VII Latin American Symposium on Circuits and Systems (LASCAS) 2016

The Influence of Feature Vector on the Classification of Mechanical Faults using Neural Networks

Denys Pestana-Viana*, Rafael Zambrano-López†, Amaro A. de Lima*†,
Thiago de M. Prego*†, Sergio L. Netto† and Eduardo A. B. da Silva†
*PPEEL/DIPPG, Federal Center of Tech. Edu. Celso Suckow da Fonseca (CEFET-RJ) - Rio de Janeiro, Brazil
Email: denys.pestana@gmail.com
†PEE/COPPE, Federal University of Rio de Janeiro - Rio de Janeiro, Brazil
Email: {rafael.lopez,amaro.lima,thiago.prego,sergioln,eduardo}@smt.ufrj.br

***Abstract*—This paper investigates the problem of automatic detection of rotating-machine faults based on vibration signals acquired during machine operation. In particular, two new signal features, namely the kurtosis and entropy, are considered along with main spectral peaks to discriminate between several machine conditions: normal operation, (vertical and horizontal) misalignment, unbalanced load and bearing faults. Moreover, the inclusion of one set of three accelerometers for each roller bearing associated to the system acquiring more vibration signals also affects the generation of feature vector and is part of our proposal. In order to evaluate the rotating machine fault classification, a database of 1951 fault scenarios with several different fault intensities and rotating frequencies was designed and recorded, taking into consideration the specificities of the proposed machine learning task. The artificial neural networks recognition system employed in this work reached 95.8% of overall accuracy, showing the efficiency of the proposed approach.**

***Keywords*—Rotating machine, pattern classification, fault diagnosis, feature extraction.**

I. Introduction

The industrial scenario experienced a huge structural and organizational changes in recent years, requiring a high standard of quality in their processes. On large companies with a great amount of rotating machines and industrial equipment, it is vital to avoid halting production due to mechanical faults. In order to accomplish this aim, it is necessary to invest in the fault anticipation methods to be conjugated with predictive maintenance practices. Due to the gradual intensity increase and the random process characteristics in rotating machinery fault signal, it is difficult to human beings observe its early occurrence, making the pattern recognition a powerful tool for dealing with this kind of data and appropriated to this diagnosis task [1]–[4].

This work extends the research of [5], proposing a new feature vector adding the kurtosis and entropy measures, which are originally designed to address the bearing fault problems, and also with one more set of accelerometers, i.e., one for each roller bearing. In order to validate the proposal a new database was created, considering the characteristics to be analysed. The idea is to develop a pattern recognition system by using Artificial Neural Networks (ANN) [6] to discriminate among the simulated rotating machine failures addressed in the database and showing the effectiveness of the feature vector proposal in the system accuracy.

The paper is structured as follows: Section II describes the rotating machine fault emulation system, which is used in the database design and recording, detailed in Section III. In Section IV, the proposed feature vector is explained and compared with the original implementation. Section V shows the efficiency of the proposed approach applied to an even more complex recognition task than the one presented in [5]. Finally, in Section VI, the conclusions of the work are stated.

II. System Description

The system consists in a vibration simulation workbench, known as RotorKit Alignment Balance Vibration Trainer (ABVT) commercialized by SpectraQuest. This equipment is designed to study the dynamic of motors with shaft supported by roller bearings, allowing the fault simulation of unbalanced mass, axes misalignment and problems in roller bearings, among others.

The vibration signals are acquired from accelerometers assembled on the bearings support measuring the signals in 3 orthogonal directions: axial (x axis), tangential (y axis) and radial (z axis), as shown in Fig. 1.

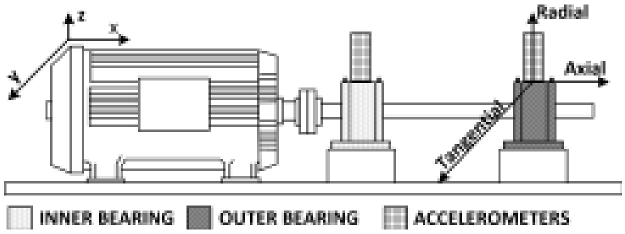

Fig. 1. Measurement directions of accelerometers and their position on the bearings. The axes x, y and z are corresponding to the directions axial, tangential and radial with respect to the shaft. The set of accelerometers positioned in inner bearing is labelled as set 1 and in outer bearing is labelled as set 2.

All simulation scenarios consider that the 2 bearings are simultaneously monitored by 2 separated sets of accelerometers, as indicated in [7]. Fig. 1 details the inner and outer bearings, the closest and the furthest bearings from the motor, respectively, and the accelerometer sets assembled on the top of the bearings. Four operating situations were emulated using the system described:

1) Normal (without faults): It is a no fault situation (without the presence of operating situations 2, 3, 4, described below), i.e., the disk load and shaft are perfectly balanced and aligned, and no defective roller bearing is used in the system.
2) Mass unbalance: This situation occurs due to an asymmetric mass distribution around the rotational axis. This fault is implemented by using the disk load varying the weights of the load attached and the rotating frequency.
3) Shaft parallel misalignment: The misalignment scenario consists in dislocating the motor shaft in parallel directions to the load shaft, i.e., shifting the motor into the radial or tangential directions.
4) Fault in roller bearings: Most of the bearing faults are originated from its main components: outer race, cage, rolling elements

(balls) and inner race.

The acquisition system has 2 bearings (the inner and outer bearings), being only one defective at a time, i.e., if the inner is faulty the outer is normal (has no fault), and vice versa, as can be observed in Fig. 1. This kind of fault is not apparent by itself, it requires a certain mass unbalance to make it evident in spectral analysis. This configuration was implemented by applying 3 different types of bearing faults related to cage, outer race and ball defects, in both bearing positions individually, and also varying unbalance mass and rotation speed.

The vibration signals were obtained from 2 sets of accelerometers each one associated to one bearing and measuring in 3 directions (axes x, y and z). The set assembled on the inner bearing (set 1) support generates the signals labelled as $s_{x_1}(n)$, $s_{y_1}(n)$ and $s_{z_1}(n)$, respectively to each axis. Consequently, the signals related to the outer bearing (set 2) are named $s_{x_2}(n)$, $s_{y_2}(n)$ and $s_{z_2}(n)$, following the same rationale. A trigger signal, $s_t(n)$, acquired a part from the vibration signals allows the rotating speed estimation compounding a 7-channel data signal with the vibration data from the 2 sets of accelerometers plus the trigger information. All the signals were recorded using sampling frequency of 50 kHz for 5.0 s, making a total of 250000 samples for one channel signal.

III. Database

The database is expected to emulate all operational scenarios under study. In our case, all possible-rotating machine faults and system condition variations containing enough information to characterize and discriminate the faults. In this work the developed database covered the following operating conditions:

- Normal: In this class, no fault is implemented and 49 different rotation speeds are applied within the interval $10 < F_r < 60$ Hz, where F_r is the rotation frequency of the machine.
- Unbalance: Several scenarios with unbalancing weights of $[6, 10, 15, 20, 25, 30, 35]$ g are considered, each with $[49, 48, 48, 49, 47, 47, 45]$ different rotation speeds, respectively, in the range of $10 < F_r < 60$ Hz, making a total of 333 mass unbalance scenarios are addressed.
- Misalignment: This condition involves both scenarios, the horizontal and vertical shaft misalignments. The vertical misalignment signals includes displacements of $[0.51, 0.63, 1.40, 1.90, 1.27, 1.78]$ mm, each with $[51, 50, 50, 50, 50, 50]$ different rotation speeds, respectively, totalling 301 scenarios of vertical misalignment. The horizontal misalignment signals include displacements of $[0.50, 1.00, 1.50, 2.00]$ mm, each with $[50, 49, 49, 49]$ different rotation speeds, respectively, totaling 197 horizontal misalignment scenarios, making a total of 498 misalignment scenarios addressed with rotation frequencies in the range of $10 < F_r < 60$ Hz.
- Bearing fault: The scenarios consider the faulty bearing in 2 different positions, inner or outer bearing, where both are subjected to bearings with the following faults: cage, outer race and ball. The signals obtained with faulty bearing at the inner bearing position include [cage, outer race, ball] faults with unbalancing weights of $[0, 6, 20, 35]$ g for all faults. The amounts of rotating speeds for each fault are $[49, 48, 49, 42]$, $[49, 49, 49, 37]$ and $[50, 49, 49, 38]$, respectively, totaling 558 bearing fault scenarios at the inner bearing position. The bearing fault signals at outer bearing position include the same 3 faults, with the same 4 unbalancing weights, and with $[49, 48, 49, 41]$, $[49, 49, 49, 41]$ and $[49, 43, 25, 38]$ rotating speeds, grouped by fault respectively, totaling 513 bearing fault scenarios for the outer bearing position. Combining both bearing positions (inner and outer), a total of 1071 bearing faults scenarios are addressed. The rotating speed is within the range of $10 < F_r < 60$ Hz for this whole operational condition.

The entire database comprises a total of 1951 different fault scenarios for 4 different operational conditions. 49 of which from the normal class, 333 from unbalance class, 498 from misalignment class and 1071 from the bearing fault. This database is named 1951-signal database.

IV. Feature Extraction

The techniques for feature extraction consist in estimating the rotating speed from $s_t(n)$, and obtaining spectral and statistics information from the 6 vibration signals. The adopted measures associated to the faults in order to discriminate them are:

- Rotating frequency, (F_r), estimation: The technique to estimate the rotating frequency is described in [5], and it consists in calculating the N_{DFT}-point DFT of $s_t(n)$, $S_t(k)$, where k is the frequency bin associated to a frequency $f_i = \frac{k_i F_s}{N_{\mathrm{DFT}}}$ Hz and F_s is the data acquisition sampling frequency. The frequencies of the 4 highest magnitude peaks are obtained and their associated lowest frequency is chosen as the estimated rotating speed, F_r.
- Entropy: It is a measure of randomness or unpredictability of a random variable [8]. The entropy H of a discrete signal $x(n)$ is given by (1).

$$H = \sum_{i=1}^{C} p_i(x) \log\left(\frac{1}{p_i(x)}\right), \tag{1}$$

  where $p_i(x)$ is the signal Power Density Function (PDF), which is obtained through a kernel density estimator by applying Parzen-Rosenblatt window method [9] to the input signal $x(n)$, and C is the number of points associated to the PDF. The entropy is calculated for the 6 vibration signals associated to the accelerometers sets, $s_{x_1}(n)$, $s_{y_1}(n)$, $s_{z_1}(n)$, $s_{x_2}(n)$, $s_{y_2}(n)$ and $s_{z_1}(n)$, generating H_{x_1}, H_{y_1}, H_{z_1}, H_{x_2}, H_{y_2} and H_{z_1}, respectively.
- Kurtosis: It is a dimensionless statistical measure defined by the normalized fourth statistical moment of a signal, as shown in (2). It indicates the shape deformation of a PDF in relation to a Gaussian PDF.

$$K = \frac{E[x(n) - \mu]^4}{\sigma^4}, \tag{2}$$

  where, $E[\cdot]$ is the expected value, μ is the mean and σ is the standard deviation of $x(n)$. The kurtosis calculation is applied to the 2 sets of accelerometers for 3 directions, making a total of 6 measures, K_{x_1}, K_{y_1}, K_{z_1}, K_{x_2}, K_{y_2} and K_{z_2} for the axes (x_1, y_1, z_1) and (x_2, y_2, z_2) related to accelerometer sets 1 and 2.
- Spectral analysis: A N_{DFT}-point DFT of each signal acquired from the accelerometers, $s_{x_i}(n)$, $s_{y_i}(n)$ and $s_{z_i}(n)$ for the sets of accelerometers $i = 1, 2$ is calculated, generating their spectral representation $S_{x_i}(k)$, $S_{y_i}(k)$ and $S_{z_i}(k)$ for $i = 1, 2$, where n represents discrete-time index and k represents frequency bin.

A. Feature vector

Compared to the technique presented in [5], the proposed one differs by applying one more set of accelerometers, extracting twice more vibration signals and calculating 2 additional measures, namely kurtosis and entropy, which are expected to be helpful in discriminating bearing faults.

In order to generate the feature vector, the first step is to estimate the motor rotating frequency, F_r and its first 2 harmonics, i.e., $2F_r$ and $3F_r$, then to calculate their related frequency bins $k_j = \frac{f_j N_{\rm DFT}}{F_s}$ for $j = 1, \ldots, 3$, where $\{f_1, f_2, f_3\} = \{F_r, 2F_r, 3F_r\}$. The second step is to obtain the magnitude from the vibration signals frequency response in F_r and its harmonics, i.e., $S_{x_i}(k_j)$, $S_{y_i}(k_j)$ and $S_{z_i}(k_j)$ for $i = 1, 2$ and $j = 1, 2, 3$, making a total of 12 measures. The third step is to calculate the statistic measures, kurtosis and entropy, from the discrete-time vibration signals, i.e., the 6 kurtosis K_{x_i}, K_{y_i} and K_{z_i} for $i = 1, 2$ and also the 6 entropy amounts H_{x_i}, H_{y_i} and H_{z_i} for $i = 1, 2$. The final step is to combine the measures in a feature vector, V_f, using the 3 magnitude harmonics, kurtosis and entropy from the 3 axes of the 2 accelerometers plus rotating frequency, which can be represented by:

$$\begin{aligned} V_f \;=\; & \{F_r, S_{x_i}(k_j), S_{y_i}(k_j), S_{z_i}(k_j), \\ & K_{x_i}, K_{y_i}, K_{z_i}, H_{x_i}, H_{y_i}, H_{z_i}\}, \end{aligned} \quad (3)$$

for $i = 1, 2$ and $j = 1, 2, 3$, achieving a 31 dimensionality vector, while the feature vector published in [5] reached dimension 10, because it only uses the accelerometer set 1 and no statistical measures, being represented by

$$V_{10}^1 = \{F_r, S_{x_1}(k_j), S_{y_1}(k_j), S_{z_1}(k_j)\}, \quad (4)$$

for $j = 1, \ldots, 3$. Just for comparison purposes 2 other feature vector are defined: V_{16}^1 and $V_{19}^{1,2}$.

$$\begin{aligned} V_{16}^1 \;=\; & \{F_r, S_{x_1}(k_j), S_{y_1}(k_j), S_{z_1}(k_j), \\ & K_{x_1}, K_{y_1}, K_{z_1}, H_{x_1}, H_{y_1}, H_{z_1}\}, \end{aligned} \quad (5)$$

for $j = 1, 2, 3$, making a dimensionality 16 vector for the accelerometer set 1. It can be regarded as V_{10}^1 adding kurtosis and entropy.

$$V_{19}^{1,2} = \{F_r, S_{x_i}(k_j), S_{y_i}(k_j), S_{z_i}(k_j)\}, \quad (6)$$

for $i = 1, 2$ and $j = 1, 2, 3$, making a 19-dimension vector for the accelerometer sets 1 and 2. It is equivalent to implement V_f without kurtosis and entropy measures.

V. Experimental Results

The fault classification experiment consists in adopting similar procedure presented in [5] to evaluate the system ability of fault discrimination by adding the kurtosis and entropy measures, not only for normal, unbalance and misalignment classes, but also for bearings faults. These tasks used as classifier Multi-Layer Perceptron ANN with input layer with the same size of input feature vector, one hidden layer with the amount of neurons approximately equals to the input layer size (obtained empirically), and the output layer with number of neurons equals to the number of classes to be discriminated.

A great difference in the number of elements of the classes can affect drastically the ANN classifier performance. In order to prevent it, the low number of elements of 'normal' class of the 1951-signal database, will be artificially increased, making the 49 elements of 'normal' class to be expanded in seven new signals by just adding 7 different realizations of a 10 dB-SNR Gaussian white noise in the original signals totalling 343 elements for the 'normal' class. The expanded database is composed of 343 signals from 'normal' class, 333 from the 'unbalance' class, 498 from 'misalignment' class and 1071 from 'bearing' fault, being named as 1951-signal-increased database.

The original 1951-signal database was divided into 3 disjoint sets with approximately 70%, 10%, and 20% of the signals for training, validation and test, respectively. Each of the sets is chosen to represent the data with maximum variability from rotating frequency and fault intensity. The ANN learning (training) process was performed by applying the Levenberg-Marquardt backpropagation algorithm and the validation set, which is employed to avoid ANN to become excessively specialized on the training signals thus loosing its generalization capability [5]. The database expansion is applied after the sets separation, in order to prevent the expanded signals originated from the same original one populating different database sets.

All tables related to the fault classification experiment present the classification performance X/Y, also called confusion matrix, for the test data from the 1951-signal-increased database, where X represents the recognized signals and Y is the total number of signals for the target class under analysis.

A. Influence of kurtosis and entropy

It is known that kurtosis and entropy are measures traditionally applied in bearing fault diagnosis. However their usage in discriminating 'normal', 'unbalance' and 'misalignment' classes has presented a quite improvement in recognition performance.

Tables I and II only differ by the use of kurtosis and entropy. Table I implements the framework described in [5], i.e., performs a classification task using V_{10}^1 with an ANN of 10-10-3 neurons in the input, hidden and output layers, respectively. Table II applies V_{16}^1, which is V_{10}^1 adding kurtosis and entropy, for the classification task with an ANN of 16-16-3 neurons. Although V_{16}^1 is not the proposed feature vector, it is the fair one to be compared to V_{10}^1. Comparing tables I and II, it is clearly noticed that the use of kurtosis and entropy improves each individual class performance and consequently the overall system accuracy from 81.2% to 94.8%, reaching a relative increase of about 14% in performance. The total classification indicated in Table II did not change significantly using the accelerometers set 2. However, the set 1 was chosen for comparison reason, being equivalent scenario adopted in [5].

TABLE I
CONFUSION MATRIX USING FEATURE VECTOR V_{10}^1 DESCRIBED IN [5].

Class	Target		
	Normal	Unbalance	Misalignment
Normal	**48/63**	2/60	6/90
Unbalance	0/63	**44/60**	3/90
Misalignment	15/63	14/60	**81/90**
Total (%)	**81.2**		

A similar rationale is considered in Tables III and IV, since they also differ by the use of kurtosis and entropy. Table III performs the classification task with an ANN of 19-19-3 neurons and feature vector $V_{19}^{1,2}$, which is equivalent to V_{10}^1 for both, inner and outer, accelerometer sets. Table IV implements the classification task using the proposed feature vector, V_f, which is $V_{19}^{1,2}$ adding the kurtosis and entropy measures, and an ANN of 31-35-3 neurons. Observing the performances of tables III and IV, once again the individual classes and the system accuracies presented significant improvements.

TABLE II
CONFUSION MATRIX USING FEATURE VECTOR V_{16}^{1}, WHICH IS V_{10}^{1} ADDING KURTOSIS AND ENTROPY.

Class	Target		
	Normal	Unbalance	Misalignment
Normal	**56/63**	0/60	2/90
Unbalance	0/63	**58/60**	0/90
Misalignment	7/63	2/60	**88/90**
Total (%)	**94.8**		

The relative total increase is approximately 9%, which represents the improvement in accuracy from 90.1% to 99.1%.

TABLE III
CONFUSION MATRIX USING FEATURE VECTOR $V_{19}^{1,2}$.

Class	Target		
	Normal	Unbalance	Misalignment
Normal	**52/63**	0/60	5/90
Unbalance	0/63	**55/60**	0/90
Misalignment	11/63	5/60	**85/90**
Total (%)	**90.1**		

TABLE IV
CONFUSION MATRIX USING FEATURE VECTOR V_f, WHICH IS $V_{19}^{1,2}$ ADDING KURTOSIS AND ENTROPY, MAKING A 31-DIMENSION VECTOR.

Class	Target		
	Normal	Unbalance	Misalignment
Normal	**63/63**	0/60	0/90
Unbalance	0/63	**58/60**	0/90
Misalignment	0/63	2/60	**90/90**
Total (%)	**99.1**		

The influence of the 2 sets of accelerometers, one set for each roller bearing, can be observed by comparing table I against III, and table II against IV, since their only difference is the usage of vibration signals of 2 sets of accelerometers, instead of only one. Similar behaviour shown in previous comparisons is also observed, indicating that the 2 sets of accelerometers also affect positively the system performance. The improvement was from 81.2% to 90.1%, considering tables I and III; and for tables II and IV, it was from 94.8% to 99.1%, making relative performances increase of around 10% and 4%, respectively.

B. Addition of bearings faults to the classification task

The idea is to apply the proposed feature vector, V_f, which reached an excellent performance in the 3-class task discrimination, in a 6-class recognition system. This more complex classification task consists in discriminating among 'normal' (C1), 'unbalance' (C2), 'horizontal misalignment' (C3), 'vertical misalignment' (C4), 'outer bearing' (C5) and 'inner bearing' (C6) faults. The ANN designed for this task has 31 neurons in the input, 35 in the hidden and 6 in the output layers.

The results shown in table V indicate an overall classification performance of 95.8%, suitable for the problem of mechanical faults classification, making the method proposed effective for the recognition of failure patterns under analysis in this work.

VI. CONCLUSION

This paper proposed a modification in feature extraction technique compared to the one adopted in [5], differing mainly by the usage of 2 sets of accelerometers, one for each roller bearings, and adding

TABLE V
CONFUSION MATRIX USING FEATURE VECTOR V_f FOR DISCRIMINATING NORMAL, UNBALANCE, HORIZONTAL AND VERTICAL MISALIGNMENT, AND OUTER AND INNER BEARING FAULTS.

Class	Target					
	C1	C2	C3	C4	C5	C6
Normal (C1)	**63/63**	0/60	0/36	0/54	0/93	0/101
Unbalance (C2)	0/63	**60/60**	1/36	0/54	0/93	1/101
Horiz. Misal. (C3)	0/63	0/60	**30/36**	1/54	0/93	1/101
Vert. Misal. (C4)	0/63	0/60	1/36	**53/54**	0/93	0/101
Outer Bear. (C5)	0/63	0/60	2/36	0/54	**89/93**	4/101
Inner Bear. (C6)	0/63	0/60	2/36	0/54	4/93	**95/101**
Total (%)	**95.8**					

2 new measures, kurtosis and entropy. To evaluate this approach, a new database with 1951 fault scenarios was developed. The proposed feature vector applied to the 3-class recognition problem (normal, unbalance and misalignment classes) described in [5] have reached the maximum system performance of 99.1% and maximum relative improvements of 14% and 10% for the applicability of the measures, kurtosis and entropy, and the 2 sets of accelerometers, respectively.

Even in a more complex discrimination task, a 6-class problem, where the misalignment is divided into horizontal and vertical, and the inner and outer bearing-fault categories are added, the proposed feature vector applied to the classification system achieved 95.8% of overall accuracy. The addition of kurtosis and entropy with one more set of accelerometers have shown to be essential to improve the fault classification accuracy of mechanical systems, even if no roller bearing fault is addressed.

The investigation of kurtosis and entropy individual contributions, the adoption of other statistical analysis, and techniques to deal with imbalanced database problem, as RusBoost and SMOTEBoost, will be considered for future work.

REFERENCES

[1] A. Widodo and B.-S.Yang, "Support Vector Machine in Machine Condition Monitoring and Fault Diagnosis," *J. Mechanical Systems and Signal Processing*, vol. 21, pp. 2560–2574, Jan. 2007.

[2] B. Samanta and K. R. Al-Balushi, "Artificial Neural Network Based Fault Diagnostics of Rolling Element Bearings using Time-domain Features," *J. Mechanical Systems and Signal Processing*, vol. 17, pp. 317–328, Mar. 2003.

[3] N. S. Vyas, D. Satishkumar, "Artificial Neural Network Design for Fault Identification in a Rotor-bearing System," *J. Mechanism and Machine Theory*, vol. 36, pp. 157–175, Feb. 2001.

[4] A.J. Hoffman and N.T. van der Merwe, "The Application of Neural Networks to Vibrational Diagnostics for Multiple Fault Conditions," *J. Computer Standards & Interfaces*, vol. 24, pp. 139–149, June 2002.

[5] A. A. de Lima, T. M. Prego, S. L. Netto, E. A. B. da Silva, R. H. R. Gutierrez, U. A. Monteiro, A. C. R. Troyman, F. J. C. Silveira and L. Vaz, "On Fault Classification in Rotating Machines using Fourier Domain Features and Neural Networks," *Proc. IEEE Latin American Symposium on Circuits and Systems (LASCAS)*, Cusco, Peru, Mar., pp. 1–4, 2013.

[6] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prenctice-Hall, Upper Saddle River:NJ, 1999.

[7] G. D. White, *Introduction to Machine Vibration*, DLI Engineering Corp., Bainbridge Island, WA,1995.

[8] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal* vol. 27, pp.379–423, July, 1948.

[9] E. Parzen, "On estimation of a probability density function and mode," *Annals of Mathematical Statistics*, vol. 33, pp. 1065–1076, Sept., 1962.