c_4 = \cdots = c_n = 0$

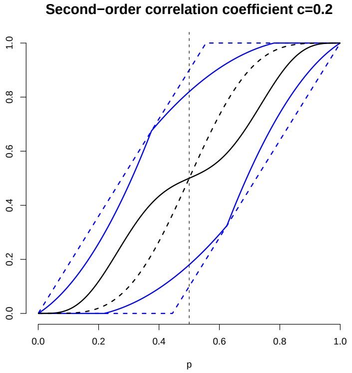




$\mathcal{B}_n$ is such that $0 < c < \frac{1}{n-1}$ for $p \approx 1$ and $0 < c < \frac{2}{n-1}$ for $p \approx 0.5$

Bounds on the probability of at least 5 successes in 9 trials





Voting power

In a 'one person, one vote' election with two alternatives a vote is decisive if it breaks a tie. With $n+1$ voters, the probability of a tie equals $C_n^{\frac{n}{2}} \pi_{\frac{n}{2}}$

For an odd n, $c_3 = c_4 = \cdots = c_n = 0$ and $(p, c) \in \mathcal{B}_n$

$$V_n^k(p, 0) = C_n^{\frac{n}{2}} p^{\frac{n}{2}} (1-p)^{\frac{n}{2}}$$

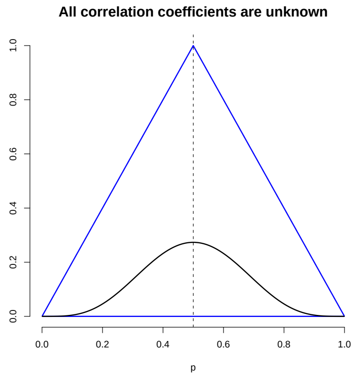
$$V_n^k(p, c) = V_n^k(p, 0) + \frac{nc}{4} C_n^{\frac{n}{2}} p^{\frac{n-2}{2}} (1-p)^{\frac{n-2}{2}} \left(\frac{n(2p-1)^2}{2} + 2p(1-p) - 1 \right)$$

Bounds on $V_n^k(c, c_3, \ldots, c_n)$ when all correlation coefficients are unknown

$$0 \leq V_n^k(c, c_3, \ldots, c_n) \leq 2 \min\{p, 1-p\}$$

For given n, p and c, bounds can be found by linear programming

Bounds on voting power



Penrose - Banzhaf and Straffin power measures

A general formula for Total Criticality

$$\sum_{T \subset N : i \notin T} \pi_T [\mathbf{w}(T \cup \{i\}) - \mathbf{w}(T)] \quad + \sum_{S \subseteq N : i \in S} \pi_S [\mathbf{w}(S) - \mathbf{w}(S \setminus \{i\})]$$

Bazhaf measure of voting power:

Assumption: $p = 0.5$ and $c = 0$

$$\pi_T = \pi_S = \frac{1}{2^{|N|}}$$

Straffin measure based on Homogeneity assumption:

Assumption: $p \sim U[0,1]$ (unconditionally $Corr(V_i, V_j) = 1/3$)

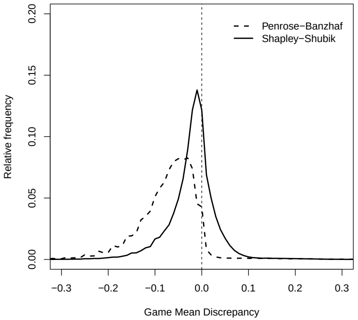
$$\pi_T = \frac{|T|!(|N| - |T|)!}{(|N|+1)!} \text{ and } \pi_S = \frac{|S|!(|N| - |S|)!}{(|N|+1)!}$$

The average Game Mean Discrepancy (GMD)

n	# Games	# Stoch. Models per Game	Measure of Power Penrose-Banzhaf	Shapley-Shubik
2	5	1,0000	-0.002	-0.002
3	19	6,518	-0.048	-0.022
4	167	4,883	-0.104	-0.037
5	7,580	3,943	-0.139	-0.040
6	7,828,353	3,304	-0.154	-0.034

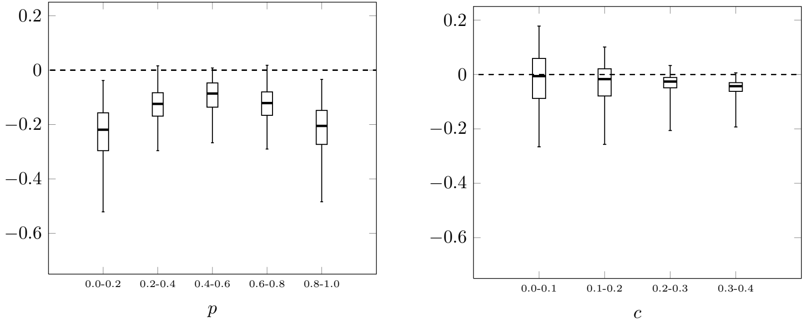
- The negative mean GMD implies that, on average, any voting power analysis carried out using the standard techniques will apportion too much power to the players. The standard versions of the Banzhaf and Straffin measures overestimate power
- The Straffin measure is closer to the probability of being critical than the Banzhaf measure

The GMD for $n = 6$



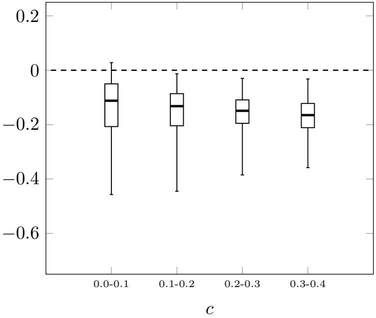
The right panel shows the GMD for both measures. The left panel shows the GMD for the Straffin measure when the assumption of the binomial model hold. This is ideal for the Banzhaf measure

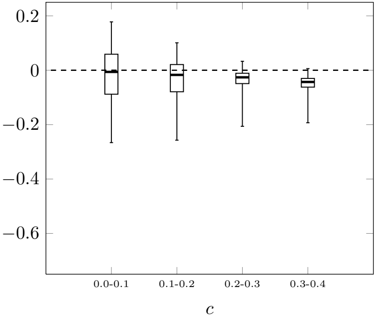
The GMD by intervals of p for $n = 6$



The median GMD shown by the horizontal bars in the middle is closer to zero for the Straffin measure (right panel) than for the Banzhaf measure (left panel). Both measures improve as they approach the midrange of $p \in [0.4, 0.6]$. The Straffin performs marginally better at small distances away from $p = 0.5$. The discrepancies for the Straffin measure remain smaller even for p close to 0 or 1

The GMD by intervals of c for $n = 6$





The Straffin measure (right panel) is superior to the Banzhaf (left panel) in all correlation ranges. Like Banzhaf, the Straffin measure becomes less accurate with increasing correlation. But the variance of the errors decrease with increasing correlation

Conclusions: Condorcet Jury

- The effect of correlation on the jury's competence is negative for voting rules close to simple majority and positive for voting rules close to unanimity
- If the individual competence is low, it may be better to hire one expert rather than several. In all other cases simple majority rule is the optimal decision rule. A jury operating under simple majority rule will not necessarily benefit from an enlargement, unless the enlargement is substantial. The higher the individual competence, the sooner an enlargement will be beneficial
- Correlation-robust voting rules minimizes the effect of correlation on collective competence by making it as close as possible to that of a jury of independent jurors. The optimal correlation-robust voting rule should be preferred to simple majority rule if mitigating the effect of correlation is more important than maximizing the accuracy of the collective decision
- For a given competence, compute the bounds to a jurys competence as the minimum and maximum probability of a jury being correct
- Using the generalized binomial distribution, we generalize the Condorcet Jury Theorem by allowing heterogeneity of experts, positive correlation between the votes, and qualified majority rules. For analytical tractability, we assume that any two votes correlate with the same correlation coefficient. The conventional wisdom holds that the groupthink or bandwagon effects diminish the collective competence. We show that this effect can be positive or negative, and provide sufficient conditions for it to have a certain sign

Conclusions: Voting power

- We can assess the magnitude of numerical error or bias in the Banzhaf measure that occurs when equiprobability and independence assumptions are not met. The probability bias is more severe than the correlation bias. Common positive correlation biases the measure upwards, common negative correlation downwards
- Despite the Banzhaf measure being a valid measure of *a priori* voting power and thus useful for evaluating the rules at the constitutional stage of a voting body, it is a poor measure of the actual probability of being decisive at any time past that stage
- Derive a modified square-root rule for the representation in two-tier voting systems that takes into account the sizes of the constituencies and the heterogeneity of their electorates. Since in a homogeneous electorate the votes are positively correlated, the larger and the more homogeneous the electorate, the less power a vote has
- Develop realistic voting scenarios that reflect the preferences of the voters via a correlation matrix. Then generate a consistent joint probability distribution and compute the probabilities of interest
- Compute the bounds to voting power as the minimum and maximum probability of the voter being decisive
- Simulations of all possible monotonic voting games with up to six players show that both the Banzhaf and Straffin measures tend to overestimate voting power when the votes are positively correlated. In most voting scenarios, the Straffin measure is closer to the probability of criticality than the Banzhaf measure

Literature 1

Parameterizations

- Bahadur, R.R., "A representation of the joint distribution of responses to n dichotomous items", in H. Solomon (ed.), Studies in item analysis and prediction, Stanford University Press, pp. 158–168
- George, E. O., Bowman, D., "A full likelihood procedure for analyzing exchangeable binary data", *Biometrics*, 1995, 51, pp. 512–523
- Zaigraev, A., Kaniovski, S., "A note on the probability of at least K successes in N correlated binary trials", *Operations Research Letters*, 2013, 41(1), pp. 116–120

Bounds

- Zaigraev, A., Kaniovski, S., "Exact bounds on the probability of at least k successes in n exchangeable Bernoulli trials as a function of correlation coefficients", *Statistics & Probability Letters*, 2010, 80(13–14), pp. 1079–1084
- Di Cecco, D., "A geometric approach to a class of optimization problems concerning exchangeable binary variables", *Statistics & Probability Letters*, 2011, 81(3), pp. 411–416

Literature 2

Condorcet Jury Theorem

- Kaniovski, S., Zaigraev, A., “Optimal jury design for homogeneous juries with correlated votes”, *Theory and Decision*, in print
- Kaniovski, S., “Aggregation of correlated votes and Condorcet's Jury Theorem”, *Theory and Decision*, 2010, 69(3), pp. 453–468
- Kaniovski, S., “An invariance result for homogeneous juries with correlated votes”, *Mathematical Social Sciences*, 2009, 57(2), pp. 213–222

Voting power / Square root rule / U.S. Supreme Court

- Kaniovski, S., Leech, D., “A behavioral power index”, *Public Choice*, 2009, 147(1), pp. 17–29
- Kaniovski, S., “The exact bias of the Banzhaf measure of power when votes are neither equiprobable nor independent”, *Social Choice and Welfare*, 2008, 31(2), pp. 281–300
- Kaniovski, S., “Straffin meets Condorcet. What can a voting power theorist learn from a jury theorist?”, *Homo Oeconomicus*, 2008, 25, pp. 1–22
- Kaniovski, S., Das, S., “The robustness of the Penrose-Banzhaf and Shapley-Shubik measures of power to positive correlation”, *Working Paper*, 2012